MOBILE SATELLITE CONFERENCE IMSC'90

Communications Canada

Communications Research Centre
Ottawa, Canada

NASA

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Mobilesat: Expanding the Communications Horizon

Mobilesat: Télécommunications sans Frontières

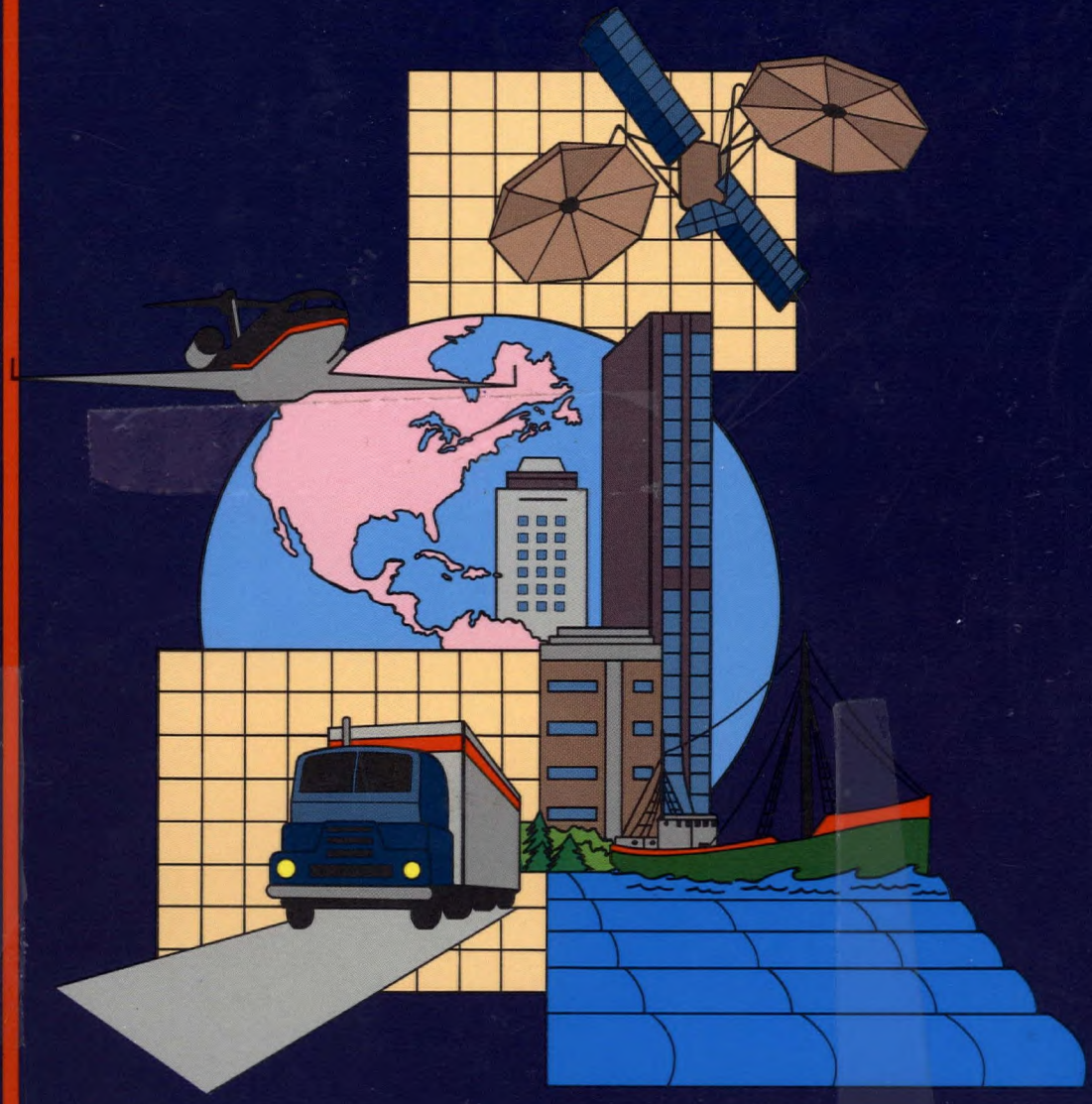
Les synopses sont disponibles en français

Proceedings of the Second

International Mobile Satellite Conference

IMSC '90

Ottawa, Ontario, Canada
June 17-20, 1990



JPL Publication 90-7

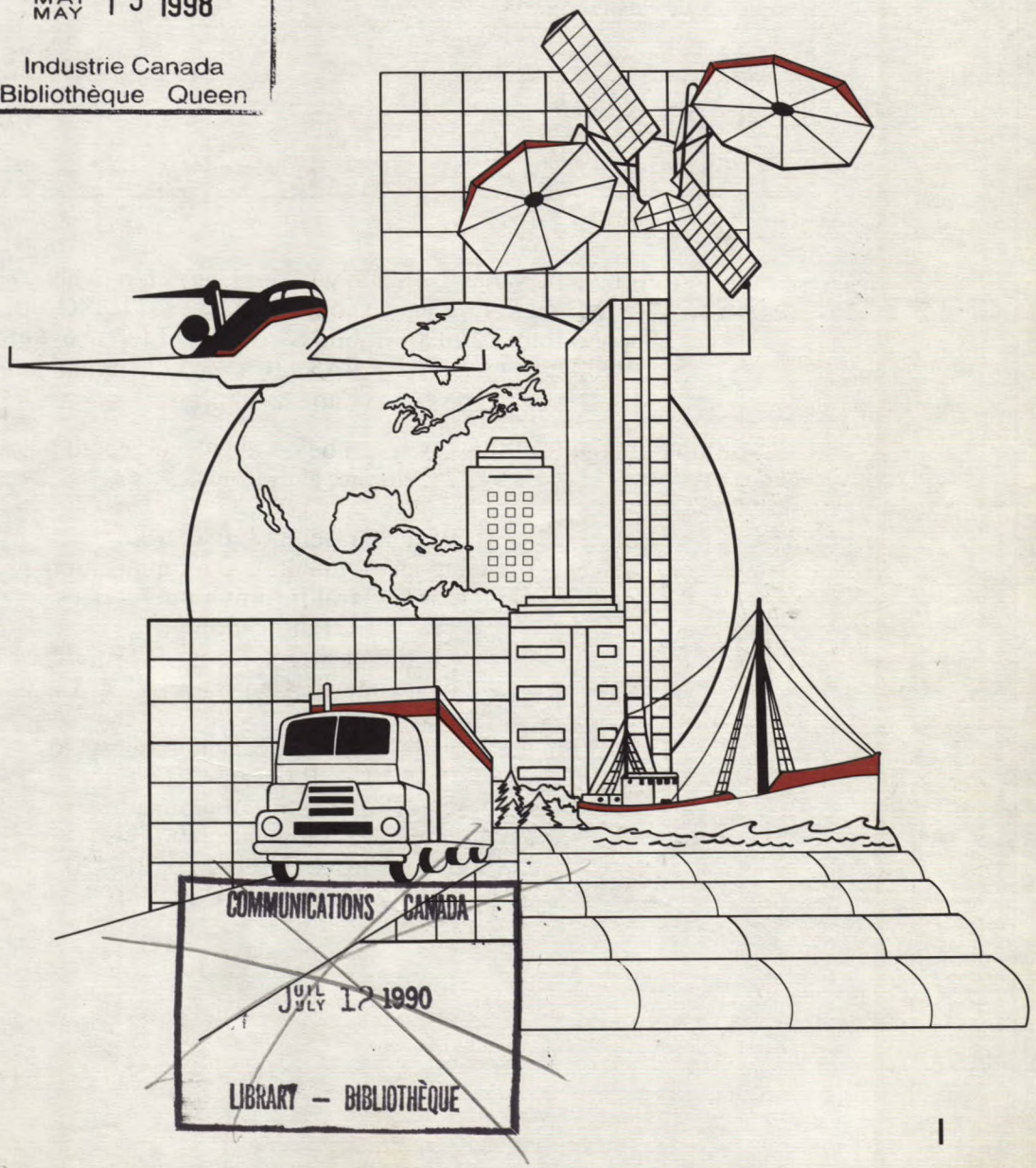
Proceedings of the Second

International Mobile Satellite Conference

(2nd : 1990 : Ottawa)

IMSC '90

Ottawa, Ontario, Canada
June 17-20, 1990



The cover graphic shows the convenience of wide area mobile access provided by mobile satellites.

Graphic by Debora Morton and Charlene Fortier, Addventures Ottawa, from an idea by Peter Mitchell, Communications Research Centre, DOC, Shirley Bay, Ottawa.

The IMSC '90 color logo was designed by the Graphics Department, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.

Artwork and Layout
Addventures, Ottawa
Printed by
National Printers, Ottawa

Cataloguing Data

International Mobile Satellite Conference (2nd : 1990 : Ottawa)

Proceedings of the Second International Mobile Satellite Conference, Ottawa, Canada, June 17-20, 1990. Compiled by R.W. Huck and Dr. W. Rafferty. Co-sponsored by the Canadian Department of Communications and the National Aeronautics and Space Administration of the United States of America.

1. Artificial satellites in telecommunications-Congresses. 2. Mobile communication systems-Congresses. I. Huck, R.W. II. Raffery, Dr. W. III. Canada. Dept. of Communications. IV. United States. National Aeronautics and Space Administration.

769 p.; 28 cm
ISBN 0-9694639-0-1
JPL Pub. No. 90-7.

TK 5104 621.38/0422

Additional copies of this book can be obtained, subject to availability, at no charge, as follows:

In Canada, send requests to:
Canadian Department of Communications
Director General Information Services
Re: JPL Publication 90-7
300 Slater Street, Room 1960
Ottawa, Ontario, Canada K1A OC8

In the U.S. and elsewhere, send requests to:
MSAT-X Program Office
Jet Propulsion Laboratory
Re: JPL Publication 90-7
4800 Oak Grove Drive, MS 238-420
Pasadena, California, USA 91109

Communications Canada
Communications Research Centre
Ottawa, Canada

NASA
National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Mobilesat: Expanding the Communications Horizon

Mobilesat: Télécommunications sans Frontières

Proceedings of the Second

International Mobile Satellite Conference

IMSC '90

Ottawa, Ontario, Canada
June 17-20, 1990

Co-sponsored
by the Canadian Department of Communications
and the National Aeronautics and Space
Administration of the United States

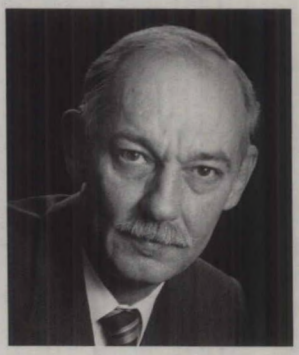
Compiled by:
R.W. Huck, Technical Committee Co-Chairman, DOC/CRC
Dr. W. Rafferty, Technical Committee Co-Chairman, NASA/JPL

Edited by:
D.H.M. Reekie, Conference Organizer, DOC/MSAT

OTTAWA CONGRESS CENTRE
Ontario
CENTRE DES CONGRÈS D'OTTAWA

IMSC - '90 : Steering Committee Members

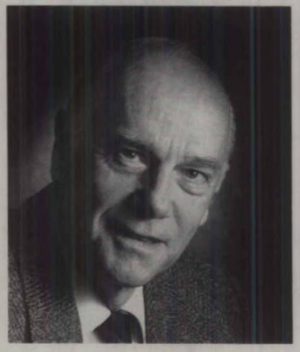
R. Arnold



D. Buchanan

L. Polite

D.H.M. Reekie



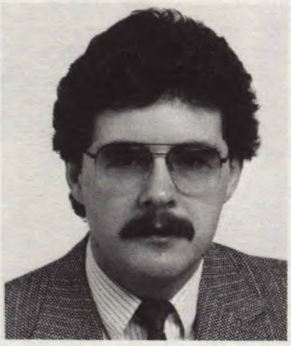
K. Cross

E. van Velthuizen

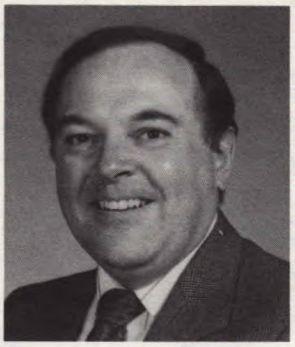
The conference was set up and managed by the IMSC '90 Steering Committee.
Shown are:
R. Arnold - Conference Co-Chairman, NASA
L. Polite - Conference Organizer, NASA/JPL
K. Cross - Finance Committee Chairman
D. Buchanan - Conference Co-Chairman, DOC/MSAT
D.H.M. Reekie - Conference Organizer & Publicity Committee Chairman, DOC/MSAT
E. van Velthuizen - Executive Assistant, Administration.

The Co-Chairmen of the Technical Committee, Dr. W Rafferty NASA/JPL and R.W. Huck, DOC/CRC are shown on the right.

A message from the IMSC '90 Steering Committee



Dr. W. Rafferty



R. W. Huck

In the late 1970s, NASA and the Canadian Department of Communications identified the potential role of satellites to provide a diversity of services to mobile users. NASA established a research and development mobile satellite program at JPL and at the same time DOC initiated a similar program with CRC. Both MSAT programs have developed a broad range of enabling technologies through independent and joint efforts. The technology development has had significant industry participation which has contributed to the realization of mobile satellite services. Currently in the United States and Canada, MSS has been licensed and will be in service in the early 1990s. In parallel, other countries have MSAT programs. Of particular note, is the INMARSAT Maritime Service which began 11 years ago, and has successfully demonstrated the feasibility of mobile satellite services.

The last few years has truly been an exciting time to be working in mobile satellite communications. The advancements in communication technologies and their electronics is bringing satellite communications to the consumer level. The growing demand and expectations of the mobile user has and will continue to exert considerable pressures on the industry to provide affordable services. Not only has this placed substantial demands on technology research and development, but also on the development of policy and regulation in both the domestic and international arena.

We would like to warmly welcome everyone attending the International Mobile Satellite Conference '90. The timeliness, and the interests of participants in IMSC '90, is well demonstrated by the large number of papers. Now is an excellent time to review technology, to consider standards and interoperability, to review spectrum issues, and to look at the various services being developed. We are particularly pleased with the future-looking papers which discuss alternate orbits, new frequency bands, and new services such as personal communications.

On behalf of the sponsors of this conference, the NASA Jet Propulsion Laboratory and the Canadian Department of Communications Research Centre, we would like to thank the session chairmen, the authors who so generously offered papers, the exhibit sponsors and all the participants. We would also like to thank members of our respective organizations who have given considerable time and effort to the planning of the overall conference and the sessions. In particular, Mr. Hugh Reekie of DOC and Ms. Lynn Polite of NASA/JPL have filled key roles.

We wish you all a good and productive three days at IMSC '90 where you can meet and interact with your colleagues and participate in the process of "Expanding the Communications Horizons", the theme of this conference.

Our Keynote Speakers

David Golden

Andrea Caruso

David Golden was born in Sinclair, Manitoba, and obtained his law degree at the University of Manitoba. He served as Deputy Minister of Defence Production for the Canadian Government from 1954 to 1962.

Prior to accepting the position of the newly-formed satellite company, Telesat Canada in 1969, he served as President of the Air Industries Association of Canada.

The degree of Honorary Doctor of Laws has been conferred on Mr. Golden by both Carleton University and the University of Manitoba. Since 1980, he has held the position of Chairman of the Board of Telesat Canada.

He is Director of many companies, including Telesat Mobile, Inc.

David Golden is the guest speaker at the Banquet on Tuesday, June 19, 1990.

Andrea Caruso obtained his law degree from the University of Rome, and started his career with the Italian Ministry of Posts and Telecommunications. In 1952 he became responsible for External Affairs with MPT, and in particular for relations with international organizations, including ITU, EEC and CEPT.

In 1964, he joined the International Telecommunications Union. Later in 1968, he became Director of International Affairs at Telespazio, Rome, and in that capacity represented Italy at Intelsat in Washington, D.C.

In 1977, Mr. Caruso became Deputy Director General of Intelsat. In 1982, he became Secretary General of the interim Eutelsat organization in Paris, becoming it's Director in 1985, when the formal Eutelsat structure was created. He retired from Eutelsat in 1989.

Andrea Caruso is the guest speaker at theclosing luncheon of the Conference on Wednesday, June 20, 1990.

Session Index

Session 1
Implementation of Current Mobile Satellite Systems

1

Session Chairman - *Alan E. Winter,* Microtel Pacific Research, Canada
Session Organizer - *Robert W. Huck, DOC*

MOBILESAT, Australia's Own

Michael Wagg
MOBILE COMMUNICATIONS BUSINESS

AUSSAT PTY LTD
GPO Box 1512
Sydney Australia
Ph: 61-2-238 7800
Fax: 61-2-238 7803

ABSTRACT/INTRODUCTION

Australia will be introducing a dedicated Mobile Satellite Communications System following the launch of the AUSSAT-B satellites in late 1991. The Mobile Satellite System, MOBILESAT, will provide circuit switched voice/data services and packet-switched data services for land, aeronautical and maritime users. This paper overviews the development program being undertaken within Australia to enable a fully commercial service to be introduced in 1992.

SYSTEM

The two AUSSAT-B satellites are HS-601 spacecraft being built by Hughes and due for launch on the Chinese Long March rockets. The primary purpose of the AUSSAT-B satellites is to provide replacement Ku-Band capacity for the current AUSSAT-A satellites with the major secondary function of providing L-Band capacity for mobile services. Each satellite will have a single 14 MHz L-Band transponder with associated Ku-Band back-haul providing 48 dBW of usable EIRP into a single beam covering the whole of Australia and its surrounding waters.

The ground infrastructure will consist of a network management system, to be provided in a redundant configuration from two locations, gateway stations to provide access to the public switched network and base stations for connection to private networks as well as the customer mobile terminals. Figure 1 schematically represents the system.

MARKET

The market for MOBILESAT services within Australia is targetted primarily at servicing the rural and remote areas of the country in a complementary nature to urban services such as cellular telephony and UHF/VHF private mobile radio systems. Australia is a vast country (~ 7.6 million square kilometres) with a small population (16 million) which is concentrated in less than ten major urban areas centres. Despite the population concentration in the urban areas the Australian export economy revolves around industries such as agriculture (sheep/wool, cattle/beef, wheat, etc), mining (coal, iron ore, etc) and tourism which, in general, are located in the remote regions of the country where the telecommunications infrastructure (both fixed and mobile) is quite often rudimentary or in some cases non-existent.

The urban population has rapidly adopted the mobile services provided by the cellular mobile telephone system with currently an 8% growth rate per month. The services was introduced in 1987 with, as of March 1990, over 10 users per thousand population, representing one of the most rapid take-ups of the technology throughout the world. The system is targetted at covering around 80% of the population during the next few years (currently 70% coverage) but this only represents around 1% coverage of the area of the country.

Market surveys have found that users in the rural and remote regions of the country also desire and have a need (it could be argued that the

need is greater than that of their urban cousins) for a quality mobile communications service with features equivalent to that available to urban dwellers. As well as showing a definite requirement for a satellite based mobile communications system, the market surveys identified a variation with the potential markets in the US and Europe where the long distance transport market dominated the user population. The trucking user does not dominate the Australian market but rather a range of service industries requiring a mix of public access and private network telephony applications dominated the usage requirements. Figure 2 shows the predicted mobile voice terminal market whilst table 1 summarises the usage breakdown.

Sector	Percentage
Sales	19.7%
Trades	18.1%
Agriculture	13.0%
Construction	9.8%
Pastoral	8.2%
Road Transport	7.2%
Mining	6.2%
Others	17.8%

Table 1
MOBILESAT Telephony Market Segmentation

The market within Australia has indicated a need for data only services but significantly at a terminal price around 50 - 60% of a voice terminal. As well a strong desire for a combined voice and packet-data capability led to the development of packet-data services using the voice signalling channels.

MARKET EXPECTATIONS

The market expectations for equipment pricing and service charges have generally been of pricing which is a little greater than that available from the urban cellular telephone system. This is due, in the main, to the perceived high technology nature of the service and, more importantly, to the added functionality, in particular the ubiquitous coverage capability of a satellite system. The expectations for a service to be introduced in 1992 are as follows:

Mobile Telephony Terminal
A$5000 (US$3750)
Mobile Radio Terminal
A$5000 (US$3750)
Mobile Data Terminal
A$2500 (US$1875)

Voice, Per minute air-time charge
+ Monthly Access Fee
A$1.20 (US$0.90) + A$0.0 to A$0.80 (US$0.60)
+ A$50 (US$37.50)

Data, Monthly Fee
A$30 - A$50 (US$22.50 - US$37.50)

These figures compare to ~A$1000 for a car mounted cellular telephone, ~A$3000 for a hand-held cellular telephone, ~A$3000 for a HF radio and A$35 per month plus A$0.70 per minute for long distance cellular charges in Australia.

The current knowledge of the system costs and perceived terminal prices suggests that the terminal prices should be achievable if not a little bit high at the commencement of service. As well, the usage charges, owing to the efficiencies of the digital voice implementation, are readily achievable and offer scope to be competitive with long distance cellular services.

TECHNICAL FEATURES

The MOBILESAT system developed for Australia shares many of the features of the systems being developed elsewhere as well as those of terrestrial mobile communication systems. However a number of key technical features have been adopted to address the market requirements and to utilise the full potential of the available technology.

A common signalling channel architecture has been adopted to accommodate developments in modulation and processing technologies and allow for a variety of communication standards.

Indeed the first release of the system specification has adopted both analog ACSSB and digital voice modulation as schemes being permitted to operate over MOBILESAT. Further, the features and the reserve capacity of the signalling system have been used to implement a packet-data communications service operating in conjunction with the circuit-switched voice service. This feature will provide for services such as position location in association with a voice service and will allow for a range of enhanced capabilities to be added to the voice services. This capability will also allow low cost stand alone packet-data terminals to provide for messaging and telemetry style applications.

The channel plan is based on 2.5 kHz channel spacing in a non-paired channel format, and will initially support up to 1,500 circuits using digital voice modulation with 5 kHz channels, at a threshold C/No of less than 45 dBHz. An additional 3 dB fade margin will be used to provide a high quality, robust voice service that will operate under severely shadowed satellite links.

The MOBILESAT system has adopted a three try (two repeat) strategy in the call initialisation procedure to accommodate for potential shadowing in the communications link. The three try strategy has been demonstrated to be the optimum scheme to provide a highly reliable signalling channel in the propagation environment of the mobile satellite service.

The voice modulation will be based on a nominal 4800 bps encoding algorithm. At the time of writing (March 1990), Australia's TELECOM Research Laboratories were undertaking an evaluation program, for Australia's MOBILESAT system and for INMARSAT's-M system, to determine the most approporiate digital encoding algorithm. The joint MOBILESAT/INMARSAT evaluation is aimed at achieving a common standard.

The MOBILESAT system will support a digital facsimile service based on the INMARSAT-M specifications, but operating at up to 4.8 kbps and incorporating an ARQ protocol to ensure error free transmission. The system design allows for an easy transition to a Store and Forward facsimile service in the future.

The public access gateway stations will directly interconnect with the ISDN network and provide an advanced, high capacity link into the PSTN, whilst low capacity minor base stations will support customised private networks.

IMPLEMENTATION

The MOBILESAT program is focussed on providing services using the AUSSAT-B satellites from mid 1992. However, the user community has an immediate need for the services that a mobile satellite system can offer so a parallel program is underway to provide an early entry MOBILESAT service. The close parallels with the Canadians with their MDS and the US with AMSC, Qualcomm and GEOSTAR where services are being implemented prior to a dedicated mobile satellite system, has been recognised. The early entry MOBILESAT service uses an unmodified INMARSAT-C product and will operate via the hub station provided in Perth, Australia by OTC Ltd, Australia's international common carrier and INMARSAT signatory.

The target market sectors for the early entry product are the remote monitoring and control industry, long distance transport and the remote services sector. The following applications are being implemented:

- interface of the INMARSAT-C system into a water resource management system through development of an interface box between the INMARSAT-C terminal and remote monitoring and control devices and modification of head office control software.

- monitoring of valuable cargos using INMARSAT-C interfaced to a position location device (TRANSIT or GPS) in the long distance transport and security industries.

- use of INMARSAT-C to provide the communications link from a lap-top PC to a head-office computer for applications such

as records access for remote area medical clinics provided by Australia's Royal Flying Doctor Service.

The program for implementation of the AUSSAT-B MOBILESAT service, at the time of writing, is at the stage of evaluation of the ground infrastructure components of the MOBILESAT system. Formal responses have been received for all components of the ground system, (Network Management System, Gateway/Base Access Stations and Mobile Terminals). The schedule for the implementation program is; contracts for infrastructure in mid 1990 with a service introduction by June 1992.

As a large proportion of the MOBILESAT market will be using a public-switched service, a close technical relationship has developed between AUSSAT, Australia's national satellite service provider and TELECOM Australia, Australia's monopoly PSTN and cellular service provider (as of April 1990). The interface issues between the terrestrial system and the mobile satellite system have been resolved and a mechanism for providing the ground infrastructure has been developed which allows for co-operation between the organisations in the technical areas and competition in the provision of services.

AUSTRALIAN INDUSTRY

The Australian decision to provide a Mobile Satellite Service was relatively late when compared to the deliberations and research and development undertaken in various parts of the world. However various sectors of industry throughout Australia have learnt from the international developments and will be in a position to provide the ground infrastructure for the MOBILESAT system (largely software control) as well as compete in the provision of MOBILESAT terminals, both within Australia and worldwide. Indeed, the timing of the MOBILESAT system being some twelve to eighteen months in advance of the dedicated North American systems has been recognised as an opportunity by both the Australian and international industry to utilise Australia as a test bed for a range of technical and service developments.

CONCLUSIONS

Mobile communications represents one of the fastest growing areas of the telecommunications industry within Australia. The MOBILESAT system is a vital component within the mobile telecommunications product range providing services to the economically important rural and remote areas of Australia. The activities within the MOBILESAT program are at the forefront of service developments throughout the world, tailored to the Australian market, but also with an eye to developments and opportunities throughout the world.

The MOBILESAT development team recognises that satellites will play an important role in the development and provision of ubiquitous mobile and personal communications services and that the MOBILESAT program is but one step in the evolution towards the now 'not-so-futuristic' "Dick Tracy" or "Star Trek" communicators.

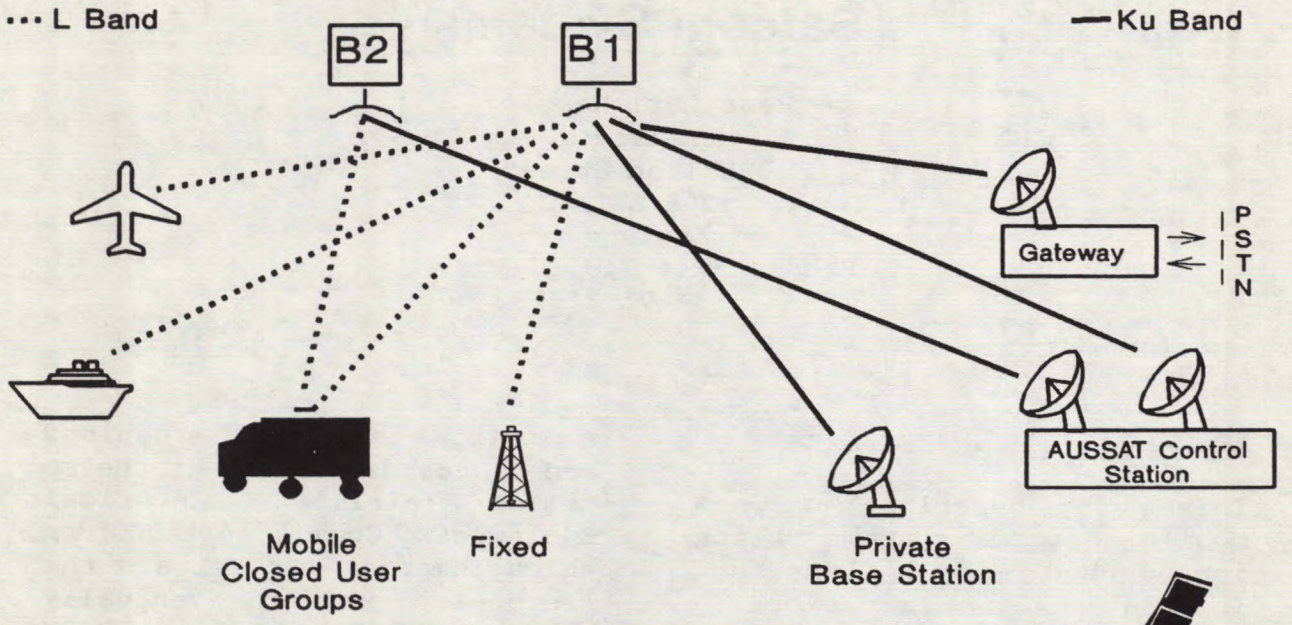


Figure 1
MOBILESAT Configuration

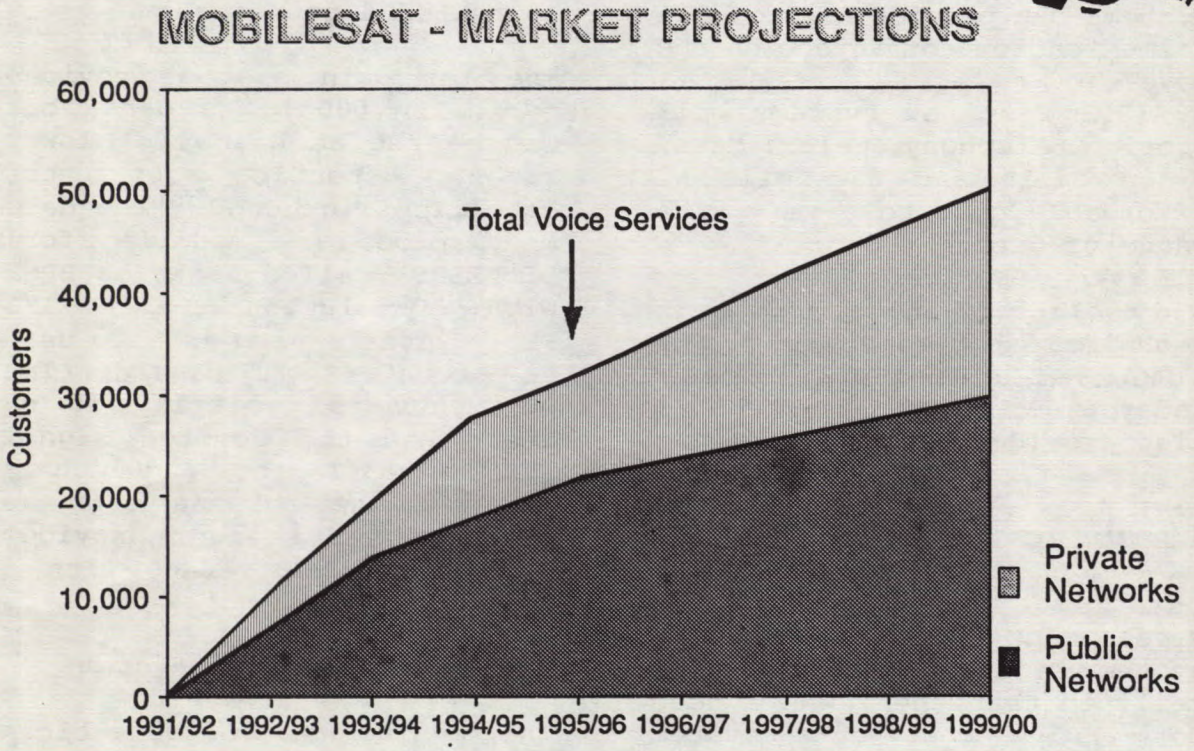


Figure 2
MOBILESAT Voice Terminal Market

Implementation of Inmarsat Mobile Satcom Systems

Hans-Chr. Haugli
Inmarsat
40 Melton Street
London NW1 2EQ
England
Phone: + 44 71 387 9089
FAX: + 44 71 383 3204

ABSTRACT

This paper describes the new mobile satcom systems being implemented by Inmarsat. Inmarsat has traditionally provided professional communication to ships and is now actively implementing new systems for use on land, in the air and at sea. These new systems can provide near global communication for anyone on the move.
By 1993 four new systems will provide telephony, telex, data, group calls and facsimile at affordable cost to a very wide range of users.

A table giving the main technical parameters for Inmarsat Aeronautical, Inmarsat-B, C and M systems is also provided.

AERONAUTICAL SYSTEMS

Inmarsat started the development of an Aeronautical system in 1985. It was soon realised that there was a need for a low cost data system using an omnidirectional antenna and a medium gain antenna for voice and higher speed data. The aeronautical community is well organised and standards are set in the so called Airline Electrical Engineering Committee (AEEC) in which Inmarsat put forward the proposals which eventually became ARINC standard 741. Further information on the Aeronautical system is provided in separate contributions.

Low antenna gain system

The low gain system provides typically 600 bit/s data, but can be set at higher or lower rates depending on the satellite link C/N0. The data modulation is a modified form of DEBPSK called Aviation BPSK with convolutional rate 1/2 FEC. Interleaving is also used to counter multipath. The multipath is typically 10 dB less than the desired signal and typical fading bandwidth is 400 Hz. The antenna gain is around 0 dBi and provides hemispherical coverage with a single antenna.

Medium antenna gain system

The medium gain aeronautical system is based on an antenna with around 12 dBi gain. This was highest gain considered practical and requires a

surface area of approximately 0.25 square metre using a phased array antenna.
This antenna supports rates up to 9.6 kbit/s used for voice, data and fax services.
Again convolutional rate 1/2 coding and interleaving is used, to conserve bandwidth OQPSK is used.
Some of the challenges with this system has been design of phased array antennas with low phase shift as function of beam steering. The design of OQPSK demodulators operating at Es/N0 of 0 dB was also required.

Protocols

The highest amount of engineering has been spent on the access control and protocols. The system can provide a wide range of bitrates in a very flexible manner. Multichannel operation can also be provided. This complexity and flexibility has been considered essential to provide a system which is likely to be used for the next 20 years.

INMARSAT-C SYSTEM

This system provides two way data at 600 bit/s from a small low cost terminal. The system can provide store and forward telex, data calls, electronic mail, position reporting, fleet management and safety services. An interesting application is Supervisory Control And Data Acquisition (SCADA) where remote control and monitoring is done via satellite.
Inmarsat-C is using a hemispherical antenna and operates in half duplex. Inmarsat-C terminals are designed for installation on both vehicles and small ships. They are also available as briefcase terminals small enough to be carried as hand luggage on aircrafts.

Optimising channel performance

The maritime channel suffers from multipath typically 7 dB weaker than the direct satellite signal and with a fading bandwidth as low as 0.7 Hz. The system is designed to provide very reliable (99 %) message transfer and this is accomplished by using long interleaving (8.64s). Modulation is CPSK using rate 1/2 convolutional encoding. To obtain good performance in landmobile channels which suffer from short, but frequent link interruption a repeated long unique word is used to establish carrier reference quickly after interruptions. The channel format does not have CW or bittransition preambles and this has been demanding for the burst demodulators.
The use of channel state information by the convolutional decoder has enabled the system to tolerate blockages up to 3 s in a 8.6 s frame. Without the channel state information only 1.5 s could be tolerated.
The system is designed for operation down to 5 degrees elevation angle, but many reports from ships shows that the system works down to 1 degree. There has even been reports of operation below -0.5 degrees. It seems that there are certain elevation angles which have high attenuation. They repeat every 0.4 to 0.5 degrees below the Brewster angle.

Testing in the landmobile environment has also shown good results. The combination of reasonably high bitrate,

interleaving, channel state information, ARQ and non directive antenna gives a good throughput even in quite bad channels.

Protocols

The largest amount of effort has been spent on the protocols. The handling of a large number of users, ARQ and store and forward buffers is a very complex task. Inmarsat equipment is made by many manufacturers and specifications can surprisingly often be interpreted different ways. The use of a pre-operational system has enabled Inmarsat to iron out many of these potential problems.

INMARSAT-B SYSTEM

The Inmarsat-B system is the second generation Inmarsat-A system used by some 11,000 ships worldwide. It significantly improves bandwidth and power efficiency and is also designed to work with the Inmarsat-3 spotbeam satellites. The system can provide voice, data, fax, group calls and telex. Both data and fax can operate at rates up to 9.6 kbit/s which much faster than through Inmarsat-A. Near toll quality voice is provided using a 16 kbit/s algorithm. The algorithm is transparent to voiceband data and fax at rates up to 2.4 kbit/s.
The antenna requirements are identical to Inmarsat-A's and this simplifies terminal design for the manufacturers.
The system is designed to provide high (99%) link availability and satisfy the requirements for the maritime distress system (GMDSS).
The narrow antenna beamwidth attenuates multipath even at low elevation angles and allows good frequency reuse between satellites.
The Inmarsat-B system will be attractive to professional maritime users who want high quality and low usage charges.
The Inmarsat-B and Inmarsat-M systems have common protocols and can share Access Control and Signalling Equipment at the fixed side, thereby greatly reducing cost.

INMARSAT-M SYSTEM

The Inmarsat-M system is designed for large new land-mobile and maritime markets and to fill the gap between Inmarsat-C and B. The system will provide communications quality voice, group calls, fax and duplex data at 2.4 kbit/s.
The largest market segment is landmobile which imposes a number of constraints. Vehicle installation limits the antenna hight and the diameter should not be much greater than 0.5 m.
Frequency re-use can be improved by having a narrow beamwidth in Azimuth. This limits the antenna gain to around 13 dBi. The maritime version uses 2 dBi higher antenna gain to provide a 95% link availability with the same forward link satellite EIRP.
The low antenna gain and the need for low usage charges dictates high power efficiency.
The limited amount of spectrum dictates high bandwidth efficiency.

Voice coding selection

Subjective voice quality is strongly dependant on the voice coding rate. As there are no real standards which sets the minimum acceptable quality for mobile satcom system connecting to the PSTN, the selection of

the bitrate was difficult. The rapid development of voice coding algorithms did not make the selection easier. Inmarsat considered seriously 9.6 kbit/s, 4.8 kbit/s, 2.4 kbit/s and ACSSB voice codecs before selecting 4.8 kbit/s. It was considered that during the system development time quality of such codecs would almost reach the quality of the 9.6 kbit/s algorithm used for Aeronautical public correspondence. An evaluation programme was set up to select the best codec and this was scheduled to be as late in the programme as possible to take advantage of the latest developments. Initially some 20 organisations planned to submit hardware for testing, of these only 8 did deliver and 5 of these were tested under a wide range of conditions. Inmarsat and Aussat joined forces during the testing which was done by Telecom Research Laboratories in Australia. The sensitivity to bit errors of the voice coding algorithm varies significantly within a voice frame. Selective FEC is used to increase robustness and to save power. The FEC increases the bitrate to 6.4 kbit/s, but maintains quality even at a 4 % bit error rate.

Optimising channel performance

The landmobile channel is characterised by short link interruptions due to blockage of the line-of-sight path to the satellite. Typical blockage duration is 50 ms, but varies directly with vehicle speed. A significant number of link interruptions last 1 second. Short interruptions are overcome by using speech prediction and repetition. This bridges gaps up to 60 - 90 ms which are in majority. Longer interruptions result in squelching.

In order to minimise the effect of blockages longer than 60-90 ms the modem reacquires very fast, typically within 60 ms for interruptions lasting up to 1 s. The signal format has short (60 ms) subframe with a CW and Unique Word preamble. A typical demodulator stores 25 ms of the signal and use different algorithms to estimate course frequency, fine frequency, phase and bit timing. As a subframe only contains 480 bits and the clock accuracy is 10E-5 the clock phase can only drift 1.8 degree. The demodulator therefore only needs to determine the best clock phase. The Inmarsat-M system can provide data and fax at 2.4 kbit/s. The objective is to provide a bit error rate of less than 10E-5 on a Gaussian channel. The data are encoded using a 3/4 convolutional code and sent twice after encoding with a time difference of 120 ms. The demodulator combines both transmissions before convolutional decoding. Interleaving is not used in order to reduce delays which is important with duplex data used interactively.

Table 1: Inmarsat systems technical characteristics

Characteristics	Inmarsat-B	Inmarsat-C	Inmarsat-M	Aeronautical	
Satellite Service Type	MM (LM)	MM/LM	MM/LM	AM	AM
Typical antenna gain (dBi)	20	1	12/14	12	0
Typical antenna example	Dish	Quadr. helix	Backf/Lin arr	Phased array	Quadr helix
Typical antenna size	1 m dia	100x 25mm cyl	0.4m d/0.5m l	0.5 X 0.5 m	100x 25 mm
MES figure of merrit (dB/K)	-4	-23	-10/-12	-13	-26
MES EIRP (dBW)	33	13	25/27	26	14
Signalling rate/mod forw link	6k/BPSK	1.2k/BPSK	6k/BPSK	600/ABPSK	600/ABPSK
FEC type/rate	Conv/ 1/2	Conv/ 1/2	Conv/ 1/2	Conv/ 1/2	Conv/ 1/2
Signalling rate/mod ret link	24k/OQPSK	1.2k/BPSK	3k/BPSK	600/ABPSK	600/ABPSK
FEC type/rate	Conv/ 3/4	Conv/ 1/2	Conv/ 1/2	Conv/ 1/2	Conv/ 1/2
Voice coding rate (bit/s)	16k APC	NA	4.8k CELP	9.6k	NA
User data rate (bit/s)	9.6k	600	2.4k	9.6k	300
Comm channel rate/modulation	24k/OQPSK	1200/BPSK	8k/OQPSK	21k/OQPSK	600/ABPSK
Interleaving time (s)	NA	8.64	(0.12)	0.04	0.67
Forw link satellite EIRP (dBW)	16	21.4	17	21	24
Comm channel C/N0 (dBHz)	49	37	42	48	36
Channel spacing (kHz)	20	5	10	17.5	2.5
Scheduled service date	1992	1990	1992	1990	1990

A Second Anniversary Operational Review of the OmniTRACS® - The First Two-way Mobile Ku-band Satellite Communications System

**Irwin M. Jacobs, Allen Salmasi,
Klein S. Gilhousen, Lindsay A. Weaver, Jr.,
Thomas J. Bernard**

QUALCOMM, Incorporated
10555 Sorrento Valley Road
San Diego, CA 92121
United States
Phone: (619) 587-1121
Fax: (619) 452-9096

ABSTRACT

A novel two-way mobile satellite communications and vehicle position reporting system currently operational in the United States and Europe is described. The system characteristics and service operations are described in detail and, technical descriptions of the equipment and signal processing techniques are provided.

INTRODUCTION

Within the past two years, a unique two-way mobile satellite communications and position location/reporting system, called the Omni–TRACS® system, has been introduced in the United States and Europe. This system, operates on a secondary basis in the 12/14 GHz Band, (Ku-band), which has been allocated to the Fixed Satellite Services (FSS) on a primary basis. Developed by QUALCOMM, Inc., the system is the first operational domestic mobile Ku-band satellite service to provide two-way messaging and position reporting services to mobile users throughout the United States. Blanket authority ("license") to construct and operate a Ku-band network of mobile and transportable earth stations and a Hub earth station was granted from the United States Federal Communications Commission (the "FCC") in February of 1989 for an application filed in December of 1987 [1]. Using a pair of Ku-band transponders on a domestic FSS satellite as described below, the system is capable of serving a population of user terminals ranging between 40,000-80,000 users depending on the average length of the messages being transmitted and the frequency of transmissions. As demand for the system grows additional Ku-band transponders can be leased from the satellite operator(s) and modularly added to the system to increase the overall capacity of the system.

The system began operational tests in January, 1988 when a mobile terminal was driven from the west coast to the east coast of the United States and back in constant communication with a Hub facility. Fully operational system providing for commercial services began in June, 1989. Terminals have since been installed in vehicles ranging from 18-wheel tractor-trailers to mini-vans, marine vessels and automobiles to bi-cycles. Operation has been very successful in all kinds of environments from wide open Western freeways to the concrete canyons of New York City.

As of June, 1990, in the U.S., over 9,000 mobile terminals are in operation using a pair of existing Ku-band transponders aboard GTE Spacenet Corporation's GSTAR-I satellite with another backlog of approximately 6000 mobile terminals scheduled for shipments. A European version of the OmniTRACS system, Eutel-TRACS®, is currently operational in Europe using Ku-band transponders aboard two EUTELSAT satellites. Typical users include those involved in public safety, transportation, public utilities, resource extraction, construction, agriculture, national maintenance organizations, private fleets, and others who have a need to send and receive information to vehicles, marine vessels or aircraft enroute. As vehicles travel across the U.S., they move out of range of

conventional land-based communication systems. This satellite-based system eliminates the range problems inherent with land-based systems in the U.S., creating a true nationwide network enabling users to manage mobile resources efficiently and economically.

SYSTEM AND SERVICE DESCRIPTION

Figure 1 shows the end-to-end system overview. The Ku-band mobile satellite communications network has three major components:

1. A Network Management Facility (NMF) for controlling and monitoring the network.

2. Two Ku-band transponders, aboard a U.S. domestic satellite located at 103 degrees West Longitude.

3. Two-way data-communication and position-reporting mobile and transportable terminals.All message traffic passes through the NMF. The NMF contains a 7.6 meter earth station including modems (the "Hub") for communication with the mobile terminals via the satellite, a Network Management Center (NMC) for network monitoring and control as well as message formatting, processing, management and billing, and land-line modems for connection with Customer Communications Centers (CCCs)[1]-[4]. The OmniTRACS NMC supports two-way data messaging, position reporting, fleet broadcasting, call accounting and message confirmation, and a host of other services.

System description

The system, as used in the U.S., uses two transponders in a single Ku-band Satellite. One transponder is used for a moderate rate of 5-15 kb/s continuous data stream from the Hub to all the mobile terminals in the system. The system users can also use the terminals in a transportable or fixed mode. Messages are addressed to individual mobile terminals or to groups of mobile terminals on this channel. The antenna utilized for the mobile terminal is vertically polarized on both receive and transmit. This necessitates the use of two transponders, one horizontally polarized and the other vertically polarized, for the mobile service. To provide for the secondary status of the return link transmissions (mobile terminal to the Hub) a combination of frequency hopping and direct sequence spread spectrum waveforms are utilized, together with low power, low data rate transmissions.

For the forward link, the system uses a triangular FM dispersal waveform similar to that used by satellite video carriers, resulting in interference properties similar to television signals but with substantially less energy outside of a 2 MHz bandwidth.

A second transponder on the same satellite is used by the return link. Each mobile terminal has a low transmit power level (+19 dBW EIRP). This power level allows data rates on the return link ranging from 55 to 165 b/s, which is dynamically adjusted depending on available link margin for each individual return link trans-

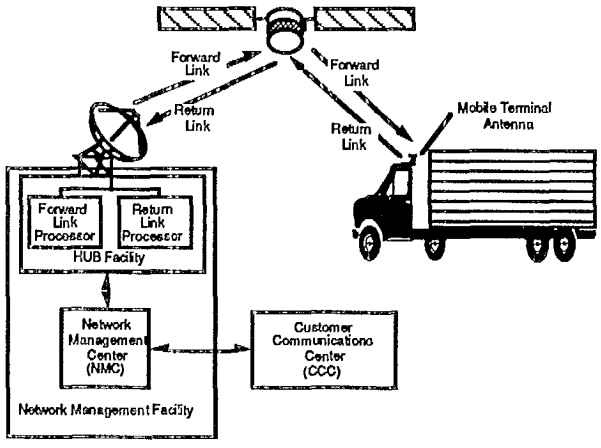


Fig. 1. Ku-Band Mobile Satellite Communications System Block Diagram

mission from the terminal. The antenna pattern of the mobile terminal is rather broad and, therefore, to mitigate for any potential for interference on the return link, several techniques are used:

1. Direct-sequence spread-spectrum techniques are used to spread the instantaneous power spectral density of each mobile uplink over a bandwidth of 1.0 MHz.

2. Frequency hopping and FDM techniques are used to ensure that the power spectral density produced by the combination of all active mobile terminals is uniformly spread over a bandwidth of up to 54 MHz, which can be adjusted to a narrower or wider bandwidth to optimize the return link throughput capacity.

3. The transmissions of the mobile terminals are very carefully controlled. A mobile terminal will not transmit unless commanded to do so, either as a direct request (acknowledgement, report, etc.) or as a response to a carefully defined—and limited—group poll. This polling technique controls the number and frequency location of mobile transmitters at all times so that the level of interference can be tightly regulated. Furthermore, reception of the command through the forward link at appropriate signal levels implies that the antenna is correctly oriented for transmission.

4. Back-up satellite frequency plans and bandwidth can be downloaded "over the air" to the mobile terminal providing for automatic switch-over to back-up transponders and/or satellites.

As a result of the above techniques, a network consisting of tens of thousands of mobile terminals causes no unacceptable interference to adjacent satellites. Initially for U.S. operations, LORAN-C derived position information in the mobile terminal was made available to the dispatcher or driver on a scheduled basis or on-demand. Alternatively, and for operations in other parts of the world, radio-determination providing for radio-location and radio-navigation satellite services may be provided either through the use of the Global Positioning System (GPS) or by direct range measurements through two or more Ku-band satellites depending on the cost of GPS receivers and availability/coverage provided by GPS as well as other Ku-band satellites. In April, 1990, QUALCOMM introduced the first provision of vehicle position location through the use of two Ku-band satellites and direct range measurements, the QASPR® system, while eliminating LORAN-C receiver circuit card from the mobile terminal, to obtain navigation accuracy of less than 1000 feet throughout the continental United States and without the error conditions prevalent in the existing LORAN-C system due atmospheric interference and the lack of radio signal coverage. The QASPR system uses the OmniTRACS TDMA timing signal formats in the forward and return link directions plus an auxiliary, low power forward link signal through a second Ku-band satellite (Ranger) to derive distance values. The distance values are then converted into the mobile terminal's latitude and longitude in real time as messages or acknowledgments are received at the NMC.

THE MOBILE SATELLITE TERMINAL

Figure 2 shows a functional block diagram of the mobile terminal. A microprocessor implements all of the signal processing, acquisition and demodulation functions. The antenna has an asymmetric pattern optimized for operations in the U.S. (approximately 40° 3 dB beamwidth in elevation and approximately 6° beamwidth in azimuth). It is steerable in azimuth only. A low-noise amplifier and conventional down-conversion chain provide a signal to the microprocessor for acquisition, tracking and demodulation. During transmission, an up-conversion and spreading chain provide a signal in the 14-14.5 GHz band to the 1.0 Watt power amplifier. This signal is transmitted via the steerable antenna that has a maximum gain of +19 dBi for a total transmit power of +19 dBW.

Whenever the mobile unit is not in receive synchronization, it executes a receive acquisition algorithm until data from the satellite can be demodulated. At this point, the antenna is pointed towards the satellite and messages can be received from the Hub. When commanded by the Hub, the mobile may start transmission of a message. The terminal is half-duplex, and transmissions are done at a 50% duty cycle to allow for continued antenna tracking of the received downlink signal. If at any time during a transmission the receive signal is lost, the terminal

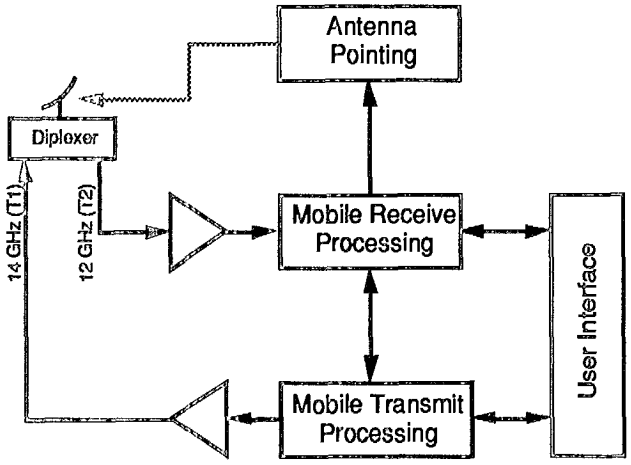


Fig. 2. Mobile Terminal Block Diagram

ceases transmission to prevent interference from being generated.

Description of return link

At the lowest return-link data rate, binary data at 55.1 bits per second is rate 1/3 convolutionally encoded to produce code symbols at a rate of 165.4 symbols per second. These code symbols are used five at a time to drive a 32-ary FSK modulator at a rate of 33.1 FSK baud. A 50% transmit duty cycle produces an FSK symbol period of 15.1 msec. The tones out of the FSK modulator are direct-sequence spread at a rate of 1.0 Megachip per second for an instantaneous bandwidth of 1.0 MHz. This 1.0 MHz bandwidth signal is then frequency hopped over up to a 54 MHz bandwidth. To maximize system capacity in areas with good satellite G/T, a 3.0x data rate of 165.4 b/s is also provided. This is implemented by a three times FSK symbol rate.

Return link power density

Figure 3 shows the mobile terminal main lobe antenna pattern in azimuth keeping elevation at maximum gain. This figure assumes a nominal boresight gain of 19 dBi. The sidelobes of this antenna are asymmetric and non-uniform, but stay below -12.0 dB relative to boresight gain. Table 1 shows the maximum transmit power density link budget for the return link with 250 units transmitting simultaneously. Table 1 combined with the transmit antenna pattern of Figure 3 produces the EIRP power density shown in Figure 4. Also shown in Figure 4 is the U.S. Ku-band guideline for inbound transmissions out of VSATs, which indicates that the aggregate EIRP (measured in dBW/4 KHz) generated by 250 mobile units transmitting simultaneously is below the U.S. inbound transmission guideline for one VSAT.

Description of forward link

Binary data at variable rates of 4960.3 to 14,880.90 bits per second is rate 1/2 block encoded to produce code symbols at rates of 9,920.6 to 29,761.8 symbols per second, respectively. These code symbols are used to drive a BPSK modulator at rates of 9,920.6 to 29,761.8 PSK symbols, respectively. A triangle wave FM dispersal waveform is then applied, resulting in similar co-ordination properties to video signals for coordinating with adjacent satellite transponders or cross-polarized co-channel transponder on the same satellite. The power density calculations are shown in Table 2.

Coordination for the forward link waveform is easier than coordination of a video signal. The frequency band in the range of ±1.0 MHz about the center of the dispersal waveform will contain relatively high instantaneous power densities, but the frequency bands outside of this ±1.0 MHz range will drop off rapidly—even faster than video signals. Figure 5 shows the suggested coordination mask for the forward link signal compared with the coordination mask suggested for video signals.

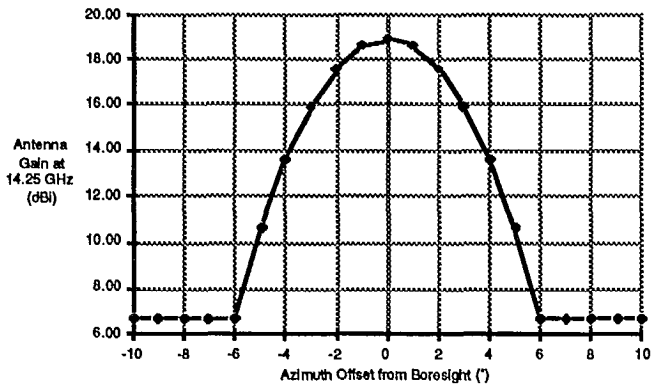


Fig. 3. Mobile Terminal Transmit Antenna Gain (Azimuth Cut)

Table 1. System Return Link Power Density

Max. Tx Power	1.26 Watt	1.0 dBw
Max. Tx Antenna Gain		19.0 dBi
Occupied BW	48,000,000 Hz	-76.8 dB/Hz
FCC Reference BW	4,000 Hz	36.0 dB-Hz/4kHz
Number of Uplinks	250	24.0 dB
Tx Duty Cycle	50%	-3.0 dB
System EIRP Density		0.2 dBW/4kHz

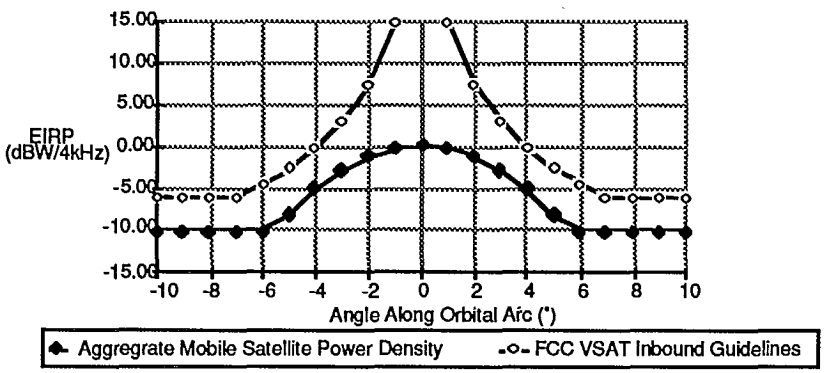


Fig. 4. Return Link Power Density Compared to the U.S. VSAT Inbound Power Density Limits

Table 2. Forward Link Power Density

Satellite Transmit EIRP		44.0 dBW
Occupied Bandwidth	2,000,000 Hz	-63.0 dB/Hz
Reference FCC Bandwidth	4,000 Hz	36.0 dB-Hz/4kHz
Transmit Power Density		17.0 dBW/4kHz

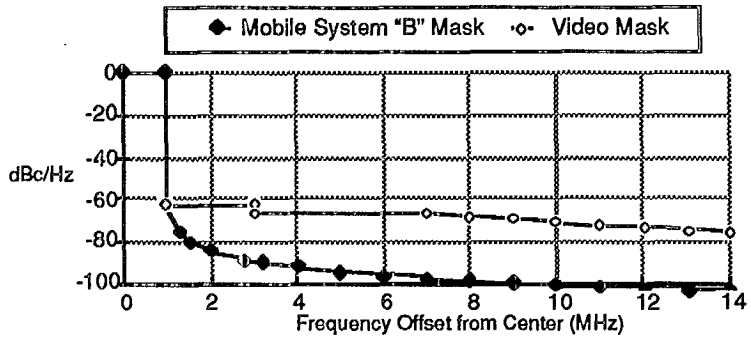


Fig. 5. Suggested Coordination Mask for the System Forward Link vs. Video Coordination Mask

SYSTEM ENHANCEMENTS

Since initial implementation of the system, the need for a number of enhancements has been recognized and the following have been implemented:

a. TrailerTRACS™—a nominal hardware addition which provides for automatic identification and position reporting whenever a trailer is attached or dropped from a power unit (e.g., tractor) equipped with the system.

b. Driver Pager—a local paging unit which is operated by the system and which notifies the driver (up to 1000 yards from the vehicle) when an important message is waiting for him in the truck—can also be used as an alarm for "wake-up".

c. Message Return Receipt—provides the ability to have the system notify central operations or dispatch whenever a message has actually been read--analagous to certified mail with a return receipt.

d. Panic Button—provides the ability for a driver to hit one key or button, thereby initiating a "panic" message which is intercepted by the Network Management Center, causing a telephone call to be initiated to appropriate personnel to inform them of serious trouble at the location of the truck so that authorities can be dispatched.

SYSTEM BENEFITS

More than 50 truckload motor carriers and others have implemented the OmniTRACS system in part or all of their fleets and approximately 100 firms are currently evaluating the system. System benefits currently being experienced are: a) reduced telephone and other telecommunications expenses, b) increased revenue miles, c) reduced "deadhead" miles, d) reduced driver turnover, e) reduced accidents, and f) increased dispatch efficiency.

REFERENCES

1. **The U.S. Federal Communications Commission (FCC)** February 14, 1989. Memorandum Opinion, Order and Authorization in the Matter of QUALCOMM, Inc. Application for Blanket Authority to Construct and Operate a Network of 12/14 GHz Transmit/Receive Mobile and Transportable Earth Stations and a Hub Earth Station.

2. **Antonio, F.P.,Gilhousen, K.S., Jacobs, I.M., Weaver, Jr., L.A.** 1988. Technical Characteristics of the OmniTRACS - the First Operational Mobile Ku-band Satellite Communications System. Proceedings of the Mobile Satellite Conference, JPL Publication 88-9, Pasadena, California, pp. 203-208.

3. **Salmasi, A.** 1988. An Overview of the OmniTRACS—the First Operational Mobile Ku-band Satellite Communications System. Proceedings of the Mobile Satellite Conference, JPL Publication 88-9, Pasadena, California, pp. 63-68.

4. **Bernard, T. J.** February, 1990. Logistics Benefits From Two-Way Satellite Tracking. Defense Transportation Journal.

Domestic Mobile Satellite Systems in North America

Muya Wachira
Telesat Mobile, Inc.
P.O. Box 7800, Ottawa, Canada
K1L 8E4
Phone: 613-756-5601
FAX: 613-746-2277

ABSTRACT

Telest Mobile Inc.(TMI) and the American Mobile Satellite Corporation (AMSC) are authorized to provide mobile satellite services(MSS) in Canada and the U.S.A. respectively. They are developing compatible systems and are undertaking joint specification and procurement of spacecraft and ground segment with the aim of operational systems by late 1993. Early entry (Phase 1) mobile data services are offered in 1990 using space segment capacity leased from Inmarsat. This paper gives an overview of these domestic MSS with an emphasis on the TMI component of the MSAT system.

INTRODUCTION

Canada and the United States have long recognized the need for reliable mobile communication services to all areas of the continent.

There is an extensive need for improved and additional land mobile service communication capabilities. In both Canada and the United States, there are large coverage gaps in services provided by terrestrial systems.

There is also a need for improved aeronautical mobile communication services in all areas of North America and adjacent over-water areas. These needs extend to both safety and non-safety communications, and are analogous to the land mobile service needs discussed above. The ICAO Future Air Navigation Systems (FANS) Committee envisions aeronautical mobile safety communication evolving to the use of satellite services thus allowing a significant reduction of the present ground infrastructure.

Canada and the United States have begun initiatives to develop compatible mobile satellite service systems to serve aeronautical, land, and maritime users. These system developments are being pursued by Telesat Mobile Incorporated (TMI) and the American Mobile Satellite Corporation (AMSC) in Canada and the United States, respectively. The objective of these developments is to have a compatible overall system in which there are separate but similar spacecraft operated respectively by TMI and AMSC, with mutual backup. The companies are in the contract negtiation stage for the spacecraft procurement and have issued an RFP for the specification of the ground segment. These systems are planned to be in operation from late 1993.

SYSTEM DESCRIPTION

User Requirements

From the several market studies carried over the past few years to identify the types of services are required and the size of the market for a wide area mobile satellite system, the services identified include the following:

(1) Mobile Radio Service (MRS) which provides voice and data communications between a mobile terminal and a base station or between mobile terminals in a closed user group.

(2) Mobile Telephone Service (MTS) - a full duplex voice communications service which provides direct access for mobile terminals to the Public Switched Telephone Network (PSTN).

(3) Mobile Data Service (MDS) - a service which provides a communications path for the bi-directional transfer of data between a mobile

terminal and a data hub station in a packet-switched mode.

(4) Aeronautical Service may be provided by some of the above categories, but it has its own international system definition (as described by ICAO's FANs committee) and is therefore categorized separately. Applications include Air Traffic Services (ATS), Aeronautical Operational Communications(AOC),Aeronautical Administrative Communications (AAC), and Aeronautical Public Correspondence(APC). Over Canada and the U.S., satellite aeronautical communications will be over MSAT; however, aircraft will be able to switch to international space segments.

Other services, such as Position Location Service, Wide Area Paging and Transportable Communications Services, will be provided as the needs are better defined.

Some services, such as those pertaining to safety of life and regularity of flight for aeronautical services(ATS and AOC) or distress and safety for maritime services, require higher priority than other communication services.

TMI estimates the potential demand for MSS services at 300,000 to 450,000 mobile terminals in Canada by the year 2000. Data is expected to constitute 50% to 60% of this potential demand. Personal communication via satellite is expected to start in the next few years and possibly to become the dominant market beyond the year 2000.

System Configuration

The service coverage area is the whole of the North American land mass and coastal waters up to the 300 km limit plus the islands of Hawaii and Puerto Rico. Optional coverage includes Mexico and the Carribean, and the international flight information regions controlled by Canada and the U.S.A.

The system has been designed to serve a large number of users, each with low activity of an intermittent nature. The satellite resources (channels) are therefore best utilized by assigning them on a demand basis. The demand access capability is satisfied by a Priority Demand Assignment Multiple Access (PDAMA) system residing in the Group Controller(GC). Mobile Earth Terminals (METs) can be connected among themselves, to base Feederlink Earth Stations (FESs), PSTN or data network subscribers (via a gateway FES), hence providing a flexible communication network.

The main elements of the system are:

(1) The space segment which includes the satellite and the Satellite Control Centre (SCC).

(2) The ground segment consisting of FESs(gateway, base, and hub stations), and METs.

(3) The Network Control System(NCS) which includes the Network Operating Centre(NOC) and Network Control Centre(NCC) housing the Group Controller(GC) .

(4) A Signalling system to interconnect the elements of the system.

Communication with other networks will be done by the NOC.

The configuration of the MSAT system is shown in Figure 1.

The GC can allocate circuit capacity on a per call basis for voice and circuit-switched data. It can allocate a channel (or channels) to be used for packet-switched data under the control of the data hub. A MET typically places a call on a signalling channel (request channel). The GC uses a signalling channel (assignment or control channel) to assign a free communication channel to that MET and a FES. At the termination of the call the terminal relinquishes the communication channel which then becomes free for reassignment by the GC to another call. For AMS(R)S service frequency and power capacity will be allocated in distinct portions which will be used in channels consistent with ICAO SARPS.

The MSAT system will have multiple spot beams in order to increase the satellite EIRP, and to allow for frequency reuse since the available spectrum is limited.

Forward links from FESs will use SHF uplinks to the satellite and L-band downlinks from the satellite to the METs. In the other direction, reverse links from the METs will use L-band uplinks to the satellite and SHF downlinks from the satellite to the FESs. Thus all commun-

ications to and from the METs are at L-band (uplinks in 1631.5-1660.5 MHz and downlinks in 1530-1559 MHz, as allocated by WARC-MOB'87), while all transmission to and from the FESs and the network control centre will use 13/11 GHz uplinks and downlinks. In Canada 13 GHz band is allocated on a co-primary basis to Very High Capacity Microwave (VHCM) links. The 11 GHz band is also used by fixed terrestrial services. Therefore, effective use of SHF spectrum is also very important.

In the normal mode of operation, the uplink signals from 1.6 GHz band beams will have their frequencies translated at the satellite into the SHF downlink beam. Similarly, the uplink SHF channels will be translated in frequency to the 1.5 GHz band downlink. As well, a segment of these SHF channels will be cross-strapped to form a direct SHF-SHF link. These SHF-SHF links are intended mainly for signalling, data exchange among gateways, and for NCS communication purposes.

PROVISION OF SERVICES

The system operators have the capability to operate as common carriers, owning and operating the satellites and NCSs that provide priority demand assigned multiple access to satellite power and spectrum resources as explained below. As the single point switching and control centre, the NOC/NCC will assign these resources to the multiplicity of users in a flexible manner. End users could obtain service from service providers who may be institutionally independent of the space system operator.

The MSAT system will have " variable dynamic spectrum partitioning" capability which will allow flexibility in assigning spectrum to different services. This capability will allow the assignment of exclusive spectrum for aeronautical safety communication services which will be varied depending on need. In the event that peak AMSS (R) safety traffic is anticipated to exceed then allocated bandwidth, additional bandwidth and power will be made available on a realtime pre-emptive basis. The aeronautical operators will notify the MSAT NOCs ahead of time of their requirements to make sure that resources are available when required without having to preempt other services except in extreme situations.

Authorities responsible for aviation safety will arrange with the system operators for the operational capacity they require and will set up their own network using their own signalling characteristics, protocols and control channels. Their NOC/NCC(s) may operate independently of the MSAT NCS.
Air Traffic Services (ATS) which includes air traffic control, and Aeronautical Operational Control (AOC) are the only communications functions that comprise AMSS(R) with the right of priority and immediate preemptive access. Other aeronautical communications, Aeronautical Administrative Communications and Airline Passenger Correspondence have the same status as similar non-aeronautical communications. The control capability of the GCs will also enable priority access for non-aeronautical safety and emergency communications.

Communications other than aeronautical safety communication services will be provided to end users through a variety of service provider arrangements under the control of the NOC.

An MSS telephone subscriber may be served by his local telephone company or radio common carrier who may own a FES at the telephone exchange. When a call is placed by the vehicle or to the vehicle, the request will be relayed from the vehicle or FES through the satellite to a GC on a request channel. The GC will use the request channel to assign a communication channel pair to the vehicle and the FES. The call will then be set up by the exchange. When the "on hook" signal is received at the GC over the request channel, the communication channel pair will be returned to the pool of available channels for reassignment to another call.

A government agency, utility or large trucking company may set up its own network of feeder link earth stations for its direct call setup and communications with its vehicles. Operation of the network will require lock to the control channels of the GCs in order to comply with the requirement for priority and preemptive access for essential safety communication services.

SYSTEM PERFORMANCE REQUIREMENTS

For optimum utilization of the spectrum, most communications will be established using Single-Channel-Per-Carrier (SCPC), Frequency

Division Multiple Access (FDMA) with nominal 5 kHz channel spacing as compared with the wider 30 kHz channels adopted for most terrestrial systems. This 5 kHz channel requirement places a constraint on the possible speech coding/modulation techniques that can be used in MSAT. The baseline analog voice modulation scheme considered by TMI is Amplitude Companded Single Sideband (ACSSB) and the baseline digital version is 4.8 kbps Codebook-Excited Linear Predictive Coding (CELPC) using 16-QAM Trellis Coded Modulation. Other suitable coding/modulation techniques may be employed.

Voice performance is dependent on the inherent performance of the modulation schemes used, propagation characteristics of the links, noise and interference, the practical limits of the satellite RF power per voice channel as well as the practical limit of the earth terminal G/T's that can be achieved considering the dimensions of the vehicular antenna. Mobile Radio Service (MRS) voice channels will be designed to operate with an unfaded carrier-to-noise density (C/No) of around 51 dB-Hz at the L band. For users who do not experience shadowing, e.g. for aeronautical and marine applications, the C/No requirements may be relaxed by a few dBs.

For an MRS terminal with a G/T of -15 dB/K, these voice channel performance requirements translate into a satellite EIRP of around 30 dBW per voice channel for ACSSB, as shown in Table 1.
The G/T performance of the satellite at L-band has a signicant impact o the coast of the mobile earth terminals. A reasonably high satellite G/T permits the operation of the METs will resonable RF power levels. The MSAT satellite L-band G/T is specified at 2.8 dB/K.

Capacity of a first generation Canadian system should be high enough to serve, during the peak traffic period, at least 60 thousand equivalent voice users. Assuming voice activation with an activity factor of 0.4, traffic of 0.0106 Erlang per average user, and a grade of service of 2%, the 60,000 users translate into approximately 710 assigned channels, for voice traffic. Of course, the traffic will not be homogeneous--various channels and blocking rates will be used.

Since the actual distribution of users may not be known until some considerable time after the start of service, the system has to be able to accommodate a concentration of users in some parts of coverage area. The space segment is being designed to allow redistribution of system capacity.

For the land mobile voice channel performance requirements, some tests have indicated that for ACSSB, the 51 dB-Hz may degrade to around 40 dB-Hz before the MRS channel becomes unusable. With the given noise budget, this translates into a downlink margin of about 10 dB. Signal level degradation for mobile terminals operating in the 1.5/1.6 GHz band is mostly due to multipath and blocking by obstacles, such as trees, and a full characterization of these will not be available until some time after the start of the service. The link margin required for a given grade of service may therefore have to be modified after the launch of the satellite to accommodate the blockage characteristics of a particular service area. Hence the space segment design will permit operation with some level differences among the active channels. Some places in the coverage area enjoy relatively high elevation angles and/or have little or no shadowing. For these areas, the link margin requirements could be reduced, hence lower satellite EIRP per channel would be needed. Further, other modulations requiring less EIRP's may be developed in the future.

MSAT SPACE SEGMENT

The size, power, and configuration of the MSAT satellites are determined by the performance requirements of the system. The spacecraft outlined in here is only one of several configurations being considered. It is taken as a baseline for system capacity and cost estimation.

The baseline configuration provides the coverage of the North American service area at L band by five beams with an Edge Of Coverage (EOC) gain of 32 dBi. Two other beams are inluded for Hawaii, Puerto Rico and Mexico. See Figure 2. Other configurations considered have as many as 11 beams. A single SHF transmit/receive beam serves to cover the combined Canadian-US service area, with an EOC gain of 25 dBi (Figure 2). The MSAT aggregate EIRPs for the basic area are approximately 55 dBW and 36 dBW for L-band and SHF, respectively.

To accommodate a variety of users with different requirements, hence different modulations, power levels, etc., the satellite will have wideband linear transponders i.e. there will be no channelization at the satellite. Intermodulation performance of the transponder is a key requirement. A minimum of 22 dB C/Im is required for the fully loaded transponder. Another important parameter is the interbeam co-channel isolation which determines the minimum frequency reuse factor achievable. Co-channel C/Is of 20 dB in the forward link and 22 dB in the reverse link have been specified.

Because the same 1.5/1.6 GHz spectrum is sought by many administrations and international organizations, there is no guarantee that the satellite will have a contiguous segment of spectrum to meet its needs. It is most likely that several pieces of spectrum will be available all across the band after coordination with other users of the band. The system therefore, should have enough flexibility such that any piece of spectrum in the band can be switched into any beam. This permits an assignment of system capacity in multiples of a basic unit of spectrum among the beams, by means of a ground commandable switching network.

The main space segment characteristics are summarized in Table 2.

MSAT GROUND SEGMENT

The major components of the ground segment are the Feederlink Earth Stations (Gateway, Base, and Data Hub) and the Mobile Earth Terminals.

The Gateway FESs will interface with the PSTN to allow for communication between users and the PSTN subscribers. The gateway FES will support voice and data transmissions. For PSTN data services, such as Data-Route and Datapac, protocol translators will be required at the gateways. The gateways will operate in the 13/11 GHz bands and they will be located at sites across Canada to ensure optimum usage of the terrestrial and satellite networks aimed at minimizing user cost.

The Base FESs are defined as the dispatch centres or private fixed earth stations for user groups subscribing to MRS or other private services. It is assumed that every MRS user group will consist of at least one Base FES and a number of mobile earth terminals. It is expected that up to 95 percent of MRS traffic will be between Base FESs and mobile terminals,and only 5% will be MET-to- MET. The base FESs will operate in the 13/11 GHz bands.

The Mobile Earth Terminals will be under complete control by the GC. They will operate primarily in the full-duplex mode, and the transceivers will have a frequency agile design. The terminal will operate in the 1.5/1.6 GHz band. It will have a DAMA processor which will enable it to receive and respond to instructions from the GC communicated via the signalling channels. For positive control, the MET will have a signalling channel receiver (SCR) separate from the main communication channel receiver so that it will receive the signalling channel at all times (even when transmitting). This offers several advantages including:

- the availability of a continuous signal for mobile antenna tracking and power control;

- positive derivation of frequency from a common frequency reference;

- access denial for terminals not meeting performance specifications;

-support of specialized network features such as display of calling number, call waiting etc.

- the means to continuously control the MET for reassigning to open channels or for preemption.

The MET will not be able to transmit unless it is locked to the control channel. An option that may be required of some classes of METs will be the incorporation of built-in sensors that measure key performance parameters that could be read out on command over the control channel by the GC.

TMI expects that the majority of METs will be integrated voice and data terminals. That is a single terminal will be capable of providing a combination of mobile voice, mobile telephone, and mobile data services

Antenna gain of the mobile terminals has a direct effect on one of the most cost-sensitive parameters of the system - the downlink(satellite) EIRP. Design and qualification of a cost-

effective high gain vehicular antenna which is acceptable by the user community from a cost, performance, and an esthetic point of view is therefore of considerable importance.

NETWORK CONTROL SYSTEM

The Network Control System will be the single point control for all system operating functions except aviation safety. The functions include:

System frequency reference
Communication channel assignment
Priority assignment of channels
Network performance monitoring
User equipment performance monitoring
Turn-off of malfunctioning user equipments
Preemption of bandwidth needed for safety reassignment
Recording of time, bandwidth and power used per call
Recordkeeping and billing.

The NOC will be responsible for overall supervision, maintenance and long-term resource planning of the system. It will communicate directly with the GCs and the Aviation NOCs in order to receive information concerning traffic statistics and AMS(R)S needs. The NOC will compute the partition between voice, data, and signalling channels to optimize performance and will execute the partitioning required by the aviation safety system. Furthermore, it will be responsible for optimizing the routing tables used by the GC to route the MTS calls to the proper gateway stations. This information will be transmitted to the GC. The NOC will also gather billing information from the GC and the gateways.

The Group Controller coordinates, controls, and maintains the activities within different Control Groups, with particular emphasis on the short-term activities and events. The principal objectives of the GC are to achieve maximum channel availability in performing these functions, and in conjunction with the NOC to make the best use of the system resources. An efficient signalling system is required which can use the minimum number of channels while still providing an acceptable performance, such as call set up times. The GC will also transmit frequency and timing information to other network entities.

The Canadian GC(s) will operate a network consisting of approximately 80,000 to 160,000 users at the point of system saturation for the first generation system. It will be connected to all FESs and METs via packet-switched signalling channels. The United States GCs must accommodate an even larger number of users and will have functionally similar signalling channels.

All communication control functions associated with the network, such as call processing, must be performed by a set of highly reliable fault-tolerant computers compatible with the other network elements to ensure maximum system reliability.

The Network Control System is described in detail in companion papers elsewhere in this conference.

PHASE 1 MSS SYSTEMS

Phase 1 Mobile Data Services will be introduced in 1990 by both TMI and AMSC, using leased space segment capacity, in order to build up a customer base for future MSAT services and to gain experience in the operation of a domestic MSS system. TMI's system, which went into operation in May 1990, is mainly a two-way store and forward digital messaging system. It is capable of collecting regular position reports(e.g. every 15 minutes) and exchanging a number of messages each day with up to 6,000 vehicles in the initial implementation. Details of TMIs phase 1 MDS system are given in a companion paper in this session. The AMSC Phase 1 MDS system, commencing later in 1990, is also described in a paper in this session. TMI is also planning on a Phase 1 Mobile Voice System using Inmarsat's space segment beginning in 1991.

CONCLUSION

This paper has described the baseline concept of the MSAT system that will provide a variety of mobile satellite services beginning late 1993. The North American MSS systems are being developed jointly by TMI of Canada and AMSC of the U.S.A. Procurement of the space segment hardware will commence in mid 1990, and that of the ground segment hardware in mid 1991. The Phase 1 MSS, offering mobile data services via leased space segment, started in May 1990.

PARAMETER	UNIT	FORWARD LINK	REVERSE LINK
UPLINK		4.7m FES Ant. to Mobile	Mobile to 4.7m FES Ant.
Frequency	MHz	13200	1650
Uplink EIRP/Voice Act. Carr.	dBW	47.2	14.5
Path Loss	dB	206.8	188.7
Satellite G/T	dB/K	-3	2.8
Req'd Flux Density/Voice Carr.	dBW/m2	-115.7	-148.4
C/No Thermal	dB-Hz	66.0	57.2
DOWNLINK			
Frequency	MHz	1545	11300
Req'd EIRP/Voice Act. Carr.	dBW	30.0	8.6
Path Loss	dB	188.2	205.8
Receive Terminal G/T	dB/K	-15.0	28.2
C/No Thermal	dB-Hz	55.4	59.6
INTERFERENCE (C/I)			
Intermod & Energy Spread			
Uplink	dB	32	30
Downlink	dB	22	22
Interbeam Co-channel			
Uplink (Incl. interstitial spac.	dB		27
Downlink (Inc. interstitial sp	dB	25	
Other Sources		23.5	22.7
Total Interference (C/I)	dB	18.4	18.3
Total C/Io	dB-Hz	53.1	53.1
Total Unfaded C/No+Io	dB-Hz	51.0	51.0

Table 1. Reference Circuit Link Budgets for ACSSB Voice

Orbit location	106.5°W, Canadian 62°W, 101°W, 134°W U.S.
Satellite class	2,500 kg
Payload mass	350 kg
Payload power	2.0 kW
Frequency bands	RX 1631.5-1660.5 MHz TX 1530-1559 MHz
Transponder bandwidth	29 MHz
Number of beams	5 at L-band (+option for Mexico) 1 North American SHF
Polarization	RHCP for L band linear VP Receive HP Transmit for SHF
Antenna size	two 5m reflectors for L-band
EOC G/T	2.8 dB/K 1.6 GHz -3 dB/K SHF
Eclipse protection	50% of busy-hour traffic
Service life	10 years minimum

Table 2. Space Segment Characteristics

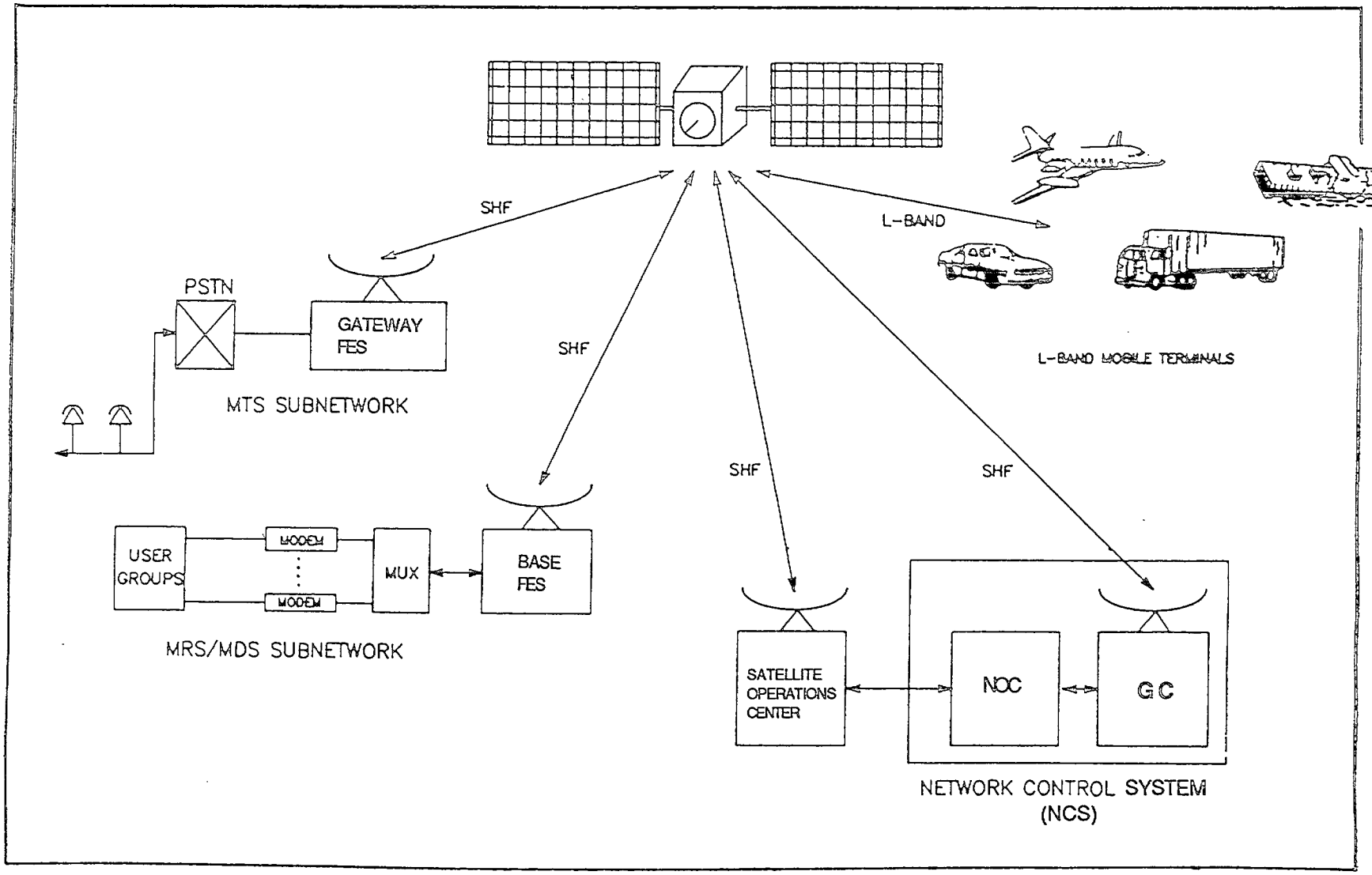


FIGURE 1: SYSTEM CONFIGURATION

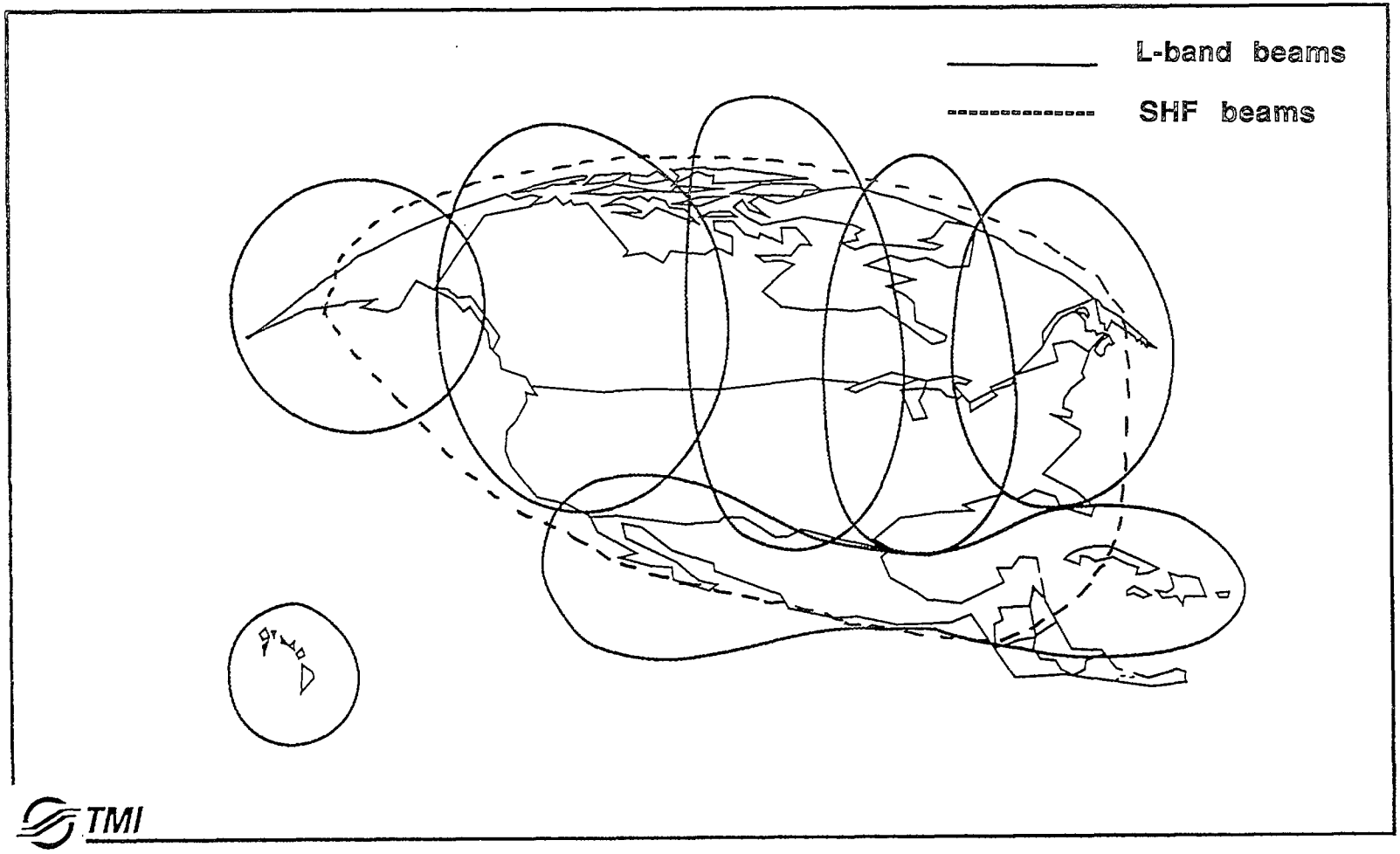


Figure 2 L-band beam coverage pattern

The American Mobile Satellite System

W. B. Garner
American Mobile Satellite Corporation
1233 20th Street NW
Suite 301
Washington, DC 20036
Phone: 202-331-5858
FAX: 202-331-5861

During 1989, American Mobile Satellite Corporation, (AMSC), was authorized to construct, launch and operate satellites to provide Mobile Satellite Services to the US and Puerto Rico. AMSC has undertaken three major development programs to bring a full range of MSS services to the US. The first program is the space segment program. It will result in the construction and launch of the satellites, and the construction and installation of the supporting ground telemetry and command system. The second program is the ground segment program. It will result in the specification, design, development, construction and installation of the Network Control System necessary for managing communications access to the satellites, and the specification and development of ground equipment for standard circuit switched and packet switched communications services. The third program is the Phase I program. This program will provide low-speed data services within the US prior to availability of the AMSC satellites and ground segment. This article describes the present status and plans for these three programs, and also provides an update on related business arrangements and regulatory matters.

SATELLITE PROGRAM

AMSC has been authorized by the Federal Communications Commission to construct, launch and operate three satellites for the purpose of providing general mobile satellite services to all fifty States, Puerto Rico and the Virgin Islands. The technical characteristics of these satellites are described in AMSC's application to the FCC of February, 1988 and the Modifications to the application submitted in December, 1989 at the FCC's request. The significant submittals in the Modification were a detailed transponder plan which included details on channelization by means of sub-band filtering, a complete set of antenna patterns for all three orbit positions, detailed link budgets for a wide range of services, and a description of how priority and preemption would be provided for aviation safety and regularity of flight services through the Network Control System. Authority was requested to construct the satellite to be able to operate in the bands 1530-1545 MHz and 1631.5-1646.5 MHz in addition to the bands 1545-1559 MHz and 1646.5-1660.5 MHz authorized by the FCC. This request was made in order to be able to provide backup services to Canada. Authority was also requested to operate single polarization in 200 MHz of feeder-link

spectrum rather than dual polarization in 100 MHz of spectrum as authorized. This request was made to eliminate a serious uncontrollable cross polarized interference situation caused by foreign systems illuminating the spacecraft at L-band.

It is expected that a satellite contract will be let by mid 1990, with first delivery in time for a late 1993 launch. Operations are planned to begin the first quarter of 1994. Options for additional spacecraft will cover failures and growth needs. The Canadians will launch a similar satellite in the same time frame so that backup capability will be in place.

NASA will provide the first AMSC launch under a barter agreement. In return for the launch, AMSC will provide an equivalent value in services to US government agencies during the first two years of operation. NASA has already put in place a program to involve government agencies as users of the services.

GROUND SEGMENT PROGRAM

The first objective of the ground segment program is to specify, design, develop and bring to operational readiness a ground system capable of providing standardized circuit switched and packet switched services by the time the satellites are ready for operation. A second major objective is to have available a supply of Mobile Earth Terminals (METs) and Feeder-link Earth Stations (FESs) and service providers by the time operations commence. A third major objective is to have in place the Network Control System that will manage all communications access to the AMSC satellites. In order to achieve these objectives, AMSC, in cooperation with TMI, has established a two stage development program for circuit switched services and the Network Control System. In addition, a program is being established for the development of standard packet switched services.

Network Control System and Circuit Switched Services.

The Network Control System and Demand Period circuit switched services are being developed under the same program since they are closely related. The Network Control System provides management of all communications access to the AMSC satellites. Physically it consists of a Network Operations Center (NOC) and two Network Control Centers (NCC). The NOC provides top level operations, administrative and network management services while the NCCs provide real-time management of Demand Period circuit switched services. The NCCs interact with the METs and FESs using Demand Period circuit switched services by means of satellite signalling channels designed for that purpose.

The NCS and circuit switched service standards are being developed under a two stage contracting program. The first Design contract is for the system definition and technical specification for the NCS, METs and FESs supporting circuit switched services. The RFP for this contract was released in April 1990, with the goal of being under contract by the end of June, 1990. The contract is planned to be completed by the 2nd quarter of 1991. Based upon the specifications developed under the Design Contract, a Construction Contract will be let by mid 1991. This contract will result in the design, construction, installation and operational readiness of the NCS. It will also result in the availability of METs and FESs capable of providing circuit switched services to the standards developed under the first contract.

As part of these programs, the specifications for METs and FESs will be available for review and comment as they are being developed. Two Industry Participation reviews will be held during the Design Contract period to familiarize manufacturers and service providers with the specifications and to obtain comment. The resulting specifications will be issued as AMSC Technical Standards at the end of the first stage, and updated as the second stage development program proceeds.

Packet Switched Services

Packet Switched service standard specifications will also be developed. These services will include end to end two-way Virtual Circuit packet services of general application, including both switched virtual circuits and permanent virtual circuits. Message store and forward service will be included, building on the low speed data capabilities of Phase I. Currently the functional requirements for packet switched services are being defined. From these functional requirements the detailed system design and technical specifications will be developed. These specifications will be available as AMSC Technical Standards. A contract will be let for the development, construction and implementation of a set of equipment to support these Standards in time for service when the satellites are operational.

PHASE I SERVICE

The AMSC satellites will be available for service in early 1994. Recognizing the long time gap between the present and that date, AMSC has inaugurated a program to supply a low speed, two way, store and forward message service in the interim. The purposes of this program are several fold. First, the program will permit the Corporation to develop a customer base prior to full service. Second, it will provide a means for understanding customer needs and aid in development of products to meet those needs once the full service is available. Third, it will provide a means for AMSC to gain hands-on operational experience of an operating system prior to going operational with the much larger, more complex AMSC satellite system.

Phase I Service is a two-way, low speed store and forward messaging service between omni-directional mobile terminals and an AMSC owned and operated Network Operations Center located in Washington, DC. Users fixed site data equipment will access the NOC via dial-up lines or via public data network services. Messages to mobiles are delivered at once, either to individual mobiles or to groups of mobiles. Messages from mobiles may be generated at random, in response to polls from the NOC, or on a scheduled basis. Mobiles may also transmit emergency messages that are given delivery priority by the NOC. Mobiles will have a Data Terminal Unit for entering and receiving messages in a variety of ways. An auxiliary data port is available for connecting other asynchronous devices. A Loran-C position location receiver is offered as an option.

Service will be provided via space segment leased on the MARISAT spacecraft located at 106.5 degrees West longitude. The COMSAT Southbury, Connecticut INMARSAT earth station will be used to access the satellite. The coast earth station and the AMSC NOC will be interconnected via leased lines. Forward direction transmissions will be via one or more packet multiplexed, continuous, 1200 sps BPSK channels.

Return transmissions will be via several Slotted Aloha or Reservation TDMA 600 sps BPSK channels. The Service will be transferable to the AMSC satellites without modification of the mobile terminals or changes in user operations. The service will continue to be supported after full service commences. The mobiles and the NOC will have an INMARSAT-C interoperable mode for backup purposes. This mode will be such that transfer can take place automatically and transparently to users of the system.

A contract has been placed for construction of the NOC, coast earth station equipment and an initial quantity of mobile terminals. The system will begin initial operation in November, 1990 and be in full service by the second quarter, 1991. Mobile terminal specifications will be available for use by other manufacturers in that time frame.

BUSINESS AND REGULATORY MATTERS

AMSC now has eight shareholders, each of which has contributed 5 million dollars in equity to the Corporation. The firm of Donaldson, Lufkin and Jenrette has been retained to help obtain additional financing needed to construct the system. AMSC and TMI have entered into a Joint Operating Agreement that provides for mutual satellite backup services, cross-leasing of satellite capacity and joint development of core communications ground equipment standards and specifications to promote commonality and interoperability. AMSC and NASA are working on an agreement for the barter exchange of services for a launch.

On the regulatory side, the MARISAT lease for Phase I Service has been approved under a Special Temporary Authority. Several items were pending at the time this article was submitted in April. Among these items is final approval of the satellite license and approval of the Phase I Service earth stations applications. AMSC has also applied for authority to operate in the lower portion of the mobile satellite bands. This application has been deferred pending an FCC Notice of Proposed Rule Making that would allocate these lower bands to generic MSS services domestically, similar to the allocation in the upper portion of the bands.

Access to additional spectrum is viewed as a must to meet long term, world-wide projections of service demands. The totality of projected needs is well in excess of that currently available. Because of the low discrimination of mobile terminals, frequency reuse by orbit separation is severely limited. Further limiting the available spectrum is the fragmentation caused by international allocations based on type of service (Land, Maritime, Aeronautical). This fragmentation results in inefficiency by leaving unused spectrum in one band while other bands saturate. The result of all these limitations makes sharing of spectrum difficult and an equitable distribution of the spectrum between systems a major issue.

The 1992 World Administrative Conference (WARC 92) offers an opportunity to improve the situation. The adoption of generic allocations would remove the fragmentation by service type limitation. Additional allocations for generic MSS should be vigorously pursued, with MSS system operators, service providers, manufacturers and users working together to make those responsible for spectrum allocations aware of, and supportive of, MSS needs.

Mobile Data Services

David J. Sward
General Manager, National Systems
Telesat Mobile Inc.
333 River Road, P.O. Box 7800
Ottawa, Ontario K1L 8E4
Canada
Tel: (613) 746-5601 Fax: (613) 746-2277

SYNOPSIS

Telesat Mobile began commercial operations on December 15, 1988 with the signing of a shareholder agreement between Telesat Canada, Canadian Pacific Ltd. and the C. Itoh Group of Japan. TMI's mission is to construct and operate a commercial mobile satellite system in Canada. This will be done in two distinct phases.

Firstly, TMI will introduce mobile data services in May of 1990. A contract has been awarded to Canadian Astronautics Ltd. supported by Gandalf Systems Ltd. and Hughes Network Systems for the supply of hub equipment and 3000 mobile data terminals. Over-the-satellite tests will commence in January 1990. The mobile data service will provide full two-way digital messaging, automatic vehicle location, and fleet management services.

The second phase is to construct, launch and make operational the MSAT satellite and associated network control facilities. This paper will focus on Phase I, i.e., the implementation of a mobile data service in Canada. In addition to a technical description, the paper will provide information on markets and applications.

The full text of the paper is in the Appendix ..... page A-15

Implementation of Geostar's RDSS System

Ronald J. Lepkowski
Geostar Corporation
1001 22nd Street, N.W.
Washington, D.C. 20037 U.S.A.
Phone: (202)-778-6008
FAX: (292)-223-6155

ABSTRACT

The Geostar® system began its initial operations in 1988 and was the first domestic satellite system to provide regular service to mobile users within the United States. This paper provides an overview of Geostar's radiodetermination satellite system (RDSS) concept and its development by Geostar, with a focus on the current operational status of Geostar's interim RDSS system and services.

INTRODUCTION

Geostar was created in 1983 to develop RDSS. Mobile satellite systems (MSS) are designed to provide a communications link between a mobile user and base earth station. Such MSS links can be used to provide position reports based on position fixes determined from a separate radionavigation system. RDSS, on the other hand, integrates radionavigation, radiolocation and messaging within a single satellite system and transmission protocol.

Geostar received its RDSS license from the Federal Communications Commission in 1986[1]. Frequencies for RDSS were allocated throughout the world by the 1987 World Administrative Radio Conference for the Mobile Services[2].

Geostar is phasing the implementation of its RDSS satellite system. Before launching its own satellite relays, Geostar commercially provided a satellite-based position location and reporting system using the low-orbit Argos satellites in 1987. That phase of Geostar's service development provided an early demonstration of the benefits of a nationwide, satellite service to the long-haul trucking industry.

With the launch of its RDSS relay on board the GTE Spacenet III satellite in 1988, Geostar began its current phase of operations with a position reporting service. In 1989, a two-way messaging capability was introduced using conventional 4/6 GHz domestic satellite transponders for messaging to the mobile terminals.

CURRENT OPERATIONS

Geostar is currently providing over 65,000 position reports and messages each day to over 70 commercial and government customers. The current service provides users with a position reporting and two-way messaging capability. Presently, positions are determined using Loran-C. A detailed description of Geostar's initial system is provided in a paper presented by Robert D. Briskman at the AIAA 12th International Communications Satellite System Conference entitled, "Geostar Initial RDSS System."[3]

Mobile units transmit a burst containing position and status information, as well as an alphanumeric message of up to about a hundred characters, at 1618.25 MHz in the band allocated for RDSS. The transmission is a direct sequence, spread spectrum signal, with a duration of 20 to 80 milliseconds per transmission packet. This signal consists of 15.625 kilobit per second data, rate-1/2 encoded for forward error correction, spread by a direct sequence pseudo-random noise code operating at 8.000 megabits per second. The spread signal is BPSK modulated and occupies about 16 MHz of

bandwidth. The RDSS relay on board the host domestic fixed satellite retransmits these signals to Geostar's central earth station and operations center in Washington, D.C., using the conventional 11.7-12.2 GHz domestic satellite band.

A C-band (4/6 GHz) link using commercial fixed satellite transponders provides communications from Geostar central to the mobile terminals. This outbound link to the mobile terminals is a continuous signal, which is framed and contains message packets addressed to mobile users. This signal consists of 1200 bit per second data, rate-1/2 encoded for forward error correction, spread by a direct sequence pseudo-random noise code operating at 1.2288 megabits per second. The spread signal is BPSK modulated. Two such signals can be transmitted over the same transponder. Geostar is currently using an uplink earth station provided by GTE Spacenet to transmit this signal to the satellite, with landline connections between that earth station and Geostar's operations center in Washington, D.C.

Geostar recently extended its coverage capabilities with the addition to its system of RDSS relays on board the Gstar III satellite. That satellite carries an RDSS relay covering the continental United States as a backup to the RDSS relay on Spacenet III, as well as a second RDSS relay covering Mexico, Central America, the Caribbean and portions of South America. Figure 1 illustrates Geostar's current coverage capabilities.

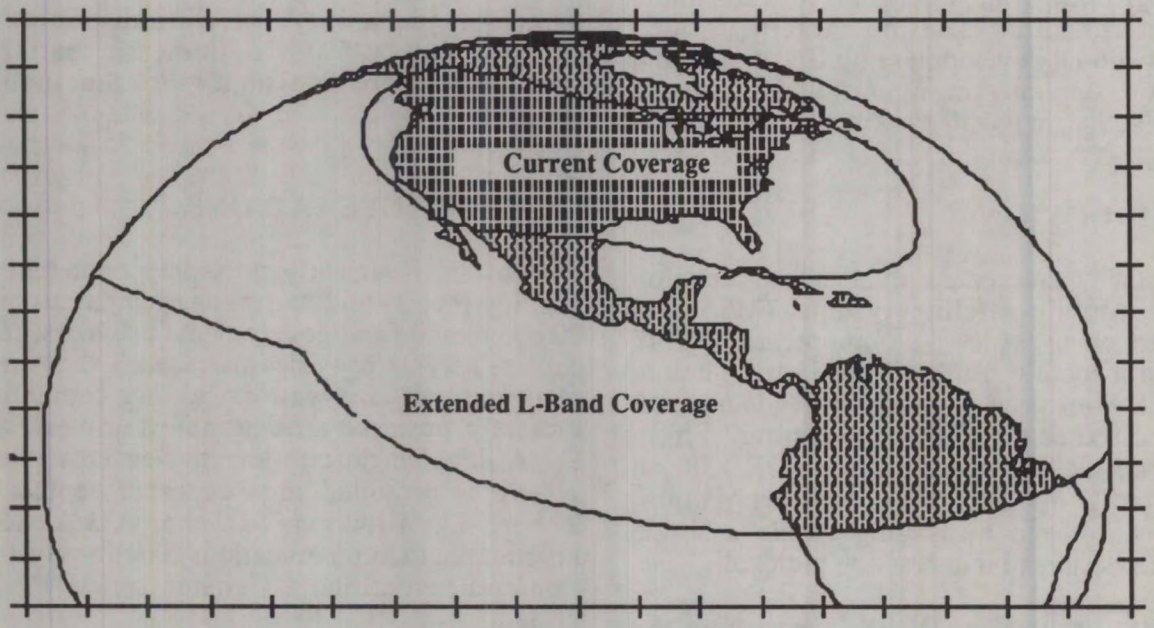


Figure 1. Geostar System Coverage

The launch of Gstar III in September 1988 was anomalous, in that it was left in an initial 16 by 36 thousand kilometer, 16-hour orbit after the firing of its apogee kick motor. However, by means of a long and complicated set of maneuvers, the on board thrusters were used to successfully boost the satellite into geosynchronous orbit. It became operational in October 1989, and is expected to have four or more years of operating lifetime.

Mobile user equipment is being provided by Sony (*2-Wayfarer*™) and Hughes Network Systems (*SkyRider*™). The mobile terminal consists of a small, omnidirectional antenna unit mounted on the outside of the vehicle, a receiver/transmitter and mobile data terminal

installed inside the vehicle, and a keyboard/display unit mounted near the driver. For the Sony unit, the external transmit/receive antenna unit is 3.5 inches high and weighs 2.9 kilograms, with the internal electronics unit occupying approximately 825 cubic inches (11.7" by 7.1" by 9.9") and weighing 9.6 kilograms. A small external Loran-C whip antenna is also mounted on the vehicle. The keyboard/display unit is backlit, displays four 40-character lines, and can be programmed with various menus and preformatted message layouts. The keyboard/display is approximately 11" by 5.9" by 1.4" and weighs 1 kilogram. The Sony unit also includes a sensor port and an RS-232C port to interface with other ancillary equipment. The Hughes mobile terminal is similar in size, weight and configuration, and provides the same capabilities as the Sony unit.

Geostar's current system is being used for a wide variety of applications to all types of vehicles, including trucks, trains, aircraft and boats. Software is available for displaying position and status information and for two-way communications using personal computers. For large fleets, software has also been developed to interface Geostar's service directly with mainframe or minicomputers at the subscriber's headquarters. Real-time interfaces between Geostar and its customers utilize SNA LU 6.2 and Communications Manager protocols, and Geostar's operations center can be interconnected with the subscriber's headquarters through leased lines, the IBM Information Network or the Telenet™ point-to-point service. Geostar will also support the X.25 packet switched protocol, and VSATs can also be used to connect a subscriber's headquarters with Geostar's operations center. Geostar supports Electronic Data Interchange, and RDSS user terminals can be interfaced directly with refrigeration, trailer and security devices for automatic transmission of alarms.

RDSS SYSTEM DESIGN

Geostar is licensed by the Federal Communications Commission to construct, launch and operate three dedicated RDSS satellites by 1993. These satellites will be built by GE-Astro and launched on the space shuttle. They will provide service to users in both the S-band (specifically, 2483.5-2500 MHz) for satellite-to-mobile terminal transmissions, and the currently used L-band RDSS frequencies (1610-1626.5 MHz) for mobile terminal-to-satellite transmissions[4]. Position determination will be accomplished automatically by ranging through these multiple satellites. A detailed technical description of Geostar's RDSS system is provided in a paper presented by L.O. Snively and W.P. Osborne at the AIAA 11th Communications Satellite Systems Conference entitled, "Analysis of the Geostar Position Determination System."[5]

The general operation of Geostar's dedicated RDSS system is illustrated in Figure 2. This RDSS system will integrate the position determination and messaging functions into a single operation. In the RDSS system, a continuous outbound signal is transmitted to provide time reference marks and an outbound time division multiplex data stream. To determine the position of a mobile unit, the mobile unit retransmits one of the time reference marks, adding its unique identification code and other information, through two or more geosynchronous RDSS satellites. The position of the mobile unit is calculated at the central earth station from the round trip propagation times through three satellites, or by the round trip propagation times through two satellites and altitude information obtained from a digitized terrain map or on-board altimeter. The calculated position and any other information addressed to the user is transmitted back on the outbound data stream between the system time reference marks. The positioning accuracy of Geostar's dedicated RDSS system is expected to be better than fifty meters.

A miniaturized terminal is currently under development for government customers which will measure 5.6 by 3.1 by 1.6 inches and weigh 22 ounces, excluding battery pack. This user terminal includes a keyboard display containing a 2-line by 39 character display and 3-row by 10-column alphanumeric keypad. The unit also includes the L-band transmitter, S-band receiver, omnidirectional patch antennas, and internal processor module. Geostar expects that similar user terminals will also be commercially available to the general public.

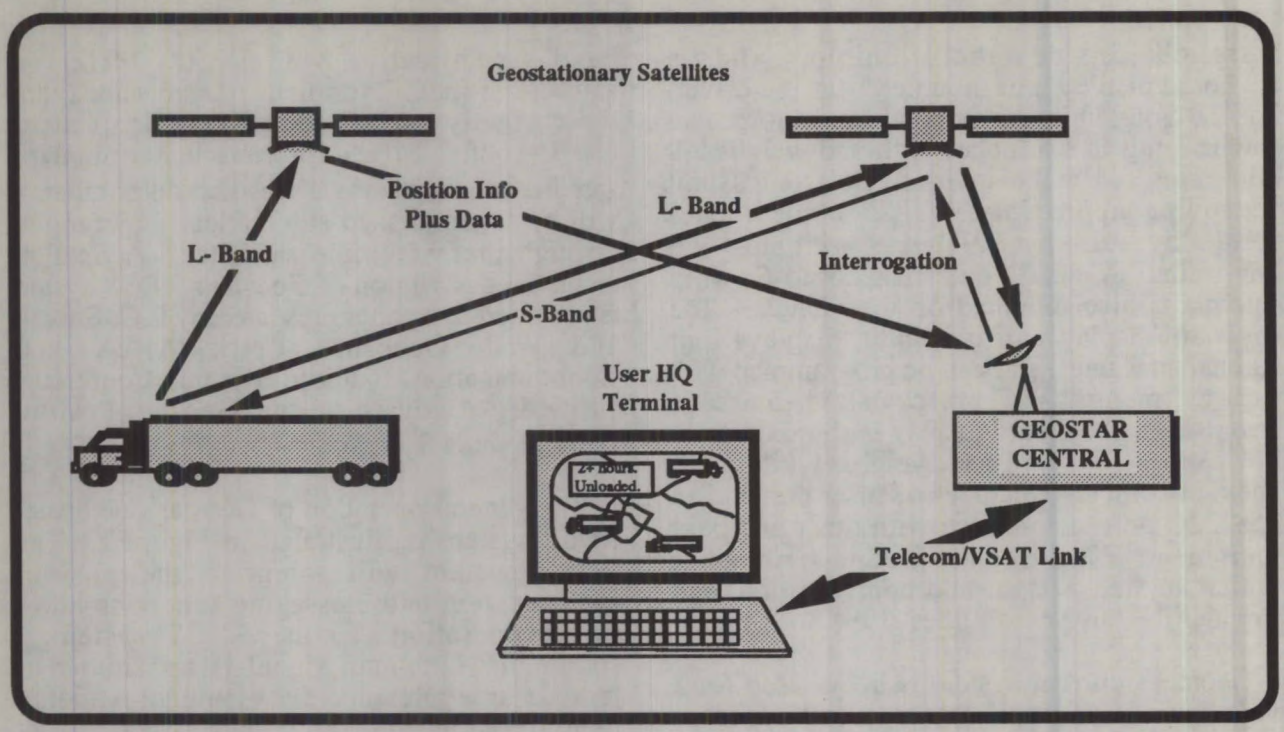


Figure 2. RDSS System Architecture

INTERNATIONAL RDSS

Significant progress is being made in developing RDSS in other parts of the world. In Europe, Locstar is developing an RDSS system that will be interoperable with Geostar. In March 1990, Locstar increased its capitalization from 100,000,000 French Francs to 676,000,000 FF (approximately $17.4 and $117.7 million, respectively), with 42 European partners from 13 countries. Locstar plans to launch two RDSS satellites in 1992 covering Europe and the Mediterranean basin. Mobile terminals are being developed by three consortia, lead by MAN, SAGEM and Techniphone. An RDSS processing and control center is being established in Marseilles, France.

Geostar is also establishing a Mexican affiliate to provide a position reporting and messaging service in Mexico using a combination of Geostar's RDSS relays and the Morelos system. In addition, the second generation of AUSSAT satellites will carry RDSS relays, and Geostar is working to establish an Australian RDSS venture.

REFERENCES

1. **Federal Communications Commission** 1986. *Memorandum Opinion, Order and Authorization*, Mimeo 6144, released August 7, 1986.

2. **International Telecommunication Union** 1987. *Final Acts of the World Administrative Radio Conference for the Mobile Services*, Geneva: ITU.

3. **Briskman, R. D.** 1988. Geostar® Initial RDSS System. *AIAA 12th International Communications Satellite System Conference.* Crystal City, VA; Paper 88-0862.

4. **Geostar Corporation** 1985. *Geostar Satellite System Compendium of Application and Technical Information.* [filed with the Federal Communications Commission on April 5, 1985].

5. **Snively, L. O. and Osborne, W. P.** 1986. Analysis of The Geostar Position Determination System. *AIAA 11th Communications Satellite Systems Conference.* San Diego, CA; Paper 86-0606.

Session 2
Future Mobile Satellite Communication Concepts

Session Chairman - *Joseph Pelton,* University of Colorado, USA
Session Organizer - *William Rafferty,* JPL

Integration between Terrestrial-Based and Satellite-Based Land Mobile Communications Systems

Antonio ARCIDIACONO*
European Space Agency
Telecommunications Directorate
10 rue Mario Nikis
75738 Paris Cedex 15, France

1. INTRODUCTION

The growing interest in Mobile Satellite Systems (MSS) worldwide has underlined the need for a definition of the possible role of MSS in the more complex panorama of land mobile services. Entrepreneurs stand ready to develop regional mobile satellite systems. Nevertheless potential MSSs would only serve a niche segment of the overall mobile telecommunications market (see Fig. 1), with a potential market share strongly depending on the relative penetration of MSS with respect to terrestrial-based mobile telecommunications services.

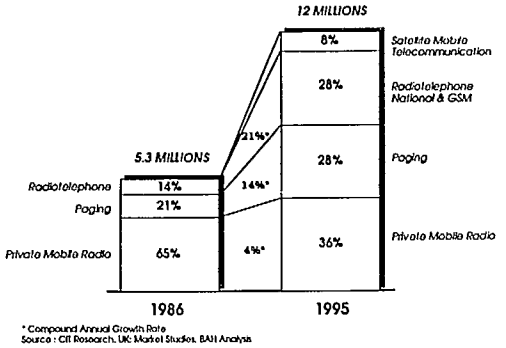


Fig.1. Growth projections for mobile communicating in Europe

On the basis of technical and economical factors, several promising solutions have been analyzed to establish the relative market penetration of MSS. One of those resides in the integration between terrestrial-based and satellite-based MSSs. Another resides in the research of a possible synergy among fixed, transportable and mobile satellite voice/data services/technologies.

*Now with Eutelsat, Tour Maine Montparnasse, Office of the Director General, 75755 Paris Cedex 15, France. The ideas expressed in this paper are those of the author. Responsibility for the contents resides with the author and does not belong to the organizations.

The market segment to be addressed has been clearly identified in the low traffic density areas (see fig. 2). The complementarity with existing and future mobile services is then evident whilst at the same time a possible synergy between mobile and fixed (or transportable) satellite services could be demonstrated. In the long run mobile and Personal-VSAT stations will only differ because of the different frequency allocation and the different market segment addressed while at network level the same tele-services and bearer services are likely to be provided. Several studies have recently been performed on these topics in Europe and this paper is intended to summarize some results which have been obtained. In particular there are two main areas in which the interest has been focused.

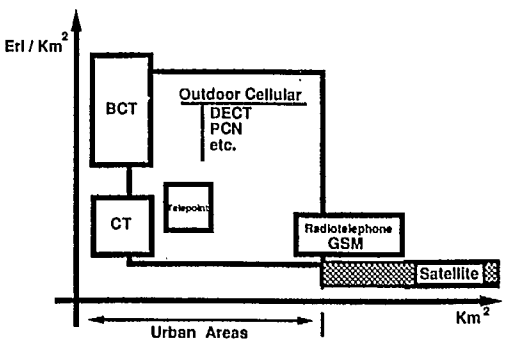


Fig 2. Traffic density vs. mobile market segmentation

The first one is the possible "network" integration between the first European generation of the Digital Cellular Mobile System (GSM) and a Land Mobile Satellite System (LMSS) for the provision of public LMSS services.

The GSM (and in the long run the UMTS: Universal Mobile Telecommunications System) will ultimately provide one common cellular system for the whole of Europe. However,

while the implementation of GSM starts in 1991, it may take quite a few years until the coverage is more or less complete, and it is unlikely that this service will be available outside the CEPT countries: Hence, there is a market for a mobile telephone system which complements the GSM system and

- provides complete coverage of Europe, land and maritime neighbouring areas;
- requires little installed equipment so that a service can be provided in a short time;
- has the flexibility to adjust to a changing geographical distribution of the users to keep up with a build-up of the GSM system, and reduces the early and overall investment for the GSM implementation.

These requirements can be optimally satisfied by a satellite-based land mobile radio system: the entire area of interest can be covered from one location in space, hence all the above features are provided automatically. The usefulness of such a system can be enhanced by making the two systems as compatible as possible in terms of services, protocols, etc. This can lead to a system which requires very little hardware in addition to the GSM terminal and which selects the best link automatically and autonomously, hence the user, in theory, needs not even be aware of the fact that the connection is completed through a satellite link.

Lastly, this system will provide a satellite extension of the public cellular mobile system at the expense, for customers interested in a full satellite coverage, of the use of a dual mode satellite/GSM terminal.

On the other hand, if we look at non-public mobile communication services, the same philosophy can be applied even if due to the reduced communication quality and availability requirements, a different system architecture can be optimized. Due to the power and bandwidth constraints in the LMSS environment, the optimized access scheme for satellite closed user network applications has been found to be a system using a synchronized CDMA accessing scheme.

This kind of accessing scheme has the advantage, in terms of frequency reuse, of at least a factor of 2.5 over an FDMA based satellite system. Furthermore, the same accessing technique used in a mobile cellular (or trunked PMR) system has several advantages over the classic TDMA scheme. CDMA requires fewer cells, lower power levels, a simpler frequency planning and an increase in capacity of up to 20 times that of the existing techniques. Finally, in terms of service provision CDMA techniques can provide a complete set of services without any change in the lower network layers. This capability allows a larger flexibility in the provision of different services tailored to the different user requirements. The easy implementation of features like link-activation (voice activation, connectionless data links, etc.), a localisation and a very short call set-up time together with an efficient use of spectrum give to this accessing technique further advantages to become a good candidate for the provision of both terrestrial based and satellite based mobile and fixed telecommunication services.

For all these reasons the possibility of integration of a terrestrial based and satellite based system seems to be of great interest and also in this field several studies and experimental activities have been performed in Europe and all over the world.

2. THE PROVISION OF VOICE/DATA SERVICES IN THE FUTURE REGIONAL EUROPEAN LMSS

More and more interest has recently been given to the need for a European standard compatible with GSM for public services, and for the standardization of PMR satellite services.

The development of these standards is aimed at supporting the growth of a European regional system (covering also Eastern countries, North African countries and, in general, all countries in the Mediterranean basin) in order to complement the future public and private terrestrial based systems.

An initiative has been taken by ESA to fund the development of an L-band payload (with Ku-band feeder links) called EMS, to be launched in 1993. This payload is intended, as the first choice by ESA, to be operated by Eutelsat to provide digital MSS voice/data communications. A possibility for the embarkation of this payload is the Italsat I-F2 Satellite (1993). The spare to the EMS payload could be provided through the ESA-Artemis-LLM payload (1994) which will also be provided to experiment L-band spot beam technology in Europe.

It has also been considered in the future to provide a better coverage (elevation angles > 45°), of the countries in Northern Europe, utilising satellites in highly-inclined elliptical orbits (Archimedes (ESA), Loopus (D), Sycomores (F), Elmsat (I), etc.) For those systems the possibility of

utilising the Ku-band (or even mmwave) frequencies is under study.

2.1 Public service

It is recognized that the GSM system will ultimately provide a common system for the whole of Western Europe.

As outlined in the introduction, several requirements can be optimally satisfied by a satellite-based land mobile radio system (LMSS): the entire area of interest can be covered from one location in space. The usefulness of such a compatible system could be enhanced by making the two systems compatible in terms of services and eventually protocols. This can lead to a system which requires some hardware in addition to the GSM terminal and which selects the best link automatically and autonomously, hence the user needs not even be aware of the fact that the connection is completed through a satellite link. The proposed integration can be obtained by designing the LMSS/GSM integrated system without any impact on the already existing GSM specifications.

It can be assumed that the most important use of the space segment is when a user in a zone not covered by the GSM cellular system is considered. The possibility of a mobile terminal able to select which of the two link characteristics, i.e. GSM or satellite, is better at a fixed moment, could be envisaged, thus comparing, for example, the two levels of the signal received from the base station. This is possible for a standard GSM terminal since the central station periodically forwards managing hand-over and location information. However, some additional modifications should be considered in this case as in other cases. For example, in the terrestrial system each base station broadcasts some data on its neighbours. This information is used in the handover procedure. Since handover from the satellite can be to any one of a very large number of terrestrial cells, the satellite cannot provide this information.

A more realistic approach could be to allow handovers between the terrestrial and the satellite system but not vice-versa. Furthermore, considering the significant cost involved in defining and designing a new telecommunications system both at terminal and network levels, research on the maximum commonality between the protocols of the two systems is fundamental.

For this reason, the mobile terminal could be conceived with a view to minimizing the impact in terms of cost/complexity on the design of a re-usable GSM/LMSS terminal. Similarly, the maximum commonality at protocol level (at least for signalling) could permit the use of the same switching network developed for the GSM system.

To summarize, it is in principle possible to obtain the maximum commonality both in terms of protocols and eventually in terms of terminal hardware between the LMSS and the GSM. However, should the re-usable terminal architecture be simplified, if a different protocol or hardware were assumed for the LMSS system, the requirement of the maximum commonality would be abandoned in favor of the reduced cost/complexity of the dual mode terminal. This is due to the fact that a different distribution of the overall business opportunity has to be considered in the terrestrial-based and in the satellite-based systems. In the terrestrial GSM, only 30% of the overall system cost can be allocated to the terminals' cost and the remaining 70% is dedicated to the base stations (50%) and switching centres (20%). On the other hand, in a satellite-based system the largest share of the overall system investment cost corresponds to the terminals' cost amounting to more than 50% of the overall LMSS cost.

The integration of terrestrial and satellite-based mobile systems will be further taken into account in the evolution towards UMTS (Universal Mobile Telecommunication Services). During this work the particular characteristics of satellite-based mobile systems will be taken into account to ensure effective interwork- ing. A base for these investigations can be found in the work of CCIR 8/14, CCIR 8/13, COST 227 and ESA.

2.1.1 Identification of critical areas for the implementation of integrated GSM/LMSS systems

The terrestrial and satellite land mobile systems have been considered separately up until now and, therefore, their specifications have not been harmonized. The difficulty in the harmonization of specifications of the two systems seem, at first glance, to depend on the physical macroscopic differences between the terrestrial-based and satellite-based mobile systems. In particular, these physical limitations can be found in the signal attenuation, propagation delay, multipath channel characteristic, etc.

Moreover, difficulties in the harmonization between the two systems depend on the delay in the frequency allocations of LMSS as

compared with cellular systems.

A fully integrated approach based upon the harmonization of all the technical solutions adopted in the two systems hardly seems feasible. In an attempt to rationalize the integration possibilities a scale of possible levels of harmonization between the two systems has been proposed to CCIR IWP 8/14.

This classification follows a "Russian doll" structure where each possible level of integration includes the basic concepts of the previous one. We could have, in the following order: Geographical, Services, Network and finally System Integration (see fig. 3).

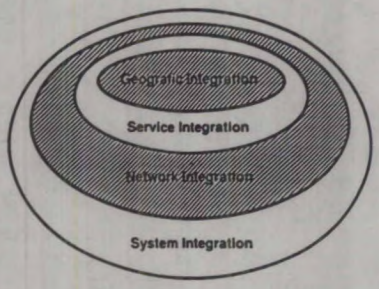


Fig.3. GSM/LMSS Integration levels

In the first case (Geographical Integration), the GSM and LMSS systems are independently conceived and based on different techniques and do not provide the same or compatible services as they could at the second level of integration. In this second case (Service Integration), the LMSS will only be capable of supporting a subset of the services available with a degraded service. Nevertheless, from the user point of view already at this level of integration a unique interface can be provided which has an appropriate protocol conversion.

However, even more interesting is the possible integration at the third level and, in a next generation, possibly at the fourth level (UMTS).

In the case of network integration, the fundamental requirement is to enable the fixed user (connected with the PSTN or ISDN) to call the mobile without having to select the call routing (satellite or terrestrial). A single calling number identifies the mobile. As already explained, this solution results in economies with regard to infrastructure due to the common utilisation of network functions.

At the same time, the maximum integration at network level should have an impact on communication and signalling protocols which should be as similar as possible to those of GSM. Finally, the last level of integration will envisage the system integration in a single network where the coverage areas provided by the LMSS are regarded as one (or more) cells of the original GSM system. This solution should include all the advanced system features such as rerouting of live calls between either satellite cells or terrestrial cells.

2.1.2 Space segment scenarios

As far as the space segment is concerned, two different satellite P/Ls are at present under development:

- satellite P/L with single beam (Eurobeam)
- satellite P/L with spot beams

The two listed alternatives have different operational capabilities and not all the GSM functions can be implemented in each of them.

Option 1.

In the first case, we assume a transparent satellite with a single beam covering Europe. The link with mobiles is in L-band and the link with fixed earth stations is in Ku-band. This kind of space segment corresponds to the European Mobile System (EMS) payload (see fig. 4) which should fly in 1993 [1]. In this case only three kinds of services have been considered:

- low data rate LMSS services
- voice/data PMR services (up to 9.6 kbit/sec)
- voice/data public LMSS services

The possibility of integrating a subset of GSM network services with these kinds of EMS services seems reasonable but only at the first or, at most, second level of integration (geographical or service integration, see sect. 2.1.1)

The main characteristics of the referenced EMS L-band payload are:

Forward Link
Total EIRP 42dBW (45dBW on Italsat-F2)
3 channels: 4+4+3 Mhz

Return Link
G/T - 2.0 dBK (-1.2dBK on Italsat-F2)
11 channels: 11x1 Mhz

Option 2.

In this case, we assume a transparent satellite with multi-spot beams to and from the mobiles (see fig.5) and a global beam to and from the fixed earth stations.

This solution corresponds to one of the experimental payloads which will be embarked on the ESA-Artemis satellite to be flown in 1994.

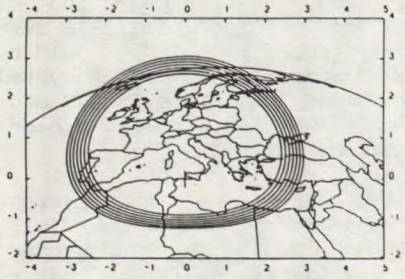


Fig.4. EMS coverage (L-band)

Fig.5. Artemis coverage (L-band)

In the case of the provision of a public service, we can consider for this space segment two possible system architectures: a central earth station for the system communicating with all mobiles or multiple earth stations in the Eurobeam (as many as the number of beams).

In the first case the use of a single fixed earth station has the disadvantage of creating long terrestrial tails. In the second case those ground tails can be reduced. However, if the mobiles are generally far from their base stations the tails may actually be longer than with a single earth station.

The main characteristics of the referenced ESA-Artemis L-band payload [1] are:

Forward Link
Total EIRP 51dBW (to be shared across the 6 beams)
7x1 Mhz channels*

Return Link
G/T - 4.6 dBK
7x1 Mhz channels

*over the MSS allocated frequencies

2.2 LMSS for Digital Trunked PMR services

In addition to a public system providing voice/data services, the need for a business voice/data service has been identified as a natural improvement and follow-on of the already existing low data rate LMSS service (e.g. Euteltracs).

Several study activities are in progress in order to define an optimized network architecture for the provision of a satellite digital trunked PMR service. The possibility of accepting a reduced speech quality (2.4 : 4.8 Kbit/sec. voice coding) together with the maximization of the spectrum utilisation (synchronized CDMA techniques, etc.) should provide an economically viable service. Two other main functional features for the proposed system are the provision of a VSAT receiving station at customer premises together with the reduction in the call set-up time (bypassing any PSTN connection).

2.2.1 CDMA techniques for a terrestrial/ satellite system.

Historically CDMA was mainly considered for applications requiring low susceptibility to interference and/or information privacy. For other applications (civil applications) the utilisation was considered to be limited due to the performance degradation arising from "self-noise", caused by the hardware complexity for the implementation of CDMA receivers (mainly due to the code acquisition and tracking procedures), and to the complexity of the compensation of near-far problems and finally to the poor spectrum efficiency when compared with conventional accessing techniques.

At European level an attempt to introduce CDMA techniques for land mobile communications was made by Philips (NL) and SEL (D) which presented to the GSM committee ("Groupe Spécial Mobile" which selected the first generation of a European Digital Mobile System) two proposals based on a CDMA concept. Finally, partly due to the mentioned criticalities, neither of those two proposals was retained and the final decision was taken for a TDMA/TDM system.

Only recently in Europe [2] and in the United States [3] a new interest has been shown to the implementation of CDMA based land mobile communication systems. This new interest is fundamentally due to the introduction of a new CDMA implementation concept which can be called synchronised CDMA.

In Europe this interest has mainly been

originated by ESA which has been working for several years on CDMA techniques and which recently has devoted a strong interest in the development of the synchronized CDMA system concept. Moreover ESA is also supporting the European industry in a technological effort aimed at the development of a series of systems, with a common technological basis, for the provision of fixed, transportable and mobile services.

Similarly, in the United States a slightly different synchronized CDMA concept has been introduced by Qualcomm [3] in the case of satellite based communications systems.

Only a few months ago the same technique was experimentally applied to digital cellular radio communications [4] obtaining an improvement in capacity of 20 times, compared with the present cellular techniques.

The big advantages of these CMDA concepts are represented by the reduced carrier-to-interference ratio which can be accepted by a CDMA system (allowing the reuse of the same frequency in every cell or in every satellite spot), by the less stringent requirements in terms of required Eb/No (because of the more powerful codes which can be used and because of reduced fading margin) by the exploitation of the voice duty cycle (allowing active spectrum and power reuse) and by the implementation of a sophisticated power control algorithm able to cope with the different signal levels from different users because of their location or because of different service requirements (the system is inherently able to dynamically provide different services at different data rates for different users).

Concerning the specific cellular implementation CDMA allows also soft transition from the already existing analog cellular networks to the digital network because of the inherent robustness to interference of the CDMA system and because of the low level of interference generated by each user vis-a-vis the existing analog network. In practice it is possible to double the system capacity reusing the same frequencies at the beginning of the new service and then, as the demand for CDMA service grows, some band segments can be removed from analog service and dedicated to the CDMA service. Finally, very accurate position location together with privacy can be provided by the cellular system because of the spread-spectrum nature of CDMA.

Therefore taking into account all those advantages and the common nature of the used technology in the terrestrial-based and satellite based scenarios, there is an opportunity to develop an interworking network for fixed, transportable and mobile applications to provide voice/data services to closed user groups. The idea could be to have on one side a common CDMA system for terrestrial based and satellite-based PMR applications in which a simplified hand-over philosophy can be applied.

Only hand-over from the terrestrial network to the satellite network will be considered and, due to the long propagation delay and the slightly different nature of the terrestrial based and satellite based systems, this hand-over will be performed breaking for a few hundred milliseconds the conversation/datalink in order to abandon the terrestrial network and set up a call in the satellite network.

On the other hand, it is possible to reuse the same CDMA technique/technology in the case of fixed and transportable voice/data services (Personal-VSAT) for closed user groups, and an eventual integration of the different services in a single trunked PMR network should be further investigated.

Finally, in order to reduce the number of terrestrial tails and connections to a minimum the PMR satellite network architecture has to be based on a distributed relevant private hub-station configuration. In practice feeder link access to the network should be provided through private VSAT hub stations (located at customer premises) while only the network coordination functions should be centralized at the satellite operator/service provider premises (see fig.6).

With reference to the first two options given in sect. 2.1.2 an analysis has been recently performed [5] in order to give an assessment of the capacity which can be obtained using a synchronized CDMA access technique with the two above-mentioned space segment configurations:

	EMS (42dBW)	LLM (51dBW)
Total number of available channels	480	3200

Starting from the capacity calculation in the case of the EMS payload it is possible to address a nominal market of at least 75,000 users in Europe. This figure could virtually be doubled if the Italsat I-F2 is the selected space segment.

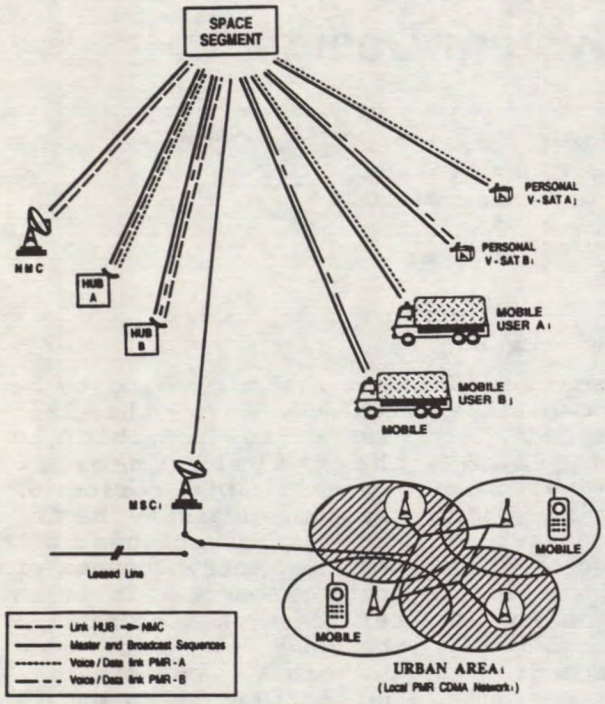


Fig.6. A CDMA-based terrestrial/satellite architecture

Therefore, considering on the present cost for the development, launch and operation of such a small piggy-backed payload (1/4 of a medium size satellite including in-orbit sparing) there is the possibility of offering a voice/date service to closed user groups at a cost well under 500 ECU/year/user.

This figure does not include the investment for the mobile terminals and for the mini-hub station.

As already mentioned, more than 50% of the overall system cost for an MSS is due to the investment for the procurement of mobile terminals.

For this reason the support in the development of appropriate (in financial and technical terms) user terminals (mobile and fixed) is at least as important as the funding and development of an appropriate space segment. Several activities have been started by ESA in this domaine, and a few units for the experimentation of future personal satellite services will be available in 1991.

3. CONCLUSION

In order to increment the economical attractiveness of future personal (fixed, transportable and mobile) satellite services a new access technique has been proposed. In 1991 the experimental viability of such a technique will be experimentally tested and trialed in Europe. Should the present performance estimation be confirmed by the experimental results, the large-scale industrialization of personal satellite communication terminals shall be initiated with the aim of developing integrated closed user network services.

Concerning the development of public mobile satellite services the same technology could be applied limiting to network level the integration with the terrestrial mobile system.

A more detailed analysis will be necessary to define a satellite/terrestrial integrated mobile system (reusing the same protocols) aimed at the development of the future UMTS services.

REFERENCES

[1] **G. Perrotta, F. Rispoli, T. Sassorossi** "Payloads development for European Land Mobile Satellites" IMSC '90, Ottawa (Canada), June 1990

[2] **R. De Gaudenzi, R. Viola,** "A synchronous Code Division Multiple Access System for High Capacity Mobile Satellite Systems" IMSC '90, Ottawa (Canada), June 1990

[3] **K.S. Gilhousen, I.M. Jacobs, R. Padovani, L.A. Weaver Jr.,** "Increased Capacity Using CDMA for Mobile Satellite Communication" IEEE-J-SAC, January 1990

[4] **Qualcomm, Pactel Cellular,** "CDMA Cellular the Next Generation" San Diego, November 1989

[5] **A. Vernucci, S. Arenaccio, R. Padovani, A.Arcidiacono,** "Performance, Implementation and Network Management Techniques of a European CDMA-Based Land Mobile Satellite Systems" Submitted for approval to IEEE-Globecom 90, San Diego, December 1990.

Spectrum and Orbit Conservation As a Factor In Future Mobile Satellite System Design

Robert R. Bowen,
Department of Communications, Canada
300 Slater St., Ottawa, Ontario, K1A 0C8
Phone: 613-998-3974
FAX: 613-958-0567

ABSTRACT

Access to the radio spectrum and geostationary orbit is essential to current and future mobile satellite systems. This access is difficult to obtain for current systems, and may be even more so for larger future systems. In this environment satellite systems that minimize the amount of spectrum-orbit resource required to meet a given traffic requirement are essential. Several spectrum-conservation techniques are discussed in this paper, some of which are complementary to designing the system at minimum cost, others less so. All may need to be implemented to the limits of technological feasibility if network growth is not to be constrained because of the lack of available spectrum-orbit resource.

1. INTRODUCTION

Spectrum and orbit conservation in the context of the mobile satellite service is the business of ensuring that the frequency bands and the geostationary orbit (GSO) available to mobile satellite system operators are used effectively. Spectrum-orbit conservation can perhaps be considered within the broader theme of environment conservation, but with a difference: while environmental problems such as water and air pollution, deforestation, etc. degrade the environment in a tangible way, the radio spectrum and GSO is not "destroyed" in the same direct way; it is in theory completely reusable at any future time. But when we introduce the telecommunication engineer's ongoing problem of being a slave to his own history, next year's and next decade's systems having to be compatible with those of the last decade, it is quite possible to degrade the spectrum-orbit "environment" over a long period of time, with system designs having short-term economic advantages but using larger than necessary amounts of the spectrum-orbit resource. It is in the mobile-satellite system operator's long-term interests to use good spectrum and orbit conservation techniques, whether they be to ensure coordination of his present system or to ensure spectrum is available for his follow-on system.

This paper attempts to find ways of balancing the objectives of designing mobile satellite systems that are both cost-effective and that make good utilization of the available spectrum-orbit resource. These are in some instances complementary objectives, and in other cases divergent. In the latter case they are likely to be followed only through the process of setting radio regulations at the international (ITU) level, the national level (in Canada through policies and regulations set by the Department of Communications), and by coordination agreements among mobile satellite system operators through the Article 11 (Radio Regulations) process.

In the 1 to 3 GHz band, where there are a large number of users and potential users, and where many systems and services are at an early stage in their development, it is particularly important to weave good spectrum and orbit conservation techniques into the design of systems. Another reason for paying particular attention to such techniques at this

time is that WARC-92 will likely set the basic international radio regulations for the mobile satellite service in the 1-3 GHz band for the next decade at least.

Spectrum-orbit conservation can perhaps be considered from two different aspects: sharing the resource between the networks of a given service, the mobile-satellite service in this case, and the sharing of the resource between the mobile-satellite service and other radio services. The latter may or may not be possible, but we must know whether it is possible by the time we go to WARC-92. The first question will be considered in section 2, the second in section 3.

2. SPECTRUM-ORBIT SHARING AMONG MOBILE SATELLITE SYSTEMS

Unless or until spectrum-orbit conservation is a basic design objective of the satellite-system designer, his job is essentially to design the system with the required capacity and required signal quality at the minimum possible cost. Perhaps the first step in examining spectrum-orbit sharing among mobile-satellite systems is to look at that design problem from the perspective of the system noise budget. The overall link carrier-to-noise ratio (c/n) is given by the equation

$$\frac{c}{n} = \left\{ \left(\frac{c}{n}\right)_u^{-1} + \left(\frac{c}{n}\right)_d^{-1} \right\}^{-1} \qquad (1)$$

where (c/n)u and (c/n)d are the uplink and downlink carrier to noise ratios respectively. The uplink ratio (c/n)u can be specified by the equation

$$\left(\frac{c}{n}\right)_u = \frac{p \cdot g(o) \cdot l \cdot f \cdot h(o)}{kTB} \qquad (2)$$

where p is the earth-terminal transmitted power,
g(ϑ) is the earth terminal antenna gain at an angle ϑ degrees off boresight,
l is the uplink free-space loss,
f is the fading loss that must be included in the design,
h(ϕ) is the gain of the spacecraft antenna at an angle ϕ degrees off boresight,
k is Boltzman's constant,
T is the uplink effective noise temperature, and
B is the uplink noise bandwidth, approximately equal to the signal necessary bandwidth.

A similar equation exists for $(c/n)_d$ to go into equation (1). The mobile satellite system designer's task is to balance the choice of p, g, h, T, f, and B to minimize the system cost subject to a specified (c/n), taking into account such factors as satellite weight budget, number of earth terminals, traffic growth over the systems design life, etc.

2.1 One Spectrum-Orbit Conservation Choice: Signal Bandwidth

The most obvious interface between the system designer and the frequency manager is the choice of the bandwidth parameter B of Eqn (2). If n channels are to be accommodated in a given area, the required bandwidth of the system is nB; if the total available bandwidth is fixed, n can only be increased by decreasing B. However, in decreasing B by choosing a different type of modulation, the system designer is confronted by a tradeoff between bandwidth and power to transmit a given signal of given post-detection quality (S/N). Power is costly, especially on the spacecraft, and so one must speak already of the "cost" or "value" of spectrum conservation. Further, decreasing B at the expense of requiring a higher carrier to interference ratio (c/i) may or may not be a good choice from an overall spectrum-orbit conservation perspective, because higher (c/i) values mean larger satellite spacings, and possibly reduced potential for interservice sharing.

2.2 A Second Spectrum-Orbit Conservation Measure: Spacecraft Antenna Discrimination

To understand a series of spectrum-orbit conservation measures

one should look at the carrier-to-interference equations that quantify the interference mechanism between two satellites. These are similar to the (c/n) equations (1) and (2) above. The uplink $(c/i)_u$ carrier-to-interference equation for two networks using the same frequency in the same direction is

$$\left(\frac{c}{i}\right)_u = \left(\frac{p}{p'}\right)\left(\frac{g(o)}{g'(\theta)}\right)\cdot\left(\frac{l}{l'}\right)\left(\frac{f}{f'}\right)\left(\frac{h(o)}{h'(\phi)}\right) \qquad (3)$$

where the superscript (′) refers to a parameter of the interfering network, θ is the separation angle on the GSO between the interfered-width and the interfering satellite, and (ϕ) is the angle off boresight of the interfering Earth-station as seen from the interfered-width satellite location.

The spacecraft antenna discrimination $\{h(o)/h(\phi)\}$ may be enough to provide the necessary $(c/i)_u$, with or without assistance from other factors of Eqn (3). A particular case of Eqn (3) is with $\theta = 0$, i.e. the interfering and interfered-with networks being on the same spacecraft. This is the situation with multi-beamed spacecraft with enough isolation between beams to permit frequency reuse. This frequency-conservation measure may be complementary to the objective of minimizing overall system cost if larger antenna gain h(o) and lower downlink power p in the downlink counterpart to Eqn (2) results in a lower spacecraft cost to provide a given EIRP. However, the provision of steep antenna-gain rolloff characteristics and high values of $\{h(o)/h(\phi)\}$ may require satellite costs greater than that expended simply to provide the required satellite EIRP at minimum cost. Thus again these is a need to quantify a "value" or "cost" to spectrum conservation, in this case through frequency reuse.

2.3 A Third Spectrum-Orbit Conservation Measure: Earth Terminal Antenna Discrimination

Another factor in Equation (3) is the Earth terminal antenna discrimination factor $\{g(o)/g(\theta)\}$. It may be possible to design an antenna with enough earth-terminal antenna discrimination to permit frequency reuse of the GSO from another satellite away. Given that the mobile terminal must operate while moving, some combination of mechanical and/or electronic (phased array) steering would be required. This technique is at least partially complementary to the objective of minimum-cost design in that a higher gain antenna, and therefore an antenna with greater discrimination, will permit lower satellite transmit powers and consequently lower satellite weight and cost.

2.4 A Fourth Spectrum-Orbit Conservation Measure: Network Homogeneity

If Earth terminal antennas can be designed with enough discrimination that frequency reuse at orbit separations say 30° to 60°, then it becomes important to minimize the inhomogeneity between the networks so that this angle can be minimized. This inhomogeneity is expressed in Eqn (3) by the parameter (p/p'). If two networks have significantly different transmitted power levels, for one or another reason, then the necessary separation angle to protect the low-power network is greater than that necessary to protect the high-power network. However, the angle between them has to be large enough to protect both, and so must be the larger of the two. To minimize this angle, the two networks should be designed with $p_1 \simeq p_2$. There may be variations from this when this factor is combined with others mentioned above, but the trend should be to avoid large differences between p_1 and p_2. This hasn't been considered seriously to date because Earth terminal antenna discriminations are not yet large enough to allow frequency sharing at less than "over-the-horizon" separations. Hopefully, this will change, and when it does p_1/p_2 inhomogeneities will be a significant factor in spectrum-orbit utilization.

3. SPECTRUM SHARING BETWEEN NETWORKS OF DIFFERENT SERVICES

There are in theory a large number of possibilities that might be

considered here, but two are particularly attractive:

i) sharing between different mobile-satellite services, i.e. the aeronautical mobile-satellite service (AMSS), the land mobile satellite service (LMSS), and the maritime mobile-satellite service (MMSS); and

ii) sharing between these satellite services and the terrestrial fixed and mobile services.

3.1 Spectrum Sharing Between Mobile Satellite Services

In the ITU Radio Regulations the frequency bands 1530 to 1559 MHz and 1626.5 to 1660.5 MHz are divided into a number of sub-bands allocated to various combinations of AMSS, LMSS, and MMSS on a primary and secondary basis. In contrast, according to Canadian spectrum policy document SP 1530, these bands are allocated to the composite Mobile-Satellite Service (MSS) with two exceptions:

i) the bands 1545-1548 MHz and 1646.5-1649.5 MHz are allocated to the aeronautical mobile satellite service exclusively, (and this excludes air public correspondence), to provide a firm base for the development and implementation of air-traffic-control systems by satellite; and

ii) aeronautical mobile satellite traffic, (excluding air public correspondence) must be provided a means of real-time priority or interrupt in the higher bands 1548-1559 MHz and 1649.5-1660.5 MHz, in recognition of the fact that ATC traffic needs very fact response from the telecommunications network that it uses.

There are three reasons for taking this approach that are related to the objective of conserving or making better utilization of the radio spectrum and the GSO:

i) in a relatively new service (or services) such as the MSS, it is not obvious that the division of capacity requirements should be made in a particular way. The atent requirements of the AMSS, LMSS, and MMSS may evolve at different rates and to different extents from that foreseen in dividing the band between the three services. The more generic approach allows the evolution of system development to accommodate the different services as they emerge.

ii) the requirements for AMSS, LMSS, and MMSS vary in different geographical areas. For instance, there is very little demand for LMSS capacity in mid-Atlantic, and similarly very little demand for MMSS capacity on the Canadian prairie. This becomes important as the systems evolve from the earlier global-beam systems to those using multi spot beams with frequency reuse, discussed in section 2.2.

iii) the different types of MSS traffic have definite diurnal peaks, and if these peaks differ for the different types of traffic more efficient use of a given block of spectrum can be made by combining the services in a larger composite network. As an example, the trans-Atlantic air travel has definite diurnal patterns in eastern North America, with traffic to North America peaking in early afternoon local time and Europe-bound traffic peaking in the evening.

This combining of mobile-satellite services in a "generic" system to make more efficient use of the spectrum is complementary to the system designer's objective of designing a minimum-cost system, in that in designing a larger composite system considerable economies of scale in satellite design can be obtained, and perhaps economies of producing larger numbers of similar Earth terminals. As well, it allows the system operator to increase the utilization of his facility by integrating the different types of traffic.

3.2 Sharing Between Mobile-Satellite and Terrestrial Services

WARC-MOB-87 recognized the need to review the ITU allocations to mobile-satellite services in the near future; presumably that will take place at WARC-92. There are, however, no unused frequency bands in the 1 to 3 GHz range of the radio spectrum. Agreement for additional allocation to spectrum for mobile-satellite services would be eased considerably if these services could share on a simultaneous-use basis with terrestrial services such as the fixed and/or the mobile services.

Sharing is not as convenient as an exclusive band, but it may be feasible, particularly in the MSS downlink. In that direction interference is between the transmitting spacecraft and the receiving terrestrial station, and from the associated transmitting terrestrial station and the receiving Earth station. In this sharing arrangement the key may be in using directive Earth-station antennas, seen in section 2.3 as providing spectrum and orbit conservation for a completely different reason. If higher gain directive Earth station antennas can be used this would reduce the required power-flux-density on the ground from the satellite, thereby easing the interference into terrestrial receivers. Moreover, such directive antennas would reduce the interference from transmitting terrestrial stations. Perhaps sharing arrangements in the MSS downlink can be agreed upon.

In the bands used for the MSS Earth-to-space link sharing may be more difficult, because the satellite receiver is subject to interference from the composite of all the terrestrial transmitters in its coverage area. Sharing may be possible with fixed systems, because antennas of fixed systems should not be pointed at the geostationary orbit. However, sharing with transmitting terrestrial mobile systems would be more difficult and may not be possible.

4. SUMMARY

A number of spectrum and orbit conservation techniques involving the design and operation of mobile-satellite systems have been described. Some of these may be considered at WARC-92, others more appropriately considered in the normal activities of the CCIR. But in the final analysis they can only be put into effect if spectrum and orbit conservation is fully integrated into the design and operation of a mobile-satellite system, not put together after the system has been designed and is about to be "coordinated" under Article 11 of the Radio Regulations. The mechanisms are there; some are complementary to minimum-cost design, others less so. But the health of the mobile-satellite industry over the longer time-frame depends on effective available spectrum and orbit conservation techniques being implemented.

Personal Communications: an Extension to the Mobile Satellite

Murray Epstein, François Draper
Scotgroup Enterprises Inc.
2300 Laurentien, Montreal, PQ
H4R 1K3, Canada
Phone: 514-337-7222
FAX : 514-332-5228

INTRODUCTION

A great statesman once said: "Information is power." Timely information can be even more powerful. In today's mobile society, it is increasingly more difficult to reach individuals in a timely manner. Telephone tag can be a businessman's pastime, or perhaps his nightmare.

Technologies are converging, becoming more sophisticated, and many are evolving towards personalised communications, reaching an individual by name rather than a vehicle, an office, or a terminal on a desk. Technologies have been developed such as paging, cellular telephony, cordless telephony, data terminals, personal computers (with modems) and the widespread proliferation of facsimile, similar to the PC boom several years ago, to name a few.

Today, Industry has developed tools that can reach us by telephone in our car, in our boat, on the convention centre floor. We can send weather updates, stock market quotations, and feed a message to a FAX machine that will be transmitted to someone's pocket or belt on a device that accepts 80,000 characters in its memory, no wires attached. We can monitor alarms, shop at home, move information around the world in seconds, even monitor a patient's heartbeat while he continues his normal daily chores. Our imagination is the only limit to these applications.

Telephony has been the most preferred technology to date and the most utilised. However, even in the better developed urban areas, covering less than 10% of the Canadian land mass, portable (cellular) telephones are sometimes impractical, often unaffordable to the average consumer. The problem is even more compounded when we consider the rural and underserved areas of our country, the remaining 90%. **"Mobile satellites may hold the answer."**

TERRESTRIAL SERVICES: URBAN APPLICATIONS

In this rapidly changing world, we can no longer speak of one technology - one application. Technologies such as "computers," once thought of as large mainframes, have permeated our everyday lives not only at the work level but in the home. They are qualified as "personal" and have become consumer products. They are utilised in a multiplicity of applications, new ones being continuously created.

The same phenomenon is occurring in communications; they are becoming more personalised. New applications are being developed through the combination and manipulation of two or more existing ones and the emergence of any one technology will ensure the metamorphosis of many existing ones. Such is the evolution of services today.

These developments are evident in the more urbanised, densely populated areas, covering 90% of the population and only 10% of the land mass. In these areas, due to blockage, fade margins, high interference, and low sensitivity and signal levels, satellite signals can only be captured with very large, high gain antennas, impractical for transportable and mobile applications. Hence, terrestrial systems provide the means to personalise communications to the individual. Although analog systems are making way for digital ones, such services as cellular telephony, cordless telephony, paging, data services one and two-way, facsimile, etc. will continue to be satisfied by terrestrial systems for some time to come.

MESSAGING BY SATELLITE: RURAL AND UNDERSERVED AREAS

The advent of mobile satellites will push Communications beyond the traditional urban boundaries, well into the rural roots of Canada. It will optimise the "service coverage area," extending it to all parts of the country. We have a resource based economy, yet we function in large urban locales, the population living in a 5,000 kilometer corridor, huddled near the U.S. border, occupying 10% of our land mass. We transport our goods along this corridor yet we can communicate with these various vehicles for only 20% of their journey. They are the lifelines of our Industries, our food stuffs and our economy. The question that remains unanswered is: "Will mobile satellites be affordable?"

The high capital investments and the heavy operating costs of these satellite systems will make the use of voice circuits difficult to afford. On the other hand, the high compression techniques and near real time needs of data circuits make them more practical and more cost-effective. Thin route, satellite to device, "Data Messaging" (paging) becomes a solution to the non-urban user needs, one-way or two-way.

INTEROPERABILITY: RURAL AND URBAN

As high as the benefits of existing systems may be, one critical element is missing: INTEROPERABILITY between them, that is the ability to operate devices in both rural and urban areas; thereby realising the true needs of the user, full ubiquitous services in both environments.

The marriage between these converging technologies will develop new and exciting services to better meet the user requirements. Where satellites cannot offer mobile, personal communications services to subscribers, in the 12 major Canadian corridors, covering 10% of the territory, terrestrial systems will emerge dominant. While in the remaining 90% of the land mass, where capital investments for terrestrial systems would be ludicrous, satellite services will dominate. The key, however, is found in the ability to operate the same device in both

areas, offering complimentary rather than competitive services.

Downconversion

Initially, it is envisaged that the messages will be received with a standard L-band/MSAT receiver (Mobile Earth Terminal) and antenna, and be downconverted from L-band (1.6 GHz) to an Intermediate Frequency (IF), encoded, and retransmitted at VHF (150 MHz), low band UHF (420 MHz) or high band UHF (931 MHz); the bands where pagers presently operate (Figure 1). This downconverter/repeater will be a low power, standard stability, digital only, low cost transmitter. Outputs will be in the 100mw to 1W range. The subscriber would then be able to receive the message within a radius of 1/2 mile from the downconverter, with excellent penetration.

The L-band receiver would be either stationary (repeater on land) serving remote communities, transportable or located in a vehicle (truck, car, ship or plane). A truck driver, for instance, could leave his vehicle when stopped, and continue to receive all his messages within a reasonable distance from that vehicle, a problem that presently exists since he is out of that vehicle longer than he is in. This service would likely utilise existing alphanumeric pagers, synonymous with one-way data devices.

When in a major corridor, the device could be disconnected from the downconverter, and simply operate through existing terrestrial systems, as they do today. The advantage of this process is that no major technological leap is required.

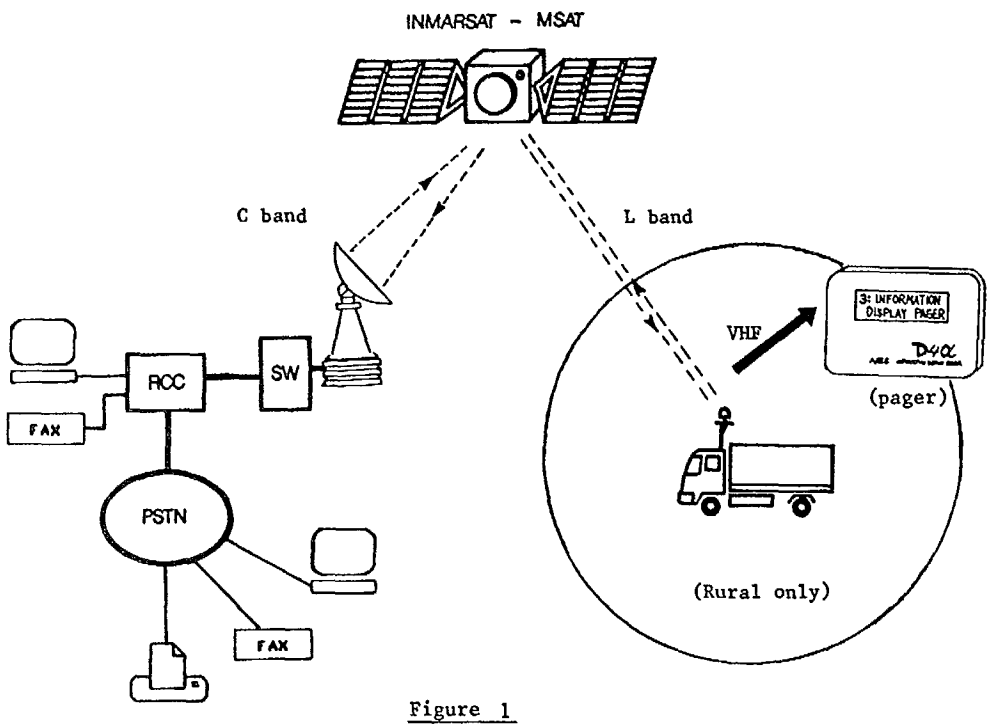


Figure 1

Dual-mode

The next development will be a dual-mode or dual-band data device, capable of transmitting/receiving L-band messages directly from the satellite while in rural areas and transmitting/receiving messages from terrestrial systems when in urban areas (Figure 2).

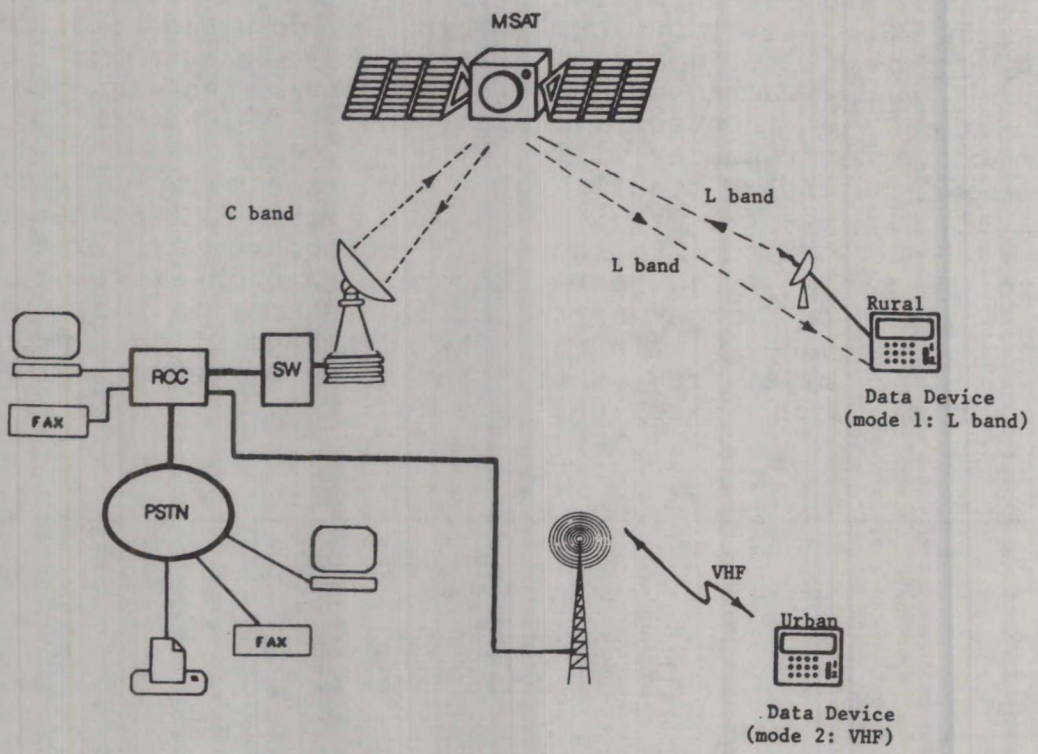


Figure 2

The downconverter, in effect, becomes part of the device. By then the satellite antenna gains will need to be greater and the device gain and sensitivity equally higher. The service will be truly portable and likely two-way. It is perceived, however, that the data device would require an external antenna (folding) in order to transmit to the satellite while in rural areas. Reception only from the satellite would not need the external antenna, nor would it be needed in urban (terrestrial) areas. Rather than paging, we might think of this service as two-way data or full messaging.

The subscriber will no longer need to carry the "Mobile Earth Terminal." This mobile data device will be dual-band scanning or perhaps manually switchable in order to conserve battery life.

Single band

Further into the future, we can conceive of a single-band data device operating at L-band only, directly with the satel-

lite, two-way and with terrestrial systems when necessary in order to reduce the cost. As in cellular telephones, the data device would identify its position, through a control channel, so that the switch could locate the subscriber and transmit messages to it by the most economical means (by satellite or terrestrially). (Figure 3)

CONCLUSION

As time progresses, the customer demands are far more universal: integrated, simple to operate, cost-effective services, with technology virtually transparent to the operator.

Industry will be in a position of providing those necessary services to meet the subscriber

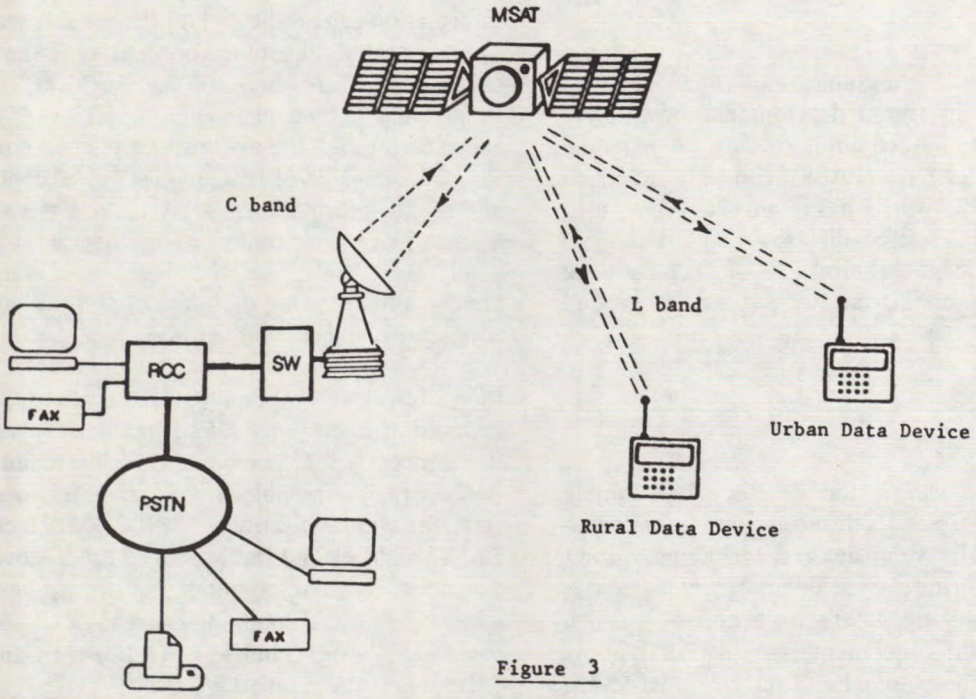


Figure 3

The subscriber now has a truly interoperable data device, including position location. It is fully two-way data and by now at relatively high speeds. The subscriber no longer carries any external antenna and the technology has evolved to provide an efficient, cost-effective, practical device. Although much more costly, the subscriber has a digitised-voice option on his device that he uses when required.

needs. Our resource based industries, transportation and utilities in the more rural and underserved areas will require quality and affordable services that can only be supplied via satellite. Yet these same services will need to be satisfied, to the same subscribers, operating the same devices in terrestrial environments. The convergence of technologies is showing the way. One answer to those needs will be one and two-way, interoperable, Data Messaging.

A Satellite-Based Personal Communication System for the 21st Century

Miles K. Sue, Khaled Dessouky, Barry Levitt, and William Rafferty

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109, U.S.A.
Phone: (818) 354-3927
FAX: (818) 393-4643

ABSTRACT

Interest in personal communications (PCOMM) has been stimulated by recent developments in satellite and terrestrial mobile communications. A personal access satellite system (PASS) concept has been developed at JPL which has many attractive user features including service diversity and a handheld terminal. Significant technical challenges have been addressed in formulating the PASS space and ground segments.

1.0 INTRODUCTION

The 1980's were clearly the decade when mobile satellite systems (MSS) advanced from initial concepts to practical system designs, technology development, and interim service demonstrations. Now, at the beginning of the 1990's, expectations are high for the successful implementation of MSS, in its many forms, at the national and international levels. While MSS is eagerly awaited on many fronts, interest in extending MSS to the personal level is already gaining significant momentum.

Although the implications of personal communications (PCOMM) will only become clearer to both users and the technologists alike as the last decade of the 20th Century unfolds, early concepts are emerging which will greatly influence current thinking on the shape of telecommunications in the 21st Century. The socio-economic consequences of PCOMM will inevitably become a major topic of discussion and, in real terms, will heighten the competition between fiber/wire, terrestrial and satellite communications.

The last decade of the 20th century might well become synonomous with the true arrival of the information age predicated on the reliable transfer of unprecedented quantities of data between diverse users located anywhere in the world [1]. Such a claim, and attendant requirements, can only be achieved through the existence of telecommunication resources capable of reaching individuals with useful and timely information. However, it is not sufficient to just place a communications device in a user's hand: the device, i.e the link, must meet user demands by allowing data access and exchange in a competitive and cost-effective manner.

In an effort to respond to this challenge, the Jet Propulsion Laboratory (JPL) has been investigating the concept of a personal access satellite system (PASS) [2,3] to enable individual users to share in satellite communications (SATCOM) technology. PASS would extend the primarily urban coverage of terrestrial cellular systems by providing similar services to less populous areas that are not commercially or technically practical for land-based communications networks.

Many innovative services could be supported by a PCOMM system of this type, including

- o direct personal voice and data
- o personal computer file transfer
- o data base inquiry and distribution
- o low-rate broadcast (voice, data, video)
- o telemonitoring and control
- o disaster and emergency communications

The objectives of the PASS program are to develop and demonstrate system concepts and high-risk technologies for a personal SATCOM system. Ka-band, with downlinks at 20 GHz and uplinks at 30 GHz, has two unique features compared to the lower frequencies currently in use in mobile satellite communications. First, this band permits small user

terminals suited for PCOMM, particularly for hand-held operation. Second, ample bandwidth is currently available in that band. This should permit a system of significantly higher capacity than in currently utilized lower frequency bands. In turn, this should reduce eventual terminal cost, despite the higher frequency, and make it as affordable at the personal level as satellite terminals at lower frequencies. Thus, the potential for both enhanced user services and the development of new technology at higher frequencies have spurred the move to Ka-band.

Migration to high frequencies is certainly fraught with its unique difficulties and risks. The goal of this article, then, is to identify these challenges, and to present the early results of the research aimed at overcoming the hurdles to a cost effective realization of PASS.

In the following, the PASS concept and basic elements of its system design are first highlighted. Next the key challenges and risk areas are identified along with some possible approaches to resolving them. The relevant early research results are explained and their implications addressed. Finally, the present status and future plans are discussed.

2.0 PASS SYSTEM CONCEPT AND BASIC DESIGN FEATURES

PASS is a satellite-based PCOMM system that will offer users freedom of access and mobility. Equipped with a handheld or laptop terminal, a subscriber would have access to a host of voice and data services anywhere within the range of the associated satellite transponder. The system would be capable of handling data rates ranging from less than 100 bps for emergency and other low-rate services, to 4.8 kbps for voice communications and hundreds of kbps for computer file transfers.

As illustrated in Figure 1, PASS connects a network of private or public service providers with a large community of individual subscribers. The major elements of PASS include one or more satellites, a network management center (NMC), tracking, telemetry and command (TT&C) stations, supplier stations, and user terminals. The NMC and TT&C stations govern the operation of the system. As presently conceived, the user equipment falls into three categories: a basic personal terminal (BPT), an enhanced personal terminal (EPT), and telemonitors. The EPT is similar to today's very small aperture terminals (VSATs), whereas the BPT is a compact personal terminal that provides users greater freedom and mobility. The telemonitors are used for remote data collection and monitoring.

The fundamental elements of the PASS design rest on the utilization of a geostationary (GEO), bent-pipe satellite transponder with multiple, fixed spot beams to provide simultaneous up- and downlink coverage to users in the contiguous United States (CONUS). In addition, a single CONUS beam connects the satellite with the supplier terminals. A high power commercial satellite bus is also assumed. The multiple access techniques and concomitant modulation and coding schemes to be chosen on the forward (supplier-to-satellite-to-user) and return links need to support the highest possible overall system capacity without unduly complicating the user terminal. The BPT itself must be small (hand-held in size) and, as a minimum, capable of stationary operation. Ambulatory (talk-while-you-walk) operation is a desired option.

The basic features of the PASS design are highlighted in Table 1. The characteristics of the PASS satellite are given in Table 2 and the requirements for the BPT are listed in Table 3. Also to aid in placing the technological challenges in perspective, an abbreviated representative link budget is given in Table 4 for a data link requiring a 10^{-5} BER; this implicitly assumes the use of time division multiple access (TDMA) in the forward direction and single channel per carrier (SCPC) frequency division multiple access (FDMA) on the return. (Multiple access issues will be addressed in more detail later.)

3.0 HIGH-RISK ENABLING TECHNOLOGIES

Several high-risk enabling technologies have been identified. Some of these technologies are system architecture specific while others are not. The key enabling technologies are:

- o low-cost, compact, high-gain, tracking user antenna
- o low-cost user terminal frequency reference
- o MMIC transmitter
- o high-gain, low-noise MMIC receiver
- o VLSI-based integrated vocoder/modem
- o efficient multiple-access schemes
- o multi-beam satellite antenna and beam forming
- o Robust, power-efficient modulation and coding

Table 5 compares the state-of-the-art performance and PASS requirements for several key technologies. Timely development and validation of these technologies are essential to the successful implementation of PASS.

4.0 ADDITIONAL TECHNOLOGICAL CHALLENGES

In addition to the high-risk technologies described above, the PASS strawman design reveals a number of other challenges that are equally critical.

User Terminal Radiated Power Level

The transmitter and antenna of the user terminal need to produce an effective isotropic radiated power (EIRP) of about 17 dBW. One combination that can produce the required EIRP is 0.25 W transmit RF power and a 23 dBi antenna gain. An important consideration in determining the transmitter parameters and limitations is that the near- and far-field microwave energy levels comply with established safety standards.

System Reliability and Service Quality

The strawman design employs a combination of uplink power control on the forward link and adjustable data rate in both directions to combat rain attenuation. When increased uplink power from the supplier fails to fully compensate for rain degradation, the data rate can be reduced to close the link. This could conceivably result in a reduction of service quality, or even the suspension of certain services during severe rain conditions. Additional measures, such as the use of satellite on-board processing, could improve system reliability and service quality.

Non-Uniform Subscriber Distribution

Since the users are not likely to be uniformly distributed over CONUS, the available network capacity will be under-utilized unless this factor is properly accounted for in the design of the satellite. While this problem is common to all systems employing multiple spot beams, the large number of these, and correspondingly small footprints exacerbate this problem for PASS. If an acceptable adaptive power management scheme can be found, the improvement might be significant.

While these challenges are not necessarily show stoppers, they could be design drivers or result in serious operational constraints, performance degradation, and/or system capacity reduction.

5.0 SOLUTIONS

A number of studies have been performed in the past year to address these challenges. These efforts are intended to improve performance, increase capacity, alleviate operational constraints, and reduce the burden on the spacecraft and ground terminals. Some potentially promising remedies have been identified while other options that once seemed attractive have been eliminated.

5.1 OPTIMIZED MULTIPLE-ACCESS SCHEME AND SATELLITE DESIGN

Economical viability of a PASS-type system is a direct function of user terminal cost, which in turn is inversely proportional to system capacity, i.e., to the number of users who can be supported by the system. As mentioned earlier, one of the fundamental reasons for migrating to Ka-Band is the availability of bandwidth. A study has been performed to determine the bottlenecks limiting system capacity, and to determine the most effective design approach to ameliorate capacity limitations.

With a preset multi-beam spacecraft antenna architecture, and a user terminal of given capabilities, it is found that choices of multiple access technique, modulation and coding schemes, spacecraft total RF power, spacecraft link power allocation, channel rates and number of channels are all interrelated [4]. Analysis shows that the most serious bottleneck exists on the forward downlink to the user. Consequently an efficient TDMA scheme has been adopted for the forward link. On the return link it is found that either FDMA or CDMA (using direct-sequence spreading, i.e, SSMA) could be used effectively for maximum capacity depending on the nature of the traffic and the size of the satellite. FDMA is best with data traffic while CDMA is more suitable in a voice dominated system, particularly for a higher powered satellite. Table 6 summarizes some of the key results [4]. Capacities ranging between half and a full order of magnitude more than an L-band system could be achieved [5]. This requires an order of magnitude increase in bandwidth relative to L-band. This is, however, one of the primary reasons for a leap to Ka-band.

A result that appears to be particularly promising is the use of SSMA on the return link. Powerful convolutional codes and exploitation of voice activity combine to result in substantial capacity increases [4]. The use of SSMA can also realize the benefits of instant access to the system, minimum network control, and position determination. It also can make more feasible ambulatory operation by taking advantage of the inherent multipath rejection capability of SSMA.

Additional information on the proposed SSMA design can be found in [4,6,7]. A more definitive study will be conducted in the near future.

5.2 ALTERNATIVE ANTENNA COVERAGE CONCEPTS

Different CONUS cellular configurations have been studied as a means of alleviating the burden on user terminals and more effectively matching the satellite resources to the traffic demand arising from the previously discussed non-uniform user distribution. As stated earlier, PASS is more sensitive to traffic variations from cell to cell because of the relatively large number of spot beams. One way to alleviate this problem is to employ interbeam power management to dynamically adapt to traffic variations. Scanning/switched beams and hybrid fixed/switched beams are more amenable to such schemes by permitting variable dwell times. Results of initial studies indicate that while these approaches utilize the satellite capacity more efficiently, the benefits come at the expense of increased satellite complexity and user terminal EIRP. Some possible disadvantages include: increased complexity of the antenna beam forming network, increased message delay, increased user transmitted data rate and radiated power. At this point, these offsetting disadvantages appear to outweigh the potential benefits. Consequently, other methods of mitigating the possible effects of traffic variation are being explored.

5.3 THE USE OF NON-GEOSTATIONARY ORBITS

The potential advantages of elliptical and circular non-GEO orbits for PASS have been examined with the objective of reducing the user terminal EIRP requirements. Low-earth orbits (LEOs) have several potential advantages over their GEO counterparts: higher elevation angles and hence less multipath and rain attenuation, less space loss, and lower launch costs.

Analyses indicate that non-GEO orbits are not desirable for PASS because of the following negative factors: the large number of satellites required to provide continuous CONUS coverage, more complicated spacecraft antenna pointing requirements, increased satellite handover complexity, and ultimately, the small savings in link power requirement. It should be noted, however, that a combination of GEO and non-GEO satellites could be used to extend coverage to higher latitudes which is a consideration for global coverage applications.

5.4 USER TERMINAL RADIATION CONSTRAINTS

Many studies have concluded that potential harm to humans from microwave energy, including millimeter waves, is strictly due to thermal insult [8]. Radiation at 30 GHz is generally less difficult to manage than at L band or UHF. This is primarily due to its minimal penetration of human tissue, typically .77 mm. Studies have also found that because of the superficial nature of the exposure (i.e., similar to visible light) the eye, particularly the cornea, is the primary area of concern; this is because it lacks blood circulation which drains deposited heat. The ANSI standard for frequencies above 1.5 GHz is 5 mW/cm^2 averaged over a 6 minute period, which includes a factor of safety of 10 or more. Recent studies at 30 GHz (see references in [9]) have indicated that incident densities up to 100 mW/cm^2 did not cause any harm. Judicious PASS terminal design, however, dictates maintaining the average radiation level below the 5 mW/cm^2 in both the near and far fields. Preliminary computations on a 25 dBi antenna [9] have shown that the maximum radiation density is 153 x P mW/cm^2, where P is the average radiated power in watts. This occurs at a distance of 8 cm from the aperture and drops logarithmically with distance. With 0.2 W radiated power and 35% voice activity, the average maximum radiated density is about 10 mW/cm^2. This indicates the possible need for some additional transmit power restrictions if the ANSI standard is to be strictly followed. By exploiting a combination of duty cycles, call duration, and antenna pointing this problem can be safely resolved. The design of the user terminal will take this into consideration.

5.5 USER TERMINAL FREQUENCY REFERENCE

In mobile SATCOM systems, the dominant frequency uncertainty is due to Doppler. However, in the PASS environment, the motion of the user terminal is relatively insignificant so the critical frequency uncertainty component is the user terminal frequency reference.

Studies indicate that a demodulator frequency error equal to 10% of the bit rate (assuming binary modulation) will result in about a 0.5 dB performance degradation [e.g., 10]. For the baseline PASS data rate of 4.8 kb/s, this implies a receiver frequency stability of 1.6×10^{-8}.

A temperature-compensated (quartz) crystal oscillator (TCXO) could satisfy this frequency stability requirement, but it would exceed the cost (about $100) and power consumption (about 50 mW) constraints on the user terminal. The most viable alternative appears to be a microprocessor-compensated crystal oscillator (MCXO) operating at a fundamental frequency of 10 MHz that was recently developed for the U.S. Army [11]. Because of spectral purity deficiencies in this device and the need to operate at 20 and 30 GHz, the MCXO would have to be accompanied by a multiplier and phase-locked loop (PLL) clean-up circuit to meet the PASS specifications.

6.0 FUTURE PLANS

PASS is a satellite-based communications system designed to provide a variety of services ranging from low bit rate PCOMM to high bit rate computer file transfer. Media competition for the low-rate personal applications is already emerging in the form of terrestrial (microcellular) PCOMM Networks (PCNs). This will most likely lead to the integration of space and terrestrial networks, forcing each to play an optimized telecommunications role, which will ultimately benefit the user. Telecommunications in the 21st century will be characterized by diversity of services; choice of media; and user-transparent, optimized information routing. In recognition of these trends, a two-pronged approach has been adopted for the PASS Program with the following objectives.

The first objective is to continue the 20/30 GHz PASS system study and technology development with the goal of advancing Ka-band technology in general, and Ka-band mobile/personal technology in particular. Enabling technologies targeted for development are: user antenna; user terminal components (vocoder/modem, transceiver, MMIC front-end, and frequency reference); modulation and coding; rain compensation techniques; and multiple access schemes. Currently, the intention is to incorporate these technologies, to the extent possible, into a mobile terminal for use with NASA's ACTS. This terminal is being explored by JPL as way to demonstrate future Ka-band mobile applications.

The expansion of cellular phones suggests that they will play a significant role in personal communications in the 21st century. Considering this fact, and general telecommunications trends (technical and economic), the second objective is to specifically address the roles of communication satellites in PCOMM and to devise system concepts for an integrated satellite/ground PCOMM network. Such a network will provide choice of media and route selection. The key to integrating the characteristically and architecturally different space and terrestrial communications networks lies in networking protocol compatibility. The ultimate objective is to devise a system concept capable of providing PCOMM to the user using a truly universal personal terminal.

ACKNOWLEDGEMENT

The authors wish express their appreciation to their colleagues at the Jet Propulsion Laboratory who have contributed significantly to the work described in this paper: P. Estabrook, M. Motamedi, and V. Jamnejad. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. "Telecommunications 2000", NTIA, 1989.

2. M. K. Sue, A. Vaisnys, and W. Rafferty, "A 20/30 GHz Personal Access Satellite System Study," 38th IEEE Vehicular Tech. Conf. VTC-88, Philadelphia, June 15-17, 1988.

3. P. Estabrook, J. Huang, W. Rafferty, and M. Sue, "A 20/30 GHz Personal Access Satellite System Design," International Conf. on Comm., June 11-14, 1989, Boston Ma.

4. K. Dessouky and M Motamedi, "Multiple Access Capacity Trade-offs for a Ka-Band Personal Access Satellite System," IMSC'90, these proceedings.

5. C. Wang, T. Yan, and K. Dessouky, "Performance of DA/FDMA Architecture Proposed for MSS," Proc. of the Mobile Satellite System Architecture and Multiple Access Techniques Workshop, JPL Publication 89-13, March 1989.

6. T. Yan and C. Wang, "An Alternative Resource Sharing Scheme for Land-Mobile Satellite services," IMSC'90, these proceedings.

7. M. Motamedi, M. K. Sue, "A CDMA Architecture for a Ka-Band Personal Access Satellite System: Complexity and Capacity," the 13th International Communication Satellite Systems Conference, March 11-15, 1990, Los Angeles, CA.

8. L. Heynick, "Critique of the Literature on Bioeffects of Radiofrequency Radiation: A Comprehensive Review Pertinent to Air Force Operations," USAFSAM-TR-87-3, June 1987.

9. K. Dessouky and V. Jamnejad, "Radiation Levels of the Ka-Band Mobile Terminal," Internal JPL Document: Interoffice Memorandum 3392-90-016, February 21, 1990.

10. P. Wittke, P. McLane, P. Ma, "Study of the Reception of Frequency-Dehopped M-ary FSK," Research Rep. 83-1, Queens University, Kingston, Ontario, Canada, March 1983.

11. S. Schodowski, R. Filler, J. Messina, V. Rosati, J. Vig, "Microcomputer-Compensated Crystal Oscillator for Low-Power Clocks," U. S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ 07703.

Table 1. Salient Features of PASS

OPERATING FREQUENCY	
UPLINK:	30 GHZ
DOWNLINK:	20 GHZ
COVERAGE CONCEPT	
SAT/SUPPLIERS:	CONUS BEAM
SAT/USERS:	142 SPOTBEAMS
GENERIC SERVICES	VOICE AND DATA
DATA RATES	
FORWARD:	UP TO 100 KBPS (BPT) UP TO 300 KBPS (EPT)*
RETURN (NORMAL):	4.8 KBPS (BPT)
RAIN COMPENSATION	
FORWARD:	UPLINK POWER CONTROL & VARIABLE DATA RATE
RETURN:	VARIABLE DATA RATE

* For EPT, the stated data rate includes built-in margin for rain compensation.

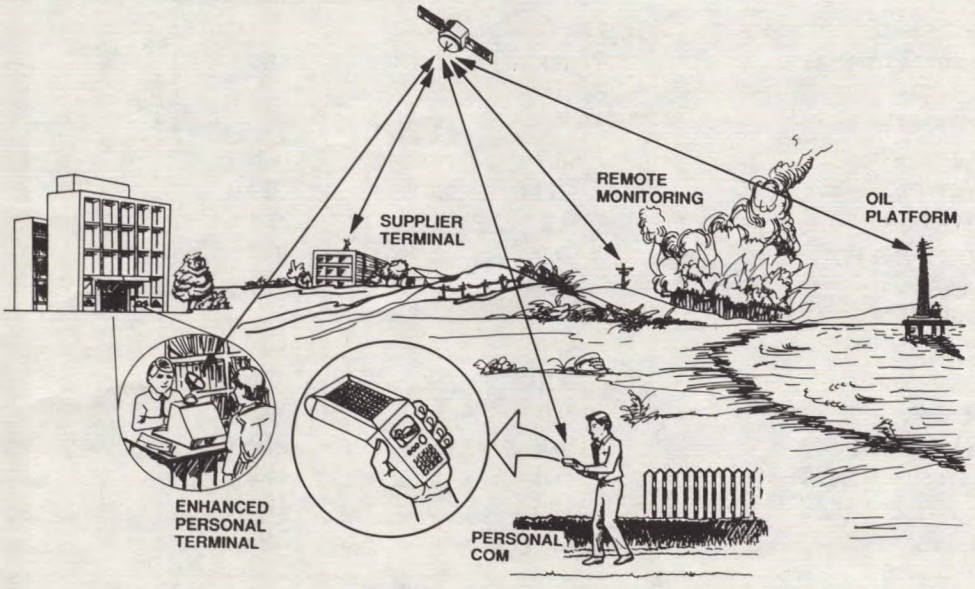


Figure 1. Personal Access Satellite System Concept

Table 2. Summary of Satellite Design

SPOTBEAM	
ANTENNA SIZE (TRANSMIT)	3 M
(RECEIVE)	2 M
NUMBER OF SPOTBEAMS	142
ANTENNA GAIN	52.5 DBI
ANTENNA BEAMWIDTH	0.35 DEG
SYSTEM G/T	23.4 DB/K
AVERAGE EIRP/BEAM	55 DBW
CONUS BEAM	
ANTENNA GAIN	27.0 DB
ANTENNA BEAMWIDTH	7.7 DEG
SYSTEM G/T	- 1.2 DB/K
EIRP	39 DBW
SATELLITE MASS (GTO)	7300 lb
SATELLITE POWER (EOL)	3.4 kW
	(for 520 RF watts)

Table 3. Design Requirements for the BPT

ANTENNA GAIN @20 GHZ	19.3 DBI
ANTENNA GAIN @30 GHZ	22.8 DBI
ANTENNA TRACKING/COVERAGE CAPABILITY	
AZIMUTH	360.0 DEG
ELEVATION	15-60 DEG
RECEIVE G/T	-9.0 DB/K
TRANSMIT POWER	0.3 W
NORMAL DATA RATES	
RECEIVE	100 KBPS
TRANSMIT	4.8 KBPS
OTHER REQUIREMENTS	
SIZE	HAND-HELD
MODEM	VARIABLE RATE

Table 4. Strawman Link Budgets for Basic Personal Terminal
(Data Link with BER requirement of 1E-5; No Rain)

	FORWARD (SUPPLIER-SAT-USER) IN DB	RETURN (USER-SAT-SUPPLIER) IN DB
DATA RATE, KBPS	100	4.8
UPLINK:		
EIRP,DBW	60.7	16.8
PATH LOSS, DB	-214.0	-214.0
RX G/T, DB/K	-1.2	23.4
RCVD C/NO, DB-HZ	69.9	46.9
DOWNLINK:		
SAT EIRP,DBW	57.0	6.4
PATH LOSS, DB	-210.5	-210.5
RX G/T, DB/K	-9.0	30.3
D/L C/NO, DB-HZ	58.8	50.3
OVERALL C/NO, DB-HZ	57.4	44.2
REQ'D C/NO, DB-HZ	54.5	41.3
MARGIN, DB	2.9 (1.0*)	2.9 (1.1*)

NOTE: * ESTIMATED 1-SIGMA VALUE

Table 5. Comparison of State-of-the Art Performance and PASS Design Requirements for Selected Key Technologies

TECHNOLOGIES	PASS ASSUMPTIONS/ REQUIREMENTS	RELEVANT EXISTING CAPABILITY/ DEVELOPMENT GOAL
DEVICE/COMPONENT		
LNR NF @20 GHZ	3.0 DB (S/C) 3.5 DB (USER TERMINAL)	1.5 DB HEMT LOW NOISE DEVICE 3.5 DB LNR BEING DEVELOPED
LNR NF @30 GHZ	3.0 DB (S/C)	2.0 DB HEMT DEVICE 5.0 DB LNR (ACTS)
HPA EFF. @ 20 GHZ	50% @ 5W	15% SSPA (<=5W) 40-50% TWT
HPA EFF. @ 30 GHZ	20-30% @ 0.3 W	5-15% @ 1-2W (HEMT SSPA) 15-20% @ .2W (HEMT SSPA) 35% @ 250mW BEING DEVELOPED FOR PLANETARY APPLICATIONS
MULTIBEAM ANTENNA AND FEED		
ANT SIZE @ 20GHZ	3M	3.3 M (ACTS)
ANT SIZE @ 30GHZ	2M	2.2 M (ACTS)
NO. SPOTBEAMS	142	<=10
USER TERMINAL ANTENNA		
BPT TRACKING ANT. GAIN @ 20/30 GHZ	19/23 DBI	
USER TERMINALMINIMIZATION		
TERMINAL SIZE & TECHNOLOGIES	HAND-HELD MMIC FRONT END VLSI MODEM/CODEC	MMIC ARRAY, RX/TX MODULES, CHIP-SIZE MCXO CODEC ON 1 BOARD

Table 6. PASS Capacities for Different Multiple Access Scheme Choices (adapted from [4]) p83 129
(Voice Links Assumed with BER = 1e-3 and a VOX factor of 0.35)

ACCESS SCHEMES	LINK Eb/NO dB	CODING (CONVOLUTIONAL)	CAPACITY (#CHANNELS) RETURN	FORWARD	BANDWIDTH (MHz) UP-LINK	DN-LINK	TOTAL	SAT RF POWER SPLIT(TOT/F/R)
FDMA (RET)/ TDMA (FWD)	3	R=1/2, K=7	8072	8216	111.8	110.3	222.1	410/390/20 W
"	2.3	R=1/3, K=7	9483	9653	183.4	181	364.4	410/390/20 W
CDMA (RET)/ TDMA (FWD)	2.3	R=1/3, K=7	10143	8452	65.1	183	248.1	410/335/75 W
"	1.5 (R)/ 2.3 (F)	SUP. ORTH.(K=10)/ R=1/3, K=7	10142	9331	69.1	180	249.1	410/375/35 W
FDMA (RET)/ TDMA (FWD)	3	R=1/2, K=7	10493	10433	142.3	143	285.3	520/494/26 W
"	2.3	R=1/3, K=7	12328	12258	233.6	234.6	468.2	520/494/26 W
CDMA (RET)/ TDMA (FWD)	2.3	R=1/3, K=7	12171	10565	78.6	206.4	285	520/425/95 W
"	1.5	R=1/3, K=9	13894	13523	99.8	258.2	358	520/470/50 W
"	"	"	19069	11411	85.1	225	310.1	520/390/130 W
"	1.5 (R)/ 2.3 (F)	S. ORTH.(K=11)/ R=1/3, K=7	17851	11411	101.8	356.3	458.1	520/460/60 W

Note: For a 10:1 data-to-voice traffic ratio, a 1.4 s average message delay, 2% voice blocking probability, 90 s/call/user/hr, 1000 bits/data message, one channel can servean average of 100 users

Future Mobile Satellite Communication Concepts at 20/30 GHz

S.K. Barton
University of Bradford
Bradford
W. Yorks BD7 1DP
U.K.

Tel: 44-274-733466
Fax:44-274-305340

J.R. Norbury
Radio Communications Research Unit
Rutherford Appleton Laboratory
Chilton, Didcot, Oxon OX11 OQX
U.K.

Tel: 44-235-445508
Fax: 44-235-445753

ABSTRACT

The outline design of a system using ultra small earth stations (picoterminals) for data traffic at 20/30 GHz is discussed. The picoterminals would be battery powered, have an RF transmitter power of 0.5 W, use a 10 cm square patch antenna and have a receiver G/T of about -8 dB/K. Spread spectrum modulation would be required (due to interference considerations) to allow a telex type data link (< 200 bit/s data rate) from the picoterminal to the hub station of the network and about 40 kbit/s on the outbound path. An Olympus type transponder at 20/30 GHz could maintain several thousand simultaneous picoterminal circuits. The possibility of demonstrating a picoterminal network with voice traffic using Olympus is discussed together with fully mobile systems based on this concept.

INTRODUCTION

Satellite-communications system at 20/30 GHz offer the exciting possibility of ultra small pocket size earth stations. Microterminals or VSAT (very small aperture terminals) have been established for some years in the 4/6 and 12/14 GHz frequency bands. However, these systems are fixed and cannot be considered to be either portable or easily transported. This paper summarises certain aspects of an ESTEC sponsored study (1) concerned with ultra small satellite terminals at 20/30 GHz, and makes some projections for future mobile systems for these frequencies.

One of the options identified in the ESA study, which might be considered to be portable rather than mobile, was a very small "picoterminal" with antenna sizes of about 10 cm square and beam widths near 10 degrees. Practical portable earth stations, the size of a thick paperback book are now technically possible through developments in VLSI circuits and MMIC technologies. Networks with several thousand simultaneous picoterminal circuits could be supported by the Olympus transponder, albeit with very low data rates for each picoterminal.

A picoterminal network could offer an alternative to the land mobile satellite systems now being considered at L-Band (1.5 GHz). The provision of 2 GHz bandwidth at 20/30 GHz could support a very much larger user community and offer the cost effectiveness of volume production.

THE PICOTERMINAL CONCEPT

The picoterminal concept (Figure 1) was conceived as a purely portable earth station, about the size of a thick paperback book (i.e. 10 cm by 20 cm and some 3-4 cm thick) for telex type data traffic.

These picoterminals would be part of an overall network controlled through a central hub station. However interference from satellites operating in adjacent orbit slots (2 degree spacing) automatically imposes some anti-jam modulation scheme.

Data rates are necessarily low, through power constraints both at the picoterminal and on the satellite. In the immediate future, commercial picoterminal systems would be confined to low data rate telex type messages of up to 200 bit/s.

The practicalities of producing a 20/30 GHz picoterminal depend very much on the achievable performance from monolithic microwave integrated circuits (MMIC). Power levels of near 1W with efficiencies greater than 5% have been reported in Japan (2). The current state of the art for low noise HEMT and MESFET indicate noise figures near 2 dB for 20 GHz. (3).

A printed array antenna has been suggested for the picoterminal. The feeder losses to the individual patches limit the practical peak gain to near 30 dB.

Propagation constraints at 20/30 GHz can be quite severe. Although a 99% availability would be acceptable for the picoterminal-satellite link, a better availability of 99.8% would be necessary on the hub to satellite section.

In Europe at 30 GHz, 99% availability can be achieved with a 4.3 dB margin whereas 14.5 dB is needed to reach the 99.8% level.

Table 1 summarises the parameters used in the subsequent traffic analysis. All the figures chosen are slightly conservative and already achieved in terms of practical systems.

Table 1

Power delivered to antenna	=	-6 dBW
EIRP	=	19.2 dB
System noise temperature	=	29.4 dBK
Antenna gain (edge of coverage	=	21.7 dB
G/T at 20 GHz	=	-7.7 dB/K
E_b/N_o outbound	=	10 dB
E_b/N_o inbound	=	8 dB

Traffic capacity of an Olympus transponder supporting a picoterminal network

The purpose of the calculation is to determine the maximum traffic capacity, which the transponder can support. The 20/30 GHz Olympus transponder parameters have been used where appropriate.

Outbound link budget. It is assumed that the hub station can always provide sufficient power to drive the transponder to a point representing 4 dB below input saturation. Uplink power control at the hub station is anticipated during fading conditions.

The link budget in Table 2 indicates the increasing traffic capacity which can be supported with reducing transponder gain. Taking the first column (maximum transponder gain) of Table 2 as an example, the values of carrier/noise temperature for the uplink $(C/T)_{up}$ and intermodulation $(C/T)_i$ are indicated.

This level of input signal produces an EIRP 48.1 dBW on the downlink (20 GHz) at the edge of coverage. Using the 99% case (i.e. 2.7 dB rain attenuation), the C/N_o achieved at the picoterminal receiver is 56.2 dBHz. Thus a data rate of about 42 Kbit/s can be supported at the highest gain setting with an E_b/N_o = 10 dB.

Some form of spread spectrum modulation is essential to provide protection against interference from similar systems in adjacent orbits.

If the bit rates supported in Table 2 are halved, then the channel will maintain the same equivalent E_b/N_o when the level of interference (I_o) equals the noise level (N_o). Then the relative total level of the interfering signals determines the minimum spreading chip rate. It can be shown that chip rates of 3, 4.6 and 5.4 Mcps, are required to maintain data rates of 21, 31 and 37 kbps on the outbound path in the presence of interference from similar networks in adjacent orbit slots.

Table 2

Outbound Capacity from Single Hub

Input power for saturation	-114.0	-111.0	-108.0	dBW
(C/T)up	-152.6	-147.9	-144.2	dBW/K
(C/T)i (inter modulation)	-151.4	-149.7	-149.0	dBW/K
Signal eirp (EOC)	48.1	49.8	50.5	dBW
Free space loss	-210.0	-210.0	-210.0	dB
Atmospheric loss (99%)	-2.7	-2.7	-2.7	dB
Picoterminal G/T	-7.7	-7.7	-7.7	dB/K
(C/T)dn	-172.3	-170.6	-169.9	dBW/K
(C/T)tot	-172.4	-170.7	-170.0	dBW/K
(C/No)	56.2	57.9	58.6	dBHz
Data rate	46.2	47.9	48.6	dBHz
Data rate	42.2	62.3	73.3	Kbps

Inbound traffic capacity. A similar calculation can be performed to derive the inbound path capacity. However some basic differences should be emphasised. The picoterminal will transmit in the presence of many other simultaneous transmissions. A realistic case assumes that the particular picoterminal is both faded by 4.3 dB and mispointed, whereas all the other systems are unfaded and on boresight. Even with this worst case and the down link to the hub station faded by 7.4 dB, the full transponder can support 1230, 3670 and 8540 simultaneous picoterminal transmissions at data rates of 440, 297 and 180 bps respectively, when received by a hub station of similar G/T to TDS-6. (Transponder gain settings are identical to those of Table 2). It is again necessary to use spread spectrum to overcome the adverse interference from other picoterminal networks, which could produce an interference level 5.4 dB higher than the total power of networks under consideration. Chip rates of 20.4, 40.7 and 58.9 Mcps are necessary to maintain individual picoterminal data rates in the region of 220, 150 and 90 bps respectively. However the combined effect of other interfering networks also reduces the number of terminals in any one network by a factor of 3.5.

Chip rates in excess of a few Megahertz are undesirable and splitting the 40 MHz into several bands is a practical alternative to reduce the code rates.

In a practical system both the inbound and outbound channel could share the same transponder, although the power sharing would need to be controlled carefully to maintain the correct traffic balance. More details of this picoterminal concept are contained in references 1 and 4.

Voice traffic calculations on Olympus. The above traffic calculations apply to a fully operational system which would have to coexist with other similar systems in adjacent orbits. However for the purposes of demonstration, many of the conservative margins imposed on these calculations could be removed for a pure demonstration exercise.

On the outbound path the Eb/No of 10 dB could be reduced to 7 dB as a low error rate is not necessary for a voice channel. A further gain of 3 dB could be achieved from fairly simple coding techniques.

Further gains could also be achieved by operating only in non fading conditions for demonstration (rainfall only occurs for less than 5% of the time in most of Europe). The margin improves by a further 2.5 dB. The accumulated gain of 8.5 dB on the outbound link budget is equivalent to a factor of just over seven in traffic capacity.

Thus the data rates at the three gain settings increase to about 300, 442 and 520 kbit/s. Even if this data rates is reduced to accommodate the spread spectrum modulation, then the total capacity is more than enough for at least 10 speech channels at 9.6 kbit/s.

Following the same logic and assuming the inbound link budget can be improved by 1) reducing the Eb/No level to 7 dB (1 dB gain), 2) using coding providing a further 3 dB, 3) operating in unfaded condition (3.7 dB), and 4) reducing the pointing misalignment by 1 dB, the overall improvement is then 8.7 dB or a factor of 7.4 in the supported data rate.

A data rate of 3.4 kbit/s is marginal for a speech channel. An increase in the picoterminal EIRP to 0.5 W delivered to the antenna seems desirable to bring the data rate above the 4.8 kbit/s required for a reasonable speech channel.

The demonstration system could be operated using conventional time division multiplexing (TDM) on the outbound path and single channel per carrier (SCPC) on the inbound route. However this would not demonstrate the advantages of spread spectrum random access systems, particularly on the inbound path. If a ten voice-channel system is envisaged on Olympus, using a picoterminal network, then one compromise solution of TDM on the outbound path could be combined with a spread spectrum system for the return signals from the picoterminals to the hub station.

A ten voice channels system would require 96 kbit/s on the outbound path, equivalent to 32% of the transponder capacity. The remaining capacity could be allocated to the inbound link which could support up to 400 channels at 6.6 kbit/s by extrapolation from the previous traffic calculation. If a spreading code modulation system is used, the channel throughput rate reduces. However with only ten channels operating simultaneously and in the absence of interference from other systems, then the reduction in throughput rate is much lower than the 50% calculated above. For instance a processing gain of 25 dB allows ten channels to operate with a reduction in the date rate of ony 14%, ie more than adequate to support a 4.8 kbit voice channel.

These rather tentative calculations suggest that the Olympus satellite could provide a useful demonstration for a voice channel network operated through a number of these very small portable picoterminals.

MOBILE APPLICATIONS AT 20/30 GHz

The above discussion of the picoterminal concept has concentrated on portable rather than mobile applications. To progress the concept to the mobile application requires consideration of a number of new constraints, most of which are related to propagation and antenna aspects of the system.

A truly mobile system should in principle operate with a specified systems performance, anywhere in the coverage area. This requirement is difficult for any mobile satellite system, as propagation constraints are always present. However increasing the frequency of operation of a satellite system from the currently utilized bands around L Band (1.5/1.6 GHz) to 20/30 GHz does not impose insurmountable propagation problems. The building blockage problems only deteriorate marginally. However the problems imposed by vegetation are significantly worse with acceptable transmission through anything more than a small bush being very difficult to accommodate with any realistic link margin (5).

However, the greatly increased rain attenuation in the millimetre bands does not contribute as much to the system outage time as would be expected. Only a 4 dB margin is necessary to provide a 99% service. This availability is much higher than expected from building, terrain and vegetation effects, even at the much lower L-Band frequencies.

Mobile antenna considerations do require a higher level of accuracy than those currently being developed at L-Band. The maximum proposed gain of steerable antennas at L-Band is about 12 dB, producing 3 dB beam widths of about 50°. The antennas proposed for the picoterminal are equivalent beam widths near 10°, ie requiring nearly an order of magnitude increase in the search and tracking system, when compared with the L-Band situation.

It would be a poor engineering compromise to increase the antenna beam width to alleviate these antenna system problems, as the high achievable gain for practical sized antennas is one of the main advantages of 20/30 GHz mobile systems. The incentive should be to reduce the beam width even further by increasing the antenna dimension thus increasing the traffic capacity of the overall system. It is debatable whether the ultimate limitation will be the antenna size which can be readily mounted on a mobile or the constraints imposed on the search and tracking mechanism.

The antenna steering configuration for these millimetric bands will probably follow the trends at lower frequencies, where combinations of patch arrays and electro mechanical steering have been developed for systems which have to accommodate the range of relatively low elevation angles (10° to 50°), associated with geostationary systems.

Although a number of other mobile system problems which arise from doppler shift, intermittant signal blockage and multipath occur, the portable picoterminal concept could be extended to provide conventional mobile services.

NON-GEOSTATIONARY ORBITS

There seems to be an even more over-riding case to investigate non-geostationary orbits for 20/30 GHz than at lower frequencies for certain parts of the world - especially Europe. Firstly the propagation margin for rain could be reduced even further. Typically the rain occurs in the first 3 km of the troposphere. Thus 3 mmh^{-1} of rainfall rate (the level exceeded for 1% of the time) produces an attenuation of less than 0.6 dB at 30 GHz on a vertical path. If satellites using the highly elliptical Molniya or Tundra orbits (Figure 2) are considered, then the elevation angle to the satellite is always above 50° for all of Europe. (Figure 3). The rain attenuation is still less than 2 dB and many of the building, terrain and vegetation attenuation problems disappear as a result of the much higher elevation.

The antenna tracking problem on the mobile is also reduced as a vertically pointing antenna only needs to track about ± 4 beam widths to cover the complete range of situations possible. With this configuration (ie the near zenith inclined orbit system) the hand held portable becomes a practical possiblity. For a geostationary system using a high gain antenna, it is necessary to have a sophisticated phased array tracking antenna and to know approximately where the satellite is situated to avoid blockage from both the operator and buildings. For the inclined orbit system however, the operator only needs to point the antenna vertically upwards. A simpler tracking system or a wider beam antenna would lock onto the satellite transmission much more easily.

CONCLUSION

The concept of the picoterminal network operating at 20/30 GHz through an Olympus type transponder has been demonstrated to support several thousand low data rate channels for telex traffic in an operational environment. This concept of the very small portable earth terminal can also be extended to demonstrate a voice traffic system, again through an Olympus type transponder. However an operational system would need to be supported through a much more powerful satellite such as ACTS (6). The portable concept can be extended to fully mobile applications. However geostationary orbit configurations are a poor

compromise with inclined orbit systems offering even more advantages for the 20/30 GHz mobile satellite systems than their lower frequency counterparts.

REFERENCES

1. **Barton S K and Norbury J R,** 20/30 GHz Microterminal Study, ESA Contract Report CR(X) 2715.

2. **European Microwave Conference,** June 1987, Special Session on Japan, pp 87-112. Papers by
   (a) **Akaike M,**
   (b) **Fukuta M and Harachi X,**
   (c) **Hori S,**

3. **Smith P M et al,** Millimetre Wave Low Noise and Power HEMT's; MILCOM 86, Monterey, California, paper 35.3, 1986.

4. **Barton S K and Norbury J R,** Ultra Small Earth Stations/Picoterminals for 20/30 GHz Sattelite Communication Systems, Olympus Utilization Conference, Vienna, Austria 1989, pp 3-8.

5. **CCIR Report 236-6,** Volume V, pp 96-101, Geneva 1986.

6. **Naderi F M and Campanella S J,** NASA's Advanced Communications Technology Satellite (ACTS), AIAA, 12th International Communications Satellite Systems Conference, Virginia, 1988.

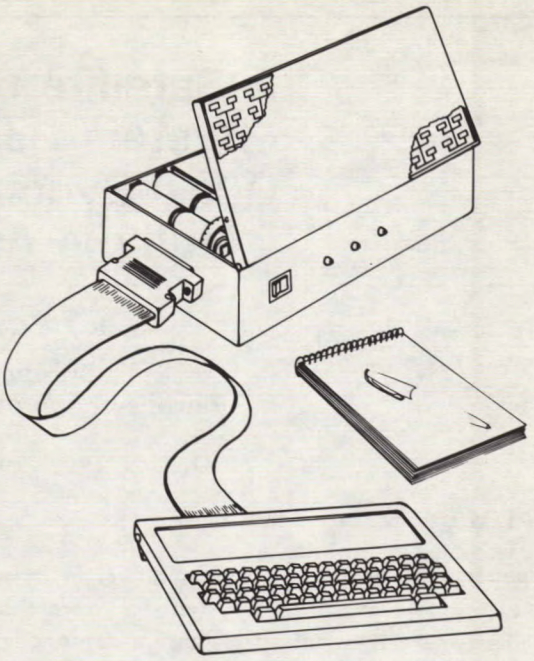
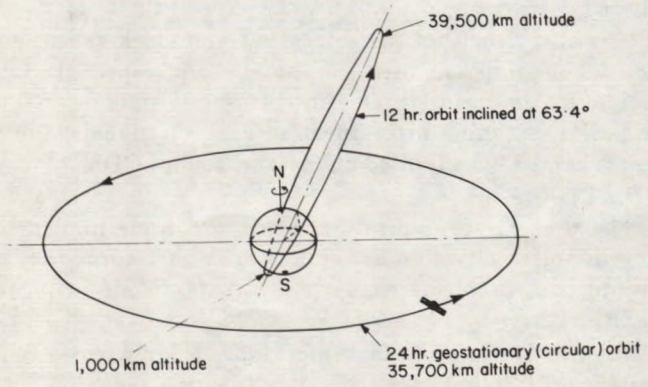


Figure 1: Picoterminal



Molniya (12 hour) and Geostationary orbits

Figure 2

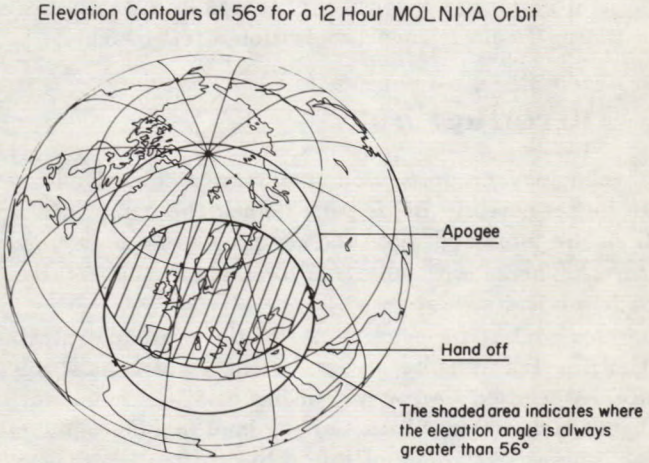


Figure 3

A Satellite Mobile Communication System Based on Band-Limited Quasi-Synchronous Code Division Multiple Access (BLQS-CDMA)

R. De Gaudenzi, C. Elia, R. Viola

European Space Agency (ESA)
European Space Research and Technology Centre
RF System Division
P.O. Box 299, 2200 AG Noordwijk The Netherlands

Abstract

In this paper we discuss a new approach to code division multiple access applied to mobile system for voice (and data) services based on Band Limited Quasi Synchronous CDMA (BLQS-CDMA). The system requires users to be chip synchronized to reduce the contribution of self-interference and to make use of voice activation in order to increase the satellite power efficiency.

In order to achieve spectral efficiency, Nyquist chip pulse shaping is utilized with no detection performance impairment.

The synchronization problems are solved in the forward link by distributing a master code, whereas carrier forced activation and closed loop control techniques have been adopted in the return link.

System performance sensitivity to non-linear amplification and timing/frequency synchronization errors is analyzed.

1 Introduction

As technology progresses offering services for mobile users with higher satellite EIRP per channel, the bandwidth limitations are more and more the key problem of high capacity services. This is especially true for L-band, where currently mobile services are supposed to operate.

In this contest an efficient use of the available spectrum, optimizing the satellite access system, is of capital importance. A selected access system has also to account for the difficult environment characterizing land mobile communications, where users are frequently subjected to deep shadowing and multipath conditions.

ESA is now embarking in activities centered around deployment of a new satellite system dedicated to Europe Mobile Service (EMS). This satellite will provide business voice/data services in the framework of private land mobile networks (PMR).

EMS will have an initial development phase based on a global beam coverage of Europe and subsequently it will be upgraded with a multi-beam antenna system.

In this framework studies to optimize the access scheme and the network are presently carried out. Code division multiple access is one of the candidates.

Considering the hostile environmental conditions, CDMA was always regarded as an attractive but inefficient access scheme For satellite mobile services. Techniques for increasing the CDMA efficiency were recently introduced [3]. The key idea is to reduce the voice duty cycle by employing carrier voice activation which reduces the average number of contemporary users. Moreover CDMA allows frequency reuse by polarization discrimination, being the cross-polarization effects reduced by the code diversity. In a multi-beam environment further advantages have been predicted in ref. [3].

A more substantial advantage which has never been exploited in traditional asynchronous CDMA systems is the drastic reduction of self-noise interference when synchronous CDMA is used [6]. Synchronous CDMA means that every user is using a common clock reference and on this basis the signature sequence transmission is aligned.

The attractive performances of S-CDMA described in ref. [6], refers to the unfiltered case. In practical system spectral compactness is of primary relevance and shall be achieved with minor end-to-end performance degradation.

In the following we show how the use of Nyquist chip pulse shaping allows to maintain the same performances of the unfiltered case.

It is also clear that perfect carrier and clock synchronization among different users can not be kept especially in the mobile environment [7]. Due to the user carriers incoherency and relative timing jitter affecting different signals, the system is named band limited quasi-synchronous CDMA (BLQS-CDMA).

The use of synchronization yields to simple implementation of voice activation in the forward and return link, thus avoiding all problems connected with fast code acquisition and tracking.

Voice activation is in fact important also in a synchronized system providing power savings and further reduction of the self-interference.

2 System Overview

Our reference system is a private mobile network composed of a transparent satellite transponder, a distributed feeder link

and a large number of mobile users distributed over a vast geographical area (see Fig. 1). The access system is CDMA both in forward and return link.

As depicted in Fig. 1, the system is composed by a network coordination station (NCS), a certain number of fixed earth stations (FES) connected with the ground network or directly placed at corporate user location. FES's provide two way voice/data communications with the mobile user terminals (MTs).

Each MT has the capability to transmit/ receive compressed digital voice, data, signalling information. A special code is reserved to the master station called *master code* (MC) [5], [6]. This code (not necessarily of the same type of user codes), is broadcasted to every MT and is used for frequency and time synchronization purposes. The signalling low data rate information is direct sequence spread by the MC and transmitted from the NCS in TDM fashion to the MT's

A master-to-user signal power ratio in the range of [5-10] dB can be set without significant loss of capacity. This is extremely beneficial for the network performances in case of shadowing. In fact, the mobile terminal drastically reduces traffic channel reacquisition time by increasing master channel availability,. Moreover master signal level monitoring at MTs provides a tool for open loop power control.

To each mobile terminal two codes are assigned, one for the I ($c_p(t)$) and one for the Q channel ($c_q(t)$).

Return link synchronization is achieved through MT's time/frequency comparison with the master reference. MT's synchronization errors are detected at the FES, then coded and looped back to the corresponding MT's together with the useful information.

In the following paragraphs a description of the elements composing the system will be given, detailing the mechanisms generally described in this section.

3 System Description

3.1 Modulation Technique

The selected modulation scheme is basically a direct sequence QPSK with independent I and Q signature sequences. *Ad hoc* chip shaping, satisfying the Nyquist criterium, is employed in order to obtain a band-limited signal. This point is of major relevance because more traditional filtering approach can affect the S-CDMA small sensitivity to self-noise interference. The use of DS/QPSK with independent I-Q codes is essentially dictated by system robustness requirements to non ideal conditions (i.e. non-linear amplifier, non perfect synchronization etc..).

The simulated spectrum at the satellite transponder input and output for a reduced scale (M=7, roll-off=0.4) BLQS-CDMA is presented in Fig. 2 and 3. For sake of comparison Fig. 2 also reports the spectrum of the signal with rectangular shaping. As HPA output filter an elliptical filter with two sided bandwidth of 1.5 the chip rate and stop band attenuation of 40 dB has been used. Observe the compactness and absence of sidelobes for both input and output signal's spectrum.

3.2 Synchronization Considerations

Each FES synchronizes its signature sequence start epoch, clock and carrier frequency using the received MC. In order to keep start epoch alignment, each station compares the phase of its own signature sequence looped back by the satellite with the NCS master reference. Measured timing error will be used to correct possible drifts. A similar concept apply for the FES carrier frequency and code timing control.

Delay compensation in MT terminals is instead performed by the FES's in closed loop to keep the mobile terminal as simple as possible.

3.3 Network Coordination Station Description

The NCS is dealing with :

- network master reference broadcasting and signalling,
- dynamic signature sequence allocation to MT's.

To fulfill to the above mentioned tasks the NCS requires a dedicated communication link with the FES.

3.4 Fixed Earth Station Description

A FES schematic block diagram is given in Figure 4. The received signal after IF conditioning feeds a bank of spread spectrum demodulators differing only in the signature code used for signal despreading.

Two of them named *echo demodulator and master demodulator* / are dedicated to FES synchronization errors detection.

The demodulated signals enter the Central Processor Unit which performs all necessary baseband processing such as:

- baseband demodulated user signals routing,
- signalling messages decoding,
- FES synchronization to the master reference,
- MT's synchronization error detection and coding

The digital baseband signal to be transmitted to the mobile users is composed by a synchronized set of different signature codes.

The resulting SS IF signals with equal power and carrier frequency are combined and up-converted to the transmission frequency prior final amplification. Signature sequence timing and carrier frequency is controlled by the so called *timing/frequency control unit.*

3.5 Mobile Terminal Description

In Figure 5 a principle block diagram of the user terminal is presented.

The RF-Front End of the *j-th user* delivers the IF signal which contains all the DS/SS signals plus the master code.

The master code is acquired and tracked by the dedicated sub-systems depicted in Figure 5. The on-time master code replica is used for despreading the master data stream which is routed to the modem central processor. Time epoch and

clock information supplied by the master code tracking loop drive the user I and Q codes replica generator.

The analog vocal signal on the transmitter side is first compressed, and then convolutionally coded. The I and Q symbol streams are logically added to the respective codes and used for carrier PSK modulation.

The up-link signal switch is controlled by the speech activity detector which is intended to interrupt the transmission during speech pauses.

The demodulator is being developed using full DSP approach and will include FIR chip pulse shaping filter equally shared between the transmitter and receiver side in order to achieve signal band-limitation.

4 System Performance Evaluation

The key novelty of the system consists in the use of a synchronous multiple access. Therefore the user BER evaluation based on the average SNR [2], or on the Gaussian interference approximation [1], is not anymore directly applicable. In ref. [8] it has been shown the validity of Gaussian interference approximation upon appropriate second order moment derivation.

Detailed derivations for the probability of error for the synchronous and quasi-synchronous cases are reported in [6], [7] respectively.

It has been shown that for S-CDMA the probability of error for the $j-th$ user is well approximated by:

$$\begin{aligned} P_e^j &= \frac{1}{2}[Q\left(\sqrt{\frac{\frac{2E_b}{N_0}}{1+V\mu_j^p\frac{(M-1)}{2G_p{}^2}\frac{E_b}{N_0}}}\right) \\ &+ Q\left(\sqrt{\frac{\frac{2E_b}{N_0}}{1+V\mu_j^q\frac{(M-1)}{2G_p{}^2}\frac{E_b}{N_0}}}\right)] \\ Q(x) &= \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp(-\frac{x^2}{2})dx \end{aligned} \quad (1)$$

where the factor μ_j^l named *quadratic average cross-correlation factor* is defined as

$$\begin{aligned} \mu_j^l &= \frac{1}{(M-1)}\sum_{i=1,\, i\neq j}^{M}[C_{i,j}^l]^2 \\ C_{i,j}^l &= \sum_{k=1}^{N} c_{k,i}^l c_{k,j}^l \quad l=p,\, q \end{aligned} \quad (2)$$

where M is the number of users, T_b is the data bit duration, T_c is the chip duration, E_b/N_0 is the energy per bit to thermal noise density ratio, $G_p = T_b/Tc$, $c_{k,i}^{p,q}$ is the $k-th$ chip of the $i-th$ p, q sequence of length N and V is the voice activation factor.

From formula (1) we see that the effect of the so called CDMA *self-noise* can be reduced increasing the factor G_p (usually named *spreading factor*), and/or reducing the average quadratic cross-correlation coefficient μ_j^l.

In particular in approximated formula (1) we observe the self noise effect reducing with $\mu_j {G_p}^{-2}$. For Gold codes $\mu_j = 1$ and it can be shown a self noise reduction with respect to the asynchronous CDMA equal to G_p [6].

Furthermore, thanks to the system chip coherency, we can minimize for a given number of users the average quadratic cross-correlation factor μ_j^l, defined by eqn. (2) and characterizing the family of signature codes used.

The selection of code family for the proposed system is the trade-off result between the number of required quasi-orthogonal codes and their cross-correlation properties around the origin.

New code families matched to this particular application, having good cross-correlation properties around the origin and a larger size than Gold codes, are presently being investigated.

4.1 Numerical Results

System simulation is based on a time domain approach. The number of simultaneous users has been limited to 7 simply to limit the computation burden. The processing gain G_p has been consequently reduced to 15.5 to be representative of a large scale loaded CDMA system. The BER evaluation technique is based on the semi-analytical approach already used in a previous work [8].

It is easy to show that theoretical results for unfiltered S-CDMA system from [6], also apply to the case of BL-CDMA when the optimal receiver structure like the one drawn in Fig. 6 is used.

This theoretical result has been confirmed by simulation as shown in Fig. 7 where the theoretical BER for rectangular shaped chips (continuous line), is compared with simulations results (dots) obtained making use of Nyquist shaped chip pulses (roll-off factor $\alpha = 0.4$).

The previous results are in very good agreement with the theoretical calculation and Fig. 7 clearly shows the self-noise reduction achieved through system synchronization.

In order to assess the system robustness to the non linear characteristics of the high power amplifier, extensive simulations have been carried-out.

Fig. 8 also shows the simulated BER (cross) using a typical satellite L-band solid-state amplifier driven with 2.5 dB of output back-off. Minor BER degradation is observed around the nominal working point of 4 dB E_b/N_0.

The results obtained are expected to be even better employing an higher processing gain.

Simulated BER sensitivity to the HPA output back-off at two different working points (dots and cross) is displayed in Fig. 8. Those results are slightly better than those already presented in [8] for the case of rectangular shaped chips (* and +).

Moreover the effect of the satellite HPA output filter has been analyzed by simulation. Using either an 8 poles Butterworth with two-sided bandwidth 1.5 the chip rate or an elliptical filter like the one described in section 3.1, no practical difference with the results of Fig. 8 has been observed.

Another point of major interest is the BER sensitivity to timing/frequency synchronization errors. To assess the impact of combined timing/frequency errors we have assumed an uniform frequency uncertainty of 100 Hz representative of

mobile motion and local oscillator instability, and we have derived the BER as a function of the maximum chip timing jitter. The calculation has been performed supposing uniform time jitter distribution and Gold codes signature sequences.

Fig. 9 compares theoretical (continuous line) and simulation results for rectangular chips (dots) with those obtained by simulation using two different roll-off factors (cross and stars).

In ref. [7] smooth BER dependence on frequency uncertainty range has been demonstrated.

5 Capacity Considerations

It can be shown that for the maximum number of users in the forward link ($[M_{CDMA}]_{fw}$) the following formula applies:

$$[M_{CDMA}]_{fw} \simeq \frac{[\frac{N_0}{E_b}]_{req} - [\frac{N_0}{E_b}]_{ut} - X_{PIM}}{VR_b\{\frac{L_d L_f K}{P_s G_s [\frac{G}{T}]_m} + \frac{S_{IM}}{(NPR)W_s} + \frac{\rho a_f L_{nl}}{W_c}(\frac{\overline{\mu} R_b}{W_c} + \frac{1}{K_{rice}})\}} \quad (3)$$

where:

L_d	:	Path loss at the edge of satellite coverage.
L_{nl}	:	Satellite transponder non linear amplifier loss.
L_f	:	Average forward link fading loss.
ρ	:	Polarization discrimination factor.
P_s	:	Satellite available output power.
$[\frac{G}{T}]_m$	:	Mobile receiving antenna G/T.
G_s	:	Satellite transmitting antenna gain.
a_f	:	Satellite antenna discrimination factor.
K	:	Boltzmann constant.
$[\frac{E_b}{N_0}]_{req}$	:	Required $\frac{E_b}{N_0}$ at the demodulator input.
$[\frac{E_b}{N_0}]_{ut}$	:	Uplink Hub to satellite $\frac{E_b}{N_0}$.
X_{PIM}	:	Passive intermodulation noise factor.
R_b	:	User bit rate.
S_{IM}	:	Satellite amplifier intermodulation spreading factor.
NPR	:	Satellite amplifier noise-to-power ratio.
$\overline{\mu} = \frac{1}{M}\sum_{i=1}^{M}\mu_i$	:	Global average quadratic cross-correlation factor.
W_c	:	Spreading codes chip rate.
W_s	:	Satellite channel bandwidth.
K_{rice}	:	Rician fading factor.
V	:	Voice activation factor.

More detailed description about the above parameters can be found in reference [4].

The return link formula is not shown because when including the return link propagation model it becomes of not trivial explanation, therefore out of the scope of the present work.

Observe in eqn. (3) that $\overline{\mu}$ (*global quadratic cross-correlation factor*), is in general dependent from $[M_{CDMA}]_{fw}$ hence eqn. (3) has to be solved with respect to the variable $[M_{CDMA}]_{fw}$. For particular code families like Gold it can be shown that $\overline{\mu} = 1$ hence eqn. (3) becomes of immediate application with an upper bound given by N+2 (Gold family size).

6 Conclusions

We have presented a satellite mobile communication system based on band limited quasi synchronous CDMA (BLQS-CDMA). Advantages of maintaining network synchronization at chip level have been shown to consist in an easy mobile terminal synchronization even in presence of voice activation and shadowing together with CDMA self-noise reduction. This fact implies improved link transmission quality and, for the return link, reduced sensitivity to the near-far effect due to self-noise effect mitigation. Chip pulse shaping has been introduced to achieve signal band limiting. System performances have been verified through time domain computer simulation including channel non-linearity, filtering and imperfect timing and frequency synchronization. Finally a system capacity expression has been derived and its meaning and range of applicability discussed.

References

[1] A. J. Viterbi, *"Spread Spectrum Communications-Myths and realities ,"* IEEE Communication Magazine, No.5 May 1979

[2] M.B. Pursley, *"Performance Evaluation for Phase-Coded Spread-Spectrum Multiple Access Communication-PartI and Part II,"* IEEE Transactions on Communications, VOL. 25, No. 8, August 1977.

[3] I.M. Jacobs, K.S. Gilhousen, L.A. Weaver, K. Renshaw, T. Murphy, *"Comparaison of CDMA and FDMA for the Mobilestar System ,"* Proceedings of the International Mobile Satellite Conference, Pasadena California, May 3-5 1988

[4] R. De Gaudenzi, C. Elia, R. Viola, *"High Efficiency Voice Activated CDMA Mobile Communication System Based on Master Code Synchronization: System Analysis,"* ESA Working Paper 1543, September 1989.

[5] R. De Gaudenzi, R. Viola, *"Système de Communication à Access Multiple par Repartition à Codes avec Porteuse Activée par la Voix de l' Usager et Synchronisation par Code* ESA Patent Filed in France, August 1989.

[6] R. De Gaudenzi, R. Viola, *"High Efficiency Voice Activated CDMA Mobile Communication System Based on Master Code Synchronization,"* In the proceedings of the IEEE Global Telecommunications Conference GLOBECOM, Dallas, Texas, November 27-30 1989.

[7] R. De Gaudenzi, C. Elia, R. Viola, *"Performance Evaluation of Quasi-Synchronous Code Division Multiple Access (QS-CDMA) for Satellite Mobile Systems,"* To be presented at the IEEE Global Telecommunications Conference GLOBECOM, San Diego (CA), December 2-5 1990.

[8] R. De Gaudenzi, R. Viola, *"A Novel Code Division Multiple Access System for High Capacity Mobile Communication Satellites,"* ESA Journal Vol.13, No. 4, December 1989.

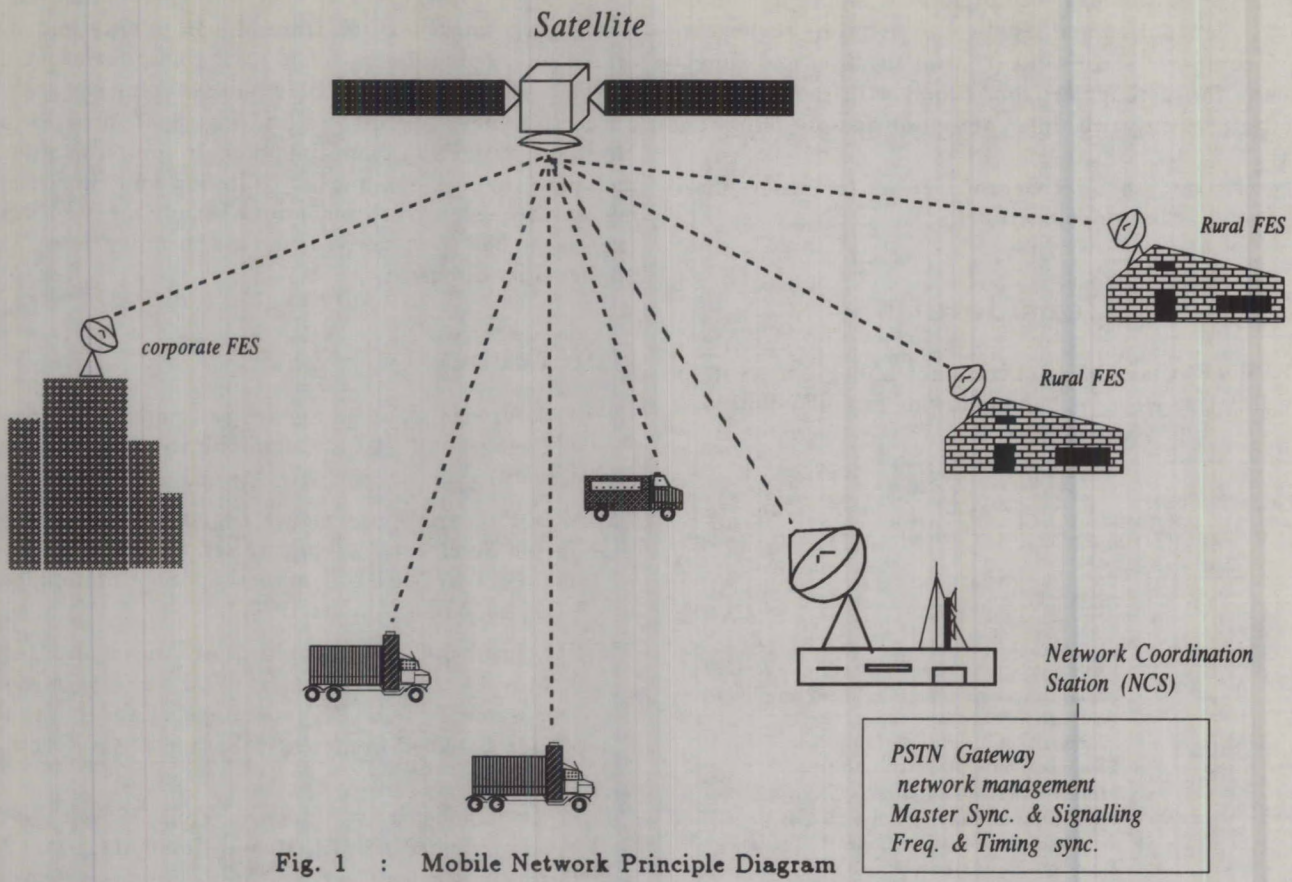


Fig. 1 : Mobile Network Principle Diagram

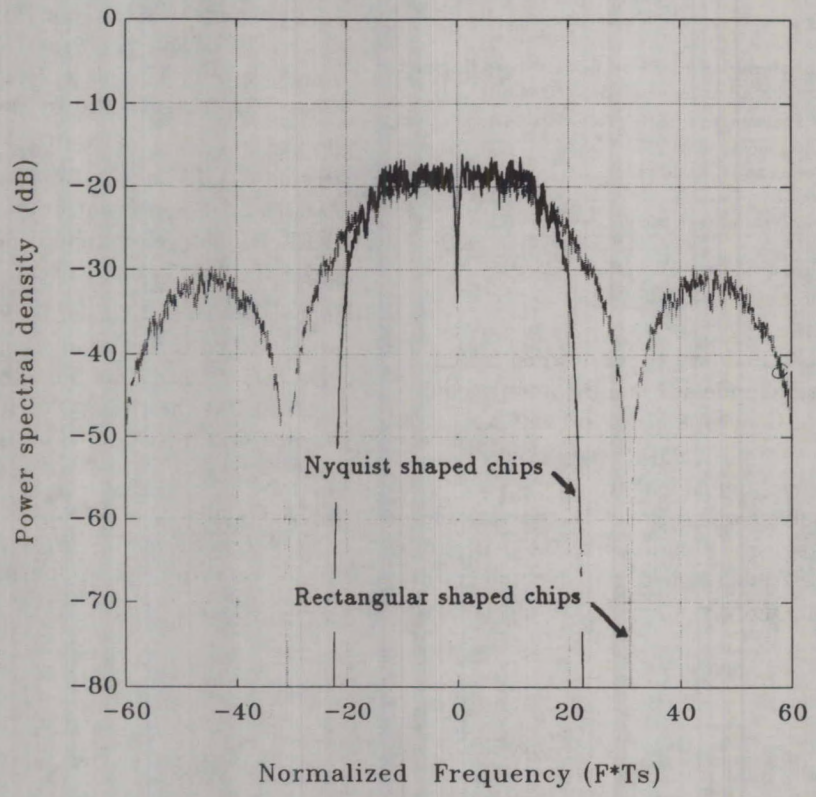


Fig. 2 : BLQS-CDMA spectrum: satellite input

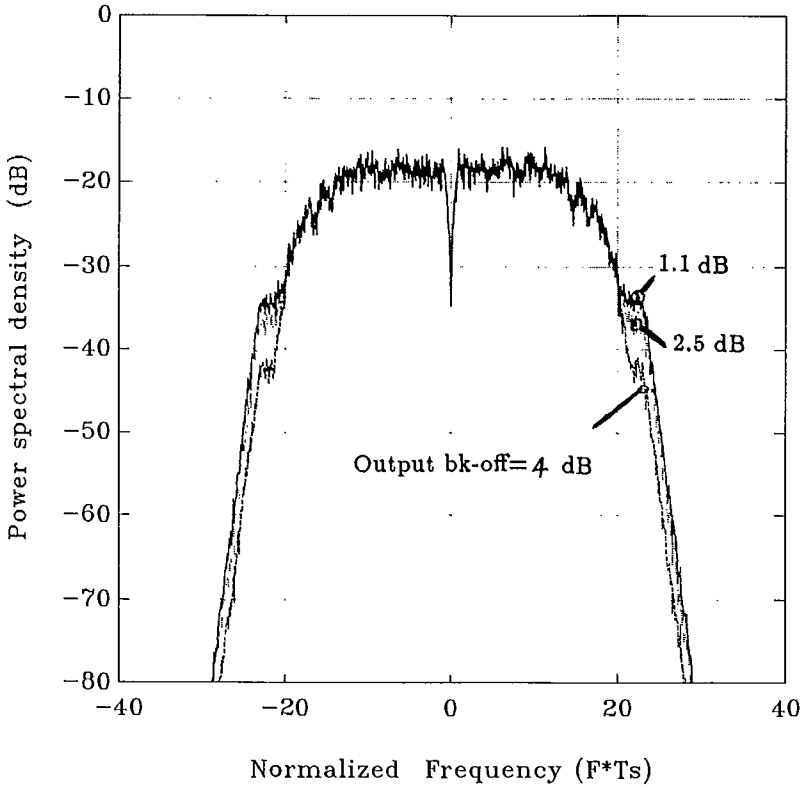


Fig. 3 :
BLQS-CDMA spectrum:satellite output

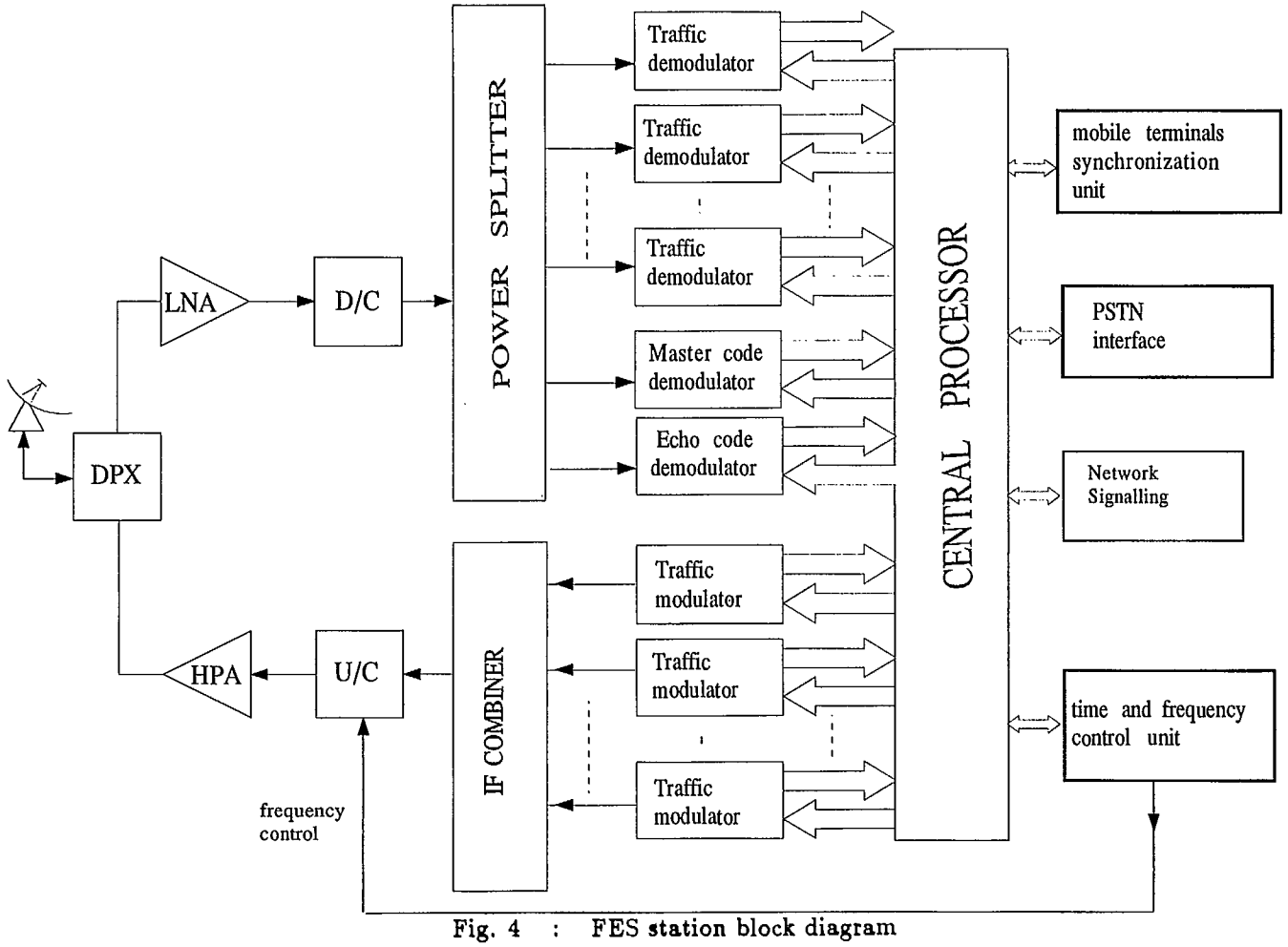


Fig. 4 : FES station block diagram

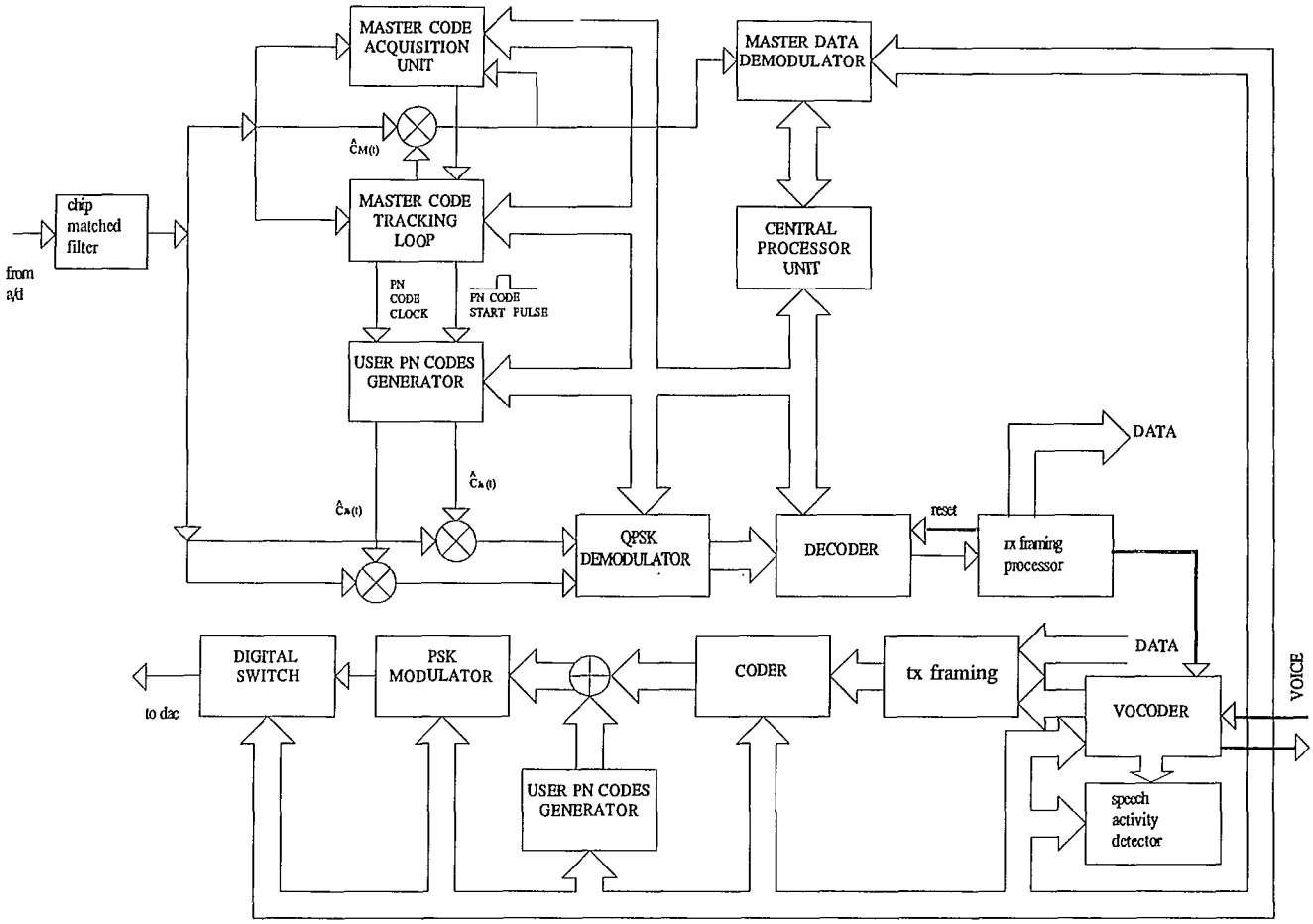


Fig. 5 : Mobile user terminal block diagram

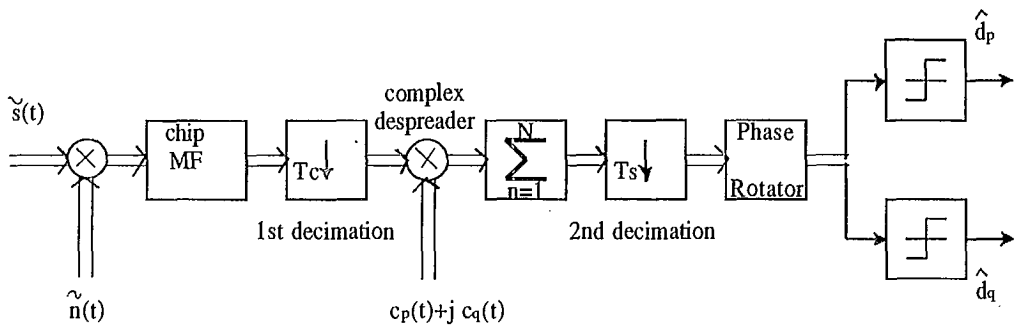


Fig. 6 : BLQS-CDMA demodulator block diagram

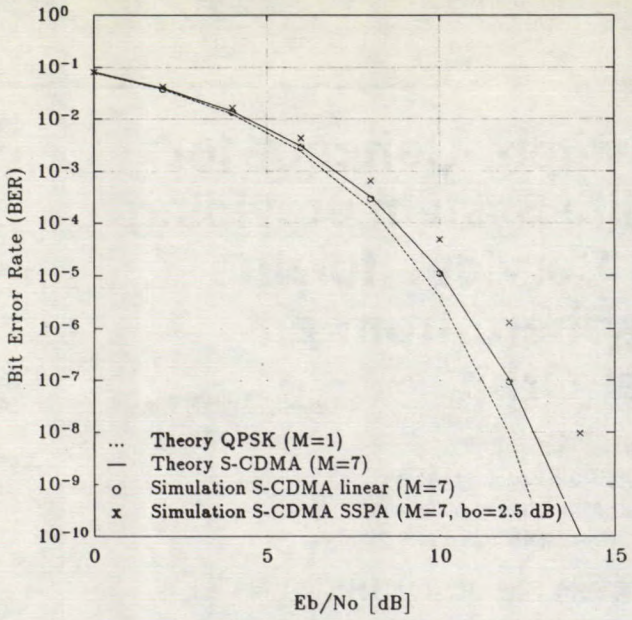


Fig. 7 : Probability of error vs. E_b/N_0, $G_p = 15.5$, $M = 7$

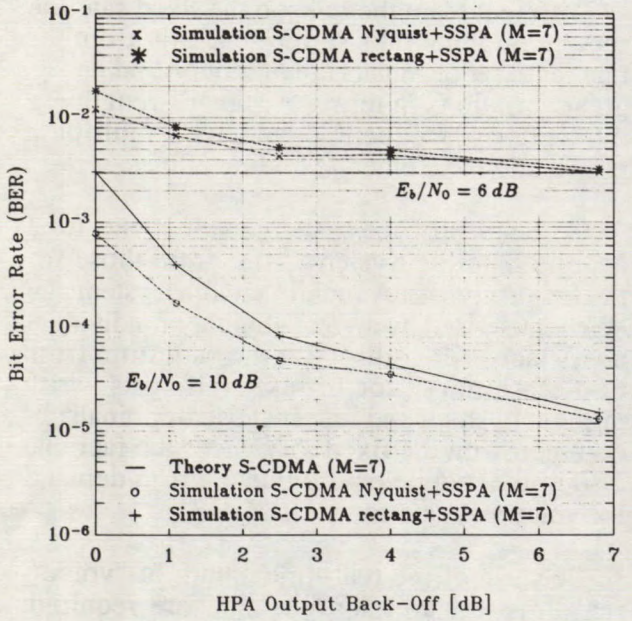


Fig. 8 : Probability of error vs. SSA back-off, G_p=15.5, $M = 7$

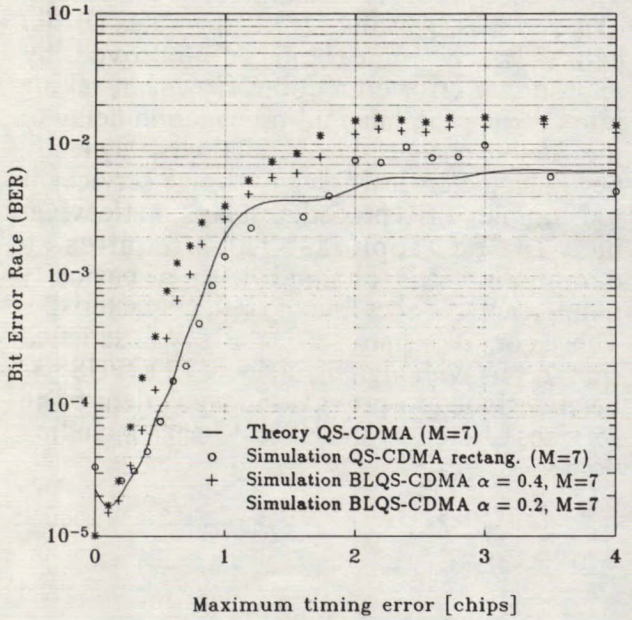


Fig. 9 : Probability of error sensitivity to timing jitter

LOOPUS Mob-D: System Concept for a Public Mobile Satellite System providing Integrated Digital Services for the Northern Hemisphere from an Elliptical Orbit

H.Kuhlen, P.Horn
MBB - DEUTSCHE AEROSPACE (DASA)
P.O. Box 80 11 69; D 8000 Munich 80, FRGermany
Phone +49 - 89 - 607 23542 FAX: +49 - 89 - 607 28998

ABSTRACT

The market for Mobile Communications Services is presently in a process of revolutionary development not only in terms of quantity but also in terms of quality of services. Sophisticated end user equipment allows intelligent signalling and comfortable user interfaces at reasonable prices.

In 1988 the Groupe Spécial Mobile (GSM), a group jointly sponsored by European PTTs and industry, submitted a set of standards for a fully digital mobile communications system to be implemented in the 90s in at least 18 European countries. However, providing Mobile Services at any place within Europe implies a major investment on one hand and a considerable implementation period on the other. Satellite systems, inherently, would not only offer an immediate availability in the entire anticipated service area once they are implemented, but also provide options to extend a regional service into a European or even worldwide accessability.

This paper will introduce a new concept for a satellite based public mobile communications system, LOOPUS Mob-D, where most of the "classical" problems in mobile satellite systems are approached in a different way. The LOOPUS system will offer a total capacity of 6000 high rate channel in three service areas Europe, Asia and North America, covering the entire Northern Hemisphere, with a set of GSM compatible mobile services eventually providing the "office-in-the-car".

This paper we will briefly highlight special characteristics of the LOOPUS orbit and the communications network architecture.

SYSTEM REQUIREMENTS

Different from the links in the fixed satellite service, mobile links suffer generally from the fact that user terminals are in motion and hence, great variations in terms of signal strength are induced by frequency selective multipath propagation and other degrading effects.

Consequently, the system requirements for a Mobile Satellite System differ from those for the fixed services. A mobile satellite system can be characterised by a great number of individual users accessing directly the satellite from non-stationary locations. All this with preferably small and low cost user terminals. In common with the fixed services is the fact that also in the mobile environment a firm demand for voice and data services exists.

Because of the real-time nature of "voice", real-time transmission channels are required. This is considerably less critical in the case of data services. Absolute Bit Error Rates (BER) can to a certain extend be improved by appropriate error correction. However, taking into account the inherent propagation delay of the satellite channel the overall time delay really becomes the critical issue for voice services if too complex FEC procedures, e.g. interleaving, has to be applied. This requires a communications channel with a basically uninterrupted direct line of sight characteristic. These conditions are particularly well satisfied from orbits with high inclination. LOOPUS has been designed to provide the equivalent of all services offered by the GSM [1] hence including voice .

The maximum number of user channel has been evaluated as 2000 32kbit/s (full-rate) or 4000 16kbit/s (half-rate) respectively. The channel granularity reflects a 32kbit/s channel for a voice signal with speech encoding i.a.w. a Regular-Pulse Excitation combined with Linear Predictive Coding (RPE-LTP) resulting in bitrate of 13kbit/s, i.e. 260bit per 20ms. A detailed description of this encoding process is given in refs.[2,3]. This capacity is determined by a number of system parameters i.e. the size of the satellite (thermal dissipation), the DC provisions of the solar generator, the efficiency of the DC/RF conversion in the power amplifier and, last but not least, by the overall gain of the spacecraft and user antenna.

One of the interesting aspects of an inclined elliptical orbit is that the satellites appear to be quasi-stationary during a long period around the apogees. It actually remains visible for several hours at a very high elevation angle for the user terminals particularly for those in high northern lattitudes. This fact led to the name for this program LOOPUS: (Quasi-Geostationary) **L**oops in **O**rbit **O**ccupied by **U**nstationary **S**atellite.

Due to the relative movement of the spacecraft with respect to the Earth the track of the subsatellite points creates two cross-over points. The elapsed travelling time of each satellite between the one closer to the apogee, through the apogee itself and back to the same cross-over point takes about 8 hours of 14.4 hours for an entire orbit. The coherency in the travelling of all nine satellites in their orbits creates for an earth bound observer the effect shown in Fig. It seems as if all satellites would follow each other on the same ground track. One satellite after the other becomes "visible" until after 71h 46min the number one satellite shows up again.

By appropriate selection of the six parameters defining such orbits, the hand-over from the descending to the ascending satellite, referred to the cross-over points, can be handled by one earth station. It is further ensured that any of the five resulting apogee loops is occupied by exactly one satellite at a time ref.[4]. Since all satellites are only active transmitting within the defined loop, the remaining time of the orbit can be used for attitude and orbit corrections by means of an electrical ion thruster system.

System Design Rationale

Not particularly for mobile systems, but here this aspect is more critical than in fixed systems, is the generic conflict of illuminating large service areas such as one third of the Northern Hemisphere with sufficiently high power flux density (PFD). Large coverage and high PFD are basically conflicting requirements. As a result from a first analysis on maximum antenna size, a footprint of about 1000km in diameter has been determined as optimum. This spot corresponds to an antenna gain of 38.6dBi which would also be in line with the requirements from link calculations.

In terms of frequency selection, the restrictions induced by the rather limited bandwidth of the L-band for mobile satellite services, the investigations for LOOPUS have led to a different solution. It is expected that in the mid 90s the entire L-band is almost worldwide occupied by geostationary satellites, particularly by 2nd and 3rd generation INMARSAT. An interference analysis showed that frequency re-use (e.g. by polarization decoupling) can not be guaranteed under all mobile link propagation conditions. We, therefore, came to the conclusion, that the Ku-band provides a good compromise between sufficient bandwidth for further services on one side and a reasonable link attenuation on the other.

Today, the Ku-band already receives much attention mainly in the field of direct broadcasting and VSAT services. The growing market for Ku-band equipment in consumer technology has therefore not only significantly increased their availability at reasonable prices but also improved their performance (noise figures).

The application of Ku-band implies the use of high gain antennas for mobile terminals. This requires on one side a minimum tracking capability of the mobile antenna but on the other side offers the advantage of inherent frequency re-use due to large angular decoupling between satellites on inclined orbits from those on the geostationary orbit.

SYSTEM CHARACTERISTICS

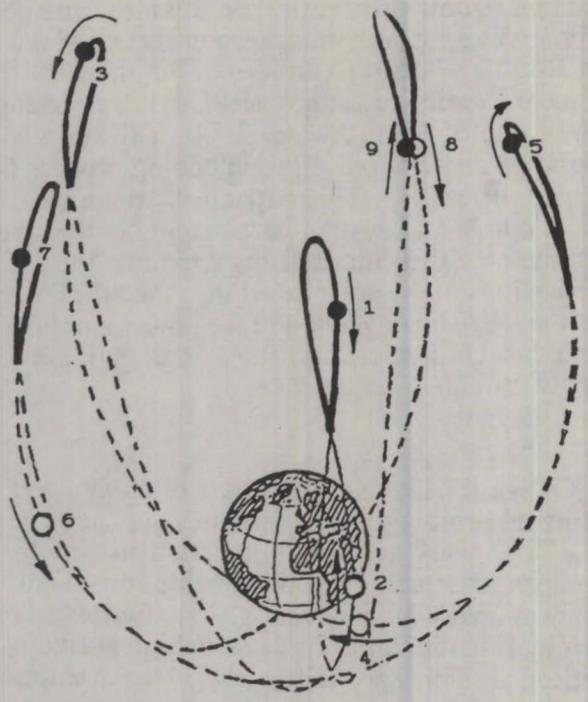
Highly-inclined Elliptical Orbit

Main characteristic of the full scale system is that nine satellites will eventually travel along three highly inclined (63.4°) elliptical orbits (HEO), three on each orbit. The selected orbit configuration with heights of 41450km and 5784km for the apogee and perigee respectively provides a simultaneous 24h-visibility in three major economic regions of the Northern Hemisphere: Europe, North America and Asia. The large potential visibility of the entire Northern Hemisphere and the resulting Earth coverages due to the high inclination is shown in Fig.

However, to cover the anticipated service area would still require 75 spots of this size. It is obvious that this approach for a fixed spot configuration is not only technologically but also economically unfeasable. Moreover, many of the provided spots would quite probably never require any communications at all. On the other side also areas of less frequent communications should not be discarded from the beginning (e.g. distress communications). A comprehensive market survey particularly for the LOOPUS system has not yet been performed. Comparison with or extrapolations from the terrestrial mobile market is inherently incomplete because the provisions of the potential service areas offered by this system are unique and are therefore hard to compare to other systems.

The conflict of high gain antennas with a large service area is solved in the LOOPUS case by means of an electrically steerable (flying) spotbeam implemented by an active phased array antenna. Thus, only one physical spot is used but capable to be directed and dwell on each required spot position on demand. Hence, no area is excluded per se but also no energy is wasted for unused spots.

Two active phased array antennas, one for receive one for transmit, are controlled by an on-board computer, allowing flexible beam hopping in accordance with a permanently updated, i.e. demand oriented, scan pattern.



Communications System Architecture

A number of up to 40 MSCs have been considered for each of the three regions to provide multiple access points to the terrestrial networks(ISDN or PSTN). This would provide one MSC in any significant area avoiding long and costly landlines. For a mobile call from Manitoulin Island e.g. into Ottawa it would be uneconomic to downlink via the Vancouver network and route by landline into Ottawa. Functionally, these MSCs are gateways between the satellite system protocol used on the space links and the required signalling protocols of the connected networks.

The operation of the satellite payload is performed in close cooperation with the Network Control Station (NCS). In the NCS all network management functions are implemented such as the

- configuration management
- fault management
- performance management
- security management
- accounting management

Also the special functions, Home Location and Visitor Location Register (HLR, VLR), are implemented in the NCS.

The spotbeams and the link concepts have briefly been outlined before. Now, Fig. shows the elements of the system architecture in one of the three defined regions. The satellite plays an active role in the operation of the system. It actually interconnects the mobile subscriber as well as the Mobile Switching Centers by providing all communications links.

The Communications System

The characteristics of the mobile channel, taking into account the high nominal elevation angle for mobile terminals, is expected to be dominated by direct propagation components. Particulary the degradations introduced by multipath propagation are less significant than those in L-band.

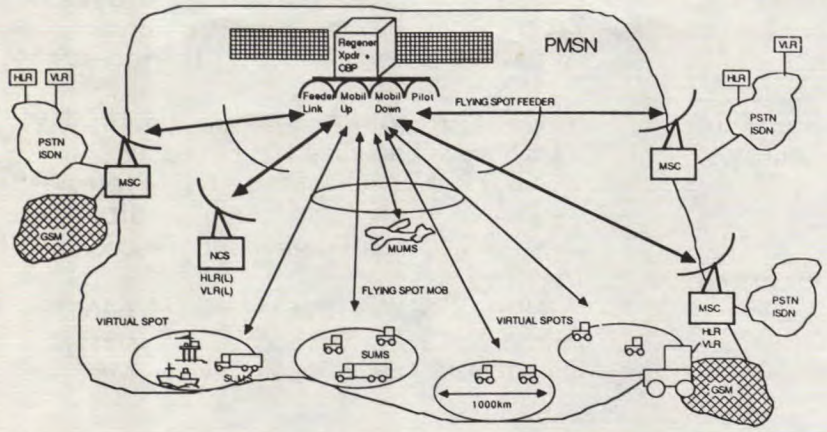
The flying spot concept makes a traditional FDMA approach impossible. Therefore, the existence of a time "slicer" (flying spot) leads to the application of TDMA, correlated and synchronized with the hopping of the beam. The resulting high bitrate for the indicated system summes up to 69Mbit/s, i.e. 2048 channel @ 32kbit/s each. The actual capacity of each virtual spot position is directly related to its dwell time.

Admittedly, at first sight, the high bitrate has a frightening impression due to its extraodinary magnitude. Theoretically, there is no difference in applying TDMA or FDMA to a given communication channel, it is simply not more than an exchange of time for frequency. However, a number of technical aspects do require a closer look before a final conclusion on the applicability can be taken. It is out of the scope of this paper to discuss the technical implications of using high bitrates, but a few remarks shall be given at this point.

The system is operated in Ku-band, i.e. 14GHz. At this frequency the ratio bitrate vs. bandwidth is moderate, hence power amplifiers and other frequency selective devices only have to provide a moderate bandwidth capability. As a further consequence, the unavoidable Doppler shift due to the relative movement of the spacecraft to the users becomes negligible against the necessary bandwidth of the broadband signal. The applied powerful modulation scheme is Rectangular Spectrum Modulation (RSM ref.[5]) providing excellent spectrum efficiency, fast acquisition combined with technological feasability.

The intelligent payload comprises a regenerative transponder with processing and routing capabilities and active phased array antennas. As mentioned before, these two antennas of $1m^2$ each are reserved for the reception and transmission of the mobile signals. A third combined receive/transmit antenna provides four spots for the Mobile Switching Centers.

All links are interconnected on baseband level, i.e. channel by channel, in the on-board switching processor. Control and the demand updating of the dynamic timeplan is performed in the NCS and permanently communicated to the on-board processor. The execution of the timeplan is then performed by the on-board processor. The updated timeplan reflects the actual change in demand in terms of channel capacity, dwell period and beam scan pattern of the system. With the availability of intelligent on-board facilities, many other functions in support of the overall system operation can be performed leading to autonomy in the overall system operation. This includes time management, present position calculation, information dissemination, initial access coordination monitoring and others.



PLMN - Public Land Mobile Network
MSC - Mobil Switching Center
NCS - Network Coordination Station
HLR - Home Localizing Register
VLR - Visitor Localizing Register
SUMS - Single User Mobil System
MUMS - Multiple User Mobile System

Flying Spot Antenna System
=> 93 Virtual Spots at 1000km diam
40 MSCs
Total System Capacity: 2000 Channel

SYSTEM ARCHITECTURE

OF ONE LOOPUS Mob-D PMSN

From looking at the overall topology of the LOOPUS system it becomes obvious that due to the excellent inherent coverage of practically any area of the Northern Hemisphere including the Northpole the offered services can be comprehensive. The clear cut separation of land, aero or maritime mobile diminishs in such a system environment. It has been pointed out that originally the orbits have been optimized for a coverage of the Northern Hemisphere. An extension into the Southern Hemisphere is hampered by the fact that the re-use of the Ku-band frequencies is achieved only as long as sufficient angular decoupling from systems in the geostationary orbit is ensured.

SYSTEM OPERATIONS

A beacon signal is transmitted via a global antenna from each active satellite to enable automatic tracking of the ground antenna. By receiving this signal, the system availability, system timing, direction of satellite and other system information can be obtained. In addition this beacon also serves as a return channel to a service request from a user terminal. The user terminal is inhibited to transmit during times that the beacon signal is not present. This beacon transmits on a separate frequency. It is important to note that the beacon can be received everywhere in the entire region and not only in the areas where the spot beam dwells.

A call can be initiated at any time the user sees the "ready" message on his terminal. After dialing, a short initial access burst carrying origin and destination information is transmitted from the user terminal on a separate frequency (slotted ALOHA coordinated by the beacon). This request is routed to the NCS. The NCS will in response acknowledge the request, check the authentication of the user card and will eventually allocate a channel in the on-board switch together with an update in the antenna scan pattern control. As a consequence, the spot will either extend its dwell time for another channel period or a new virtual spot position will be included in the pattern for the time of communication. At the same time the NCS determines the optimum gateway, i.e. the Mobile Switching Center closest to the desired destination of the call.

In analogy to the GSM, the call, its duration, service category and so on will be registered in the Visitor Location Register of the satellite system. The VLR function is located in the NCS. All information is then automatically communicated to the Home Location Register of the caller. So, one bill where all calls are listed is received at the end of the month.

Any satellite user terminal receives its network identity only after a smart card has been inserted. The insertion, even without an immediate call set-up, can be communicated to the system if desired and will create an update in the VLR. From that moment on, the user can be reached by any other subscriber in the system or caller from any other telephone in the world.

CONCLUSIONS

In this paper the HEO satellite based mobile communications system LOOPUS has been briefly described. It has been shown how the LOOPUS mobile subscriber can, via satellite, either communicate (voice, data, FAX, etc.) directly to each other or have access to any national public network (ISDN or PSTN) within three defined service areas: Europe, Asia and North America. Through the use of a sophisticated signalling in line with recommendations of the GSM group home and visitor location register (HLR/VLR), any subscriber in the entire LOOPUS system (i.e. Northern Hemisphere) would be within reach of any public telephone throughout the world.

REFERENCES

1. **GSM Recommendation 02.02:** Bearer Services supported by a GSM PLMN
2. **Vary, P. et.al.** 1988. Speech Codec for the European Mobile Radio System. Proceedings International Conference on Acoustics Speech and Signal Processing, ICASSP-1988, pp. 227-230
3. **GSM Recommendation 06.10:** Full Rate Speech Encoding and Decoding
4. **Nauck, J.; Günther, H.; Plate, H.** 1987 A new type of orbits for INMARSAT's 3rd generation, Proceedings IAP-87-481
5. **United States Patent** No: 4 646 323 Feb. 24, 1987. Method and System for Digital Data Transmission; Inv.: K. Meinzer

Use of Elliptical Orbits for a Ka-Band Personal Access Satellite System

Masoud Motamedi and Polly Estabrook
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, California, 91109
Tel: (818) 354-8041
FAX: (818) 393-4643

ABSTRACT

This paper examines the use of satellites in elliptical orbits for a Ka-band personal communications system application designed to provide voice and data service within the continental U.S. The impact of these orbits on system parameters such as signal carrier-to-noise ratio, roundtrip delay, Doppler shift, and satellite antenna size, is quantized for satellites in two elliptical orbits, the Molniya and the ACE orbits. The number of satellites necessary for continuous CONUS coverage has been determined for the satellites in these orbits. The increased system complexity brought about by the use of satellites at such altitudes is discussed.

INTRODUCTION

JPL is exploring the potential and feasibility of a Personal Access Satellite System (PASS) designed to offer the user freedom of access and mobility via the use of handheld and portable satellite terminals [1,2]. The currently conceived Ka-band system uses a geostationary satellite to support services such as voice, data, and video between a group of service providers, or suppliers, and users. The choice of Ka-band ensures ample spectrum for system growth and will reduce component weight and size thus enabling small user terminal size. Users and suppliers can be located anywhere within the continental U.S. (CONUS), hence the satellite must provide interconnectivity within CONUS. This is accomplished in the preliminary design by employing two CONUS coverage antennas and two multibeam antennas on the satellite. Alternative system designs are being investigated to enhance user capacity and/or reduce the system complexity [3]. The use of circular non-geostationary orbits has been studied and presented in [4]. This paper reports on the impact of using satellites in Elliptical Orbits (EO) on the PASS system design.

Two elliptical orbits, the Apogee at Constant time-of-day Equatorial (ACE) orbit and the Molniya orbit, are characterized in this paper. The selection of these orbits is motivated by the desire for CONUS visibility from one satellite. Thus signals between geographically separated earth stations within CONUS can be relayed from one satellite. This bypasses the need for intersatellite links. For these orbits, system parameters such as signal carrier-to-noise ratio, roundtrip delay, Doppler shift, and satellite antenna size, are discussed and the number of satellites required to provide continuous CONUS coverage is calculated. The relative advantages and disadvantages arising from their use are discussed as they relate to the PASS application. Finally a comparison between system characteristics obtained with elliptical and previously reported circular orbits is presented.

MOTIVATION FOR NON-GEOSTATIONARY ORBITS

Use of satellites in Non-Geostationary Orbits (NGO) is motivated by the possibility of reducing the EIRP and G/T requirements on the user terminals, lowering signal delay through the satellite, reducing satellite antenna size, supporting a global communication system, and decreasing the fade margin and blockage requirements for mobile vehicle applications [5]. Lastly the use of NGO satellites permits the consideration of a greater range of launch vehicles which may permit lower launch costs due to the use of simpler launch vehicles or the launch of several satellites per vehicle.

While NGOs offer a number of attractive features, other factors must be considered such as: the number of satellites and their control mechanism; the use of tracking antennas on Earth; the existence of large Doppler shifts; and the variations in link characteristics as the satellite moves accross the sky. In addition, the design of the NGO satellite will need to cope with radiation effects due to increased radiation exposure from the Van Allen radiation belt as well as support a complex antenna pointing mechanism. Finally questions of possible interference between geostationary and NGO satellites must be resolved (see [6] for details).

To date satellites in non-geostationary circular or-

bits have been proposed to provide a global mobile communications link at L-band [7]; EOs have been proposed to offer primary coverage in Europe for mobile users at L-band [5], to offer global coverage for personal and mobile users at Ku-band [8], and, to offload traffic from GEO satellites at peak traffic hours for fixed users in the U.S. at C- and Ku-bands [6,9].

ORBIT PARAMETERS

The ACE orbit is an elliptical equatorial orbit with five revolutions per day. It is sun-synchronous and highly eccentric (eccentricity = 0.49). The satellite is at the same point in its arc at the same time each day. The apogee and perigee of this orbit are 15,100 km and 1,030 km, respectively. This orbit has been studied extensively by Price et al. [6,9].

The Molniya orbit is a highly elliptical orbit at an inclination of 63.4°. With a perigee of 426 km and an apogee of 39,771 km, the satellite spends most of its orbital period ascending to its apogee. The maximum coverage period for CONUS is attained when the orbit's apogee is placed over the center of CONUS.

The following equations relate the period of the satellite's elliptical orbit around the Earth and the satellite velocity to the orbit's geometry. The former, τ_s, is given by:

$$\tau_s = \frac{2\pi a^{\frac{3}{2}}}{\sqrt{G * M}} \tag{1}$$

where

$$\begin{aligned} G &= 6.67 * 10^{-8}\ \mathrm{cm^3/gm\ sec^2}, \\ M &= 5.976 * 10^{27}\ \mathrm{gm}, \end{aligned}$$

and a is the major semiaxis of the ellipse, defined to be one half of the sum of apogee and perigee. The satellite velocity, v_s, can be expressed as:

$$v_s = \left[2G * M \left(\frac{1}{r_E + h} - \frac{1}{2a}\right)\right]^{\frac{1}{2}} \tag{2}$$

where r_E is the radius of the Earth, 6379.5 km, and h is the height of the satellite above the Earth. Table 1 gives the period, velocity and roundtrip signal delay (τ_d) when satellites in the ACE and Molniya orbits are at their apogees.

LINK CHARACTERISTICS

Doppler Shift

The Doppler shift of the signal is proportional to the velocity vector of the satellite relative to the Earth's motion. Specifically it is proportional to the component of this relative velocity vector which lies in the direction of the receiving earth station, $v_{rel_{in}}$. The Doppler shift, $f_{Doppler}$, can be written as:

$$f_{Doppler} = \pm \left(\frac{v_{rel_{in}}}{c} f_c\right) \tag{3}$$

where c is the speed of light and f_c is the signal frequency. To determine the Doppler shift, $v_{rel_{in}}$ must be found from the velocity of the satellite, the inclination angle of the satellite's orbit with respect to the Equator, and the angle between the satellite and the user terminal on Earth.

The velocities and locations of a satellite and an earth station can each be decomposed in terms of their three orthogonal components in the geocentric equatorial coordinate system. The origin of this coordinate system is the center of the Earth, its x-axis points towards the Sun and the z-axis coincides with the Earth's axis of rotation. The velocity vector, $\vec{v_E}$, of an earth station at location $\vec{P_E}$ is:

$$\begin{aligned} \vec{v_E} &= r_E\ [\sin(\omega t) + \cos(\omega t + \beta)]\ \cos\theta_l\ \hat{i} \\ &\quad + r_E\ [\cos(\omega t) - \cos(\omega t + \beta)]\ \cos\theta_l\ \hat{j} \\ &\quad + 0\ \hat{k} \end{aligned} \tag{4}$$

where ωt denotes the angular distance moved in period t (ω being the angular velocity of the Earth), β is the angle from the x-axis at $t = 0$, and θ_l is the latitude of the earth station. The position ($\vec{P_s}$) and the velocity ($\vec{v_s}$) of a satellite can be written similarly. They are omitted from the text due to length of their expressions but can be found in [10].

The relative velocity between the earth station and the earth station and the satellite can be written as:

$$\vec{v_{rel}} = \vec{v_s} - \vec{v_E}. \tag{5}$$

The vector between the earth station and the satellite is known as P_{ES}. The component of $\vec{v_{rel}}$ along the unit vector in the direction of $\vec{P_{ES}}$ sets the Doppler velocity, i.e.

Table 1: Characteristics of the ACE and Molniya Orbits

Orbit	Altitude km	Period hrs	v_s km/hr	τ_d msec
GEO	35,784	23.93	11069	239
ACE	15,100	4.8	11152	252
Molniya	39,771	12	5590	100

$$v_{rel_{in}} = \frac{\vec{v_s} \cdot \vec{P_{ES}}}{|\vec{P_{ES}}|} - \frac{\vec{v_E} \cdot \vec{P_{ES}}}{|\vec{P_{ES}}|}. \quad (6)$$

The Doppler shift of the carrier frequency, f_c, can be obtained by finding the velocity and position vectors of the satellite and the earth station. The Doppler shift is zero when both $\vec{v_s}$ and $\vec{v_E}$ are perpendicular to $\vec{P_{ES}}$. For both the ACE and the Molniya orbits, this occurs when the satellite is at its apogee for those earth stations at the same longitude as the satellite. Alternatively, the Doppler shift will be maximized for earth stations directly under the satellite as the satellite ascending to its apogee or descending from it. The maximum Doppler shift at 30 GHz is approximately 300 KHz and 600 KHz for satellites in the ACE and Molniya orbits, respectively.

These large Doppler shifts require a compensation mechanism or a modulation scheme capable of tolerating wide deviations. Doppler compensation techniques would be straightforward if all the communications were done by a central station. The central station would use a set algorithm to change the pilot frequencies going to the user terminals and the downlink frequencies to the supplier stations. However in PASS, where users for a given supplier can be located in different beams, frequency tracking and compensation would put a big burden on the Network Management Center (NMC). The NMC would have to keep track of the position (beam number) of all the users and change the inbound and outbound frequencies accordingly. This would require large guardbands and a real time frequency calculation for all of the active users and suppliers of the system. The implication of the Doppler shift on the complexity of the channel assignment routine is detailed in Appendix A.

Propagation Loss

The use of elliptical orbits leads to a changing path range between the satellite and the users. Pt. A in Fig. 1 depicts the moment at which all of CONUS is visible from the EO satellite. At this point the range from the earth station (at the closest edge of coverage) to the satellite, d_{min}, sets the minimum propagation loss. Pt. B depicts the satellite at its apogee; at this point the range from the earth station (at the farthest edge of coverage) to the satellite, d_{max}, sets the maximum propagation loss. Pt. C depicts the satellite at the last moment at which all of CONUS is visible. The propagation loss, L_P, can be calculated according to:

$$L_P = \left(\frac{4\,\pi\,d}{\lambda_c}\right)^2 \quad (7)$$

where λ_c is the wavelength corresponding to the carrier frequency f_c and d varies from d_{min} to d_{max}. The propagation loss along with its variation compared to GEO is shown in Table 2. For the ACE orbit, the improvement in the path loss ranges from 7.6 dB to 20.6 dB. For the Molniya orbit, propagation loss increases relative to GEO by 0.9 dB at orbit apogee.

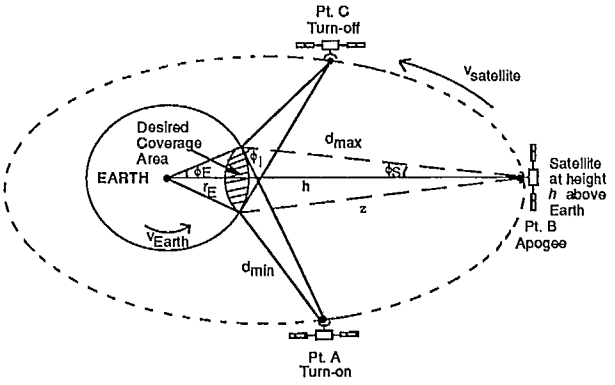


Figure 1: Earth station and satellite geometry.

Table 2: Variation of Propagation Loss

Orbit	Pt. A*		Pt. B†	
	30 GHz	ΔL_P‡	30 GHz	ΔL_P
GEO:	213.1dB	0dB	213.1dB	0dB
ACE:	192.5dB	20.6dB	205.5dB	7.6dB
Molniya:	186.0dB	27.1dB	214dB	-0.9dB

* 20° elevation, at turn-on, for earth stations at the farthest edge of coverage.
† 20° elevation, at apogee, for earth stations at the center of coverage.
‡ $\Delta L_P = L_{P_{GEO}} - L_{P_{NGO}}$.

CONUS Coverage Antenna

The preliminary PASS system design utilizes two CONUS beam antennas on the satellite for all communication between the supplier and the satellite. The size of the receive antenna is 0.1m and that of the transmit antenna is 0.15m; both have a gain of 26.9 dBi. In order to determine the effect of operating with EO satellites, the gain of a CONUS covering antenna is calculated for satellites at the apogee of their orbits. The satellite is assumed to have a tracking antenna which is trained on CONUS whenever the latter is visible. This scheme is illustrated in Fig. 1. As the satellite moves from Pt. A through Pt. B to Pt. C, the antenna tracks CONUS (shown by the hatched area). Antenna gain is calculated when the satellite is over the middle of the U.S., i.e. Pt. B in Fig. 1.[1]

[1] Antenna gain for an equatorial orbit could be slightly larger as the beamwidth required to see CONUS from the Equator is

Antenna gain can be written as:

$$G_{ant} = \rho \frac{4\pi h^2}{2\pi h^2(1-\cos\phi_S)}, \quad (8)$$

where the numerator represents the surface area of a sphere of radius h and the denominator is the surface area into which the radiated power is directed and ρ is the aperture efficiency. For CONUS coverage, ϕ_S, ϕ_l, and ϕ_E in Fig. 1 are denoted by ϕ_{S_C}, ϕ_{l_C}, and ϕ_{E_C}, respectively. ϕ_{l_C} is then the minimum elevation angle required for CONUS coverage.

ϕ_{S_C} can be expressed in terms of the minimum elevation angle on Earth to the satellite from:

$$\phi_{S_C} = 90^\circ - \cos^{-1}\left(\frac{r_E \cos\phi_{l_C}}{r_E + h}\right). \quad (9)$$

To find the minimum elevation angle necessary to see a satellite from CONUS, seven points on the perimeter of CONUS are defined. They are listed in Table 3. The elevation angle, ϕ_l, is calculated for each location by determining the angle between that location and the satellite. Calculation of ϕ_l and θ_{ES} for any point and a satellite at any position is given in [4].

Here we calculate the gain of a CONUS antenna specifically for the case of a satellite in an inclined orbit such that it passes over the center of the US, i.e. its latitude is 40° N and its longitude is 95° W. Satellites in the inclined orbits under consideration will have optimal CONUS coverage at this point. Satellites in equatorial orbits will, of course, never pass over a latitude of 40° N but will have optimal coverage when their longitude is 95° W; the gain of their CONUS antennas will at that point be slighly greater than that of their inclined orbit counterpart as the antenna beamwidth will be slightly narrower.

Once the elevation angles are calculated for each of the locations in Table 3, then ϕ_{l_C}, or the minimum value of ϕ_l, can be found. ϕ_{S_C} can then be calculated from Eq. 9 and the gain of a CONUS antenna can be found from Eq. 8.

ϕ_{l_C}, ϕ_{S_C} and CONUS antenna gain G_C (for $\rho = 0.5$) are given in Table 4; ϕ_{E_C} is not given as it does not vary with the satellite's height. The minimum CONUS elevation angle, ϕ_{l_C}, for a satellite at 35,784 km (equivalent to the height of a geostationary satellite) can be seen from Table 4 to be 64°.

In Table 4 the 0-3 dB power beamwidth of the CONUS antenna, ϕ_{S_C}, increases from 3.8° to 6.7° for the ACE orbit and decreases to 3.6° for the Molniya orbit. Thus, the gain of the CONUS antenna falls from 26.6 dB to 19.1 dB for the ACE orbit and increases to 27.5 dB for the Molniya orbit. The change in CONUS antenna gain relative to a geostationary satellite, ΔG_C, is given in the table.

less than that required when the satellite is directly over the center of CONUS.

Table 3: Cities Considered to Bound CONUS

City	Latitude	Longitude
Bay of Fundy, Maine	47.2°	-68.0°
Key Largo, Florida	25.0°	-80.5°
Brownsville, Texas	26.0°	-97.0°
San Diego, California	32.5°	-117.0°
Seattle, Washington	49.0°	-123.3°
Bottineau, North Dakota	49.0°	-100.0°
Center of USA	40.0°	-95.0°

Table 4: Satellite CONUS Antenna Characteristics

Altitude	CONUS Satellite Antenna			
	ϕ_{l_C}	ϕ_{S_C}	$G_{C\|_{\rho=0.5}}$	ΔG_C†
35,784 km (GEO)	64°	3.8°	26.6 dB	0 dB
15,100 km (ACE)	57°	9.3°	19.1 dB	-7.5 dB
39,771 km (Molniya)	65°	3.4°	27.5 dB	+0.6 dB

†$\Delta G_C = G_{C_{NGO}} - G_{C_{GEO}}$.

Antenna diameter can be calculated from the standard gain equation ($\sqrt{\frac{G_{ant}}{\rho}} \cdot \frac{\lambda_c}{\pi}$) and the ensuing size reduction compared to GEO operation can be found. The results are given in Table 5.

Multibeam Antenna Gain

As currently envisaged, the PASS satellite uses two multibeam antennas (MBAs) for communication between the user terminals and the satellite. Each MBA has a beamwidth of 0.35° and uses an 142 beam feed network to cover CONUS. In the preliminary PASS design, the gain of both the transmit and receive MBAs is 52.5 dBi and their efficiencies are both taken to be 0.45, corresponding to a reflector diameter of 2m at 30GHz and 3m at 20GHz. As the satellite orbit height decreases, the spot beams must continue to cover the same area in CONUS. Therefore the required MBA beamwidth will increase from 0.35° and the gain of the MBA will decrease with decreasing satellite altitude.

In reference [4] the gain of a multibeam antenna, $G_{MBA_{NGO}}$, used in a NGO satellite is derived in terms of the gain of the multibeam antenna of a GEO satellite, $G_{MBA_{GEO}}$ and the ratio of the satellite heights. It can be written as:

$$G_{MBA_{NGO}} = \left(\frac{h_{NGO}}{h_{GEO}}\right)^2 G_{MBA_{GEO}}. \quad (10)$$

Antenna diameter and the ensuing size reduction compared to GEO operation are given in Table 5 when the satellites are at the apogees of their orbits for $G_{MBA_{GEO}} = 52.2$ dBi and $\rho = 0.45$.

Impact on Link Equations

The PASS system is asymmetrical: the user terminal equipment, designed to be handheld and portable, is less powerful than the supplier station, a fixed earth station with a 4m antenna. In both the link from the supplier to the user and the link from the user to the supplier, the channel between the user terminal and the satellite determines the overall carrier to noise of the received signal.

The change in downlink C/N and uplink C/N are given in [4]. They can be written as (in dB):

$$\Delta\frac{C}{N}_{down} = \left(G_{MBA_{t_{NGO}}} - G_{MBA_{t_{GEO}}}\right) + \left(L_{P_{down_{GEO}}} - L_{P_{down_{NGO}}}\right) \quad (11)$$

and

$$\Delta\frac{C}{N}_{up} = \left(G_{MBA_{r_{NGO}}} - G_{MBA_{r_{GEO}}}\right) + \left(L_{P_{up_{GEO}}} - L_{P_{up_{NGO}}}\right) \quad (12)$$

where the subscripts t and r refer to the transmit and receive gain of the MBA. Substituting Eqs. 7 and 10 into Eqs. 11 and 12, the latter can be shown to reduce a common expression, d being the range to the satellite:

$$\Delta\frac{C}{N} = \frac{h_{NGO}}{h_{GEO}}^2 \frac{d_{GEO}}{d_{NGO}}^2. \quad (13)$$

NUMBER OF SATELLITES REQUIRED

ACE orbit

Approximately eight satellites are required to provide continuous CONUS coverage from this orbit. The total coverage period and thus the number of satellites can be calculated from finding the turn-on and turn-off points of the satellite assuming that the satellite antennas have their boresights directed towards the center of CONUS during the coverage period. As depicted in Fig. 2 satellites in the ACE orbit can be phased in order to give a 24hr CONUS coverage. These satellites, with their orbit apogees at 0° latitude and -95° longitude, would turn their transponders on at the first point that $\phi_{l_{min}} = 20°$ and turn them off at the next $\phi_{l_{min}} = 20°$.

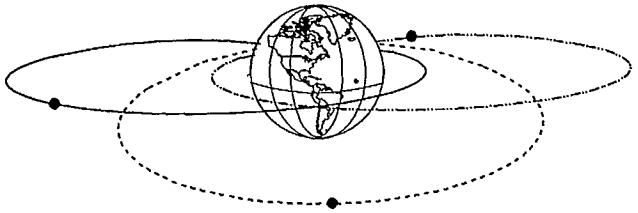


Figure 2: Satellite location for three satellites in three ACE orbits.

Molniya orbit

The satellites in the Molniya orbit are phased in the same fashion as the ACE satellites. Approximately three satellites, with their apogee at 37° latitude and -95° longitude, are required to provide continuous CONUS coverage from this orbit.

ADVANTAGES AND DISADVANTAGES

The relative advantages and disadvantages to using NGO satellites for the PASS application can be measured in terms of the parameters mentioned in the introduction. These parameters are given in Table 5 for the three NGO satellite altitudes studied in [4] at which CONUS coverage from one satellite is possible and for the two elliptical orbits considered. The propagation loss listed, L_{P_1}, represents the variation in propagation loss from the time when the satellite first becomes visible at $\phi_{l_{min}}$ to the time when the satellite is overhead, $\phi_{l_{max}}$. The link margin is calculated from Eq. 13 for satellites at their apogees for earth stations with $\phi_l = \phi_{l_{min}}$.

As can be seen from Table 5, the gain in signal power brought about by the reduced propagation loss when NGO satellites are used is outweighed by the loss in satellite antenna gain. For the PASS design, the degradation in uplink C/N means that both suppliers and users will have to transmit higher EIRPs; the degradation in downlink C/N will require both suppliers and users to increase their receive G/T. This most certainly means that the size of the user antenna could not be reduced: directional antennas will still be necessary as in the preliminary design. Moreover the user and supplier antennas will require satellite tracking mechanisms for both circular or elliptical orbits. Both user and supplier transceivers will need to implement techniques to compensate for the varying Doppler shift of the signal which at 20 GHz and 30 GHz is substantial, up to 344 KHz for satellite altitudes of 5,143 km. Lastly the variation in the link will necessitate some modification to the present rain fade

Table 5: GEO and NGO Satellite Parameter Comparison ($\phi_{l_{min}} = 20°$).

Orbit	Incl. Angle Ψ	Prop. Loss ΔL_{P_1}‡	Link Margin	$N_{sat.}$	Roundtrip Delay (Overhead)	Max. Doppler Shift at 30 GHz	Satellite Antenna † at 20 GHz CONUS	MBA
Circular Orbits								
GEO	0°	0	0	1	239 msec	0.0 KHz	14.4cm	3m
20,182 km	45°	1.4 dB	-0.6 dB	8	135 msec	131 KHz	8.2 cm (43%)	1.7 m (44%)
10,353 km	45°	2.3 dB	-1.5 dB	18	69 msec	216 KHz	4.35 cm (70%)	0.87 m (71%)
5,143 km	45°	3.5 dB	-2.7 dB	48	34 msec	344 KHz	2.3 cm (83%)	0.43 m (86%)
Elliptical Orbits								
ACE	0°	6 dB	-1.7 dB	8	252 msec	300 KHz	6.1 cm (58%)	1.3m (58%)
Molniya ($\phi_{l_{min}} = 41°$)	63.4°	5 dB	-0.4 dB	3	100 msec	600 KHz	16.0 cm (-11%)	3.3m (-11%)

† Reduction in antenna size compared to GEO is given in parenthesis.

control scheme wherein both users and suppliers utilize the signal strength of the pilot to determine the rain fade condition in their uplink and downlink.

A comparison of the orbits under consideration shows that the fewest number of satellites required for continual CONUS coverage is the Molniya orbit which requires only three. However this orbit has the highest maximum Doppler shift, the widest range in path loss, and requires one dB more margin compared to the geostationary orbit. The 20,182 km circular orbit has far less Doppler shift and negligible path loss change with less link margin degradation but requires 8 satellites. Lower orbits require greater numbers of satellites, have greater degradation in the link margin, possess higher variations in the path loss, and are characterized by high Doppler shifts; however they have greater advantages in terms of signal delay and reduction in satellite antenna size.

CONCLUSION

This paper has studied the effects on the link characteristics of satellites in two elliptical orbits. The number of satellites necessary for continuous CONUS coverage has been determined for these orbits. The relative advantages and disadvantages of using EO satellites for the PASS application are measured in terms of the following parameters: the impact on the signal uplink and downlink, the roundtrip signal delay, Doppler shift, reduction in the size of satellite antennas, and the number of satellites required to cover CONUS continually.

The advantages of using satellites in the studied orbits do not appear to outweigh the increased system complexity especially for the Personal Communication Satellite System application considered here.

ACKNOWLEDGMENTS

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

APPENDIX A. IMPLICATION OF DOPPLER SHIFT ON THE COMPLEXITY OF THE CHANNEL ASSIGNMENT ROUTINE

To a stationary observer, the frequency of a moving transmitter varies with the transmitter's velocity. If a stationary transmitter's frequency is at f_T, the received frequency f_R is higher than f_T when the transmitter is moving toward the receiver and lower than f_T when the transmitter is moving away from the receiver. This change in frequency, or Doppler shift, is quite pronounced for low orbiting satellites and compensating for it requires intense frequency assignment and tracking.

For a non-geosynchronous satellite system linking supplier stations to users scattered around CONUS, direct supervision of the frequency assignment by a network management routine is required. The man-

agement routine has to ensure that no two signals in any link overlap, that is for any two signals centered at f_n and f_m in a given link,

$$|f_n - f_m| > BW, \quad (14)$$

where BW is the required data bandwidth.

For the PASS forward link specified in the preliminary design [1,2], all the suppliers will have to compensate for the Doppler shift by tracking the incoming pilot from the satellite thus ensuring that the same nominal frequency is received by the satellite from all suppliers transmitting via a TDM/TDMA channel to a particular beam.

Denoting the signal destined to beam b by $F_{up}(b)$, onboard the satellite the signal is frequency shifted using a constant frquency $F_{for}(b)$ and is transmitted using beam b. This signal $F_{down}(b)$ received by user n, U_n, is in the form of:

$$F_{down}(b, U_n) = F_{up}(b) + F_{for}(b) + \text{Doppler}(U_n) \quad (15)$$

where $\text{Doppler}(U_n)$ is the Doppler shift due to the motion of the satellite as observed by the n^{th} user. The Doppler shift varies from one user to the other as a function of their positions. Following the channel assignment of Eq. 14 the network controller has to ensure that in an area where frequency reuse is not implemented, for any beam number p and q ($p \neq q$), and and i and j ($i \neq j$),

$$|F_{down}(p, U_i) - F_{down}(q, U_j)| > BW_{forward}, \quad (16)$$

thus requiring wide guard bands or a constant changing of the TDMA channels center frequencies based on the location of the users in communication.

In the return direction where SCPC channels are assigned to users for transmission, the assignment procedure includes selection of an even greater number of center frequencies. In this link, assuming that Doppler correction is done on the inbound channel, means should be taken to correct the transmit frequency so that overlap at the satellite are avoided. That is at the satellite for any k and l ($k \neq l$) in a region without frequency reuse:

$$|F_{user_k} - F_{user_l}| > BW_{return}. \quad (17)$$

Furthermore, the downlink from the satellite to the suppliers will require frequency correction based on the received pilot by the suppliers.

References

[1] Sue, M. K., Vaisnys, A. and Rafferty, W., *A 20/30 GHz Personal Access Satellite System Study,* 38th IEEE Vehicular Technology Conference, 15-17 June, 1988, pg. 453- 460.

[2] Estabrook, P., Huang, J., Rafferty, W., and Sue, M. K., *A 20/30 GHz Personal Access Satellite System Design,* IEEE International Conference on Communications, 11-14 June, 1989, pg. 7-4-1 to 7-4-7.

[3] Sue, M. K., Dessouky, K., Levitt, B., and Rafferty, W., *A Satellite-Based Personal Communication System for the 21st Century,* Inter. Mobile Satellite Conference, 17-20 June, 1990, these proceedings.

[4] Estabrook, P. and Motamedi, M., *Use of Non-Geostationary Orbits for a Ka-Band Personal Access Satellite System,* AIAA 13th International Communications Satellite Systems Conference, March 11-15, 1990, Los Angeles, California, pg. 14 to 24.

[5] Stuart, J.R., Norbury, J.A., and Barton, S.K., *Mobile Satellite Communications from Highly Inclined Elliptic Orbits,* AIAA 12th International Communications Satellite Systems Conference, March 13-17, 1988, Virginia, pg. 535 to 541.

[6] Price, Kent M., Doong, Wen, Nguyen, Tuan Q., Turner, Andrew E., and Weyandt, Charles, *Communications Satellites in Non-Geostationary Orbits,* AIAA 12th International Communications Satellite Systems Conference, March 13-17, 1988, Virginia, pg. 485 to 495.

[7] Richharia, M., Hansel, P.H., Bousquet, P.W., and O'Donnell, M., *A Feasibility Study of a Mobile Communication Network using a Constellation of Low Earth Orbit Satellites,* IEEE Global Telecommunications Conference, November 27-30, 1989, Texas, pg. 21.7.1 - 21.7.5.

[8] Nauck, J., Horn, P., and Göschel, W. ,- *LOOPUS - MOBILE - D - A New Mobile Communication Satellite System,* AIAA 13th International Communications Satellite Systems Conference, March 11-15, 1990, Los Angeles, California, pg. 886 to 899.

[9] Ford Aerospace, *The Use of Satellites in Non-Geostationary Orbits for Unloading Geostationary Communications Satellite Traffic Peaks, Volumes I and II,* NASA publication, CR-179-597.

[10] *Orbital Flight Handbook Part 1 - Basic Techniques and Data,* NASA SP-33 Part 1, National Aero. and Space Admin., Washington, D.C., 1963, pg. III-24.

Multi-carrier Mobile TDMA System with Active Array Antenna

Ryutaro Suzuki, Yasushi Matsumoto and Naokazu Hamamoto
Communications Research Laboratory, MPT
4-2-1 Nukui-Kitamachi, Koganei-shi
Tokyo 184, Japan
Phone: +81-423-21-1211
FAX: +81-423-24-8966

ABSTRACT

A multi-carrier TDMA system is proposed for the future mobile satellite communications which include multi-satellite system. This TDMA system employs the active array antenna in which the digital beam forming technique is adopted to control the antenna beam direction. The antenna beam forming is carried out at the base-band frequency by using the digital signal processing technique. The time division duplex technique is applied for the TDM/TDMA burst format, in order not to overlap transmit and receive timing.

INTRODUCTION

First generation land mobile satellite communication will be introduced with the mechanically tracking antenna and the 4.8 kbps digital voice terminal. For the future mobile satellite communications a multi-satellite system which uses Molniya or Loopus orbit is proposed for effective coverage of the high latitudes[1,2]. In this system, the satellites are moving in a highly elliptic orbit, and the switch over of the satellites is essential for continuous coverage. This will require automated satellite tracking, antenna control etc. Digital signal processing based modulation/demodulation, error correction are already in use in satellite systems. An extension of this technique to perform the antenna control and beam forming operations will be attractive for the future mobile satellite communication system.

This paper proposes a mobile satellite communication system which employs the digital beam forming (DBF) antenna[3] and TDM/TDMA as access techniques. The time division multiplex (TDM) is adopted for the forward link, and time division multiple access (TDMA) is adopted for the return link. The mobile terminal with the active array antenna can track more than two satellites simultaneously which allows easy and uninterrupted switch over of satellite links. However, the disadvantage of the active array antenna is the complexity of the antenna system in which many components such as the low noise amplifiers (LNAs), the high power amplifiers (HPAs) and the diplexers are integrated. In the proposed system, TDM/TDMA with time division duplex technique is adopted for the satellite link between the mobile earth station and HUB station, in order to remove the diplexer from each mobile antenna element.

TDM/TDMA SYSTEM

The simplified configuration of the mobile satellite communication system is shown in Fig. 1. In the earth station, the LNAs and the HPAs are connected to each antenna element. While receiving, the output of LNAs are down-converted and combined in the IF section before delivering to the demodulator. For transmission, the modulated signal is up-converted and sent to each antenna element through its HPA.

The transmit and receive signal power difference for usual satellite communication is more than 100 dB. Therefore, a good diplexer is needed to get high isolation ratio between transmit and receive port for the full-duplex communication system. However, it is quite difficult to get high isolation between the

transmitter and the receiver for the active array antenna system in which all RF component should be installed near the antenna elements.

The proposed TDM/TDMA system with time division duplex technique will overcome this difficulty. The features of this system are as follows.

1) Because of alternating operation, isolation between the transmitter and the receiver is needed only to protect the LNAs. Hence, RF switches can be use instead of the diplexers.
2) In a TDMA system, required transmitting power is proportional to the time compression ratio and usually higher than that of SCPC. However, the total power in the proposed system is divided to each HPA of the antenna array element. Hence, low power HPA can be used.
3) Transmitting and receiving frequency can be fixed in the TDMA system. Therefore, there is no need to employ a channel synthesizer.

DIGITAL BEAM FORMING ANTENNA

Fig. 2 shows the configuration of the phased array antenna for the receiver. In the case of the conventional phased array antenna, the antenna beam direction is controlled by using phase shifters which is connected to each antenna element, as shown in Fig. 2a. Therefore, the phase shifters should have low insertion loss. The resolution of the phase control is limited by the control bits of the phase shifters.

In the case of an active array antenna system, the antenna beam forming is carried out at IF stage of the receiver, as shown in Fig. 2b. There is no restriction on the insertion loss of the phase shifters. Also, the multi-beam function is easily available, because another phase shifters with summing port can be added without any loss of the signal quality. Furthermore, in the case of the DBF antenna, the antenna beam can be formed with high flexibility since amplitude compensation is possible.

The features of the DBF antenna are as follows.

1) The signals received at each antenna element are digitized and manipulated to change its phase and amplitude with high resolution. There is no limit to the resolution of the phase and amplitude control.
2) Satellite tracking is carried out by using the software control. Therefore, there is no need to prepare a special equipment for the satellite tracking function.
3) The multi-beam function can be achieved by adding calculation routine for another beam forming.
4) Difference of characteristics among each RF components can be compensated easily by using software without any dedicated circuitry.

For the land-mobile satellite communications, the size of the antenna system should be small enough to install on the roof of the vehicle. Therefore, the LNAs and HPAs in a DBF antenna system should be integrated just under the antenna elements, and high integration technique is required to combine all units. At present, required complexity of LSI is not available. Further developments are necessary for the digital LSI technique to realize such DBF antenna system. Then, the functions such as the antenna control, beam forming, modulation/demodulation, coding/decoding can be performed using digital signal processing.

LINK BUDGET FOR PROPOSED SYSTEM

In Japanese ETS-V satellite, a similar TDM/TDMA system has been designed. Table 1 presents a typical link budget for the system. For the calculation of the link budget, following conditions were assumed.

1) Maximum satellite power is used for the TDM link.
2) 4.8 kbps digital voice is used for this communication system.
3) Seven element active array antenna is adopted for the mobile antenna, because of its simplicity.
4) Separate antennas are used for transmission and reception respectively to reduce the feeder loss.

In Table 1, G/T for the mobile earth station is estimated to be -12 dBK at an elevation angle of 45 degrees. For this case, link budget estimate shows 60 dBHz of the carrier to noise power ratio (C/No). Therefore, the transmission rate of up to 64 kbps is available for the forward link. This means ten voice

channels are available for the link. For the return link, assuming same transmission rate as the forward link, the total transmitting power estimate of the mobile terminal is 25 W. This value is divided among each HPA connected to antenna array elements. As a result, 4 W HPA will have sufficient output power for each antenna array element.

Table 1. Link budget of the TDM/TDMA system

Forward link (TDM)		
satellite TX (ETS-V)		
HPA output power	11.0	dBW
Feeder & DIP loss	3.9	dB
Antenna gain	23.5	dBi
Satellite EIRP	30.6	dBW
Propagation loss	187.7	dB
Mobile station RX (Active array antenna)		
Antenna array element gain	4.0	dBi
System noise temperature	240	K
Antenna element G/T	-19.8	dBK
Number of array elements	7	
Coupling loss	1.0	dB
Total G/T	-12.4	dBK
Down-link C/No	59.1	dBHz
Total C/No	59.1	dBHz
Required C/No (64 kbps)	55.0	dBHz
Link margin	4.1	dB
Return link (TDMA)		
Mobile station TX (Active array antenna)		
Antenna array HPA output (3.5 W)	5.4	dBW
Antenna element gain	4.0	dBi
Coupling loss	1.0	dB
Total antenna gain	11.5	dBi
Total TX power (24.5 W)	13.9	dBW
Total EIRP	25.4	dBW
Propagation loss	188.3	dB
Satellite RX (ETS-V)		
Antenna gain	24.0	dBi
Feeder & DIP loss	3.1	dB
Satellite G/T	-5.5	dBK
Up-link C/No	60.2	dBHz
Total C/No	60.0	dBHz
Required C/No	55.0	dBHz
Link margin	5.0	dB

For Molniya, Tundra, or Loopus orbits the elevation angles at the mobile terminals are in excess of 50 degrees at all times. For such elevation angles shadowing is minimal and required fade depth is 4 dB and 10 dB for 95% and 99% link availability[5]. A dedicated fade margin of 5 dB and moderate adaptive FEC (e.g. RCPC constraint length 7, concatenated coding etc) with interleaving can easily provide highly reliable link.

TDM/TDMA FRAME FORMAT

An example of TDM/TDMA frame format is shown in Fig. 3. The short frame length results in low frame efficiency because of its overhead of the guard time, synchronization and identification of the burst signals. On the other hand, the delay time of the demodulation is proportional to the frame length. Therefore, a frame length of 120 mS is selected to minimize the delay time.

In addition, the time division duplex mode is employed for the mobile terminal. Hence, the TDM and TDMA frame formats should be selected such that transmit and receive timing of each mobile earth stations do not overlap. Each mobile earth station receives the link control data and the voice data from satellite/HUB station in the TDM frame, and transmit the voice data to the satellite/HUB station using the TDMA frame format. The overlap between the transmit and receive voice signal is avoided by changing the order of the burst signal of the voice data in the TDMA frame. However, it is possible that one mobile station transmits the signal at the same time as the link control signal is being received. This collision will cause the loss of link control information. In order to avoid this problem, the link control signal is transmitted twice in the TDM frame.

There is a allocation of 0.5 mS of guard time between the burst signals and 5.5 mS for the initial acquisition. Each Mobile earth station transmit the initial acquisition signal with predicted timing calculated from the position data of the earth station.

SYSTEM CONFIGURATION

Fig. 4 shows the configuration of the mobile earth station. It is interesting to observe that the antennas, which consist of seven elements of the microstrip disk antenna are separated for the transmitter and the receiver. Therefore, there is no need to protect the low noise amplifier from the transmitting power damage.

Also, transmission frequency is fixed for the TDMA system, and channel synthesizer is not needed. To eliminate the IF section, direct L-band modulation can be considered.

At the receive side, the LNA outputs are down-converted using common local carrier,

and fed to the beam forming circuit. Received signals from each of seven antenna elements are amplitude/phase compensated before being summed. Therefore, the signal to noise power ratio of the received signal is improved at the output of the beam forming circuit. Finally, this signal is stored in the frame buffer memory, and demodulated.

CONCLUSIONS

The authors have proposed the multi-carrier TDMA system with active array antenna for the future mobile satellite communications. An example of TDM/TDMA system and link budget calculation is given. An analysis of the signal search algorithm in low C/N condition is left as a subject for further studies.

REFERENCES

1. **Norbury, J. R.** 1988. Land Mobile Satellite Service Provision from the Molniya Orbit - Channel Characterization. In *Proc. of the 4th Int. Conf. on Satellite System for Mobile Communication and Navigation,* pp.143-146.
2. **Dondl, P.** 1986. Digital Transmission in a Quasi-Geostationary Satellite System. In *Proc. of the 7th Int. Conf. on Digital Satellite Communication,* pp. 641-.
3. **Steyskal, H.** 1987. Digital beamforming antennas An introduction. *Microwave Journal.* Jan pp107-124.
4. **Vogel, W. J.** 1990. Mobile Satellite System Propagation Measurements at L-band Using MARECS-B2. *IEEE Trans. on Antennas and Propagation.* Vol 38, No 2, pp. 259-264.

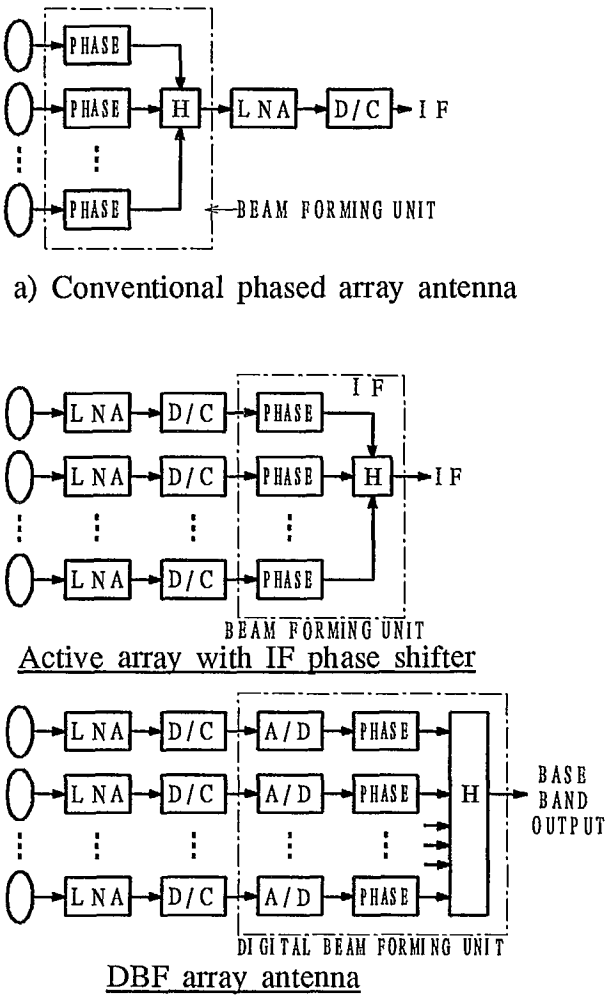


Fig. 2. Configuration of the phased array antenna

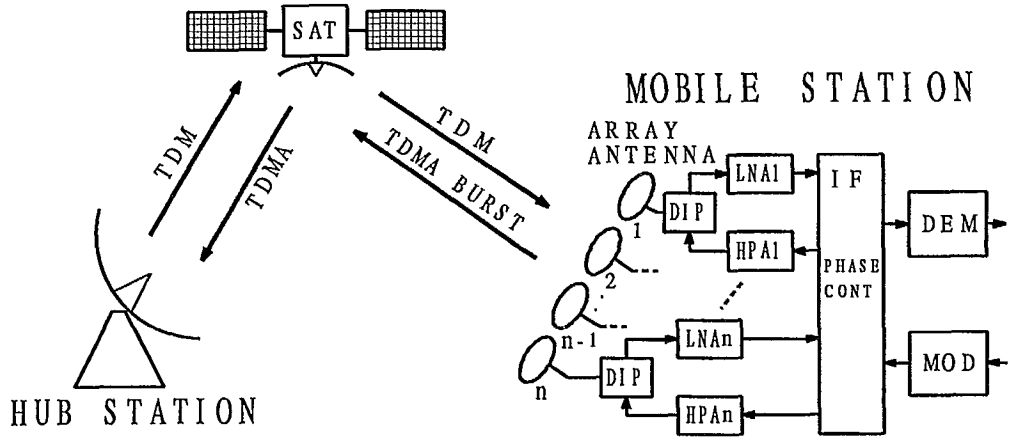


Fig. 1. Simplified configuration of TDM/TDMA satellite communication system

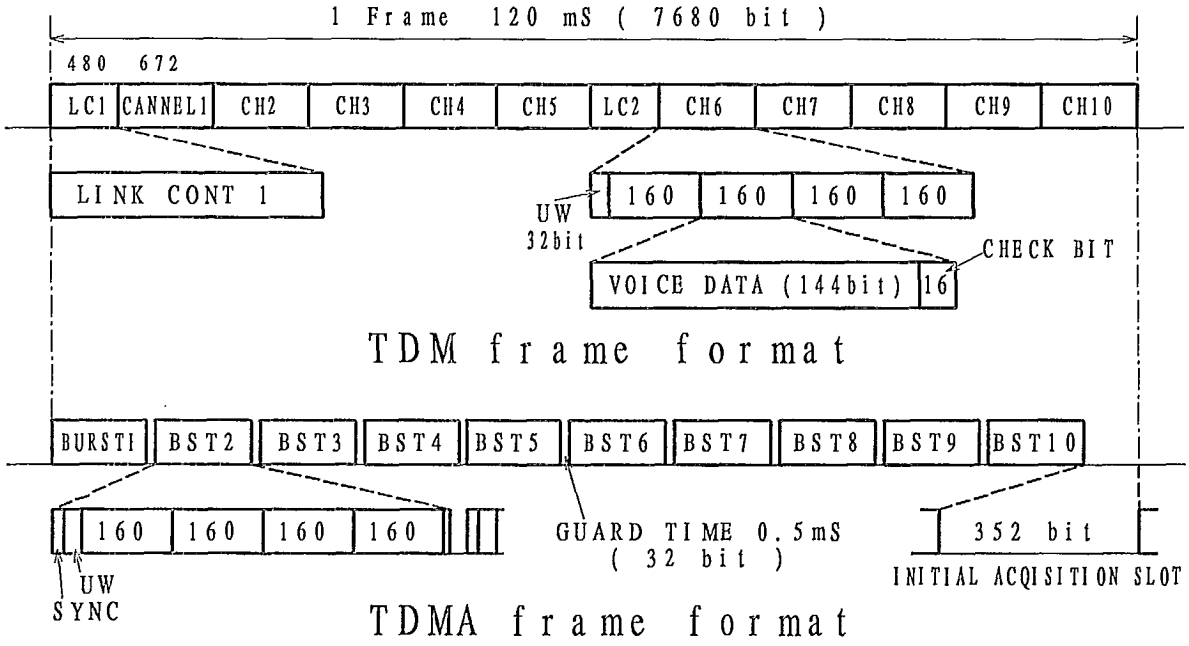


Fig. 3. Example of TDM/TDMA frame format (without FEC)

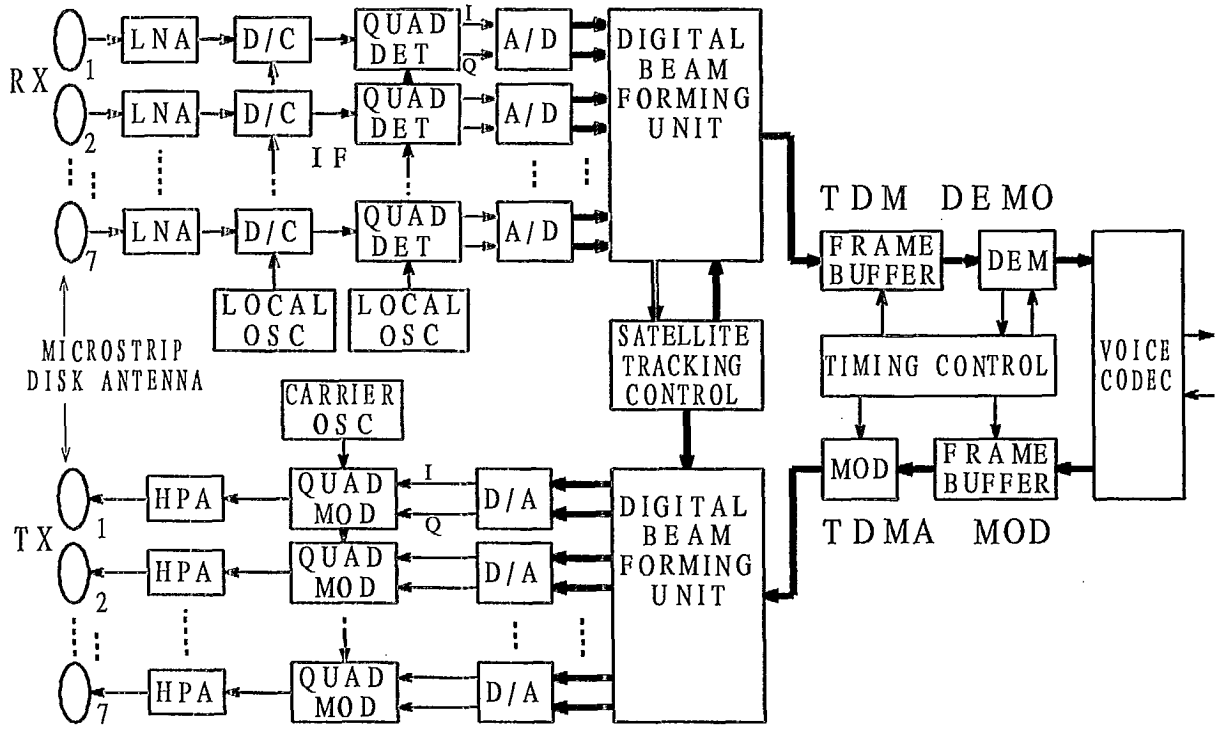


Fig. 4. Configuration of TDM/TDMA mobile earth station

Session 3
Specialized Systems

Session Chairman - *Tony Bundrock,* Telecom Research Labs, Australia
Session Organizer - *Arvydas Vaisnys,* JPL

A Land Mobile Satellite Data System

John D.B.Kent
Canadian Astronautics Limited
1050 Morrison Drive
Ottawa, Ontario K2H 8K7
Phone: 613-820-8280
FAX: 613-820-8314

ABSTRACT

The Telesat Mobile Incorporated (TMI) Mobile Data System (MDS) has been developed to apply satellite technology to the transportation industry's requirement for a fleet management system. It will provide two-way messaging and automatic position reporting capabilities between dispatch centres and customers' fleets of trucks. The design has been based on the Inmarsat L-Band space segment with system link parameters and margins adjusted to meet the land mobile satellite channel characteristics. The system interfaces with the Teleglobe Des Laurentides earth station at Weir, Quebec. The signalling protocols have been derived from the Inmarsat Standard C packet signalling system[1] with unique trucking requirements incorporated where necessary.

SYSTEM DESCRIPTION

The system topology is illustrated in Figure 1. and the HUB equipment connectivity in Figure 2. Fleets are connected to the central HUB via terrestrial (leased lines, dialup facilities or X.25 data accesses) or VSAT systems. The system connectivity maps each dispatch centre to a fleet or subfleet(s) of mobiles. Each mobile has an association with only one dispatch centre. The mobile earth terminal (MET) identifier then becomes the only routing indicator needed to address or identify the originator or addressee of a one-to-one message.

Four classes of messages are utilized in the MDS: general messages of up to 121 bytes for outbound messages and 129 bytes for inbound messages, coded messages which are a set of 32 pre-selected messages, 32 pre-formatted message with blank fields to be filled in by the operator and broadcast messages. Fleet or subfleet broadcast messages from the dispatch centre in the outbound direction are routed to groups of METS whose association is established when the MET logs-on the system.

SYSTEM ELEMENTS

The major hardware elements in the HUB system are the STARMASTER system switch (SS), redundant VAX 3300 Network Control Processors (NCP) and Single Board Computers (SBC) Satellite Protocol Processors (SPP), the PACX 200 Remote Switch and the Channel Units. Messages from dispatch centres are routed from the incoming terrestrial circuits to the on-line NCP via the SS. The NCP contains the dispatch centre records, fleet and terminal information and system configuration data. Historical records are also maintained in the

NCP for billing and network management purposes. Incoming messages are processed by the NCP and forwarded to two queues (Priority and Routine) to the SPP. The NCP also maintains the position reporting register and assigns channels, slots and slot intervals to the MET's at log-on for position reporting messages.

The two elements of the HUB are interconnected by four 9.6 kbps terrestrial routed lines, backed up by a single redundant line. The remote switch configuration is controlled by the NCP and lines and channel unit grouping are changed via a 2.4 kbps dialup orderwire.

PROPAGATION AND SYSTEM MARGIN

The land mobile satellite channel is characterized by a line-of-sight (LOS) component, a multipath-path component and shadowing or blockage of the LOS component. The LOS combined with the multi-path form the fast fading component which follows a Rician distribution and the shadowing results in a slow fading component with a log-normal distribution. Figure 2[2] shows the cumulative probability distribution function for typical land mobile satellite paths in the Ottawa, Ontario area.

The limited satellite EIRP available at L-Band (21.6 dBW) and the relatively wide bandwidth available per channel (5 KHz), in relation to the channel data rate, indicated that coding should be utilized to correct errors introduced by the rapidly changing channel characteristics. To improve the performance of the coding the outbound frame is interleaved in a 64x162 matrix thus spreading burst errors throughout the frame. The selected coding system was rate 1/2 convolution coding with a constraint length of 7 with 3 bit soft decision Viterbi decoding. The inbound frame is not interleaved due to its short duration (527 ms); however, it is coded as per the outbound channel. The margin and data rates on the inbound channels are adjusted for the lack of interleaving.

FRAME STRUCTURE AND MESSAGING

The outbound and inbound channel and packet structure is shown in Figure 4. A single outbound TDM channel is used on the forward link with a frame period of 8.64 seconds. A 128 bit unique word indicates the start of frame and is followed by a 112 bit bulletin board which includes the system identifier and network information. The inbound channel descriptors (ICD) provide the acknowledgement status of the previous frame (ACK/NACK) register and the reservation status of the 14/28 slots in the position reporting and messaging channels three frames ahead. This is necessary due to the depth of interleaving and the time required for decoding and descrambling. The message packet length is established by the length of the message. In the event of a message overlapping two frames, the message is carried over to the first information packet of the next frame. If the frame has excess capacity, the unused space is padded out with dummy zeros. The outbound message includes the reservation of an inbound channel and slot on which the MET sends the message acknowledgement.

Optimization of the system throughput is accomplished by utilization of a combination of slotted Aloha and reservation Aloha for inbound messages. The inbound channels, up to a total of 14 messaging and position reporting channels, are divided

into 14/28 slots depending on the channel data rate of 300 or 600 bps. The channels are divided into position reporting and general messaging channels which are subdivided into retry and priority message slots. Initial log-on to the system and the first packet of a single or multi-packet message utilize a slotted Aloha protocol with subsequent packets sent in reserved slots controlled by the satellite protocol processor (SPP) through the ICD. If a fade or other loss of the inbound packet occurs, indicated by a cyclic redundancy code (CRC) error, the ICD is set to NACK and the reservation is extended for another packet by the SPP. The MET acknowledges the NACK by repeating the lost packet and continuing the message on a three frame multi-slot basis adding a slot for the lost packet. The relatively long message delay for a multi-packet inbound message is indicated in Figure 5, which assumes no CRC errors for the inbound packets. An ARQ system is used in conjunction with the FEC with the last two bytes of packets reserved for the CRC.

Position report information is forwarded from the Loran-C board (latitude and longitude to .01 degree along with a two digit relative reliability indicator) to the MET processor board on the schedule established at log-on. This information is sent in the reserved slots and channel every 104, 208, 416 or 832 frames corresponding to position information every 15, 30, 60 or 120 minutes approximately. If the position report is not received by the HUB, indicated to the MET by the ICD status the MET takes no further action. If the user has subscribed to a retry service then the HUB interrogates the MET requesting a position report and assigning one of a number of reserved slots in a position reporting channel for the retry.

Broadcast messages to fleets and subfleets are not acknowledged by addressees to avoid congestion on the inbound channels. They are repeated at least three times in successive outbound frames to ensure reception by vehicles subject to nominal propagation conditions.

DEMODULATOR PERFORMANCE OBJECTIVES

The MET demodulator operates in a synchronous half duplex mode at 1200 sps. During transmission the MET must retain frequency and bit timing so that it can re-acquire the outbound TDM channel after the first unique word. With additive white Gaussian noise and a specified received phase noise the demodulator performance shall be:

C/No dB-Hz	Pe xE-05
33.0	7.8
33.5	2.0

Initial acquisition of the carrier with a frequency offset at the demodulator of +/-850 Hz shall be achieved over a range of C/No of 33.0 to 35.0 dB-Hz within 25 seconds with a probability of failure not greater than 0.01.

The inbound TDMA burst demodulators will operate at 600/1200 sps (HUB selectable) with 632/316 TDM symbol periods per burst. The burst to burst frequency uncertainty after frequency correction at the Des Laurentides earth station shall not be greater than +/- 650 Hz and the maximum rate of change of frequency during a burst shall not exceed +/- 65 Hz per second for six seconds. The packet error probability (PEP) for a 15 byte packet shall be less than the

following at 300 bps:

C/No dB-Hz	PEP
32.3	0.1
34.6	0.02

Digital signal processing is used in all modems and provides the flexibility to rapidly reconfigure the network and reduces the hardware required for redundancy.

SUMMARY

The MDS system is currently undergoing system and acceptance tests and is scheduled to be in operational service in the Spring of 1990. The satellite link performance has proven to be quite robust and to meet the messaging objectives even under relatively dense shadowing. The Loran-C position reporting accuracy has been excellent in rural areas and the quality indicator has alerted to possible errors in high noise environments.

Final test results will be presented at the conference and overall system performance will be discussed.

REFERENCES

1. **INMARSAT**, Standard-C, System Definition Manual, Release 1.3, July 1989.

2. **BUTTERWORTH, J. AND MATT, E.** 1983. The Characterization of Propagation Effects for Land Mobile Satellite Services. IEE Conference Publication No. 222.

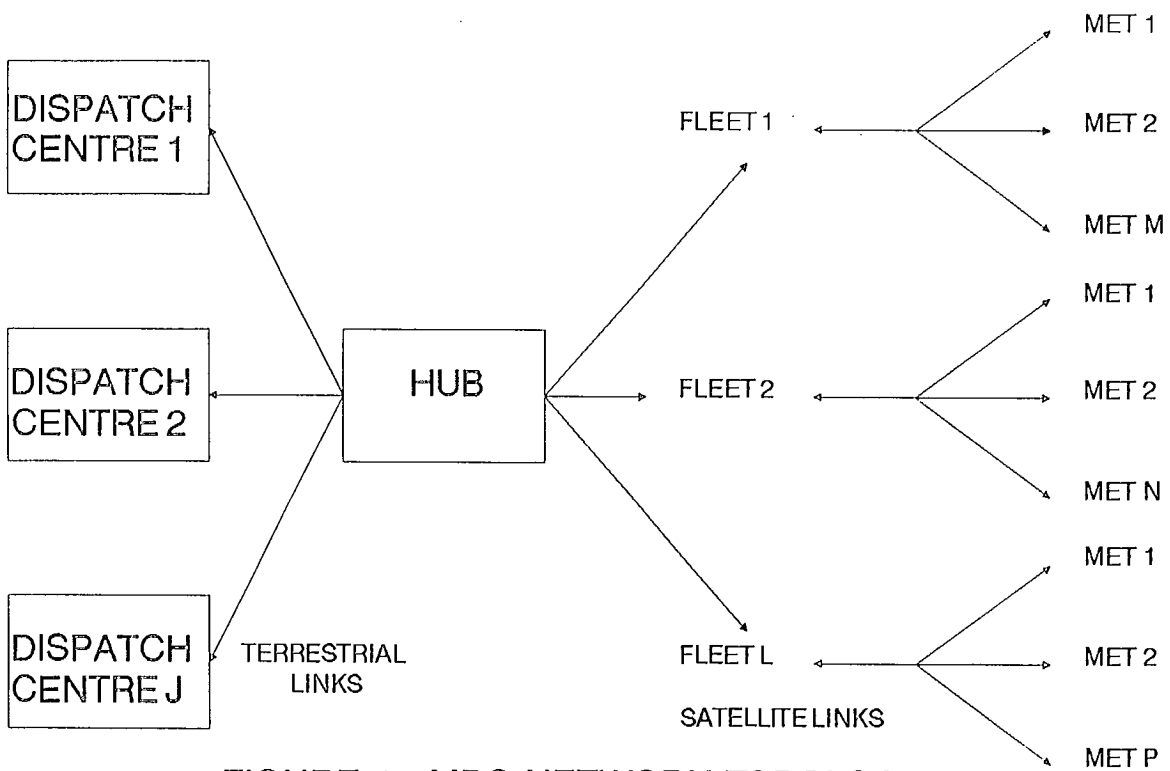


FIGURE 1 MDS NETWORK TOPOLOGY

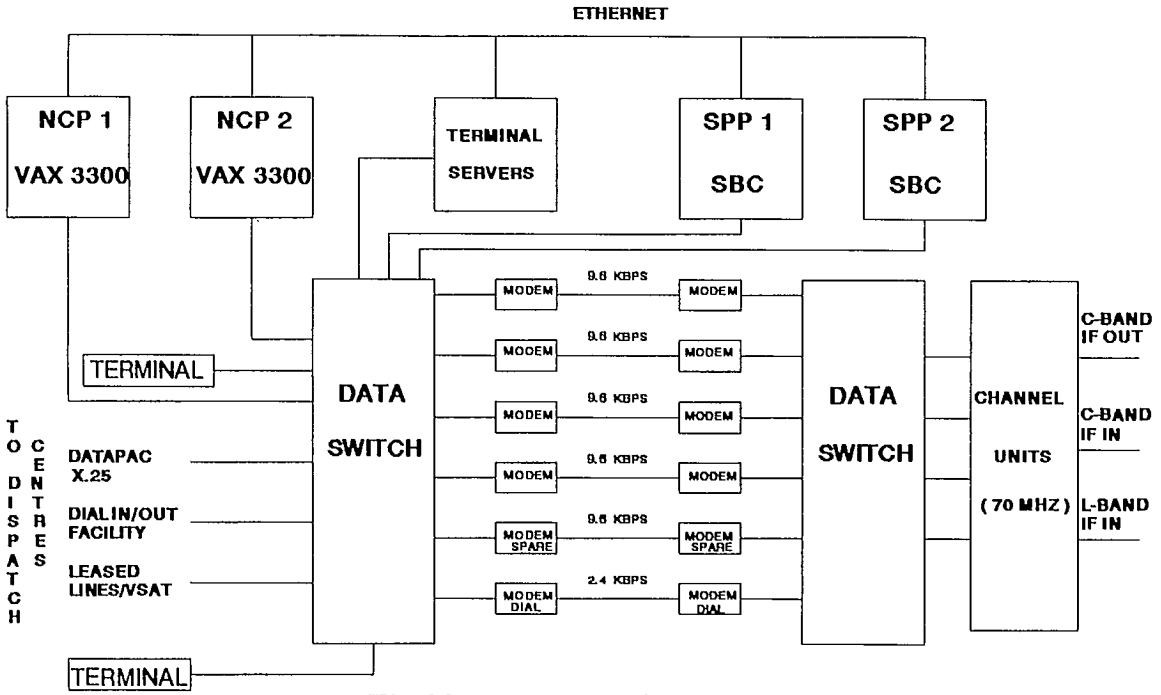


FIGURE 2 MDS HUB

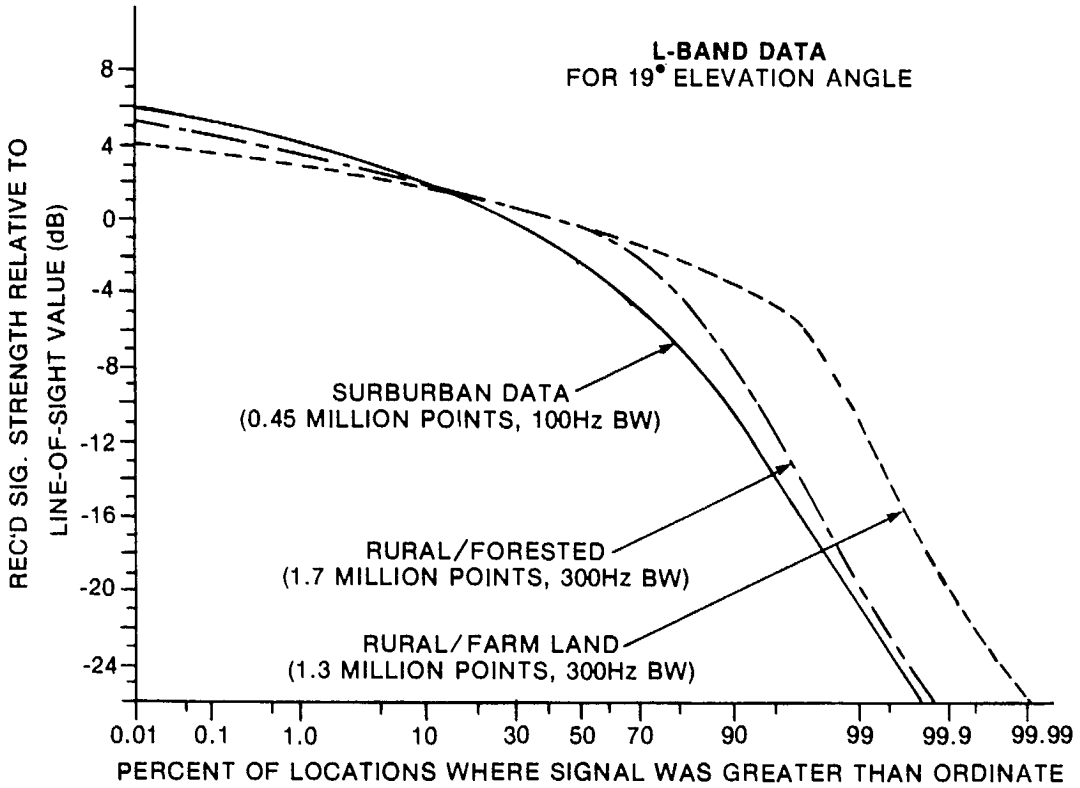


FIGURE 3 L-BAND PROPAGATION

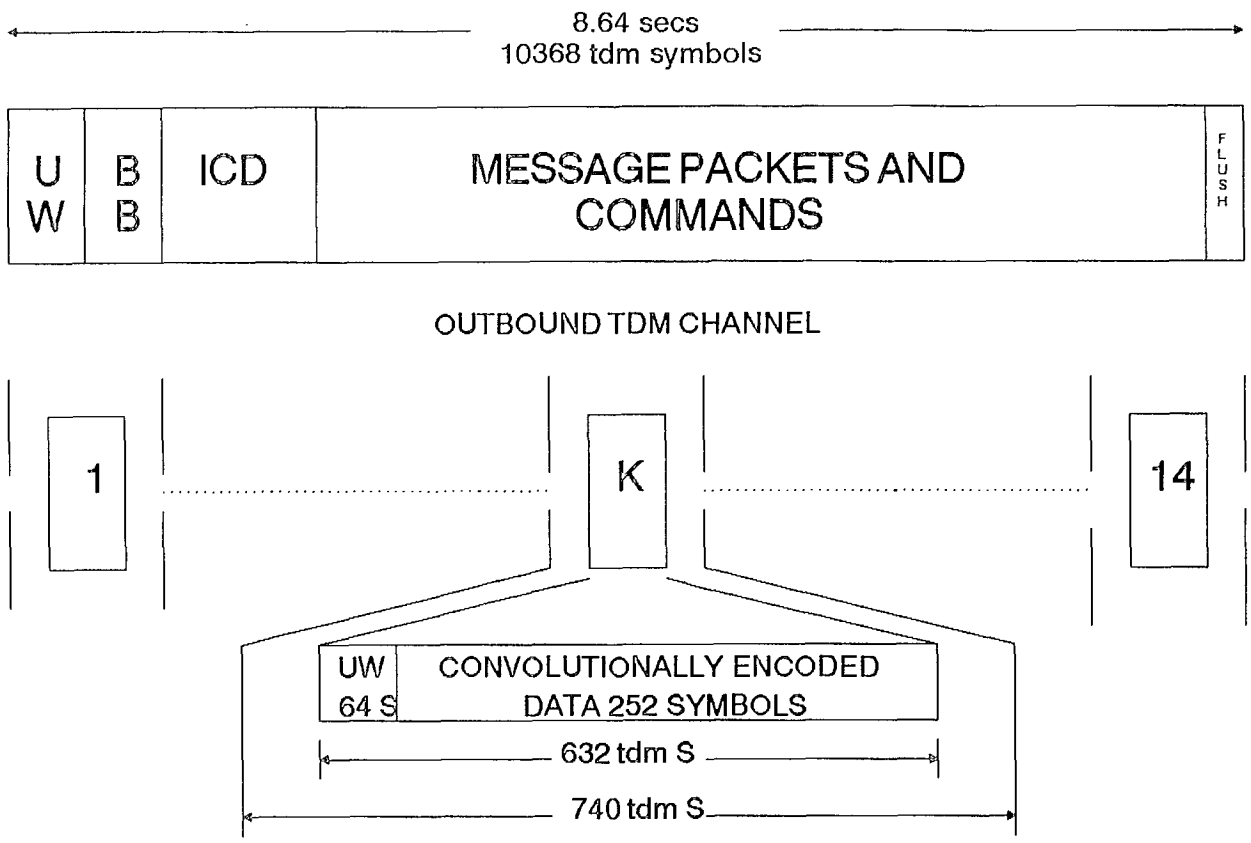


FIGURE 4 CHANNEL AND PACKET STRUCTURE

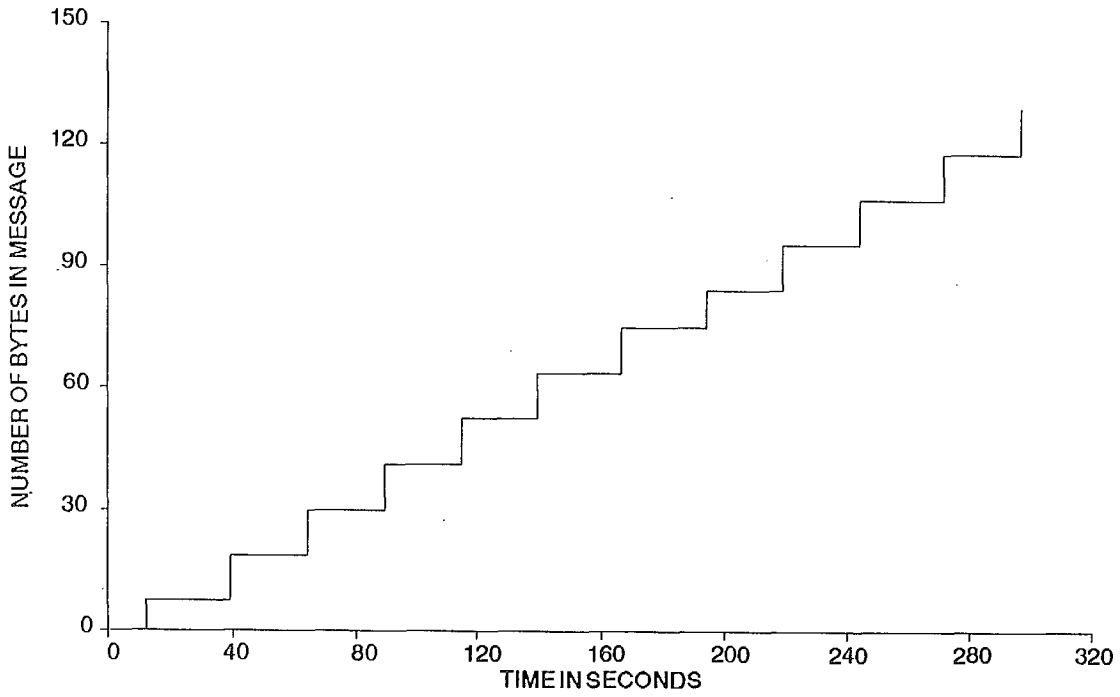


FIGURE 5 INBOUND PACKET TRANSFER TIME

Hybrid System of Communication and Radio Determination Using Two Geostationary Satellites

Shingo Ohmori, Yasushi Matsumoto and Eihisa Morikawa
Kashima Space Research Center
Communications Research Laboratory
Ministry of Posts and Telecommunications
Kashima, Ibaraki 314, Japan

Masayoshi Wakao
R & D Center for Radio Systems
Hirakawa, Chiyoda, Tokyo, Japan

Abstract

Communications Research Laboratory (CRL) of the Ministry of Posts and Telecommunications has developed a new hybrid satellite system which can provide both communications and positioning services in one system using two geostationary satellites. The distinctive feature is that location information can be provided by transmitting and receiving ranging signals over the same channel as communications through two geostationary satellites.

1. INTRODUCTION

Mobile satellite services are classified into two categories, one being "communications", providing voice, message, data transmission and so on, and the other being radio determination which gives information about location. In the past, these two satellite systems have been developed and implemented independently. However, in recent years, new satellite systems, which function both for communication and radio determination, have been studied by various organizations around the world.

Communications Research Laboratory (CRL) of the Ministry of Posts and Telecommunications has developed a new hybrid satellite system which can provide both communications and positioning services in one system using two geostationary satellites. The distinctive feature is that location information can be provided by transmitting and receiving ranging signals over the same channel as communications through two geostationary satellites.

The two types of terminals for the hybrid system developed by CRL are as follows;

Type A: in which one channel is used for both communication and positioning. A PN code (1023 chip rate) is used as a ranging signal and is transmitted over a single channel by OQPSK modulation.

Type B: in which the entire frequency bandwidth of SCPC channels is used. Communication and ranging information at 5 kbps, is spread by a PN code (255 chip rate) into a 1.2 MHz bandwidth and is transmitted by BPSK modulation.

This paper gives an the outline of system and the preliminary results.

2. EXPERIMENTAL SYSTEM

The experimental system consists of two geostationary satellites, ETS-V (150 E) and INMARSAT (180 E), a base earth station (Kashima Space Research Center) and mobile earth stations [3]. The frequencies between the satellites and mobiles are 1.6/1.5 GHz. In order to determine the position of a mobile precisely, ranging accuracy is an essential factor, which depends on the performance of the test equipment and the prediction accuracy of satellite orbits. In order to evaluate system performance, the following two ranging systems will be evaluated in the experiments.

(1) Two-way Ranging

As shown in Fig.1, a ranging signal is transmitted from the base earth station to a mobile station through the ETS-V satellite. The mobile station transmits the signal, which is repeated on IF or base band signal levels, to the base earth station through the ETS-V and INMARSAT satellites. The base station measures the time delay of reception between the respective signals from each satellite. In this system, the base station can determine the position of mobiles, so the mobiles are not required to have any complex functions for position determination. However, with this method satellite capacity is not efficiently used, because the mobile has to access to two satellites simultaneously.

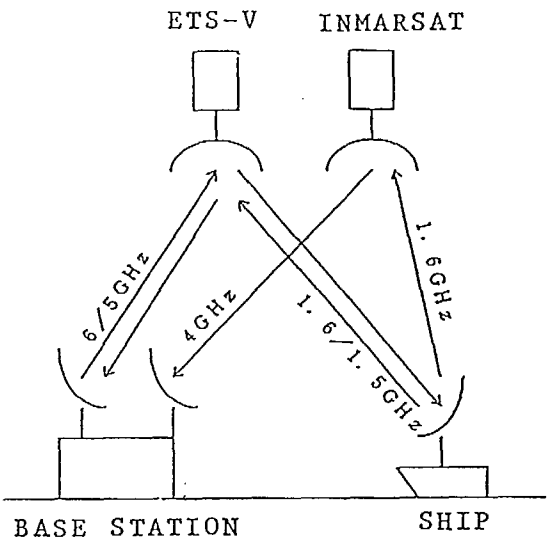


Fig.1 Two-way ranging.

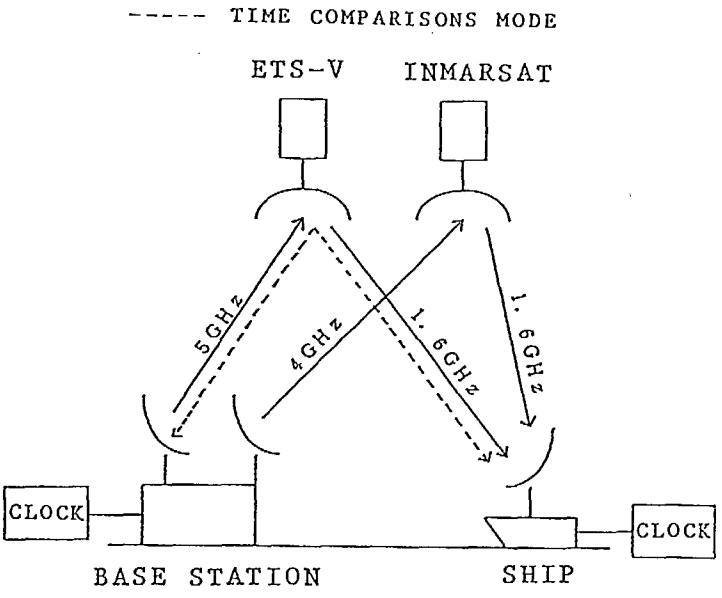


Fig.2 One-way ranging.

(2)One-way Ranging

As shown in Fig.2, the base station transmits the ranging signal, which is synchronized with a highly accurate clock, to a mobile station through the ETS-V and INMARSAT satellites. The mobile station measures the time delay with its clock, which is calibrated at appropriate intervals by the accurate clock through the ETS-V link. In this system, a mobile station can find its own position using the information of satellite orbits, which is provided by the base station through the satellite link.

Although, in this system, the mobile station has to have an additional function, whereby it can refer to the accurate time transmitted by the base station's master clock and adjust its own clock. Accordingly, this system can use a transponder more efficiently than the previous Two-way system.

3. EQUIPMENT CONFIGURATION

Two types of test equipment have been developed for the experiments with the hybrid satellite communication system. The first one is called the SCPC method, which uses one communication channel in a Single Channel Per Carrier (SCPC) access system for both communication and radio determination. The second one is called the SS method, which uses a Spread Spectrum technique to transmit and receive both the communication and ranging signals .

Both methods use digital modulation and demodulation techniques, and have ranging capabilities using a PN (pseudo-random noise) code. The Main characteristics of both methods are shown in Table1.

Table 1. Method of modulation and Parameters of ranging signals

	SCPC	SS
METHOD OF MODULATION	24kbps OQPSK	5kbps BPSK
SYMBOL RATE OF PN SIGNALS	12kbps	5kbps * 1.275Mbps
PN SIGNAL LENGTH	1,000 chips, 88.33msec	500 chips,100msec * 255 chips,0.2msec **

* PN SIGNALS FOR ELMINATING AMBIGUITIES
** PN SIGNALS FOR SPREADING SIGNALS

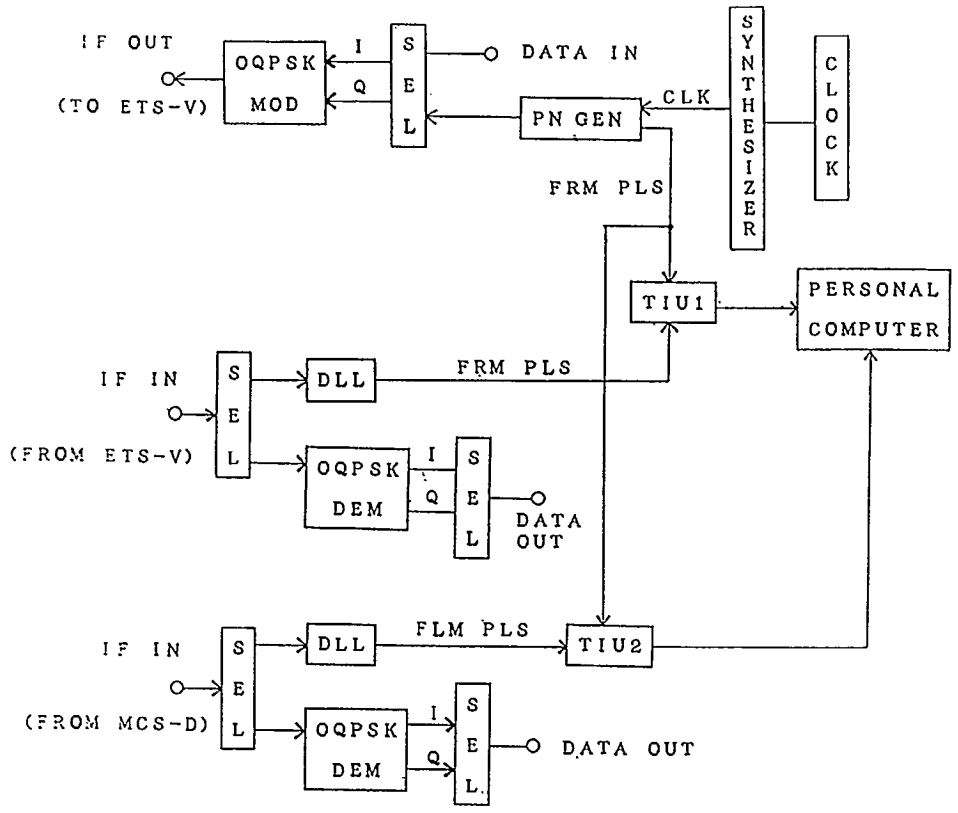


Fig.3 Base station block diagram.
(SCPC ranging system)

(1)SCPC method

Fig.3 is a block diagram of an SCPC terminal, which is connected to the RF section of the base station on IF signal levels. The PN code with 1023 chips is used for the ranging signal, which modulates and demodulates the carrier through a 24kbps-offset QPSK modulator (MOD) and demodulator(DEM). The I and Q channels of MOD/DEM are used for the communication signal and the ranging signal respectively. Synchronization of the received signal is locked by a Delay Lock Loop (DLL) with a Matched Pulse Detector (MPD), which consists of a digital correlation detector. The received PN signal is correlated with only 31-chip length of MPD, which allowing very fast acquisition of the initial carrier. The theoretical probabilities of mis-detection and false-detection of received PN signals are shown in Fig.4. If an appropriate threshold level of detection is chosen, for example 25-bit, the probabilities of mis- and false-detection are expected to be as negligible as 10-7 even in the case where Bit Error Rate (BER) is 10-2. A Time Interval Unit (TIU) measures the time delay of the received ranging signal, which is transmitted by the PN generator unit of the base station. The data are processed by computer to decide the position of a mobile station, taking account of the data of satellite orbits.

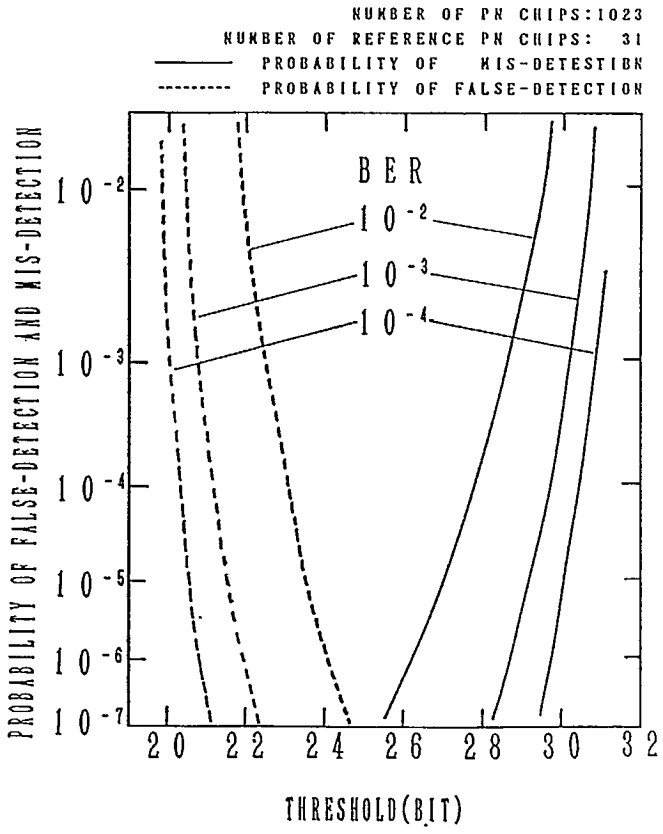


Fig.4 Probability of false- and mis-detection of matched pulse.

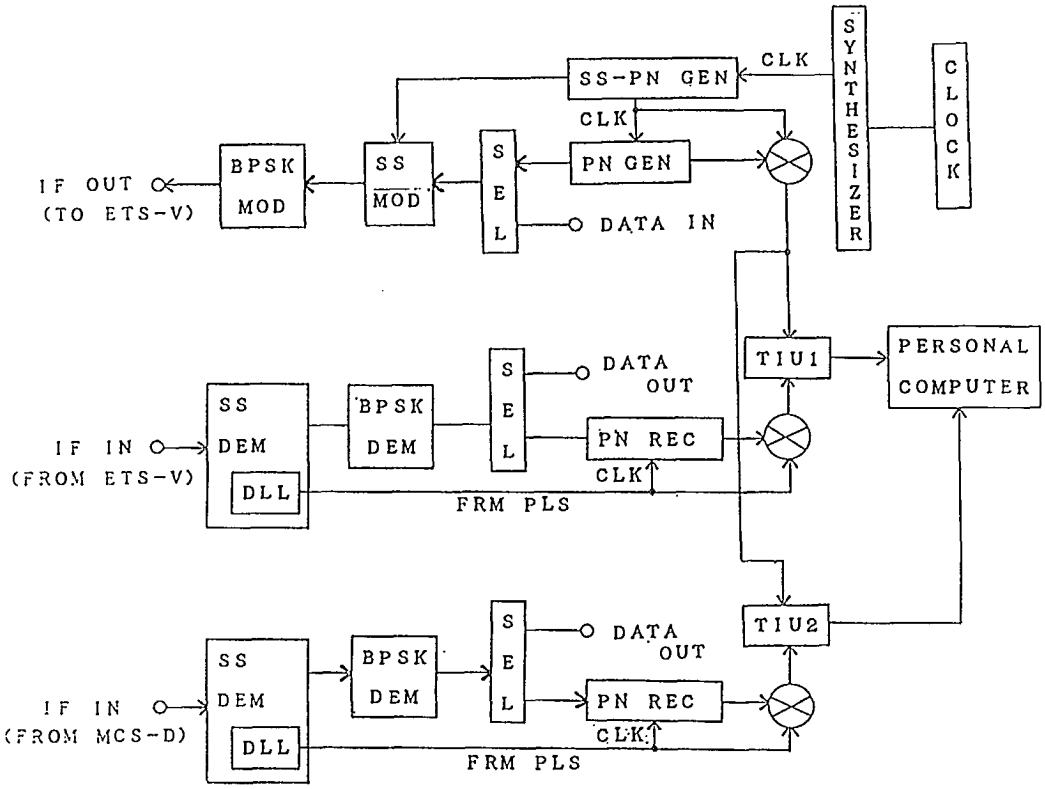


Fig.5 Base station block diagram.
(SS ranging system)

(2) SS method

Fig.5 is a block diagram of the SS test terminal, which is connected to the base station on IF signal levels. A base band signal of 5kbps is spread over a 1.2MHz frequency bandwidth by an SS modulator, and is transmitted through a BPSK modulator. The number of chips and length of frame of the PN code for the SS are 255 and 0.2msec, respectively. This method has some ambiguity in deciding the time delay, because the time delay through a satellite link is greater than that of a frame length of the PN code. In order to eliminate the ambiguity, a 5kbps PN code is transmitted first, and when the initial acquisition is completed, voice and/or data can be transmitted instead of the 5kbps PN code through the same channel. At a receiving section, the DLL is used to synchronize with the received PN code of 1.275Mbps, which is used for spreading signals. The Time Interval Unit (TIU) measures the time delay of both PN signals, which are used for signal spreading and ambiguity elimination respectively, and a computer calculates the position of a mobile station with taking account of the information of satellites.

4. Positioning Accuracy

(a) Ranging Accuracy

The major ranging error is caused by the tracking error of the DLL in a receiving unit. The standard deviation of phase detection error is given by the following equation [5].

$$\frac{\delta}{f_0}=\sqrt{\frac{B_L N_0}{2P}}$$

δ : standard deviation of phase detection error

f_0 : clock frequency of ranging signal

B_L : equivalent noise bandwidth of DLL

Fig.6 shows the theoretical results of phase detection error of the DLL in the SCPC and SS systems, calculated by the above equation.

The SCPC system has the advantage of a simple configuration, but the ranging accuracy is worse than that of the SS system because of its ranging signal's lower clock frequency. On the other hand, the SS system is affected by other communication carrier signals in the SS band which reduce its channel quality C/No, the ratio of carrier to noise density. Fig.7

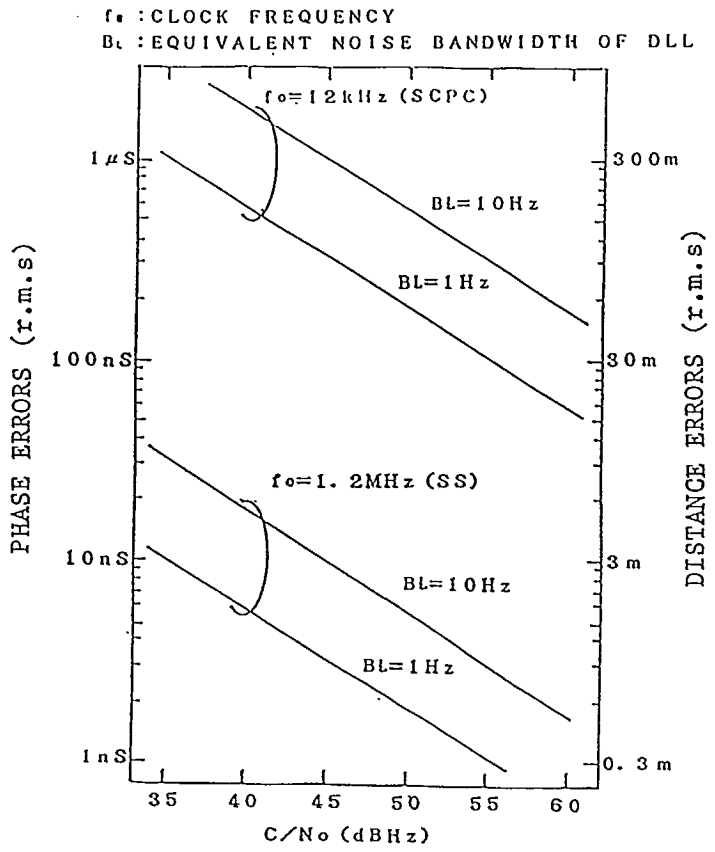


Fig.6 Received C/No versus phase errors of DLL.

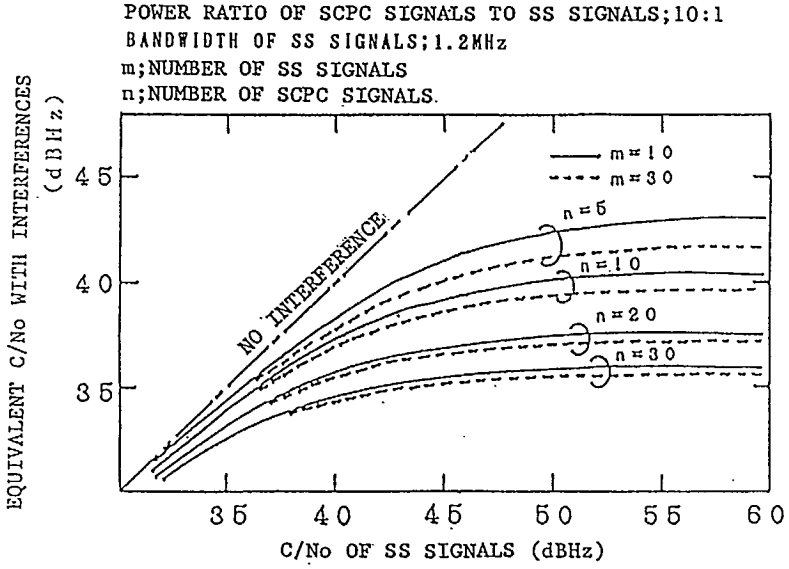


Fig.7 C/No for SS ranging system with interferences.
(Levels of SS and SCPC signals are equivalent.)

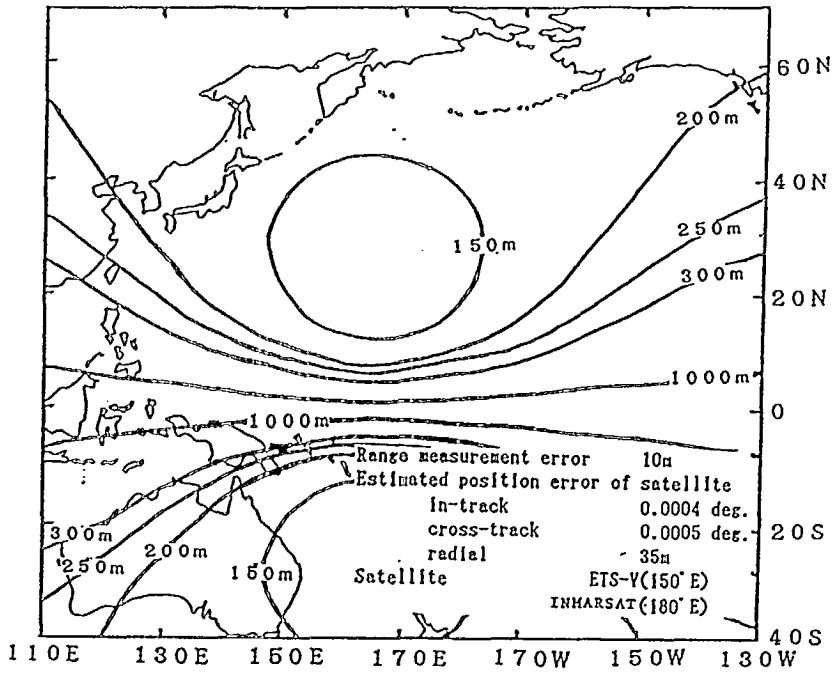


Fig.8 Accuracy of radio determination of this system.

shows the relation between the C/No and an equivalent C/No with interference signals affecting the SS channel.

The link budget of the SCPC system predicts the channel quality of C/No to be about 47-58 dB, which will give a ranging accuracy of several tens of meters as shown in Fig.6. In the SS system, the expected C/No with interference signals is about 35-45 dB, which gives a ranging accuracy of about several meters.

(b) Positioning Accuracy

Positioning accuracy depends on both accuracy of ranging and satellite location information. Fig.8 shows an example of the positioning accuracy using ETS-V and INMARSAT satellites. The theoretical prediction is calculated under the assumption that ranging accuracy is 10m, the ambiguity of satellite location is 0.004 degrees in latitude, 0.0005 degrees in longitude and 35m in radius. In low and middle latitude areas, a positioning accuracy of about several hundred meters is expected as shown in Fig.8. These figures are obtained without using any reference earth stations, but it is expected that greater accuracy can be obtained using reference stations which can eliminate many ranging bias errors.

5. CONCLUSION

A hybrid satellite system, which functions both for communication and radio determination, is proposed and the system configurations of two different methods are described. The proposed system can provide not only ordinary communications such as voice and data, but a radio determination service using a single communication channel. Radio determination and communication experiments using ETS-V and INMARSAT satellites are planned to start from the end of 1989. The experimental results are expected to contribute to establishing new satellite services, which have dual missions of communications and radio determination.

REFERENCES

[1] G.K. O'Naill;"The Geoster Radio Determination Satellite System", Telecommunication Journal, Vol.53, 1987.
[2] Y.Hase et al.;"Experimental Mobile satellite System using ETS-V", IEEE Denshi Tokyo, No25, 1986.
[3] Miura et al.;"Code ranging system using ETS-V and INMARSAT system", Technical Report of IEICE of Japan, SAT88-10,1988.
[4] W.Gill;"A Comparison of Binary Delay-Lock Tracking Implementations", IEEE Trans., Vol.AES-2, No.4, 1986.

Position Reporting System Using Small Satellites

B. Pavesi (1), G. Rondinelli (1),
F. Graziani (2)

(1) ITALSPAZIO, Roma, Italy
(2) Scuola di Ingegneria Aerospaziale -
Università "La Sapienza", Roma, Italy

ABSTRACT

The demand for mobile communications is growing in the world. Satellites systems can be a candidate solution to supply some of these services. In this context, a system able to provide position reporting and monitoring services for mobile applications represents a natural complement to GPS navigation system.
This service could be managed by public organization or by private groups. Private organizations can operate through Closed User Groups (CUG) and optimize the use of the own assigned satellite capacity.
In order to reduce the costs and the time necessary to implement the system, small satellites can be used, launched in piggy-back mode or by small launcher.
The system architecture is defined on the basis of the communications requirements derived by user needs, allowing maximum flexibility in the use of channel capacity and very simple and low cost mobile terminal.
The payload is sketched outlining the blocks modularity and the use of qualified hardware. The global system capacity is also derived.
The spacecraft characteristics are defined on the basis of the payload requirements.
A small bus optimized for Ariane IV, Delta II vehicles based upon the modularity concept is presented. The design takes full advantage of each launcher with a common basic bus or bus elements for a mass production.

1. INTRODUCTION

Satellite mobile communications together with position reporting are requested for several reasons and constitute a valuable market.
The aim of this paper is to propose a low cost system architecture which can be convenient for private networks, for land mobile and aircraft applications.

Land mobile users are requesting:

. two-way data messages,
. localization and position reporting,
. alarm services for security and safety of transported goods,
. emergency,
. interactive services.

Target customers are trunking organizations, tour-operators, governmental agencies, surveillance and rescue service providers.

The system architecture adopts message transceivers suitably integrated with localization equipment. The transceivers communicate with the control gateway stations, sending and receiving messages to/from the satellite.

The key element for a successful application of this scheme is the high-value service, when provided in a cost-effective manner. Low costs are associated to the concept of private use. Thus, data messages, position determination and position reporting can be a promising business for closed user groups operating through their own private networks.

Direct access to data bases, together with vehicle management and control are additional features offered by this system. Automated or on demand updating messages with vehicle geographical location can be beneficial in controlling ad-hoc dispatching, and in scheduling travel management.

The flexibility of the system and the possibility of enhancing the capacity according to the communication traffic or the number of mobiles are other significant advantages. This modularity will allow gradual investments for private networks, which can begin their operations dimensioned as needed.

For the aircraft the most demanding service is the support to ATC (Air Traffic Control) with ADS (Automatic Dependent Surveillance) services.

The adoption of satellite-based system for air traffic ADS is mainly driven by the need to improve the Communication, Navigation and Surveillance services. According to ICAO requirements, the ADS will support aircraft position monitoring with high frequency of updating. The requirements for ADS service are:

a) the ADS data acquisition has a minimum periodicity of 10 sec, with a message length of 144 bits;
b) typically on long route application an ADS message is requested every 5 minutes with a message length of 330 bits (including basic ADS report, extended ADS report and associated ADS report data).

For case a) 1 message of 144 bits each 10 sec is equivalent to an information of 864/minute. Up to 40 users can be served by one channel rated at 600 bps.

For case b) 1 message of 330 bit each 5 min. allows to serve more than 500 users through one channel, rated at 600 bps.

The traffic we assumed considers 85% of the traffic due to the high frequency updating and the remaining for the complete message updating each 5 min. In this way, we can consider that each channel can serve 36 users with updating message every 10 sec, and 85 users with updating messages every 5 min that is equivalent for each channel to a 1 min. of updating for a message of 330 bits for 100 users each hour.

2. SYSTEM CHARACTERISTICS

A system capable to provide position determination service, position reporting and data message services can be implemented in a selected zone using a reduced constellation of satellites. The proposed configuration is a constellation of two small satellites positioned in Tundra orbits plus 3 small geostationary satellites.

All satellites are provided with navigation packages, GPS compatible.
In addition, the GEO satellites are provided with communication payloads for position reporting and data message services. The revenues of the overall system are obtained through the communication services.

The major system features are:

- complementary to GPS-NAVSTAR;
- use of same GPS-NAVSTAR navigation message structure, but with different sets of PRN codes;
- unlimited number of users for position determination;
- gap filler in zones of poor GPS availability;
- 20 to 50 thousand users for position reporting at 600 bps;
- TDMA scheme for position reporting.

We envisage the following organization for position reporting:

- use of one GEO-satellite for aircraft services in connection with user's terminals of enhanced type,
- use of two GEO-satellites for land mobile services.

The HEO satellites will carry the navigation payload only.

Satellite Payload Configuration Outline

The payload is composed of:

- one transparent transponder receiving at C or Ku-band (or another band to be specifically allocated) the navigation signal from a feeding station part of a fiducial system supporting the service and retransmitting at L-band, at the same frequencies of operation of the GPS-NAVSTAR;
- one transparent transponder receiving messages from mobile users at L-band (in the range allocated to mobiles) and retransmitting down to the earth stations at C or Ku or another band and vicecersa.

The major navigation payload parameters, for a margin of about 5 dB in GPS signals reception, are:

	GEO	TUNDRA
. Satellite EIRP	26.8 dBW	28.7 dBW
. Antenna gain (EOC)	16 dB	16.5 dB
. HPA RF power	12 W	16.6 W

The navigation payload mass is about 35.9 Kg with a power consumption of 60 W for GEO payload and 74 W for HEO payload.
Only C/A code is envisaged for this type of civil service. Thus, the frequency used will be 1.57545 GHz

The major parameters of the data service are:

		AIRC/FWD	AIRC/RET	LMOB/FWD	LMOB/RET
. N. of carriers/ transponder		2	20	4	30
. Access/modul.		TDMA/ QPSK	TDMA/ QPSK	TDMA/ QPSK	TDMA/ QPSK
. BW per carrier		81 KHz	8 KHz	30 KHz	4 KHz
. Total bandwidth		162 KHz	160 KHz	120 KHz	120 KHz
. Power per carrier	L-band	40 W		15 W	
	C-band		0.5 mW		0.16 mW
. Total power	L-band	80 W		60 W	
	C-band		10 W		5 W
. Output back-off		1 dB	3 dB	2.2 dB	6 dB
. Satur. power	L-band	100 W		100 W	
	C-band		20 W		20 W
. Ch.s per carrier		40	4	15	2
. Channels		80	80	60	60

A two way service is established with:

- L-band, for down links from the traffic control centers (Forward link),
- C-band, for down links from the users (Return link).

The position reporting payload mass is about 40 Kg, excluding the antenna system which is part of the navigation payload while the power consumption is about 285 W.
The frequencies of L-band are in the same range of aeronautical and land mobile communications. The frequencies of C-band (or Ku-band) are in the conventional band allocated to feeder links.
The payload budget for navigation and position reporting are as follows:

PAYLOAD BUDGET FOR NAVIGATION SERVICES

	MASS (Kg)	POWER (WDC)	POWER (WDC)
	GEO-HEO	GEO	HEO
Amplifiers	3.5	35	47
Receiver section	5	10	10
Up-down converter IF transp. section Filters and switches	10	10	10
Antenna system L-band	12		
C-band	2.5		
Margin	3	5	7
Total	36	60	74

PAYLOAD BUDGET FOR POSITION REPORTING SERVICES

		MASS (Kg)	POWER (WDC)
		GEO	GEO
Amplifiers	L-band	14	180 +
	C-band	7	60
Receiver section & dipl.		5	10
Up-down converter IF transparent section Filters and switches		10	10
Antenna system	L-band		
	C-band		
Margin		4	25
Total		40	285

The overall payload mass is estimated 76 Kg about, and the power consumption is about 345 W (DC).

The earth segment for navigation service will be composed by:

. Navigation fiducial network: four stations
. Master control station: one main center, at least
. Mobile terminals: GPS receivers, or similar

The typical features of the ground segment for the message communications are simplicity and low cost. A large scale production of the mobile terminals, will allow the low cost characteristics. The major characteristics of data message communication earth segment are:

GROUND STATIONS GATEWAY

. Frequency	C-band
. Antenna size	3 m diameter
. Antenna gain	39/42.6 dBi Rx/Tx
. EIRP (for 1 carrier)	40 dBW
. G/T	12.1 dB/K

MOBILE TERMINALS DATA MESSAGE

	AIRCRAFT	LAND MOBILE
. Frequency	1.5 GHz Rx, 1.6 GHz Tx	1.5 GHz Rx, 1.6 GHz Tx
. RF Tx power	15 W	10 W
. Antenna gain	3 dBi	3 dBi
. G/T	-21.5 dB/K	- 21.5 dB/K

The number of user's terminals will be several thousands.

3. THE USE OF LOW COST SMALL SATELLITE

In order to implement these systems it is necessary to reduce the high costs involved with satellite development and launch, a small bus capable to fly at marginal cost on a piggy-back mode can provide an alternative access with respect to other bigger satellite families.
A such kind of satellite shall be developed, tested and launched in a short period with a low launch cost, it shall be also lightweight, manufactured with mass production techniques, having some relaxation in failure rate and reliability requirement.

In the following table the principal cost reduction areas are presented for a small satellite with respect to a typical allocation cost for a communication satellite. As it can be noted the major areas candidate for cost reduction are administration, system engineering, product assurance, a launch.

	ALLOCATION
Program Administration	16.4%
System Engineering	8.2%
Product Assurance	10.3%
Assembly Integration test	10.5%
TT&C and DH	3.5%
AOCS	6.3%
Propulsion	6.3%
Solar Array	3.8%
Power Control	4.1%
Structure	5.2%
Thermal	1.3%
Communications Antenna	8.1%
Communications Transponder	16.0%
	100%

Launch cost with respect to satellite cost	100%
Launch Insurance with respect to Satellite cost	25%

Cost Reduction Areas

The administrative costs can be reduced by negotiating contract conditions directly with the equipment vendors. Co-responsibility and incentive can be used to reduce sub-system contract costs.
Further reductions are possible incorporating several functions in single units, with reduction of number of contracts to be managed.

The system engineering costs can be reduced involving the system team during all activities across the program, to cope also some specialistic efforts and minimizing the subcontracted items.

Another area for cost reduction concerns the design, manufacturing, integration, testing techniques and product assurance procedures.
The primary reasons for high cost of satellites are the small number of units produced and the high single-satellite reliability required.
It is evident that systems employing larger number of moderate capacity platforms can be more economical than high performance platforms, reducing a significant amount of production cost.
The current methods to enhance reliability are extensive testing and equipment multiple redundancy. In the case of a small satellite, the low level of interconnection and the units integration allow to significantly reduce the test procedures, taking also into account limited redundancies.

As regards the launch the primary payload usually does not exactly fit to the launcher capability, so additional mass is available for secondary payloads. Taking also into account that the Ariane and Delta II launchers can increase the length of fairing and make available small additional volumes for a piggy-back passenger sandwiched between the launcher interface ring and spacecraft adapter (Fig. 1). It is clear the possibility of launching at a marginal cost the additional passenger.

4. SMALL BUS FOR GEO AND HEO MISSION

For the basic configuration of "MINISTAR" satellite, a sun pointing attitude concept has been selected, that with some modifications permits to support both HEO and GEO missions (Fig. 2).

The satellite design adopts an inter-passenger approach.
The external structure is dimensioned to support the upper passenger during the launch and have in the upper part a ring for the attachment of primary passenger adapter and in the lower part a circular junction with the launcher.
In the lower side of spacecraft there is a racks where the tanks for pressurization are placed. At the center of the structure is positioned a central tube that support the bapta for the payload platform and the bipropellant motor. Four vertical plates attached at the tube support the propulsion tanks.

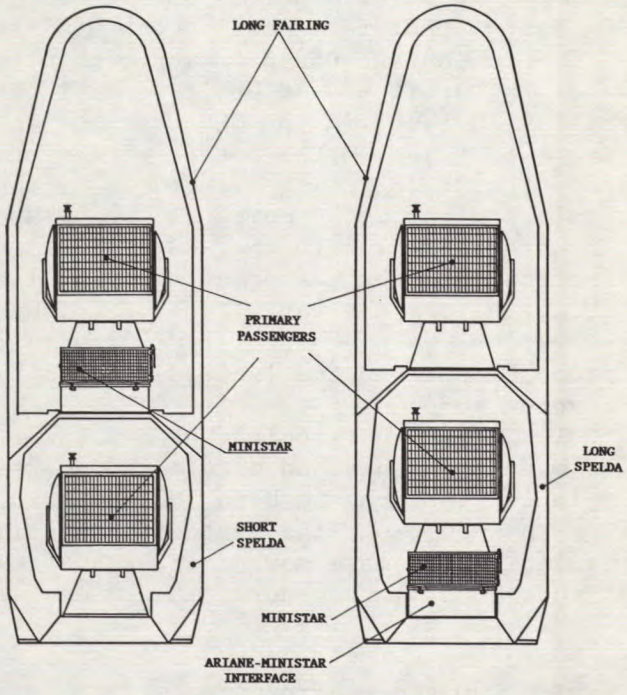


FIGURE 1 - "MINISTAR" IN ARIANE IV LAUNCH CONFIGURATION

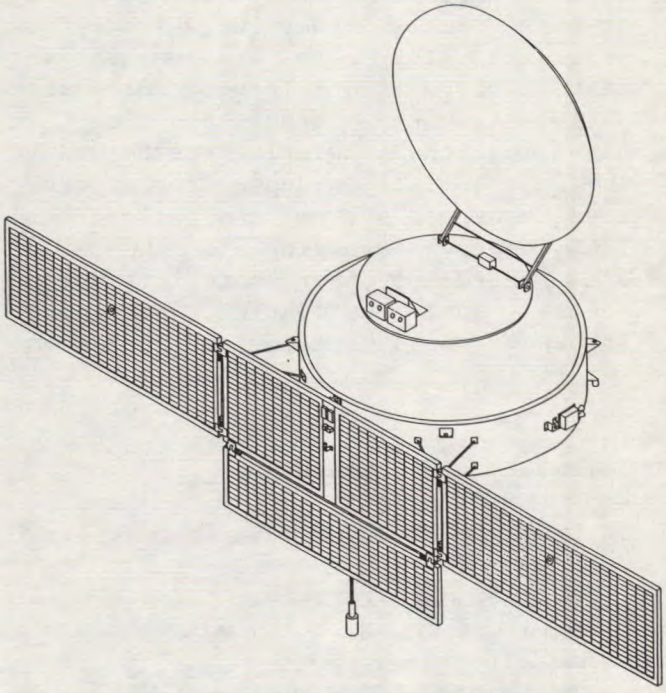


FIGURE 2 - MINISTAR CONFIGURATION

The solar panels, once deployed, are rigidly fixed to the spacecraft and during the launch are wrapped around the spacecraft structure.
The payload is positioned on a platform rotating at one revolution per day and allows to accommodate a 1.8 m reflector communication antenna.

The main differences between HEO and GEO are the antenna pointing mechanism and the attitude and orbit control system. In fact, in the GEO case the antenna is fixed on the rotating platform and the yaw pointing is assured by the control loop using momentum wheel and thrusters, while in the HEO case it is necessary a further pointing mechanism which allows the variation of antenna elevation with an angular displacement proportional to the latitude of earth covered zone.

The relative velocity between service module and payload module is one revolution per day in both GEO and HEO. In the HEO case, the velocity rate is variable and make more complex the attitude control procedure.

System Management Processor

The satellite functions and operation coordination are performed by an on-board central computer which can be programmed for each mission and for each configuration. This microcomputer derives from the current technology, it has a multitask operating system and it is substantially a supervisor of attitude control system and perform the TT&C functions managing also the periodic control of other satellite subsystems and activating particular sequences as separation and apogee motor firing.

Telemetry and Command

It performs three main functions:

- the telemetry function (a reduced rate in the order of 200 bps could be used in this case),
- the command function,
- the tracking function.

In stationary conditions, the payload antenna system, is used. In the emergency and launch conditions, when the attitude of the satellite is not fixed, the TT&C will operate through the omnidirectional antenna.

Propulsion

The motor used for the apogee maneuvers will be a liquid propellant.
This seems a suitable solution in the case of "MINISTAR", even if the use of liquid bipropellant technology will lightly increase the cost.

The first advantage is that multiple burns are possible, permitting higher injection accuracy. Besides, it is not necessary to spin-up the satellite at high spin rate during the maneuver.

The second advantage is that bipropellant facilitates the accommodation of the tankage structure into a small structure.

A third advantage resides on the motor control electronic unit, which remains the same while the tanks can be tailored for each mission and for each spacecraft weight.

Attitude Control

The main element for attitude reconsititution are a two axis infrared earth sensor mounted on the payload platform and a sun sensor positioned on the solar panel. This sensing principle, together with a significant momentum bias provided by a momentum wheel positioned along S/N direction, provides the nucleus of the attitude control system.
The momentum bias needs to be adjusted in length and direction in regular intervals of days. To execute the correction maneuvers 6 thrusters (12 for redundancy) are adopted also for orbit control.

Thermal

The thermal control of "MINISTAR" uses only passive techniques. Electric heaters with automatic control capability could also be included. This means that the temperatures of the components have to be maintained within their limits by:

. controlling conductive and radiative heat paths,
. selecting a suitable geometrical configuration of the equipment,
. selecting thermal coating (i.e. thermo-optical properties of the surfaces),
. using thermal insulations for selected zones of the satellite in order to reduce both heat flow to or from component and temperature fluctuations of component due to time varying external heat flux.

In addition, electric heaters could be required for some particular component (e.g. battery).
A design margin of 10 degrees centigrade has to be maintained between operating in orbit temperature limits of the equipments and the limits over which they are to be qualified.

Power and Mass Budgets

In the following tables are presented some preliminary estimation of mass and power for the selected satellite options.

SATELLITES POWER BUDGETS

	GEO	HEO
	COMMUNICATION & NAVIGATION	NAVIGATION
Payload	345 W	74 W
Propulsion	12 W	12 W
System manag. processor	25 W	15 W
Telemetry and command	20 W	20 W
Attitude control	22 W	12 W
Thermal	7 W	3 W
Harness loss	5 W	3 W
Battery charge	60 W	-
Margin	20 W	10 W
TOTAL	516 W	149 W

SATELLITES MASS BUDGETS

	GEO	HEO
	COMMUNICATION & NAVIGATION	NAVIGATION
Payload	76 Kg	36 Kg
Electric power	42 Kg	10 Kg
Structure	50 Kg	25 Kg
Propulsion	16 Kg	16 Kg
System manag. processor	18 Kg	14 Kg
Telemetry and command	10 Kg	10 Kg
Attitude control	20 Kg	18 Kg
Thermal	5 Kg	3 Kg
Mass margin	30 Kg	25 Kg
Dry spacecraft mass	267 Kg	157 Kg
Propellent	26 Kg	10 Kg
Apogee motor expendable	118 Kg	50 Kg
TOTAL	411 Kg	217 Kg

These mass and power budgets are derived considering:

. for the GEO satellites: Ariane IV launch, 50 m/s per year for station keeping, 5 years of life plus 1 year of margin;

. for HEO satellites: Delta II launch, 30 m/s per year for station keeping (Ref. 1, 2, 3), 5 years of life plus 1 year of margin, no eclipse during operational phase.

Economic analysis

A cost/revenue model has been applied to this configuration. The case of small satellite allows the following economical considerations:

. low cost per Kg of the payload,
. quick saturation of the available capacity,
. low cost of the launch,
. two operational Tundra satellites plus one spare satellite, launched all together with a single launch with Ariane,
. three GEO satellites launched in Ariane piggy-back,
. reduction of the system organization cost due to the low complexity of the management of the overall mission.

The assumptions for cost of the organizations are:

Inflation rate	3.5%
Discount rate	
. Capital investor	9%
. Service provider	9%
. Ground segment prov.	9%
Profit	
. Space Corporation	10%
. Service Provider	14%
. Ground Segment prov.	12%
or	
. Unified corp. earnings (Space Corp.+Serv. Prov.)	24%
. Banking system	9%
Cost	
. Administration (MAU/yr)	0.05
. Investment cost (MAU)	2.5
. Operation cost (MAU/yr)	0.15

The revenues are derived as output from:

. number of channels used or number of users, according to the service penetration profile during the mission,
. cost per unit, which is the cost per number of bits transmitted each second.

The navigation service is offered free of charge while the channel occupancy is charged with a cost per message.

The navigation and position reporting system recovers the cost from the message transmissions. The messages mainly concern with the position of the user, but other information can be transported by the system. The system provides a two-way message capability.

The references costs are:

. 0.7 USD per 1 Kbit message,
. 0.1 USD per messages of blocks of 256 bits, (Geostar).

In making the comparison, we consider the channel cost per year as results from the model runs which is 0.315 MAU/ch/year and consider a channel utilization of 6 hours per day during the working days, which corresponds to 93960 min/year.

For the 1 Kbit messages we obtain: 0.093 AU and 0.081 AU, respectively for the case of full insurances of the system elements and the case of absence of insurances.
The Space company and the Service-providers are considered as different organizations.

In the case of only one Actor, the cost per 1 Kbit message is 0.084 AU.

ACKNOWLEDGMENT

The authors wish to acknowledge the contribution and the support of Dr. A. Teofilatto President of Italspazio and the advice and suggestion of Dr. G. Barresi Director of Italspazio.

REFERENCES

[1] "Low Cost Station Keeping Maneuvers for a Small-Satellites Constellation in Tundra Orbits". A. Cramarossa, G. Barresi, G. Rondinelli, F. Graziani Proceedings of the 2nd Annual AIAA/USU Conferences on Small Satellites, Logan, UTAH, U.S.A. September 1988.

[2] "Orbit Control for a Regional Navigation System Based on Tundra Orbit". G.Rondinelli, A. Cramarossa, L. Caporicci, F. Graziani AIAA-89-3619, AIAA Guidance, Navigation and Control Conference. Boston, Massachusetts, U.S.A. August 1989.

[3] "Orbit Acquisition and Control Strategy for Small Satellites in Inclined Eccentric Orbits". G. Rondinelli and F. Graziani. CNES International Symposium on Space Dynamics, Toulouse, France, 6-10 November 1989.

[4] "The Land Mobile Satellite Service: Road and Railway Communications". K.P. Galligan, C. Lluch, B. Pavesi, A. Tuozzi. IEE: 4th International Conference on Satellite System for Mobile Communications and Navigations, London 1988.

[5] "Archimedes Mission Aspect". P. Palmucci, B. Pavesi. IEE, Colloquium "Highly Elliptical Orbit Satellite Systems, London 1989.

Satellite Sound Broadcasting System Study, Mobile Considerations

Nasser Golshan
Jet Propulsion Laboratory,
California Institute of Technology,
4800 Oak Grove Drive
Pasadena, Ca 91109-8099
Phone: 818-354-0459
FAX: 818-393-4643

ABSTRACT

There is an ongoing study at JPL to investigate a Satellite Sound Broadcast System in the UHF or L bands. The study considers program reception both by portable and mobile receivers. This paper reports on the mobile reception part of the study. Existing propagation and reception measurements for the Land Mobile Satellite channel are used with proper interpretation to evaluate the signalling, coding, and diversity alternatives suitable for the system. Signal attenuation in streets shadowed by buildings appear to be around 20 db, considerably higher that the 10 db adopted by CCIR. With the marriage of proper technologies, an LMSS class satellite can provide substantial direct satellite audio broadcast capability in UHF or L Bands for high quality mobile and portable indoor reception by low cost radio receivers. This scheme requires terrestrial repeaters for satisfactory mobile reception in urban areas. A specialized bandwidth efficient spread spectrum signalling technique is particularly suitable for the terrestrial repeaters.

INTRODUCTION

There has been considerable international effort in the areas of system studies, system development and regulatory work towards enabling a Satellite Sound Broadcast System (SSBS) [1]-[5]. An important milestone will be the 1992 World Radio Administrative Conference (WARC 1992) consideration of frequency allocation for such a service.

There is an ongoing effort at JPL to study a SSBS for reception in the UHF or L bands by portable and mobile receivers. Portable reception considerations have been discussed in [4]. The current paper presents preliminary results of the study concerning mobile reception. A flexible multilayered system is developed with each layer relatively independent of internal workings of other layers to facilitate the orderly evolution of the system, and allow adaptation of technologies from other communication systems either existing or under development. Important system issues include: state of the art digital audio coding, power and bandwidth efficient channel coding and modulation techniques, anti multipath signalling and diversity techniques, and finally the space segment.

AUDIO CODING

There has been considerable work and progress in digital coding of 20 KHz audio due to ISO (International Standards Organization) activity towards achieving bit rates under 256 kbps for Compact Disk (CD) quality stereo music. Systems developed towards these objectives use either sub-band or transform coding and rely on the noise masking threshold of the human auditory system for efficient adaptive bit allocation schemes [6]-[9]. With these technologies, digital coding of music is feasible at bit rates in the order of 200 kbps to 256 kbps. Such systems can find immediate application in the recording industry, studio facilities, as well as in terrestrial and satellite audio transmission systems. One can safely assume that official or defacto standards for CD quality audio at bit rates of 256 kbps (4 times the basic 64 kbps ISDN channel) will be firmly in place before SSBS will be implemented. The use of this standard for high quality SSBS is one obvious option. Development of a more compact bit rate around 192 kbps (3 time the basic 64 kbps) is also likely. Based on the properties of the human ear, it is unlikely to achieve significantly lower bit rates without impairing the desired CD quality for this application. Thus we will assume bit rates of 192 kbps to 256 kbps as the lower and upper limits for CD quality audio.

One can significantly reduce the required bit rate by choosing a monophonic system and or allowing degradation of the audio quality in bandwidth and dynamic range. A typical bench mark is monophonic FM quality audio rated 4 on the subjective Mean Opinion Score (MOS) at 64 kbps [7], using the same audio coding techniques discussed above. We have

adopted the 64 kbps as the upper limit for a monophonic SSBS system. To allow for future advances in the technology, 48 kbps is used as the lower limit for such a systems. At present technology one can expect an audio quality somewhat lower than monophonic FM at the 48 kbps. There is already a CCITT (standard G.722) audio coder optimized for high quality speech (slightly higher than 4 on the MOS quality scale) at the 64 kbps digital rate. The quality is slightly below 4 (on the MOS scale) for music signals. However one cannot rule out the possibility that the same system can be modified to give satisfactory performance for music.

Based on the status of audio coding technology, we will consider two grades of service quality for SSBS:

- Digital broadcasting of monophonic audio with bit rates in the 48 kbps to 64 kbps.

- High fidelity digital Stereo Broadcasting at bit rates in the 192 kbps to 256 kbps range with quality approaching that of CD. The four fold increase in bit rate requirements from monophonic to CD quality audio is partially due to stereophonic requirements (factor of 2). The higher dynamic range and wider spectrum of CD quality audio result in an additional doubling of the bit rate requirements.

MOBILE RECEPTION AND PROPAGATION CONSIDERATIONS

There is a significant body of experimental, theoretical, and modelling work on the Land Mobile Satellite Service (LMSS) channel. Available information includes experimental data obtained with the ATS-6, Marecs, and ETS satellites as well as a number of terrestrial simulations of the LMSS channel using balloons, towers and aircraft. [10]-[22].

The coarse structure of the LMSS channel is determined by the intermittent blockage of the line of sight by roadside objects (shadowing). Such shadowing can be extensive in urban areas. It is illustrative to look at mobile reception propagation experiments with the ATS-6 [10].

The ATS-6 experiments provide information on the excess path Loss for the LMSS channel as a function of local environment, vehicle heading, and link frequency (860 MHz, 1550 MHz) in several cities with satellite elevation angle ranging from 19 degrees, in Chicago, to 43 degrees, in San Francisco. At the frequency of 860 MHz and satellite elevation angle of 32 degrees, the median signal attenuation exceeded 20 db for 4% of sectors for example in urban Denver. Each sector has been chosen as a "few hundred wavelengths long" (order of 100 meters). The shadowing effect is particularly sensitive to the street bearing. The median signal attenuation exceeded 20 db for 10% of sectors in the streets running perpendicular to the satellite azimuth. Measurements obtained in the same environment at a frequency of 1550 MHz indicate the same phenomenon but at attenuation levels a few db's higher.

The ATS-6 experiments show that severe urban shadowing effects have only a modest dependence on satellite elevation angle over the measured range(19 degrees to 45 degrees). This can be explained by the fact that severe shadowing is generally occurring in streets at right angle to the satellite azimuth, which remain shadowed anyway within these range of satellite elevation angles. Satellite elevation angles need to reach around 60 to 70 degrees to reduce this type of shadowing to an insignificant percent of the sectors [5]. Thus to provide urban coverage in urban setting one has three alternatives:

- Provide massive link margins (about 20 db) to compensate for shadowing losses and use a signal structure consistent with the fine structure of the shadowed signal. An example of this system is given by CCIR [23] defined as Advanced Digital System II, albeit with a link margin of only 10 db for 99% sector coverage. Such a link margin is insufficient and should be increased to at least 20 db based on the experimental results discussed above.

- Keep the satellite elevation angle very high (at least 60 to 70 degrees). This option is feasible with geostationary orbits only for locations with very low latitudes. For other locations one needs a non geostationary orbit. Under this scenario one can provide a link margin typical of a Rician channels, with short shadows mitigated by diversity techniques.[5].

- Provide enough link margin for a Rician channel, augmented with terrestrial retransmission to fill in urban shadowed areas. Short term signal blockage in rural and suburban areas can be mitigated by diversity techniques. One such system has been proposed based on the Eureka DBS-audio system developed in Europe [1]. We will explore this alternative using spread spectrum signalling techniques for the urban coverage.

Shadowing effects in rural and suburban areas have been studied extensively [10]-[22]. Roadside trees have

been identified as the most significant cause of signal blockage with shadowing depths as high as 20 db for the 1% probability level for some severely shadowed routes in Australia [12]. Tree trunks and branches are the major cause of deep shadowing, while the foliage on the trees has a modest additional contribution of 1 to 3 db's of attenuation [12]. Tree shadowing in SSBS can be mitigated by the use of diversity techniques to overcome deep but short shadows caused by trunks and branches. The same solution would also mitigate shadowing caused by utility poles. A few db's of extra link margin would also be needed to cover the shallow but long foliage shadowing. Rural and suburban overpasses, tunnels, and occasional buildings would result in deep and wide shadows which cannot be combated with this combined scenario. However such shadowing would be infrequent in rural areas and has to be tolerated.

We need to consider fade duration statistics to evaluate the effectiveness of various diversity techniques against tree shadowing. Measurements taken in central Maryland [11] with a helicopter based transmitter to simulate a satellite signal at 30 degrees elevation angle show some interesting results for tree shadowing. These measurements show that in the case of fades deeper than 5 db, 12 % lasted longer than 5 wavelengths (1 meter), and only 2% lasted more than 20 wavelengths (20 x 20 cm=4 meters).

Obviously, antenna diversity with separation of a few meters, or time interleaving with a length of about one second can mitigate a significant number of tree shadowing fades. Frequency diversity will not work because these fades are due to shadowing and not due to the multipath phenomenon. Antenna diversity is particularly attractive on the basis of link budget, but has the drawback that not all car manufactures may want to incorporate such schemes in their cars. Time interleaving does not have this problem, but would put a few db's of burden on the link budget. Within the scope of our design, modest additional link margin is not a significant problem as we are providing excess link margin of 10 to 12 db's for portable reception inside buildings anyway. Another problem with time interleaving is its sensitivity to the speed of the mobile. With a car stopped within a shadow, time interleaving does not work. However this problem is somewhat mitigated in the rural and suburban case where stoppage is less frequent than urban areas. Given these facts, one can include time interleaving as a system solution and leave antenna diversity as an elective option that the car manufacturer or the mobile user can exercise. Quantitative study of time interleaving parameters, including the FEC coding structure to be used, is currently being carried out at JPL. Antenna diversity improvements will also be considered.

In urban areas, where there is frequent deep and lengthy shadowing of the satellite signal, satisfactory reception will be practical with the help of terrestrial repeaters. Suburban areas would be borderline cases between urban and rural areas as candidates for terrestrial boosters. Next we will look at a tentative system architecture for such a hybrid system.

SYSTEM ARCHITECTURE AND LINK BUDGETS

The proposed system architecture is based on the use of a satellite with modest link margins to provide country wide coverage, with a signal structure to mitigate deep short fades encountered due to roadside tree trunks and branches and such other objects as utility poles and traffic signs in an otherwise Rician channel. Detailed link budget and satellite power and spectrum requirements are given in Figure 1 for a typical configuration to cover the CONUS with 4 beams of 3 degrees width. Sixteen programs per beam of digital sound at 48 kbit is assumed in this table. Based on indoor portable reception requirements [4], 10 watts of radiated satellite power per program has been assumed which results in 5.5 db of surplus link margin for mobile rural/suburban reception under unfavorable conditions.

For the typical system of Figure 1, the spectrum requirement is 1.54 MHz per beam resulting in 4.61 MHz per CONUS if the first and fourth beams share the same frequency. Let us name the three frequency blocks F1, F2, and F3, and the four beams B1, B2, B3, B4. Then the frequency assignments will be (F1,B1), (F2,B2),(F3,B3), and (F1,B4). Within each beam the two unassigned spectrum blocks can be used by spread spectrum terrestrial boosters to provide urban coverage. Figure 2 gives detailed analysis on the number of spread spectrum channels that can be supported in each urban area by such a scheme using terrestrial boosters. The analysis is based on the use of a specialized spread spectrum technique developed for the urban LMSS channel [24]. This spread spectrum coding system combines a very low-rate almost orthogonal convolutional code with a pseudo noise (PN) sequence. The very low rate code provides FEC capability with very low decoder Eb/No threshold, 3.65 db's, for example for a BER of 1.0E-5 for rate 1/32 coding. The coder also acts to spread the bit rate, in this case by a spreading factor of 32. The PN sequence has a chip rate equal to the symbol rate of the coder and merely acts to randomize the output symbols. As coding is almost orthogonal, decoding implementation is very

Figure 1. LINK BUDGET FOR RURAL/SUBURBAN MOBILE RECEPTION

MONOPHONIC DIGITAL SOUND AT 48 KBPS
Transmission frequency= 1 GHz
FDM, Same frequency reused every 3rd beam.
R=1/2 , Conv. code, soft decoding, BER=1.0E-4

AUDIO LINK BUDGET	UNITS	Fav. valu	Unfav.
Inf. bit rate/program	KBPS	48.00	48.00
Inf. bit rate/program	DB	46.81	46.81
XMITTER PWR per program	WATTS	10.00	9.00
XMITTER PWR per program	DBW	10.00	9.54
FREQUENCY	GHZ	1.00	1.00
SATELLITE ANTENNA DIA	M	7.00	7.00
XMIT ANT GAIN	DB	34.71	34.71
ANT BEAMWIDTH	DEG	3.00	3.00
EIRP	DBW	44.71	44.25
RANGE	KM	40000.00	40000.00
FREE SPACE LOSS	DB	184.48	184.48
RCV ANT DIA	M	0.23	0.18
RCV ANT GAIN	DB	5.00	3.00
RCV ANT BEAMWIDTH	DEG	91.70	115.58
RCVD SIG PWR	DB	-134.77	-137.24
RCV SYST TEMP	K	200.00	400.00
RCV SYST TEMP	DB	23.01	26.02
RCV SYST G/T	DB	-18.01	-23.03
No	DBW/HZ	-205.59	-202.58
C/No	DBHz	70.82	65.35
Eb/No Available At Beam Center	DB	24.01	18.53
Eb/No At Beam Edge	DB	21.01	15.53
Implementation loss	DB	1.00	2.00
Req. Eb/No	DB	6.00	8.00
Surplus link margin	DB	14.01	5.53
Spot Beams/CONUS		4.00	4.00
Numb of Prog/Spot Beam		16.00	16.00
SAT Radiated Power/CONUS	Watts	640.00	576.00
Code Rate		0.50	0.50
Modulation Rate QPSK,	Bit/Hz	0.50	0.50
Bandwidth/Program	KHz	96.00	96.00
Bandwidth/SpotBeam	MHz	1.54	1.54
Bandwidth/CONUS	MHz	4.61	4.61
Spectral Efficiency/CONUS	Bit/Hz	0.67	0.67

Figure 2. CHANNEL CAPACITY CALCULATIONS FOR URBAN TERRESTRIAL SYSTEM

spread-spectrum signaling with FEC.,

					units
conv. code constraint length	7	7	7	7	
code rate	R=1/32	R=1/32	R=1/32	R=1/32	ratio
spread factor	32	32	32	32	ratio
assumed Eb/No ratio, excluding interference	20.00	20.00	20.00	20.00	db
assumed No/Eb ratio, excluding interference	0.01	0.01	0.01	0.01	ratio
program channels/frequency block	8	12	8	12	channels
spread spectrum pilot channel/ frequency block	1	1	1	1	channels
interfering channels/frequency block	8	12	8	12	channels
Ni/Eb	0.25	0.38	0.25	0.38	ratio
Eb/Ni, signal to interference ratio	4.57	2.91	4.57	2.91	ratio
combined noise and interference to signal ratio	0.26	0.38	0.26	0.38	ratio
available Eb/(No+Ni)	5.85	4.15	5.85	4.15	db
decoder Eb/(NI+No) threshold for BER=1.0E-4	2.85	2.85	2.85	2.85	db
data margin for BER=1.0E-4	3.00	1.30	3.00	1.30	db
decoder Eb/(Ni+No) threshold for BER=1.0E-5	3.65	3.65	3.65	3.65	db
data margin for BER=1.0E-5	2.20	0.50	2.20	0.50	db
number of spread spectrum chanells/frequency block	8	12	8	12	channels
number of frequency blocks /location	2	2	2	2	
number of spread spectrum chanells/location	16	24	16	24	channels
satellite broadcast programs/location	16	16	16	16	programs
terrestrial channels surplus to satellite programs	0	8	0	8	channels
bit rate/Channel	48.00	48.00	256.00	256.00	kbps
chip rate	1536.00	1536.00	8192.00	8192.00	kcps
spectrum per frequency block	1.54	1.54	8.19	8.19	MHz
frequency blocks/CONUS	3	3	3	3	
spectrum/CONUS	4.61	4.61	24.58	24.58	MHz

simple. As the decoder has a very low Eb/No threshold, the channel can tolerate the interference from a relatively large number of other spread spectrum channels using the same spectrum. In the example of Figure 2, twelve spread spectrum channels with equal power can share the same spectrum. Excluding spread spectrum interference, the signal to system noise ratio, Eb/No, is assumed to be 20 db. One also should account for the interference from the other 11 program channels and from one pilot channel. We note that the signal to interference ratio, Eb/Ni, is given by the spread factor, 32, divided by the number of interfering channels, 12, resulting in Eb/Ni=2.67 or 4.26 db's. There is a slight signal degradation due to system noise. The overall signal to noise ratio, Eb/(No+Ni), is 4.15 db's, giving us a data margin of 0.5 db above the decoder threshold of 3.65 db's for a BER of 1.0E-5.

For this specific system configuration, the two unused frequency blocks in each beam footprint can support 24 terrestrial spread spectrum program channels, resulting in eight surplus programs over the 16 satellite based programs. This extra capacity can be used to support local programs. One can also operate the spread spectrum system below capacity just to retransmit the satellite based programs. This back off of the spread spectrum capacity results a further improvement in the data margin to 2.2 db's over the 3.65 db threshold for BER of 1.0E-5.

The spread spectrum channels can use multiple repeaters on the same frequency to provide ideal coverage of urban areas. Rake type receivers can combine signals received from multiple transmitters resulting in very effective diversity techniques. Such a system has been demonstrated in the US for cellular mobile telephony applications [25]. A completely different system, that does not use spread spectrum signaling, developed by Eureka project in Europe has somewhat similar capabilities as long as the multiple signals received from different repeaters do not differ in path length by more than 5 kilometers. The spread-spectrum system discussed here does not suffer from this limitation and can provide very flexible coverage of urban areas.

Figure 3 shows system requirements for a number of broadcast quality options for CONUS coverage.

CONCLUSIONS

ATS-6 experiments at 860 MHz and at satellite elevation angles in the range of 19-43 degrees indicate typical 20 db median signal attenuation for urban streets shadowed by buildings. Signal attenuation is a few db's higher at 1.5 GHz. The 10 db link margin adopted by CCIR [23] for urban shadowing needs to be accordingly revised.

With the marriage of proper technologies an LMSS class satellite can provide substantial direct audio broadcast capability in UHF or L Bands for high quality mobile and portable indoor reception by low cost radio receivers. This scheme requires terrestrial repeaters for satisfactory mobile reception in shadowed urban streets. A specialized spread spectrum technique developed for the urban LMSS channel [24] is particularly suitable for the terrestrial repeaters.

ACKNOWLEDGEMENT

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

Figure 3. TYPICAL SYSTEM PARAMETERS FOR A NUMBER BROADCAST QUALITY OPTIONS

Carrier frequency = 1 GHz	Almost mono FM	Mono FM	Almost CD	CD	
Digital sound quality comparable to:					units
Bit rate/Channel	48	64	192	256	kbps
programs /locations	16	16	16	16	channels
Urban booster channels/location	24	24	24	24	channels
Surplus urban booster channels	8	8	8	8	channels
Spectrum requiremen/CONUS	4.61	6.14	18.44	24.6	MHz
Satellite radiated RF power/program	10	13.4	40	53.4	watts
Satellite radiated RF power/CONUS	640	854	2560	3414	watts
Surplus mobile link margin, Fav.	14	14	14	14	db
Surplus mobile link Margin, unfav.	5.5	5.5	5.5	5.5	db

REFERENCES

1 G. Waters, and F. Kosamernik, " Plans and Studies in the EBU for Satellite Broadcasting of Sound Radio" proceedings 13th AIAA international Communication Satellite Systems Conference, Los Angles, CA, March 1990, pp 176-185.

2 J. E. Miller, "Application of Coding and Diversity To UHF Satellite Sound Broadcasting Systems" IEEE Transactions on Broadcasting, Vol. 34, No. 4, December 1988, pp 465-475.

3 T. Rogers, "Some important Advances in international Direct Audio Broadcasting "proceedings 13th AIAA international Communication Satellite Systems Conference, Los Angles, CA, March 1990, pp 205-208.

4 N. Golshan, and A. Vaisnys, "Satellite Sound Broadcasting System, Portable Reception " proceedings 13th AIAA international Communication Satellite Systems Conference, Los Angles, CA, March 1990, pp 186-204.

5 D.K. Banks and D. Robson, " The Use of Orthogonal FDM for Sound Broadcasting by Satellites in Highly inclined Orbits to Overcome Multipath Fading " paper presented at the 13th AIAA international Communication Satellite Systems Conference, Los Angles, CA, March 1990.

6 J.D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal Selected Areas in Communications, Vol. 6, No. 2, pp 314-323, February 1988.

7 N.S. Jayant, "High Quality Coding of Telephone Speech and Wideband Audio " To appear in IEEE Journal on Selected Areas in Communications.

8 E.F. Schroeder , H.J. Platte, and D. Krahe, " MSC Stereo Audio Coding with CD-Quality and 256 KBit/sec" IEEE Transaction on Consumer Electronics, Vol. CE-33, No. 4, November 1987, pp512-519.

9 G. Theile, G. Stoll, and M. Link, "Low bit-rate coding of high-quality audio signals, An Introduction to the MASCAM system" EBU Review-Technical, No. 230, August 1988, pp71-94.

10 G. C. Hess, " Land Mobile Satellite Excess Path Loss Measurements ", IEEE Transactions " IEEE Transactions on Veh. Tech , Vol VT-29, No. 2, May 1980, pp 290-297.

11 J. Goldhirsh and W.J. Vogel, "Mobile Satellite System Fade Statistics for Shadowing and Multipath from Roadside trees at UHF and L-Band" IEEE transactions on Antennas and Propagation, Vol 37, No.4, April 89, pp 489-498.

12 W. Vogel and J. Goldhirsch, "Propagation Limits for Land Mobile Satellite Service" proceedings 13th AIAA international Communication Satellite Systems Conference, Los Angles, CA, March 1990, pp 564-574.

13 W.J. Vogel and J. Goldhirsh, "Fade Measurements at L-Band and UHF in Mountainous Terrain for Land Mobile Satellite Systems" IEEE Transactions on Antennas and Propagation, Vol. 36, No.1, January 88, pp 104-113.

14 J. Goldhirsh and W.J. Vogel, "Roadside Tree Attenuation Measurements at UHF for Land Mobile Satellite Systems" IEEE Transactions on Antennas and Propagation, Vol AP-35, No. 5, May 87, pp 589-596.

15 W. J. Vogel, et. al., "The Australian Experiment with ETS-V " Proceedings of NAPEX XIII, June 1989.

16 W.L. Stutzman, "Mobile Satellite Propagation Measurements and Modeling: A Review of Results for Systems Engineers " Proceedings of the Mobile Satellite Conference, May, 1988 pp 107-117.

17 R.M. Barts and W.L. Stutzman, "Propagation Modeling for Land Mobile Satellite Systems" Proceedings of the Mobile Satellite Conference, May 1988 pp 95-100.

18 D.C. Nicholas, "Land Mobile Satellite Propagation Results" Proceedings of the Mobile Satellite Conference, May 1988. pp 125-131.

19 R. J. C. Bultitude, "Measured Characteristics of 800/900 Mhz Fading Radio Channels with High Angle Propagation through Moderately Dense Foliage ", IEEE Journal on SAC February 87.

20 K. Dessouky and Loretta Ho, "Field Measurements in Msat-X", Proceedings of NAPEX XIII, June 1989, pp 18-21.

21 A. Bundrock, and R. Harvey, "Propagation Measurements for an Australian Land Mobile Satellite System " Proceedings of the Mobile Satellite Conference, May 1988, pp 119-124.

22 Jan Van Rees, "Measurement of the Wide-Band Radio Channel Characteristics for Rural, Residential, and Suburban Areas" IEEE Transactions on VT, Vol. 36, No. 1, February 1987.

23 CCIR, "Chapter 6, Technical Information to define the practical system parameters for satellite sound broadcasting (Recommendation No. 2 of WARC ORB-85," CCIR Doc. JIWP/ORB(2)/90-E, December 1987, Geneva, Switzerland.

24 A.J. Viterbi, "Theoretical Foundations for CDMA: Processing Gain, Interaction of FEC and Spreading Codes and Ultimate Performance in Mobile Satellite Communications" Proceedings of The Mobile Satellite System Architectures and Multiple Access Techniques Workshop, March 1989.

25 PacTel Cellular and Qualcomm, Inc "CDMA Cellular-The next generation" Qualcomm, Inc, San Diego, CA, November 1989.

Multiple Access Capacity Trade-offs for a Ka-Band Personal Access Satellite System

Khaled Dessouky and Masoud Motamedi

California Institute of Technology
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91109
Phone: 818-354-8041
FAX: 818-393-4643

ABSTRACT

System capacity is critical to the economic viability of a personal satellite communication system. Ka-band has significant potential to support a high-capacity multiple access system because of the availability of bandwidth. System design tradeoffs are performed and multiple access schemes compared with the design goal of achieving highest capacity and efficiency. Conclusions regarding the efficacy of the different schemes and the achievable capacities are given.

1.0 INTRODUCTION

The telecommunications infrastructure of the 21st Century will very likely be characterized by a diversity of services and a choice of media. In anticipation of the future needs in communications, the Jet Propulsion Laboratory (JPL) is exploring the potential and feasibility of a Personal Access Satellite System (PASS) which is intended to offer the user freedom of access and mobility [1,2].

The telecommunications industry of the future will undoubtedly witness fierce competition. The different systems will have to provide their benefits to the users in a cost-effective and efficient manner. Crucial to the economic viability of a satellite communication system targeted to the individual user is competitive and affordable user equipment. The importance of the reduction in cost achieved through the economies of scale cannot be over-emphasized. One of the primary reasons for selecting Ka-band for PASS is the availability of a considerable amount of bandwidth, easily an order of magnitude more than at L-band or UHF. This, in a successful system design, should translate into proportionally larger capacities, and in turn would translate into lower costs.

This paper addresses the issue of system capacity. Different multiple access scheme combinations are considered and compared. Tradeoffs of system parameters are performed to achieve highest capacity (in number of channels) and optimize efficiency (in channels/Hz). The implications of the results and comparisons are explained and conclusions regarding the efficacy of the different schemes are given.

2.0 SYSTEM PARAMETERS

The design of a PASS architecture is an intricate process that involves a multitude of factors. A first set of design parameters includes satellite RF power, overall system bandwidth, link performance specification, coding gains, voice activity, and ultimately, overall system capacity. Another set of factors that could be considered include number of beams on each satellite link, user EIRP, user receive G/T, basic terminal types and associated data rates. This latter set of parameters is tied directly or indirectly to the capabilities of the user terminal. Since those capabilities have evolved through a study of soon-to-be-available or projected Ka-band technologies [2], it is felt that design optimization should, at least at this stage, focus only on the former group of parameters. In addition to avoiding a radical impact on PASS, this also renders the multiple access design problem tractable.

In 1988 a system architecture utilizing a hybrid Time Division Multiple Access (TDMA)/ Frequency Division Multiple Access (FDMA) was investigated [2,3]. The architecture called for TDMA in the forward direction (from Suppliers to Users), and FDMA in the return direction (from Users to Suppliers). An alternative architecture employing

Random Access Code Division Multiple Access (CDMA) was studied in 1989 [4]. The CDMA schemes considered in [4], as well as here, employ direct-sequence spreading, and could therefore be referred to also as Spread Spectrum Multiple Access (SSMA). Direct-spreading provides added benefits in the personal or mobile environment including multipath rejection and position determination.

Based on the chosen design approach, a set of basic system architecture constraints is common to all of the access schemes considered. These include the use of the satellite as a bent-pipe repeater, with a CONUS beam for the satellite/supplier side, a multi-beam antenna on the satellite/user side, and with a fixed set of parameters such as gain and G/T. The basic user terminal and supplier station also have pre-selected specifications. Table 1 contains a summary of the key system parameters that have been kept fixed. On the satellite/user links, frequency re-use is employed on the 142 beams so that only 9 frequency bands are used in covering CONUS.

Table 1. Summary of Pre-Set PASS Parameters

GENERAL	
OPERATING FREQUENCIES	
UPLINK	30 GHZ
DOWNLINK	20 GHZ
COVERAGE CONCEPT	
SAT/SUPPLIERS	CONUS BEAM
SAT/USERS	SPOTBEAMS
BASIC PERSONAL TERMINAL	
G/T	-9.0 DB/K
EIRP	16.8 DBW
BASIC DATA RATE	4800 BPS
SATELLITE	
SPOTBEAM ANTENNA	
NUMBER OF SPOTBEAMS	142
ANTENNA GAIN	52.5 DBI
SYSTEM G/T	23.4 DB/K
AVERAGE EIRP/BEAM	55 DBW
CONUS ANTENNA	
ANTENNA GAIN	27.0 DB
SYSTEM G/T	-1.2 DB/K
EIRP	39 DBW
SUPPLIER STATION	
G/T	60.7 DB/K
EIRP	30.3 DB

3.0 LINK CHARACTERIZATION

The factors of satellite RF power, system bandwidth, link performance/coding performance and system capacity are all tied together through a set of link budget equations, complimented with bandwidth and capacity computations. Two link budget equations, one for the forward and one for the return, are needed for each multiple access scheme. Each pair of forward and return budgets is tied together through the key constraint of limited overall satellite RF power. Occasionally, the satellite power used on the two link directions could be traded effectively to increase overall capacity, or to balance the forward and return capacities. Unfortunately, in many circumstances the gains achieved are limited due to the constraints placed on the system.

A simple approach to understand the various situations existing on the different links is to consider the basic equation relating the received bit signal to noise ratio to the down-link and up-link carrier to noise ratios, and in the case of CDMA, to the added mutual interference. This equation can be written as

$$\left\{\frac{E_b}{N_0}\right\}^{-1} = \left\{\frac{P_{rd}}{R_b N_0}\right\}^{-1} + \left\{\frac{P_{ru}}{R_b N_{0u}}\right\}^{-1} L_d + \left\{\frac{R_c}{(M-1)R_b}\right\}^{-1} \quad (1)$$

where E_b denotes the received energy per bit, N_0 is the one-sided thermal noise power spectral density. P_r is the received signal power with the second subscript d or u denoting down-link or up-link (at the satellite), respectively. L_d is the loss that the transponded up-link signal plus noise experience by going through the satellite and the down-link environment. R_c is the chip rate, R_b is the bit rate, and (M-1) are the simultaneous interfering users. For FDMA the third term is simply dropped.

When the first term in the right hand side of (1) dominates, link performance is limited by the thermal noise on the down-link. Similarly, when the second term dominates, performance is limited by up-link thermal noise. Finally, in a CDMA system, if the third term dominates, link operation is mutual interference limited. The inverse of each term (i.e., the quantity between parentheses) can be regarded as an effective signal to noise ratio (SNR) for either the down-link, up-link, or mutual interference. Naturally, the lowest SNR drives the attainable E_b/N_0. An "efficient" system design generally requires more or less equal contributions from the three terms. In a system such as PASS this is rarely

achievable due to the constraint on the overall RF power. This power is not only used to amplify the uplink signal but also the uplink noise, which can become significant in a wide band system. Although this "power robbing" type of effect is not explicitly shown in (1), it is manifested in a drop in P_{rd} (an increase in the first term of (1)) when bandwidth is increased to accommodate a higher chip rate; so an attempt to reduce the third term results in increasing the first. Consequently, an "optimal" chip rate can be found to maximize link performance under the given power constraints.

4.0 PERFORMANCE TRADEOFFS

The strawman design of PASS described in [2,3] has utilized a 6000 lb-class satellite providing 410 watts of RF power. This satellite along with a somewhat larger satellite with roughly 25% more RF transmit power capability are selected for the tradeoff. Different satellite powers are considered since system performance in severely noise limited conditions inherently favors FDMA. In fact, it will become clear from the discussions to follow that a satellite sized for an FDMA architecture does not generally support an optimal CDMA design. This can be intuitively derived from (1) since there is one more term that the system/satellite designer has to contend with in CDMA.

In the tradeoffs link BER specification is taken to be either 10^{-3} or 10^{-5}. The usual performance for voice channels is 10^{-3}. However, it is possible that a more stringent 10^{-5} requirement be placed on data links. The E_b/N_0 needed is determined by the BER and the choice of a suitable coding scheme consistent with either the FDMA, TDMA or CDMA approach.

One of the inherent advantages of CDMA is that coding gain can be achieved without further expanding the bandwidth. Consequently, powerful codes can be applied without that usual penalty. Convolutional codes of rates R=1/2 and R=1/3 and constraint lengths K=7 and K=9 are considered. Lowering the code rate from 1/2 to 1/3 with K=7 is achieved at only a minor cost/complexity increment to the user terminal. The more powerful code with R=1/3 and K=9 is included for its lower E_b/N_0 requirement of about 1.5 dB for a BER of 10^{-3} in additive white Gaussian noise. As will be seen, CDMA capacity is quite sensitive to this E_b/N_0 requirement. A novel set of codes known as super-orthogonal codes could be selected to obtain this 1.5 dB performance [5]. These codes have a rate of $2^{-(K-2)}$, and in a spread spectrum system the code rate is taken to be the ratio of the bit to chip rates. Reportedly [5], reasonable K values for such a system are 10 to 12. The impact of having the symbol rate equal to the chip rate on the robustness of CDMA needs a close look; however, the main advantage of this coding scheme is the reduced E_b/N_0 requirement. Hence, for the purposes of our tradeoffs the R=1/3, K=9 code and the super-orthogonal codes are generally equivalent.

For FDMA both rate 1/2 and 1/3 codes with K=7 and 9 are considered. Super-orthogonal codes are not applicable since they would expand the bandwidth by at least 256 times.

The bandwidth requirements computed in what follows are based on some assumptions. The baseline beam and frequency plan mentioned above, and described in detail in [2,3,4], is assumed. For CDMA a channel bandwidth is taken to be twice the chip rate. For FDMA, twice the symbol rate is used plus 5 kHz of guard band is allowed. No allowance is considered for intermodulation products avoidance in either scheme since the satellite HPA is assumed to operate in the linear region. This is necessitated by the baseline multiple beam/FDM architecture common to either the FDMA or CDMA strategies.

4.1 Lower Power Satellite Trades

The forward and return link capacities for the 410 W satellite (in terms of number of 4800 BPS users) are given in Table 2. A host of coding choices and service types (voice or data) is provided.

We start by observing that under the given power and bandwidth constraints the CDMA/CDMA approach cannot compete with either FDMA/TDMA or CDMA/TDMA. A close examination of the link budgets and the applicable terms in (1) reveals certain inherent limitations in the design problem. For the CDMA/CDMA entries number 9 or 10 of Table 2, the down/up/mutual SNR's (R.H.S. of (1)) are 11/10/20 for the return and 5.5/96/21.7 for the forward, where the numbers are in ratio. The optimal satellite power allocation was found to be 375/35 for the return/forward directions. This clearly shows that system performance is severely thermal noise limited on the forward down-link; which is indeed the segment that requires most of the satellite power. Numbers for the FDMA/TDMA baseline [2,3] (corresponding to entry 1 in the table) are 10/22 for the down/up

return and 7.6/93.7 for the down/up forward. Obviously, the baseline design is also power limited on the forward down-link. This bottleneck on the forward down-link is further exacerbated in the CDMA design, particularly when both directions use CDMA. This is because some satellite power has to be set aside to combat mutual interference and to amplify a wider up-link noise band; the forward down-link becomes even more power starved.

For CDMA to be a viable candidate the solutions involve one or more of the following: 1) use CDMA only on the return link and TDMA on the forward, this results in bandwidth and power savings by eliminating the spreading on the forward link if random access is not needed; 2) increase the satellite power to enable the multiple access needs of CDMA while not aggravating the thermal noise bottleneck in the forward down-link; 3) reduce the received power requirements at the user such as with the use of R=1/3, K=9 or super-orthogonal codes.

The comparison of FDMA/TDMA with CDMA/TDMA is also given in Table 2 for the 410 W satellite. For data links FDMA is clearly the proper choice based on the number of channels and the required bandwidth. This is seen by comparing entries 1 and 4 versus 11 for a BER requirement of 10^{-5}, and 5 versus 12 for a BER of 10^{-3}. Roughly the same data channel capacity is obtained at half the bandwidth (compare entries 5 and 12). Because of the voice activity factor (0.35) CDMA excels in a voice dominated system; as channels are added the bandwidth requirement does not change-- whereas it increases substantially for FDMA (compared to data only). Entries 3 and 6 for FDMA and 13 for CDMA clearly demonstrate this fact. A higher number of voice channels per Hz is obtained with CDMA, even without the more powerful codes requiring only 1.5 dB E_b/N_0. Going to R=1/3, K=9 or super-orthogonal codes (entries 7, 14 and 15) CDMA's advantage in channels/Hz increases further.

An interesting tradeoff can be seen in entries 15a,b. The power savings realized on the return link can be transferred to the forward direction to boost its capacity. Because the forward link is so power thirsty, the gains obtained in this manner are not large. Alternately, the capacity of the forward can be left fixed, and dramatic gains shown on the return (entry 15.a). This will be illustrated further in the case of a 520 W satellite.

4.2 Higher Power Satellite Trades

The situation with 520 W RF power on the satellite is quite interesting because it ameliorates the power bottleneck on the forward down-link. Steps similar to above are followed to optimize the RF power distribution between the forward and return directions and to optimize the chip rate and bandwidth. The results are shown in Table 3.

The first observation is that increased RF satellite capability notwithstanding, FDMA is still the better choice for data. The situation becomes quite different for a system dominated by voice users (entries 3 and 8 for example). Roughly three times as many users as the all data case can be supported with CDMA at no extra cost. The equivalent increase in channels for FDMA is achieved at a three fold increase in bandwidth. Table 3, entries 3 and 8, give the net results for the same total bandwidth of 285 MHz. The results evince a slightly higher CDMA capacity in the forward direction and a 16% advantage for the return.

It is interesting to note here that power limitations on the forward link persist (SNR break-ups for entry 8 are 4.4/76.7 forward and 25.5/10/10.1 return). Increasing the satellite power beyond 520 Watt would predominantly improve the forward capacity. As mentioned earlier, an alternate approach is the use of a more powerful code on the CDMA return and transferring some power to the forward link. This is achieved with either the R=1/3, K=9 or the super-orthogonal codes as demonstrated in entries 9 and 10. In particular, entry 9.a when compared to entry 3 shows CDMA capacity advantages of 8% on the forward and 16% on the return, together with an 8% savings in total bandwidth. Alternately, the forward capacity can be maintained as in entry 8 and all of the performance savings used on the return to realize a 43% advantage over FDMA (entry 9.b versus 3). In fairness it should also be mentioned that super-orthogonal decoding is likely to be more complex than typical Viterbi decoding.

The final step in the CDMA vs. FDMA comparison centers around allowing a higher overall PASS system bandwidth. The bandwidth is allowed to exceed the "magical number" of 285 used above. This comparison is relevant here since there is enough satellite power to use the extra bandwidth. As the code rate is reduced to 1/3 in FDMA the bandwidth leaps from 285 to 468 or 631 MHz, depending on the code used, to support the

additional voice channels now feasible. For CDMA, either R=1/3, K=9 or super-orthogonal codes with K=11 are used. A comparison of entries 10 and 11 for CDMA with 4 and 5 for FDMA demonstrates a considerably higher efficiency in channels/Hz for CDMA.

5.0 CONCLUSIONS

The following conclusions can be drawn in the context of the above results and discussions.

- Due to power limitations on the satellite, and for bandwidth efficiency considerations as well, CDMA should not be used on the forward link. TDMA should be used.

- For a system dominated by data users, FDMA is superior based on lower bandwidth requirements. FDMA in general can support a higher number of channels if performance is very power limited (e.g., a thermal noise limited performance on the forward down-link if a 10^{-5} data BER is the predominant requirement in the system).

- For a system dominated by voice users CDMA is superior; it generally requires less bandwidth than FDMA, or can support a higher number of users for a given bandwidth.

- The increases in capacity with the lowering of the user E_b/N_0 requirement is more significant for CDMA than for FDMA. Alternately, increased satellite power is more advantageous for CDMA in the sense that it can be used more efficiently than in FDMA.

- Since future trends are for lower E_b/N_0 requirements and higher satellite RF powers, CDMA appears to be a stronger candidate for a state of the art system (provided that a significant proportion of the traffic is voice).

- Overall capacities that are half to a full order of magnitude higher than at L-band [6] are achievable. However, a concomitant increase in overall system bandwidth of about an order of magnitude is experienced. This bandwidth requirement is one of the primary reasons for migrating to the uncrowded Ka-band region.

ACKNOWLEDGEMENT

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The work presented here is part of a wider effort to which Barry Levitt and Miles Sue have had many valuable contributions.

REFERENCES

1. M. K. Sue, K. Dessouky, B. Levitt and W. Rafferty, "A Satellite-Based Personal Communication System for the 21st Century," IMSC'90, these proceedings.

2. M.K. Sue, ed., "Personal Access Satellite System Concept Study," Internal JPL Document: D-5990, Jet Propulsion Laboratory, February 1989.

3. M.K. Sue, A. V. Vaisnys, and W. Rafferty, "A 20/30 GHz Personal Access Satellite System Study," Proceedings of the 38th Vehicular Technology Conference, June 1988.

4. M. Motamedi, and M. K. Sue, "A CDMA Architecture for a Ka-Band Personal Access Satellite System," 13th AIAA International Communications Satellite Systems Conference, March 1990.

5. A. Viterbi, "A New Class of Very Low Rate Convolutional Codes with Application to Spread Spectrum Multiple Access," pre-print, April 1989.

6. C. Wang, T.-Y. Yan and K. Dessouky, "Performance of DA/FDMA Architecture Proposed for MSS," Proceedings of the Mobile Satellite System Architectures and Multiple Access Techniques Workshop, JPL Publication 89-13, March 1989.

TABLE 2. COMPARISONS OF PASS CAPACITIES FOR DIFFERENT MULTIPLE ACCESS DESIGNS FOR A SATELLITE WITH 410 W TOTAL RF POWER
NOTE: VOICE ACTIVITY FACTOR (VOX) FOR VOICE CHANNELS TAKEN TO BE 0.35

ENTRY #	ACCESS SCHEMES	LINK PARAMETERS BER	Eb/NO dB	CODING (CONVOLUTIONAL)	SERVICE TYPE	CAPACITY (# CHANNELS) RETURN	FORWARD	BANDWIDTH (MHz) UP-LINK	DN-LINK	TOTAL	SAT POWER,W SPLIT (F/R)	COMMENTS
1	FDMA (RET) / TDMA (FWD)	1.00E-05	4.5	R=1/2, K=7	DATA	2000	2036	27.7	27.3	55	390/20	
2	"	1.00E-03	3	R=1/2, K=7	DATA	2825	2875	39.1	38.6	77.7	390/20	
3	"	1.00E-03	3	R=1/2, K=7	VOICE	8072	8216	111.8	110.3	222.1	390/20	
4	"	1.00E-05	4	R=1/3, K=7	DATA	2244	2284	43.4	42.8	86.2	390/20	
5	"	1.00E-03	2.3	R=1/3, K=7	DATA	3319	3378	64	63	127	390/20	
6	"	1.00E-03	2.3	R=1/3, K=7	VOICE	9483	9653	183.4	181	364.4	390/20	
7	"	1.00E-03	1.5	R=1/3, K=9	VOICE	11401	11605	220	218	438	390/20	
8	CDMA (RET) / CDMA (FWD)	1.00E-05	4	R=1/3, K=7	DATA	852	1278	150	100	250	350/60	
9	"	1.00E-03	2.3	R=1/3, K=7	DATA	2414	2272	225	225	450	375/35	
10	"	1.00E-03	2.3	R=1/3, K=7	VOICE	6897	6491	225	225	450	375/35	
11	CDMA (RET) / TDMA (FWD)	1.00E-05	4	R=1/3, K=7	DATA	1704	1775	50.9	203	253.9	295/115	Rc=1.40625M; Rb=60K
12	"	1.00E-03	2.3	R=1/3, K=7	DATA	3550	2958	65.1	183	248.1	335/75	Rc=1.25M; Rb=100k
13	"	1.00E-03	2.3	R=1/3, K=7	VOICE	10143	8452	65.1	183	248.1	335/75	Rc=1.25M; Rb=100k
14	"	1.00E-03 1.00E-03	1.5 (R)/ 2.3 (F)	SUP.-ORTH(K=10)/ R=1/3, K=7	VOICE	10142	9331	69.1	180	249.1	375/35	Rc=1.228800M Rb=110.4k
15.a	"	1.00E-03	1.5	R=1/3, K=9	VOICE	10142	11072	78.3	185	263.3	375/35	Rc=1.25M; Rb=131k
15.b					VOICE	16228	9720	76	219	295	330/80	Rc=1.5M; Rb=115k

TABLE 3. COMPARISONS OF PASS CAPACITIES FOR DIFFERENT MULTIPLE ACCESS DESIGNS FOR A SATELLITE WITH 520 W TOTAL RF POWER
NOTE: VOICE ACTIVITY FACTOR (VOX) FOR VOICE CHANNELS TAKEN TO BE 0.35

ENTRY #	ACCESS SCHEMES	LINK PARAMETERS BER	Eb/NO dB	CODING (CONVOLUTIONAL)	SERVICE TYPE	CAPACITY (# CHANNELS) RETURN	FORWARD	BANDWIDTH (MHz) UP-LINK	DN-LINK	TOTAL	SAT POWER, W SPLIT (F/R)	COMMENTS
1	FDMA (RET) / TDMA (FWD)	1.00E-05	4.5	R=1/2, K=7	DATA	2600	2585	35.3	35.4	70.7	494/26	
2	"	1.00E-03	3	R=1/2, K=7	DATA	3673	3652	49.8	50	99.8	494/26	
3	"	1.00E-03	3	R=1/2, K=7	VOICE	10493	10433	142.3	143	285.3	494/26	
4	"	1.00E-03	2.3	R=1/3, K=7	VOICE	12328	12258	233.6	234.6	468.2	494/26	
5	"	1.00E-03	1.5	R=1/3, K=9	VOICE	16630	16536	315	316.5	631.5	494/26	
6	CDMA (RET) / TDMA (FWD)	1.00E-05	4	R=1/3, K=7	DATA	2070	2130	58	226	284	350/170	Rc=1.5625M (R) Rb=70k (F)
7	"	1.00E-03	2.3	R=1/3, K=7	DATA	4260	3698	78.6	206.4	285	425/95	Rc=1.40625M;RB=125k (R) (F)
8	"	1.00E-03	2.3	R=1/3, K=7	VOICE	12171	10565	78.6	206.4	285	425/95	Rc=1.40625M;RB=125k
9.a	"	1.00E-03	1.5 (R)/ 2.3 (F)	S. ORTH.(K=10)/ R=1/3, K=7	VOICE	12171	11411	80	182	262	470/50	Rc=1.2288M;Rb=135k
9.b					VOICE	15011	10565	75.4	181	256.4	425/95	Rc=1.2288M;Rb=125k
10.a	"	1.00E-03	1.5	R=1/3, K=9	VOICE	13794	13523	99.8	258.2	358	470/50	Rc=1.7578M;Rb=160k
10.b					VOICE	19069	11411	85.1	225	310.1	390/130	Rc=1.5328M;Rb=135k
11	"	1.00E-03	1.5 (R)/ 2.3 (F)	S. ORTH.(K=11)/ R=1/3, K=7	VOICE	17851	11411	101.8	356.3	458.1	460/60	Rc=2.4576M;Rb=135k

Feasibility of an EHF (40/50 GHz) Mobile Satellite System Using Highly Inclined Orbits

G.Falciasecca (*), A.Paraboni (**),
M.Ruggieri, F.Valdoni, F.Vatalaro (***)

(*) DEIS, Universita' di Bologna (ITALY)
(**) DIE, Politecnico di Milano (ITALY)
(***) DIE,Universita' di Roma"Tor Vergata"
Via E.Carnevale, 00173 - Roma (ITALY)
Phone:+ 39 - 6 - 24990464; Fax:+39 - 6 -2490519

ABSTRACT

The pan-European GSM cellular system is expected to provide service to more than 10 million users by the year 2000. This paper indicates the feasibility of a new satellite system at EHF (40/50 GHz) to complement at the end of the decade the GSM system or its descendants, in order to provide additional services at 64 kbit/s, or so. The main system aspects, channel characteristics, technology issues, and both on-board and earth terminal architectures are highlighted in the paper. Based on the performed analyses, a proposal has been addressed to the Italian Space Agency (ASI) and the process is advanced aimed to the implementation of a national research plan.

INTRODUCTION

The pan-European L-band terrestrial cellular system (GSM) will provide mobile users with digital voice and low bit rate data services starting from early 90's. According to market forecasts, at the end of the century the GSM system will serve a population between ten and twenty millions terminals all over Western Europe. However, up to now this system is not intended to transparently provide the mobile user with 2B+D ISDN channels. Even a small demand of ISDN services (few percent of the overall GSM traffic) can motivate the deployment of a new satellite system to be integrated with the GSM system. Main objectives of such a new system are: (1) to provide higher data rates (64 kbit/s and possibly n x 64 kbit/s), for both high speed data and facsimile, video services and high quality voice, (2) to widen coverage to Eastern Europe, to the Middle East Countries and to Northern Africa, also filling possible temporary holes of the GSM system coverage, and (3) to better cope with the networking needs of private users in small regional areas.

Being the L-band very congested, it is mandatory to exploit higher frequency bands. Our feasibility study identified the EHF (40/50 GHz) frequency band as the most suited candidate in a ten year perspective. Putting into account technological trend data for RF components at EHF, some solutions have been proposed for a new mobile satellite system. This system utilizes Highly Inclined Orbits (HIO) with the apogee on the zenith of central Europe to minimize the probability of link obstruction.

The present paper highlights problems and possible solutions, both at system and subsystem level, that characterize the proposed satellite system. On the basis of the feasibility study, a proposal [1] was addressed to the Italian Space Agency (ASI) and the process is advanced aimed to the implementation of a national research plan.

SYSTEM ASSESSMENT

EHF Channel Model

In mobile systems the acceptable link unavailability generally ranges from 1 to 5 percent. The unavailability time can be

estimated through the evaluation of the blockage probability due to obstacles (buildings, plants, etc.), the probability of severe multipath conditions, and the excess attenuation probability due to atmospheric effects.

Some experiments on blockage at 1.5 GHz showed that the unavailability time at 43° elevation is about 0.2 % on motorways and more than 50% in an urban environment [2]. Therefore, on the one hand the unavailability figure of a HIO system at EHF can be almost fully allotted to the atmospheric effects in rural areas. On the other hand, in metropolitan areas an adequate grade of service can only be provided under particular circumstances.

The 40/50 GHz mobile radio channel mainly suffers from rain and snow effects, clouds and fog, water vapour, molecular absorption and tropospheric turbulences. With respect to these channel impairments the 1 to 5 % unavailability range still lacks deep investigations. Therefore, to the purpose of the present preliminary analysis, some measured data, collected in the frame of the Sirio and the OTS satellite programmes, have been elaborated and extrapolated in frequency (Tab. 1) [3]. Acceptable performance is expected even at 50 GHz with reasonable power margins.

	SITE / UNAVAILABILITY % / ATTENUATION (dB)	SPINO D'ADDA		FUCINO		LARIO	
		0.1	1	0.1	1	0.1	1
40 GHz	RAIN	20	4.3	15.9	6.4	23.6	4.4
	CLOUDS	0.6	0.4	0.4	0.3	0.6	0.4
	WATER VAPOUR	0.6	0.6	0.4	0.4	0.6	0.6
	OXIGEN	0.5	0.5	0.4	0.4	0.5	0.5
	SCINTILLATION	0.4	0.4	0.3	0.3	0.4	0.4
	TOTAL	22.1	6.2	17.4	7.8	25.7	6.3
50 GHz	RAIN	26.6	5.4	21.1	8.6	31.4	9
	CLOUDS	1.1	0.8	0.9	0.6	1.1	0.8
	WATER VAPOUR	0.9	0.9	0.6	0.6	0.8	0.8
	OXIGEN	4.5	4.5	3.2	3.2	4.4	4.4
	SCINTILLATION	0.5	0.5	0.4	0.4	0.5	0.5
	TOTAL	33.6	12.1	26.2	13.4	38.2	15.5

Tab. 1: Estimated values of excess attenuation at 40 GHz and 50 GHz.

The use of highly directive earth terminal antennas that is compatible with vehicular applications at millimetre waves strongly reduces the multipath effects. Therefore, in the limit the channel can be c onsidered either AWGN or fully obstructed by ostacles. An exception to this will presumably be the case of personal communications where the channel turns out to be again a Ricean channel.

The preliminary link design indicates the following achievable performance [1]: (1) system availability between 95% and 99 % ; (2) BER $\leq 10^{-6}$; (3) capacity of about 1000 channels with less than 1kW class satellites.

HIO Implications

The joint adoption of satellites located on non-geostationary HIOs and of millimetre waves presents several implications at system level to be carefully evaluated. Among them the following two are here considered: zooming effect and Doppler effect.

During the activity period of a satellite (between two successive satellite hand-overs) the on-board antenna footprint dimension continually varies (zooming effect). In case of Continental coverage with multibeam antennas, pointing of each single beam must be independently controlled by electronic means.

The zooming effect can be partially compensated by providing the on-board antenna with a phased array feed subsystem, so to keep constant beam boresight directions. Therefore, individual beam coverage zones have variable dimensions, being minimum at hand-over and maximum at apogee.

In order to totally compensate the zooming effect, the beam radiation function must be controlled. This can be achieved either by time varying the beam gain so that it is maintained constant at the edge of beam coverage, or by keeping constant the power density at edge of coverage.

In the first case the on-board antenna main reflector is normally over-dimensioned, and is used with maximum illumination efficiency at hand-over. In the second case the power density is constant at edge of coverage: this can be achieved by statically shaping the beam to put into account the variation of the free space loss between the hand-over point and the apogee.

The Doppler effect is primarily due to the satellite movement and secondarily to the movement of the vehicle. The contribution due to the vehicle is about 3 kHz at 50 GHz for an elevation angle of 50° and vehicle velocity of 100 km/h.

By assuming the vehicle located in a given site, the satellite contribution can be split into a first component in the fixed link and a second component in the mobile link. Both these components can be evaluated as a function of the frequency and of the orbital parameters. At 50 GHz the maximum Doppler shift ranges from about 90 kHz to about 400 kHz, depending on the choice of orbital parameters. Due to the above extremely high values of Doppler shifts, some measures have to be taken.

In case of a transparent satellite, the fixed earth station can provide a pre-compensation in the forward link. To avoid frequency compensation on the mobile, in the return link the correction of the up-link and the compensation for the down-link frequency are performed on-board.

A possible solution is to provide the Doppler compensation with respect to the centre of the beam, thus accepting a residual Doppler that increases towards the beam edge. The maximum value of the residual at 50 GHz ranges from about 3.5 kHz to about 6.5 kHz, depending on orbit selection (case of 500 km beam diameter).

SYSTEM DESIGN

On-board Architecture

A possible payload configuration is shown in Fig.1. This configuration assumes frequency division coverage of the mobile link, a transparent transponder on-board and the use of the 20/30 GHz frequency band in the fixed link.

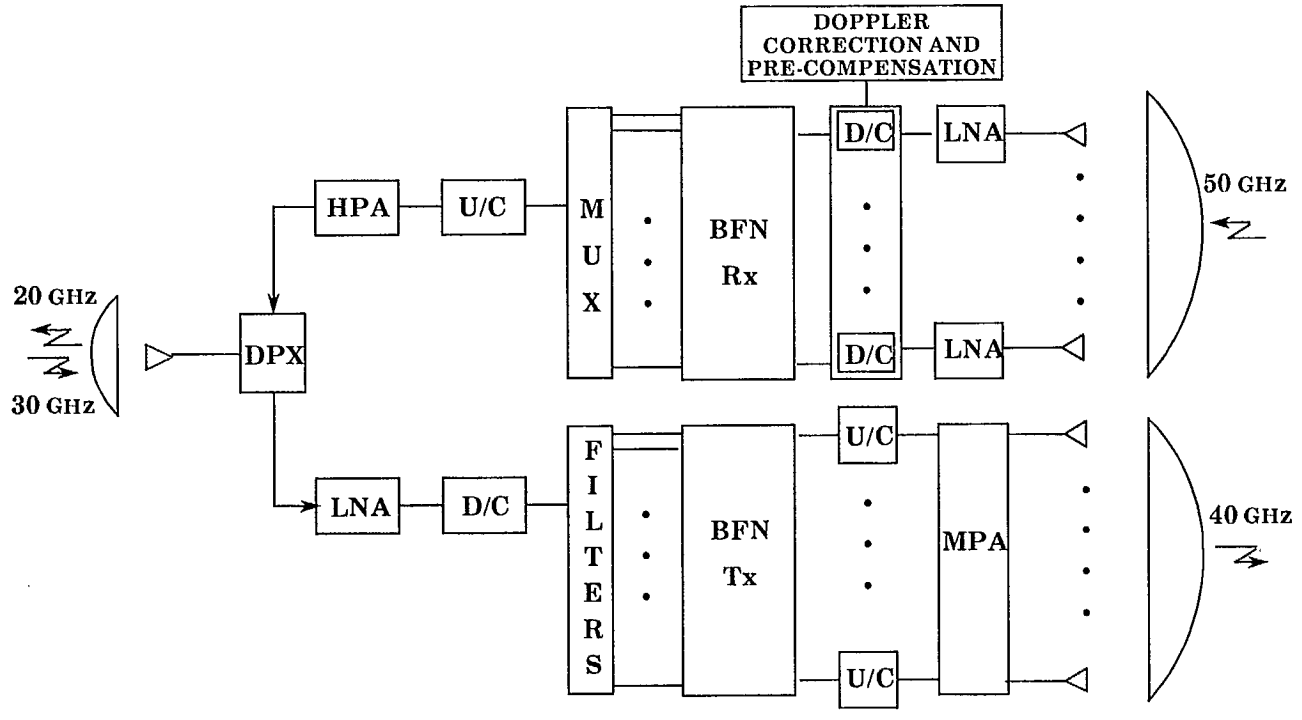


Fig.1: Payload configuration.

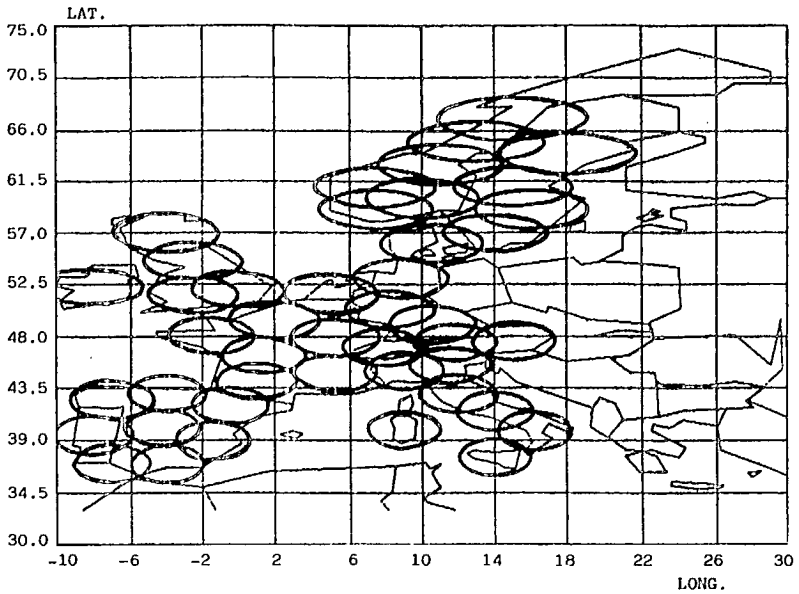


Fig.2: Mobile-links multibeam coverage of Europe.

To achieve the required multispot coverage of the mobile link, a suitable on-board antenna configuration consists of a reflector and a phased array feed subsystem located in the focal region. A single antenna could be used for both transmission and reception, however, due to the antenna limited size (diameter: about 0.5m), separate reflectors for the mobile up-link and down-link are assumed. The European coverage at 40/50 GHz is achieved with about 40 spots (Fig.2), approximately 500 km diameter each

The 40 GHz high power section is configured as a Multi-Port Amplifier (MPA) [4]. This configuration seems particularly advantageous in the proposed system thanks to: (a) efficient use of payload power in the presence of variable traffic loads; (b) graceful degradation of performance in case of failures of amplifying modules. Furthermore, when a set of low-power amplifiers with adjustable gain is adopted together with the MPA, the additional capability of a dynamic spot power reallocation results.

Most of the challenging aspects of the proposed payload architecture are related to the multibeam antenna subsystems and to the high power section. In particular, a MPA is envisaged with a number of ports equal to the number of radiating elements, adopting 40 GHz SSPA modules.

Link budgets show the need for 150 mW per carrier for an optimized selection of orbital parameters. This figure seems compatible with near-term technological outcomes.

Mobile Terminal

A focal aspect of the proposed system is the need for a high gain earth terminal antenna equipped with automatic tracking. Among possible solutions [5], active direct-radiating phased arrays with electronic steering seem to be the most promising in the envisaged program time scale. Two separate square array antennas to be mounted on the same panel have been considered, one for the reception (900 radiating elements), Fig.3, and one for transmission (400 elements). Antenna gains of about 37 dBi and of about 33 dBi result for reception and for transmission, respectively.

An overall radiated power of about 6 W is obtained by no more than 15 mW per radiating element. This figure is compatible with near-term power capabilities of monolithic amplifiers with FET devices at 50 GHz.

A low noise receiver of 500 K noise temperature is suitable for link budget needs. This figure is expected to be achieved at 40 GHz in the near future from HEMT monolithic amplifiers.

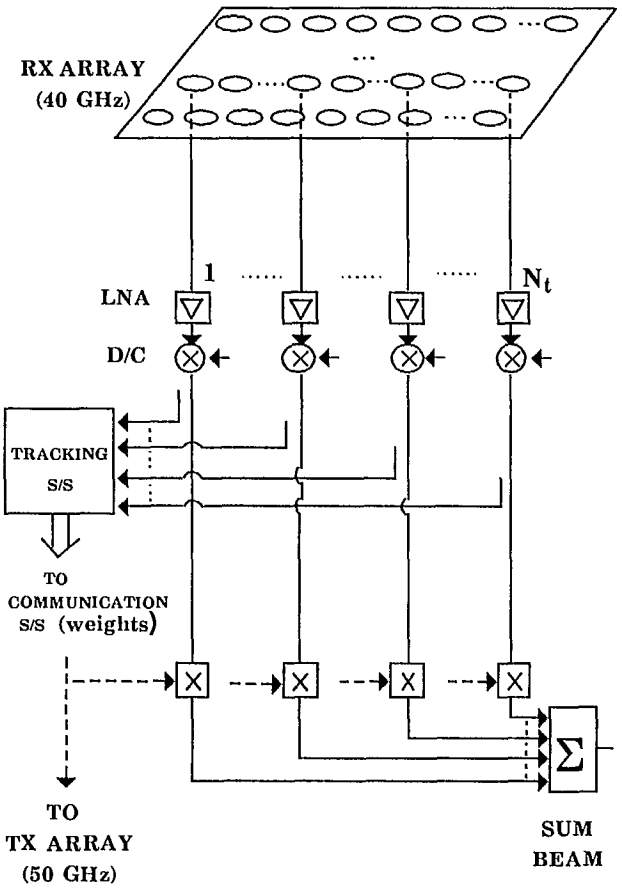


Fig. 3: Schematic of the mobile terminal receive antenna subsystem.

CONCLUSIONS

The satellite system here illustrated can be reasonably deployed within the end of the decade. Early technology development, experiments on channel characteristics and system definition can lead to a system capable to succesfully complement services provided by existing and near future terrestrial cellular systems.

REFERENCES

1. **G. Falciasecca, A. Paraboni, M. Ruggieri, F.Valdoni, F. Vatalaro:** "Proposta di Missione Tecnologica-Sperimentale in Vista di Futuri Servizi Radiomobili Via Satellite a Copertura Europea in Gamma EHF", Universita' di Roma "Tor Vergata", Int. Report, Nov. 1989.
2. **E.Lutz:** "Land Mobile Satellite Channel - Recording and Modelling", 4th Int. Conf. on Sat. Systems for Mobile Comms and Nav., IEE-294, 1988, pp.15-19.
3. **A.Paraboni:** "Propagation Problems Regarding Millimeter Waves (40/50 GHz)", Int. Report, Politecnico di Milano, N.88-023, 1988.
4. **S.Egami, M.Kawai:** "An Adaptive Multiple Beam System Concept", IEEE J. on Sel. Areas in Communications", vol. SAC-5, N.4, May 1987, pp.630-636.
5. **M.Shimada et al.:** "A Millimeterwave Mobile Satellite Communication System for Personal Use", 4th Int. Conf. on Sat. Systems for Mobile Comms and Nav., IEE-294, 1988, pp.52-56.

A New Digital Land Mobile Satellite System

Philip Schneider
Geostar Messaging Corporation
1001 22nd Street, NW, Suite 550
Washington, D.C. 20037 USA
Phone: (202) 887-0872
FAX: (202) 223-6155

ABSTRACT

Pressures on the limited spectrum currently available for mobile satellite communications, as well as the progress being made in other areas of advanced telecommunications networks, such as ISDN, provide the driving force for fully digital mobile satellite systems to serve future markets. For example, INMARSAT is in the process of converting its long-established analog Standard-A service into its digital Standard-B service to increase system capacity and reduce the cost of service to users. Other new INMARSAT services, including Standard-C and Standard-M services, as well as aeronautical satellite services, will be digital in nature.

The same is true for coming national services. Although some satellite systems will retain both analog and digital services, such as the one proposed by the American Mobile Satellite Corporation (AMSC), others such as the digital land mobile satellite service (DLMSS), proposed by the Geostar Messaging Corporation (GMC), are dedicated to the use of only digital transmission techniques.

This paper begins with a description of the different digital services planned to be carried over existing and planned mobile satellite systems, and compares them with analog services in terms of bandwidth and power efficiency.

This comparison provides the rationale for the establishment of a DLMSS service to utilize frequencies that are currently available but not yet assigned to a domestic mobile satellite system in the United States. In response to a petition filed by GMC, the FCC has issued a Notice of Proposed Rulemaking (NPRM) for the allocation of additional frequencies for mobile satellite services. GMC currently has an application pending at the FCC for the use of those frequencies for digital mobile satellite communications, and this paper reviews the status of that DLMSS proceeding.

The paper focuses on the expected advantages of digital transmission techniques in accommodating additional mobile satellite systems in this portion of the spectrum, and how such techniques can fully satisfy voice, data and facsimile mobile communications requirements in a cost-effective manner. It also provides a detailed description of the system architecture of the DLMSS service proposed by GMC, and the market potential which is intended to be addressed by DLMSS.

The final element of this paper will be a discussion of the phased introduction of new mobile satellite services, such as DLMSS. Since future mobile satellite systems are likely to be developed by start-up companies who lack the financial backing of large, entrenched carriers, long-term strategies are needed to guide the development and evolution of such future DLMSS systems by entrepreneurial companies.

INTRODUCTION

The Geostar Messaging Corporation (GMC), a wholly owned subsidiary of the Geostar Corporation, was created to serve the public with a full range of two-way communications services, including facsimile, voice, and full-scale electronic mail.

Geostar currently provides satellite-based positioning, supplemental messaging, and sensor monitoring to a host of commercial and governmental customers through its subsidiary, Geostar Positioning Corporation. Mobile users and operations centers use the patented GEOSTAR® System to determine vehicle

positions and to communicate with them quickly and cost-effectively. The heart of the Geostar System is the timely coordination of mobile transceivers using proprietary satellite transponders with the facilities at Geostar Central. Owners of truck fleets in particular have discovered that this service vastly improves efficiency. Mobile assets can be more effectively managed when headquarters has a real-time way to track their positions, receive status reports, and direct their movements.

THE MOVE TOWARD SPECTRUM EFFICIENT SYSTEMS

The limited spectrum available for mobile satellite communications and the development of low cost digital technology for use in communications systems have driven the industry towards all-digital mobile satellite systems. Voice compression devices, initially developed to permit secure voice transmission for government users, are now available to provide high quality speech channels at 9600 bits per second and good quality at 4800 bits per second. Light weight, portable facsimile machines and lap-top computers are now available for use in mobile environments. Terrestrial digital switching and transmission facilities exist everywhere. The next generation of cellular systems will be digital rather than analog, resulting in considerable investment in the development of speech compression, processing and transmission devices that will become available for other uses.

INMARSAT is in the process of converting its long-established Standard-A service into its digital Standard-B service. They will also be offering two other all-digital services, Standard-C, 600 bit per second, and Standard-M, 4800 bit per second.

The initial mobile satellite systems were based on analog FM transmission. The INMARSAT System's Standard-A terminals are in full use for voice and data transmission. Initially, domestic mobile satellite systems considered use of ACSSB transmission to reduce the bandwidth needed for a voice channel, but advances in voice compression technology are encouraging the development of all-digital mobile satellite systems.

In the INMARSAT system, the bandwidth required for a maritime telephone channel has decreased from 30 kHz (50 kHz channel spacing) for its Standard-A service to 8 kHz (10 kHz channel spacing) for its Standard-M service. In its initial application to the Federal Communications Commission, AMSC proposed both ACSSB and digital transmission schemes. AMSC's digital transmissions ranged from a 2400 bit per second emergency voice service using 5 kHz channels up to 16 kilobit per second toll quality service using 20 kHz channels. The latter service appears to be somewhat similar to the INMARSAT Standard-B service.

Although GMC's system will have as much flexibility as those of INMARSAT and AMSC to provide a wide variety of services to its users, GMC has focused its initial system design on providing a 4800 bit per second voice and facsimile service, and a 1200 bit per second data service. These two data rates can accommodate most of the anticipated customer applications of MSS, and can easily be expanded to meet other customer needs. Moreover, such channels can be accommodated within 5 and 7.5 kHz channel spacings.

With GMC's newly proposed Digital Land Mobile Satellite Service (DLMSS), mobile users anywhere within the United States will be able to use GMC's system to communicate directly with other mobile units, and to access newly emerging digital voice and data networks, including facsimile, through simple mobile or portable terminals.

PROPOSED SATELLITE SYSTEM

Geostar's digital mobile satellite communications system will use on-board signal processing and switching techniques. Two high-capacity, dedicated, GMC digital satellites will provide high-quality, reliable digital communications services to mobile stations throughout the United States. Using eight spot beams covering all fifty states, these DLMSS satellites will make maximum use of available spectrum.

In addition to these dedicated, high-capacity satellites, GMC will also construct a lower

capacity, interim DLMSS system carried on board another satellite or on a "Smallsat" to provide an early demonstration of this technology and to provide an early market entry. GMC is currently evaluating various launching alternatives. Interim service may also be provided using leased INMARSAT capacity.

Radio Frequency Plan

GMC's digital mobile satellite system architecture will consist of the following basic elements:

- a space segment operating in the 1.5/1.6/20/30/ GHz bands;
- a master central earth station;
- mobile and portable earth stations directly accessing the space segment;
- multiple ground stations throughout the United States where the satellite system interconnects with terrestrial public data and voice networks;
- a central telemetry, tracking and command earth station.

DLMSS mobile earth stations will use the 1530 - 1544 MHz and 1626.5 - 1645.5 MHz bands (i.e. L-Band frequencies). Control and links from the master control earth station will be implemented using K-Band feeder links (See Figure 1). These frequencies are particularly well-suited for digital land mobile communications from a propagation point of view. In addition, these bands have been allocated internationally in part for digital mobile satellite communications, which make them ideal for the development of this type of satellite system. The on board processing of DLMSS satellites will allow one-hop connections between mobile terminals, as well as flexible and widely distributed access to terrestrial networks.

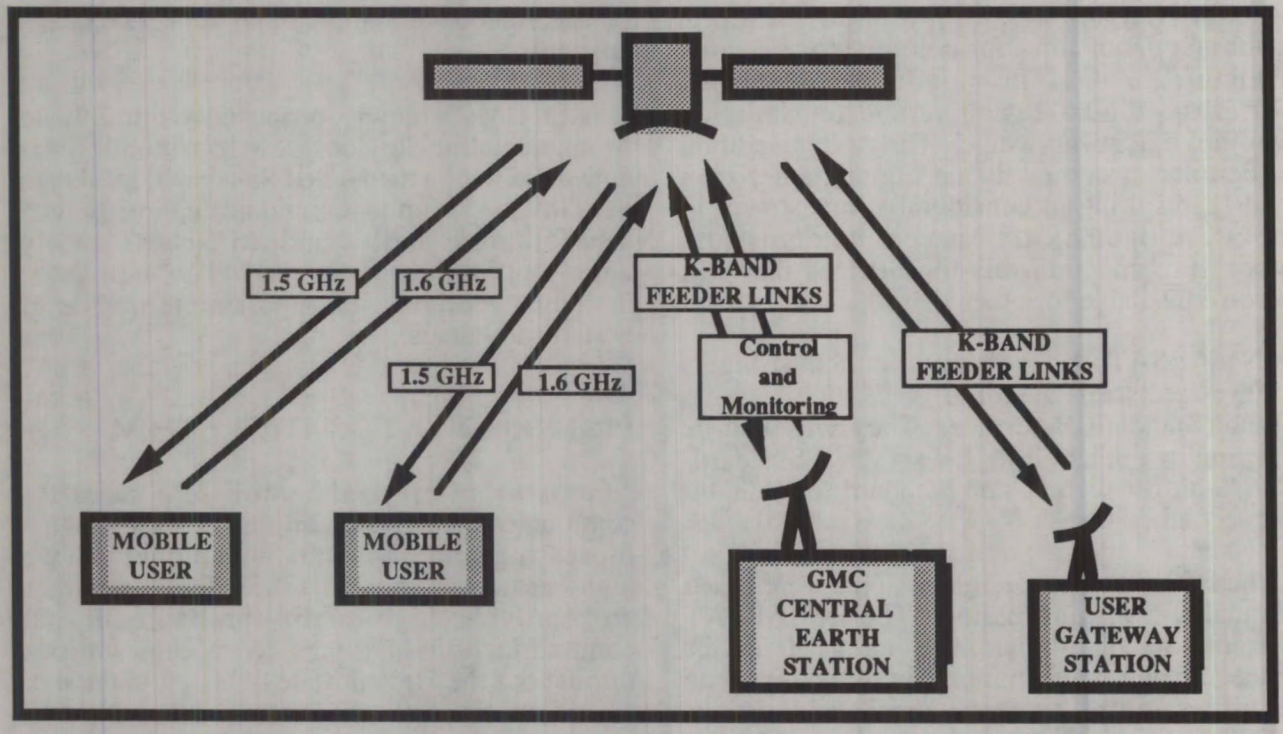


Figure 1. GMC Digital Land Mobile Satellite System

Coverage Areas

One of the principal features of the GMC digital LMSS system will be the use of proven spot beam technology. The system design currently being considered for the digital LMSS system will achieve effective spectrum and orbit utilization by using eight spot beams to cover all 50 states.

The service area will include the contiguous 48 states, Alaska, Hawaii as well as inland waterways, the Great Lakes, and coastal waters adjacent to the United States. It is expected that antenna optimization can be accomplished in such a way as to provide coverage of Puerto Rico and the United States Virgin Islands.

Within this basic service area, GMC will optimize beam parameters and pointing to best match projected traffic distribution within a given area.

System Services

There is a growing demand for a mobile communications system to provide digital communications. GMC will employ user units for portable, in-vehicle and fixed applications, and will provide initial transmission rates between 1200 and 4800 bits per second in order to meet this demand. DLMSS will provide facsimile and communications between personal computers, as well as a digital voice messaging capability. Initial units will be portable or vehicle mounted, and eventually GMC expects to be able to match the miniaturization of current handheld cellular telephone sets in its DLMSS user terminals.

Several general categories of services will be offered initially over the Geostar DLMSS system. Generally, it is expected that all of these services can be provided over any of the DLMSS user terminals.

Digital Data Transmissions. The DLMSS system will provide two-way data communications services, either on a packet switched or circuit switched basis.

Facsimile. Mobile facsimile will be a significant new capability that will be offered over the Geostar DLMSS satellite system. A mobile facsimile unit can be installed in a car or truck, and transmissions sent to and from that unit, while it is on the road. Such use is expected to add new dimensions to mobile communications. For example, bills of lading can be automatically transmitted to truckers while en route. Documents prepared in the field can be faxed to office headquarters to initiate order servicing. Maps and other graphic materials, such as updated architectural and engineering diagrams, can be easily sent to field workers.

Digital Voice. Although Geostar's mobile satellite system will be optimized for data transmission, recent developments have made it technically and economically possible to encode voice communications into a digital transmission format at transmission rates of 4800 bits per second or less. Geostar is convinced that the signal quality of its digital voice service will be suitable for user applications in the mobile environment. Moreover, as technology advances, the quality of such compressed digital voice transmissions will improve.

Digital Paging and Dispatch. Geostar expects that it will be able to develop handheld mobile units in conjunction with DLMSS equipment manufacturers. This will permit a much greater capability for paging and dispatch services than those currently available from terrestrial paging companies. Such a service offered over a DLMSS satellite would be intrinsically nationwide in nature. While the size of DLMSS paging terminals will be larger than conventional tone and alphanumeric pagers, their two-way transmission capability for acknowledgement and response, together with their high data transmission capability, will make them attractive to many paging customers.

Remote Data Collection and Monitoring. The Geostar DLMSS will support transportable or fixed data terminals for data collection from remote sites. The units will be built to withstand harsh environmental conditions and will be designed for long periods of unattended operation. This service will be particularly useful in security and environmental data collection applications.

Data and Voice Message Store and Forwarding. In many cases, real time two-way or circuit switched communications are not

required, or the called party's line may be busy. In this case, Geostar's DLMSS system will provide users with the capability to store the message at the central earth station for forwarding at a future time when the connection can be completed.

Emergency Data and Voice Communications. DLMSS provides a flexible, efficient means of providing immediate communications for emergency teams operating in areas not served effectively by terrestrial means. Particular situations requiring such communications capabilities include forest fire fighting, medical team operations in remote areas, and disaster response teams. Such emergency situations require both data transmission capabilities, such as for transmission of medical or other telemetered data, and voice communications for effective consultation between the on-site and headquarters personnel responsible for responding to the emergency.

Potential Markets

Automotive. Automobiles are becoming more and more electronically sophisticated. Soon, automobiles will be fitted with complex electronics that will perform increasingly complex tasks. On-board navigation systems have already been announced, and personal computers are likely to be installed in the near future. The Geostar DLMSS system will complement these developments by adding full, two-way digital communications capabilities to the increasingly complex array of electronics options available for installation in automobiles.

The trucking and railroad industries have already begun to install advanced electronics systems for surveillance and control of their operations. The Geostar DLMSS system will further enhance these capabilities.

Lap-top personal computers. As the employment of personal computers for professional and personal use increases, there is a growing need for the ability to access computer resources (i.e. databases and communications networks) when away from the home or office. An integrated lap-top computer/DLMSS terminal will solve this problem. Salesmen in the field will be able to directly enter sales orders into headquarters computer data bases, and authors can directly transmit manuscript revisions to their agents from retreats in the country. GMC is convinced that the flexibility and low cost of DLMSS will greatly extend access to personal computer resources for professional and individual use.

Maritime. GMC anticipates there will be a growing need for digital communications for both commercial and pleasure vessels on inland waterways, on the Great Lakes, and in coastal waters. Maritime operations are becoming increasingly dependent on sophisticated electronics systems. Moreover, there is an increasing need to be able to transmit significant amounts of data in the maritime environment. Such capabilities will be particularly useful for transmitting the most up-to-date oceanographic and weather charts to highlight changes in navigational conditions, for relay of oceanographic and marine research data, and to ensure efficient commercial maritime operations.

Rural/remote data. The Geostar DLMSS system will provide an economical means for data communications in rural and remote areas. These areas generally lack an effective infrastructure of terrestrial communications facilities.

DLMSS will readily provide high-quality, reliable data and voice capabilities for all of the rural and remote areas of the country, including Alaska. In addition to mobile and transportable applications, DLMSS is ideal to provide a reliable, economical means of two-way digital communications to rugged terminals designed to operate for long periods of unattended operations. Applications include monitoring and control of remote, unmanned facilities and equipment, such as transcontinental pipelines, and collection of environmental, meteorological and hydrographic data from remote sites. A particular value of DLMSS would be the two-way transmission capability which also allows security alarms to be transmitted to the central monitoring site in the event of intrusion or equipment malfunction, and control of interrogation signals to be transmitted to the remote, unattended site.

FEDERAL COMMUNICATIONS COMMISSION

In response to a petition filed by Geostar, the FCC issued a Notice of Proposed Rulemaking (NPRM) on March 5 to allocate additional frequencies for mobile satellite service. Geostar currently has an application before the FCC for a license to construct and operate a DLMSS system using those frequencies. GMC believes that it is the Commission's policy to spur the rapid introduction of new technologies and services that encourage development of competitive, efficient telecommunications services markets. The underlying rationale for the Commission's actions has been the desire to have the American consumer enjoy the full benefits of competitive markets in the choice, quality, and price of telecommunications services, and rely on such markets as a mechanism to support the public interest goals of the Communications Act.

PHASE IN TO FULL SERVICE

Geostar Messaging Corporation, with its DLMSS system, will combine with the RDSS service of Geostar to make the company a full service satellite communications company capable of providing ocean to ocean "seamless" voice, data and position reporting services that no other single entity can match. Unlike terrestrial systems which can be started in local areas with modest investments, a satellite system takes one large investment at the front end to get started and many years to recover the investment. Geostar has therefore developed a phase-in plan to allow the revenues to build as increasingly larger investments are made.

It is important for GMC to get into business by early 1992 to take advantage of the "window of opportunity" that will exist during the next five to seven years because:

1) No other U.S. satellite system is now in operation providing voice, facsimile and data services.

2) Cellular radio coverage now exists only in the MSA's, i.e., the large metropolitan areas. Most of the country is not covered and will not be for at least five years and possibly ten.

3) Cordless telephone CT-2, and Personal Communications Networks (hand-held cellular) have not yet arrived in the U.S.

4) Facsimile, imaging and digital technology are becoming the accepted method of managing decentralized resources.

The GMC all-digital full voice, facsimile and data two-way communications program is timed to start service in 1992. This can be accomplished by leasing capacity from INMARSAT. GMC will then build a "smallsat" dedicated satellite that can provide an order of magnitude increase in capacity. For the long term a large multi-beam satellite will be required to handle the traffic and provide adequate power. GMC would require that additional capacity in the middle '90s.

Transitioning from Analog to Digital Communications: An Information Security Perspective

By Richard A. Dean
Department of Defense,
Fort Meade, Md 20755

ABSTRACT

This is an age of revolution in communications technology and systems. While this revolution might simply be described as a transition to digital communications technologies, these fragmented (post Bell System), market driven developments create a totally new environment to plan, influence and develop systems for the U.S. Government's secure voice users.

Secure communications depend on the availability of either end-to-end analog connections or end-to-end digital connections. The uncontrolled mixture of analog and digital links in the public switched network (and its extensions), as will happen in this digital transition, can inhibit end-to-end encryption. The introduction of security into digital systems such as Mobile Satellite, Digital Cellular Telephone, Integrated Services Digital Network (ISDN), Land Mobile Radio and other systems represents both a challenge to our existing infrastructure and an opportunity to capitalize on new technology.

This paper is the author's attempt to summarize the governments perspective on evolving digital communications as they affect secure voice users and approaches for operating during a transition period to an all digital world.

NEW DIGITAL COMMUNICATIONS

Virtually every component of the communications infrastructure will be influenced by digital technology. The critical systems affected by this transition are summarized below.

Integrated Services Digital Network (ISDN) represents the largest and most visible digital service on the horizon. ISDN will offer many new features for data and secure voice users on the Public Switched Network (PSN). The menu of services will allow backward compatibility with analog POTS service for data modems, facsimile, and secure voice users. In fact the flexibility of ISDN will facilitate some of the connection problems for other networks as described later.

Mobile communications which includes Cellular, Satellite, Land Mobile Radio, and portable phones represents today's fastest growing communications market [1]. The growth of this market and the pressure for bandwidth conservation has led to the introduction of low rate compressed voice in the range of 4-13 kbps. Because these low rate digital voice links will not support a conventional modem, and the networks do not carry the digital signal to the far end, this voice compression will challenge our ability to secure

these links. Mobile communications also represents a potential threat to sensitive U.S. communications as these radio communications will be readily available in the clear. Backward compatibility to the government's STU-III secure telephone and other secure systems will likewise be a challenge. Several systems are in various stages of development including Digital Cellular telephone, Land Mobile Radio, Mobile Satellite, and INMARSAT. Each of these systems has the potential of becoming a valuable secure communications link but each will require inclusion of features to enable secure interoperability.

INTEGRATED ARCHITECTURE

The recent explosion of new digital systems has forced a more comprehensive look at at interoperability on these systems. While each system has a unique protocol set, developing custom solutions for each new network is neither wise nor feasible. It is clear that the challenge is to conceive a plan that will enable transparent, communications among the dominant future digital systems. The interoperability challenge is to facilitate a smooth transition by maintaining backward compatibility with existing secure analog equipments. These existing and future systems consist of STU-III on existing analog POTS, Digital Cellular Telephone, Mobile Satellite, and ISDN (with a potential new terminal for that media). Managing the transition from a STU-III based (analog end-to-end) system to an ISDN based system with a large mobile user population represents a major challenge. One must either make all new terminals backward compatible at some expense, replace existing secure voice equipment with new equipment, or face the problem of building high volume gateways for interoperability. In this paper it was decided by the author to try to move forward to embrace new digital techniques for future compatibility rather attempt to maintain continued analog operation. While a centralized Gateway solution is presented, various decentralized solutions will be practical and desirable to implement in specific instances. A family of compatible solutions are in fact possible in concept and such a set of solutions might be necessary to resolve the anticipated bulge that will appear in the 1994-98 timeframe when we are in the midst of the transition from analog to digital. Key to the success of such an architecture, however, is to develop a common interface with mixed digital systems that will allow transparent end to end digital operation.

In looking at alternatives for a global connection architecture, there is an option to resolve the interface problem with a distributed solution or with a centralized solution. A distributed solution deals with the interface problem by a translation for compatibility at each end terminal or as close to that terminal as possible. A centralized solution performs the translation at one central location. Both approaches have advantages. The approach presented here shows an evolution from a distributed solution to a centralized solution using a Gateway.

The architecture is presented in three phases consistent with developments in the associated usage and availability of the

Cellular, Mobile and ISDN services.

The architecture for Phase 1 is shown in Figure 1. This is a distributed solution using STU-III modem pools (shown as M-3) to perform analog to digital transition for Mobile Satellite and Cellular. The STU-IIIB uses a direct (black) digital output for mobile applications. It is shown for the 1990-93 time frame where the extent and geographic distribution of this solution is limited. This approach maintains the digital nature of the mobile network and its interface is the same as most other digital users.

An architecture for Phase 2 is shown in Figure 2 for the timeframe 1994-1998. This represents the peak of the analog to digital transition where it is expected that there will be large numbers of both digital and analog secure systems connected. In this case the Gateway is the primary place where the mobile systems transition to analog STU-III terminals. This is performed by establishing a digital connection (switched or dedicated) from the Mobile and Cellular interfaces to the Gateway. The M-3 modems shown in figure 1 are relocated at the Gateway. The digital interface shown as G is a simple (assumed standard) digital interface to the Mobile and Cellular interfaces. Figure 2 also shows an early ISDN network. The early ISDN terminals are assumed to be STU-III interoperable (direct or virtual) and include a STU-III modem. This is provided so that the Gateway traffic and resulting blockage is limited. ISDN will also provide a simple and cheap switched connection from the Mobile and Cellular interfaces and the Gateway. Figure 2 also shows a STU-IIIB (a STU-III with BLACK digital output) coupled directly into ISDN with an ISDN adapter. The incorporation of this adapter allows direct digital communication of these STU-III's over ISDN with mobile STU-III's and with ISDN terminals. The use of analog modems in the ISDN terminal and the use of ISDN adapters on STU-III's provide users options to enhance their grade of service and limit traffic through the Digital Gateway.

An architecture for Phase 3 is shown in Figure 3. This is the architecture for the period 1999 and beyond. In this era ISDN is almost universally available. Mobile and Cellular operation is supported by direct end-to-end BLACK connections directly to ISDN terminals via a simple (assumed standard) digital interface. The Digital Gateway continues to operate, however, as an interface to the remnant of analog STU-III users. These may be people at remote locations or international service where end-to-end digital is unavailable. The M-3 modems are removed from the ISDN terminals and can be used at the Gateway.

The centralized solution has several inherent advantages. This solution lumps most of the interoperability problems into a single solution where it is assumed that economies of scale will reduce overall costs. Our requirements to each of the numerous communications carriers is uniform and standard. We simply want to pull out the embedded 4.8 kbps stream from their system and port it to the Gateway on a standard digital link. We eliminate the need for a custom, government owned M-3 modem pool at geographically and organizationally diverse facilities. We trade off this widely dispersed logistics problem against the

communications costs to connect to a centralized, more efficient facility. Transition to an all digital environment is clear. When direct end-to-end digital connections can be made from the mobile interface to the ISDN user, the central Gateway is bypassed. Finally, when the bulk of the users are connected to ISDN on either an ISDN terminal or a STU-III with a terminal adapter, the gateway can be used to extend service to the remnant of analog STU-III users.

MOBILE SATELLITE INTERFACE

The proposed interoperability scheme requires a common set of features be provided by mobile cellular and satellite systems. A preliminary evaluation of the systems function reduces these to the following:

1. Recognition by the host system of the STU-III capability at the mobile unit.
2. Enable the insertion of an encrypted data stream in place of the 4.8 kbps speech data within the packet structure and removal at the other end (mobile or PSN interface) with transparency.
3. Include the control signaling to allow either the mobile STU-III or the PSN STU-III to switch from clear voice to secure voice and back at will. In the case of the PSN interface, a 2100 Hz tone is the signal used to bring in the STU-III modem. (Phase 1)
4. Support a STU-III modem pool at the PSN entry. (Phase 1)
5. Include the interface and control necessary for direct transfer of the 4.8 kbps data to a remote facility by either a dedicated or switched fractional T1 link or ISDN when available. (Phase 2)
6. Perform all system signaling out of band after the call has been established.
7. Maintain synchronization and bit integrity for the 4.8 kbps stream after the connection is established.
8. Incorporate the Fed. Std. 1016 4.8 kbps CELP voice coder as the standard speech algorithm for both clear and secure.

These functional requirements are judged to be straightforward to implement if identified in the early design phase of the mobile system. These features are also similar to the requirement of many other data a facsimile users who need to interoperate on mobile systems.

CONCLUSIONS

A concept for secure voice interoperability on emerging digital systems was presented. The transition period from analog to digital systems is handled by modem pools and a central Gateway. Long term compatibility is satisfied by an end to end connection via ISDN. The interface requirements for digital systems such as Mobile Satellite were presented and are straightforward and similar in nature to the needs of other data and facsimile users.

REFERENCES

1. M. Bonatti et al., "Telecommunications Network Design and Planning", IEEE Journal on Selected Areas of Communications, Vol 7, Number 8 Oct 1989

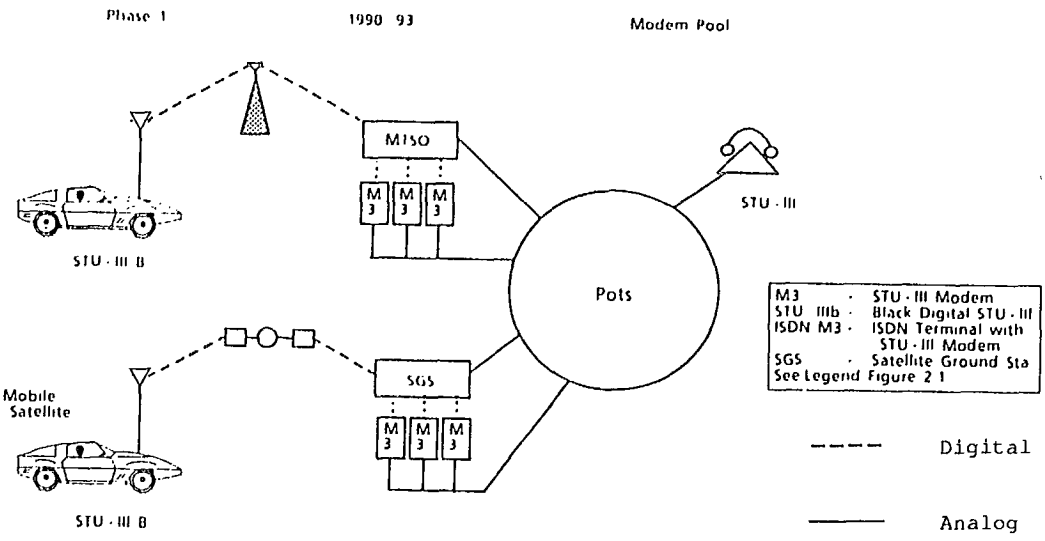


Figure 1 Phase 1, Distributed Solution with Modem Pool

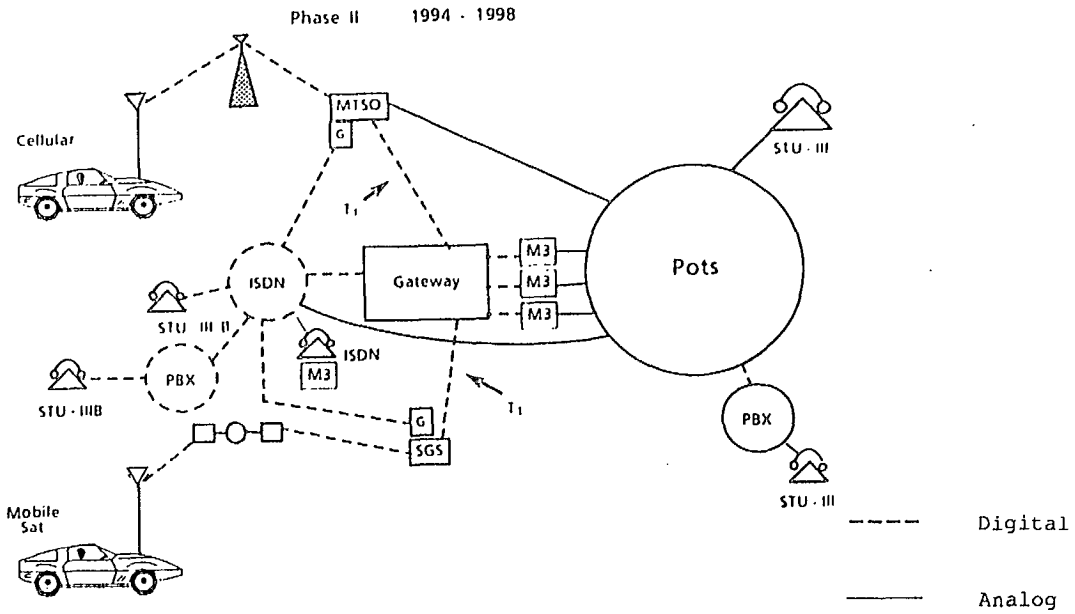


Figure 2 Phase 2, Mixed Analog/Digital Network with Gateway

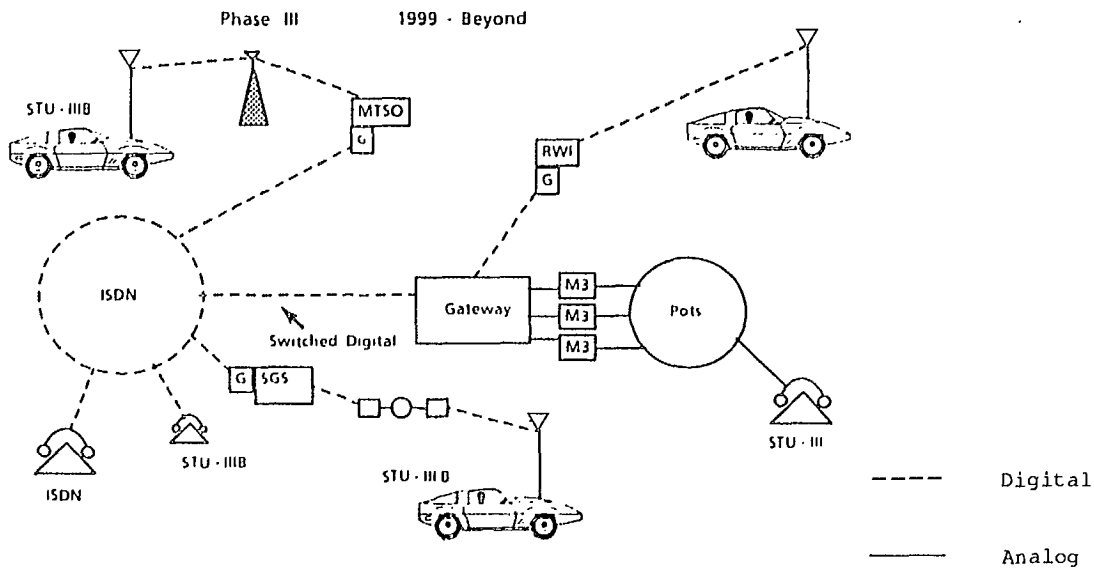


Figure 3 Phase 3, Digital Network with Analog Enclave

Session 4
Aeronautical Applications - I

Session Chairman - *J. Stuart Muirhead,* Air Canada, Canada
Session Organizer - *Jack Rigley,* DOC

4

Aeronautical Satellite System Test and Implementation

Keith Smith
Inmarsat - Aeronautical Services Division
40 Melton Street, Euston Square
London, England NW1 2EQ
United Kingdom
Tel: 44-71-387-9089 Fax: 44-71-387-2115

SYNOPSIS

This contribution reports on various aspects of the test and implementation of the new worldwide aeronautical satellite system. The system will be implemented in 1990 following several years of system design and equipment development. The full range of services and equipment will not be available immediately, but rather there will be a progressive phase-in over the first two years of service.

Data services using a low-gain aircraft antenna are planned to be introduced first, and tests are currently under way using 'interim' avionics and ground earth station equipment. Service is expected to begin on this interim basis about April/May 1990, and toward the end of 1990 equipment to the full technical standard under development by Inmarsat and the Airlines Electronic Engineering Committee is expected to be available.

Voice service is expected to begin on a full commercial basis about May/June 1990, requiring high-gain aircraft antennas.

In view of the complexity of the aeronautical system, and the fact that the system design is new and in some respects still evolving, it has become clear that the initial releases of airborne and ground equipment may, at a relatively early stage, need to be augmented or even replaced with equipment which meets an updated specification.

This contribution explains Inmarsat's approach to testing and authorization of equipment intended to make use of its satellites in this evolving situation. The approach is designed to permit progressive implementation of services while avoiding the need for large-scale retrofitting of equipment which has already been introduced into service.

The full text of this paper was not available at press time

A High Gain Antenna System for Airborne Satellite Communication Applications

M. Maritan and M. Borgford
Canadian Marconi Company Ltd.
415 Legget Drive
Kanata, Ontario
K2K 2B2, CANADA
Phone: 613-592-6500
FAX: 613-592-7427

ABSTRACT

A high gain antenna for commercial aviation satellite communications is discussed. Electromagnetic and practical design considerations as well as candidate system implementations are presented. An evaluation of these implementation schemes is given, resulting in the selection of a single top-mounted aerodynamic phased array antenna with a remotely located beam steering unit. This concept has been developed into a popular product known as the Canadian Marconi Company CMA-2100. A description of the technical details is followed by a summary of results from the first production antenna.

INTRODUCTION

The commercial aircraft industry has a rapidly growing requirement for an aeronautical satellite communication (SATCOM) system to provide data and voice services to airline fleets. These services will allow safer, more efficient air travel throughout the world.

The SATCOM system is comprised of an Aircraft Earth Station (AES), a Satellite, and a Ground Earth Station (GES). The AES portion of the system for commercial aviation is defined by Aeronautical Radio Incorporated (ARINC) Characteristic 741 and is comprised of a Satellite Data Unit (SDU), A Radio Frequency Unit (RFU), a Diplexer/Low Noise Amplifier (DIP/LNA), a High Power Amplifier (HPA) and a High Gain Antenna (HGA) with its associated Beam Steering Electronics. Refer to Figure 1.

Canadian Marconi Company has developed the CMA-2100; a High Gain Antenna System which conforms to the ARINC industry specification.

INDUSTRY REQUIREMENTS

Due to the complexity of the airborne system, coordination with many parties was necessary to allow successful development of the antenna system. The following identifies the major organizations with which Canadian Marconi Company has interfaced during product development.

Service Providers. Initially, the system will utilize the International Maritime Satellite Organization (INMARSAT) constellation of satellites. The INMARSAT requirements are outlined in their Aeronautical System Definition Manual (SDM).

ARINC. ARINC Characteristic 741 provides guidance on the form factor and pin assignments for the Line Replaceable Units (LRU) that make up the AES.

RTCA. Special Committee SC-165 of the Radio Technical Commission For Aeronautics is currently in the process of defining the Minimum Operational Performance Standards for Aeronautical Mobile Satellite Services.

Airworthiness Agencies. The involvement of the Federal Aviation Administration and Transport Canada is necessary to ensure compliance with industry airworthiness requirements.

Airlines. The input of the ultimate end users was considered to ensure development of a marketable product.

Airframe Manufacturers. The major airframe manufacturers were consulted to ensure structural compatibility of the antenna with the host airframes.

Avionics Manufacturers. The manufacturers of the complement of the AES were consulted to ensure

compatibility of the communication protocols between the LRUs.

PERFORMANCE REQUIREMENTS

The ideal High Gain Antenna coverage volume required by ARINC 741 is defined as 12 dB of circularily polarized gain in the upper hemisphere for all normal aircraft flight attitudes. Electromagnetic characteristics, however, are not the only criteria that must be considered during product development. The CMA-2100 development team considered many other parameters during the design of the antenna system. These are listed below along with the electromagnetic considerations.

Gain. Practical antennas will result in the specified gain value over approximately 75% of the ideal coverage volume. The achieved gain should be aircraft direction independent to allow for routing flexibility.

Multipath Rejection. Flight tests performed to date have shown multipath fading to be a problem[1,2] . An antenna whose gain falls off at the horizon would be ideal to maximize multipath rejection.

Axial Ratio. The axial ratio (AR) specification has been chosen to prevent excessive gain loss due to cross polarization with an elliptically polarized satellite feed antenna.

Sidelobe Level. To provide satellite discrimination in future constellations, sidelobe levels greater than 45° from the satellite direction should be minimized.

Beam Switched Phase Error. The phase error induced across the HGA aperature when switching from one beam to an adjacent beam should be kept to a minimum to satisfy SDU modem requirements and prevent demodulation synchronization loss.

Input Impedance. The input Voltage Standing Wave Ratio (VSWR) at the RF connector of the HGA should be minimized to prevent HPA overloading.

Airworthiness. There are many factors to consider to ensure the antenna conforms to airworthiness standards, however good engineering practices can ensure that most of these requirements are met.

Equipment Cost. Initial cost plays a large part in an airlines decision to outfit their fleet with SATCOM capability. For obvious reasons, this cost should be kept to a minimum.

Installation Cost. Depending on the design approach taken, installation of the antenna system can range from easy to very difficult. To minimize the installation effort and cost, the antenna system should employ simplicity. Ideally a single antenna with flexibility as to where the associated control electronics can be placed would meet this requirement. This would avoid the need for switching relays, microwave splitters/combiners and extra wiring.

Operating Cost. Any increase to the operating cost of the aircraft such as fuel burn penalty should be minimized. By implementing a low profile aerodynamic design, this can be accomplished.

Maintenance. Aircraft downtime is of major importance to airlines. A system that minimizes downtime when equipment replacement is necessary is desired. The antenna should be replaceable with access to only the exterior of the aircraft and the associated control electronics should be rack mounted to allow easy replacement.

Reliability. To prevent aircraft downtime, the antenna system should be reliable. In the event of a failure, the system performance should degrade gracefully. An electronically steerable phased array is ideal for meeting this requirement.

CANDIDATE SYSTEM CONFIGURATIONS

Four different design approaches were considered during the development of the CMA-2100 antenna system.

Mechanically Steered Antenna. This approach consists of a top mounted mechanically steered array of antenna elements enclosed in a large radome. Reliability aspects and fuel burn penalty make this approach unattractive.

Side Mounted Conformal. This antenna type consists of a number of crossed slot or microstrip patch antenna elements configured in a package conforming to the aircraft fuselage. ARINC 741 defines an antenna system which would incorporate two such antennas mounted on either side of the aircraft at 45° from the horizontal. System complexity, poor multipath rejection and aircraft direction dependent gain due to scanning limitations of the patch elements are significant disadvantages of this antenna design.

Multifacetted Top Mounted. This design approach incorporates a number of microstrip patch antenna arrays, each pointing in a different direction. The group

of arrays is housed in a single radome mounted on the top of the aircraft fuselage. The elements used will suffer from the same limited scanning characteristics and multipath problems as the conformal sidemounted antenna and will not maintain the low profile, conformal form factor, resulting in increased drag.

Top Mounted Wide Scanning Antenna. This design approach incorporates an array of elements mounted at a height which is a significant fraction of a wavelength above the aircraft fuselage. This array implementation will allow the antenna beam to be scanned to very low elevation angles allowing aircraft independent gain coverage, very good multipath rejection as well as a mechanical package with very desirable attributes. The fuel burn penalty is kept low by minimizing the radome height above the antenna elements.

CMA-2100 DESIGN APPROACH

Trade-off analysis has shown the top-mounted wide scanning antenna to be the approach needed to develop a marketable SATCOM antenna with good electrical performance that fulfills the industry reliability, maintenance, and cost requirements. Canadian Marconi Company adopted this approach in the development of the CMA-2100 High Gain Antenna System.

The CMA-2100 is configured in two line replaceable units; The High Gain Antenna (HGA) and the Beam Steering Unit (BSU). The HGA is housed in a single mechanical package and mounts on the top of the aircraft fuselage. The BSU is mounted in a standard avionics rack up to 100 ft. from the HGA. Only two cables are required to penetrate through the aircraft skin: one for the BSU/HGA digital communication, and one for the RF cable. Replacement of the HGA is accomplished from the exterior of the aircraft while replacement of the BSU requires access to the equipment rack inside the aircraft.

High Gain Antenna

The HGA is an electronically steerable phased array antenna which provides +12 dBic nominal gain with near hemispherical coverage. The major components of the HGA are the radiating elements, 3-bit phase shifters, 32-way power divider/combiner, antenna interface card, and mechanical package. Refer to Figure 1.

Radiating Elements. The antenna array is comprised of 32 crossed dipole elements configured in a 4 x 8 planar array. Utilization of element feed structures with negligible cross sectional area help to minimize scattering and improve low elevation angle antenna performance. Individual elements are fabricated by intersecting two microwave circuit cards, each containing the printed radiating structure and balanced feed network.

3-Bit Phase Shifters. Antenna beam steering is performed utilizing 32 phase shifters which provide the necessary electrical phase shift to each antenna element. The design is a 3 bit implementation that allows for 45° phase steps. Included in the circuit is an orthogonalization network for circular polarization. Each phase shifter is constructed on a printed microwave circuit card using conventional surface mount components and then bonded to the radiating element to form an antenna element module. The phase shifter is manufactured to high reliability standards with no plated through holes.

32-Way Power Divider\Combiner. This assembly is a bidirectional power divider/combiner circuit that feeds 32 equal phase and amplitude signals to/from the phase shifters. The design is based on a modified Wilkinson coupler power divider which compensates for the array impedances that vary with scan angle.

Antenna Interface. This circuit receives digital words from the BSU and sets the appropriate phase shifter bits, as well as communicates antenna status information back to the BSU. The design incorporates parallel circuitry dedicated to each antenna element resulting in only one element being affected by a failure (a negligible degradation of antenna RF performance). The antenna interface includes built-in-test circuitry to monitor the phase shifter operation, EMI surge protection for the digital and power busses penetrating the aircraft skin, and on board power conditioning of the d.c. power fed from the BSU. One double sided digital/analog printed circuit card with conventional leaded components controls all 32 phase shifters.

Mechanical Package. The antenna electronics are housed in an aerodynamic radome/baseplate integrated enclosure. The radome is constructed of honeycomb sandwich while the baseplate is milled aluminum. The bottom of the HGA is flat and adapts to the curved aircraft fuselage with an airframe dependant adapter plate. The approximate dimensions of the HGA are 67" x 18" x 5".

Beam Steering Unit

The BSU is used to translate antenna beam position data and beam change commands received from the SATCOM Satellite Data Unit (SDU) in a standard digital format into signals needed to select phasing of

the antenna elements. This results in the antenna beam pointing at the desired satellite. The major components of the BSU are the processor card, software, BSU interface card, power supplies and the enclosure. Refer to Figure 1.

Processor Card. This circuit provides interfaces to the SDU as well as the DIP/LNA. The design utilizes surface mount components on a multilayer printed circuit card. The processor is an INTEL 80C86.

Software. The software translates the azimuth and elevation satellite coordinates from the SDU into HGA phase shifter settings. In addition, the software also communicates with the SDU and HGA, and performs self test for the BSU and HGA. The software is written in Pascal and has been developed to RTCA DO-178A flight essential standards.

BSU Interface card. This circuit provides the RS-422 interface to the HGA, monitors the HGA DC power, and provides EMI/surge protection to the power and signal lines that penetrate the aircraft skin. This double sided printed circuit card utilizes both leaded and surface mount components.

Power Supplies. The BSU incorporates a 115 volt 400 Hz in/ 28 volt DC out power supply assembly. It also utilizes two high reliability DC to DC converters to provide the operating voltages. The BSU power supplies provide power to the HGA.

BSU Enclosure. The BSU is enclosed in an avionics industry standard 2 MCU ARINC enclosure utilizing an ARINC 600 connector. The enclosure can be rack mounted in the aircraft equipment bay and requires no forced air cooling.

CMA-2100 TEST RESULTS

Extensive measurement of the CMA-2100 electromagnetic characteristics has been performed utitilizing a fully automated roof top range. All measurements were taken on the Antenna Under Test (AUT), which is defined as the HGA mounted in a manner to simulate service usage on a 7 ft. X 9 ft. curved ground plane. The radius of curvature of the ground plane is 127 in. The measurement coordinate system is illustrated in Figure 2. A brief summary of these results follows.

Figure 3 shows the nominal boresight pattern of the AUT at 1.545 GHz. The antenna is orientated such that the scan occurs in the azimuth=135° plane. The elevation angle for this pattern is set to 90° (boresight). The 12 dBic line is superimposed upon the plot.

Figure 4 illustrates the same pattern cut as Figure 2, but the AUT beam has been electronically steered to 10° elevation. This figure shows that multipath signals arriving to the AUT at -10° elevation will be strongly attenuated by the sharp pattern rolloff beyond the 10° elevation point.

The composite scan, shown in Figure 5, of the antenna performance is a dynamic measurement that illustrates the peak gains and beam switching of the AUT. For this measurement a complete elevation scan was taken in the azimuth=135° plane while the computer utilized open loop steering to monitor the positioner location and updated the AUT beam selection to keep it pointed at the source antenna. The irregular pattern shape illustrates the beam switch points.

Extensive axial ratio measurements were performed over the full frequency range. In general the AR was < 6 dB (typically 3-4 dB) with worst case performance occurring with the antenna scanned towards the horizon. This phenomenon was expected since a semi-infinite ground plane (the aircraft fuselage) cannot support a horizontally polarized wave.

Due to the finite antenna beam width and phase quantization of the phase shifters to 45°, a finite number of beams were selected to cover the antenna coverage volume. When switching between beams, it is a requirement to maintain a minimum phase disruption. Adjacent beam switch point phase error measurements have been taken for all possible beams and over the full frequency range with the aid of automated testing. The phase errors are typically 2°-3° and in over 99% of the cases are below 10°.

REFERENCES

1. **Makita, F. 1988.** Air-ground Satellite Communications Test By Japan. *Proceedings of the 1988 RTCA Assembly*. pg 111.

2. *Mobile Satellite Reports*, January 19, 1990 pg 4.

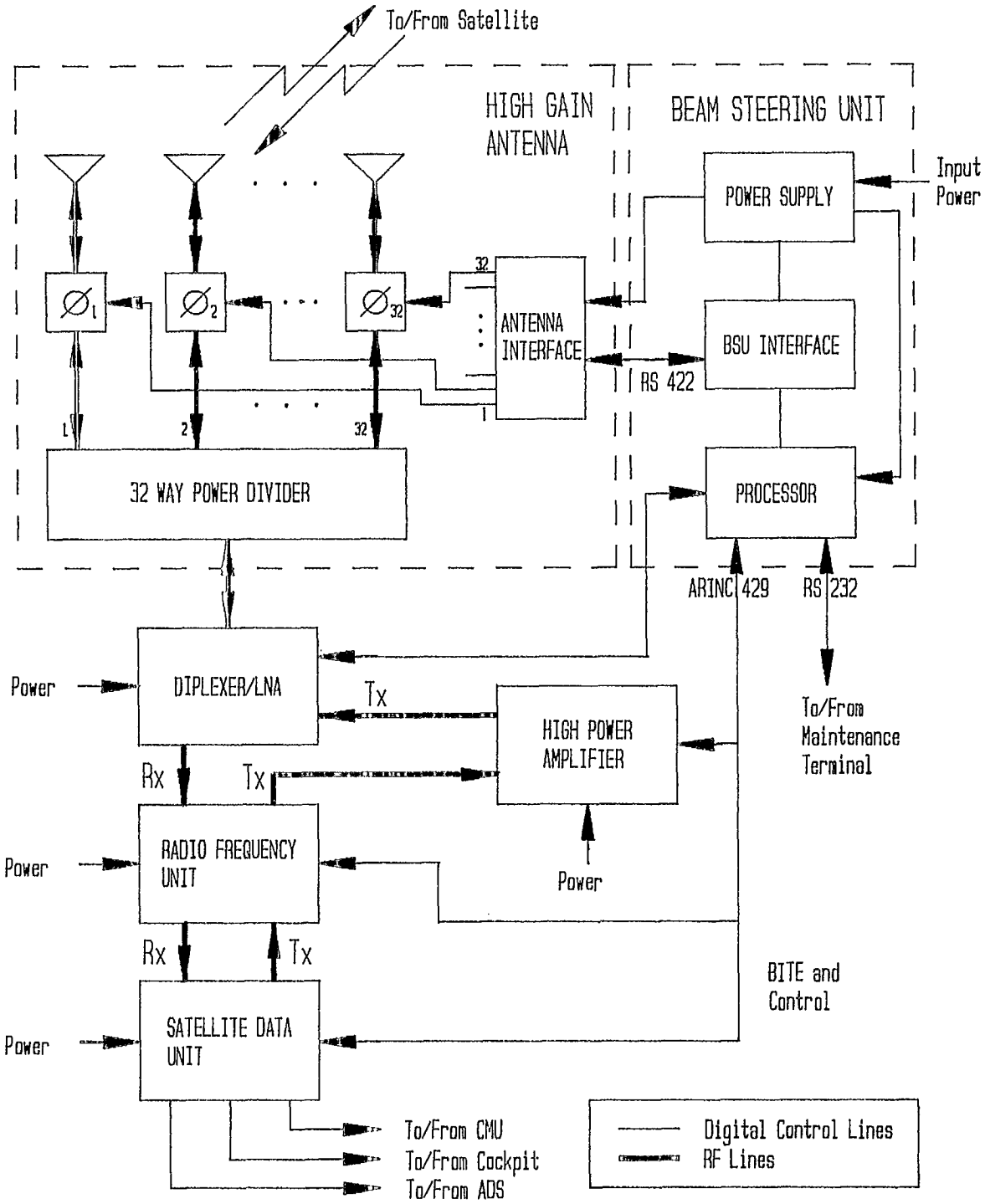


Fig 1. Typical CMA-2100 Application and Block Diagram

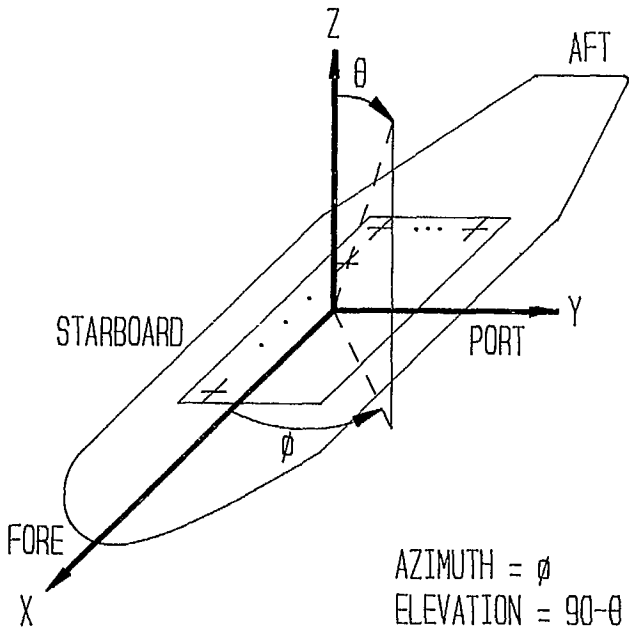


Fig. 2. Definition of AUT Coordinate System

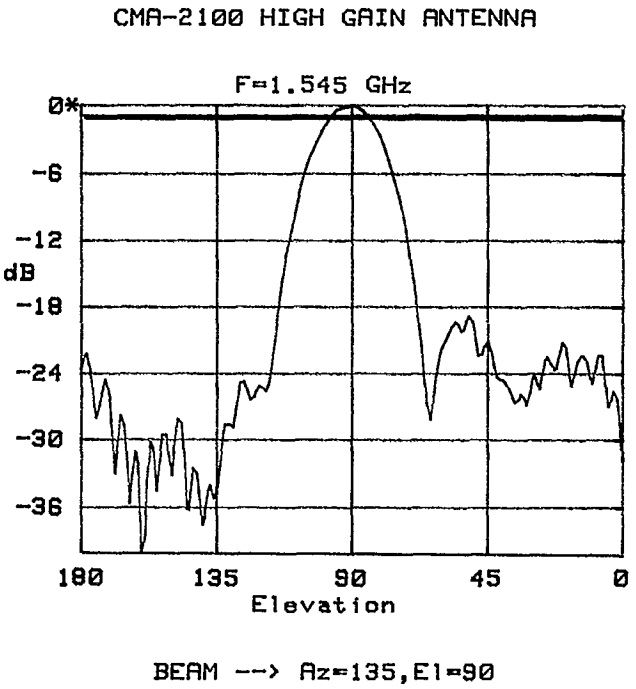


Fig. 3. Boresight Antenna Radiation Pattern

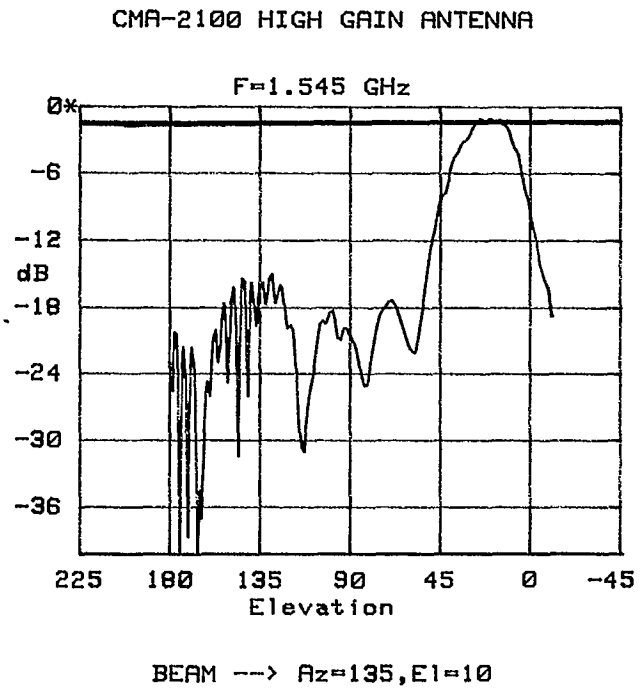


Fig. 4. Low Elevation Antenna Radiation Pattern

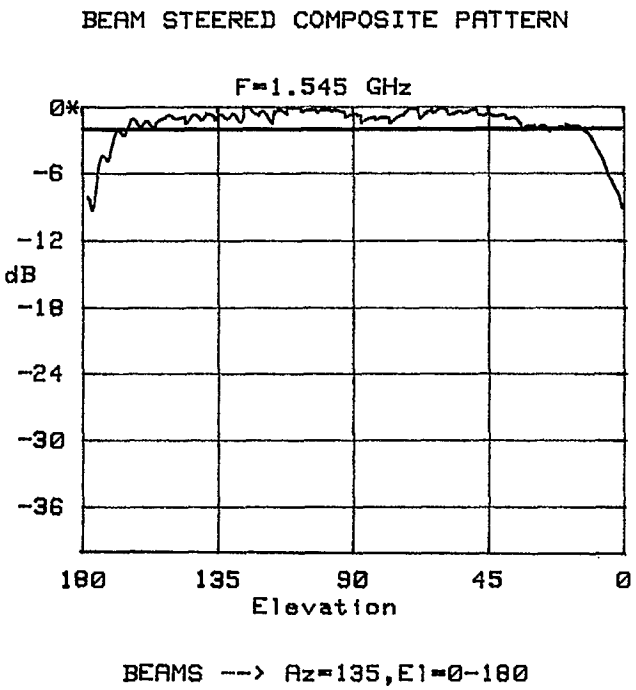


Fig. 5. Composite Scanned Antenna Radiation Pattern

An Aircraft Earth Station for General Aviation

J. Broughton, R. Matyas
Canadian Astronautics Limited
1050 Morrison Drive
Ottawa, Ontario K2H 8K7
Canada
Tel: (613) 820-8280 Fax: (613) 820-8314

R. Lyons, S. Spenler
SkyWave Electronics Limited
300 March Road, Suite 304
Kanata, Ontario K2K 2E2
Canada
Tel: (613) 592-0908 Fax: (613) 592-2104

J. Rigley
Communications Research Centre
3701 Carling Avenue
P.O. Box 11490, Station "H"
Ottawa, Ontario K2H 8S2
Canada

SYNOPSIS

The provision of aeronautical communications by satellite is gaining increasing acceptance for long range commercial aircraft. International consortia of service providers are currently procuring ground earth stations (GES's) in order to offer operational service via the INMARSAT system; some are presently conducting field trials and providing limited service in conjunction with terminal manufacturers and airlines.

Provision of a voice and data communications capability for smaller aircraft is of considerable interest in view of the very large number of such aircraft in use. Recent studies indicate a potential market of 11,500 terminals worldwide. It is anticipated that these terminals will differ in significant respects from those being developed for commercial airlines.

Canadian Astronautics Limited (CAL) and SkyWave Electronics Limited (SkyWave) are presently developing a single voice channel low cost and low weight aeronautical satellite terminal. This equipment will allow pilots and passengers to communicate with the worldwide public or private telephone and data networks through services offered by Teleglobe Canada and its international partners in France and Australia and others.

The development of a general aviation teminal poses its own set of special problems. Two issues are of paramount interest. The first is the manner in which the service can be provided in a cost-effective manner exploiting the infrastructure that is evolving for the commercial aeronautical service. The second relates to the development of an aeronautical terminal that is attractive to the general aviation community in terms of price volume, weight and capability.

This paper describes how the terminal under development by CAL and SkyWave addresses these two major issues. Topics covered include compatibility with the INMARSAT aeronautical system, design of the critical antenna subsystem, packaging philosophy, effective deployment of subsystems in the aircraft, and field trial objectives.

The full text of the paper is in the Appendix ..... page A-3

Flight Test of ARINC 741 Configuration Low Gain SATCOM System on Boeing 747-400 Aircraft

Timothy A. Murphy and Brian P. Stapleton
Boeing Commercial Airplane Group
PO Box 3707 - MS 47-31
Seattle WA. 98124-2207
(206) 655-6908
FAX (206) 655-9252

ABSTRACT

On May 23, 1989 The Boeing Company conducted a flight test of a SATCOM system similar to the ARINC 741 configuration on a production model 747-400. A flight plan was specifically designed to test the system over a variety of satellite elevations and aircraft attitudes as well as over land and sea. Interface bit errors, signal quality and aircraft position and navigational inputs were all recorded as a function of time. Special aircraft maneuvers were performed to demonstrate the potential for shadowing by the aircraft structures. During the flight test, messages were routed via a Ground Earth Station in Santa Paula, California to Collins Radio in Cedar Rapids, Iowa. These messages were also monitored by ARINC in Annapolis, Maryland through the ACARS ARINC network.

On May 20, 1989, a compass rose test was performed in preparation for a flight test. Both the compass rose test and the flight test indicated that shadowing from the tail is insignificant for the 747-400. However, satellite elevation angles below the aircraft horizon during banking maneuvers were shown to have a significantly deleterious effect on SATCOM communications.

BACKGROUND

Satellite communications will undoubtedly play an important role in future civil aeronautical operations. Presently HF communications (2-30 MHz) provides the communications link when the aircraft is beyond line of sight with VHF. HF provides SSB analog voice communications using ionospheric propagation. The FAA as well as other aviation regulatory bodies recognizes the superior reliability of SATCOM over HF with its propagation vagaries. Used wisely, satellite communications technology will enable more efficient utilization of the higher traffic density airline routes in both the Pacific and Atlantic Ocean regions.

Cognizant of the potential benefits of this technology, Boeing Commercial Airplane Group has been involved in the development of aeronautical satellite communications systems for over two decades. In more recent times Boeing has been an active participant in the industry regulatory activities associated with satellite communications. It is truly exciting that the introduction of a standard, full time, commercially available satellite communications service is imminent. As an airframe manufacturer, Boeing will be a system integrator, responsible for integrating and testing all the avionics components and finally certifying and delivering complete systems.

On May 20, 1989, in preparation for a flight test, Boeing conducted a compass rose test of the SATCOM system. This test verified that the system wiring was complete and functional as well as verifying that the installed equipment was operational.

On May 23, 1989, Boeing Avionics Engineering conducted a flight test of a SATCOM system similar to the ARINC 741 configuration[1] on a production model 747-400. This flight test was conducted using developmental engineering units as a step in design validation. Due to the nature of the aircraft and the avionics equipment, the test instrumentation was minimal. Hence the recorded performance parameters are not easily related to standard performance measures such as signal to noise ratios or physical quantities such as power. However, some qualitative observations can be made from this flight test data.

TEST DESCRIPTION

A communications link was established between the 747-400 at Paine Field, Washington and the COMSAT earth station at Santa Paula, California. The modulation used was ABPSK (Aviation BPSK) at a rate of 600 bits per second with rate one half Forward Error Correction coding (as described in the ARINC 741 characteristic). The Ground Earth Station (GES) equipment was controlled from a Collins site via a conventional telephone modem link to the Santa Paula station. Collins personnel on the ground in Cedar Rapids were able to view text messages received from the aircraft as well as compose text messages and uplink them to the aircraft. Hence, two way link operation was verified.

The aircraft was instrumented to record a number of forward link performance parameters over time throughout the flight. The data recorded were: aircraft latitude, longitude, heading and attitude (pitch and roll). In addition to these parameters, data provided by the Collins engineering Satellite Data Unit (SDU) were recorded: signal quality and bit errors. The signal quality indication is a relative measure of signal strength. The value is an indication of I squared plus Q squared as measured by the demodulator. The bit errors indication received from the SDU reported the number of (channel) bit errors detected in the 6 most recent Signal Units (or 1152 channel bits). The data collected were post processed to evaluate the effects of aircraft attitude and shadowing by aircraft structure on system performance. The aircraft position and attitude data were used to transform the satellite azimuth and elevation into the aircraft coordinate reference frame.

No return link performance parameters were recorded. In fact, very little information as to the performance of the return link was available to the flight test crew. On the forward link there is a p-channel which was monitored continuously. On the return link, the aircraft transmitted only when a text message was composed and sent. There was an acknowledgment on the control panel in the aircraft when a message had been received at the GES.

The satellite used in this test was the INMARSAT Pacific Ocean Region satellite, positioned at approximately E 180°. The GES function was performed by experimental equipment installed by Collins at the COMSAT station in Santa Paula, California.

EQUIPMENT CONFIGURATION

Figure 1 shows a block diagram of the SATCOM avionics as configured on the Boeing 747-400 aircraft. This configuration represents the low gain portion of the standard installation offered by Boeing for the 747-400 at the time of this flight test. The low gain SATCOM system design parameters are (from the ARINC 741 characteristic) a receive G/T $\geq$ -26 dB/K and a minimum transmit EIRP of 13.5 dBW. The low gain antenna is a quadrifilar helix intended to provide omnidirectional azimuth coverage from zenith down to 5° elevation angles. The antenna is housed under an aerodynamically shaped radome which is less than 6 inches high. The design goal is to achieve a minimum of 0 dBic gain for 85% of the coverage volume defined from zenith to 5° elevation angles. The low gain SATCOM system installed on the 747-400 is believed to have had a receive G/T of -24.3 dB/K (assuming a 0 dBi aircraft antenna). The transmit EIRP is believed to have just met the minimum required, 13.5 dBW.

Figure 2 shows the SATCOM provisions as offered by Boeing at the time of this engineering test. The low gain antenna is located on top centerline at body station 650 (see figure 2). Since the low gain antenna is mounted on the top of the upper deck, only the tail fin extends above the horizon as seen by the antenna. Hence the tail fin may block the direct line of sight to the satellite for some aircraft attitudes, and bearings.

COMPASS ROSE TEST

The 747-400 aircraft was positioned on a compass rose at Paine Field Washington (latitude N 47'54.1", longitude W 122'17.1"). The nominal elevation to the satellite at this location is 12.5°. The aircraft was towed in a circle with stops being made at approximately 10 degree increments in order to record the bit error rate and signal quality at each azimuth position. The signal quality and bit error indications were recorded manually from the DLC-800 controller for several seconds at each azimuth position.

Figure 3 shows the data collected during the compass rose test. The radial axis is in units of average signal quality. The angle reference is the bearing to the satellite in the aircraft reference frame. Care was taken to position the aircraft such that the tail was directly in line with the satellite. The nominal elevation was such that line of sight to the satellite was not actually

blocked and no airframe effects were evident in the data recorded.

FLIGHT TEST

Figure 4 shows the flight path chosen for the test. The flight path was designed specifically to test the system over sea and a variety of land terrain. The Pacific Northwest is at the edge of coverage for the Pacific Ocean Region INMARSAT and therefore offers an excellent opportunity to evaluate the system at its performance limits and beyond. Special maneuvers were performed at three locations, Sedar Waypoint, Moses Lake, Washington and Glasgow, Montana in order to test the satellite link at a variety of satellite bearings and elevations in three different basic multipath environments.

Near Glasgow, the aircraft adopted a heading specifically chosen to move the tail into the line of sight to the satellite. The test crew were unable to detect any harmful effects on the link performance due to tail shadowing. Also, the recorded flight data does not reflect any degradation in the link due to tail shadowing.

Figure 5 shows the elevation of the satellite, the signal quality and the bit errors recorded over a period of time spanning about half an hour. This data represents the flight maneuvers conducted over Glasgow Montana (latitude N 48'13", longitude W 106'37"). Glasgow represents the most extreme nominal operating environment included in the test with a nominal elevation to the satellite of ≈2.3°. During the orbiting maneuver, from the aircraft reference frame the satellite dipped to more than 22° below the horizon. During this extreme condition, the receiver indicated a high number of channel bit errors, messages were not received by the aircraft and messages transmitted from the aircraft were not received at the GES. However, during level flight in the vicinity, with a nominal elevation to the satellite of ≈4° (in the aircraft reference frame) messages were transmitted and received successfully.

Figure 6 shows the elevation of the satellite, the signal quality and the bit errors recorded over a period of time spanning the flight maneuvers conducted at the Sedar Waypoint (latitude N 45°30.5', longitude W 126°43.0') which is off the west coast of the state of Oregon. This geometry results in multipath from the sea for all azimuth directions. The nominal elevation to the satellite in an 'earth normal' reference frame at this location is 16.4°. It can be seen from figure 6 that the system continued to operate with a low number of bit errors when the satellite was below the aircraft reference frame horizon (as much as 10°). However, as would be expected, at some very low elevation angles, (greater than 12° below the horizon), the receiver lost lock and a high number of bit errors is indicated.

Figure 7 shows the average signal quality level and bit errors as a function of the elevation of the satellite relative to the aircraft antenna reference frame. This function was produced by averaging the data taken over the time period illustrated in figure 6.

Figure 8 shows data taken during the maneuvers conducted over Moses Lake (latitude N 47°12.7', longitude W 119°18.9') approximately in the center of the state of Washington.. Again it can be seen that the system continued to operate when the satellite was several degrees below 0° elevation relative to the aircraft antenna reference frame.

Figure 9 shows the average signal quality level and bit errors as a function of the elevation of the satellite relative to the aircraft antenna reference frame for the time period illustrated in figure 8.

CONCLUSIONS

This was a successful engineering flight test of an ARINC 741 type low data rate SATCOM system aboard a 747-400. Two way link operation was demonstrated under a variety of operating conditions. In general the SATCOM system performed far better than expected at low satellite elevation angles. There were instances where the communication link failed during banking maneuvers resulting in very low elevation angles and shadowing by aircraft structures. Furthermore, it was found that tail shadowing does not appear to be significant for the 747-400 SATCOM system.

REFERENCE

1 ARINC Characteristic 741 "Aviation Satellite Communication System", Aeronautical Radio, Inc.

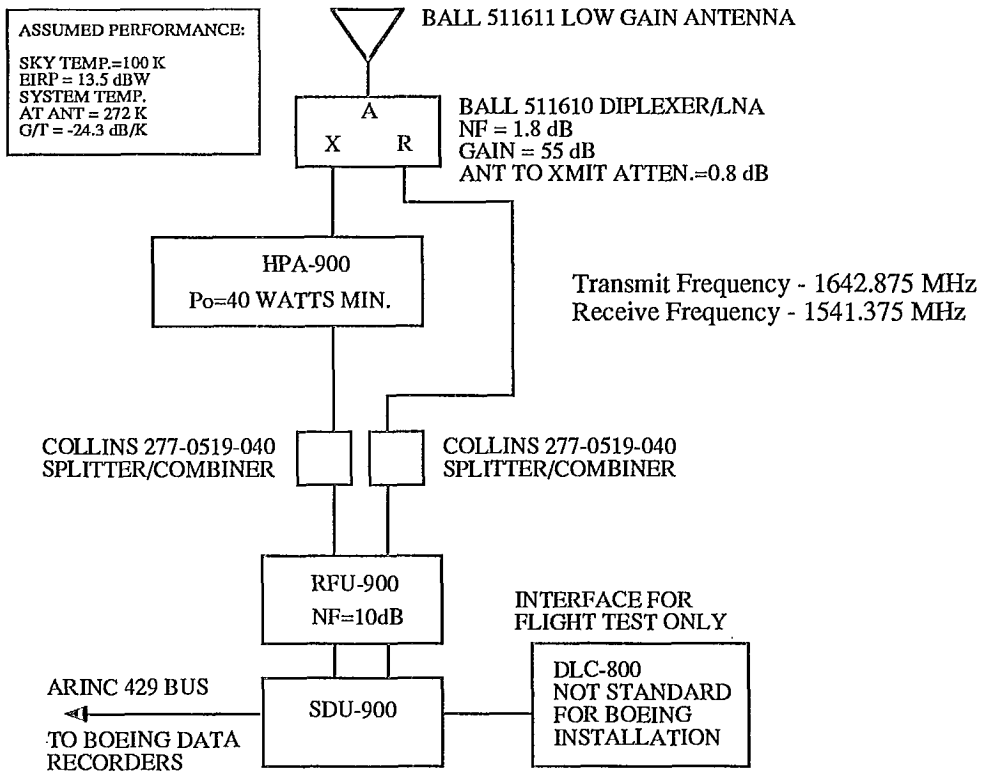


Fig. 1. Block Diagram of Low Gain Aeronautical Earth Station

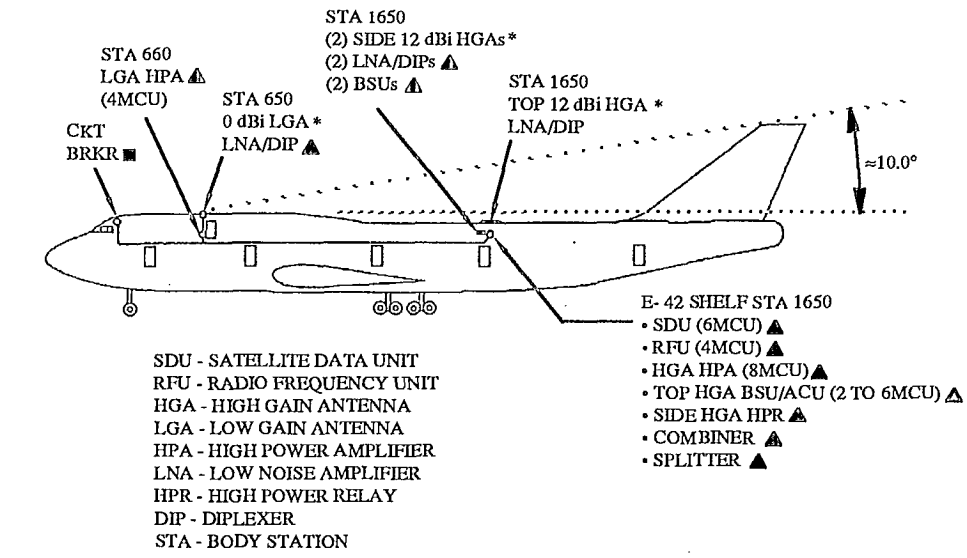


Fig. 2. Standard Boeing SATCOM Installation Provisions as of 9/89.

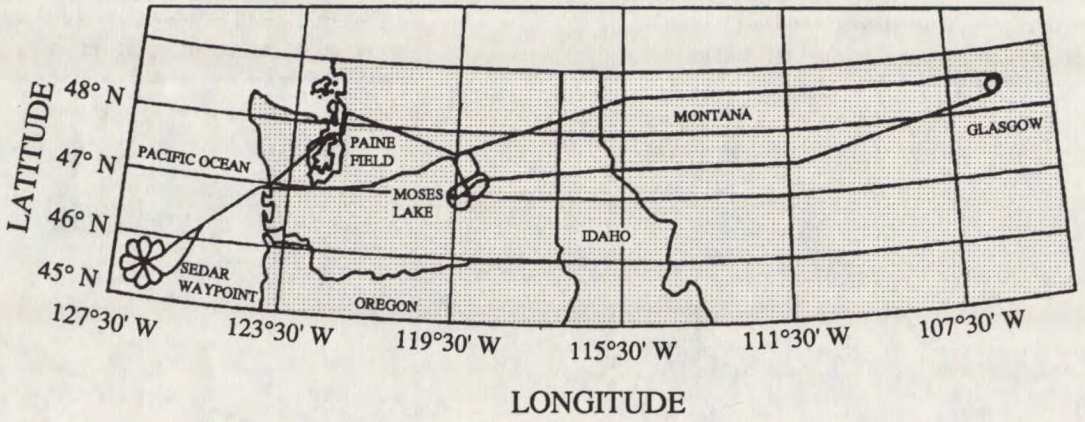


Fig. 4. SATCOM Test Flight Path

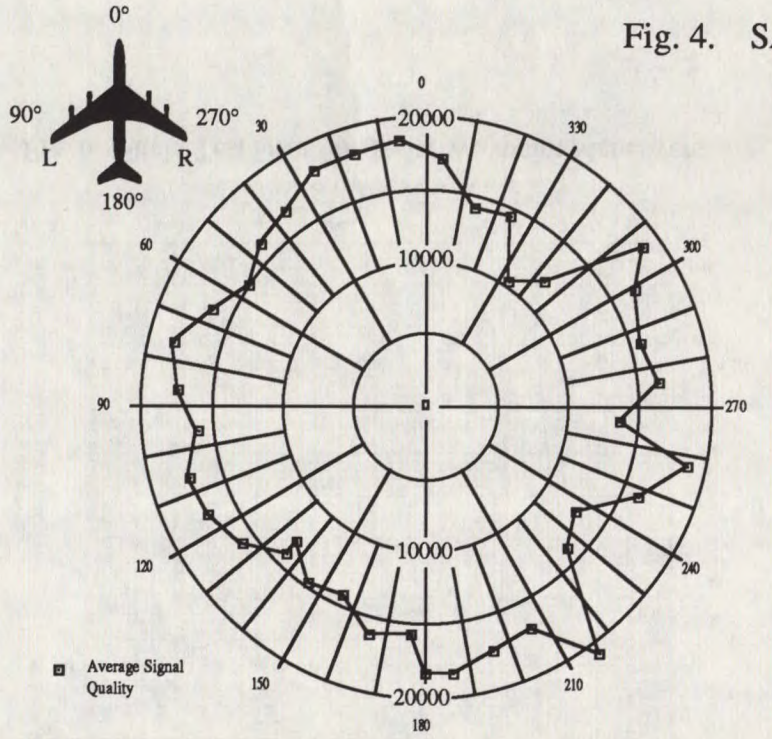


Fig. 3. Signal Quality Data From Compass Rose Test

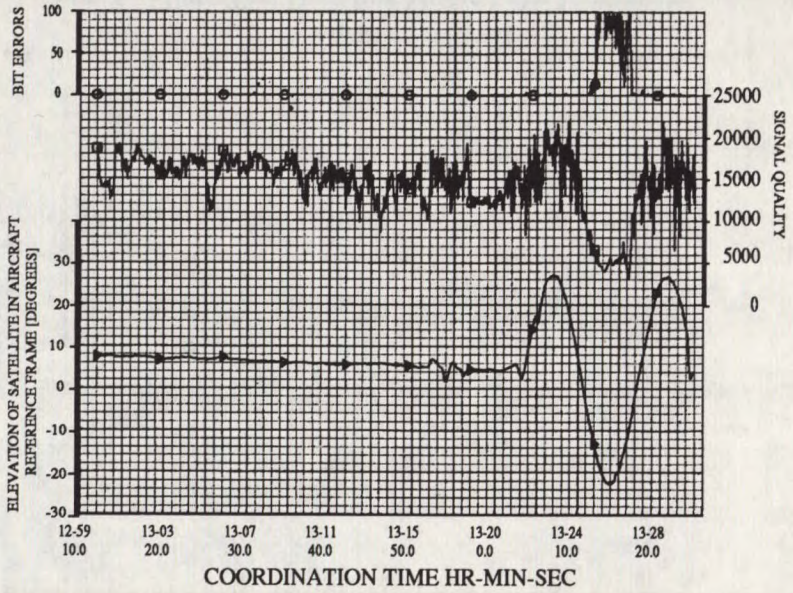


Fig. 5. Flight Test Data For Glasgow Maneuvers

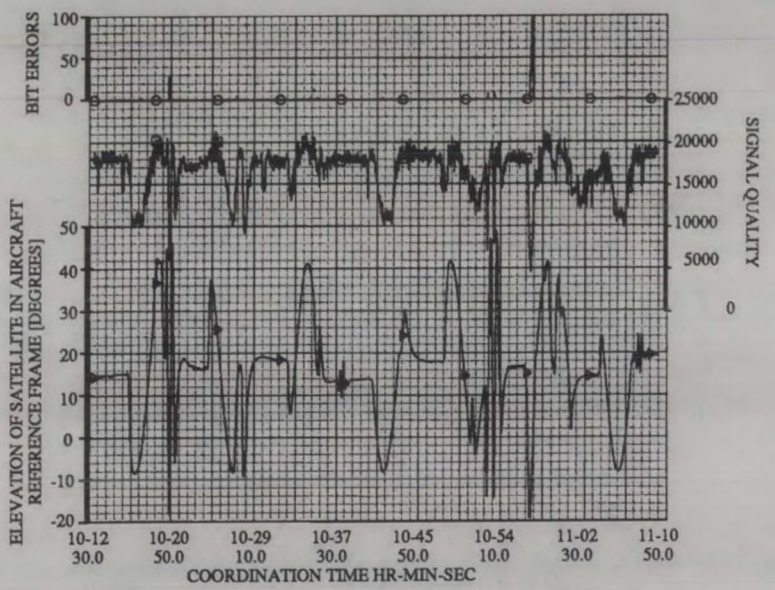


Fig. 6. Flight Test Data For Sedar Waypoint Maneuvers

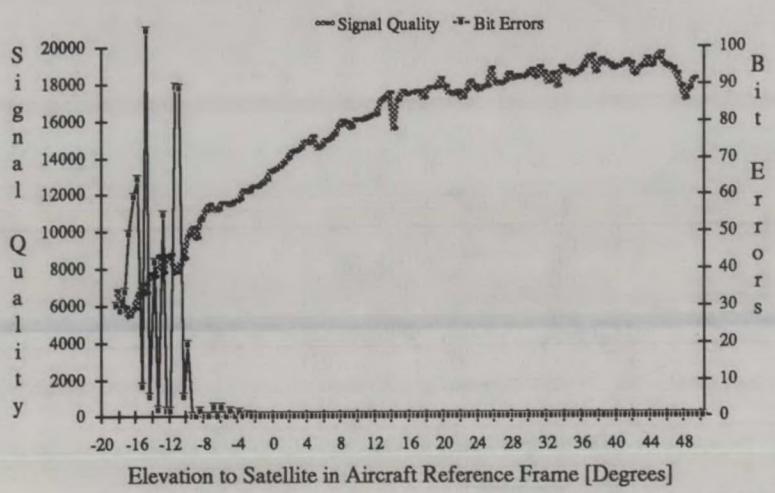


Fig. 7. Flight Test Data For Sedar Waypoint Maneuvers

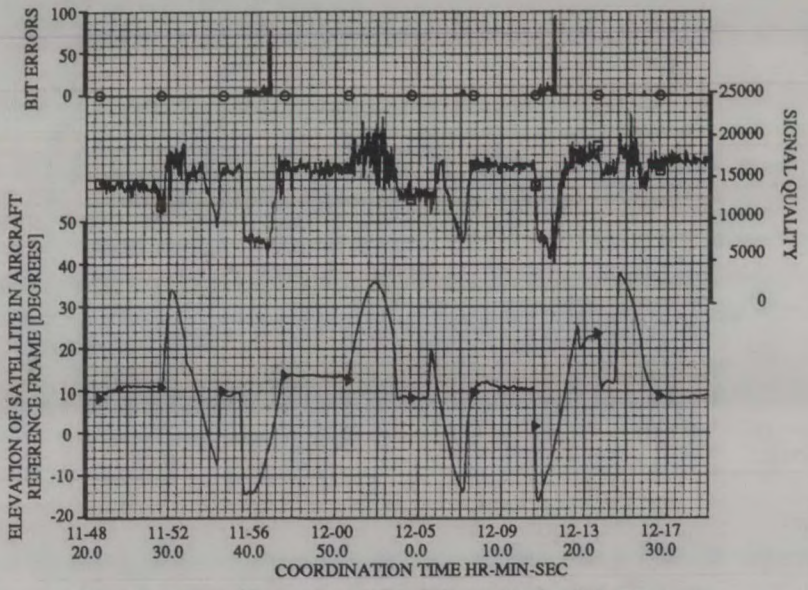


Fig. 8. Flight Test Data For Moses Lake Maneuvers

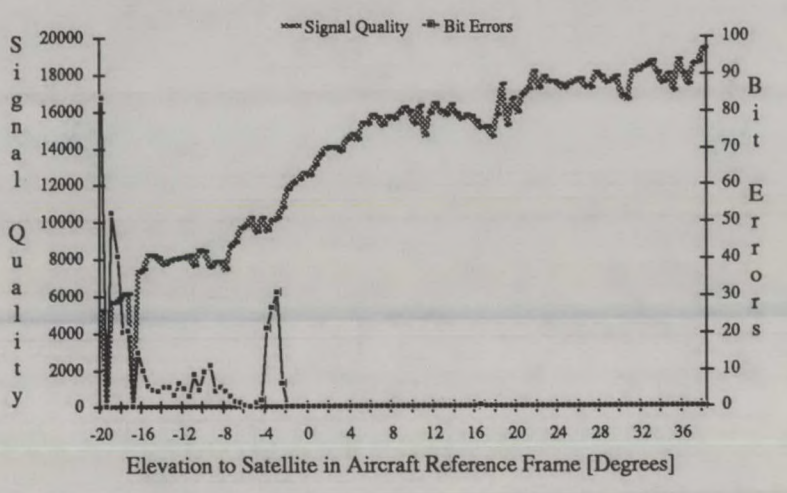


Fig. 9. Flight Test Data For Moses Lake Maneuvers

Inmarsat Aeronautical Mobile Satellite System: Internetworking Issues

Jay R. Sengupta
Inmarsat
40 Melton Street
London, UK. NW1 2EQ
Phone: +44 71 387 9089
Fax: +44 71 383 0849

ABSTRACT

The Inmarsat Aeronautical Mobile Satellite System (AMSS) provides air-ground and air-air communications services to aero-mobile users on a global basis. Communicating parties may be connected either directly or, more commonly, via interconnecting networks to the Inmarsat AMSS, in order to construct end-to-end communication circuits. The aircraft earth station (AES) and the aeronautical ground earth station (GES) are the points of interconnection of the Inmarsat AMSS to users, as well as to interconnecting networks. This paper reviews the internetworking aspects of the Inmarsat AMSS, by introducing the Inmarsat AMSS network architecture and service concepts and then discussing the internetwork addressing/numbering and routing techniques.

INTRODUCTION

The Inmarsat AMSS has a distributed network architecture composed of multiple GESs, which cooperatively perform the network management functions. Nominated GESs in each satellite region are required to continuously broadcast network management information, for the benefit of AESs wanting to register with the AMSS. The initial system will operate without any central network control facility.

Log-in function

Individual aircraft equipped with AES are required to first log-on to any GES in the network, before they may avail of the communication services offered by the AMSS. Therefore, each GES along with the AESs logged-on to it essentially constitutes an independently operating sub-network within the global AMSS. GESs are required to cooperate with each other, by forwarding call setup requests for circuit-mode service connections intended to be handled by some other GES in the same satellite region.

AMSS SERVICE CONCEPTS

Service definitions in the Inmarsat AMSS have been developed based on the principles of structures and layered system design. It has also been an objective to maximize the degree of similarity with other Inmarsat standard systems, used for maritime and land mobile applications, without compromising the unique service requirements of the AMSS. The AMSS essentially offers two basic *bearer* services to the user, namely circuit-mode and packet-mode.

Circuit-mode bearer service

The circuit-mode bearer service has been designed to be compatible with the definition of the Integrated Services Digital Network (ISDN), in order to facilitate efficient interconnection with the terrestrial telecommunications networks of the future . The circuit-mode bearer service is supported by demand assigned single channel per carrier (DA-SCPC) operation. The bearer service essentially provides a connection oriented digital circuit, which is used to support a variety of teleservices such as telephony, facsimile and data transmission using the appropriate channel interface equipment.

Packet-mode bearer service

The packet-mode bearer service has been designed to be compatible with the Open Systems Interconnection (OSI) reference model, in order to facilitate provide a common interface for interoperability with diverse applications. The packet-mode bearer service is supported by random access / reservation TDMA and TDM channels, in the return and forward directions, respectively. The bearer service essentially provides a connection oriented packet communication circuit, which is used to support a variety of teleservices such as data and telex transmission using the appropriate channel interface equipment.

In addition to offering the benefits of efficient internetworking in the short/medium term, adherence to the above design objectives is expected to offer significant benefits in the longer term due to the convergent developments in ISDN and OSI standards in the support of multi-service telecommunication networks.

The internetwork traffic routing requirement, in order to effect end-end communications, is illustrated in Figure 1.

INTERNETWORK INTERFACES

Air-ground mobile satellite communications requires the interconnection of the AMSS with other subnetworks, in order to provide the required connectivity between the end users/systems. The Inmarsat AMSS facilitates this interconnection by supporting a variety of internationally standardized network interfaces. The AES and GES are the points of interconnection of the AMSS with avionics and terrestrial networks, respectively. The interfaces supported by the AMSS for interconnection with the various different network types is illustrated in Figure 2.

AES Interconnections

In the aircraft, the AES is required to be interconnected to the avionics network, with connectivity to both cockpit and cabin communication systems. The AEEC, which is responsible for the definition of avionics equipment interface characteristics, recommends the use of ISO 8208 (CCITT X.25 PLP) and CCITT Q.931 network layer interfaces for the interconnection of the AES with the cabin and cockpit systems, respectively.

GES Interconnections

On the ground, the GES is required to be interconnected to both private and public, packet-switched and circuit-switched networks. The ICAO and CCITT are responsible for the definition of recommended interfaces for the provision of ATS/AOC and APC communication services, respectively. The ICAO recommends the use of ISO 8208 (CCITT X.25 PLP) for interconnection to the future Aeronautical Telecommunications Network (ATN) terrestrial private subnetworks. The CCITT, on the other hand, recommends the use of CCITT X.75 for interconnection to the international public switched packet data network (PSPDN), and the use of CCITT No.7 (ISUP) for interconnection to the international public switched telephone network (PSTN) and integrated services digital network (ISDN).

The Inmarsat AMSS is capable of supporting each of the aforementioned interfaces in an efficient manner, since its service have been developed on the same ISDN and OSI design principles as these interfaces.

NUMBERING/ADDRESSING PLANS

The numbering/addressing plans for the Inmarsat AMSS have been developed to facilitate end-end communications, by accommodating the following aspects of diversity in the service requirements :

a) bearer service types - circuit-mode, packet-mode ;

b) interconnecting terrestrial networks - private, public; circuit-switched, packet-switched ;

c) user service categories - ATS, AOC, AAC, APC ;

In particular, the numbering/addressing plans are required to provide the means for the routing of different traffic types, as illustrated in Figure 2.

It has been necessary to adopt several different numbering/addressing plans as described below, since a single integrated plan could not satisfy all requirements.

Circuit-mode Service Numbering

Public network numbering plan. The Inmarsat AMSS numbering plan for public correspondence circuit-mode service, as defined in CCITT Recommendation E.215, is fully integrated into the international public network plan. Each Inmarsat satellite region has been assigned a distinct country code within the global numbering plan, and the Inmarsat-aero system is distinguished from other Inmarsat standard systems by a digit immediately following.

The international telephone number format for an Inmarsat AES is : CCC T $X_1...X_8$, where,
CCC = country code for Inmarsat satellite region;
T = Inmarsat system standard designnator;
X_n =Inmarsat (aero) mobile number.

With the implementation of the ISDN Numbering Plan in the international public terrestrial network in the year 1997, the length of the above number will increase by three digits and thereby provide the capability for *direct-dialling-in* to individual terminals on-board the aircraft. Moreover, in the interim period before 1997 Inmarsat intends to utilize spare capacity within the Inmarsat (aero) mobile number field for this purpose.

Private network numbering plans. The Inmarsat AMSS provides for either transparent or non-transparent support of private network numbering plans. In the former case, the AMSS performs routing based on a *Network Type* discriminator used within internal signalling messages, and simply transports the private network number to/from the point of attachment of the private network in a transparent manner. In the latter case it is possible, at the option of individual GES Operators, for the Inmarsat AMSS to perform routing based on the private network numbering plan and thereby be transparent to the private network user.

Packet-mode Service Addressing

Public network numbering plan. The Inmarsat AMSS addressing plan for public correspondence packet-mode service, as defined in CCITT Recommendation X.121, is fully integrated into the international public data network plan. Similar to the circuit-mode service case, each Inmarsat satellite region has been allocated a Data Network Identification Code (DNIC), and the Inmarsat-aero system is distinguished from other Inmarsat standard systems by means of a following digit.

The international data address format for an Inmarsat AES is : NNNN T $X_1...X_8$ D , where,
NNNN = DNIC for Inmarsat satellite region;
T = Inmarsat system standard designator;
X_n = Inmarsat (aero) mobile number;
D = on-board terminal identifier (optional).

This addressing plan provides a limited *direct-dialling-in* capability to address individual terminals on-board the aircraft. The impact of the implementation of the ISDN Numbering Plan in the public terrestrial network in 1997, is currently under study.

Private network numbering plans. The Inmarsat AMSS provides for either transparent or non-transparent support of private network addressing plans. In the former case, the private network address is simply carried as higher layer *user data* between the private network's points of attachment. This transparent mode will be applicable for the support of Network Service Access Point (NSAP) addresses, which will be used in the future ATN. In the latter case, at the option of GES Operators, the private network address could be integrated into the AMSS DTE Addressing Plan described below.

In order to facilitate the interconnection of multiple private networks, a *Subnetwork DTE Addressing Plan* has been defined for the Inmarsat AMSS. This addressing plan is used to uniquely identify all data terminal equipment (DTE) that are directly attached to the AMSS.

INTERNETWORK ROUTING

Internetwork routing in the aeronautical mobile environment is primarily characterized by the requirement to manage the mobility of aircraft in flight. The internetwork routing techniques must therefore incorporate methods to manage the mobility of the AESs in the AMSS, which is a mutli-region distributed network. The Inmarsat AMSS is capable of supporting three different routing regimes as described below, although only the first method is currently implemented; the latter two techniques require the implementation of some form of mobile location registration facility, and the network interconnections required are illustrated in Figure 3.

Mobility management by calling users

This routing management technique does not require the AMSS or any other interconnecting subnetwork to perform any mobility management functions. It requires the calling terrestrial user to be aware of the location (region) of the called AES; the call attempt simply fails if the called AES is absent from the chosen region. This routing technique is supported by the current Inmarsat AMSS public correspondence numbering/addressing plans, which require the selection of a particular AMSS satellite region by the calling party.

This routing technique is acceptable when there are a small number of AMSS regions. However, it carries the risk of significant penalty in the proportion of unsuccessful call attempts and associated decrease in network traffic handling capacity, due to the absence of the called AES from the selected AMSS region, if the number of regions within the AMSS were to increase.

Mobility management by AMSS

This routing management scheme would require the implementation of a *Location Register* within the Inmarsat AMSS as illustrated in Figure 3, which would register the current location (satellite region) of all AESs currently logged-on to the AMSS. The Location Register would be updated by every GES, whenever an AES newly logs-in or is logged-out by it. Direct signalling links would exist between major nodes in the terrestrial networks (Network 1 in Figure 3) and the Location Register, to query the current location of the called AES. The connection would then be routed to a GES serving the appropriate AMSS satellite region. In cases where direct signalling links do not exist between the terrestrial network nodes and the Location Register the connection could first be routed to a default GES, which would then query the Location Register and reroute the connection to some other AMSS satellite region if necessary.

The signalling links and location registration functions can be implemented using CCITT No.7 (MSAP) protocols and procedures, which are currently under definition. Having implemented a central database for mobility management purposes, the same facility and signalling/control procedures can be used to accommodate other network management and value-added services, such as calling user authentication and private virtual networking.

It should be noted that, currently, there are no firm plans to implement this method of mobility management in the Inmarsat AMSS.

Mobility management by interconnecting subnetworks

This routing management scheme places the responsibility of AMSS network mobility management, on the interconnecting subnetwork. Signalling links would exist between the GESs and nodes in the interconnecting subnetwork, as depicted by Network 2 in Figure 3. The AMSS would be required to convey any changes in AES log-on status, either from the AES or the GES to their respective interconnected subnetwork node. The location registration function may be implemented either as a standalone facility, or be integrated into the routing tables of the nodes within the interconnected subnetwork.

This method of internetwork routing management has been proposed for the interconnection of the Inmarsat AMSS to the future ATN, for the purposes of

ATS/AOC/AAC packet-mode services. In this case, the Inmarsat AMSS will also support the ATN requirement of transparent transfer of Internet Protocol (IP) packets, which are used to disseminate routing information between the ATN nodes.

CONCLUSION

This paper has presented a brief description of the internetworking capabilities of the Inmarsat AMSS. The internetworking capabilities of the Inmarsat AMSS will continue to be enhanced as required, to accommodate evolutionary changes in subnetwork architectures and user service requirements. Inmarsat actively participates in the international fora to promote the development of standard methods of internetworking.

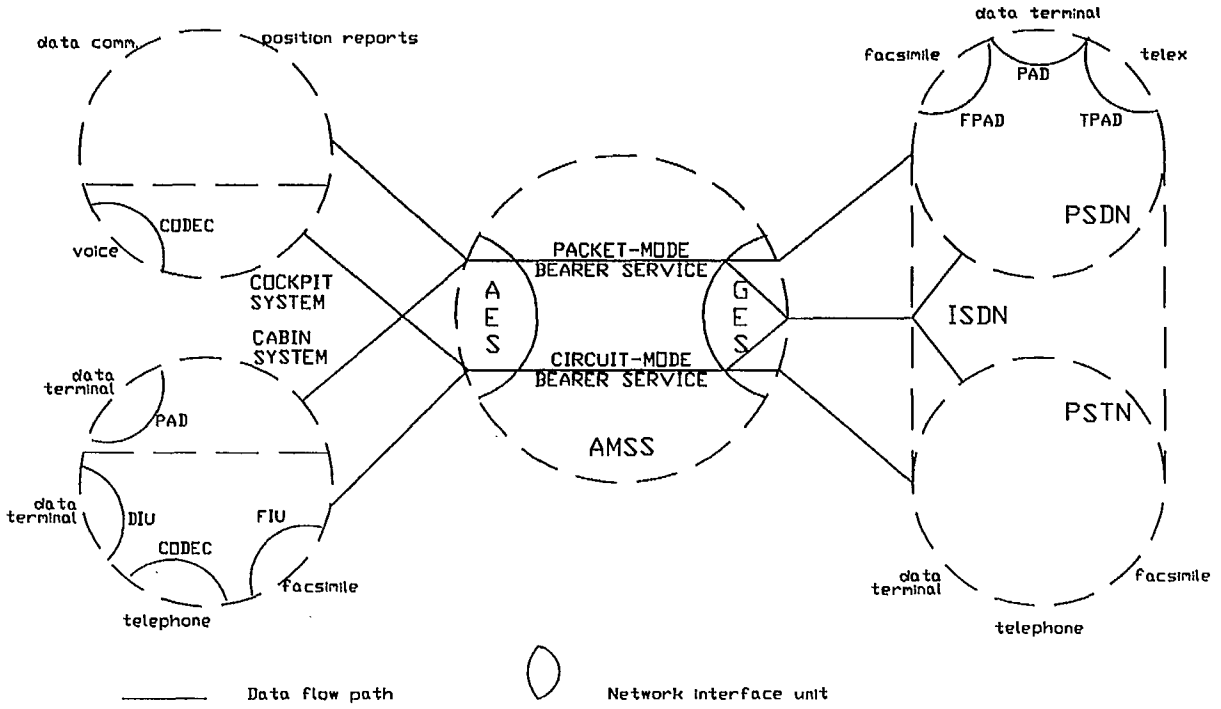


Fig. 1. Traffic Routing Requirements

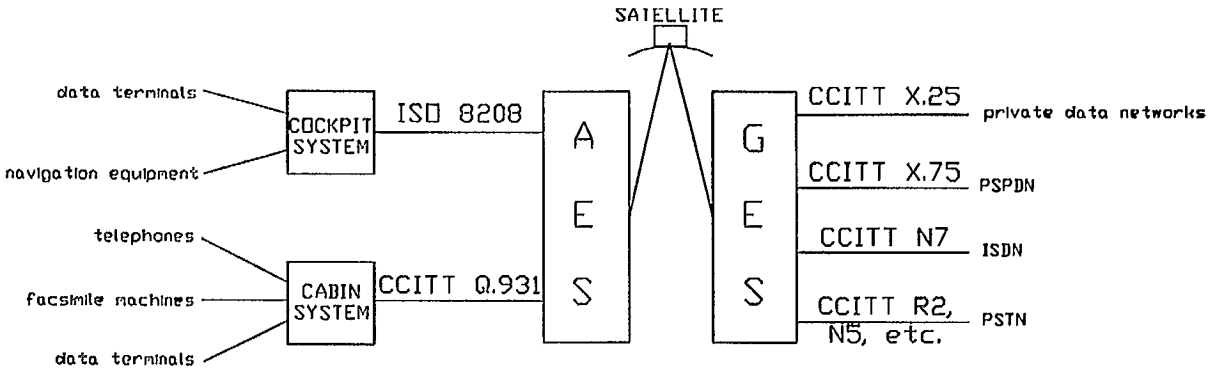


Fig. 2. Network Interconnection Scenario

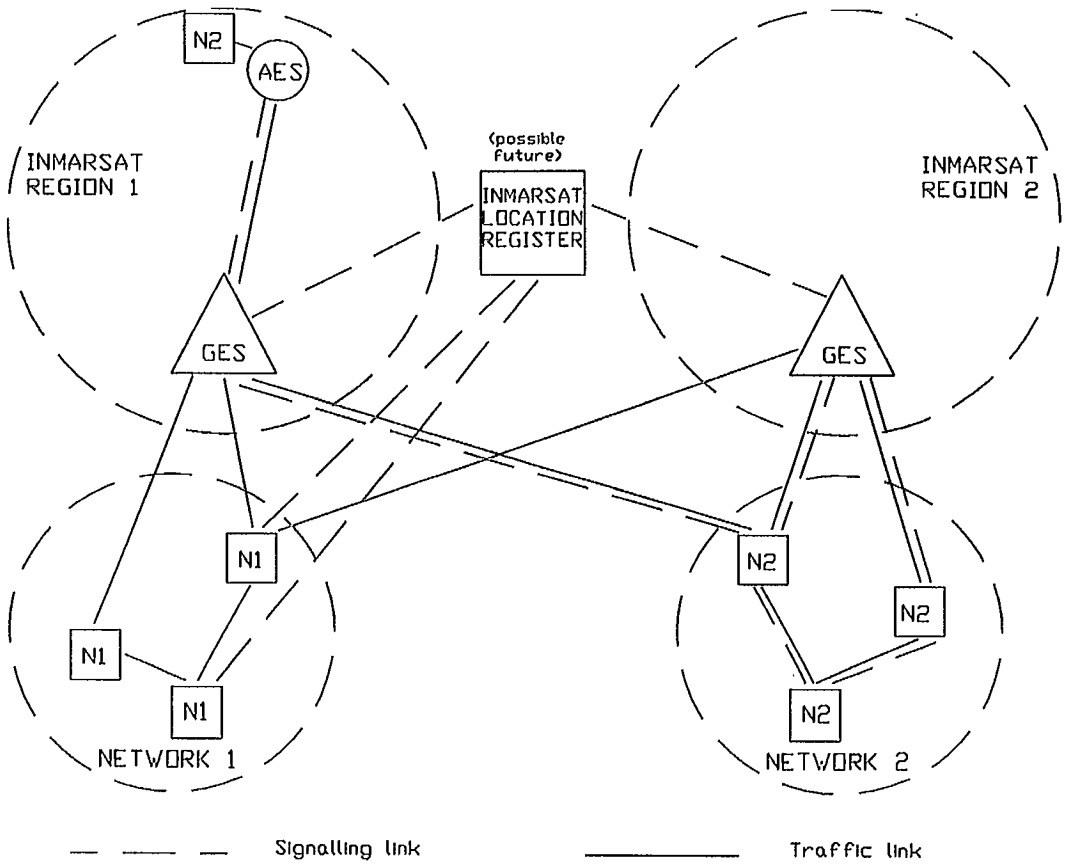


Fig. 3. Mobility Management in the Internetwork

Low Cost, Electronically Steered Phased Array for General Aviation

Peter C. Strickland
Canadian Astronautics Limited
1050 Morrison Drive
Ottawa, Ontario K2H 8K7
Phone: (613) 820-8280
FAX (613) 820-8314

ABSTRACT

This paper describes a multi-faced, phased array antenna developed for general aviation satellite communications applications. The antenna design satisfies all INMARSAT Aeronautical SDM requirements. Unique features of this antenna include an integral LNA and diplexer, integral phase shifters which are shared among the array faces, a serial beam steering interface and low manufacturing cost.

INTRODUCTION

Canadian Astronautics Limited (CAL) has developed a high gain, electronically steered antenna for aeronautical satellite communications. This development is part of a larger program which is producing a low cost terminal for General Aviation applications [1]. The program is jointly sponsored by Transport Canada, Supply and Services Canada and Communications Canada.

The antenna provides coverage in the full upper hemisphere by means of a five faced phased array. all of the array faces and control electronics are enclosed in a single compact, lightweight, aero-dynamic radome. The radome also houses the low noise amplifier and diplexer. A photograph of the antenna system with its radome removed is given in Figure 1. The diplexer, controller and low noise amplifier are shown in Figure 2.

Antenna Design

An obvious advantage of housing the LNA and diplexer within the radome is that losses are minimized. This arrangement allows a system G/T of -13dB/k to be achieved with a very compact, low cost antenna. The antenna system, without radome, is 1.27m x 0.15m x 0.31m at its base. With the configuration chosen by CAL there is no need for an additional electronics unit inside the bulkhead. All cables connecting to the antenna go directly to the transceiver unit which is mounted in an avionics equipment rack.

The antenna controller receives beam steering commands from the transceiver by means of a single serial control line. On each of the three phase scanned arrays the beam position may be stepped in 0.5 degree increments. The system incorporates two modes of beam steering. In one mode steering is derived from the inertial navigation system. In the second mode a sequential lobing algorithm is implemented. This algorithm dithers the antenna beam pointing in order to maximize a signal quality estimate obtained from the modem.

The antenna system uses a total of twelve, 4-bit phase shifters. These are shared by three of the array faces (top, port and starboard). The remaining two faces are fixed beam arrays. PIN diode switches allow rapid, hot switching between the array faces. State-of-the-art, low losses have been achieved in the switches and phase shifters. The average loss is 0.3 dB per bit.

Microstrip radiating elements have been used on all array faces. The use of a thick, low dielectric constant substrate provides broad bandwidth while maintaining a very low mass. A unique microstrip patch design results in a very broad beamwidth allowing scanning to ± 70° from broadside on each of the 12 element arrays.

Typical azimuth and elevation patterns are given in Figure 3 and 4 respectively. Low sidelobes are obtained even at large scan angles through the use of a random phase added algorithm and Taylor weighting.

[1] R. Matyas et.al., "An Aircraft Earth Station for General Aviation", in these Proceedings.

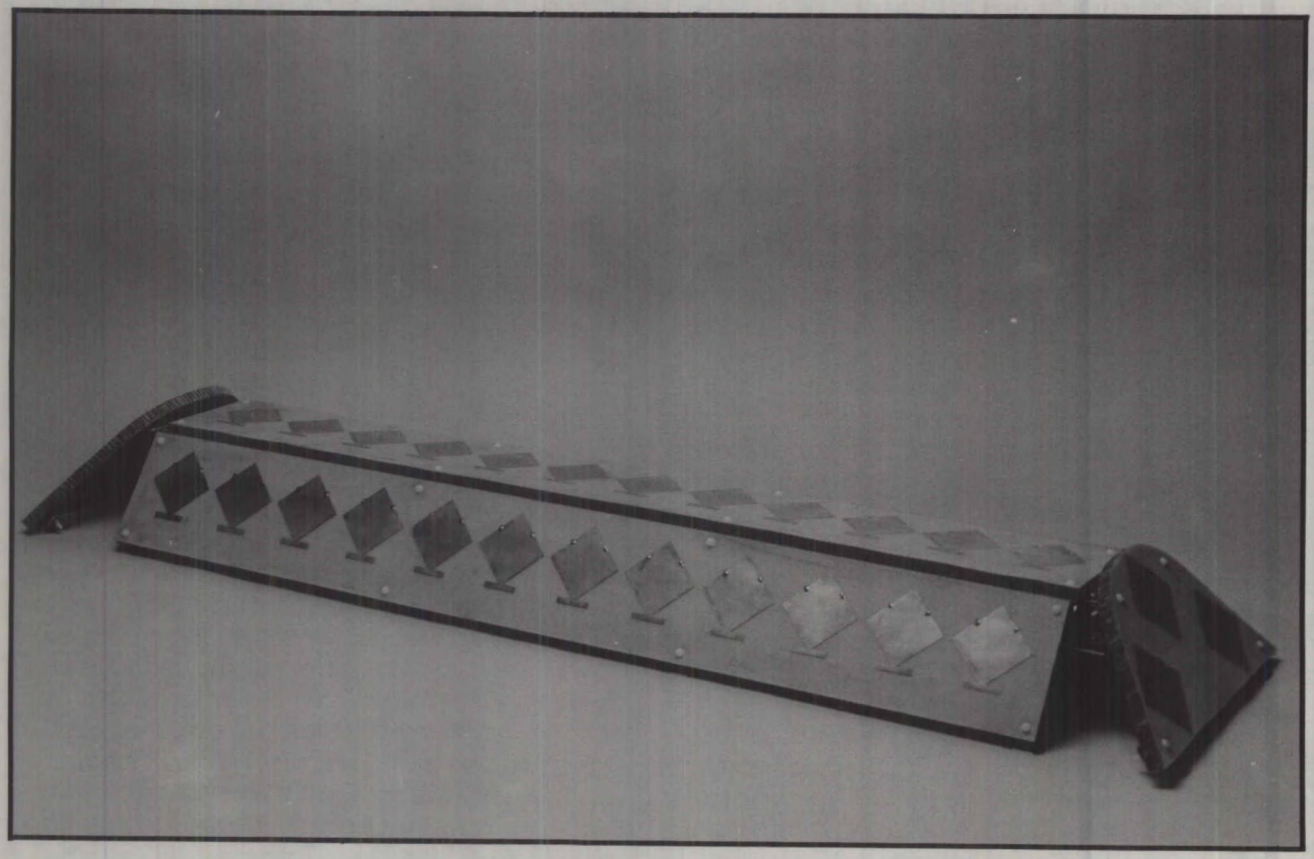


Figure 1 Antenna System Without Radome

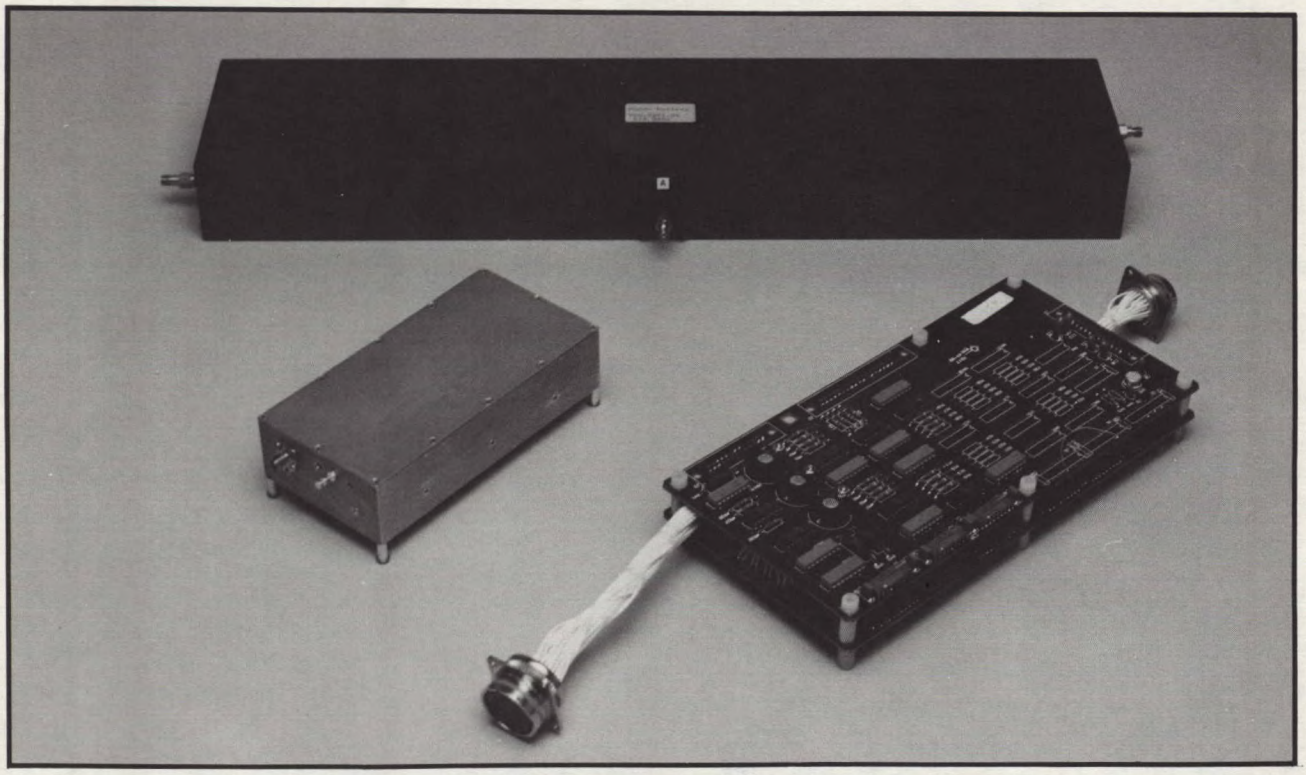


Figure 2 Diplexer, Controller and LNA

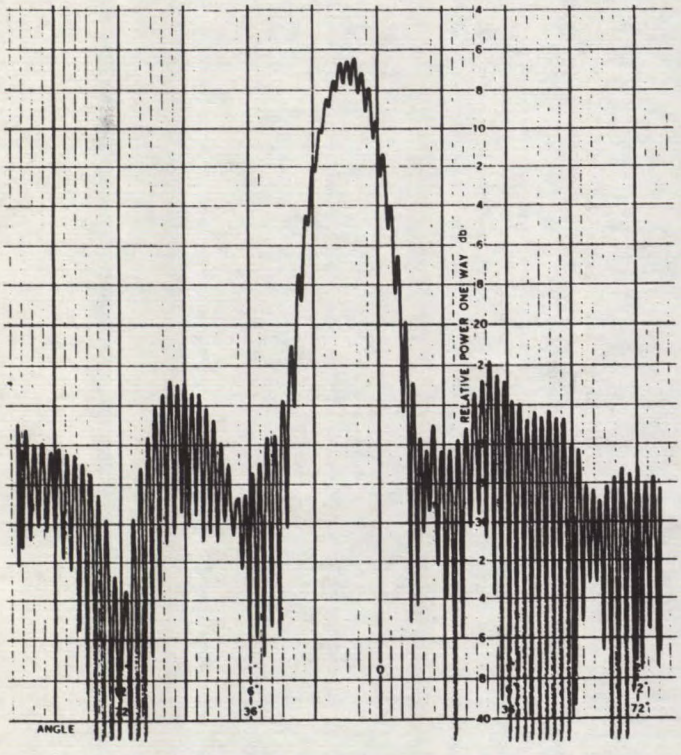


Figure 3 Azimuth Pattern of Phased Array at Broadside

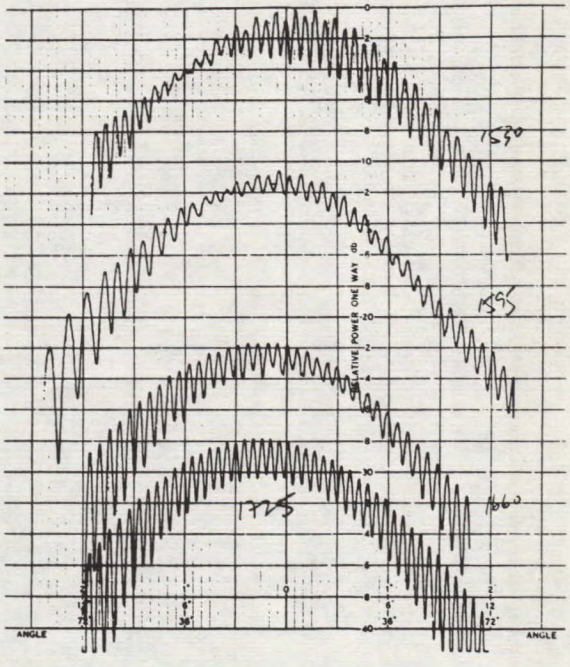


Figure 4 Elevation Pattern of Phased Array

Session 5
Modulation and Coding - I

Session Chairman - *Peter McLane,* Queen's University, Canada
Session Organizer - *John Lodge,* DOC

5

Combined Trellis Coding and Feedforward Processing for MSS Applications

D. Divsalar, M. K. Simon, T.Jedrey, N. Lay, W.Rafferty

Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109 USA

ABSTRACT

The idea of using a multiple (more than two) symbol observation interval to improve error probability performance is applied to differential detection of trellis coded MPSK over a mobile satellite (fading) channel. Results are obtained via computer simulation. It is shown that only a slight increase (e.g., one symbol) in the length of the observation interval will provide a significant improvement in bit error probability performance both in AWGN and fading environments.

1.0 INTRODUCTION

In a previous paper [1], the notion of using a multiple symbol observation interval for differentially detecting uncoded multiple phase-shift-keying (MPSK) was introduced. In particular, the technique made use of maximum-likelihood sequence estimation of N-1 (N > 2) phases rather than symbol-by-symbol detection as in conventional (N = 2) differential detection. The amount of improvement gained over conventional differential detection was shown to be a function of the number of phases, M, and the number of additional symbols (N-2) added to the observation. Furthermore, as the number of symbols, N, in the observation interval theoretically approached infinity, the performance was shown to be identical to that corresponding to ideal *coherent* detection.

In [2], this idea was extended to trellis coded modulations (TCM), in particular, MPSK. There, it was shown that a combination of a multiple trellis coded modulation (MTCM) [3] with multiplicity (number of trellis code output symbols per input symbol) equal to N-1 combined with a multiple symbol differential detection scheme analogous to that in [1] can potentially yield a significant improvement in performance, even for small N, over that corresponding to conventional trellis coded multilevel DPSK (MDPSK).

In this paper, we give further evidence of the gain demonstrated in [2] as well as extending these notions to fading channels.

2.0 SYSTEM MODEL

Figure 1 is a simplified block diagram of the system under investigation. Input bits occurring at a rate R_b are passed through a rate nk/(n+1)k multiple trellis encoder (k is the multiplicity of the code) producing an encoded bit stream at a rate $R_s = [(n+1)k/nk]R_b$. Next, the encoded bits are divided into k groups of n+1 bits each and each

[1]The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

group is mapped into a symbol selected from an $M = 2^{n+1}$ - level PSK signal set according to a set partitioning method for multiple trellis codes [3] analogous to that proposed by Ungerboeck [4] for conventional (unit multiplicity) codes. Since the MDPSK symbol rate is R_b/n, it is reasonable, from a conservation of bandwidth standpoint, to compare the performance of this system to an uncoded $M = 2^n$ level DPSK system with the identical input bit rate.

At the receiver, the noise-corrupted signal is differentially detected and the resulting symbols are then fed to the trellis decoder which is implemented with the Viterbi algorithm. In selecting a decoding metric, a tradeoff exists between simplicity of implementation and the optimality associated with the degree to which the metric matches the differential detector output statistics.

For the case of uncoded MDPSK, a metric based on minimizing the distance between the received and transmitted signal vectors is optimum in the sense of a minimum probability of error test. The specific forms of this metric for conventional and multiple differential detection were described in [1]. For conventional trellis-coded MDPSK, the metric takes on the form of a minimum squared-Euclidean distance metric. For multiple symbol detection of MTCM, the form of the distance metric is quite different. Nevertheless, as was shown in [5], by a suitable modification of the multiple trellis code design, the appropriate distance metric can be converted once again into a squared-Euclidean distance metric. The so-called "equivalent" multiple trellis code that results from this modification then becomes the key tool used for analyzing the performance of the system.

3.0 CHANNEL MODEL

The mobile satellite fading model is shown in Figure 2. A complex representation of the mathematical model for the fading channel, used to assess the performance of MTCM systems, is also illustrated in Figure 2. In this figure, F(t) represents the fading process, m(t) represents the lognormal shadowing process, N(t) is a complex Gaussian noise process, and $\omega_d = 2\pi f_d$ is the Doppler spread in units of rad/sec.

In this paper, the effect of shadowing is not accounted for. Furthermore, it is assumed that the fading is slowly varying with a normalized amplitude $\rho = |F(t)|$ which is Rician distributed, i.e.,

$$p(\rho) = 2\rho(1+K)\exp[-K-\rho^2(1+K)] \times I_0\left(2\rho\sqrt{K(1+K)}\right); \quad \rho \geq 0 \qquad (1)$$

where the parameter K is the ratio of the power in the coherent (line-of-sight and specular) component to that in the noncoherent (diffuse) component. A special case of the Rician fading model is the Rayleigh fading channel (corresponding to $K = 0$) which characterizes terrestrial mobile radio systems.

4.0 ANALYSIS MODEL

We denote a coded symbol sequence of length N_s by

$$\underline{x} = (x_1, x_2, \cdots, x_{N_s}) \qquad (2)$$

where the k*th* element [2] of $\underline{x}$, namely, x_k, represents the transmitted MPSK symbol in the k*th* transmission interval and, in general, is a nonlinear function of the state of the encoder and the nk information bits at its input. Before transmission over the channel, the sequence $\underline{x}$ is differentially encoded producing the sequence $\underline{s}$. In

[2] Here k is used as an index. As already mentioned, we shall also use k to denote the multiplicity of the code. The particular case in point should be clear from the context of the usage.

phasor notation, s_k and s_{k+1} can be written as

$$s_k = \sqrt{2P}\, e^{j\phi_k}$$

$$s_{k+1} = s_k x_{k+1} = \sqrt{2P}\, e^{j(\phi_k + \Delta\phi_{k+1})} = \sqrt{2P}\, e^{j\phi_{k+1}} \tag{3}$$

where $E_s = rE_b$ is the energy per MDPSK symbol and

$$x_k = e^{j\Delta\phi_k} \tag{4}$$

is the phasor representation of the MPSK symbol $\Delta\phi_k$ assigned by the mapper in the k*th* transmission interval.

The corresponding received signal in the k*th* transmission interval is

$$r_k = F_k s_k e^{j\theta_k} + n_k \tag{5}$$

where n_k is a sample of zero mean complex Gaussian noise with variance

$$\sigma_n^2 = \frac{2N_0}{T} \tag{6}$$

and θ_k is an arbitrary phase introduced by the channel which, in the absence of any side information, is assumed to be uniformly distributed in the interval $(-\pi,\pi)$. F_k is a sample of a complex Gaussian fading process and thus $|F_k|$ is Rician distributed.

In [1] it was shown that for uncoded MPSK the maximum-likelihood decision statistic based on an observation of N successive MPSK symbols (the present one, i.e., the k*th*, and N-1 in the past) is

$$\eta = \left| r_{k-N+1} + \sum_{n=0}^{N-2} r_{k-n} e^{-j\sum_{m=0}^{N-n-2}\Delta\phi_{k-n-m}} \right|^{\ell} \tag{7}$$

with $\ell = 1$ or 2. If (7) is used as a branch metric for the coded case then, for low SNR, $\ell = 2$ whereas, for high SNR, $\ell = 1$. Since the first phase in this sequence acts as the reference phase, this statistic allows a simultaneous decision to be made on N-1 phases in accordance with the particular data phase sequence $\Delta\phi_{k-N-2}$, $\Delta\phi_{k-N-1}$, ..., $\Delta\phi_k$ that maximizes η.

To apply the notion of multiple symbol differential detection to trellis coded MPSK, the decision statistic of (7) must be associated with a *branch* in the trellis diagram. To do this, we construct a multiple trellis code of multiplicity k = N-1. Thus, we can envision the transmitted sequence, $\underline{x}$, of (2) as being partitioned into $B = N_s/k = N_s/(N-1)$ subsequences[3], i.e.,

$$\underline{x} = (\underline{x}^{(1)}, \underline{x}^{(2)}, \ldots, \underline{x}^{(B)}) \tag{8}$$

with each subsequence $\underline{x}^{(i)} = (x_{i1}, x_{i2}, \ldots, x_{ik})$ representing an assignment to a trellis branch. Similarly, a received sequence, $\underline{r}$, of length N_s is associated with a path of length B branches in the trellis diagram. Once this association is made, computation of bit error probability for the system follows along the lines of the approach taken in [3]. The details of the analysis are presented in [2].

The total metric proposed for the multiple trellis decoder with channel state information (CSI) is[4]

$$\eta = \sum_i \left| F^{(i)} \right| \left| r_{k-N+1}^{(i)} + \sum_{n=0}^{N-2} r_{k-n}^{(i)} e^{-j\sum_{m=0}^{N-n-2}\Delta\phi_{k-n-m}^{(i)}} \right|^{\ell} \tag{9}$$

5.0 EXAMPLE - 16-STATE TRELLIS CODE

Consider a 16 state, rate 2/3 trellis coded 8PSK using conventional (N = 2) and multiple (N = 3) symbol differential detection. This code, which is optimum on the AWGN, has the transition matrix [5; Fig. 7]

[3] Since N_s is arbitrary, we can choose it such that $N_s/(N-1)$ is integer.

[4] Note that in (9), the subscript "k" on F_k has been omitted since we assume that the fading amplitude is constant along any given branch consistingf of N - 1 symbols.

$$T = \begin{bmatrix} 0 & 4 & 2 & 6 & & & & & & & & & & & & \\ & & & & 1 & 5 & 3 & 7 & & & & & & & & \\ & & & & & & & & 4 & 0 & 6 & 2 & & & & \\ & & & & & & & & & & & & 5 & 1 & 7 & 3 \\ 2 & 6 & 0 & 4 & & & & & & & & & & & & \\ & & & & 3 & 7 & 1 & 5 & & & & & & & & \\ & & & & & & & & 6 & 2 & 4 & 0 & & & & \\ & & & & & & & & & & & & 7 & 3 & 5 & 1 \\ 4 & 0 & 6 & 2 & & & & & & & & & & & & \\ & & & & 5 & 1 & 7 & 3 & & & & & & & & \\ & & & & & & & & 0 & 4 & 2 & 6 & & & & \\ & & & & & & & & & & & & 1 & 5 & 3 & 7 \\ 6 & 2 & 4 & 0 & & & & & & & & & & & & \\ & & & & 7 & 3 & 5 & 1 & & & & & & & & \\ & & & & & & & & 2 & 6 & 0 & 4 & & & & \\ & & & & & & & & & & & & 3 & 7 & 1 & 5 \end{bmatrix} \quad (10)$$

The output of the encoder enters the block interleaver. We use 128 symbol interleaving with depth of 8 and span of 8 symbol pairs (i.e., symbol pair by symbol pair interleaving). The output symbols of the interleaver are differentially encoded after which they enter the 8PSK modulator with 100% excess bandwidth root raised cosine pulse shaping. The received signal enters the demodulator with Doppler correction which is also shown in Figure 1. Every 2T-sec., three samples r_k, r_{k-1}, r_{k-2} enter the triple symbol by triple symbol deinterleaver. The output of the deinterleaver enters the multiple (double) trellis decoder whose metric is computed based on (9). The decoder structure is based on the Cartesian product TxT [2].

The system in Figure 1 for the 16 state code with multiplicity two has been simulated over an additive white Gaussian noise channel with and without the inclusion of Rician fading. We assume a Rician parameter K = 10 and Doppler spreads of 20 Hz and 104 Hz. The metric used in the simulations was that given in (9) with $\ell = 2$ and no CSI, i.e., $|F^{(i)}| = 1$ for all i. The results of the simulations are shown in Figures 3 and 4 for unfaded and faded cases, respectively. We observe from these figures that in going from N = 2 to N = 3, an improvement in E_b/N_0 performance of at least 0.75 dB is obtained for the range of bit error probabilities illustrated.

6.0 CONCLUSIONS

As was true for the uncoded case, the use of multiple symbol differential detection of trellis coded MPSK can offer an improvement in error probability performance over conventional (two-symbol observation) differential detection of the same coded modulation. Again only a slight increase in the length of the observation interval is necessary to demonstrate a significant improvement both for faded and unfaded channels.

7.0 REFERENCES

[1] **D. Divsalar, M.K. Simon**, "Multiple Symbol Differential Detection of MPSK", March issue of IEEE Transactions on Communications , 1990, pp. 300 -308.

[2] **D. Divsalar, and M. K. Simon**, "The Performance of Trellis Coded MDPSK with Multiple Symbol Detection," accepted for publication in the IEEE Transactions on Communications.

[3] **D. Divsalar, M.K. Simon**, "Multiple Trellis Coded Modulation (MTCM)," JPL Publication 86-44 (MSAT-X Report No. 141), November 15, 1986. Also in IEEE Transactions on Communications, Vol. 36, No. 4, April 1988, pp. 410-419.

[4] **G. Ungerboeck**, "Channel Coding with Multilevel/Phase Signals," IEEE Transactions on Information Theory, Vol, IT-28, No. 1, January 1982, pp. 55-67.

[5] **M.K. Simon, D. Divsalar**, "The Performance of Trellis Coded Multilevel DPSK on a Fading Mobile Satellite Channel," JPL Publication 87-8 (MSAT-X Report No. 144), June 1, 1987. Also in IEEE Transactions on Vehicular Technology, Vol. 37, No. 2, May 1988, pp. 78-91.

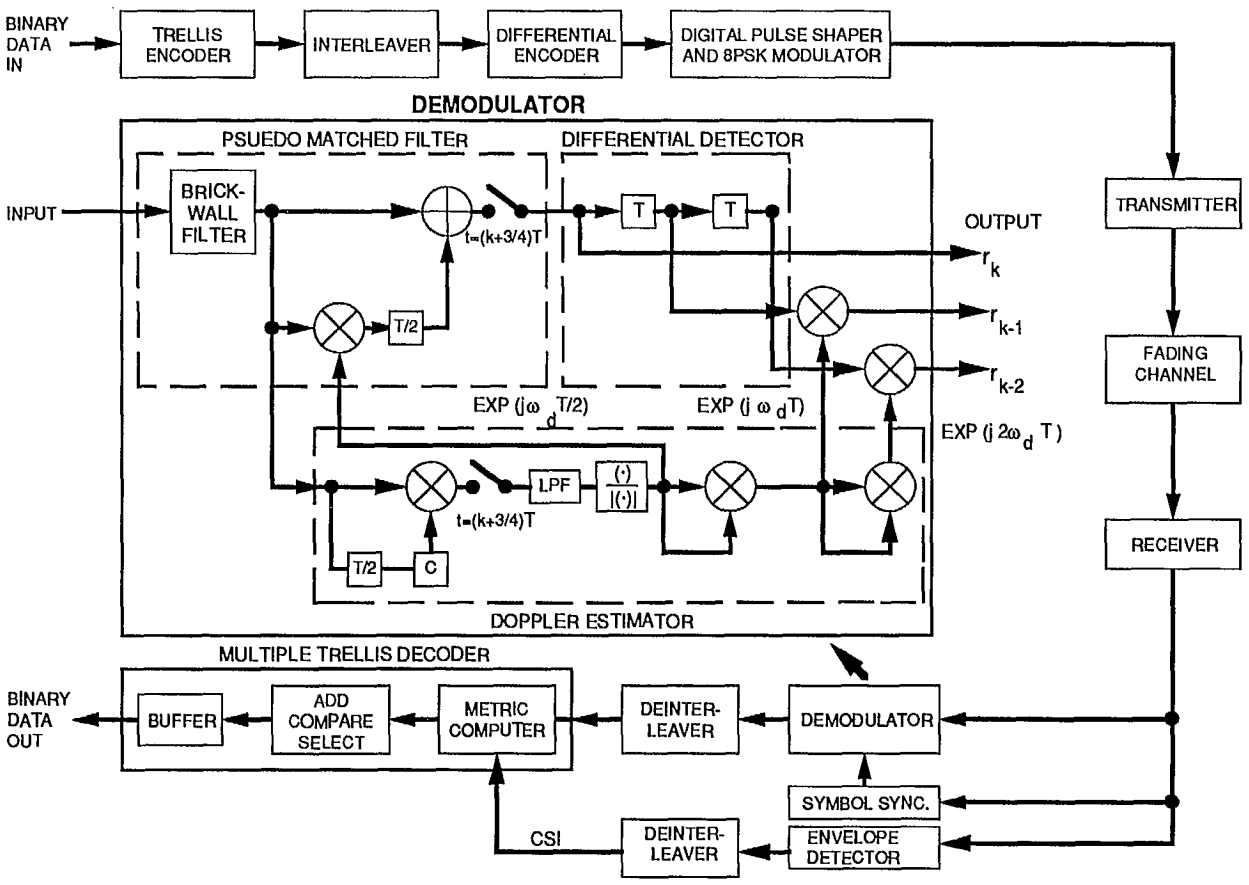
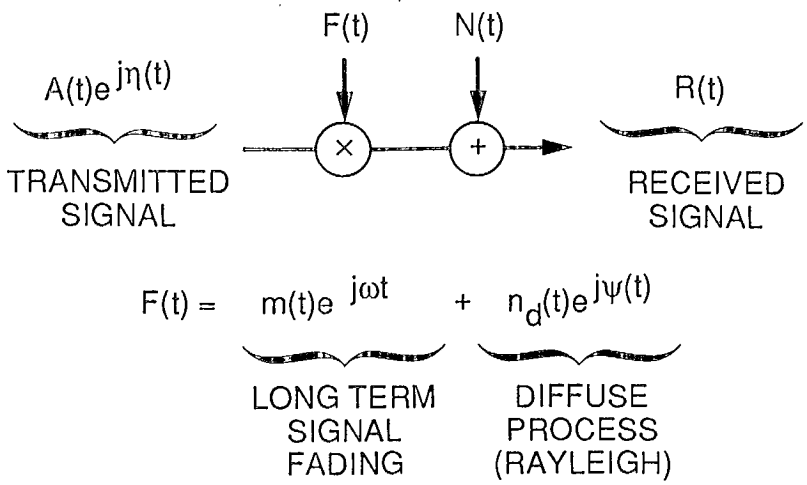
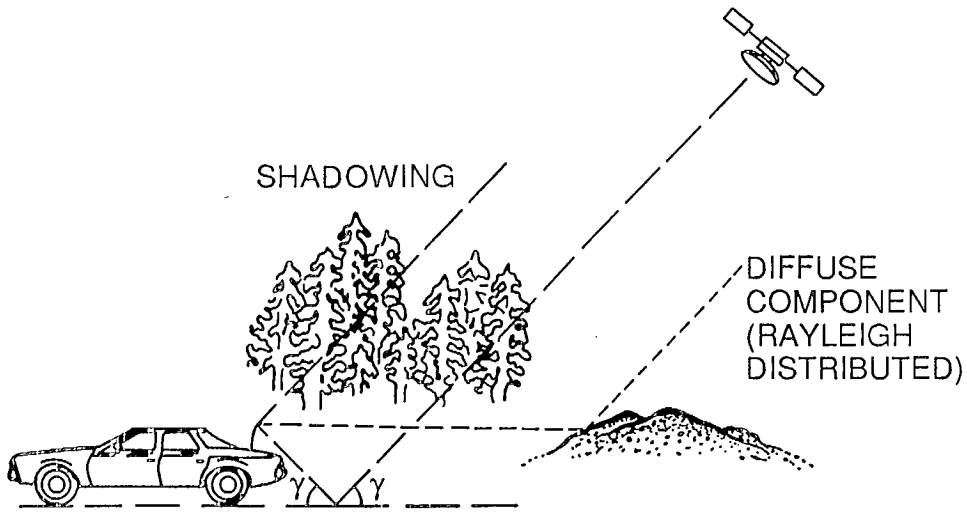


Figure 1. System Block Diagram.





COMPLEX RECEIVED FADED SIGNAL:

$$R(t) = M(t) \circ (R_{dir}(t) + R_{spec}(t)) + R_{dif}(t)$$

M(t): LONG TERM SIGNAL FADING
(LOGNORMAL DISTRIBUTED)

Figure 2. Mobile Satellite Channel Propagation Model.

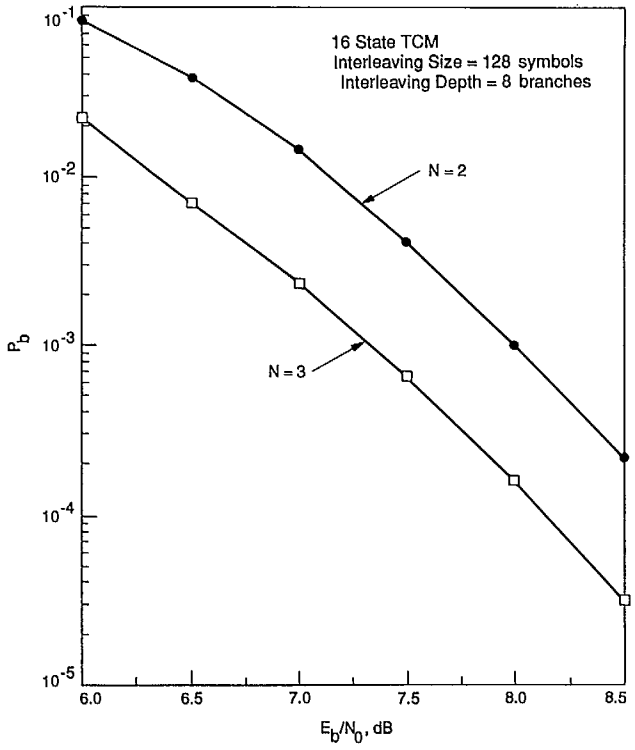


Figure 3. Simulation Results for Bit Error Probability of 16 State, Rate 2/3 Trellis Coded 8PSK With Conventional (N = 2) and Multiple (N = 3) Symbol Differential Detection (No Fading).

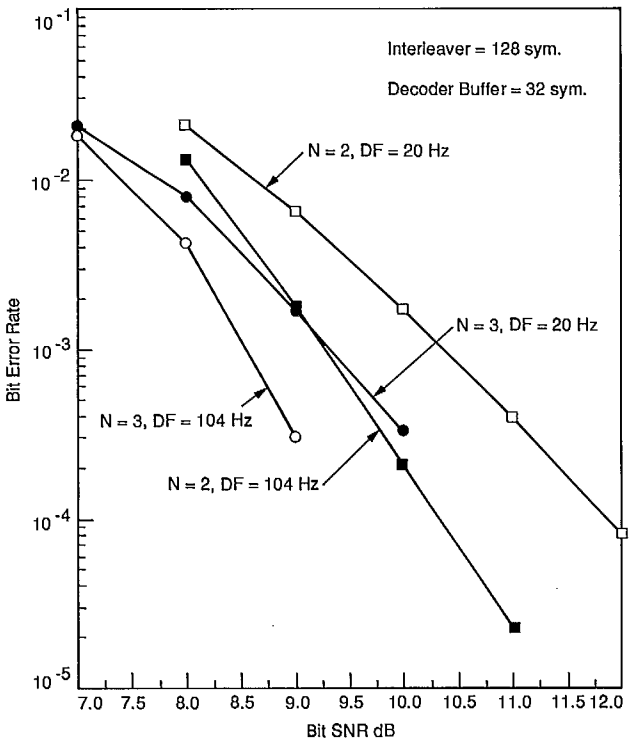


Figure 4. Simulation Results for Bit Error Probability of 16 State, Rate 2/3 Trellis Coded 8PSK With Conventional (N = 2) and Multiple (N = 3) Symbol Differential Detection (Rician Fading With K = 10).

An Implementation of a Reference Symbol Approach to Generic Modulation in Fading Channels

R.J. Young, J.H. Lodge and L.C. Pacola

Communications Research Centre, Ottawa, Canada

ABSTRACT

As mobile satellite communications systems evolve over the next decade, they will have to adapt to a changing tradeoff between bandwidth and power. This paper presents a flexible approach to digital modulation and coding that will accomodate both wideband and narrowband schemes. This architecture could be the basis for a family of modems – each satisfying a specific power and bandwidth constraint yet all having a large number of common signal processing blocks. The implementation of this generic approach, with general purpose digital signal processors for transmission of 4.8 kbps digitally encoded speech, is described.

I. INTRODUCTION

As mobile satellite communications systems evolve over the next decade, they will have to adapt to a changing tradeoff between bandwidth and power. Initial systems using existing satellites are power limited and thus power efficient modulations such as rate 1/2 coded BPSK are appropriate. Future higher power satellites are expected to be bandlimited. When these spacecraft are in service bandwidth efficient modulations such as trellis coded modulation will be more appropriate.

This evolution provides the impetus for the definition of a flexible and easily adapted modem architecture that achieves robust digital communications over fading channels. In this paper we define such an architecture that maintains a large number of common signal processing blocks while accomodating various coding rates and signal constellations. This approach to modem design will allow a modular DSP software implementation to rapidly evolve from a power efficient modulation to a bandwidth efficient modulation.

II. THE REFERENCE SYMBOL APPROACH

Robust communication in the presence of fading can be achieved by performing some form of channel estimation and compensation at the receiver. A proven method for obtaining channel estimation is achieved by transmitting a known signal along with the information bearing signal. The receiver uses this known signal to estimate the multiplicative distortions introduced by the channel and subsequently to remove them from the received information bearing signal. This technique is often referred to as feed–forward signal regeneration (FFSR). The frequency domain approach to providing a reference signal is to allocate a portion of the transmit spectrum to a pilot tone [1,2]. The disadvantages of this approach are that the position of the pilot tone can constrain the choice of modulation and/or affect the performance of the FFSR. Also, the pilot tone will disturb the envelope properties of those modulations that have been designed with approximately constant envelope for use over nonlinear channels.

Recent papers [3,4] introduced a time domain analogue of the pilot tone which has similar power and bandwidth requirements but not the disadvantages described above. In the time domain scheme a sequence of symbols that is known to the receiver is multiplexed with the coded information symbols before transmission. When synchronized the receiver demultiplexes the reference samples and uses them to obtain channel estimation and compensation. The multiplex ratio of the reference sequence is M to 1, that is, M code symbols are transmitted between each reference symbol. The reference symbol rate ($1/M$ times the code symbol rate) must be at least twice the fading rate in order that Nyquist sampling of the channel is achieved. The reference symbols use $1/(M+1)$ of the transmit power and bandwidth. This allocation is recovered when one considers that receiver can use a form of coherent detection in channels that normally require differential detection.

The reference symbols are the foundation of our generic approach to modulation and coding in fading channels. As well as providing a form of coherent detection through channel tracking they can also provide the receiver with a means of achieving its

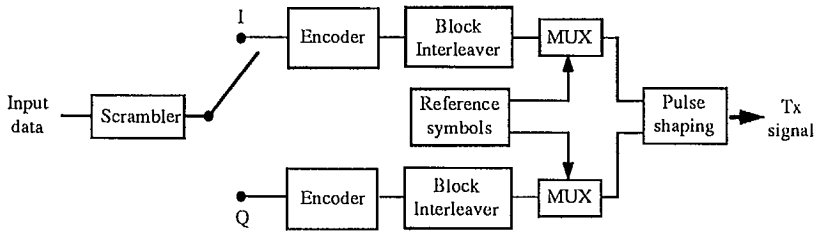


Fig.1. A block diagram of the signal processing in the transmitter.

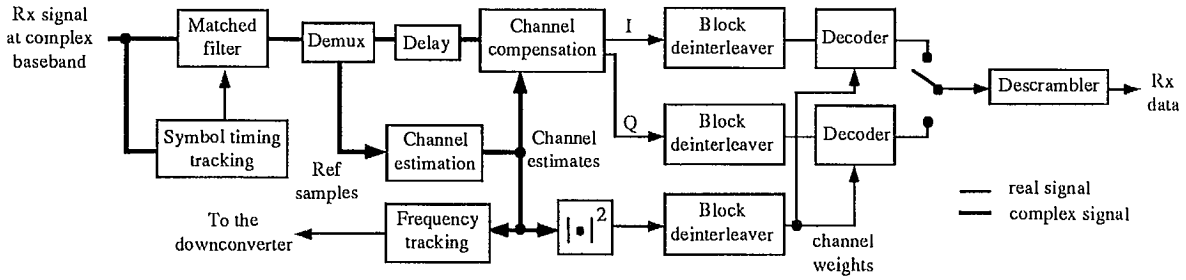


Fig. 2. A block diagram of the signal processing in the receiver.

other tracking functions (symbol timing tracking, frequency tracking and signal health tracking) independent of the signal constellation chosen. Therefore, a family of modems each with different coding rates (power efficiencies) and signal constellations (bandwidth efficiences) can be built from a common set of signal processing blocks.

A. Transmitter

The structure of the transmitter is shown in Fig. 1. The information bits are first scrambled to ensure that they have suitable randomness and then they are alternately sent to the I and Q channels. The I and Q bits are independently encoded and interleaved. Interleaving distributes the burst errors caused by the fading channels thereby improving decoder performance. The interleaver output is time division multiplexed with the reference sequence and the resulting composite symbol sequence is filtered to achieve the desired transmit spectrum.

B. Receiver

The signal processing section of the generic receiver, as shown in Fig. 2, assumes a complex baseband input signal sampled at twice the symbol rate, $f_s = 2f_r$. To recover the transmitted symbols the received signal is filtered by a filter matched to the transmitter's pulse shaping filter. The demultiplexor separates the reference samples from the code samples. The reference samples are used to estimate the channel gain and phase at the code symbol times. This channel state information is used to compensate the code samples thereby producing a "pseudocoherent" estimate of the received code sequence and to provide side information to the decoder. The I and Q code samples and the channel samples are deinterleaved and passed to the I and Q decoders. The decoded bits in the I and Q channels are recombined and descrambled to recover the transmitted data.

The signal processing to obtain the channel estimates from the reference symbols is shown in Fig. 3. The received reference samples are multiplied by the conjugate of the known transmitted sequence to produce samples of the channel at the reference symbol rate. The channel samples are lowpass filtered and then interpolated to produce estimates of the channel at each of the code sample times between the reference samples.

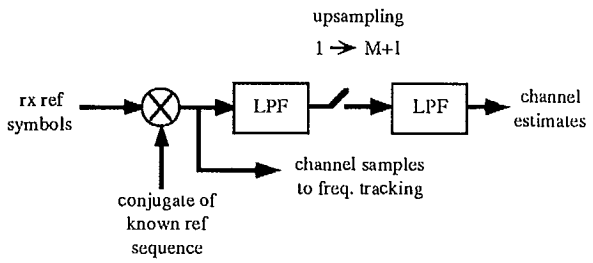


Fig. 3 Signal processing for channel estimation.

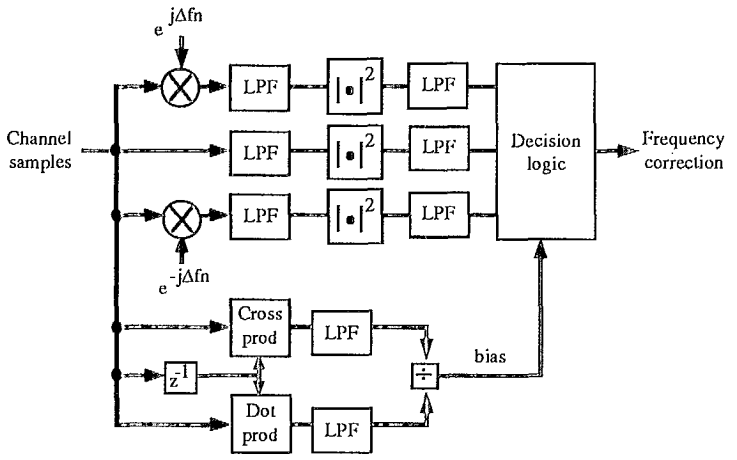


Fig. 4. A block diagram of the frequency tracking processing.

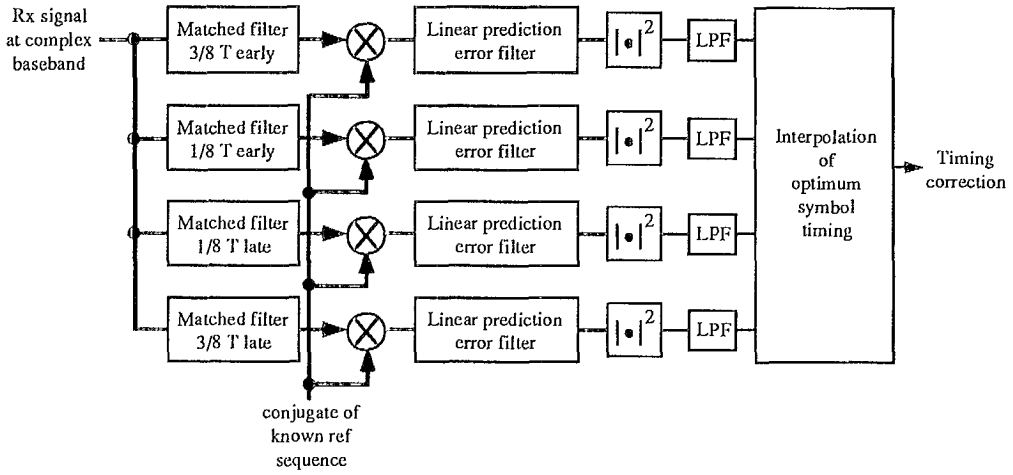


Fig. 5. A block diagram of the symbol timing tracking processing.

The receiver's frequency tracking function uses the channel samples from Fig. 3 before the noise rejection filter. The objective of the algorithm shown in Fig. 4 is to maximize the energy in the channel samples at the output of the noise rejection filter. This is done because the channel estimation and compensation process is able to make use of the multipath energy in the received signal. The channel samples are filtered in three parallel paths by frequency shifted versions of the noise rejection filter. One path contains the filter shifted by $-\Delta f$ Hz, another path contains the filter shifted by $+\Delta f$ Hz, and the final path contains the centered filter. Δf defines the resolution of the frequency tracking. The power at the output of each filter is computed and filtered. Decision logic attempts to make frequency corrections in order to maximize the power at the output of the centered filter. An additional bias signal is computed which tracks the direct path component of the channel. In static or slow fading channels the bias signal will pull the direct path towards the center of the noise rejection filter instead of allowing the received signal to wander back and forth over the maximum expected fading bandwidth.

The symbol timing tracking algorithm shown in Fig. 5 also makes use of the reference symbols. This algorithm uses an early/late detection approach to find the timing that minimizes the intersymbol interference (ISI) at the output of the matched filter. Outputs from each of a parallel bank of 4 matched filters with detection time offsets of +3/8, +1/8, −1/8, −3/8 of a symbol period respectively are obtained only during reference symbol detection periods. These outputs are multiplied by the conjugate of the known transmitted reference

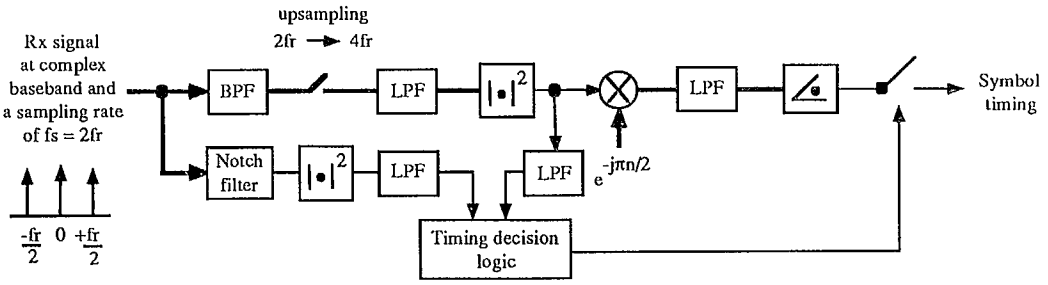


Fig. 6. A block diagram of the symbol timing acquisition processing.

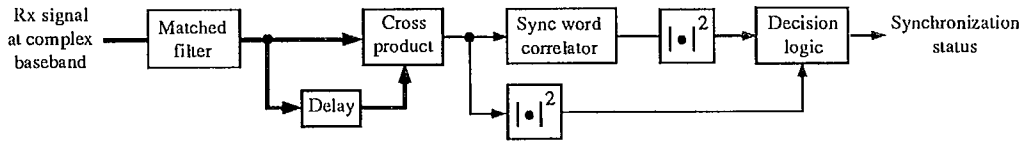


Fig. 7. A block diagram of the synchronization detection processing.

sequence. For a branch which corresponds to exact symbol timing the matched filter will detect the reference symbol without ISI and the multiplication with the known symbol will produce a signal with a bandwidth equal to the fading bandwidth. For a branch with some timing offset the matched filter output will include ISI and the following multiplication will produce a signal with a bandwidth larger than the fading bandwidth. Processing continues through a linear prediction error filter designed assuming a nominal fading spectrum. This filter acts like a highpass filter where the bandwidth of the stopband equals the fading bandwidth. The error filter output for correct timing will be small compared to the error filter output for incorrect timing. The error filter outputs are squared and filtered to produce power signals for each of the detection offsets. These power values define a detection performance versus timing offset surface where the optimum sampling time is found at the minimum. The performance surface samples are interpolated at points between $\pm 1/8$ of a symbol period timing offset to find the position of the minimum and hence the optimum timing.

C. Synchronization

The receiver requires that synchronization with the incoming sample sequence be obtained so that the demultiplexor can properly extract the reference samples and so that the block deinterleaver, decoder and descrambler can be properly initialized. Therefore, each data burst must be preceeded by a preamble that will allow synchronization to be found. Our strategy is to transmit a known timing pattern followed by a synchronization word. The timing pattern allows the receiver to obtain symbol timing before searching for the synchronization word. The signalling during the preamble was chosen to be symmetric BPSK [5] because it is well suited to transmission through non-linear channels. This choice of signalling is universally suited to preceed constant or non-constant envelope modulations and therefore fits the concept of the generic architecture.

To acquire symbol timing from the timing pattern the processing shown in Fig. 6 is used. The symbol sequence comprising the timing pattern is {1, j, 1, j, ...} and has spectral lines at $\pm f_r/2$ and 0 Hz. The received signal sampled at $f_s=2f_r$ is first bandpass filtered to extract the spectral components at $\pm f_r/2$ Hz. This filter's passband must accomodate the maximum fading bandwidth. The output of this filter is interpolated to a sampling rate of $f_s=4f_r$ and then squared. The squaring process generates spectral components at 0 and $\pm f_r$ Hz. The component at $+f_r$ Hz is down converted to complex baseband and narrowband filtered. Then if the down conversion oscillator is referenced to the assumed symbol clock the angle of the phasor at the narrowband filter output provides the symbol timing information. A timing signal presence indicator is obtained by notch filtering the discrete signal

components at the input, squaring and lowpass filtering. This produces a signal proportional to the additive noise power at the receiver input. The power in the timing signal component can be obtained by lowpass filtering the timing signal prior to the down conversion. Timing signal presence is declared when the power in the timing signal exceeds a threshold relative to the input noise power.

The synchronization detection processing, shown in Fig. 7, is engaged after symbol timing has been declared. With symbol timing already established the transmitted symbols are filtered by a matched filter and differentially detected because the channel samples are not available yet. The resulting soft decisions are passed to the synchronization word correlator. The power at the output of the correlator is compared to the total signal power in the correlator delay line and when this ratio exceeds the detection threshold, synchronization is declared. At this point the reference sample demultiplexor, the decoder, the deinterleaver and the descrambler can be initialized.

III. IMPLEMENTATION

The modem for a 4.8 kbps digital voice terminal is being implemented on TMS320C25 digital signal processors (DSPs) using the reference symbol approach. Two versions of the terminal are currently planned. The first version is bandwidth efficient and intended for operation in 5 kHz channels. The second version will be power efficient and intended for operation in 20 kHz channels. Our approach to the software implementation on the DSPs is to follow a modular structure that will allow easy migration from the narrowband to the wideband version. The only modules that will need major modification are the coding/decoding modules. The acquisition and tracking modules are independent of the signal constellation, but, minor modifications will be required for optimization at different channel rates and signal-to-noise ratios.

The modem's block structure is matched in size and synchronized to the frame structure of the voice codec. This has several advantages. In the transmitter the preamble incurs no additional delay because it can be transmitted while the block interleaver is begin filled. In the receiver, techniques can be used to smooth or mask the undesireable effects of signal blockages or deep fades. The increase in bit errors during these periods often causes the codec to output objectionable audio and sometimes to lose frame synchronization. However, the channel state information can be used by the codec interface software to substitute repeated or known silence frames during poor channel periods. As well as reducing the potential for unwanted audio output, this approach allows faster recovery from blockages because the codec's frame synchronization is maintained throughout by the modem. Of course, this scheme requires knowledge of the codec's frame format.

The digital voice terminal is capable of operating in a voice activated fashion. Each burst of voice data is preceeded by one frame of preamble. When the receiver is not currently receiving a transmission it searches continuously for the preamble. To prevent the loss of an entire transmission if the preamble is missed, the reference pattern is used as a secondary means of synchronization since it has been designed to repeat every block. A search for the reference pattern is done in parallel with the preamble search and either process can synchronize the receiver.

The narrowband version is almost complete. The transmitter is implemented using a single C25 processor although it takes less than half of the available processing power. The transmitter's encoder is a 32 state, rate 1/2 Calderbank and Mazo (C&M) trellis code [6]. Each of the input bits in the I and Q paths at 2400 bps produce a 4 level code symbol at the encoder output. The reference symbol multiplex ratio is $M=4$ which results in a reference symbol rate of 600 reference symbols per seconds. At this rate the receiver can track fading channels with bandwidths up to 200 Hz. After multiplexing, a symbol rate of 3000 symbols per second is achieved. These complex symbols which are contained in a 16 QAM constellation are filtered with 50 % square root raised cosine pulse shaping filters before transmission. The transmitted spectrum is shown in Fig. 8.

The receiver is implemented using two C25 processors. The first processor is responsible for down conversion of the received signal to complex baseband and decimation to 2 samples per symbol, matched filtering, reference symbol demultiplexing, channel compensation and estimation, symbol timing tracking, frequency tracking, symbol timing acquisition and synchronization word detection. The second processor is responsible for deinterleaving, decoding, descrambling and the codec interfacing. The simulated performance of the receiver is shown in Fig. 9 for static and $k=-5$ dB, 60 Hz Rician channels. Three different trellis codes were simulated: the 32 state C&M trellis code, an 8 state C&M trellis code and a 64 state pragmatic trellis code [7]. These results show that, in the severely fading channel, an $E_b/N_o=7.3$ dB ($C/N_o=$ 44

dBHz) is required to achieve a bit error rate of 10^{-3} using the 32 state code. A bit error rate of 10^{-3} is the threshold where most voice codecs begin to suffer performance degradation from bit errors. While these are ideal results they do include the channel estimation and compensation processing which is equivalent to carrier recovery in conventional coherent systems. Therefore, any comparison to a other coherent techniques must include their assumed carrier recovery loss.

IV. CONCLUSIONS

A flexible modem architecture has been presented which provides robust performance in fading channels through the use of reference symbols. By using the reference symbols for the receiver's tracking functions, quick migration from one modulation and coding strategy to another is possible. This architecture could be the basis for a family of modems – each satisfying a specific power and bandwidth constraint yet all having a large number of common signal processing blocks.

V. REFERENCES

[1] J. P. McGeehan, A. J. Bateman, "Phase-locked transparent tone-in-band (TTIB)", IEEE Trans. Commun., vol. COM-32, pp. 81-87, Jan. 1984.

[2] F. Davarian, "Mobile digital communications via tone calibration", IEEE Trans. Vehic. Technol., vol. VT-36, pp. 55-62, May 1987.

[3] M. L. Moher, J. H. Lodge, " TCMP – A modulation and coding strategy for Rician fading channels", IEEE Journ. Sel. Areas in Commun., vol. 7, no. 9, pp. 1347-1355, Dec. 1989.

[4] M. L. Moher, J. H. Lodge, "Time diversity for mobile satellite channels using trellis coded modulations", Globecom, 1987.

[5] J. H. Winters, "Differential detection with intersymbol interference and frequency uncertainty", IEEE Trans. on Commun., vol. COM-32, no.1, pp. 25-33, Jan. 1984.

[6] R. Calderbank, J. E. Mazo, "A new description of trellis codes", IEEE Trans. Info. Th., vol. IT-30, no. 6, Nov. 1984.

[7] A. J. Viterbi, J. K. Wolf, E. Zehavi, R. Padovani, "A pragmatic approach to trellis-coded modulation", IEEE Commun. Mag., pp. 11-19, July 1989.

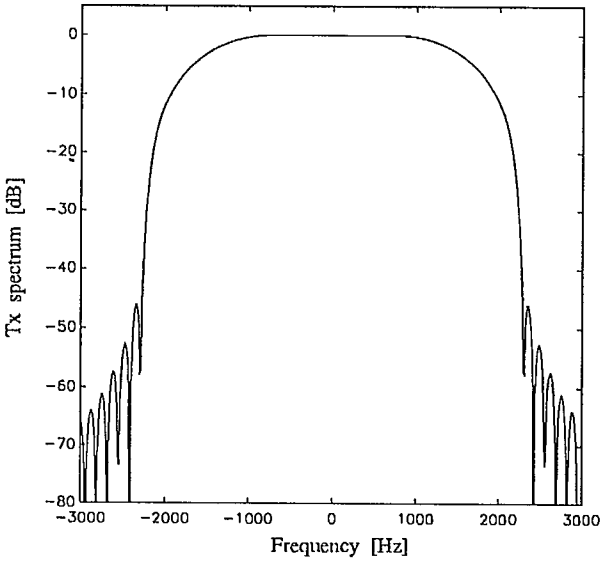


Fig. 8. The transmit spectrum of the narrowband modem.

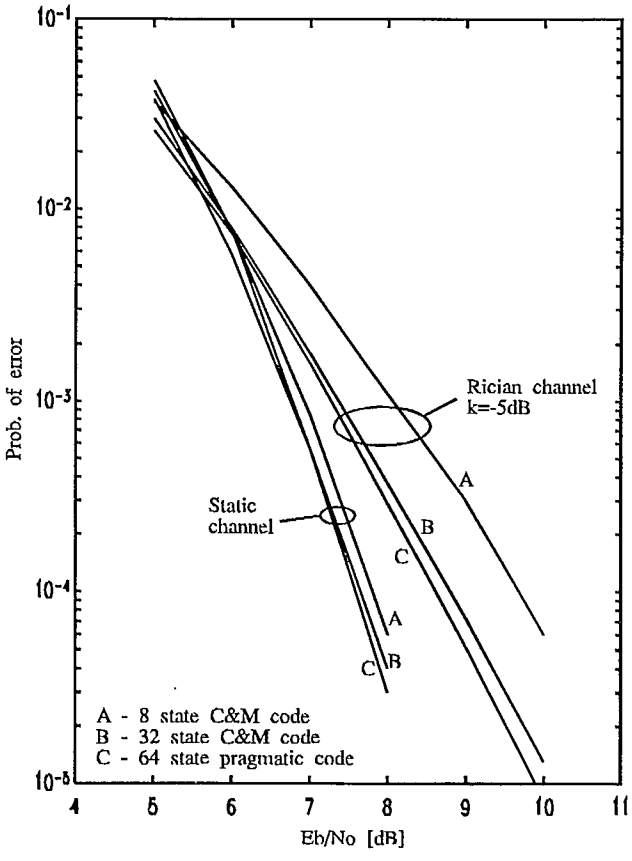


Fig. 9. The simulated BER performance of the reference symbol receiver.

Acoustic Echo Cancellation for Full-Duplex Voice Transmission on Fading Channels

Sangil Park and Dion D. Messer

Motorola Inc.
Digital Signal Processing Operations
Austin, TX 78735

ABSTRACT

This paper discusses the implementation of an adaptive acoustic echo canceler for a hands-free cellular phone operating on a fading channel. The adaptive lattice structure, which is particularly known for faster convergence relative to the conventional tapped-delay-line (TDL) structure, is used in the initialization stage. After convergence, the lattice coefficients are converted into the coefficients for the TDL structure which can accommodate a larger number of taps in real-time operation due to of its computational simplicity. The conversion method of the TDL coefficients from the lattice coefficients is derived and the DSP56001 assembly code for the lattice and TDL structures is included, as well as simulation results and the schematic diagram for the hardware implementation.

1.0 Introduction

Adaptive signal processing for echo cancellation structures has a variety of usages in *telecommunication* applications due to *multi-path* and *impedance mismatches* in communication channels. Echo cancellation is required especially for full-duplex voice transmission where the microphones and speakers are located in places such that an acoustic echo is created. One such application is a *hands-free cellular phone* which allows full duplex operation by preventing the phone from breaking into oscillations [1]. The ability to provide hands-free operation of cellular (mobile) phones offers users a safer and more convenient way to use their cellular phones while driving a car as shown in Figure 1.

In the cellular phone application, there needs to be two echo cancellers in the system, one to cancel the phone line (electrical) echo and the other to cancel the acoustic echo, which is the signal from the loudspeaker echoed back into the microphone. In this paper only the acoustic cancellation problem is considered. Figure 2 shows the model and the hardware schematic of the acoustic echo canceller. The adaptive algorithm shown as the adaptive digital filter (ADF) block in Figure 2 minimizes the error signal which is the difference between the actual transmitted signal and the estimated transmitted signal by the linear combination of the received data set. When the error terms are minimized the adaptive filter impulse response is said to have converged to the impulse response of the echo paths.

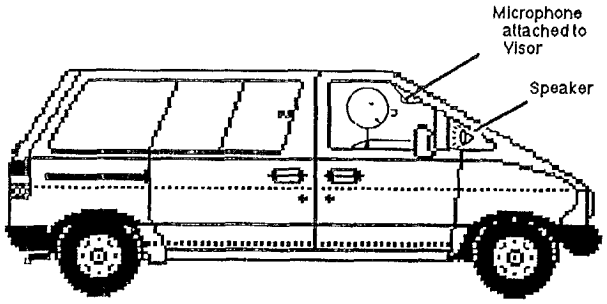


Figure 1 Depiction of a proposed hands-free cellular phone system

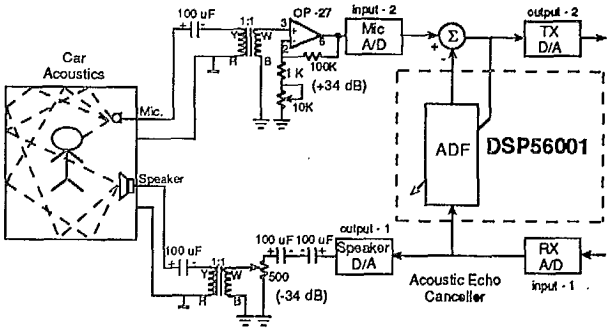


Figure 2 Block diagram of acoustic echo canceller

An implementation of the acoustic echo canceller for a *speakerphone* application to improve perfor-

mance was introduced by one of the authors [2]. The previous paper uses the DSP56200 cascadable adaptive FIR filter peripheral which implements the conventional TDL structure with the least-mean-square (LMS) algorithm for adaptation [3]. An acoustic echo canceller needs an initialization period (training) before the phone can work properly for full-duplex operation. This may require up to 5 seconds of initialization time depending on the convergence property of the adaptive algorithm. Although the TDL structure [3,4] is a simple and commonly used stochastic approximation-type algorithm, the convergence time is slow, especially if the training signal is narrowband or has band-limited spectral content. Thus, pseudo-random noise that has a broad frequency spectrum is normally used for initialization. However, this type of random noise creates problems to the user who will be able to hear it during the initialization. The user will most likely interpret this noise to be a bad (or no) connection and will hang up before initialization is complete. In adaptive filtering applications such as hands-free cellular phone, therefore, very rapid convergence of the adaptive coefficients is a requirement.

The lattice structure based on the adaptive LMS algorithm has been widely accepted for applications where rapid transient adaptation is required and/or the eigenvalues of the input signal are highly disparate [5,6]. The lattice structure can be interpreted to be self-orthogonalizing which has been shown to speed up convergence. Cellular phones are normally used inside a car which has smaller acoustic reverberation and echo paths compared to the size of a office or conference room in which a conventional speakerphone must function. Thus, a fewer number of taps, which represents the time-window for the adaptive approximation, than the conventional TDL structure can be used for the adaptive process. However, due to the computational complexity of the lattice-LMS algorithm it is hard to apply a large number of coefficients (stages) to accommodate 0.1 second (which is a time window of 800 taps) of acoustic echo delay in real-time. Thus, the lattice structure is used only for the initialization stage and the coefficients are converted to the TDL structure. The TDL structure is computationally efficient and can accommodate larger number of coefficients in real-time to cancel the long delayed echoes.

2.0 Acoustic-Echo Canceller Model

Depending on the characteristics of the car's internal acoustics, the echo may be sufficiently strong, such that this echo must be removed at the microphone input. The term used to describe the amount of echo which can be removed by the echo canceller is Echo Return Loss Enhancement (ERLE) and can be defined as [2]:

$$\mathbf{ERLE}\,(dB) \;=\; 10\log_{10}\left[\frac{\mathbf{E}\,[y\,(k)^2]}{\mathbf{E}\,[e\,(k)^2]}\right] \qquad (1)$$

where $\mathbf{E}\,[y\,(k)^2]$ and $\mathbf{E}\,[e\,(k)^2]$ are the expected values of microphone input signal power and uncanceled echo signal power, respectively, as shown in Figure 2. The desired maximum amount A goal for ERLE is 30 dB due to ambient noise which is not created by the echo itself [3].

Due to advances in CMOS process technology, inexpensive adaptive digital filters are readily available. The DSP56001 can run upto 830 taps of a TDL-LMS adaptive filter at 8 kHz samples per second with 24-bit data and coefficients. As is shown in the following section, ERLE is a function of many parameters including the number of taps and the precision of the coefficients.

3.0 Echo Cancellation Algorithms

In this section, two adaptive algorithms are described with particular emphasis on echo cancellation applications.

3.1 Adaptive TDL-LMS algorithm

Figure 3 shows a block diagram of the adaptive echo canceler model which uses a TDL structure to provide adaptive coefficient adjustment. If $h_i\,(k)$ are the filter coefficients, $R_{xx}\,(k)$ is the auto-correlation matrix of the received line signal $x(k)$ at time k, and $R_{xy}\,(k)$ is the cross-correlation vector between the received signal $x(k)$ and the echo signal $y(k)$, then the optimum filter coefficient vector that minimizes the expected value of $e^2\,(k)$ in Figure 3 is given by [3]

$$\hat{H}\,(k) \;=\; R_{xx}^{-1}\,(k)\,R_{xy}\,(k) \qquad (2)$$

where $H\,(k)$ is an N-element vector consisting of the filter coefficients at time k as

$$H\,(k) \;=\; \left[h_0\,(k)\;\; h_1\,(k)\;\; \ldots\ldots\;\; h_{N-1}\,(k)\right]^T \qquad (3)$$

and T denotes matrix transpose. The coefficients $h_i\,(k)$ are updated to minimize the error signal (residual echo), $e(k)$, which is the transmitting line signal from the echo canceler. $e(k)$ can be expressed as

$$e\,(k) \;=\; y\,(k) - H^T\,(k)\,X\,(k) \qquad (4)$$

where $X(k)$ is the input data vector given by

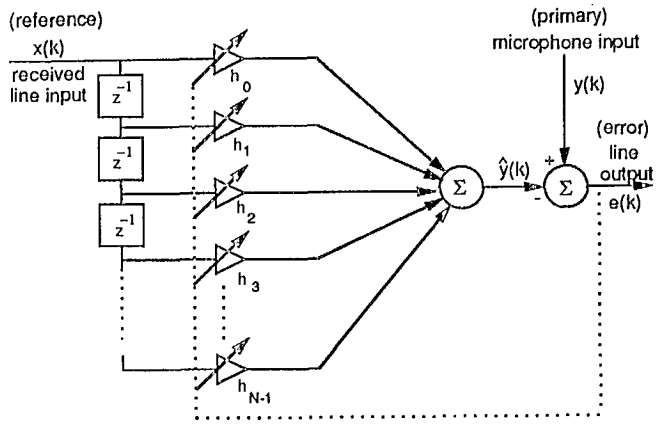


Figure 3 The tapped-delay-line (TDL) structure

$$X^T(k) = \left[x(k)\ x(k-1) \ldots\ldots\ x(k-N+1)\right] \quad (5)$$

The LMS algorithm, which is one implementation of the steepest decent method, updates the weight vector, $H(k)$, at each k via the following relation

$$H(k) = H(k-1) + \mu e(k) X(k) \quad (6)$$

where μ denotes the loop-gain factor (convergence parameter). The adaptive algorithm forces the error term toward zero. When the error terms are minimized, the adaptive filter impulse response is said to have converged to the impulse response of the echo path.

Convergence properties and stability aspects of the LMS algorithm have been well documented [4,7]. The general conditions in practice for the loop-gain factor is

$$0 < \mu < \frac{1}{tr\,[R_{xx}]} \quad (7)$$

where $tr\,[R_{xx}]$ denotes the trace of R_{xx}. The optimized DSP56001 assembly code for the TDL-LMS algorithm in (6) can be written as[8]:

```
    clr   a          x0,x:(r0)+y:(r4)+,y0 ;clear a,x0=x(n)
    move  x:(r0)+,x1y:(r4)+,y0            ;x1=x(n-1),y0=h(0)
    do    #N/2,lms                        ;do N/2 times
    mac   x0,y0,a    y0,b       b,y:(r5)+ ;a=h(0)*x(n),b=h(0)
    macr  x1,y1,b    x:(r0)+,x0 y:(r4)+,y0 ;b=h(0)+e*x(n-1)
                                           ;x0=x(n-2),y0=h(1)
    mac   x1,y0,a    y0,b       b,y:(r5)+ ;a=a+h(1)x(n-1),b=h(1)
    macr  x0,y1,b    x:(r0)+,x1 y:(r4)+,y0 ;b=h(0)+e*x(n-1)
                                           ;x0=x(n-3),y0=h(1)
lms
    move  b,y:(r5)+                        ;save new coeffs.
    move  (r0)-n0                          ;pointer update
```

where **r0** is the register pointing to the input buffer which is modulo-addressed to accommodate 768 (N) current data points. **R4** and r5 registers are pointing to the even and odd numbered current adaptive coefficients locations, respectively. The Modifier Registers, **m0**, **m4** and **m5** are set to be 767 (N-1), 383 ($N/2$-1) and 383 ($N/2$-1), respectively. This TDL-LMS algorithm requires only $2N+2$ instruction cycles per sample period. When 768 taps are used for an acoustic echo-canceler the processing requirement at 8 kHz of sampling rate is 12.3 million instructions per second (MIPS).

3.2 Adaptive Lattice-LMS Algorithm

The lattice predictor (often called as one-step predictor) structure was originally proposed by Itakura and Saito [9] for speech analysis. The one-step predictor has also been extended to a noise-canceler configuration as shown in Figure 4 [5]. If the inputs $x(k)$ and $p(k)$ are stationary, then it can be shown that the respective steady-state values of $e^2(k)$ and $v_N^2(k)$ for TDL and lattice models are the same. The filter model in Figure 4 consists of 3 stages which can be extended to M stages for mathematical analysis purpose. Its upper half (solid lines) is simply the (one-step) predictor model [6]. The lower portion (dashed lines) consists of M additional coefficients, v_l', $1 \le l \le M$. The basic idea involved in obtaining an adaptive algorithm is to continuously adjust the lattice weights $v_l(k)$ and $v_l'(k)$. The $v_l(k)$ are adjusted to minimize the instantaneous error $e_l^2(k) + w_l^2(k)$ via the one-step predictor LMS algorithm as

$$v_l(k+1) = v_l(k) + (1-\beta)\,[e_l(k)\,w_{l-1}(k) + w_l(k)\,e_{l-1}(k)] \quad (8)$$

for $1 \le l \le M$, which we refer to as the lattice-LMS equation for a one-step predictor. In practice, a convenience choice for a β is $\beta = 1-\mu$ in (7).

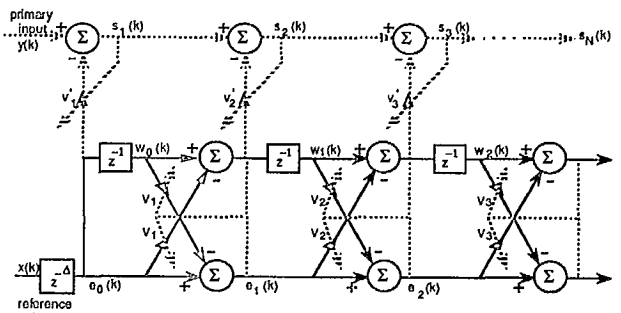


Figure 4 Lattice structure for noise-canceller

The lattice filtering computation at each stage gives successive orthogonalization process. Thus, the

successive coefficients can be optimized independent of coefficients at other stages. As a result of this orthogonalization, the convergence rate of the lattice-LMS algorithm is not restricted by the eigenvalue structure of the input signal, which is not the case for the TDL-LMS algorithm.

Next, another set of coefficients, $v'_l(k)$ in Figure 4 provides the noise-cancelling substraction paths. The individual coefficients are adjusted to minimize the filter error $s_l^2(k)$ using a technique similar to (6). Thus we have [10]

$$v'_l(k+1) = v'_l(k) - \mu\left[\frac{\partial s_l^2(k)}{\partial v'_l(k)}\right], \quad 1 \le l \le M \qquad (9)$$

where μ is a convergence parameter. Again from Figure 4 it follows that

$$s_l(k) = s_{l-1}(k) - v'_l(k)\, w_{l-1}(k+1) \qquad (10)$$

with $s_0(k) = p(k)$. Substitution of (10) into (9) leads to

$$v'_l(k+1) = v'_l(k) - 2\mu s_l(k)\, w_{l-1}(k+1) \qquad (11)$$

for $1 \le l \le M$, The DSP56001 assembly code for the lattice-LMS adaptive filter algorithm in (9)-(11) can be written as:

```
dofilt
   move  y:xdatain,b          ; read in x-input
   move  y:ydatain,a          ; read in y-input
   move  b,x:(r0)+ a,y:(r6)   ; put input in array
   move  b,x:(r2)             ; move input to memory
   do    #order,endloop
   move  b,y:(r7)             ; store previous err. in memory
   move  y:(r5),y0            ; put v' in y0 for calculation
   move  x:(r2),x1 y:(r6),a   ; put wn in x1,s state into a
   macr  -y0,x1,a y:(r4),y0   ; a=sn=sn-1-v'wn(n+1),v in x0
   move  x:(r0),a  a,y:(r6)   ; put wn in a, store sn
   move  b,y1                 ; move error into y get en-1
   macr  -y0,y1,a a,x0        ; a=wn-v*en=wn+1, wn into x0
   macr  -x0,y0,b a,x:(r0)+   ; b=en+1=en-v*wn,store wn+1
   move  a,x:(r2)             ; store wn+1
   move  x:(r3),x1            ; move 2*μ into x1
   mpy   x1,x0,a  y:(r6),y1   ; a=2*μ*wn, s state into y1
   move  a,x1     y:(r5),a    ; move a into x1, k' into a
   macr  x1,y1,a  x:(r7+n7),y0 ; v'n+1=v'+5*2*mu*wn
   move  x:(r0),x1 a,y:(r5)+  ; move wn+1 into x1, store k1
   mpy   x1,y0,a  b,x1        ; a=wn+1*en-1(n-1), e into x1
   macr  x0,x1,a  x:(r3),x0   ;a=wn-1en-1(n-1)+enwn,2μ in x0
   move  a,x1     y:(r4),a    ; move a into x1, k into a
   macr  x0,x1,a  y:(r7)+,y0  ; a=vn(new)=vn(old)+2μ(a)
   move  a,y:(r4)+            ; store vn(new)
endloop
   move  x:(r0)-,a            ; output from filter
   move  y:(r6),y0            ; output sn
   move  a,y:filtout
   move  y0,y:errout
   jmp   dofilt
```

where **r0** points to the stored filter coefficients, w_n. The buffer for **r0** is 65 ($M+1$) locations and is modulo addressed. The extra location is used because new values of the filter for the next time period are calculated before they are used in the present time period. The **r4** and **r5** registers point to v and v', respectively. Both are buffers of 64 (M) locations and are modulo addressed. The **r7** register points to the e_n values and is 128 ($2M$) locations to store two time periods of error values. The Modified Registers are used; **m0** is set to 64 (M); **m4** and **m5** are set to 63 ($M-1$); and **m7** is set to 127 ($2M-1$). This stage requires only 1280 instruction cycles per sample period, which yields a processing requirement of 10.25 MIPS at 8 kHz of sample rate.

4.0 Echo Characteristics of Car Interior

The acoustic path can be considered as a multi-reflection medium with an impulse response duration. Thus, the typical acoustics inside a car may have practically an infinite number of reflections which have different acoustical filtering effects with an exponentially decaying reverberation effect superimposed. In a typical car with reasonable acoustic treatment the reverberation time can be 0.1-0.15 seconds to reach the reverberation signal level decreased by 10 - 20 dB. However, when the car is moving, the background noise level due to the noises from engine and road may be high enough that the background noise can not be distinguished from uncancelled echoes due to an insufficient number of taps in the adaptive filters.

Echo characteristics can be measured by collecting reverberation and echo responses synchronized by an impulse output signal. The impulse can be generated in software and converted into an analog signal by a D/A converter followed by amplification to yield the audible impulse signal. Using a microphone, the residual analog signal can be converted to a digital signal by an A/D converter such as the DSP56ADC16 (16-bit Sigma-Delta A/D converter). Thus, an echo signal can be characterized by the impulse response of an acoustic chamber, or precisely, the paths from the loudspeaker to the microphone.

In this paper a simulated impulse response is used in order to characterize the convergence properties of both TDL and lattice structure. Figure 5 shows a simulated impulse response of a medium size car.

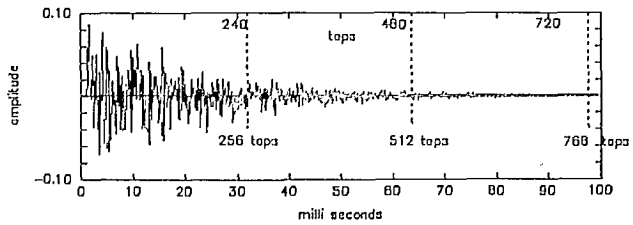


Figure 5 Impulse response of a simulated echo path

6.0 Simulation Results

A computer simulation was performed on a SUN-3/160 workstation, which modeled the system shown in Figure 2. It was assumed that the received signal was white Gaussian noise. The TDL and lattice algorithm in (7), (10) and (11) were used for the simulation. The variables in the simulation were μ, N, and M.

It has been shown that the convergence parameter μ, controls the convergence rate and the mean-square-error (MSE) of the adaptive system [4]. The constraints on the choice of μ are given by (7). The lattice-LMS and TDL-LMS algorithms are compared in terms of the MSE criteria which corresponds to the uncancelled echo. Figure 6 illustrates the MSEs of the algorithms when $\mu = 0.001$, $N = 768$ and $M = 64$. Note that the lattice-LMS algorithm converges much faster than the TDL-LMS algorithm. However, since the lattice structure used only 64 stages compared 768 taps for the TDL structure the uncancelled echo (error) is much larger than the counterpart.

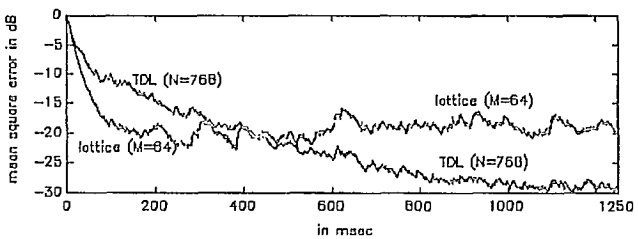


Figure 6 Mean-square errors for the lattice and TDL

After fast initialization using the lattice-LMS algorithm as shown in Figure 6, the coefficients are converted into TDL-LMS coefficients. Consider the equivalent M-TDL taps, defined in (3), from a set of M-lattice coefficients in (8) and (11). When $\hat{y}_M(k)$ is an estimate of $y(k)$ as shown in Figure 3, the corresponding error can be written as

$$s_M(k) = y(k) - \hat{y}(k) = y(k) - \sum_{i=0}^{M-1} h_{i,M} x(k-i) \quad (12)$$

where $h_{i,M}$ denotes the ith equivalent TDL tap when M is the total number of taps. Minimizing $e_N^2(k)$ with respect to the $h_{i,N}$ (new notation of h_i for the following derivation purpose), we can derive the following recursive algorithm to find a set of equivalent TDL taps using matrix bordering technique [11].

$$h_{i,L+1} = h_{i,L} + \nu'_{L+1} \alpha_{L+1-i,L} \quad , i=0,1,\ldots L \quad (13)$$

where

$$\alpha_{i,L+1} = \alpha_{i,L} + \nu_{L+1} \alpha_{L+1-i,L} \quad , i < L+1 \quad (14)$$

and $\alpha_{L+1,L+1} = -\nu_{L+1}$, $i=L+1$.

The recursion algorithm in (13) and (14) has to extend from $L=1$ through $L=M-1$ to find a set of M-TDL taps. The rest of the N-M coefficients in the TDL structure should also be initialized with zero before starting the adaptive process with the TDL-LMS algorithm.

Figure 7 shows the ERLE plots (defined in (1)). In order to smooth the output of adaptive filter the following smoothing functions were used [12].

$$E[y^2(k)] = \beta E[y^2(k-1)] + (1-\beta) y^2(k) \quad (15)$$

$$E[e^2(k)] = \beta E[e^2(k-1)] + (1-\beta) e^2(k) \quad (16)$$

where β=0.99 is the smoothing parameter. Note that the ERLE increases very rapidly at the initialization period. After the adaptive process is converted from lattice to TDL structure at $t=150$ ms, the ERLE increases slowly to the optimum solution. A total of 10,000 samples, corresponding to 1.25 seconds, were plotted to show the adaptive process when $\mu = 0.001$, $N = 768$ and $M = 64$.

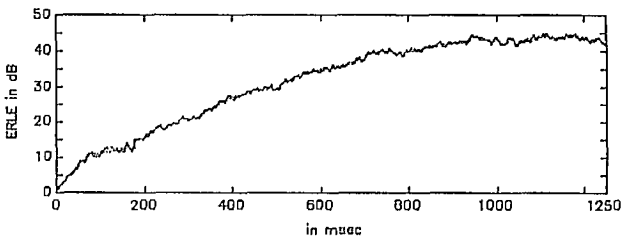


Figure 7 ERLE of the hybrid structure

7.0 Hardware Implementation Set-up

A block diagram of the hardware test implementation is shown in Figure 8. The SUN-3/160 workstation downloads assembled software into the DSP56001 Application Development System (ADS) which, in turn, controls the Ariel ADC56000 card. The ADS contains a DSP56001 general purpose digital signal processor chip which runs software in real

time. The Ariel card has dual A/D and D/A converters, which convert the signals of the loudspeaker, microphone and the receive/transmit lines. This implementation allows real-time testing of the adaptive filter concepts discussed previously.

The DSP56001 is a Harvard Architecture digital signal processor which has separate program and data memories as well as buses. It currently executes one instruction in 75 *ns* which means a 768-tap TDL-LMS adaptive filter can be performed in only 115.2 μs. This 13.5 MIPS rating is somewhat deceiving because, due to the dual buses and memories, more than one operation occurs in each instruction cycle.

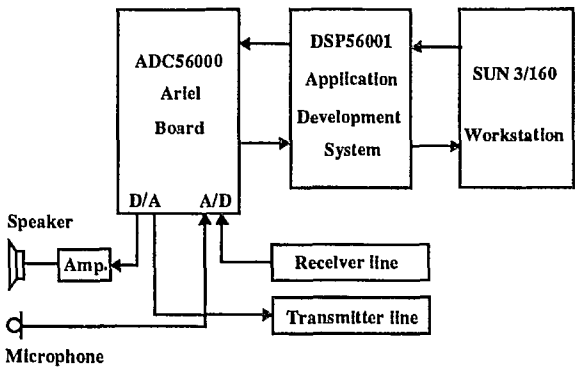


Figure 8 Hardware system set-up for experiment

8.0 Conclusions

The feasibility of implementing a full duplex hands-free cellular phone using one DSP56001 to cancel acoustic echo has been presented. Fast convergence has been achieved during the initialization stage with the lattice-LMS algorithm. After the lattice coefficients are converted to the conventional TDL structure which has 768 taps, better than 30 dB of acoustic ERLE can be theoretically achieved using a single DSP56001 by taking advantage of the 24-bit coefficient precision. The experimental set-up which will be used to verify these predictions was also described. It is hoped that sufficient and fast echo cancellation performance can be achieved by controlling the hybrid timing and the convergence parameters with this hybrid (lattice-TDL) structure.

REFERENCE

[1] K.H. Mueller, "New Digital Echo Canceller for Two-wire Full-duplex Data Transmission," *IEEE Trans. Commun.*, Vol. COM-24, pp. 956-962, Sept. 1976.

[2] S. Park and G. Hillman, "On Acoustic Echo-Cancellation Implementation with Multiple Cascadable Adaptive FIR Filter Chip," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 952-955, Glascow, Scotland, May 1989.

[3] B. Widrow, *et al.*, "Adaptive Noise Cancelling: Principles and Applications." *Proc. IEEE*, Vol. 63, No. 12, pp. 1692 - 1716, Dec. 1975.

[4] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[5] L.J. Griffiths, "A Continuously-Adaptive Filter Implemented as a Lattice Structure," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 683 - 686, March 1977.

[6] L.J. Griffiths, "An Adaptive Lattice Structure for Noise-Cancelling Applications," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 87 - 90, March 1978.

[7] T.C. Hsia, "Convergence Analysis of LMS and NLMS Adaptive Algorithms," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 667 - 670, Boston, Massachusetts, 1983.

[8] Motorola Inc., *DSP56000/DSP56001 Digital Signal Processor User's Manual*, Rev. 2, 1990.

[9] F. Itakura and S. Saito, "Digital FIltering Techniques for Speech Analysis and Synthesis," *Proc. 7th Int. Conf. Acoust.*, Budapest, pp. 261 - 264, 1971.

[10] S. Park and C. Chen, "Time-domain Algorithms for Adaptive Signal Detection Structures," *Proc. of the 22-nd Annual Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 17-19, 1988.

[11] N. Ahmed and R.J. Fogler, "A matrix bordering approach for deriving lattice models," *IEEE Trans. on Aerospace and Electronics Systems*, Vol. AES-20, No. 6, pp. 835-838, Nov. 1984.

[12] N. Magotra, N. Ahmed, E. Chael, "A Comparison of Two Parameter Estimation Schemes," *Proc. IEEE*, Vol. 74, No. 5, pp. 760 - 761, May 1986.

FEC Decoder Design Optimization for Mobile Satellitel Communications

Ashim Roy, Leng Lewi

Hughes Network Systems, Inc.
11717 Exploration Lane
Germantown, MD - 20874, USA
Ph. (301) 428-5649, Fax. (301) 428-1868

ABSTRACT

A new telecommunications service for location determination via satellite is being proposed for the continental USA and Europe, which provides users capability to find the location of, and communicate from, a moving vehicle to a central hub and vice versa. This communications system is expected to operate in an extremely noisy channel in the presence of fading. In order to achieve high level of data integrity, it is essential to employ forward error correction (FEC) encoding and decoding techniques in such mobile satellite systems. A constraint length k=7 FEC decoder has been implemented in a single chip for such systems. The single chip implementation of the maximum likelyhood decoder helps to minimize the cost, size and power consumption, and improves the bit error rate (BER) performance of the mobile earth terminal (MET).

INTRODUCTION

The Location Determination Service (LDS) provides the end users location determination capability based on satellite ranging techniques and communication capability, via satellite, to a central hub. The LDS communications system comprises three main elements: a mobile earth terminal (MET), an LDS hub and a subscriber's hub, as shown in Figure 1. In mobile satellite communication systems such as LDS, size, cost and power consumption of the MET are major system constraints. The signal to noise ratio at the MET input is typically very low and the system operates in a fading environment. Under these circumstances, it is not possible to achieve end-to-end data integrity without a forward error correction (FEC) scheme. The FEC decoder presented here is a maximum likelihood decoder and therefore, offers the highest coding gain among convolutional decoders of a given constraint length. A single chip, constraint length k=7, FEC decoder based on the Viterbi algorithm has been developed for the mobile satellite communications application described above[1].

SYSTEM REQUIREMENTS

One of the key requirements for the LDS remote terminal unit is that it must be reliable under a wide range of operating conditions. The efficient usage of VLSI technology in the MET has made it possible to achieve very high MTBF for the unit. The LDS communications system may be visualized with respect to the LDS hub, where, the inbound channel provides a link from the MET to the hub and the outbound channel provides a link from hub to the MET. In the land mobile application, the MET must operate within harsh electrical and environmental condition. The MET operating conditions are listed in Tables 1.

Decoder Requirements

The information is encoded by a constraint length k = 7 convolutional code. The encoding polynomials, G1 and G2 are:

$$G1 = 1 + X2 + X3 + X5 + X6 = (133)_o \quad [1]$$

$$G2 = 1 + X1 + X2 + X3 + X6 = (171)_o \quad [2]$$

The rate 1/2 encoder described by the above polynomials may be easily built from shift registers. The encoder register is reset to logic "zero" at the beginning of a data frame and flushed with six "zeroes" at the end of the frame. Rate 3/4 encoding is provided using the same encoder by means of puncturing the encoded symbols. A puncture pattern, which causes the minimum loss in the BER performance compared to the rate 1/2 operation, is $(110)_b$ and $(101)_b$, for the I and Q outputs respectively. In order to meet the size and power consumption requirements of the MET, the decoder for the LDS must meet the requirements outlined in Table 2.

The Viterbi decoding algorithm may be implemented by either performing a fixed number (N) of mathematical operations serially in an ALU at a speed N times the input data rate or performing these operations in parallel in N ALUs. The serial scheme minimizes the silicon area while increasing the computation speed. The parallel implementation increases the silicon area and therefore, the cost, while minimizing the computation speed. The increased gate complexity of the parallel implementation results in higher power consumption for the decoder. An alternative to either the serial or the parallel implementation is a hybrid approach, where a small number of ALUs operating in parallel at a speed higher than the input data rate perform all the operation within the duration of each input data period. The hybrid approach also allows the designer to perform the speed, power and area optimization by selecting a sufficient number of math units (ALUs) to cater for the data rates involved in the application.

THEORY OF OPERATION

The FEC decoder is implemented as a rate 1/2 decoder with the capability to accept rate 3/4 encoded data and puncture inputs. A block diagram of the decoder implementation will be presented. The decoder receives convolutionally encoded symbol pairs, either in serial or parallel format and provides one corrected information bit for each symbol pair. The symbol output of the demodulator is available as 3-bit soft decision values, which are represented in sign-magnitude form. The decoder uses the 3-bit sign-magnitude symbol values as an indication of "strength" of logic "one" and "zero" at the output of the demodulator[2].

The branch metric logic uses the soft decision values to compute the branch metrics, M_i; i = 1,4 once every information bit period. The branch metrics are the total distance of each received symbol pair from the hypothesized symbol pair and represent the probability of bit error.

For constraint length k = 7, the decoder has $2^{(k-1)}$ or 64 states. The state metric computation is an iterative process, where the state value, $S_t(p)$, for state p at time t is computed from state values, $S_{t-1}(2p)$ and $S_{t-1}(2p+1)$ for states 2p and 2p + 1 at time t-1 as defined by the equations below.

$$S_t(p) = \text{Min}[X, Y] \quad [3]$$
$$S_t(p+32) = \text{Min}[X', Y'] \quad [4]$$

where:

$$X = S_{t-1}(2p) + M_i \quad [5]$$
$$Y = S_{t-1}(2p+1) + M_j \quad [6]$$

$$X' = S_{t-1}(2p) + M_j \quad [7]$$
$$Y' = S_{t-1}(2p+1) + M_i \quad [8]$$

where:
S is the state metric value and
M is the branch metric value

The state metric values are stored in the State Metric Memory (SMM). For each present state, its originating state is stored as a pointer, in the Trellis Memory (TM). For most practical implementation of Viterbi decoder, the trellis depth is traded with the maximum data rate of operation and the BER performance of the decoder. The traceback is initiated by first computing the present best state, which has the minimum state metric value. The best state becomes the starting point, for the traceback operation, in the trellis memory. The decoder then searches backwards based on the pointer values stored in the trellis memory. Traceback is completed when the decoder has searched the entire depth of the traceback memory, and decoder outputs a decoded information bit. The decoder operation for rate 3/4 encoded symbols is identical to that for rate 1/2 symbols with two differences:

1. The trellis depth for rate 3/4 operation is longer, and
2. The dummy data contained in the punctured bit locations are ignored.

DECODER DESIGN OPTIMIZATION

The FEC decoder performs many computations during every information bit period. These mathematical operations relate to two major tasks performed by the FEC decoder:

1. Generation of the state metrics and the trellis pointers, and

2. Traceback of the trellis to provide decoded data.

These two tasks are mainly independent and, therefore, may be performed concurrently. The computation clock frequency is selected based on providing sufficient number of computation clock cycles within each data rate clock for completing these concurrent tasks. Main mathematical operations, performed to decode an encoded symbol pair, are:

State Metric Computation	Trellis Computation
64 x 1 Memory Read	Trellis Memory Write
64 x 1 Memory Write	Address Generation
64 x 2 Additions	Trellis Memory Read
64 x 2 Compare	Iterative Traceback

The design partitioning is performed with a view to select an optimal set of design parameters for the above operations. These parameters include:

Size of State Metric Memory
Size of Trellis Memory
Number of ALU Units
Computation Clock Frequency
Power Consumption
Silicon Area

State Metric Memory

The constraint length k=7 decoder has 64 states. The State Metric Memory stores the metrics for all of these states. The word length of the SMM is dependent upon number of soft decision bits and SMM overflow protection algorithm. The decoder implementation presented here employs some unique properties of the state metrics:

1. Each present state is computed from two previous states, one of which is an even parity state and other is an odd parity state. The parity, considered here is, of the state number (i.e. index of S) not of the state value.

2. The states p and $(p + 2^{(k-1)/2})$ are computed from the same set of inputs as defined by the equations above.

3. The input states to the ALU and the output statess of the ALU form even-odd parity pairs.

These properties provide the means of computing two state metrics simultaneously, with minimal increase in additional logic. The routing on the silicon and the size of the state metric memory is optimized by employing butterfly memory technique. The butterfly architecture used in the present design has two ALU units coupled to two 32 x 6 state RAMs. One RAM is used for storing state metrics for even parity states and the other RAM is used for storing state metrics for odd parity states. The storage of new state values in the butterfly arrangement is explained in Figure 2. The computation clock frequency is further optimized by keeping the state RAM bus active at all times, alternating between read and write cycles as shown in the the pipeline arrangement in Figure 3. The chip power consumption is further lowered by the use of psuedo static RAMs rather the power hungry fully static RAMs.

Trellis Memory

The trellis depth is selected to meet the BER performance of the decoder for the LDS application. In an FEC decoder, infinite trellis depth provides the highest coding gain. The land mobile channel is such that traceback path in the trellis merges and it is sufficient to provide a trellis depth which is a small multiple of the constraint length. The trellis depth requirement increases with the coding rate [3]. The trellis memory is implemented as a byte-addressable RAM. The read and write operations to this RAM are interleaved during one information bit period. The addresses for present memory read operation during the traceback is computed from the contents of the previous memory read operation. Starting at the best state, traceback logic reads the pointers at each step to determine whether the state, p, originated from state, 2p, or 2p + 1. The implementation of the trellis memory address generator, therefore, simplifies to a shift register implementation, which is initialized with the best state number. At each traceback step, the shift register is shifted left by one and the trellis pointer loaded into the least significant bit position of the shift register to generate the address of the previous state from which the present state was derived.

Power Consumption

The power consumption in a CMOS device is a function of gate complexity and the frequency of operation. A major part of the operations for the decoder are related to the state metric computations. These memory I/O and ALU operations must be repeated $2^{(k-1)}$ times every information bit period. These state metric computations are largely sequential, in that memory read is followed by addition, which is followed by comparison and finally memory write. The computation speed required to complete these operations within a single information bit period is dependent upon the number of ALU's available to complete these tasks. Increasing the number of ALU units from 1 to 2 reduces the power consumption of the decoder a considerable amount at a cost of 4% increase in gate complexity and 2% increase in silicon area of the decoder. Further increase in the number of ALU units reduces the power consumption but has a larger impact on the silicon area. This is due to the exponential increase in interconnect buses from the ALU units to the state and trellis memories. Therefore, the FEC decoder for this application was based on two ALU units generating all the state metrics and trellis pointers. The results of this optimization are presented in Table 3.

PERFORMANCE

The performance of the decoder was measured under a variety of noise environments. The BER performance of the decoder is shown in Figure 4. The present implementation of the decoder provides coding gain within 0.2 dB of a FEC decoder with infinite trellis depth. This performance has been achieved within the constraints of the LDS require-

ments. A summary of the characteristics of FEC decoder is presented in Table 4.

CONCLUSIONS

A low cost, constraint length k=7 FEC decoder for Location Determination Service has been implemented in a single VLSI chip. This device has been extensively tested and provides very high coding gain necessary for LDS applications. The low cost and low power implementation of the decoder in a single chip makes the device ideally suited to mobile satellite communication applications. The inclusion of features such as rate 3/4 operation and QPSK modulation makes the device applicable for Very Small Aperture Terminal (VSAT) applications as well. The applicability of this technology for other mobile satellite communication systems based on Inmarsat Standard-M is being investigated.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. Eryx Malcolm for valuable help during the LDS system definition phase. The authors would like to thank Dr. J. No and Mr. Y. Antia for valuable suggestions and help in simulating the performance of the decoder.

REFERENCES

[1] Forney, G.D., 1973, "The Viterbi Algorithm", Proceedings of IEEE, vol. 61, March 1973, pp 268 - 278.

[2] Yasuda, Y., et al, 1981, "Optimum Soft Decision For Viterbi Decoding", Proceedings of International Conference on Digital Satellite Communications, pp 251 - 258

[3] Yasuda, Y., et al, 1983, "Development of Variable-Rate Viterbi Decoder and its Performance Characteristics", Proceedings of International Conference on Digital Satellite Communications, pp XII.24 - XII.31

Table 1. MET Operating Environment

Temperature	-30°C to +60°C
Relative Humidity	0 to 100%
Altitude	0 to 3000 meters
Vibration @ 1kHz	+/- 5 g

Table 2. FEC Decoder Requirements

Max. Encoded Data Rate	250 Ksps
Maximum Clock Freq.	10 MHz
Constraint Length	7
Encoding Polynomials	$(133)_o$, $(171)_o$
Coding Rates	1/2, 3/4
Rate 3/4 Puncture Pattern	$(110)_b$, $(101)_b$
Demodulation Interface	BPSK, QPSK
Rate 1/2 Coding Gain	5.1 dB @BER = 10^{-5}
Rate 3/4 Coding Gain	4.2 dB @BER = 10^{-5}
Code Synchronization	Yes
Active Power Consumption	150 mW
Standby Power Consumption	10 mW
Operating Temperature	-40°C to +85°C

Table 3. FEC Decoder Optimization for LDS

ALU's	Clock Freq. MHz.	Size K mil^2	Power mW
1	25	55.1	260
2	10	56.2	105
4	4	62.5	75

Table 4. FEC Decoder Characteristics

Fabrication Technology	1.5 micron CMOS
Number of Transistors	52,000
Decoder Dimensions	5.7 mm x 6.3 mm
Viterbi Decoder Package	44-pin PLCC
Active Power Consumption	105 mW
Standby Power Consumption	3 mW
Highest Data Rate	320 kbps

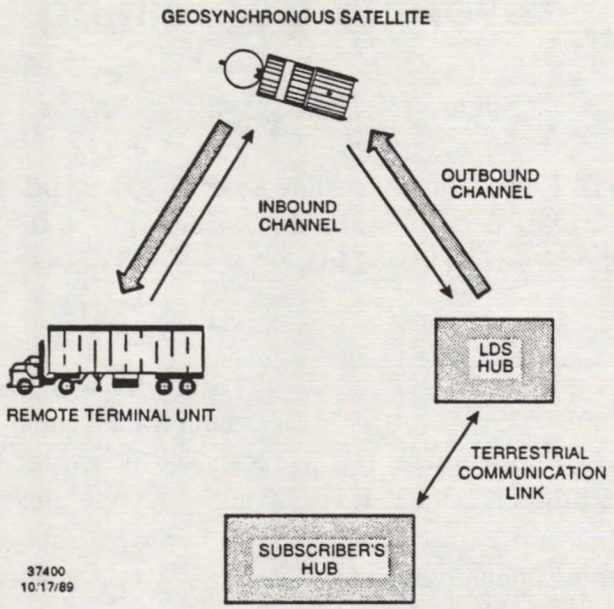


Fig. 1. LDS Communications System

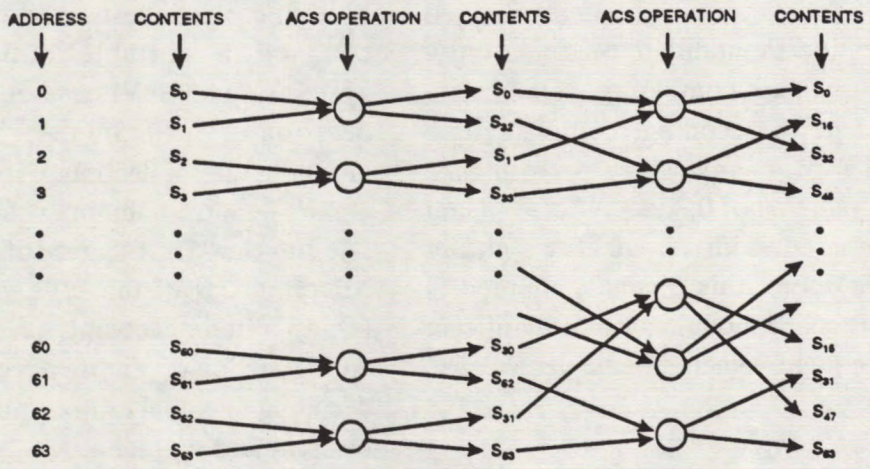


Fig. 2. Butterfly Memory Operation for k = 7 Decoder

ALU Cycle 1	ALU Cycle 2	ALU Cycle 3	ALU Cycle 4
RD (n)	WR (n-1)	RD (n+1)	WR (n)
NO-OP	ADD (n)	NO-OP	ADD (n+1)
SEL (n-1)	NO-OP	SEL (n)	NO-OP

Fig. 3 ALU Pipeline Operation

Soft-decision Viterbi Decoding with Diversity Combining

T. Sakai, K. Kobayashi, S. Kato

NTT Radio Communication Systems Laboratories
1-2356 Take, Yokosuka-shi, Kanagawa, 238-03 Japan
Phone: +81-468-59-3463, FAX: +81-468-59-3351

ABSTRACT

Diversity combining methods for convolutional coded and soft-decision Viterbi decoded channels in mobile satellite communication systems are evaluated and it is clarified that the pre-Viterbi-decoding maximal ratio combining shows better performance than other methods in Rician fading channels by computer simulation. A novel practical technique for the maximal ratio combining is proposed, in which the coefficients for weighting are derived from soft-decision demodulated signals only. The proposed diversity combining method with soft-decision Viterbi decoding requires simple hardware and shows satisfactory performance with slight degradation of 0.3dB in Rician fading channels compared with an ideal weighting scheme. Furthermore, this diversity method is applied to trellis coded modulation and significant Pe performance improvement is achieved.

1. INTRODUCTION

In a mobile satellite system, a radio terminal must stably operate in a quite low C/N and Rician fading environment. To improve bit error probability in low C/N satellite link, soft-decision Viterbi decoding has been widely adopted as a maximum likelihood decoding for a convolutional code in AWGN (additive white Gaussian noise)[1]. On the other hand, to overcome fading in a radio channel, diversity combining has been used for mobile communication and microwave transmission, and recently it is considered to be applied to mobile satellite communication systems[2]. Therefore, more improvement is expected in a Rician fading channel using soft-decision Viterbi decoding together with diversity techniques.

In previous studies on Viterbi decoding with diversity combining, several methods have been considered for various systems. For example, selection diversity in[3] or after[4] Viterbi decoding was investigated for Rayleigh fading channel in portable radio communication systems, where Viterbi decoding uses hard-decision signals. In the case of using soft-decision Viterbi decoding, it is possible to adopt maximal ratio combining before decoding[5].

In the first part of this paper, the performances of the following three methods, (a)post-Viterbi-decoding selection, (b)selection in ACS function of Viterbi decoder, (c)pre-Viterbi-decoding maximal ratio combining, are evaluated by computer simulation. As a result, it is clarified that the third method shows more improvement in bit error probability than the others. In the second part, a novel practical technique for the third method, that is maximal ratio combining, is proposed. In the proposed technique, the coefficients for weighting are derived from soft-decision demodulated signals only. The proposed diversity combining method with soft-decision Viterbi decoding requires simple hardware and shows satisfactory performance with a slight

degradation of 0.3dB in Rician fading channels from ideal weighting. Furthermore, this diversity method is applied to trellis coded modulation and significant Pe performance improvement is achieved.

2. COMPARISON OF DIVERSITY COMBINING METHODS

Three methods for soft-decision Viterbi decoding with diversity combining are investigated and their performances are evaluated by computer simulation.

(1) Post-Viterbi-decoding Selection

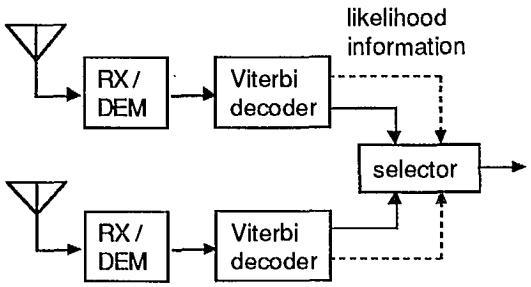
In this method shown in Fig.1(a), each signal received by diversity is individually demodulated and Viterbi-decoded. And then, decoded results are compared and selected based on the likelihood information such as path metrics provided in a Viterbi decoder. In this simulation, path metrics of the last 100 symbols are adopted as the most likelihood information for selection.

(2) Selection in ACS Function of Viterbi Decoder

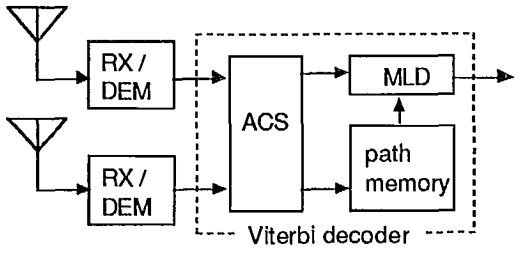
In a Viterbi decoder, the branch metrics associated with the state transitions are computed and then added to previous path metrics. The contending path metrics are compared and the path with the largest metrics is selected as a survivor. In this diversity method shown in Fig.1(b), the path metrics are computed for each diversity branch, and then compared all together at one time to select only one survivor. Therefore, this method has only one set of path-history storage.

(3) Pre-Viterbi-decoding Combining

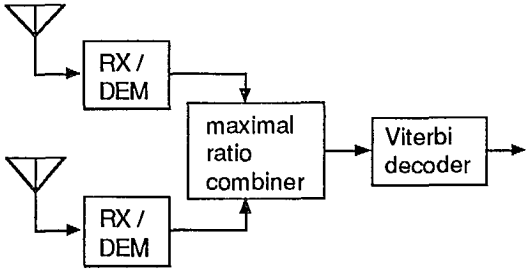
In the last method shown in Fig.1(c), conventional post-detection combining is applied to soft-decision Viterbi decoding. Soft-decision demodulated signals from each diversity branch



(a) Post-Viterbi-decoding selection.



(b) Selection at ACS in Viterbi decoder.



(c) Pre-Viterbi-decoding combining.

Fig.1 Combining methods in consideration.

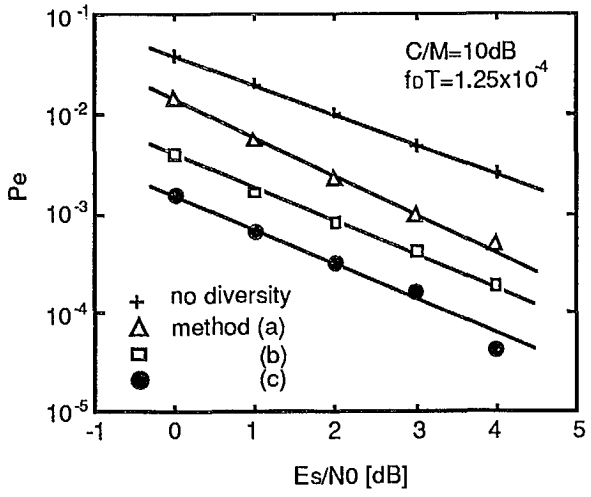


Fig.2 Comparison of combining methods with bit error probability.

are weighted for maximal ratio combining in proportion to signal to noise power ratio, and then summed. After that, they are decoded by a usual Viterbi algorithm.

The performances of the above three methods in a Rician fading channel are evaluated for QPSK coherent detection, convolutional coding and Viterbi decoding (r=1/2,K=4) by computer simulation. Simulation results on Pe performances are shown in Fig.2. In this comparison, Rician parameter C/M (the ratio of direct path signal power and diffused signal power) is set to 10dB and f_DT (f_D : maximum Doppler frequency, T : symbol period) is set to 1.25×10^{-4}. It is clarified that the method (c), pre-Viterbi-decoding combining, shows better Pe performance than the others. Similar results are obtained for other Rician parameters.

3. PRACTICAL COMBINING TECHNIQUE USING SOFT-DECISION SIGNALS

A novel practical technique for pre-Viterbi-decoding maximal ratio combining (method (c)) is proposed. The block diagram of proposed method is shown in Fig.3.

In the case of 2 branch diversity, maximal ratio combining signal R_{mc} is given by

$$R_{mc}=r_1\, s_1/\sigma_1^2+r_2\, s_2/\sigma_2^2 \qquad (1)$$

where r_i; demodulated signal, s_i; signal voltage, σ_i^2; noise power, i (=1,2); branch number.

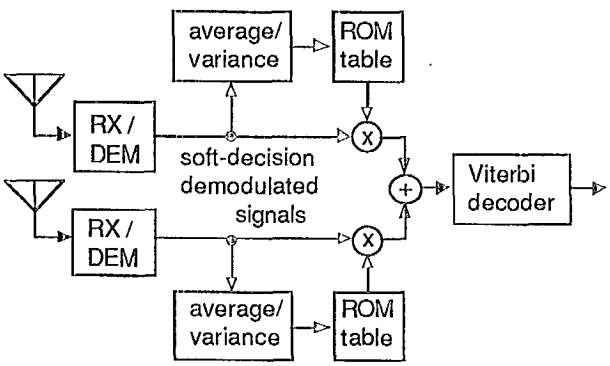


Fig.3 Weighted combining method using soft-decision demodulated signals.

In the proposed method, s_{ik} and σ_{ik}^2, which are coefficients for weighting at t=kT (T: symbol period), are estimated by calculating the average and variance of absolute value of soft-decision demodulated signals for n symbols from k-n/2 to k+n/2. In low C/N region, estimation error caused by absolute calculation is not negligible, therefore s_{ik} and σ_{ik}^2 are corrected by referring the ROM table.

4. PERFORMANCE OF PROPOSED METHOD IN RICIAN FADING CHANNEL

The performance of proposed method in a Rician fading channel depends on n, the number of observed symbols and f_D, maximum Doppler frequency. In an AWGN channel, theoretical improvement in bit error probability performance is obtained by calculating weighted coefficients, s_{ik} and σ_{ik}^2 for enough symbols. In Rician fading channels with higher maximum Doppler frequency, however, fast change of received level can't be followed if n is too large.

Relation between n, the number of observed symbols, and bit error probability in Rician channels is shown in Fig.4, where C/M=10dB and f_DT is 1.25×10^{-3}, or 5×10^{-3}. The smaller n is

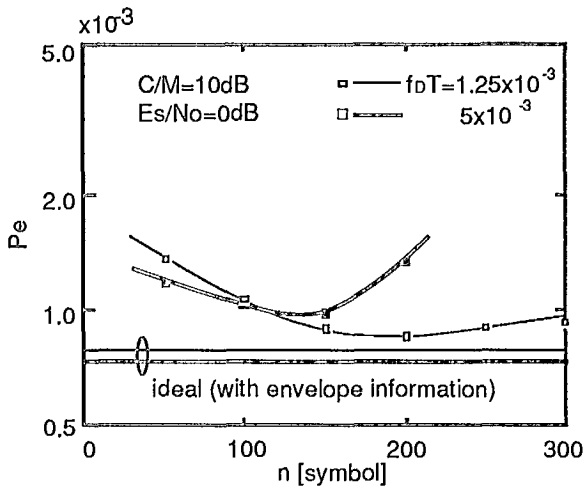


Fig.4 Relation between the number of observed symbols and bit error probability.

(shorter observation), the worse bit error probability is, because of estimation error by thermal noise fluctuation, and the larger n is, the worse it is also, because of slow response for f_D. From Fig.4, n=150 is optimum for $f_DT=5x10^{-3}$, and it also shows enough performance for $f_DT=1.25x10^{-3}$, which correspond f_D of 40Hz and 10Hz, respectively at a bit rate of 8kbps.

Improvement of bit error probability by the proposed diversity method in a Rician fading channel with C/M of 10dB and f_DT of $5x10^{-3}$ is shown in Fig.5 (n=150). The performances with ideal maximal ratio combining and equal gain combining are also shown in Fig.5. The proposed method shows satisfactory Pe performance with improvement more than 1dB compared with equal gain combining, and with slight degradation less than 0.3dB compared with ideal maximal ratio combining.

Up to the above section, received signals of two diversity branches in a Rician fading channel are supposed to have no correlation. In practical use, however, two diversity antennas can't be set up at the position with no correlation. In previous report[6] on sea surface reflection fading, coefficient of correlation is about 0.3 when two antennas are placed separating each other by a few times of wave length in vertical direction.

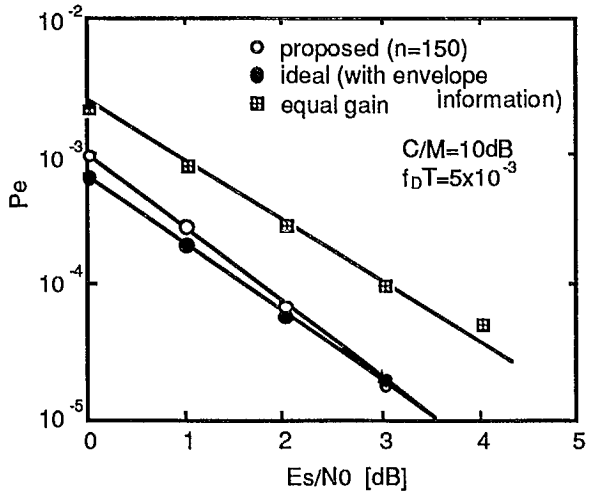


Fig.5 Bit error probability in a Rician fading channel with proposed diversity method.

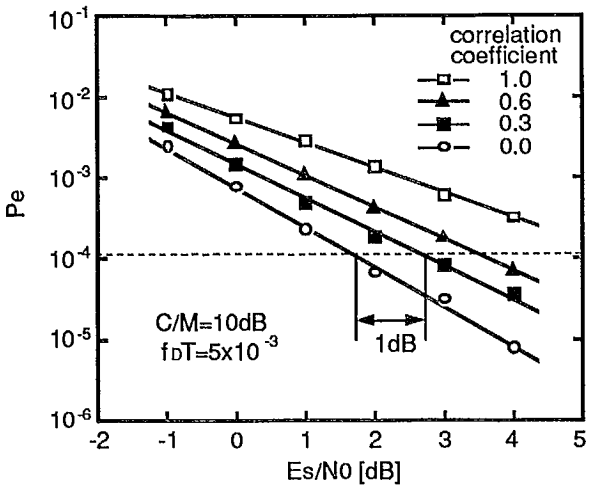


Fig.6 Relation between diversity improvement and correlation of fading.

The relation between improvement with diversity and correlation of fading is shown in Fig.6, where other parameters are as same as those in Fig.5. As seen from this figure, improvement with diversity degrades less than 1dB when correlation coefficient is less than 0.3 at $Pe=1x10^{-4}$. Therefore, this diversity method is effective in practical use such as under sea surface reflection fading.

5. APPLICATION TO TRELLIS CODED MODULATION

The proposed diversity method is applied to trellis coded modulation. In the conventional convolutional coding and Viterbi decoding, branch metrics are calculated based on Hamming distance. On the other hand, in coded modulation, it is on Euclidean distance. Therefore, maximal ratio combining for coded modulation is realized by weighting received signals in modulation-signal space, that is in Euclidean distance, instead of in soft-decision signals, that is in Hamming distance.

Simulation results of 8TCM (octal trellis coded modulation) using 16 states Ungerboeck code[7] with the proposed diversity method is shown in Fig.7, where C/M=10dB, $f_DT=8x10^{-3}$.

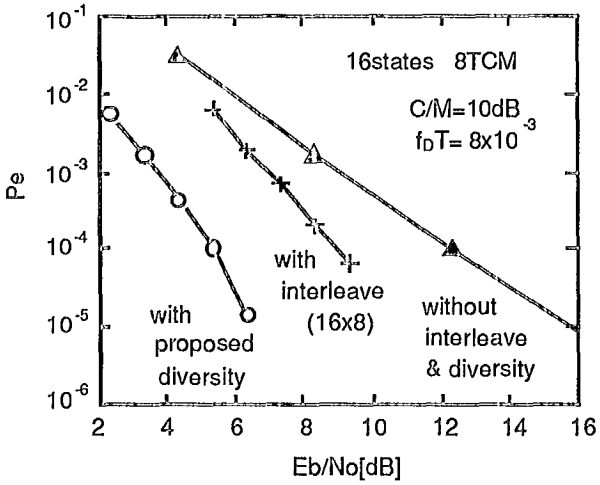


Fig.7 Bit error probability of 16 states 8TCM with proposed diversity method.

In this simulation, ideal CSI (channel state information) is assumed. The results without diversity but with interleave (16x8 symbols) and without both diversity and interleave are also shown in this figure. Improvement of required E_b/N_0 at Pe=1x10^{-4} with the proposed diversity is 7dB compared with no diversity and no interleave, and even 3.5dB compared with no diversity but with interleave.

6. CONCLUSION

In this paper, it is clarified that the pre-Viterbi-decoding maximal ratio combining shows better performance than other methods in Rician fading channels by computer simulation, and a new simplified maximal ratio combiner using soft-decision demodulated signals is proposed. The proposed method requires simple hardware and effective improvement of bit error probability is obtained in Rician fading channels. Furthermore, it also shows significant Pe performance improvement, where it is applied to trellis coded modulation.

ACKNOWLEDGMENT

The authors wish to thank Dr. H.Yamamoto and Mr. K.Morita for their encouragement and helpful guidance, and also Mr. M.Morikura, Mr. S.Kubota and Mr. T.Dohi for their useful discussion.

REFERENCES

1. **KUBOTA,S.** et al. 1989. General purpose high-speed convolutional encoder / Viterbi decoder. *The transaction of IEICE, Japan.* J72-B-1. pp.1226-1234.
2. **HARA,S.** et al. 1989. Post-detection combining diversity improvement of 4-phase DPSK system in mobile satellite communications. *The transaction of IEICE, Japan.* J72-B-2. pp.304-309.
3. **CHANG,L.F.** et al. 1987. Diversity selection using convolutional coding in a portable radio communications channel. *GCOM'87,* 21.3. Tokyo.
4. **KOHRI,T.** et al 1989. Maximum likelihood receiver for building block base station structure. *National Conf. of IEICE,* Japan. B-491.
5. **KOGA,K.** et al 1985. Signal combining method using FEC technique in diversity satellite communication. *The transaction of IECE, Japan.* J68-B. pp.1020-1028.
6. **KOZONO,S.** et al. 1981. Switch and stay diversity effect on maritime mobile satellite communication. *The transaction of IECE, Japan.* J64-B. pp.461-462.
7. **UNGERBOECK,G.** 1982. Channel coding with multilevel/phase signals. *IEEE Transaction on Information Theory.* IT-28. pp.55-67.

A Simple Implementation of the Viterbi Algorithm on the Motorola DSP56001

Dion D. Messer and Sangil Park

Motorola Inc.
Digital Signal Processing Operations
Austin, TX 78735

ABSTRACT

As system designers design communication systems with digital instead of analog components to reduce noise and increase channel capacity, they must have the ability to perform traditional communication algorithms digitally. The use of Trellis Coded Modulation as well as the extensive use of convolutional encoding for error detection and correction requires an efficient digital implementation of the Viterbi Algorithm for real time demodulation and decoding. Digital Signal Processors are now fast enough to implement Viterbi decoding in conjunction with the normal receiver/ transmitter functions for lower speed channels on a single chip as well as performing fast decoding for higher speed channels, if the algorithm is implemented efficiently. The purpose of this paper is to identify a good way to implement the Viterbi Algorithm (VA) on the Motorola DSP56001, balancing performance considerations with speed and memory efficiency.

1.0 Introduction

The DSP56001 is a digital signal processing chip which is well suited for communication applications and is also adaptable for Viterbi decoding because of it's dual data memory structure, zero overhead modulo addressing, and hardware do loop capability. The key obstacles to implementing the VA on the 56001 are: overflow in the "accumulated distance to each state" calculation, and finite memory availability for path storage. This paper discusses novel solutions to these obstacles in implementing the VA on the 56001 as well as an evaluation of the performance of the decoder using this implementation. An example code is used to help explain the concepts in this paper. The example code is the V.32 trellis, which is shown in Figure 1. Figure 2 shows the corresponding constellation and Figure 3 is a block diagram of the encoder.

This 8 state trellis is used as an example because the short constraint length is a less complicated structure to use for explanation than the K=7 codes which are popular on satellite channels. However, the performance of K=7 codes will be discussed in section 3.

2.0 Background

The Viterbi algorithm for decoding uses the structure of the trellis (i.e. the allowed transitions) and the input data to determine the most likely path through the trellis. The output for time t_0 reflects a decision made by the decoder on data received up to N time periods in the future. This means that the output for time t_0 is necessarily delayed by N time periods, or that the latency of the decoder is N time periods. N is determined by the constraint length of the code and for near-optimum decoding is 4 or 5 times the constraint length [9].

The most likely path through the trellis is determined to be that one which is a minimum distance path for the input data, or the path closest to the received data in Euclidean distance. In other words, the Viterbi algorithm minimizes the distance [1]:

$$d(r, v) = \sum_{i=0}^{N-1} d(r_i, v_i) \qquad (1)$$

where r_i and v_i are the received and the decoded signal sequence respectively.

Looking at Figure 3, there are 3 delays, (S1, S2, and S3) and the data they contain at any given time period is called the delay state in this discussion. The output (YOn, Yln, and Y2n) are referred to as the path state because they refer to the state of the path.

At each time period, every delay state in the trellis can have several paths (defined by each trellis) going into it, but only one will be the minimum distance for that delay state. Thus, the delay state with the smallest accumulated distance is the beginning point, at that time period, to trace the minimum distance path through the past N-1 time periods of the trellis. The minimum distance paths to the next delay state are then determined by evaluating the input to determine which point on the constellation in each path it is closest to, determining the Euclidean distance to each of those points, then, based on the trellis structure, and the minimum distance paths, determine the minimum accumulated distance to each delay state. So, after defining the trellis, the steps taken to decode the data are given below [1].

1) At each input compute the minimum distance path states and the corresponding Euclidean distances and store them for each path state.
2) Compute the accumulated distance to each delay state by adding the distance for each path state going into a delay state to the distance of the delay state where the path state originated, keeping the smallest of these distances and storing the path state and the delay state from which it came. Eliminate all other path states going into that delay state.
3) Find the delay state with the smallest accumulated distance and trace it back N times to read the path state, which is the output of the decoder for that time period.

Figure 4 shows the possible paths to delay state 010 for the V.32 trellis and how the minimum distance to 010 is chosen from the possible paths.

When the minimum distance path is found at each delay state, the path state taken to get there from the last delay state must also be stored (i.e., 001 in Figure 4 assuming $C+\gamma$ was the minimum) so that in N time periods, the output can be determined from the endpoint of the minimum distance path at time t_0+N. By storing the minimum distance path state (YOn, Yln, Y2n) to each delay state, as well as the delay state (S1, S2, S3) the path originated from, the most likely path can be traced. This is done by starting at the minimum accumulated distance state, going to the state it came from, and repeating this process N-1 times. That is, the minimum accumulated distance for all eight states identifies the state to be used as the starting point from which to trace back N time periods. Once the state for t_0 is found, the path taken to get to that state becomes the output of the decoder for the time period t_0. For instance, in Figure 4, if at t_0, the end point of the minimum distance path turned out to be 010 then the output of the Viterbi decoder would be 001.

In summary, at every time period, the accumulated distance to each delay state is calculated and updated and the minimum distance path state (YOn, Yln, Y2n) to each delay state is stored, as well as the delay state it came from (S1, S2, S3). This creates a history so that it is possible to trace back in time to get the correct output of the decoder.

A block diagram of the V.32 decoder showing inputs and outputs is shown in Figure 5. It can be compared to the block diagram of the encoder shown in Figure 3 to keep track of the input and output bit order. Decoding must be done by performing each decoder function in the reverse order in which it was encoded. In this case, the trellis decoding is done first and then the differential decoding is done.

3.0 Performance Parameters

The three basic parameters which affect the performance of the Viterbi algorithm are discussed in this section.

3.1 The Accumulated Distance Calculation

At every input, the accumulated distance to each state must be recomputed by adding previous accumulated distances to current path distances. Since the DSP56001 is a fixed point processor, this cannot occur continuously without resulting in an overflow problem. Thus, an alternate way to obtain the accumulated distance measurement is a weighted accumulation method which can be expressed as [10]:

$$d_{new} = \beta d_{old} + (1 - \beta)\, d_{path} \qquad (2)$$

where $O << \beta < 1$ denotes the smoothing parameter. This method (essentially a low pass filter) ensures that the new accumulated distance is a bounded arithmetic value. It has also been shown that this method gives unbiased estimates [10]. Although (2) uses all past values to compute a current accumulated distance, the value of β is directly related to the time constant, τ, which gives the number of recent past values to estimate the accumulated distance as:

$$\tau \approx \frac{2}{1-\beta} \qquad (3)$$

Using this equation, 85% of d_{new} comes from the points in the time constant, τ, and the remaining 15% is contributed by points previous to τ. In testing this implementation, values of β which fall in the range,$.9 < \beta < .99$ provided very good results in that there was no change in bit error rate (BER) with blocks of 10^4 data bits. Comprehensive tests using larger blocks of data are planned to chart the BER as β is varied over the same range. It is expected that this will produce an optimum value of β for different constraint length codes.

3.2 Path Memory Length

As stated previously, the number of time periods for near-optimum decoding is 4 or 5 times the constraint length K. The objective is to determine a path memory length which gives an optimum BER, decodes at an acceptable speed, and which conserves memory. Because of the looping capability of the DSP56001 and the modulo addressing scheme, each time period of path memory only requires 4 instruction cycles to trace. Therefore, in the case of V.32, time periods of 16 and 20 only take 60 and 80 instruction cycles respectively to determine the output when tracing through the trellis. The difference of 20 instruction cycles has a minimal affect on the total instruction cycle count needed for the decoding process. Since each time period requires only 4 instruction cycles, the extra processing time is not an issue in determining the path length.

Each additional time period does require extra memory locations for each state in the trellis. When the constraint length is short, the number of states are fewer and fewer extra memory locations are needed. As the constraint length increases to K=7, there are 2^{K-1} or 64 states. this means for each extra time period there needs to be 64 extra memory locations. In the K=7 case, the path length at 4 times K is 28 time periods and at 5 times K it is 35 time periods. For the difference of 7 time periods there would have to be 448 additional memory locations. When memory is scarce, this could be the decision factor in determining the path length used with this implementation.

Testing path memory lengths of 4 and 5 times the constraint length revealed no difference in BER performance. Again, blocks of 10^4 data bits were used for this testing. Additional testing of larger blocks of data is expected to reveal that the longer the path memory the better the performance of the decoder.

3.3 Maximum Data Rate

The Motorola DSP56001 currently operates at 27 MHz with an instruction cycle of 75 ns. 700 instruction cycles are all that is needed to decode on input symbol (4 output bits) for the V.32 case. This is only 15% of the processor capability, allowing time to perform the modem transmit and receive functions if desired. If the processor is used only as a stand alone decoder for this code, a data rate of 76Kbs can be achieved using 100% of the processor!

In the case of the constraint length K=7 codes often used for satellite channels, there are 64 delay states which pushes the processing for each input to 1300 instruction cycles. Using 100% of the processor, a data rate of 10Kbs can be achieved with present processor speeds. Ad-

vances in VLSI technology will push the processor speeds faster in the near future, allowing an even higher data rate for decoding on the DSP56001.

4.0 Summary

As shown, the Motorola DSP56001 offers a flexible solution to the Viterbi decoding task in a communications channel. Since it is programmable, it is possible to decode any number of different codes with varying constraint lengths by boot loading the software for the desired code. Using the DSP56001 can help in making designs compact by eliminating special purpose chips for decoding, echo cancellation, PLL, timing recovery, equalization, modulation and demodulation. All of these tasks can be performed on the DSP56001 for low data rate channels. As DSP's are used more and more in communication systems, Viterbi decoding as a software solution will become a necessity for efficient system designs.

The software for the example given is available on the Motorola DSP bulletin board (512-891-3771), or by contacting the authors at the address given.

REFERENCES

[1] S. Lin and D. Costello, Error Control Coding: Fundamentals and Applications, Prentice Hall, 1983.

[2] B. Sklar, Digital Communications Fundamentals and Applications, Prentice Hall, 1988, p. 319.

[3] G. Ungerboeck, "Trellis Coded Modulation with Redundant Signal Sets Part 1:Introduction," IEEE Communications Magazine 25(2) (February 1987).

[4] A. J. Viterbi, "Error Bounds for Convolutional Codes and An Asymptotically Optimum Decoding Algorithm," IEEE Trans. Inf. Theory, vol IT13, April 1967, pp. 260-269.

[5] DSP56000 Digital Signal Processor User's Manual, Motorola Inc., 1989.

[6] L. -f. Wei, "Rotationally Invariant Convolutional Channel Coding with Expanded Signal Space - Part 1:180," IEEE Journal on Selected Areas in Communications SAC-2(5) p. 661 (September 1984)

[7] CCITT, The International Telegraph and Telephone Consultative Committee, Red Book. Volume 8. 1985. p.222.

[8] A. Fagen, et Al., "Single DSP Implementation of a High Speed Echo Cancelling Modem Employing Trellis Coding," Proc. Of the Intnl. ESA Workshop on DSP Techniques Applied to Space Communications, Noordwijk, November 1988.

[9] J.A. Heller, and I.W. Jacobs, "Viterbi Decoding for Satellite and Space Communication," IEEE Trans. Commun. Technol., vol. COM19, no 5, October 1971, pp. 835-848.

[10] N. Magotra, et Al., "A Comparison of Two Parametric Estimation Schemes," Proc. IEEE, vol. 74, No. 5, pp.760-761, May 1986.

[11] E.A. Lee and D.G. Messershmitt, Digital Communication, Kluwer Academic Publishers, 1988.

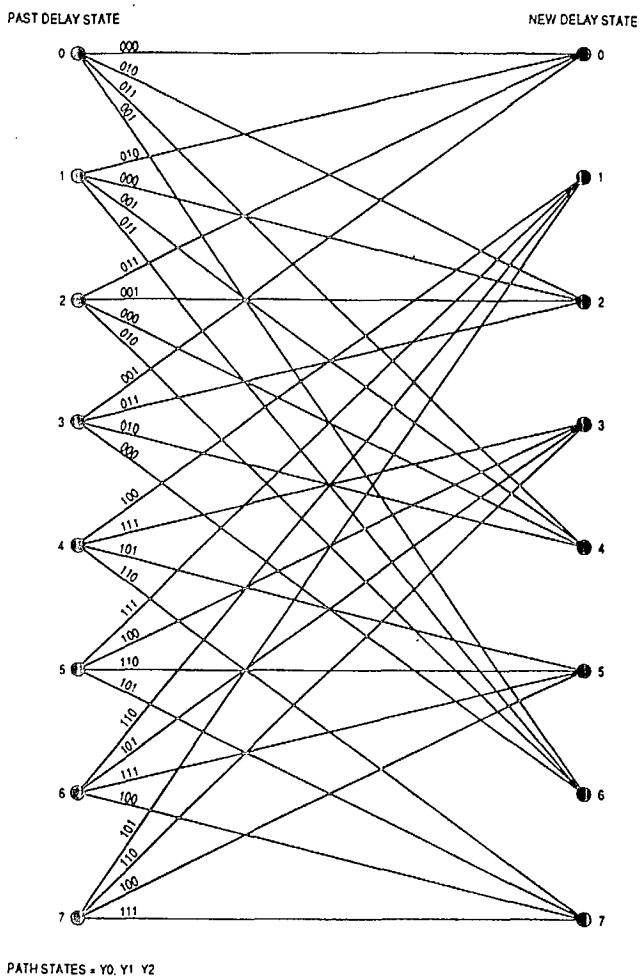


Figure 1. Trellis Diagram

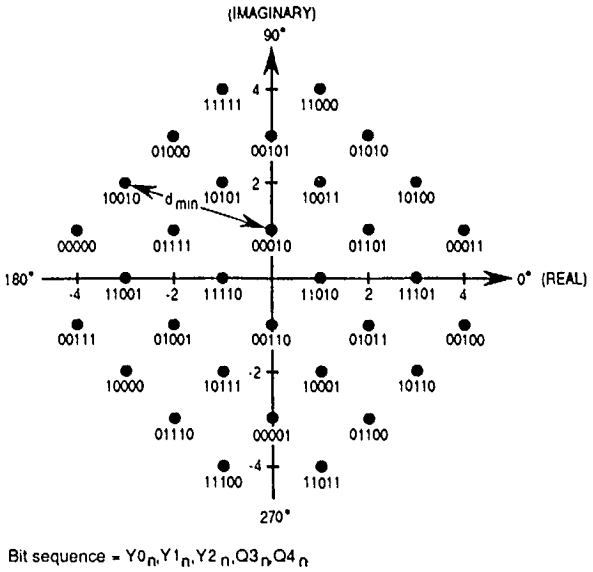


Figure 2. V.32 Constellation

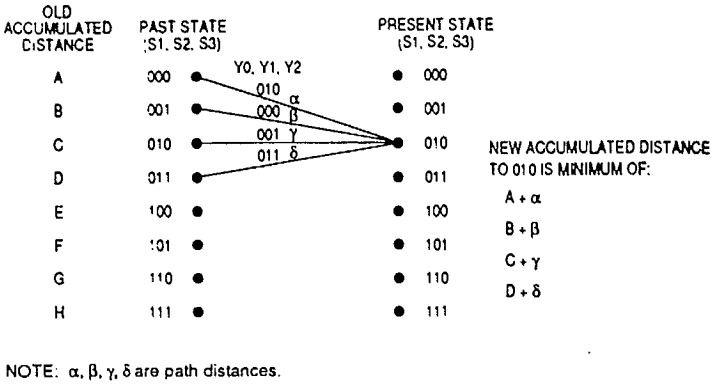


Figure 4. Possible Paths to State 010

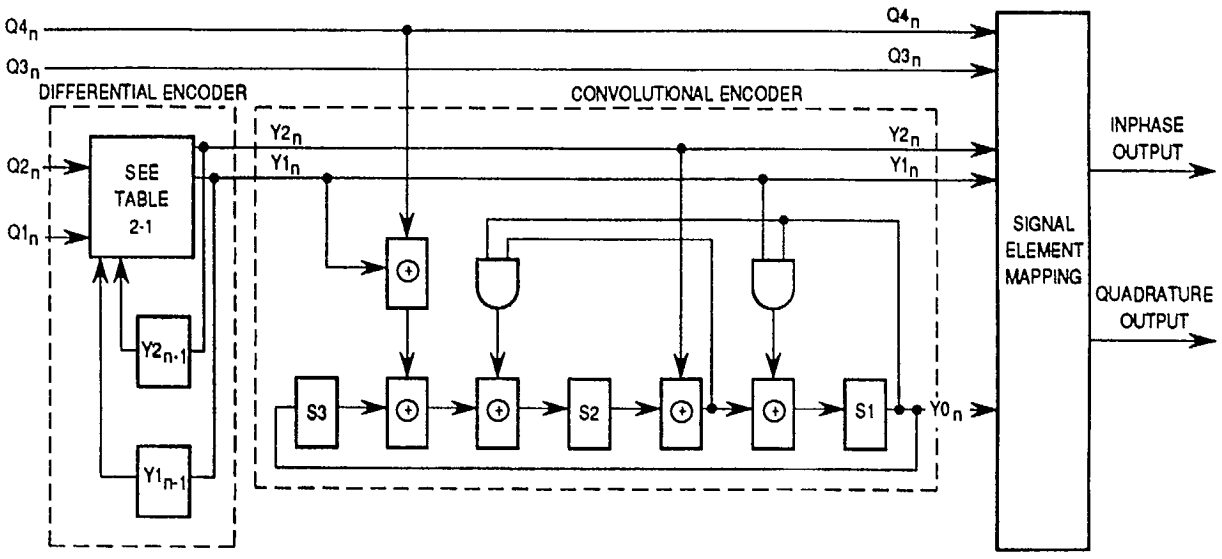


Figure 3. V.32 Encoding Diagram

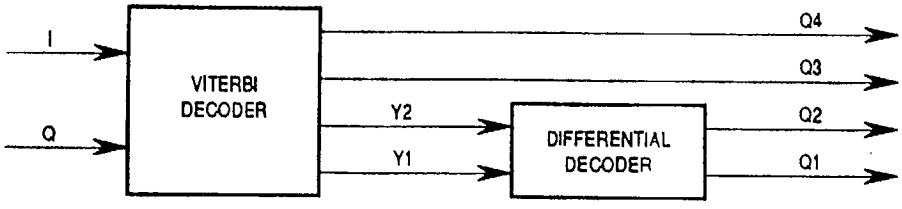


Figure 5. V.32 Decoder Block Diagram

Session 6
Propagation and Experimental Systems

Session Chairman - *Wolfhard Vogel,* University of Texas, USA
Session Organizer - *Richard Emerson,* JPL

6

Technical Development for Australia's MOBILESAT System

Kim Dinh, Nick Hart, Steve Harrison

AUSSAT Pty Ltd
GPO Box 1512
Sydney 2001, Australia
Phone: 61-2-238 7917
FAX: 61-2-238 7803

ABSTRACT

With the planned introduction in Australia of the mobile satellite service in mid-1992, MOBILESAT will be the first domestic mobile satellite system with full voice and data capability to be in operation worldwide. This paper describes the technical features which have been adopted by MOBILESAT in providing a unique system optimised for land mobile operation and the technical activities which have been carried out by AUSSAT in the past three years in supporting the development of the system.

1 INTRODUCTION

The MOBILESAT system developed for Australia shares many of the features being developed in other mobile satellite systems, as well as those of terrestrial mobile communications networks. The overall design philosophy is to provide robust, high quality services in a land mobile environment by 1992, whilst maintaining sufficient system flexibility so that future technological advances can be easily incorporated into the network. During the developmental stage, AUSSAT has also cooperated with the other mobile satellite network operators, namely INMARSAT, TMI and AMSC, with the aim of developing a common standard for the mobile satellite terminals.

In September 1989, AUSSAT and Telecom Australia released a Request for Tender for two Network Management Stations and collocated Gateway/Base stations. The successful bidder will be responsible for designing an integrated system which will optimise the interfacing protocols between the NMS, the Gateway/Base stations and the mobile terminals. The overall schedule will allow the detailed MOBILESAT system specifications to be released by 3rd Quarter 1990.

An overview of the MOBILESAT program and the market forecast for the mobile atellite services in Australia has been described in another paper at this conference[1].

In this paper, a detailed description of the technical features of MOBILESAT and related activities will be given, with particular emphasis on the following topics:

- Key network features adopted to address the market needs.
- Signalling and communications channel description.
- Experimental program carried out to support the development of the MOBILESAT system.
- Selection of the voice coding algorithm for the terminal.

2 SYSTEM OVERVIEW

The MOBILESAT system will offer, based on a single mobile terminal design, two generic mobile telephony service types, which will support voice, facsimile and full duplex data, and

an ancillary messaging service:

- The Public Telephony service will provide full duplex connection into the PSTN via a gateway station on a demand assigned basis.

- The Private Network Telephony service will offer a range of group call configurations through shared or private base stations. In addition to the full duplex capability, the terminal will also provide a press to talk facility which will enable a variety of group call types to be set up such that the talker's voice can be heard by all parties in the group.

- The Messaging service will route canned messages through the Network Management Station, using reserve capacity on the signalling channels. This will allow for a range of enhanced user features such as position reporting and vehicle monitoring.

The MOBILESAT network, as illustrated in Figure 1, will be implemented with two fully redundant Network Management Stations (NMS) at separate locations. One site will be configured as the primary NMS which will control the network whilst the other site will act as a backup. The NMS will allocate communications channels between the mobile terminals and the Gateway/Base stations on a per call basis using a Demand Assignment Multiple Access (DAMA) technique. Each NMS site will have a collocated Gateway/Base station which will interconnect into the Telecom ISDN network using 30 channel Primary Rate Access (PRA) circuits. Terrestrial lines will be used to interconnect the two sites for keeping the Network Management Stations in lock step, and for passing call processing information between the gateways and the primary NMS. Private base stations will communicate with the NMS using the mobile terminal common channel signalling scheme.

An FDMA architecture with a common channel signalling design has been adopted to accommodate future developments in modulation and processing technologies and to support a variety of communications standards. The modular system and ground infrastructure design will support the following key system features:

- Supporting a range of mobile terminal designs (e.g. mobile or fixed).

- Supporting different service types, qualities and priorities (e.g. data rates, link margins, and service priorities).

- Use of forward link power control to minimise the single carrier power.

- Providing an integrated network management facility that will control and monitor the whole network.

- Capability to support multiple common signalling channels.

- Compatibility with future multiple spot beam satellites.

- Supporting an ISDN terrestrial interface which will support advanced user features (e.g. call diversion, call charge advice, etc).

3 CHANNEL SPECIFICATIONS

3.1 SIGNALLING CHANNEL

MOBILESAT has been designed to provide a reliable and robust common channel signalling system that will meet the following criteria:

- Supporting a population of 100,000 mobile terminals in a single beam system.

- Supporting a call set up rate of 13 calls per second with additional messaging and system overheads.

- Establishing 99% of the channels within 5 seconds of call request.

- Providing system modularity to support different call types and future requirements (e.g. multiple beams).

The signalling system and the NMS will be configured to handle, on a per call basis, different communications channel types, service qualities and priorities. Also, as with the other networks, Bulletin Boards will be regularly transmitted to provide updates of the system status and

configuration. The system design will also permit the down-loading of key mobile terminal parameters from the NMS to the mobile terminals.

All signalling and messaging information will be sent using a 96 bit signal unit, protected with rate 3/4 convolutional code. A high power (5 dB link margin) outbound TDM signalling channel will be transmitted from the NMS using BPSK modulation, at a symbol rate of nominally 9600 sps (4800 and 2400 sps will also be supported). The high data rate will simplify the carrier acquisition process in the mobile terminals and will also provide sufficient capacity to support multiple repeats of critical signalling information.

A slotted ALOHA random access protocol will be used for the inbound channels. Upto thirty-two inbound channels will be allocated to a single outbound channel. Upto eight of the available inbound channels will be used in a TDMA mode with multiple repeats for acknowledging information on the outbound TDM channel. Inbound channels will operate at 2400 bps using Aviation BPSK modulation with 5 kHz channel spacing.

The performance of the signalling channels have been extensively simulated using Australian propagation data[2,3]. In combination with a straight forward channel establishment procedure, it is estimated that over 99% of the channel set up times will be less than 2.3 seconds for mobile originated calls. A complete simulation model for the overall signalling system is currently being established by Bond University, Australia, which will enable fine tuning of the overall protocol prior to the release of the final signalling specifications.

The system has been designed so that it will also support a 6 byte two way single packet data messaging capability via the NMS by using excess capacity on the signalling channels. The NMS will act as a packet data switch which will route data received on an X.25 terrestrial network port onto the outbound channel and perform the opposite service for the inbound channels. This will enable, for example, a dispatch centre to poll a mobile terminal and obtain a response in less than 2 seconds.

3.2 COMMUNICATIONS CHANNEL

AUSSAT, in conjunction with a number of Universities, has performed extensive analyses, simulations and experiments for various modulation and coding schemes for both voice and data channels. The MOBILESAT system can be configured as either power or bandwidth limited depending on the usage criteria.

The initial system will support up to 1,500 circuits using digital voice modulation with 5 kHz channels, at a threshold C/No of around 45 dBHz, including all implementation losses. An additional 3 dB fade margin will be used to provide a high quality, robust voice service that will operate under severely shadowed links. Indeed, AUSSAT has found through simulation that the capability of the voice codec to handle burst errors is the key to providing a robust service.

Simulated results by the University of Sydney, Australia, indicate that, theoretically, TCM schemes with coherent demodulation can offer 1 to 2 dB reduction in the required signal power compared to conventional schemes. However, due to the complexity associated with carrier acquisition and synchronisation in the design of the TCM demodulator, AUSSAT has adopted a more realistic approach and selected coherent QPSK modulation with 40% Nyquist filtering for maximum bandwidth efficiency. The OQPSK approach with a Class C amplifier was not considered since it did not give the required spectrum efficiency and was found to offer no other advantages over QPSK. Even though the construction of fast acquisition (10ms) four phase demodulators at low SNR is not trivial, AUSSAT is confident that good performance is possible down to at least an Es/No of 3dB.

The call signalling information between the mobile terminals and the Gateway/Base stations will be transmitted over the channels by dropping voice frames and inserting signalling packets. This will be feasible since at the start and end of the call no voice frames will be used, and during the call the dropping of a single voice frame will have no resulting degradation in the voice

quality. This simple technique will enable the channel overheads to be minimised and the channel rate to be limited to 6.6 kbps.

For the facsimile service, which will operate at upto 4.8 kbps, a similar protocol spoofing scheme to the INMARSAT specifications will be used. However, a novel ARQ scheme will also be employed to ensure error free transmission. The system design will allow easy transition to a Store and Forward facsimile service which will become available in the future. To verify the proposed scheme, AUSSAT has placed a consultancy contract to finalise the specifications and to build demonstration hardware by July 1990.

4 TECHNICAL ACTIVITIES

4.1 OVERVIEW

Over the past three years, AUSSAT has been conducting an extensive technical program to study the critical aspects of mobile satellite communications to support the development of the MOBILESAT system. This has been carried out both internally and in collaboration with other research organisations both within Australia and overseas.

One of the key activities is the experimental program which uses the Japanese ETS-V satellite as a test bed in Australia for mobile satellite experiments. This is part of a joint study between AUSSAT and the Communications Research Laboratory (CRL) of Japan.

4.2 EXPERIMENTAL PROGRAM

In order to access the ETS-V satellite, AUSSAT constructed a hub station in Sydney and equipped a van to be used as a mobile laboratory for the tests. The major experiments conducted to date include:

Propagation measurements - These were carried out in collaboration with an expert team from the Texas University at Austin to characterise the mobile satellite environment for Australia [4]. The results collected are summarised in Figure 2. This data has been used in simulation models for evaluating the signalling protocol, the voice codec and the modem performances.

Analog vs digital voice - Evaluation of the potential candidates for voice transmission over mobile satellite channel was carried out during the first half of 1989. These included both ACSSB modems supplied by Skywave, Kenwood and Bristol University, and 4.8 kbps voice codec built by the University of California at Santa Barbara (UCSB) for the Jet Propulsion Laboratory (JPL). As a result of these trials, a decision to employ a digital modulation technique was taken for the MOBILESAT system since it requires less power and its quality is comparable if not better than ACSSB.

Antenna evaluation - ETS-V has proven to be an ideal test bed for evaluating antenna performance in the real mobile environment. The antennas tested include various designs of omni-directional as well as mechanically steered (JPL, NEC, UNSW) and electronically steered (JPL) antennas.

Terminal evaluation - Complete terminal prototypes developed by NEC (digital) and Kenwood (ACSSB) for CRL were loaned to AUSSAT for evaluation in Australia in March 1989. JPL also brought their mobile van to Australia for testing in July 1989[5].

4.3 MOBILE SATELLITE HARDWARE SIMULATOR

A mobile satellite fading simulator has been constructed by AUSSAT to assist in the evaluation of the hardware in the laboratory. The simulator operates at the 70 MHz IF frequency, at which the I and Q components of the signals are modulated by the fading data. Two independent methods are employed to generate the fading data:

Measured data - a fading CW signal is received at the MOBILESAT van and is recorded in its I and Q format in the computer. The data is then calibrated and played back for the simulator.

Simulated data - a combination of a Rician model for the multipath fading and a Markov

model describing the transition between the fade and non-fade states is employed. This model produces propagation data which matches very closely to the statitical characteristics, such as fade duration, non-fade duration and fade depth, obtained from the measured data [6].

4.4 VOICE CODING SELECTION

In September 1989, AUSSAT issued a request for submission of voice codec hardware for evaluation for the MOBILESAT system. The program was conducted jointly with the INMARSAT-M voice codec evaluation, and the testing of the hardware was undertaken by Australia's Telecom Research Laboratory[7].

The MOBILESAT voice codec, including FEC, operates at 6.4 kbps, with an overall channel rate running at 6.6 kbps after the insertion of the unique word. The selected codec algorithm, in addition to providing high quality voice, must demonstrate that it can perform at BER as low as 0.01 and is able to handle fades as long as 100ms. At the time this paper was written (March 1990), it was anticipated that the final choice for the codec algorithm would be made in May 1990.

A contract has also been placed by AUSSAT with the University of Wollongong, Australia, to study various methods to combat fading and to make a recommendation for an algorithm that will improve the quality of the transmitted voice during long fades of upto 100ms. The proposed scheme will be considered for incorporation into the chosen codec to further improve its performance in the severe fading conditions.

4.5 DEMONSTRATION TERMINAL

AUSSAT is currently constructing a mobile terminal which will be fitted in a passenger car for demonstration to the customers. The terminal will employ the same codec and modem algorithms as specified for the MOBILESAT terminal and will operate at C/No of 48 dB/Hz, which is the nominal operating level for the current system. The hardware is anticipated to be completed in the third quarter of 1990 and will operate with the ETS-V satellite until the AUSSAT-B satellites are in operation.

5 CONCLUSION

As the first domestic mobile satellite system to be in operation worldwide, MOBILESAT will be charged with the challenge of being a test bed for the development of the other systems which will be implemented at a later date. Intensive study has been conducted by AUSSAT in the critical areas of mobile satellite communications in the design of MOBILESAT to make use of the current state-of-the-art technologies and to provide a robust system optimised for land mobile operation.

REFERENCES

1. Wagg,M. 1990. MOBILESAT, Australia's Own. International Mobile Satellite Conference 1990, Ottawa, Canada.

2. Irish,D. et.al. 1989. A Robust Signalling System for Land Mobile Satellite Services. Proceedings of the Mobile Satellite System Architectures and Multiple Access Techniques Workshop, March 1989, Pasadena, USA.

3. Rice,M. et.al. Modem Design for MOBILESAT Terminal, 1990. International Mobile Satellite Conference, 1990, Ottawa, Canada.

4. Vogel, W.J. et.al. 1989. Land-Mobile-Satellite Propagation Measurements in Australia using ETS-V and INMARSAT-Pacific. Published by Johns Hopkins University.

5. Dessouky,K. et.al. 1990. Field Trials of a NASA Developed Mobile Satellite Terminal. International Mobile Satellite Conference 1990, Ottawa, Canada.

6. Wakana, H. 1990. A Propagation Model for Land Mobile Satellite Communications. AUSSAT Internal Report.

7. Bundrock,A. et.al. Evaluation of Voice Codec for the Australian Mobile Satellite System. International Mobile Satellite Conference 1990, Ottawa, Canada.

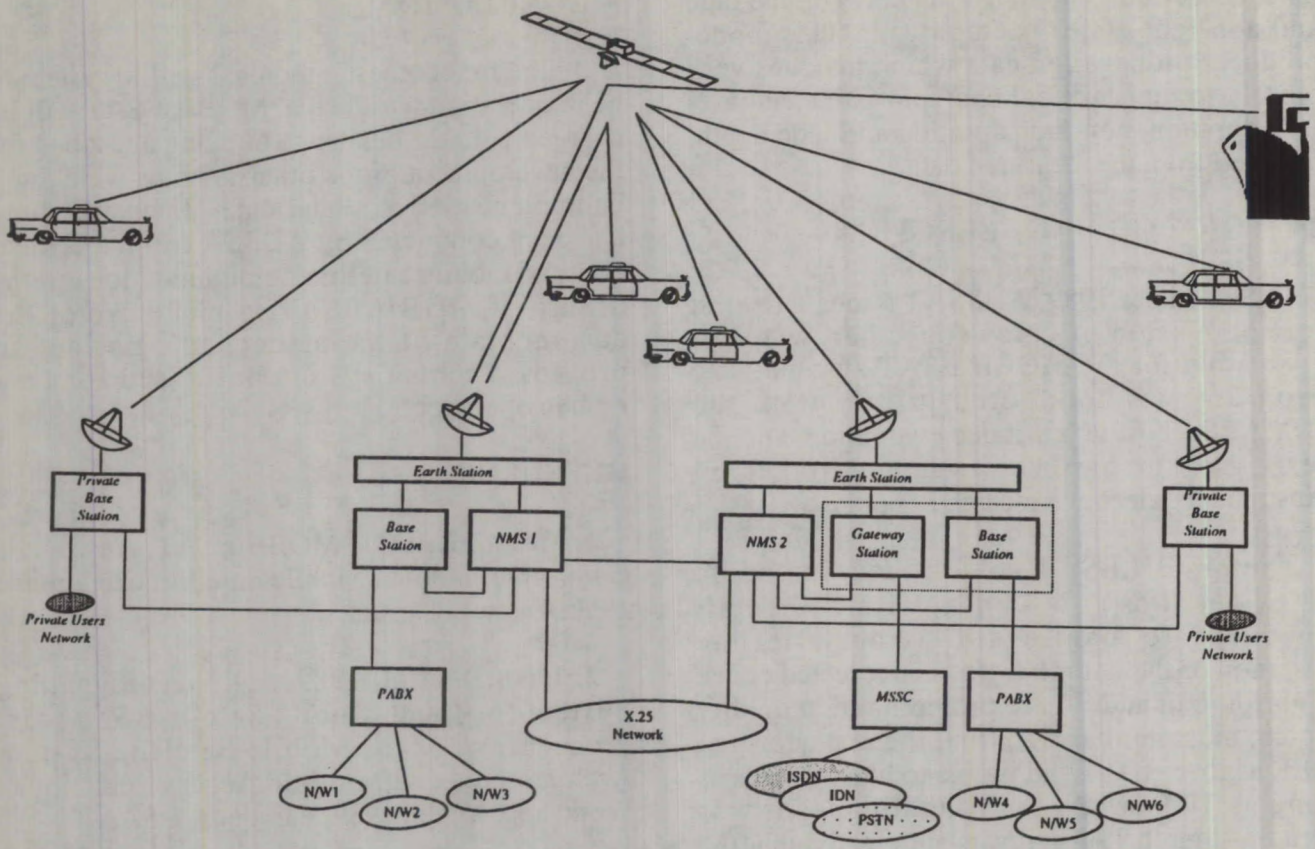


Figure 1 MOBILESAT Network Structure

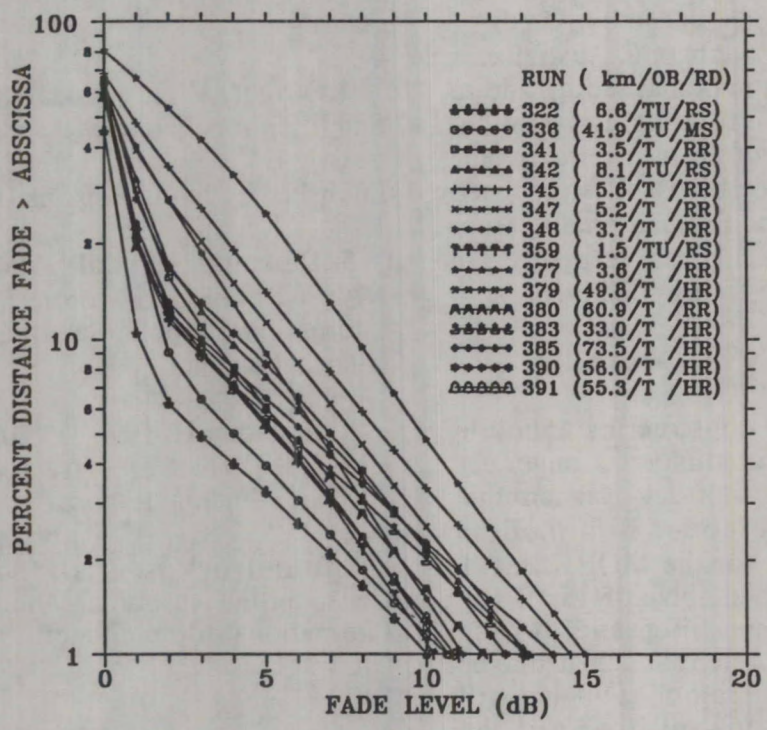


Figure 2 Propagation Data for Australia[4]

An Overview of Results Derived from Mobile-Satellite Propagation Experiments

Julius Goldhirsh
The Johns Hopkins University, APL
Johns Hopkins Road, Laurel MD 20723-6099
Phone: 301-953-5042 FAX:301-953-1093

Wolfhard J. Vogel
University of Texas, EERL
10100 Burnet Road, Austin TX, 78758-4497

ABSTRACT

During the period 1983-1988, a series of Land Mobile Satellite Service (LMSS) propagation experiments were performed in Southern United States (New Mexico to Alabama), Maryland, Colorado, and South-Eastern Australia. These experiments were implemented with transmitters on stratospheric balloons, remotely piloted aircraft, helicopters, and geostationary satellites (INMARSAT B2, Japanese ETS-V, and INMARSAT Pacific). The earlier experiments were performed at UHF (870 MHz) and the latter at both L Band (1.5 GHZ) and UHF. The general objective of the above tests was to assess the impairment to propagation caused by trees and terrain for predominantly suburban and rural regions where cellular communication services are impractical. In this paper are presented an overview of the results derived from the above experiments.

SINGLE TREE ATTENUATION-STATIC CASE

Attenuation and attenuation coefficient

In Table 1 is given a summary of the single tree attenuation results at 870 MHz (circularly polarized transmissions) based on the measurements by Vogel and Goldhirsh [1] and Goldhirsh and Vogel [2] who employed transmitter platforms on remotely piloted aircraft and helicopters and a receiving system on a stationary vehicle.

L-Band versus UHF attenuation scaling factor – static case

Ulaby et. al. [3] measured the attenuation properties at 50° elevation associated with propagation at 1.6 GHz through a canopy of red pine foliage in Michigan at both horizontal and vertical polarizations. Their measurements gives rise to an average attenuation coefficient of approximately 1.8 dB/m. Combining this result at L-band with the average value of 1.3 dB/m at UHF given in Table 1 suggests the following

$$F(f_{\mathrm{L}}) \approx F(f_{\mathrm{UHF}})\sqrt{\frac{f_{\mathrm{L}}}{f_{\mathrm{UHF}}}} \qquad (1)$$

where $F(f_L)$ is expressed in dB or dB/m. Simultaneous attenuation measurements by the authors at 1.5 GHz and 870 MHz for the mobile vehicle case showed consistency with the expression (1).

Tree attenuation versus season and path elevation angle

Figure 1 shows linear least square results of attenuation versus path elevation angle derived from measurements on the Callery Pear tree in October 1985 (full foliage) and March 1986 (bare branches)[2,4]. The best linear fit results in Figure 1 may be expressed as follows for θ between 15° to 40°

Full Foliage:

$$F_1(\theta) = -0.48\theta + 26.2 \qquad (2)$$

Bare Tree:

$$F_2(\theta) = -0.35\theta + 19.2 \qquad (3)$$

where θ is the elevation angle in degrees. The percentage rms deviations of the data points relative to the best fit expressions (2) and (3) were 15.3% and 11.1% (1.7 and 1.2 dB), respectively. We derive from (2) and (3) the average condition for f = 870 MHz, and path elevation between 15° and 35°

$$F(\text{full foliage}) \approx 1.35\text{F(bare tree)} \qquad (4)$$

which states that for the static case, the maximum attenuation contribution from trees with leaves (at 870 MHz) is nominally 35% greater than the attenuation of trees without leaves. The rms deviation about the average for the coefficient in the expression (4) is approximately .01%.

ATTENUATION DUE TO ROADSIDE TREES - DYNAMIC CASE

Empirical fade distribution model

Cumulative fade distributions were systematically derived from helicopter-mobile [2,4] and satellite -mobile measurements [5] in the Central Maryland region. This formulation referred to here as the MSS "Empirical Fade Distribution Equation" (EFDE) for P = 1 to 20% is given by

$$F_3(P,\theta) = -M(\theta)\ln P \ + B(\theta) \qquad (5)$$

where F_3 is the fade in dB, P is the percentage of distance (or time) the fade is exceeded, and θ is the path elevation angle to the satellite. Least square fits of second and first order polynomials in elevation angle θ (deg) generated for M and B, respectively, result in

$$M(\theta) = a + b\theta + c\theta^2 \qquad (6)$$

$$B(\theta) = d\theta + e \qquad (7)$$

where

$$\begin{cases} a = 3.44 \quad b = .0975 \\ c = -0.002 \quad d = -0.443 \quad e = 34.76 \end{cases} \qquad (8)$$

In Figure 2 are given a family of cumulative distributions (percentage versus fade exceeded) for the indicated path elevation angles.

L-Band versus UHF attenuation scaling factor–dynamic case

Simultaneous mobile fade measurements by the authors [2,4] at L-Band (1.5 GHz) and UHF (870 MHz) have demonstrated that the ratio of fades are approximately consistent with the square root of the ratio of frequencies formulation given by (1). The results demonstrated that for f_L = 1.5 GHz and f_{UHF} = 870 MHz

$$F(f_L) \approx 1.31F(f_{UHF}) \qquad (9)$$

The multiplying coefficient 1.31 was shown to have an rms deviation of +/- 0.1 over a fade exceedance range from 1% to 30%.

Seasonal effects–dynamic case

Seasonal measurements were performed by the authors for the dynamic case in which the vehicle was traveling along a tree-lined highway in Central Maryland (Route 295) along which the propagation path was shadowed over approximately 75% of the road distance [2,4]. Cumulative fade distributions were derived for March 1986 during which the deciduous trees were totally without foliage. These were compared with similar distributions acquired in October 1985 and June 1987, during which the trees were approximately in 80% and full blossom stages, respectively. For the "dynamic case" at a frequency f = 870 MHz and P = 1% to 30%

$$F(\text{full foliage}) = 1.24\text{F}(\text{no foliage}) \quad (10)$$

Fade reduction due to lane diversity

We examine the extent by which the fade reduces (or increases) by switching lanes for MSS configurations. For the case in which the satellite path and the tree line is to the right of the vehicle direction, the path length through the canopies of roadside trees is greater when the vehicle is in the right lane. A fade reduction should therefore be experienced by switching lanes from the right to the left side of the road. The authors measured this effect at UHF (870 MHz) and L Band (1.5 GHz) in Central MD [2,4,5]. A quantity defined as the "fade reduction, FR" is used to characterize the increase in signal power gained by switching lanes. This quantity is obtained by differencing equi-probability fade values from distributions pertaining to right and left side driving. The fade reduction at L-Band is plotted in Figure 3 for the indicated elevation angles as a function of the maximum fade as derived (for the example given) from right versus left lane driving.

At each of the elevation angles, the individual data points have been replaced by the "best fit third order polynomial" which may be expressed by

$$FR = A_o + A_1F + A_2F^2 + A_3F^3 \quad (11)$$

where FR (in dB) represents the fade reduction obtained in switching lanes from the greater shadowing configuration to the lesser one. Also, A_o, A_1, A_2, and A_3 are tabulated in Table 2 for L-band. The parameter F (in dB) represents the maximum fade level value. The "best fit polynomials" were observed to agree with FR as derived from the measured distributions to within 0.1 dB rms.

PROPAGATION IMPAIRMENT-LINE OF SIGHT COMMUNICATIONS

Multipath for a mountain environment

The results described here were obtained from MSS line of sight measurements in canyon passes in North-Central Colorado [6]. The transmitter was on a helicopter which, for each run, flew behind a receiving mobile van and maintained a relatively fixed distance and path depression angle relative to the receiving antenna. The radiating antennas on the helicopter transmitted simultaneous L-Band (1.5 GHz) and UHF (870 MHz) cw signals.

Figure 4 shows four cumulative fade distributions depicting "least square power curve fits" for the above described multipath scenario at a frequency of 870 MHz and 1.5 GHz and path elevation angle 30° or 45°. The resultant curves define the combined distribution corresponding to a driving distance of 87 km through canyon passes. Each of the best fit power curves agree with the measured cumulative distribution data points to within 0.1 dB rms. The distributions may be expressed for P = 1% to 10% by

$$P = aF^{-b} \quad (12)$$

where a and b are tabulated in Table 3 at the two frequencies and elevation angles.

Multipath Due to Roadside Trees

Similar type "line of sight measurements" as described above for mountainous terrain were also performed by the authors in Maryland along tree lined roads [4]. The resultant distribution was found to follow the exponential form for P = 1% to 50%

$$P = u\exp(-vF) \quad (13)$$

where u and v are tabulated in Table 4 and the corresponding distributions are plotted in

Figure 5. Since negligible differences in the individual cumulative distributions existed for $\theta = 30°$, 45°, and 60°, they were combined in the determination of (13).

OTHER PROPAGATION IMPAIRMENT EFFECTS

The authors have characterized a number of other propagation impairment effects for which limited space here does not allow their characterization [4,7,8]. These include distributions associated with "fade duration", "non-fade duration", "antenna space diversity", "cross polarization effects", and "high versus low antenna gain effects".

REFERENCES

1. **Vogel, W. J., and J. Goldhirsh.** 1986. Tree attenuation at 869 MHz derived from remotely piloted aircraft measurements. *IEEE Trans. Antennas Propagat.* AP-34: 1460–1464.

2. **Goldhirsh, J. and W. J. Vogel.** 1987. Roadside tree attenuation measurements at UHF for land-mobile satellite systems. *IEEE Trans. Antennas Propagat.* AP-35: 589–596.

3. **Ulaby, F. T., M. W. Whitt, and M. C. Dobson.** 1990. Measuring the propagation properties of a forest canopy using a polarimetric scatterometer. *IEEE Trans. Antennas Propagat.* AP-38: 251–258.

4. **Goldhirsh, J. and W. J. Vogel.** 1989. Mobile satellite system fade statistics for shadowing and multipath from roadside trees at UHF and L-band. *IEEE Trans. Antennas Propagat.* AP–37: 489–498.

5. **Vogel, W. J., and J. Goldhirsh.** 1990. Mobile satellite system propagation measurments at L-band using MARECS-B2. *IEEE Trans. Antennas Propagat.* AP–38: 259–264.

6. **Vogel, W. J., and J. Goldhirsh .** 1988. Fade measurements at L-band and UHF in mountainous terrain for land mobile satellite systems. *IEEE Trans. Antennas Propagat.* vol. AP-36: 104–113.

7. **Vogel, W. J., J. Goldhirsh, and Y. Hase.** 1989. *Land-mobile-satellite propagation measurements in Australia using ETS-V and INMARSAT-Pacific.* APL/JHU Tech. Rep. S1R89U-037 (Laurel, MD; The Johns Hopkins University, Applied Physics Laboratory).

8. **Hase, Y. W. J. Vogel, and J. Goldhirsh.** 1990. *Mobile-satellite measurements of L-band fade durations derived in Australia with ETS-V.* APL/JHU Tech. Rep. S1R90U-012. (Laurel, MD; The Johns Hopkins University, Applied Physics Laboratory).

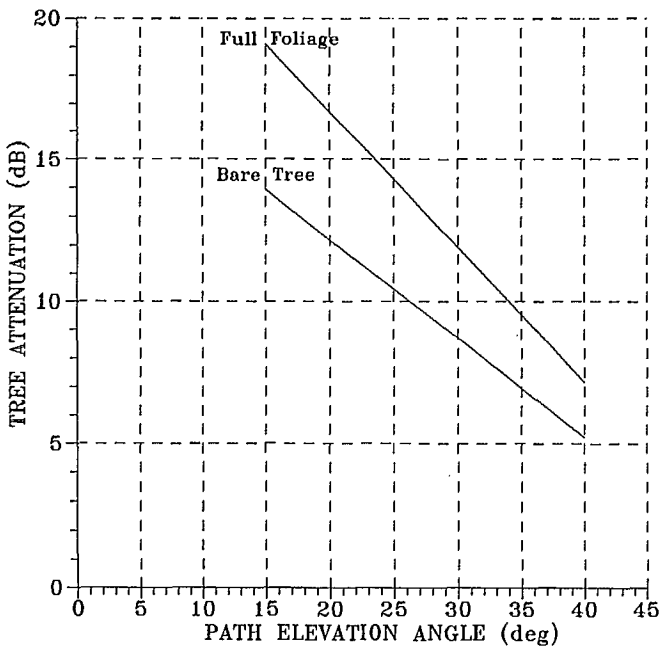


Figure 1: Least square linear fits of attenuation versus elevation angle for propagation through the canopy of a Callery Pear Tree at 870 MHz.

Table 1: Single Tree Attenuations at f = 870 MHz

Tree Type	Attenuation (dB)		Attenuation Coef. dB/m	
	Largest	Average	Largest	Average
Burr Oak*	13.9	11.1	1.0	0.8
Callery Pear	18.4	10.6	1.7	1.0
Holly*	19.9	12.1	2.3	1.2
Norway Maple	10.8	10.0	3.5	3.2
Pin Oak	8.4	6.3	0.85	0.6
Pin Oak*	18.4	13.1	1.85	1.3
Pine Grove	17.2	15.4	1.3	1.1
Sassafras	16.1	9.8	3.2	1.9
Scotch Pine	7.7	6.6	0.9	0.7
White Pine*	12.1	10.6	1.5	1.2
Overall Average	14.3	10.6	1.8	1.3

Table 2: Coefficients of fade reduction formulation (11) for lane diversity at $f = 1.5$ GHz.

El. Angle θ (deg)	A_o	A_1	A_2	A_3	dB Range
30	0.3931	0.2416	-1.420×10^{-2}	-3.389×10^{-4}	3-26
45	-1.073	0.8816	-4.651×10^{-2}	7.942×10^{-4}	3-17
60	-0.1962	0.2502	5.244×10^{-2}	-2.047×10^{-3}	3-16

Table 3: Coefficients a and b in best fit cumulative fade distribution formulation (12) for multipath in mountainous terrain.

Frequency (GHz)	El = 30°			El = 45°		
	a	b	dB Range	a	b	dB Range
0.870	34.52	1.855	2–7	31.64	2.464	2–4
1.5	33.19	1.710	2–8	39.95	2.321	2–5

Table 4: Coefficients u and v in formulation (13) describing best exponential fit cumulative fade distributions for multipath for tree-lined roads.

Frequency (GHz)	u	v	Fade Range (dB)
0.870	127.7	0.8573	1–4.5
1.5	125.6	1.116	1–6

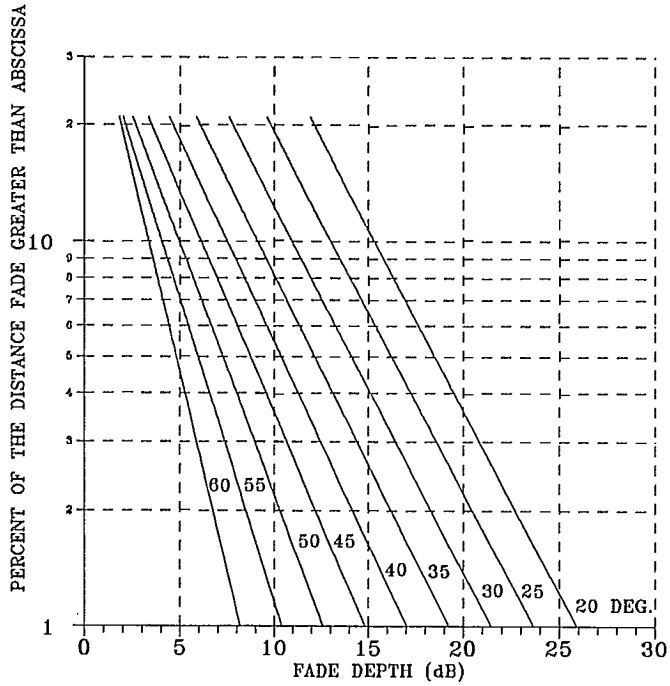


Figure 2: Cumulative fade distributions at 1.5 GHz for family of path elevation angles derived from the MSS Empirical Fade Distribution Equation (EFDE).

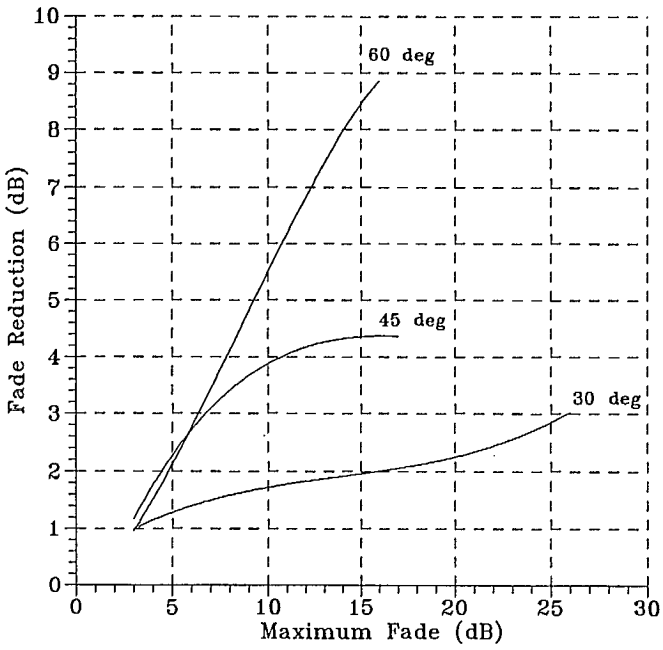


Figure 3: Best fit fade reduction at 1.5 GHz versus equi-probability attenuation at path elevation angles of 30°, 45°, and 60°.

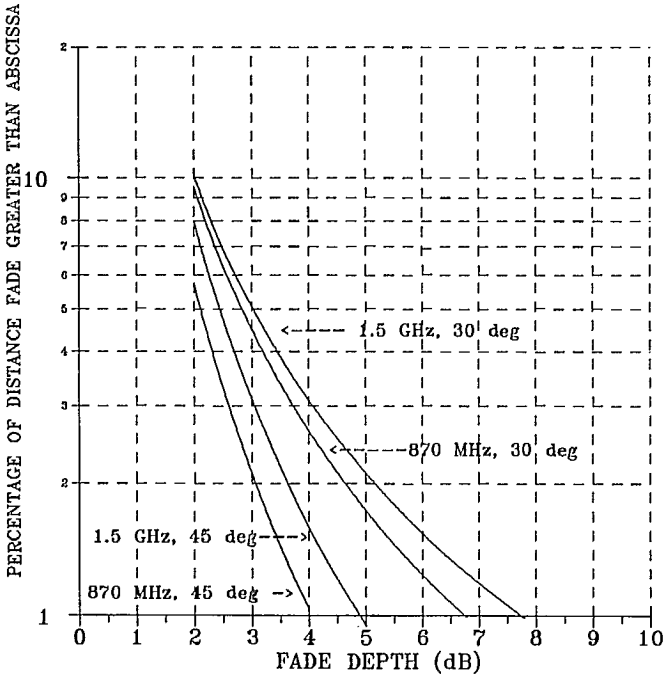


Figure 4: Best power curve fit cumulative fade distributions for line of sight configurations in mountainous terrain.

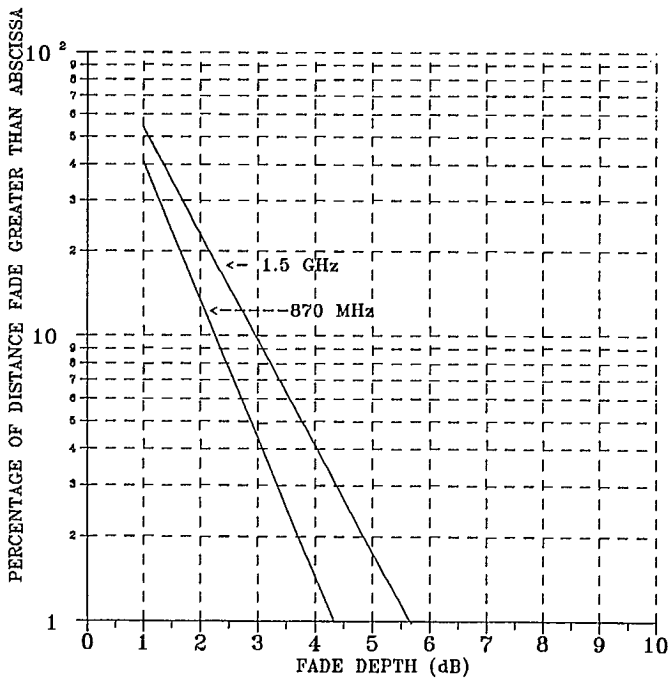


Figure 5: Best exponential fit cumulative fade distributions for line of sight configurations for tree-lined roads.

Effects of Fade Distribution on a Mobile Satellite Downlink and Uplink Performance in a Frequency Reuse Cellular Configuration

Karl Boutin, Michel Lecours, Marcel Pelletier, Gilles Y. Delisle

Electrical Engineering Department
Laval University
Quebec (Quebec), Canada G1K 7P4

ABSTRACT

In a mobile satellite system with a frequency reuse cellular configuration, significant co-channel interference can be experienced due to the antenna side-lobe level. The signal will be subjected not only to its own fading, but also to the effect of the varying degree of fading on the co-channel interferer, and this interference will behave differently in the up and in the down-link. This paper presents a quantitative evaluation of the combined effects of fades and co-channel interference on a mobile satellite link.

INTRODUCTION

In a mobile satellite system with a frequency reuse cellular configuration, significant co-channel interference can be experienced because of the satellite antenna sidelobe level.

In the downlink case, the satellite can emit simultaneously in the same frequency channel to two mobile stations located in different cells. Some of the signal intended for one station will be diverted to the other through the antenna sidelobe causing thereby some co-channel interference. Both the signal and the interference follow the same path from the satellite to the mobile station and both signal and interference will then fade together.

In the uplink case, a mobile station signal incoming at the satellite receiver can be subjected to co-channel interference from a mobile station located in an other cell. In this case, the signal and the interference fade independently. Thus one can imagine a worst case where the signal is subjected to heavy shadowing and the interference to light shadowing, or, at the contrary, a most favorable case where the signal is subjected to light shadowing and the interference to heavy shadowing.

In this paper, the impact of fades on a mobile satellite link is analyzed quantitatively. To carry on this evaluation, the ratio of signal level on noise added to interference (S/(N+I)) is calculated for an L-band uplink and downlink. The signal and the interference are represented in a probabilistic manner with their PDF's expressed by the model suggested by Loo [1,2] for heavy, light and overall shadowing.

The complementary information required to implement the evaluation of S/(N+I) is derived from estimated link parameters for a mobile satellite system. In the uplink, a C/N of 24 dB and an antenna sidelobe level of-22 dB are assumed. For the downlink case, the assumptions are for a C/N of 20 dB and a sidelobe level of-18 dB, which takes into account an expected degradation of the sidelobe isolation level due to the requirement of flexible power distribution among the different antenna beams. Results are applicable to transmissions in the L band.

It is important to note that, in this paper, the ratios and levels given correspond to power levels when given in dB (i.e.: S/N=24 db, antenna sidelobe level = -22 db), but that they correspond to signal (voltage) ratios or levels when they are given in a linear scale (i.e.: S/N=15.85, antenna sidelobe level = 1/12.58).

THEORY

According to Loo's model [1,2], the amplitude level of a signal transmitted through a land mobile satellite link can be described by the sum of two vectors. The first one represents the slow amplitude variations encountered in the channel and can be statistically described by a lognormal distribution. The second vector amplitude probability density function follows a Rayleigh law and characterizes multipath phenomena. The probability density function (PDF) of the resulting vector can be expressed from those preceding distributions to give:

$$f(r) = \frac{r}{b_o\sqrt{2\pi d_o}} \int_o^\infty \frac{1}{z} \exp\left[\frac{-(\ln z - \mu)^2}{2d_o} - \frac{(r^2+z^2)}{2b_o}\right] I_o\left(\frac{rz}{b_o}\right) dz \qquad (1)$$

where $I_o(u)$ is the modified Bessel function of zeroth order, b_o represents the average scattered power due to multipath, $\sqrt{d_o}$ and μ are the standard deviation and mean respectively of the lognormal distribution forming this model.

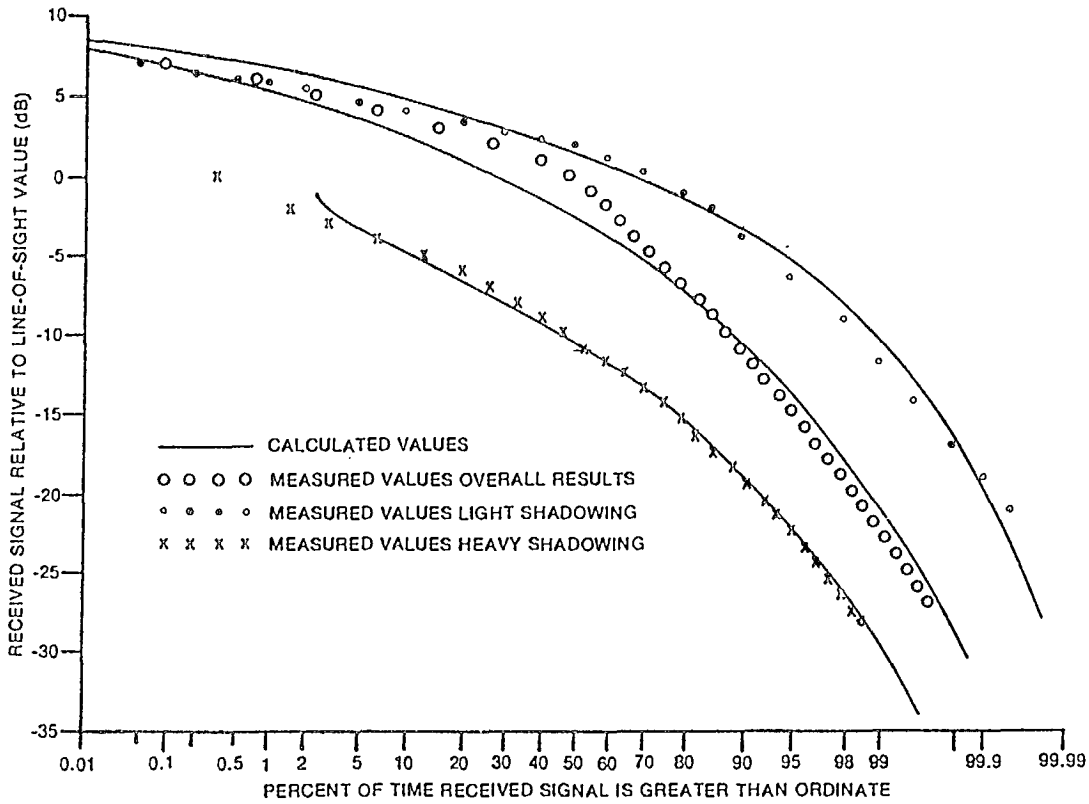


Figure 1: Comparison of measured and calculated values of the probability distribution of signal in a M-SAT channel (from Loo ([1]).

Environment	Standard Deviation $\sqrt{d_o}$	Mean μ	Multipath Power b_o
Infrequent light shadowing	0.0132	0.115	0.158
Frequent heavy shadowing	0.649	-3.91	0.0631
Average shadowing (overall results)	0.053	-0.691	0.251

Table 1: Loo's model parameters

In his work, Loo verified his model with experimental results obtained by Butterworth et al. [3,4] of C.R.C. in Ottawa. Figure 1 [1] gives a comparison of the measured values and of the model with the parameters adjusted as in Table 1 [1]. It represents the cumulative probability distribution of the received signal in a land mobile satellite link for different types of environment at the mobile.

3. UPLINK

In the uplink case, a mobile station signal incoming at the satellite receiver can be subjected to co-channel interference from a mobile station located in another cell. The signal and the interference fade independently.

The analysis below will be carried out considering a S/N ratio of 24 dB and an antenna sidelobe level of -22 dB.

With a reference level normalized to one, equation 1 can be used directly to represent the PDF of the signal level. S is then considered as a random variable following Loo distribution with the parameters b_o, d_o and μ fixed according to the channel or, more precisely, to the environment of the mobile.

$$f(S) = \frac{S}{b_o\sqrt{2\pi d_o}} \int_o^\infty \frac{1}{z} \exp\left[\frac{-(\ln z - \mu)^2}{2d_o} - \frac{(S^2+z^2)}{2b_o}\right] I_o\left(\frac{Sz}{b_o}\right)dz \qquad (2)$$

In the same way, the interference coming from neighboring cells follows a path that imposes the same type of PDF to its amplitude. But one must take into account the relative attenuation caused by the fact that this interference is picked up by the antenna sidelobe. The interference can be expressed in term of the normalized reference level which gives, on a voltage scale, 1/12.58 (-22 dB).

By applying the rules for functions of random variables, the PDF of the interference level in relation to the normalized reference level is obtained.

Considering that I = S/12.58, then

$$f(I) = \frac{12.58\,I}{b_o\sqrt{2\pi d_o}} \int_o^\infty \frac{1}{z} \exp\left[\frac{-(\ln z - \mu)^2}{2d_o} - \frac{((12.58\,I^2)+z^2)}{2b_o}\right] I_o\left(\frac{12.58Iz}{b_o}\right)dz \qquad (3)$$

It is important to note here that the parameters b_o, d_o and b_o shown above are not related to those given in equation 2 since, in the uplink case, the signal and the interference follow two different paths. It can happen however that these parameters take the same value for the signal and for the interference, for instance in the particular case where both the mobile station and the interference station are in a similar environment.

The last value to define in order to be able to implement the evaluation of the ratio of interest is the thermal noise. Considering for the uplink a S/N ratio of 24 dB, then, relative to a signal level of unity, N will be given by 1/15.85.

The ratio S/(N+I) to be evaluated is in fact a function of random variables. If we pose Z= S/(N+I), our work will then be equivalent to finding $F_Z(Z)$. The CPF should be expressed in term of the joint PDF of the random variables S and I:

$$F(Z) = \{z \le Z\} = P\{(S,I) \in D_z\} = \int\int_{Dz} f(S,I)\,dS\,dI \qquad (4)$$

To determine F(Z), one has to find the region D_Z for every Z and to evaluate the above integral.

From our previous discussion we know that S and I are independant, thus:

$$f(S,I) = f(S)\,f(I) \qquad (5)$$

With further observations we can sort out the appropriate region of integration and define:

$$F\left(Z \le \frac{S}{N+I}\right) = \int_0^{\infty}\int_0^{Z(N+I)} f(S)\,f(I)\,dS\,dI \qquad (6)$$

with f(S) and f(I) defined by equations (2) and (3) respectively. For ease of interpretation, it is preferable to evaluate the percentage of time when the ratio S/(N+I) is larger than a certain value. Thus we will calculate:

$$P\left(Z \ge \frac{S}{N+1}\right) = 1 - P\left(Z \le \frac{S}{N+1}\right) \qquad (7)$$

4. DOWNLINK

For the downlink, the evaluation is carried out considering a S/N ratio of 20 dB and an antenna sidelobe level of -18dB. Thus the thermal noise level N with reference to a signal level S of unity is 0.1, and the PDF of the random variable S is again defined in term of Loo's model by equations (1) and (2).

But for the downlink, the signal and the interference are transmitted through the same channel. The fluctuations that one signal will encounter will affect the other in the same way. Thus S and I cannot be considered as independent random variables anymore.

In fact they should be viewed as "dependent" events in that sense that the value of one variable can be directly computed when the value of the other one is known. The joint PDF of S and I will be expressed in term of their conditional PDF:

$$f(S,I) = f(I/S)\,f(S) \qquad (8)$$

The relation linking these two variables is simply the downlink sidelobe level. It states that the interference level will always be 7.94 times lower (18 dB) than the signal level. The conditional PDF takes the form:

$$f(I/S) = \delta(I - S/7.94) \qquad (9)$$

so we can define:

$$f(S,I) = \delta(I - S/7.94)\,f(S) \qquad (10)$$

F(Z) is evaluated with the same integral equation as in the previous section.

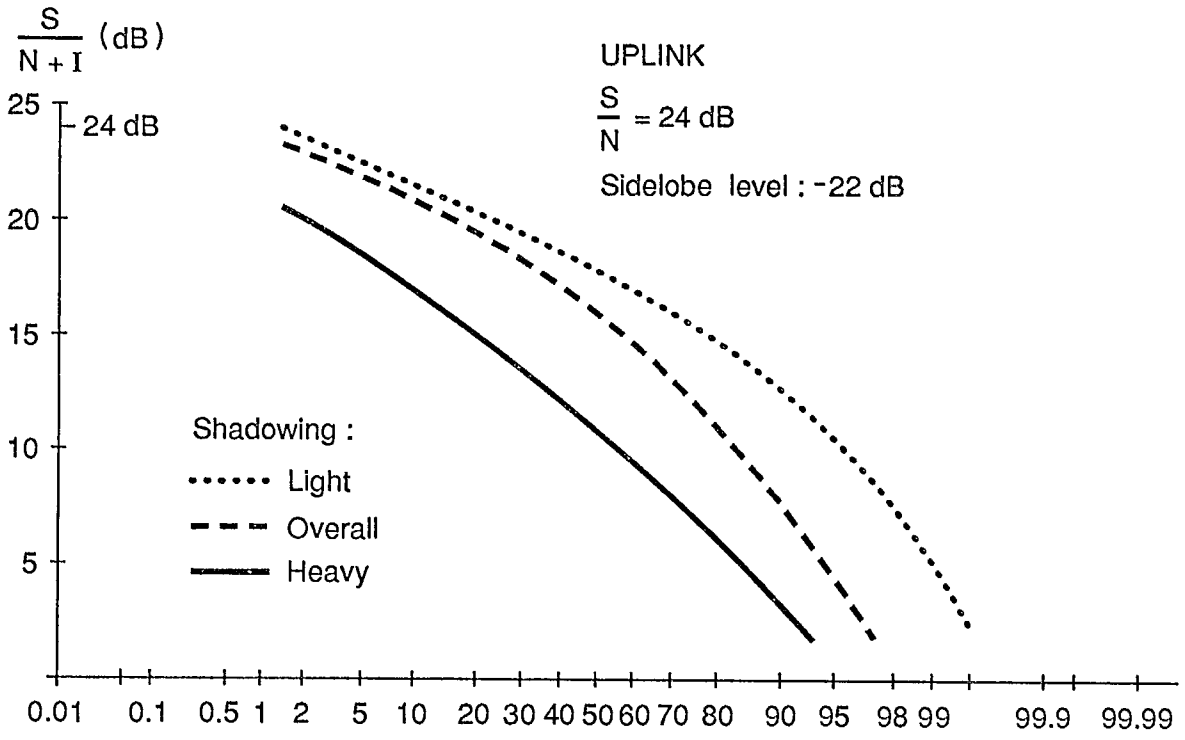


Fig. 2 - % OF TIME $\frac{S}{N+I}$ IS >THAN THE ORDINATE FOR THE UPLINK

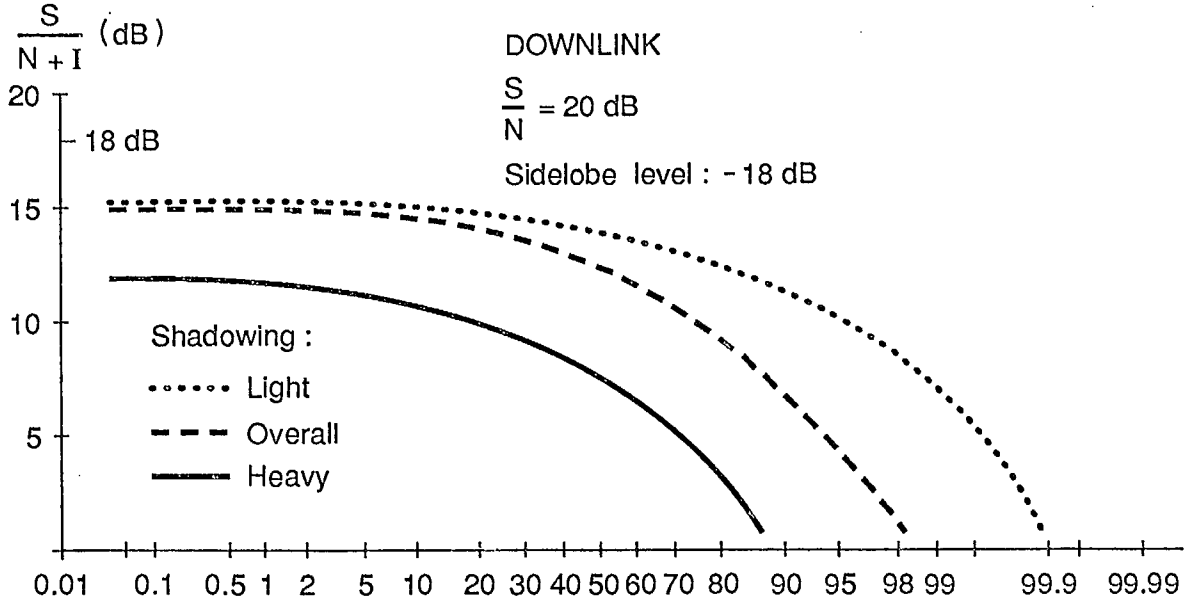


Fig. 3 - % OF TIME $\frac{S}{N+I}$ IS >THAN THE ORDINATE FOR THE DOWNLINK

$$F(Z) = \int_0^{\infty} \int_0^{Zx(N+I)} f(S,I)\,dS\,dI \qquad (11)$$

But with f(S,I) given by (10), this integral equation will be simplified to the following form:

$$F(Z) = \int_0^{\frac{N}{1/z-1/7.94}} f(S)\,dS \qquad (12)$$

The final step will be to seek the cumulative distribution of:

$$P\left(Z \geq \frac{S}{N+1}\right) = 1 - F(Z) \qquad (13)$$

RESULTS AND CONCLUSION

The results for the uplink case are illustrated in figure 2. The three curves give the percentage of time the signal to noise plus interference ratio will be exceeded for the cases where both signal and interference are subjected either to light, or to average or to heavy shadowing. Estimations are for a S/N of 24 dB (for the main signal) and a sidelobe level of -22 dB (for the interfering signal). As can be seen on the figure, for light and for average shadowing, the S/(N+I) ratio would be higher than 20 or 19 dB 50% of the time, and larger than 14 or 10 dB 80% of the time.

Figure 3 illustrates the results for the downlink. The three curves give the percentage of time a given signal to noise plus interference ratio will be exceeded for the cases where the mobile unit is submitted to light, average or heavy shadowing.Estimations are made for a S/N of 24 dB and a sidelobe level of -18 dB. As can be seen on the figure, for light and for average shadowing, the S/(N+I) ratio would be higher than 13 or 11 dB 50% of the time, and larger than 12 and 9 dB 80% of the time.

Those definitions of shadowing come from the experimental work of Butterworth [3] and give a fair representation of the kind of environment that would be imposed to the transmission channel of a canadian customer. Light shadowing and heavy shadowing correspond to a surrounding of sparse and dense tree cover respectively. The representation called average shadowing comes from the combination of the light and heavy shadowing results in Butterworth's experiment.

In this study, the performance of mobile satellite links affected by fades has been analyzed. This was achieved by evaluating the S/(N+I) ratio for an L-band uplink and downlink. The statistical representation of the signal and fade fluctuations has been assumed to correspond to Loo's model. Performance curves have been obtained for various types of channel environment and with estimated link parameters of 24 dB C/N and -22 dB antenna sidelobe level for the uplink with 20 dB C/N and -18 dB antenna sidelobe level for the downlink.

The authors wish to acknowledge the financial and scientific support of SPAR Aerospace Montreal, and in particular the advices of MM. K. Hing, R. Belanger and of Dr. H.J. Moody.

[1] C. Loo, "A statistical model for a land mobile satellite link," IEEE Trans. Vehic. Tech., vol. VT-34, no. 3, pp. 122-127, Aug. 1985.
[2] C. Loo, "Measurements and models of a land mobile satellite channel and their applications to MSK signals," IEEE Trans. Vehic. Tech., vol. VT-35, no. 3, pp. 114-121, Aug. 1987.
[3] J.S. Butterworth, "Propagation measurements for land mobile satellite system at 1542 MHz," Commun. Res. Cent., Dept. Commun., CRC Tech. Note 723, Aug. 1984.
[4] J.S. Butterworth, "Propagation measuremnts for land mobile satellite services in the 800 MHz band," Commun. Res. Cent., Dept. Commun., CRC Tech. Note 724, Aug. 1984.

A Propagation Experiment for Modelling High Elevation Angle Land Mobile Satellite Channels

M.Richharia, B.G.Evans and G.Butt

University of Surrey,Guildford,Surrey GU2 5XH.
Phone: (0483) 571281
Fax: (0483) 34139

ABSTRACT

This paper summarises the results of a feasibility study for conducting high elevation angle propagation experiments in the European region for land mobile satellite communication. The study addresses various aspects of a proposed experiment. These include the selection of a suitable source for transmission,possibility of gathering narrow and wide band propagation data in various frequency bands,types of useful data , data acquisition technique , possible experimental configuration and other experimental details.

1. Introduction

Propagation characteristics of land mobile satellite links play an important role in the design and viability of such a system for civilian use. The main problems arise due to the shadowing and multipath effects caused by the nature of environment in the vicinity of a mobile. The propagation effects worsen as the elevation angle is reduced. As a result geostationary satellites donot provide an optimum solution for land mobile communications in many parts of Europe. An alternative is to use a suitable non-geostationary orbit so that satellites appear at high elevation angles ($> \approx 50$ deg) throughout Europe (1). It is expected that the advantages gained in terms of reduction in degradation due to propagation should be large enough to warrant the added complexity in the system architecture. However,at present the advantages possible by the use of this architecture are difficult to quantify precisely since propagation data for elevation angles above $\approx$ 50 deg are very scarce (2).

In addition there is currently an interest in wide-band characterization of such a channel for applications such as satellite sound broadcasting (3). Most measurements to date have attempted a narrow-band characterization of the land mobile satellite (LMS) channel. Hence there is a need to obtain data for such applications.

The present frequency allocation for LMS services is a bandwidth of around 4 MHz in L-band.It is expected that this bandwidth will not be adequate to provide the projected demand for future LMS applications. Hence an allotment in another frequency band may be essential. The present LMS channel propagation data are limited to UHF and L bands . Therefore there is a need to expand the data-base to other possible frequency bands.

This paper is an outcome of a study performed by the University of Surrey for the European Space Agency (ESA). A part of the work was performed by Racal - Decca Advanced Development Ltd., U.K. under a sub-contract.

The main objective of the study was to assess the feasibility of conducting propagation experiments for land mobile satellite communication service at eleva-

tion angles greater than 50 deg. The study was expected to result in a recommendation for one or more experiments which can statistically quantify the elevation angle dependence of narrow-band and wide-band propagation characteristics in L (1.5 GHz),S (3 GHz),C (4/6 GHz) ,K_u (11 - 15 GHz) and K_a (20 - 35 GHz) bands. In this paper the main features of the investigation have been summarised.The details are available elsewhere (4). As a follow-on of the study an experiment is currently being developed at the university. The present status of this experiment is also briefly stated.

2. Source for transmission

A basic requirement in the design of the experiment was to identify a suitable source for transmission in the desired frequency bands at an elevation angle > 50 deg . Due to the inherent advantage of using a satellite beacon as a source,investigations were performed to select a suitable satellite. Investigations were also conducted to identify an alternative platform.

2.1. Satellite sources

The evaluation for the suitability of a satellite source was based on the following considerations,

i) Frequency of transmissions in the desired bands.

ii) Transmission bandwidth.

iii) High elevation visibility - within mainland Europe/outside mainland Europe.

iv) Passive or active (transponded) mode.

In addition to the GPS and GLONASS satellite constellations, several other satellites (all geostationary) were considered for wide-band measurements. It was noted that all these other satellites required active mode of transmission and provided visibility only outside mainland Europe.

The decision reached during the first phase progress meeting was to preclude locations outside Europe due to difficult logistics and doubts regarding the applicability of the data-base to European environments. As a result ,GPS and GLONASS satellites were the only possible satellite sources for wide-band measurements. Further considerations such as the required measurement bandwidth and the available technical data-base of individual systems led to the choice of GPS satellites as the only possible wide-band satellite source. This implied that a multi-band experiment using a satellite source was at present not possible. Moreover considering the bandwidth requirements for applications such as satellite sound broadcasting only the P-code transmissions ,which could provide bandwidths of about 4 MHz ,were considered suitable. At the time (early 1989) the availability of GPS constellation over Guildford,at elevation angles > 50 deg,was estimated to be over 6.5 hours/day. Figure 1 shows the NAVSTAR-6 visibility (> 50 deg) over Guildford as a function of elevation angle.

Similarly for narrow-band measurements about a dozen possible satellites were considered. Most of these satellites could not provide visibility of sufficient duration or were only visible outside mainland Europe. OSCAR-13 was identified as the only possible satellite source (S band) capable of providing the required visibility. However the severely limited dynamic range available from its beacon was considered inadequate. Therefore for narrow-band measurements the only possibility was to use a suitable alternative platform.

2.2. Alternative platforms

Alternative platforms can be used to transmit in any desired band and bandwidth. In addition the required receiver complexity is reduced. Some types of platforms offer the possibility of a complete control over the elevation angle - a very important parameter in this experiment. Another interesting feature was the possibility of transmitting multi-band,narrow and wide band signals,simultaneously.

The following types of platforms were considered ,

(i) Tower

(ii) Balloons of different types including

airship
(iii) Helicopter

After careful consideration airship and helicopter were retained as possible alternatives. Although a Helicopter can provide a complete control of the elevation angle, doubts were raised about its suitability in an urban environment with regards to the air-space regulations and difficulty in maintaining a constant elevation angle in situations where the mobile may only be visible for limited periods and its speed may vary (eg. due to traffic lights). It was also noted that the hardware to be used on a heliocopter would have to be suitably designed so as to minimize vibration effects. Thus further investigations were considered necessary.

2.3. Experimental configuration

A multi-band experiment with the capability of both narrow and wide band transmissions is capable of providing a variety of useful data under identical environmental and experimental conditions. Therefore this configuration was retained for a detailed study. It may be noted that at present this type of experiment can only be conducted by the use of a platform simulating satellite transmissions. It was assumed that a L-band experiment using GPS transmissions can also be conducted to increase the database and to quantify the confidence limits of the simulated measurements.

3. Main elements of experiment

The main elements of the proposed experiment are : a platform transmitting multi-band,narrow and wide-band signals which simulate the space segment and an instrumented mobile equipped with suitable receivers and data acquisition hardware,travelling at a given speed along a specified route - the ground segment. The data-acquisition was proposed to have the capability of analysing samples of data in real time but major data analysis was proposed to be performed off-line. Figure 2 shows the main sub-systems of the experiment. An elevation angle control mechanism and co-ordination elements have been included on the airborne platform together with the transmitter,antenna and data acquisition sub-systems.

3.1. Platform

As noted earlier, helicopters and airships or a combination of these two were considered as possible platforms. The main issues investigated were,
(i) Air-space restrictions
(ii) Cost
(iii) Vehicle tracking technique
(iv) Recordable data on the platform
(v) Overall feasibility/special requirements

Enquiries indicated that helicopters and airships have specific flying corridors in Guildford/London areas but these corridors could provide all types of environments. It was noted that helicopters are generally not permitted to fly below about 500 m whereas airship flights are not permitted above the same height. Both platforms can hover,permitting static measurements when necessary.

The use of a helicopter was favoured for the final phase of the study because of factors such as proven success of helicopters in previous experiments,cost consideration and its expected greater simulation accuracy .

It was however considered necessary that the antenna mount and other hardware be designed to ensure compatibility with an airship ,if this option was exercised later.

It was then necessary to identify a suitable technique for tracking the instrumented mobile for maintaining a constant geometric relationship between the mobile and the platform. Several possibilities were considered. These include the use of navigation information from GPS satellites , the use of a tracking antenna,the use of navigation information on the helicopter and the use of a visual tracking technique employed in an earlier experiment. The implementation of the visual tracking technique appeared simple, although it was noted that the available accuracy would be dependent on the skill of the pilot and may consequently vary.

The possibility of using this technique was discussed with a pilot and technical staff of a helicopter firm . The main conclusions of the meeting was that the technique seemed feasible and the tracking accuracy could be improved with careful planning. In addition the logistic problems were seen to be tractable. It was noted that any equipment to be mounted on the helicopter would have to be certified by the Civil Aviation Authority.

3.2. Frequency considerations

L (1.5 GHz), S (3 GHz), C (4/6 GHz), K_u (11 - 15 GHz) and K_a (20 - 35 GHz) bands were identified for the study. The process of obtaining the necessary frequency clearance can be very time consuming. Amateur bands exist close to these bands and were considered to offer minimum regulatory constraints for the experiment and were therefore used for defining the baseline transmitter/receiver hardware.

3.3. Transmitter and receiver design

The investigations into the design of transmitter and receiver equipment to provide narrow and wide band signals in these frequency bands has verified the feasibility of constructing the system considered [Racal study,(4)].

Since the rate at which the changes in multipath signals can be tracked during the broad-band experiment is very limited,the vehicle speeds suggested for these tests are typically below 12 % of those used for the narrow-band experiments. It might seem therefore that there was to be no advantage in developing hardware which could carry both narrow-band and broad-band reception simultaneously. However,since the cost of the additional hardware involved is likely to be a small proportion of the total cost of equipment procured,this remains a feasible approach.

The work required to fit an adjustable antenna array to the helicopter was considered to be a major cost factor and is expected to require the experience of specialised aircraft fitters. For this reason it was considered essential to carry out the work with an aircraft operating company with a particular interest or experience in developing similar externally mounted equipment.

3.4. Data types

The main types of data identified are real and quadrature components of the received signal for the narrow band transmissions and the complex impulse response for wide band transmissions. Other useful types of data identified are the elapsed time, mobile velocity and the elevation angle of platform. Some auxiliary data are useful during the analysis and interpretation of results. Suggested alternatives were voice records,key-board codes identifying broad terrain features,video records,receiver callibration and photographs. The use of video records was expected to be used in the development of a passive channel sounding technique. Such an approach using a camera and an photo-transistor has been attempted with some success before (6). Useful auxiliary data to be recorded on the helicopter include elapsed time,video/audio records,elevation angle variations (if possible),height of helicopter and other useful data if available.

3.5. Data acquisition and analysis

The use of a personal computer using a suitable interface with the receiver was the favoured data acquisition technique. The use of a tape recorder is proposed for recording the data set on the helicopter. One of the main problems in this experiment with regard to the data acquisition technique is the vast amounts of data generated.

The throughput for the wide-band measurements, without using any data reduction technique,is about 720 Mbytes/hour. This requires 21.6 Gbytes storage for a 30-hour experiment. However, it is possible to reduce the data storage requirement by using techniques such as real-time data reduction algorithms or reduction in quantization bits.

The throughput for the narrow-band measurements is much lower. Using data reduction through averaging in the K_u

and K_a bands - which contribute maximum data samples - it is possible to reduce the throughput even further. For example the throughput at 90 Km/hr reduces from 105.1 Mbyte/hr to 14.2 Mbyte/hr.The K_u and K_a band data however do not contain the doppler information.

3.5.1. Storage medium

The high throughputs expected implies that the storage device must have a large capacity. Optical disk,fixed hard disk and removable hard disk were considered.

A major constraint in the selection of the storage medium arises due to the mobile environment. Manufacturers usually do not guarantee a reliable operation in such an environment. The optical disk solution provides a likely solution in this experiment if the reliability consideration is waived. It is expected that the reliability can be improved by using ruggedized personal computer (PC) which together with the optical disk drive may be mounted on a vibration resistant mount.

Suitable software packages have been identified for both real-time data analysis and off-line data analysis. Formats for data presentation in line with previous ESA experiment (PRODAT campaign) were also identified.

3.6. Environment characterization

To optimize the measurements a representative sample of data must be obtained with a minimum of effort and cost. A careful selection of route can help to achieve this goal. Two possibilities were considered.

(i) Typical/worst case database

For each environment, a typical/worst case representative routes ($\approx$ a few kilometers) are chosen and data collected on the same route as a function of elevation/bearing angles. This should permit an accurate characterization of the channel as a function of elevation angle in a well controlled manner.

One difficulty in this approach is the lack of an acceptable definition/criterion for the selection of an environment. A typical locality in one coutry may not be representative of the whole Europe. It is felt that if 3-4 regions are carefully selected the resulting data base should provide statistics with acceptable confidence

(ii) Large area data-base

For each environment data are collected over a large area as a function of elevation angle. This approach has been used in many land mobile propagation experiments and is well suited when there is unlimited visibility ,eg.,from a satellite transmission.

The proposed solution is to use the first approach for the helicopter measurement and the second with GPS transmissions. The satellite measurements can provide unlimited data and provide confidence in the helicopter measurement.

Environment may be catagorized in several ways.For example, PRODAT results have been given for three categories ,viz.,open areas (eg.,motorways),partial obstruction (eg. suburbs of cities) and almost complete obstroction (eg.,city). For this experiment six categories were proposed. This should permit a better indication of the grade of service expected in various environments but it is noted that management of this database is more complex.

The proposed characterization and some possible locations in the UK are as follows :

(1) Dense urban - (city centre of a large city) :Central London

(2) Urban - (suburbs of a large city/city centre of a small city) : Outskirts of London,Guildford city centre

(3) Rural open - Motorways : M-3,M-25

(4) Rural/Suburban - (roads with some woods/residential areas/villages) : A-type roads around Guildford

(5) Rural dense forest : Some B-type roads around Guildford

(6) Mountainous regions : Not specified

Locations across mainland Europe can similarly be identified.

3.7. Experimental configuration

Two possible configurations for measurements are possible.

(1) Narrow-band and wide-band measurements performed on separate runs.
(2) Both the measurements performed on the same run

The second option minimizes the time duration of the experiment. Hence an investigation was performed to check the feasibility of this configuration.

It was noted that this would require simultaneously recording of 20 channels with a consequent increase in the overall throughput and added complexity in the data acquisition hardware. Further investigations showed that doppler information for wide-band measurements can be extracted only at very low vehicle speeds and in L - band (upto 10 km/h) and S-band (up to 5 km/h) only. This limitation arises due to the data reduction method used in the receiver (the sliding correlator approach). Moreover in practice relative speed variations of the same order of magnitude are possible between the mobile and the helicopter resulting in an additional doppler component . It may be expected that it will then be difficult to resolve components due to vehicle motion .

From the above discussion it is evident that only static wide-band measurements are possible for higher frequency bands ,i.e.,C,K_u and K_a bands. Even for L and S bands static measurements should be preferred due to the uncertainty of resolving the doppler components. Racal study (4) has shown that should it be decided to perform simultaneous measurements,the increase in the transmitter/receiver hardware complexity is not significant.

It is concluded that at high vehicle speeds only narrow-band measurements are possible. For static and possibly very low vehicle speeds both measurements can be done simultaneously. However, since static narrow-band measurements are of limited interest it is proposed that the two sets be performed separately to minimize the complexity.

3.8. Duration of campaign

For land mobile satellite service the events of interest depend on the immediate vicinity of the mobile. Hence it is expected that ,provided the routes are chosen with care ,a good confidence may be obtained with limited measurements. It was noted that existing databases vary from 1 Million samples to 750 Mbytes corresponding to a distance of from 50 km to over 750 kms.

For this experiment a 30 hour campaign was tentatively proposed for each region. The main purpose was to estimate the costs involved with this data-base requirements. A total of 3-4 regions were considered representative . A 20 hour period is allocated exclusively to the narrow-band measurement. This corresponds to 1000 km of data at an average speed of 50 km/h. However it can be noted that measurements over each route are repeated several times as a function of elevation angle.

A period of 10 hours was allocated for the wide-band measurement. This corresponds to 50 kms of data at an average speed of 5 km/h. The database is expected to provide an adequate base-line information to designers of future satellite sound broadcasting systems. The size of this database compares well with those of other experimenters. Moreover further data may be expected from the proposed propagation experiment using GPS transmissions.

4. Conclusions and present status

From the extensive search performed during the study,L-band P-code transmissions of the GPS satellites have been identified as the only satellite source for the broad-band measurements in the European region.

Considering the interest in multi-band - broad and narrow band channels it is concluded that a helicopter is the best platform for the experiment . The visual tracking technique appeared to be feasible but other alternatives are possible.

The investigations into the design of transmitter and receiver equipment to provide signals for the measurements

have verified the feasibility of constructing the system considered. It is concluded that the cost of the additional hardware required to perform simultaneous narrow-band and wide-band measurements is likely to be a small proportion of the total cost of equipment procured and hence remains a feasible approach. The work required to fit an adjustable antenna array to the helicopter is expected to constitute a major cost component and will require the experience of specialized aircraft fitters. For this reason it will be essential to carry out the work with an aircraft operating company with a particular interest or experience in developing similar externally mounted equipment.

A data acquisition system using a personal computer and the optical disk technology has been favoured to store the vast amounts of data which are expected to be generated in the mobile. A tape recorder is considered adequate for recording data on the helicopter. Suitable formats for presentation of the analysed result have been proposed.

A six-category environment characterization has been proposed.It is further suggested that the experiment be conducted in 3 - 4 carefully selected regions across Europe with 10 hours of broad-band and 20 hours of narrow-band measurements in each region. More data for the broad-band channel are expected to be available through measurements of transmissions from satellites in the GPS constellation.

The university is continuing further development of a scaled down version of the experiment. The experiment uses an alternative platform for transmitting CW beacons in S and lower K_u band and a mobile with a dual-band receiver and a data acquisition system using a personal computer. First results of the experiment are expected to be available in the following months.

Acknowledgement :

Major part of the work was performed under ESA contract no. 8084/88/NL/PB. Authors acknowledge valuable comments of Dr Arbesser-Rastberg (ESTEC), Dr J. Norbury (RAL,UK) and Mr I.E.Casewell (Racal-Decca Advanced Development Ltd,UK) during the course of the work.

REFERENCES

1. IEE colloquium on Highly Elliptical Orbit Satellite Systems,1989. Digest no. 1989/86 (London,24 May).

2. Richharia,M. 1989.Review of propagation aspects of highly elliptical orbits,IEE colloquium on Highly Elliptical Orbit Satellite Systems,Digest no. 1989/86 (London,24 May).

3. Advanced digital techniques for UHF satellite sound broadcasting, 1988,ed. R.J.Levey,Published by European Broadcasting Union Technical Centre.

4. University of Surrey,1989,A study on the feasibility of high elevation angle propagation experiments applicable to land mobile satellite services - Phase 1 and 2 reports,Contract no 8084/88/NL/PB.

5. Vogel W.J. and Hong S. (1988),Measurement and modelling of land mobile satellite propagation at UHF and L-band,IEEE Trans. on Antenna and Propagation,vol. 36,No. 5,P. 707-719.

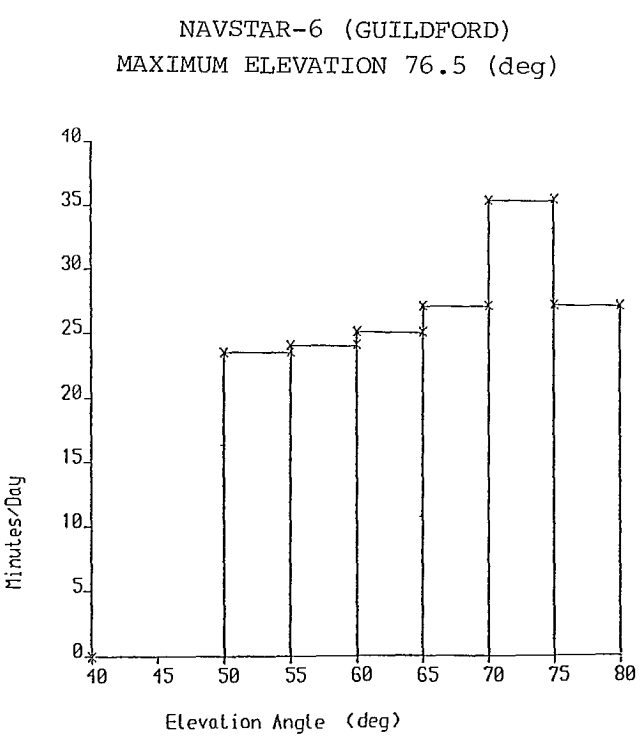


Figure 1: NAVSTAR–6 visibility/day over Guildford as a function of elevation angle

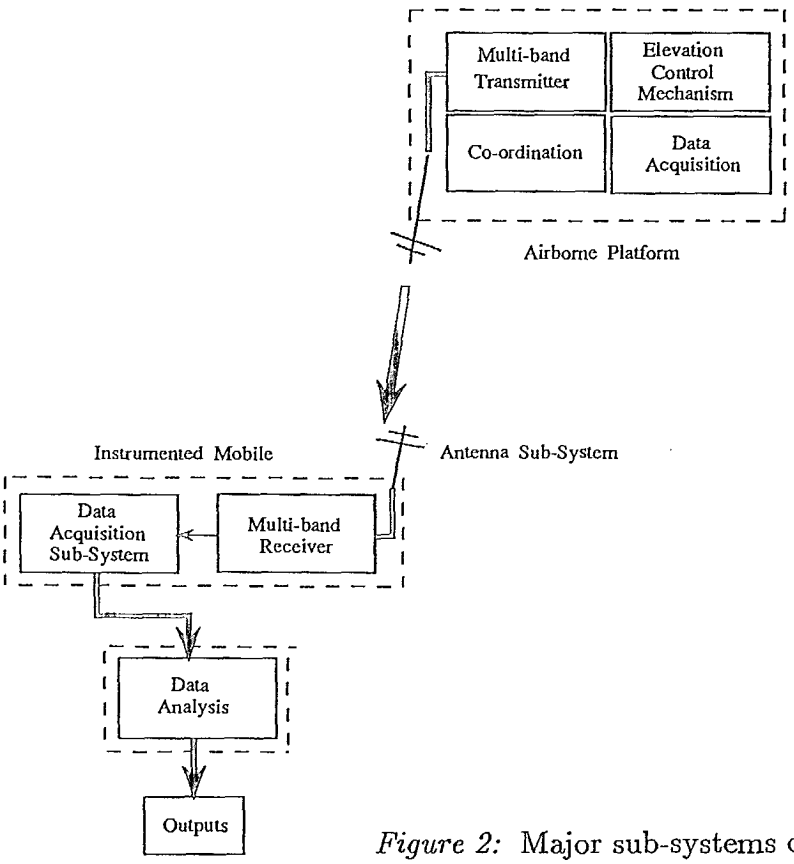


Figure 2: Major sub-systems of the experiment

Field Trials of a NASA-Developed Mobile Satellite Terminal

Khaled Dessouky, Richard F. Emerson, Loretta Ho, and Thomas C. Jedrey

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Dr
Pasadena, Ca 91109, U.S.A.
(818)354-3848
(818)354-6825 (FAX)

ABSTRACT

Various field trials have been performed to validate and optimize the technologies developed by the Mobile Satellite Experiment (MSAT-X). For each of the field experiments performed, a brief description of the experiment is provided, followed by a summary of the experimental results. Emphasis is placed on the two full scale land-mobile and aeronautical-mobile experiments. Experiments planned for the near future are also presented.

1. INTRODUCTION

The MSAT-X Task is dedicated to developing system concepts and high-risk technologies to enable the introduction of commercial land-mobile satellite communications in the United States.

MSAT-X developments have centered around the efficient utilization of the scarce resources of bandwidth, power, and orbital slots. The proposed architecture is based upon one or more bent-pipe satellite transponders; spot beams for mobile coverage; vehicular tracking antennas; low data rate speech compression; and modulation techniques which are both power and bandwidth efficient, and robust to the impairments of the mobile channel.

In this dynamic and challenging environment, field trials are imperative to validate the critical technologies and the underlying system architecture. The field trials, collectively called the Pilot Field Experiments (PiFEx), have evolved from testing individual subsystems, to end-to-end system demonstration. Accordingly, the objectives have been to: 1) evaluate the actual performance of the subsystem elements in relation to expected performance (theoretical or simulated), 2) evaluate the end-to-end performance of the system, 3) obtain data to support the enhancement of the system components and architecture, and 4) demonstrate system potential to end users and manufacturers.

To meet these objectives, a variety of field experiments have been formulated and conducted. In the following sections the ground terminal and test equipment are first described briefly. This is followed by a concise history of the PiFEx program and a more detailed description of the two full-scale aeronautical and land-mobile experiments. Experiments planned for the future are also described. Conclusions emerging from the field trials are then presented.

2. SYSTEM DESCRIPTION

Three distinct classes of equipment were used in the experiments: the MSAT-X terminal(s), a translator or satellite transponder, and the supporting test and data gathering equipment. For all experiments the terminals at both ends of the link were functionally similar and were basically copies of the MSAT-X mobile terminal. No formal network control center was required, although the required link set-up protocols were implemented in the terminals.

The major components of the mobile terminal [1] are the speech codec, the terminal processor, the modem, the transceiver, and the antenna. The speech codec provides good quality speech at 4800 bits per second (bps). The terminal processor acts as the heart of the mobile terminal, and implements all networking and control functions. The modem converts data from the terminal processor at 4800 bps into a baseband waveform, as well as demodulates a low IF from the receiver to provide 4800 bps data to the terminal processor. The transceiver up-converts the baseband waveforms from the modem to a suitable L-band transmit frequency. It also receives the signal at L-band, and down-converts the

signal to the IF required by the modem. The transceiver coherently demodulates and tracks a pilot signal and provides it to the antenna subsystem for tracking. It is also capable of using the recovered pilot as a reference for the down-conversion chain to the modem (this removes the one way Doppler from the received signal). The antennas developed for MSAT-X range from omni-directional drooping dipoles to mechanically and electronically steered arrays. Typically, the omnidirectional has been used for a reference pilot receiver (for propagation measurements) while the tracking antennas are used for data reception and transmission.

In addition to the basic communications terminal, enhancements for experimental purposes include a data acquisition system (DAS), an E_b/N_0 measurement subsystem, an audio record/playback unit, and various pieces of test equipment. The DAS performs real time analysis and displays key data for the benefit of the experimenters in the field. The DAS also records more extensive information for post-experiment analysis, such as baseband received pilot (in-phase and quadrature channels for propagation measurements), antenna pointing parameters, etc. For the mobile terminal (versus fixed) enhancements to the terminal have included a Loran C (position determining) receiver, a flux-gate compass with digital outputs, a digital speedometer, and a video camera. All of the outputs from these additional data systems, except the video camera, are also recorded by the DAS.

For field test purposes, an L-band translator was developed to allow simulated satellite tests using an antenna range or large tower.

3. PIFEX

The Pilot Field Experiments have covered a range of activities: from simulated satellite experiments using a translator and a tower as a satellite simulator, to full scale satellite experiments. A summary of these experiments is presented in Table 1.

Tower Experiments

A series of tests were performed using a 1000 foot NOAA tower in Erie, Colorado. These included antenna acquisition and tracking (Tower 1 [2]) using the JPL mechanically steered antenna; antenna tests and half duplex data transmissions using the MSAT-X modulator and the JPL and TRE steerable antennas (Tower 2 [3]); and full scale system tests (Tower 3 [4]). In Tower 3 full duplex speech and data transmissions were conducted using both the JPL and TRE antennas. Tower 1-3 served as a series of shakedown tests for the mobile terminal, and many operational obstacles and system deficiencies were identified and overcome. Indeed, further tests to be described below have shown that, in many respects, the tower set-up created an environment that was more severe than actual satellite environments.

Satellite 1a

The Satellite 1a experiment was the first of the PiFEx experiments using a real satellite, and was conducted in Santa Barbara, California during August of 1987 [5]. This experiment used a beacon from the MARISAT satellite to serve as a pilot signal for antenna tracking. The JPL mechanically steered antenna was used and a series of mobile tracking tests were run in the Santa Barbara area. This experiment was significant because it was the first time the antenna had tracked a true satellite, and in particular, with the very low elevation angle (13^o) to the satellite. The experiment verified the JPL antenna's mechanical robustness and tracking capabilities in a very demanding configuration (the JPL antenna was designed for a $20\text{-}60^o$ elevation range). In addition to verifying the antenna performance, numerous propagation results were obtained and analyzed.

MSAT-X/FAA/COMSAT/INMARSAT Experiment

The joint MSAT-X/FAA/COMSAT/INMARSAT experiment consisted of ground based experiments conducted during the first three weeks of January, 1989, and aeronautical experiments conducted during the last week of March 1989. The objectives of the experiments were to characterize the MSAT-X mobile terminal performance, for both the fixed ground and aeronautical-mobile satellite link environments. Link and equipment characterizations were performed by collecting both BER results at various signal to noise ratios (SNR) as well as evaluations of the speech link performance [6].

The experiment configuration is shown in Figure 1. The ground experiment consisted of a ground-to-ground full duplex communications link between the FAA Technical Center in Atlantic City, New Jersey, and the COMSAT ground station in Southbury, Connecticut, through the MARECS-B2 satellite. The aeronautical experiment was functionally similar to the ground segment. Experiments took place with

the aircraft stationary and with the aircraft following prescribed flight paths. In both parts of the experiment, the MSAT-X terminal was used except that passive dual helibowl antennas were used instead of the steerable antennas. Two antennas were used on the aircraft, one on each side of the fuselage attached to the inside of a passenger window. At the ground locations and during flight the elevation angle to the satellite was approximately 23^{o}.

The measured ground based BER for the forward link (FAAT.BER) and the return link (CEST.BER) are shown in Figure 2. Plotted on the same graph are the curves for simulation (SIM.BER) and laboratory hardware tests (LAB.BER), both for an additive white Gaussian noise (AWGN) channel. The experimental curve is about 0.5 dB from the laboratory measured performance for the forward link and very close to the laboratory results for the return link. The primary sources of degradation in the forward link are the operation of an automatic level control circuit in the satellite, and noisy pilot tracking at the receiver. Ground based voice links were established to demonstrate the digital speech coder. Both conversations and standard speech tapes were recorded. The voice quality was considered to be very favorable.

Once the equipment was installed in the aircraft, ground calibrations were performed to establish a benchmark for the flight tests. These calibration tests were found to differ from the previous ground-to-ground performance results by about +0.1 dB in each link due to window aperture effects.

Following the calibrations, two flight tests were performed. The flight paths followed are detailed in Figure 3. During both flights, the aircraft was flown at an altitude between eight and nine thousand feet, with ground speed that ranged from 180 to 290 knots.

Along both paths of the first flight, heavy turbulence due to severe thunderstorms was encountered. The average performance in the forward link for both legs of the flight (ACT.329) and the return link for both legs (CEST.329) are shown in Figure 4. Also shown in this figure is the average forward (ACT.328) and return (CEST.328) link performance for the aircraft ground calibration. There is approximately a 0.8-1.0 dB degradation in link performance due to the aeronautical environment. This degradation comes from several factors, including the pitch and roll of the aircraft caused by the heavy turbulence, and the change in the received Doppler (which varied on the northerly flight path from approximately 128 Hz at one end to 79 Hz at the other of the flight path).

The experiments conducted during the second flight consisted of data transmissions from the aircraft to the ground and speech demonstrations. At course changes the received Doppler changed rapidly and over a fairly wide range (e.g., +218 Hz to -223 Hz as shown in Figure 3). Residual Doppler on the order of +/-100 Hz remained after coarse compensation was performed using the transceivers. The BER results for the return link from the portion of the flight path from Charleston, South Carolina to the first major course change are shown in Figure 5 (CEST.RHS.331). For comparison, the return link performance of the ground based aircraft terminal is also shown (CEST.328). As can be seen, the performance is very close to the ground tests. The primary reason for this is that little turbulence was encountered on this portion of the flight.

During both flights, the full-duplex MSAT-X voice link was established often and used as the main, and, in fact, the only available direct voice link between the experimenters on the aircraft and in the CES. There was no perceptible difference in speech quality or intelligibility between in-flight and ground operations. In particular, the background jet noise had no significant effect on the communications.

In summary, while the link between the aircraft and the ground was more dynamic than expected, the operation of the MSAT-X mobile terminal was very close to theory/simulation and laboratory results.

MSAT-X/AUSSAT Experiment

The joint MSAT-X/AUSSAT experiment was conducted in Australia from July 17th through August 2, 1989. The experiment tested for the first time, the MSAT-X technologies and equipment in a true land-mobile satellite environment. Speech and data communications were demonstrated and tests were performed to characterize, quantitatively and qualitatively, MSAT-X system performance. Extensive data was recorded for various system parameters, as well as for vehicle antenna validation and propagation studies. A secure voice experiment for the United States National Communications System (NCS) Agency was also performed.

The experiment configuration is shown in Figure 6. The land-mobile satellite link consisted of a fixed

hub station at AUSSAT Headquarters in Sydney, the Japanese Experimental Technologies Satellite (ETS-V), and JPL's MSAT-X van. The van and hub station contained the basic MSAT-X communications terminal and other test and data-acquisition equipment. Two steerable antennas were available for use on the van: the JPL mechanically steered antenna, and the electronically steered phased-array antenna developed for JPL by Teledyne Ryan Electronics (TRE). (See [1] and the references therein.)

Mobile tests were conducted along Highway 1 between Sydney and Brisbane. Elevation angles to the satellite ranged from 51^{o} to 57^{o}. A variety of environmental conditions were encountered from clear line-of-sight to heavy shadowing. The received pilot signal power on the forward link (hub-to-van) during a clear condition is shown in Figure 7. The constant overall signal level resembles an AWGN channel. The Gaussian channel is channel present under stationary conditions or in the in the absence of multipath. In contrast, a lightly shadowed case is shown in Figure 8. Here, considerably more signal variation is seen. The probability density functions derived from the experimental data for these two environments are shown in Figure 9.

The clear channel density function was fitted with an analytically-derived Rician density. This is shown also in Figure 9. The K-factor (ratio of direct to scattered power in a multipath environment) is 16.5 dB. It should be noted that K may have actually been higher; however, the noise present on the pilot (approximately 42 dB.Hz C/N_0), and the noise inherent in the data acquisition system limited the values of observable K. This high value of K (16.5 dB) shows that the distribution approaches the Gaussian. This reiterates the observation that the clear mobile channel, with a medium gain vehicle antenna, and at a sufficiently high elevation angle, is approximately an AWGN link.

Both antennas tracking subsystems operated well and performed as expected. The modem and speech coder subsystems also performed well. Preliminary analyses have indicated that the return link performance under generally clear conditions was 8.5 to 9.0 dB for a BER of about 10^{-3}. The forward link required 8.5 to 9.5 dB E_b/N_0 for the same performance. These results show approximately a 0.2 to 1.2 dB degradation relative to hardware performance in the lab.

Overall, and in a qualitative manner as well, the MSAT-X system demonstrated good and robust performance. Good speech quality was observed for the many voice test tapes used. The robustness of the system was shown by a nearly continuous two hour voice link, during which synchronization was maintained over the range of environmental conditions. Short blockages resulted only in short bursts of codec induced silence or, seldom, garbled speech.

Detailed data analysis of the experiment is on-going and more complete results will be published in forthcoming articles.

Multipath Rejection Measurement Experiment (MRMEx)

MRMEx is scheduled for July/August 1990. It will evaluate the ability of directive antennas to discriminate against multipath signals. This will lead to an estimate of the attendant system performance improvement. The mobile terminal will be mounted on the MSAT-X van and will receive signals from a satellite while the van is driven on various roads. A variety of propagation environments will be encountered. The signal will be received simultaneously through both an omni and a directive antenna. Several tracking, directive antennas will be tested. The received in-phase and quadrature signals will be recorded for both antennas simultaneously. The data will be analyzed to determine the propagation characteristics observed with each antenna type.

4. CONCLUSIONS

Both the quantitative and qualitative results of the two latest field experiments show the viability and applicability of the MSAT-X architecture and technology. Technological risks associated with the implementation of a first generation MSS have been shown to be minimal. Commercialization of the speech codec and modem technology is already underway. Commercial implementation of mechanically steered antennas is feasible today. Some additional development work, however, remains in the area of electronically steered antennas. The greatest challenge will be in reducing the cost of the phased arrays. Through the combination of research, development and field testing, MSAT-X technology has now been shown to be a viable MSS option that can be adopted by U.S. industry.

ACKNOWLEDGMENT

NASA and JPL wish to thank the many organizations and individuals who made these experiments possible. The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. N. Lay, T. Jedrey and J. Parkyn, "Description and Performance of a Digital Mobile Satellite Terminal", IMSC'90, Terminal Equipment Session, these proceedings.

2. MSAT-X Quarterly No. 13, "Special Issue on PiFEx Tower-1 Experiments," JPL 410-13-13, January 1988.

3. J. Berner, "The PiFEx Tower 2 Experiment," MSAT-X Quarterly No. 15, JPL 410-13-15, June 1988.

4. K. Dessouky, T. Jedrey and L. Ho, "Summary of Results from the Tower-3 Experiment", MSAT-X Quarterly No. 20, JPL 410-13-20, July 1989.

5. K. Dessouky and L. Ho, "Propagation Results from the Satellite 1a Experiment", MSAT-X Quarterly No. 17, JPL 410-13-17, October 1988.

6. T. Jedrey, K. Dessouky and N. Lay, "An Aeronautical-Mobile Satellite Experiment", JPL report in preparation.

Table 1. Summary of Pilot Field Experiments

Experiment	Date	Purpose
Tower 1	Winter '87	Antenna acquisition & tracking
Satellite 1a	Summer '87	Antenna acquisition & tracking
Tower 2	Fall ' 87	Antenna acquisition & tracking & half duplex data
Tower 3	Summer' 88	End-to-end data & voice performance & demonstration
MSAT-X/FAA	Winter' 89	Fixed ground and aeronautical-mobile end-to-end
MSAT-X/AUSSAT	Summer' 89	Full scale land-mobile
MRMEx	Summer' 90	Antenna multipath rejection

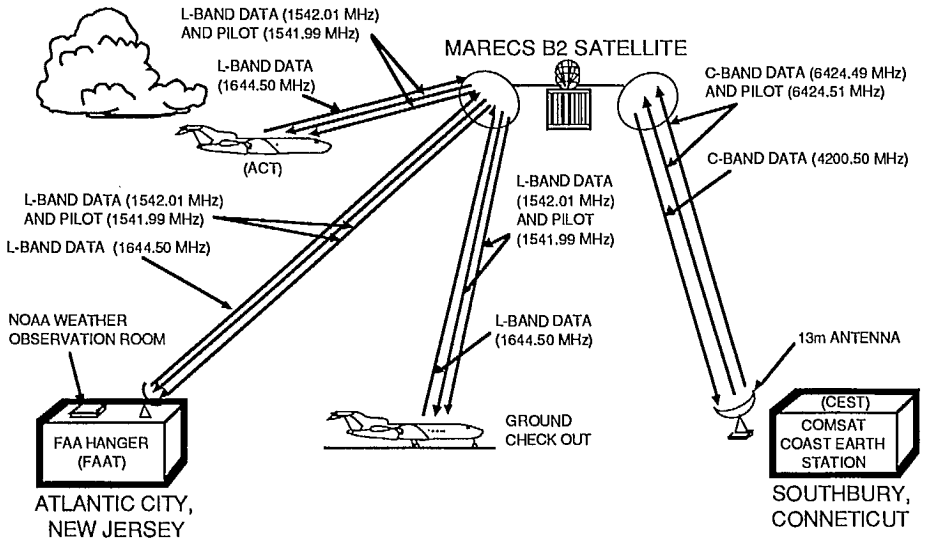


Figure 1. MSAT-X/FAA/COMSAT/INMARSAT Experiment

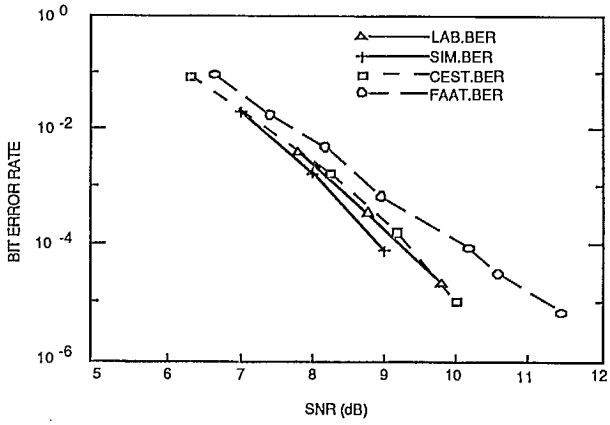


Figure 2. Ground Based Forward and Return Link BER Measurements (SNR of 8 dB = C/N_0 of 44.8 dB.Hz)

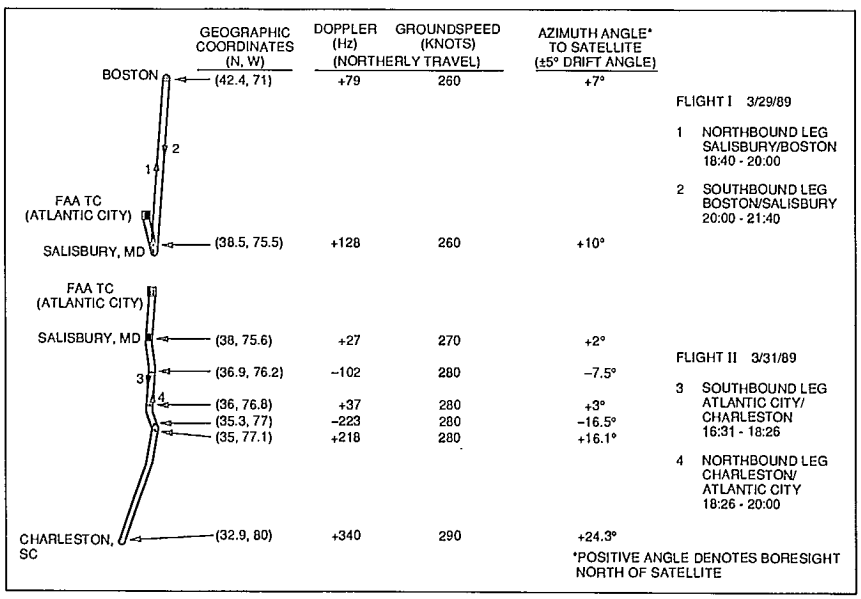


Figure 3. Flight Paths

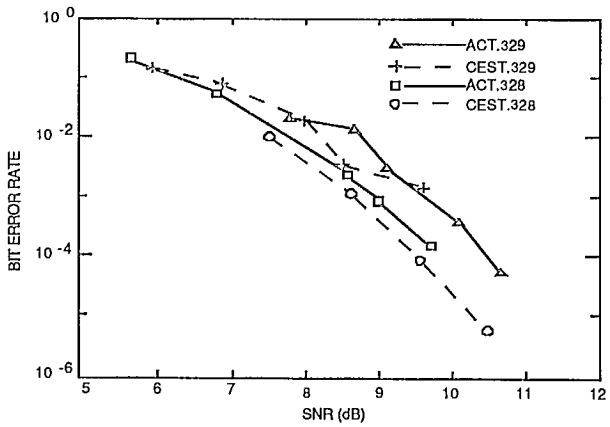


Figure 4. Flight #1 BER Performance

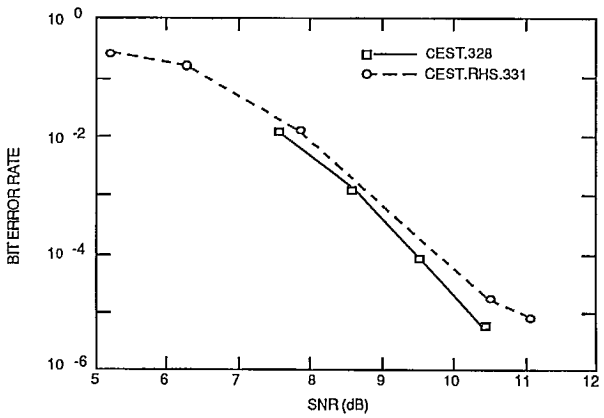


Figure 5. Return Link BER Performance from Flight #2

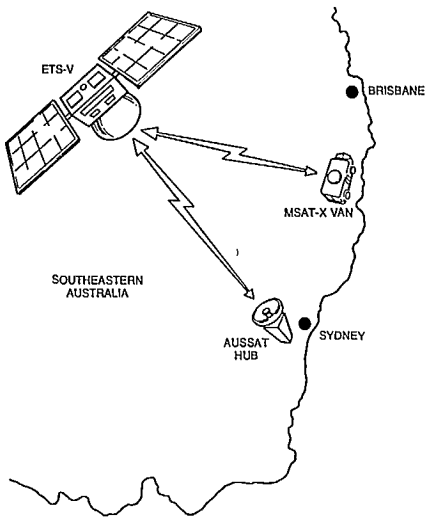


Figure 6. MSAT-X/AUSSAT Experiment Configuration

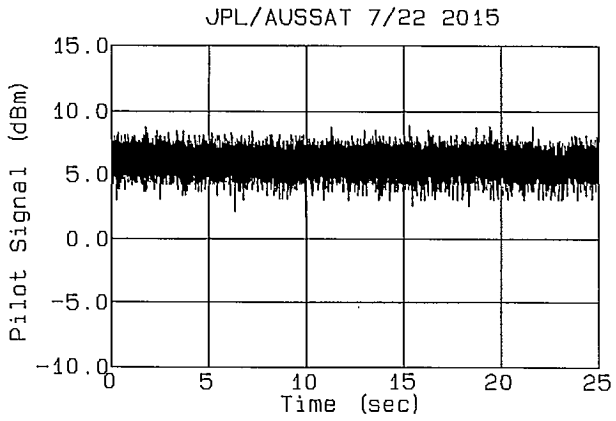


Figure 7. Received Pilot on Clear Channel

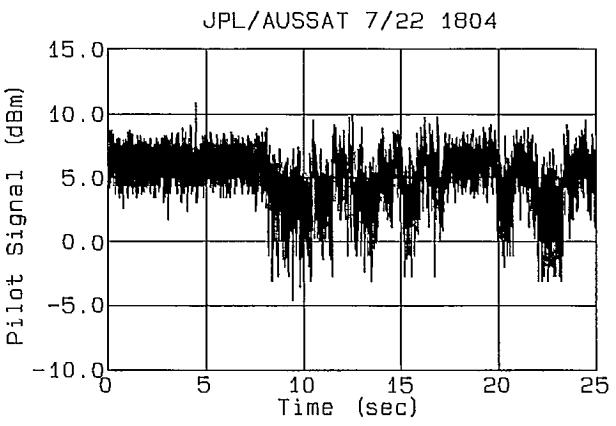


Figure 8. Received Pilot on Channel with Light to Moderate Shadowing

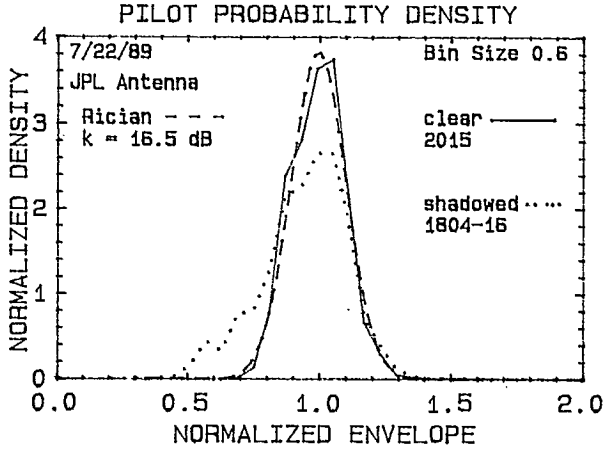


Figure 9. Probability Density Functions for Envelope of Received Pilot

Spread Spectrum Mobile Communication Experiment Using ETS-V Satellite

Tetsushi Ikegami*, Ryutaro Suzuki**, Naoto Kadowaki*
Shinichi Taira* and Nobuyasu Sato***

*Kashima Space Research Center,
Communications Research Laboratory,
Ministry of Posts and Telecommunications
893-1 Hirai, Kashima, Kashima, Ibaraki 314 Japan
Phone : +81-299-82-1211 FAX : +81-299-83-5728

**Space Communications Division, CRL, MPT
4-2-1 Nukui-kita, Koganei, Tokyo 184 Japan

***Toshiba Corporation
Komukai-Toshibacho, Saiwai, Kawasaki 210 Japan

ABSTRACT

The spread spectrum technique is attractive for application to mobile satellite communications, because of its random access capability, immunity to inter-system interference and robustness to overloading. A novel direct sequence spread spectrum communication equipment is developed for land mobile satellite applications. The equipment is developed based on a matched filter technique to improve the initial acquisition performance. The data rate is 2.4kbps and the PN clock rate is 2.4552MHz. This equipment also has a function of measuring the multipath delay profile of land mobile satellite channel, making use of a correlation property of a PN code. This paper gives an outline of the equipment and the field test results with ETS-V satellite.

INTRODUCTION

The first generation of land mobile satellite communications systems will be introduced by using analog ACSSB and/or 4.8kbps digital modulation. The spread spectrum technique is also attractive for application to satellite communications, because of its random access capability, immunity to the interference with other systems and robustness to overloading of transponders. There are some articles that treat comparisons among FDMA, TDMA and SSMA systems [1][2]. However, further discussions, based on the real land mobile satellite channel, are needed.

Recently, Communications Research Laboratory has developed a novel direct sequence spread spectrum communication equipment for land mobile satellite applications, and conducted field tests with the ETS-V satellite [3] in suburban areas. The design objectives are to shorten the initial acquisition time and to keep a good performance in fading conditions. To meet these objectives, a coherent matched filter (CMF) technique [4] in which digital matched filters are employed to make code acquisition, AFC and coherent detection of data, is used in the equipment. This equipment can also operate as a multipath measurement system making use of a correlation property of a spreading PN code. This paper gives an outline of the equipment used and the test results.

PRINCIPLES OF COHERENT MATCHED FILTER TECHNIQUE

In order to extract the desired information bit stream from a matched filter (MF) output directly, an RF signal should be down-converted to baseband by using the recovered carrier signal. However, the carrier to noise power ratio at the input to the MF is generally very low. The coherent matched filter (CMF) technique makes possible the carrier

recovery from the received low C/N SS signals. The CMF circuit is shown in Fig. 1. The principle of the carrier recovery circuit in Fig. 1 is based on that of a Costas loop. The special feature of this technique is that correlators are inserted into both the I and Q arm filters of the Costas loop.

The transmitted signal can be expressed as

$$S(t)=I(T_f)PN(t)\cos(\omega t+\psi(t)), \quad (1)$$

where, $I(T_f)=\pm 1$: information signal,
T_f : PN frame rate= data rate,
$PN(t)=\pm 1$: PN sequence,
ω : carrier angular frequency,
$\psi(t)$: carrier phase.

At the receiving side, S(t) is received and down-converted by using local carriers which are orthogonal each other. These local carriers are expressed as

$$S_i(t)=\cos\omega t, \quad (2)$$
$$S_q(t)=\sin\omega t. \quad (3)$$

By using eqs. (1),(2) and (3), the baseband signals in the I and Q arms are obtained as follows,

$$B_i(t)=(1/2)I(T_f)PN(t)\cos\psi(t), \quad (4)$$
$$B_q(t)=(1/2)I(T_f)PN(t)\sin\psi(t). \quad (5)$$

Then, $B_i(t)$ and $B_q(t)$ are sampled with the PN clock period and put into the correlators. When $\psi(t)$ changes gradually over a long period, as compared with the length of the PN sequence, it can be assumed that $\psi(t)$ is a constant ψ during one frame of the PN sequence. As a result, the correlator outputs are given as

$$C_i(m)=(1/2)R(m)\cos\psi, \quad (6)$$
$$C_q(m)=(1/2)R(m)\sin\psi, \quad (7)$$

where $R(m)=\sum_{k=0}^{N-1}PN(k)PN(m+k)I(m+k)$, (8)

$m=[t/T]$,
[ ] : Gauss bracket,
T : PN chip rate,
N : length of PN sequence.

Eq. (8) is the cross-correlation between the received PN sequence and the reference PN sequence. Since both PN sequences are the same, R(m)is the autocorrelation function of the PN sequence. The amplitude of the autocorrelation peak R(m), equals N at $m=nN$ ($n=0, 1, 2, \ldots$) and $R(m)=1/N$ at $m\neq nN$. In the same manner as for the Costas loop, an error signal is generated at $m=nN$ by multiplying C_i and C_q in eqs. (6) and (7) resulting in

$$C(nN)=-(1/8)I^2(nN)R^2(0)\sin(2\psi), \quad (9)$$

where $R(0)=R(nN)$ is the autocorrelation peak.

Since $I^2(nN)$ and $R^2(0)$ are positive and constant, eq. (9) is a function of ψ only. This phase error signal is generated once per PN frame, and controls the VCO which generates the local carrier. Consequently, the carrier recovery and coherent detection of the data can be accomplished.

OUTLINE OF SS EQUIPMENT

The SS equipment was developed based on the CMF technique, using 8bit digital correlators. Major specifications and a block diagram of the equipment are shown in Table 1 and Fig.2.

An antenna system is composed of two micro-strip patches excited in a higher mode and two patches used separately for transmitting and receiving. Each antenna has an ominidirectional beam. A linear amplifier is used for transmitting to avoid any broadening effect of the filtered BPSK spectrum. A Linear Predictive Coding (LPC) vocoder of 2.4kbps is installed as voice codec. The voice data is spread with a PN sequence of length 1023 and then modulates the carrier signal by BPSK. The spectrum is band limited to 3MHz, which corresponds to the bandwidth of the ETS-V transponder. A Unique Word (UW) of 31bits PN sequence is used instead of data when the initial acquisition is performed.

In the receiver, AFC,synchronization with the spreading PN code and coherent demodulation of the data can be achieved with the CMF technique. The received signal is down-converted into I and Q channels and digitized to 8bits at a rate of twice the chip rate. The digitized signals are then processed by digital correlators of length of 2046, twice the code length. The matched pulse appears every 2046

cycles of the clock at the correlator outputs and the data clock of 2.4kHz is recovered by detecting peaks of the matched pulse train. The product of matched pulses in both I and Q ch. has the information on phase error of the local carrier as mentioned in the previous section. The coherent detection is achieved using this error signal as in a Costas loop receiver.

The initial acquisition and AFC are accomplished as follows. There are two stages of frequency search modes during the initial acquisition, the coarse search mode and the fine search mode. In the corse search mode, the frequency slot where the correlation peak is maximum is determined by sweeping the local carrier in 1kHz steps from -10kHz to +10kHz. The correlation peak is detected by finding the maximum after recursive integration over 32 PN frames. After finding the frequency slot, the fine search mode is started at that slot. In the fine search mode, the local carrier is swept by 30Hz steps within a ±510Hz bandwidth. The sweep is stopped when the carrier recovery loop is locked to the received carrier and the UW is successfully detected. After initial acquisition has been accomplished, frequency tracking is performed by the carrier recovery loop. When the frequency drift accumulates, the local carrier frequency is shifted.

This equipment can be used as a multipath measuring receiver as well as a communication terminal. When the base station transmits the spreading PN signal without any information data, the correlator outputs of the matched filter are time correlation functions between the received PN signal which contains multipath components and the reference PN signal. As the autocorrelation function of the PN signal exhibits a peak at zero time shift and at multiple of PN length N, and is nearly zero everywhere else, this correlator output corresponds to the multipath delay profile, i.e. impulse response of the channel. Fig.3 shows the multipath measuring adapter. Recursive integrators are used to improve the S/N of the outputs.

EXPERIMENTAL RESULTS

Loop-Back Test

Figs.4(a)(b)(c) show the loop-back test results of initial acquisition time performance at transmitting frequency offsets of -8kHz, 0 and +8kHz, respectively. The initial acquisition time is defined as the time interval between when the SS signal transmission is started and when the UW is detected at the receiver. In these figures, the bars indicate the minima and the maxima and the dots indicate the average over 30 trials. The mean acquisition times are 900ms at f=-8kHz, 800ms at f=0 and 700ms at f=+8kHz, respectively, when C/No is greater than 44dBHz. These values correspond to the theoretical prediction of frequency sweep time. Therefore it can be seen that code acquisition and carrier recovery are accomplished in a very short time compared to frequency sweep. The lowest C/No where the initial acquisition succeeded is 36dBHz, the synchronization is lost within a few seconds at levels below 38dBHz, however.

Fig.5 shows the bit error performance for loop-back tests in a Gaussian noise environment. The solid curve indicates the theoretical BPSK performance with differential encoding and coherent detection. The experimental result shows good agreement with the theoretical curve: the degradation is only 0.7dB. The degradation due to the limited quantizer levels in the matched filter was considered in a computer simulation. From our simulation, a four bit quantizer is sufficient for the case of a white Gaussian noise channel. The dotted line indicates the experimental results with only a 1bit quantized correlator, which was developed previously [4]. The great advantage of using multi-bit quantized correlators can be seen.

Field Tests

The field test was performed with the ETS-V satellite (150° E). Table 2 shows an example of the link budget. The base earth station transmitted both the SS signal and a CW carrier for simultaneously measuring C/No. The equipment installed in a measuring van received both signals. The elevation angle to the ETS-V is around 47 degrees. In order to evaluate basic performance,

there was no interfering signal other than the desired SS and CW signals. The frequency separation between the SS and CW signals was 4.5MHz. During the experiment the main obstacles that caused fading or shadowing were roadside trees and utility poles. Data such as signal level, data error pattern, state of synchronization and running conditions were recorded in a data recorder.

Fig. 6 shows the cumulative time distribution of the received signal level. The dotted line indicates the theoretical Rician statistics with K, the ratio of direct component power to diffused component power, equals 12dB, and agrees well with the measured level above 10% of time. The received signal level below 10% of time does not exhibit Rician statistics and corresponds to shadowing by obstacles. Fig. 7 shows the relation between BER and C/No, averaged over 20 seconds. The solid line indicates the theoretical curve for BPSK in a Rician fading channel with K=12dB, which matches the measured cumulative distribution of the received signal level in Fig. 6. There are some points with higher BER of around 0.1 at low C/No condition. This is due to erroneous data being output during the sync loss condition caused by shadowing by obstacles. The BER performance agrees well with that of BPSK with Rician fading statistics, and the spreading of the spectrum to 3MHz has not affected the BER performance in this experiment.

CONCLUSION

The outline of the spread spectrum communication equipment developed was briefly discussed. The fundamental performances such as initial acquisition time and bit error rate were shown to be acceptable. The field test experiments were conducted in a suburban area with the ETS-V satellite. The BER performance in the field tests corresponded to that of non-SS BPSK on a Rician fading channel. However, further data collection must be needed in urban, suburban and rural areas to evaluate the effects of fading and multipath. The results of the multipath measurement using this equipment will be reported near future.

REFERENCES

1. Jacobs, I. M. et. al. 1988. Comparison of CDMA and FDMA for MOBILESTAR System. Proc. Mobile Sat. Conf. pp.303-308.
2. Proceedings of the Mobile Satellite System Architectures and Multiple Access Technique Workshop. 1989. JPL pub. 89-13.
3. Ikegami, T. et. al. 1987. Land Mobile Satellite Communication Experiments with ETS-V Satellite. IEEE VTC 87, pp.166-169.
4. Suzuki, R. et. al. 1986. Spread Spectrum Satellite Communication Terminal with Coherent Matched Filter. IEEE Globecom 86. pp.21.3.1-21.3.5.

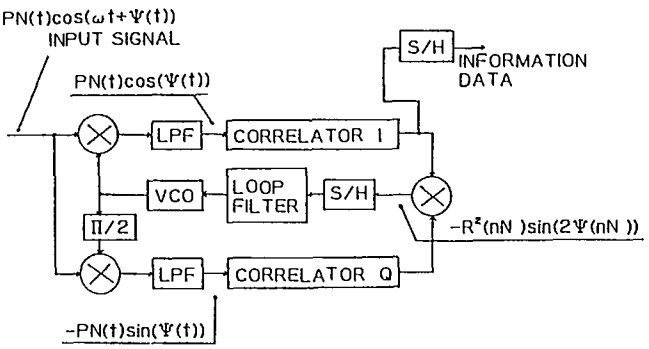


Fig. 1. Principle of CMF loop

Table 1. Major specification of SS equipment

Antenna	Higher Mode Microstrip Patch, 6dBi
RF Frequency	1648.5MHz(Tx) / 1546.5MHz(Rx)
HPA	20W Linear
CODEC	LPC Vocoder
Data Rate	2.4kbps
PN Code	M-sequence, Length 1023
PN Chip Rate	2.4552MHz
Modulation	BPSK, Differential Encoding
Demodulation	CMF, Coherent Detection
Matched Filter	8bits Digital Correlator, 2046stages
AFC Range	±10kHz

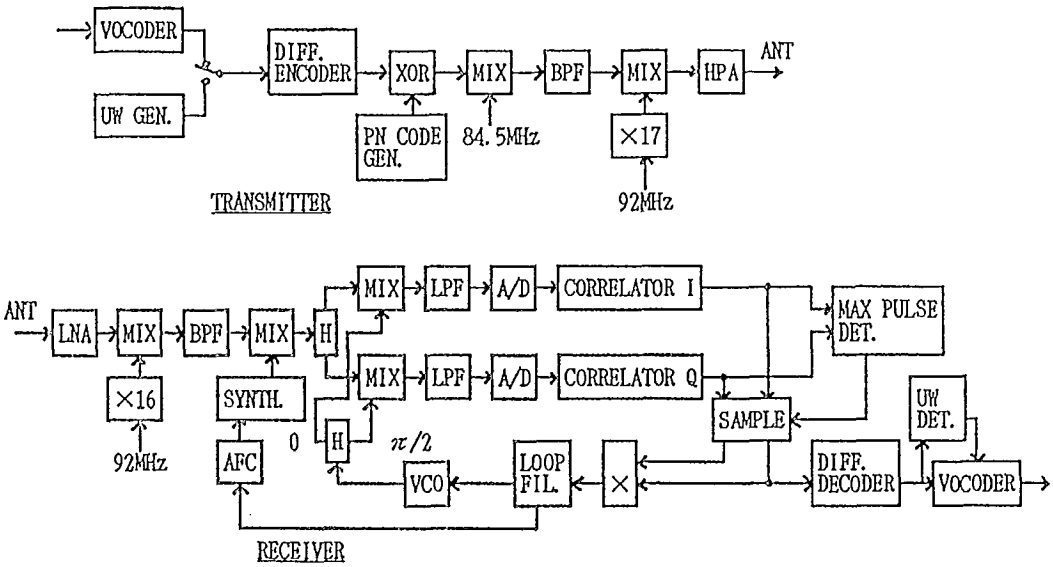


Fig. 2. Block diagram of SS equipment

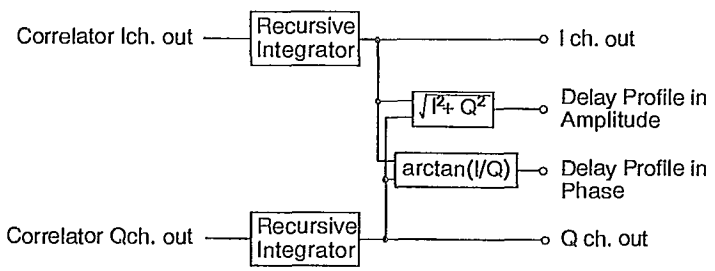
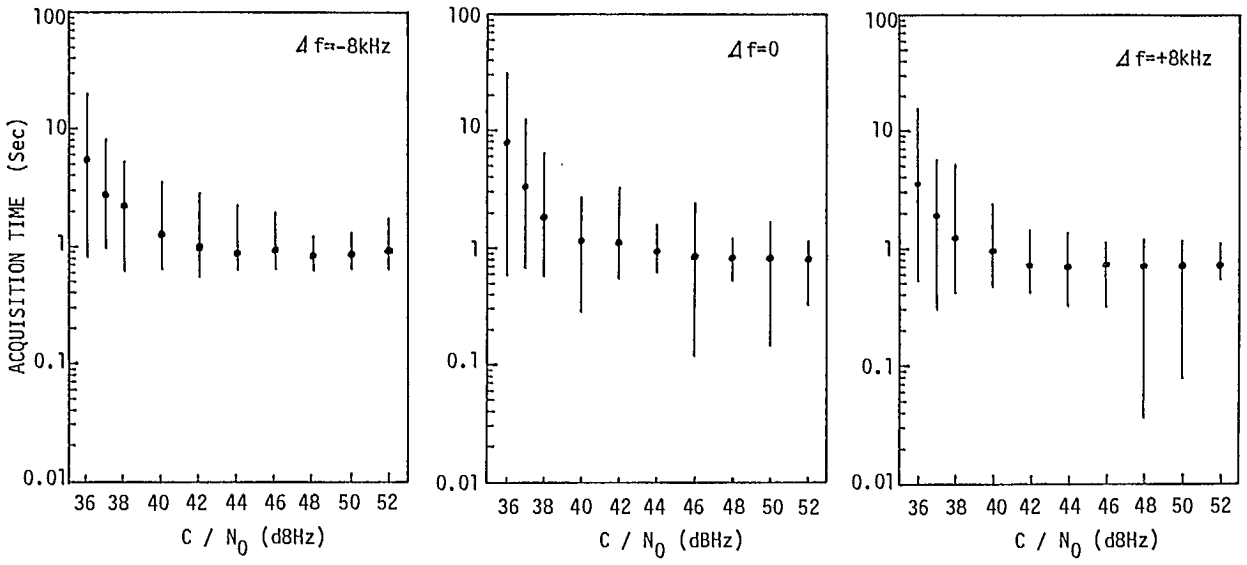


Fig. 3. Block diagram of multipath measurement adapter



(a) Frequency offset -8kHz (b) Frequency offset 0 (c) Frequency offset +8kHz

Fig. 4. Initial acquisition time performance in a loop-back test

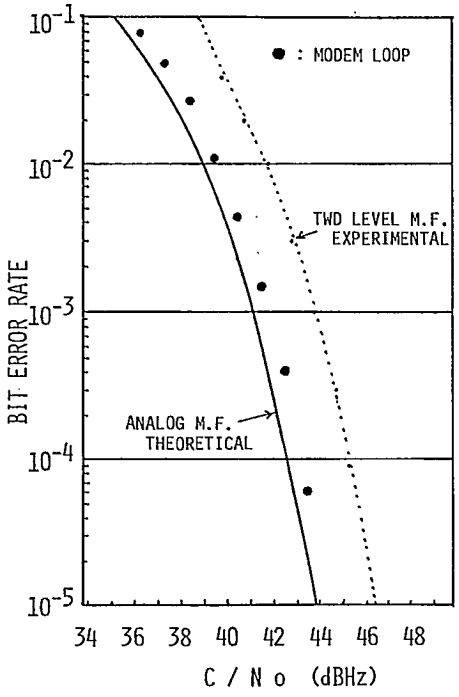


Fig. 5. Bit error rate performance in a loop-back test

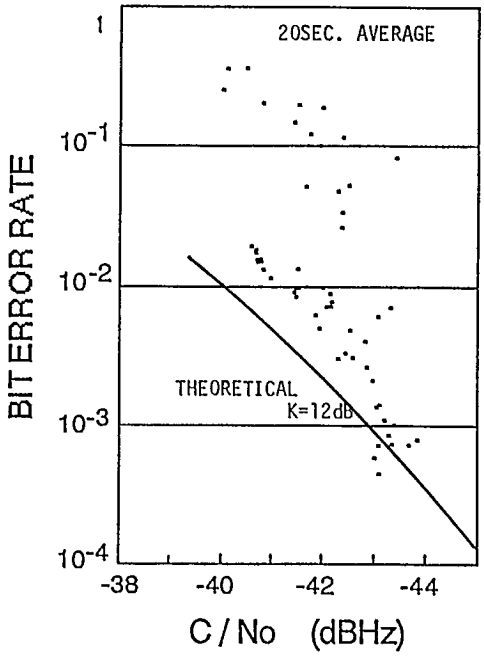


Fig. 7. Bit error rate performance in the field test

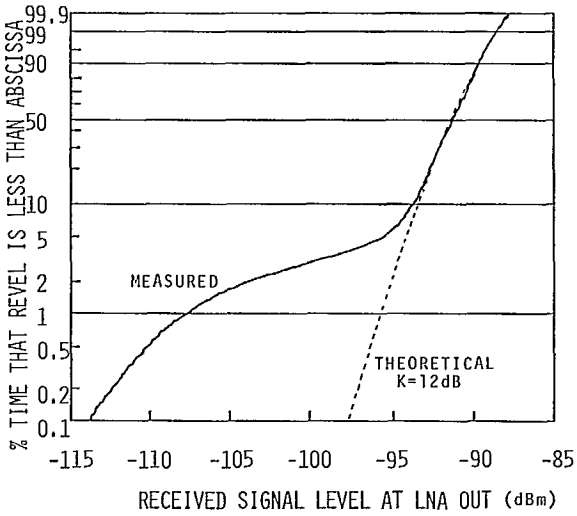


Fig. 6. Cumulative time distribution of received level in the field test

Table 2. Link budget

Base Earth Station to Mobile Earth Station	
Base station EIRP/ch.	60.2dBW
Up Link Path Loss (6GHz)	199.4dB
Sat. G/T	-8.1dBK
Up Link C/No	81.1dBHz
Transponder Gain	127.7dB
Sat. EIRP/ch.	26.7dBW
Down Link Path Loss (1.5GHz)	187.6dB
Antenna Gain of MES	6.0dBi
Feeder Loss	0.7dB
Received signal Level	-156.3dBW
MES G/T	-19dBK
Down Link C/No	48.0dBHz
Overall C/No	48.0dBHz
Required C/No	44.0dBHz
Margin	4.0dB

Mobile Earth Station to Base Earth Station	
MES Tx Power	13.0dBW
Antenna Gain	6.0dBi
Feeder Loss	0.7dB
MES EIRP	18.3dBW
Up Link Path Loss (1.6GHz)	188.2dB
Sat. G/T	-4.5dBK
Up link C/No	54.1dBHz
Transponder Gain	127.8dB
Sat. EIRP/ch.	-3.9dBW
Down Link Path Loss (5GHz)	198.2dB
Base Station G/T	32.7dBK
Down Link C/No	58.9dBHz
Overall C/No	52.9dBHz
Required C/No	44.0dBHz
Margin	8.9dB

An Ocean Scatter Propagation Model for Aeronautical Satellite Communication Applications

K.W. Moreland
Communications Research Centre
3701 Carling Ave., Ottawa, Ontario
K2H 8S2, Canada
Phone: 613-990-8287, FAX: 613-990-7987

ABSTRACT

In this paper an ocean scatter propagation model, developed for aircraft-to-satellite (aeronautical) applications, is described. The purpose of the propagation model is to characterize the behaviour of sea reflected multipath as a function of physical propagation path parameters. An accurate validation against the theoretical far-field solution for a perfectly conducting sinusoidal surface is provided. Simulation results for typical L-band aeronautical applications with low complexity antennas are presented.

INTRODUCTION

For L-band oceanic mobile satellite communications, sea reflected multipath is the most significant propagation component, especially when relatively low gain, non-directive antennas are employed. A propagation model is required to characterize the behaviour of sea reflected multipath as a function of the elevation angle to the satellite, aircraft altitude, sea state (surface roughness), aircraft antenna characteristics, aircraft velocity, and signal polarization.

Evaluating the electromagnetic field scattered by a rough surface is an extremely difficult problem. Most theoretical approaches are based on the Helmholtz surface integral, with the electric field on the surface assumed to match that present on a tangent plane at that point [1]. In general, complicated integrals that are difficult to solve are obtained.

In many theoretical treatments, attention is restricted to perfectly conducting surfaces in order to simplify the scattering solution. In the perfectly conducting case, the surface reflection coefficient has unity magnitude, independent of grazing angle. This is not appropriate for the sea surface, as is evident from the behaviour of the reflection coefficients at L-band shown in Figure 1. In most theoretical treatments, shadowing and blockage by the surface is ignored for mathematical tractability. Unfortunately, this makes low elevation angle scatter predictions suspect. The propagation model described in this paper incorporates the variability of the surface reflection coefficient as well as shadowing and blockage effects.

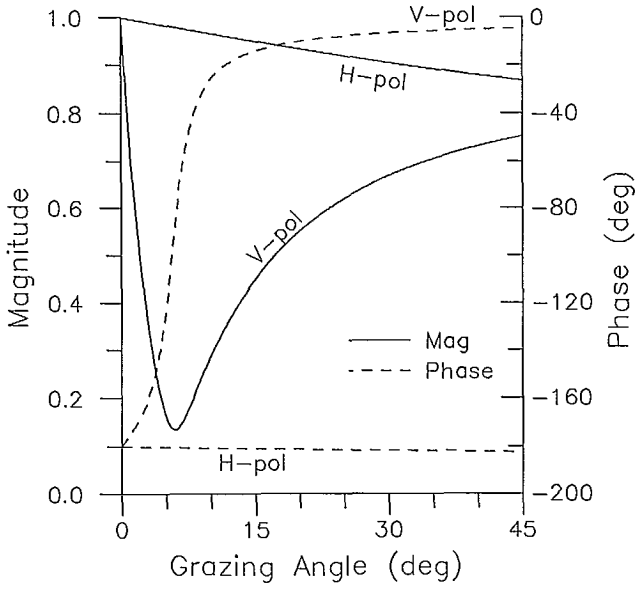


Figure 1: Sea surface reflection coefficients at L-band (λ_c = 0.2 m).

DESCRIPTION OF THE PHYSICAL PROPAGATION MODEL

A simplified view of the physical model that was implemented as a computer simulation is provided in Figure 2. Along the azimuthal direction to the satellite, a rough wave profile $z_w(x)$ is superimposed on a curved earth surface,

$$z_s(x) = z_w(x) - \frac{x^2}{2a_e} \qquad (1)$$

where a_e is the radius of the earth.

It is assumed the waveheight does not vary perpendicular to the azimuthal direction. The wave profile is described by a sum of random phase sinusoidal components of sufficient number to give reasonable agreement with tabulated sea state data. A summary of the wave profile model and appropriate sea state indices is provided in Appendix A.

The initial task in the determination of the scatter signal received at the aircraft antenna is the identification of the surface regions making the most significant contributions. These are the regions about the "specular" points, which geometrically reflect rays to the receiver (see Figure 2). Locating specular points involves finding the values of x where

$$\theta(x) = \beta(x) \qquad (2)$$

with

$$\theta(x) = \tan^{-1}\left\{\frac{z_o - z_s(x)}{x - x_o}\right\} \qquad (3)$$

$$\beta(x) = E - 2\,\varepsilon(x) \qquad (4)$$

$$\varepsilon(x) = \tan^{-1}\left\{\frac{dz_s(x)}{dx}\right\} \qquad (5)$$

$$\alpha(x) = E - \varepsilon(x) = \frac{E + \beta(x)}{2}\,. \qquad (6)$$

Here, $\varepsilon(x)$ denotes the incline of the tangent plane to the surface, $\beta(x)$ is the angle of the reflected ray, $\alpha(x)$ is the local grazing angle, the co-ordinates of the antenna are (x_o,z_o) and E is the elevation angle. The specular point search procedure is extremely computationally demanding for aeronautical applications (i.e. the active scattering region extends to the horizon range of the aircraft). An attractive feature of the simulation model is the exclusion of scattering facets where the specular point is blocked by the wave profile.

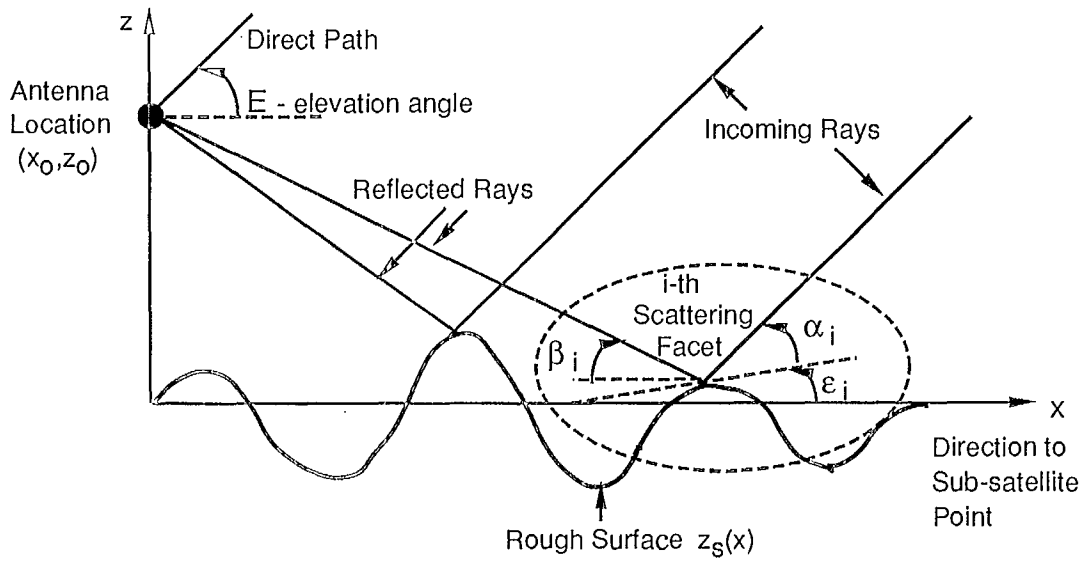


Figure 2 : Propagation model geometry

In the propagation simulation, unmodulated transmissions are considered and the H-pol and V-pol components are treated separately. The scatter signal from each contributing surface facet is accurately determined. This is accomplished by incorporating the variability of the surface reflection coefficient, aircraft antenna gains, as well as the divergence and phase shift resulting from scattering from curved surface facets. That is,

$$S_H(i) = \Gamma_H(\alpha_i)\ e^{-j\phi_i}\ G_H(B_A - \beta_i)\ A_i \qquad (7)$$

$$S_V(i) = \Gamma_V(\alpha_i)\ e^{-j\phi_i}\ G_V(B_A - \beta_i)\ A_i \qquad (8)$$

where Γ_H, Γ_V and G_H, G_V denote the surface reflection coefficients and mobile antenna gains for H-pol and V-pol, B_A is the aircraft banking angle (positive towards the surface), ϕ_i is the scatter path phase shift (relative to the direct path), A_i is the signal attenuation and phase shift factor, and β_i and α_i are the reflection and grazing angles, respectively.

The scatter path phase shift can be expressed as

$$\phi_i = \frac{2\pi\, r_i\,(1 - \cos(E+\beta_i))}{\lambda_c} \qquad (9)$$

where r_i is the distance between the antenna and the scattering facet and λ_c is the carrier wavelength.

An appropriate signal attenuation and phase shift factor for the aeronautical application, derived in Appendix C of [2], is

$$A_i = \begin{cases} \left(1 + \frac{r_i}{f_i}\right)^{-1/2} & \text{for convex facets} \\ j\left(\frac{r_i}{f_i} - 1\right)^{-1/2} & \text{for concave facets} \end{cases} \qquad (10)$$

where

$$f_i = \frac{r_c}{2}\sin(\alpha_i) \qquad (11)$$

$$r_c = \left|\frac{(1 + z_s'(x_i)^2)^{3/2}}{z_s''(x_i)}\right| \qquad (12)$$

Here, f_i is the focal length of the curved facet and r_c is the surface radius of curvature. For aeronautical applications, the magnitude of A_i is equivalent to the standard ray-optics divergence factor.

The overall scatter signals for each polarization are given by the superposition of the individual components from each contributing surface facet. These signals can be subsequently combined to account for circular and elliptical polarizations. To give some appreciation for the complexity of the aeronautical propagation simulation, there are in excess of 10,000 contributing facets for an aircraft altitude of 9.1 km, an elevation angle of 10°, and rough surface conditions.

AN IMPORTANT VALIDATION OF THE SIMULATION SOLUTION

For perfectly conducting, sinusoidal surfaces, the far-field Helmholtz integral solution is provided in Section 4.3 of [1]. The received electric field can be expressed as a scaled version of the field that would have been present in the absence of surface roughness. The scale factor ρ is a function of the angle of incidence of the incoming radiation and the scattering direction under consideration. Restricting attention to the specular direction, the theoretical scale factor can be expressed as

$$\rho_{theory} = \left|J_0\left(\frac{\phi_d}{2}\right)\right| \qquad (13)$$

where ϕ_d is the phase difference between the crest and trough specular paths, and $J_0(\cdot)$ is the Bessel function of the first kind of order zero.

The scale factor obtained with the simulation model (p.61 of [2]) is

$$\rho_{sim} = \sqrt{\frac{2}{\pi\left(\frac{\phi_d}{2}\right)}}\left|\cos\left(\frac{\phi_d}{2} - \frac{\pi}{4}\right)\right| \qquad (14)$$

The excellent agreement between the theoretical and simulation solutions for rough surface conditions $\left(\phi_d > \frac{\pi}{2}\right)$ is evident in Figure 3.

SIMULATION RESULTS

Some L-band (f_c=1.5 GHz) aeronautical simulation results for a low complexity crossed-slot antenna are presented in this section. The antenna pattern considered is presented in Figure 4. For rough surface conditions, it has been confirmed that the multipath process can be accurately modelled as a Rayleigh fading process [2]. Given the large number of contributing surface facets, this is not a surprising result. With the addition of a direct path component, a Rician channel model is appropriate for aeronautical applications. The average interference level and behaviour of the second order statistics of the multipath process are required to characterize the fading aeronautical channel .

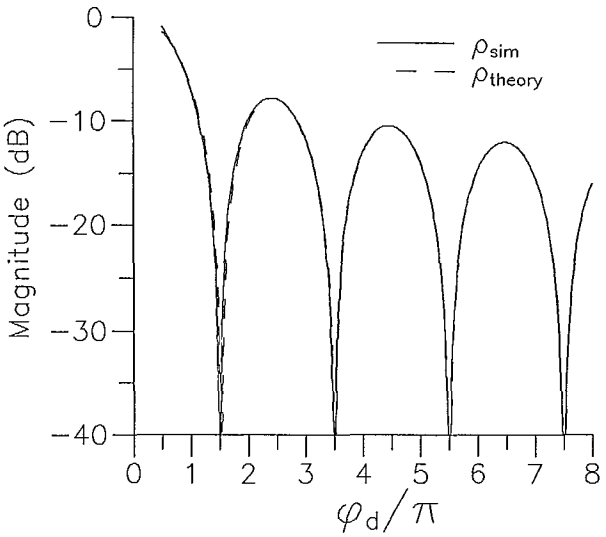


Figure 3 : Simulation and theoretical reflection factors for perfectly conducting sinusoidal surfaces.

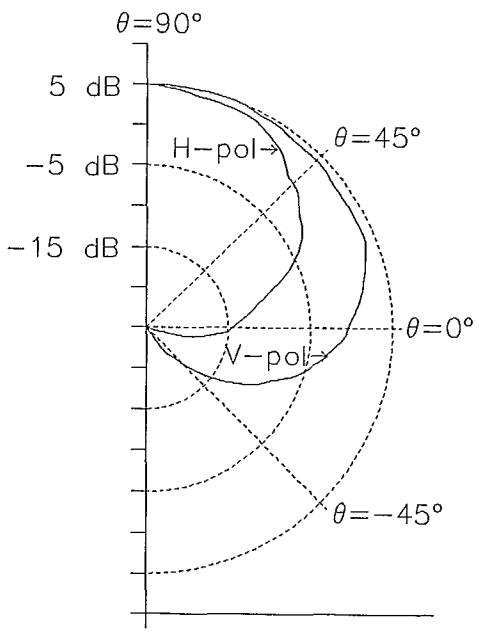


Figure 4 : Crossed−slot antenna pattern

Table 1 summarizes results in terms of signal to interference ratio (S/I) as a function of elevation angle for the crossed-slot and reference omni-antenna patterns. The aircraft altitude was 9.1 km and the surface conditions were rough (sea state index SS4(ii)). The superiority of circular polarization (C-pol) is evident in Table 1. This is a consequence of the polarization sense reversal that a surface imparts on a scattered signal, which is discriminated against by the receiving antenna.

For a circularly polarized crossed-slot antenna, it is evident that S/I is not very sensitive to elevation angle, and the values are excellent, exceeding 20dB. At low elevation angles, the crossed-slot antenna discriminates against H-pol multipath (see Figure 4), while the surface discriminates against V-pol multipath (see Figure 1). With S/I values this high, ocean scatter will

Elevation Angle	S/I [dB] For		
	H-pol	V-pol	C-pol
31°	24.6 (1.2)	16.9 (3.9)	21.6 (17.9)
20°	14.9 (1.0)	15.7 (5.9)	20.4 (13.6)
15°	11.9 (1.6)	16.5 (8.0)	21.3 (12.4)
10°	10.4 (2.7)	18.8 (11.3)	22.9 (11.7)
5°	11.2 (5.2)	21.0 (16.2)	22.9 (12.5)

(Values in parenthesis for an omni-antenna)

Table 1: S/I as a function of elevation angle for an aircraft altitude of 9.1 km and SS4(ii) surface conditions.

not have a severe impact on data communications integrity. Even for an omni-antenna, the S/I values are higher than 10 dB for circular polarization.

The simulation results presented in Table 2 demonstrate that S/I is not very sensitive to sea state and antenna height. The largest difference between the SS3 and SS5(i) results is only 0.7 dB, while the biggest discrepancy between the results for antenna heights of 9.1 km and 5 km is only 0.4 dB. A slight trend of increasing scatter power as the surface gets smoother (decreasing sea state) and as the aircraft altitude is lowered is noticed.

The power spectrum of the multipath process, commonly referred to as the Doppler spectrum, conveys information about the second order statistics of the fading process. An example Doppler spectrum for an elevation angle of 10° is presented in Figure 5. The most striking feature is the asymmetry. This can be theoretically justified (Appendix E of [2]), although the explanation is too lengthy to be presented here. The following relationship between Doppler spread D_{10} (Hz), horizontal airspeed v_x (m/s), and elevation angle E, was derived from simulation results presented in Section 3.2 of [2]:

$$D_{10} = \begin{cases} 0.23\, v_x, & E=10° \\ 0.37\, v_x, & E=15° \\ 1.15\, v_x, & E=31° \end{cases} \qquad (15)$$

Sea State	Aircraft Altitude	S/I [dB] for Polarizations H- pol	V- pol	C-pol
SS3	9.1 km	9.9 (2.4)	18.4 (11.3)	22.7 (11.3)
SS4(ii)	9.1 km	10.4 (2.7)	18.8 (11.3)	22.9 (11.7)
SS5(i)	9.1 km	10.6 (2.8)	18.9 (11.2)	23.0 (11.9)
SS4(ii)	5 km	10.0 (2.5)	18.7 (11.1)	22.5 (11.4)

(Values in parenthesis for an omni-antenna pattern)

Table 2: Effect of sea state and aircraft altitude on S/I for E=10°

Here, D_{10} is the width of the spectral region where the relative Doppler response is above -10 dB. For an aircraft speed of 600 mph, the Doppler spreads are around 60 Hz at low elevation angles (E=10°) and around 300 Hz for intermediate elevation angles (E=31°).

Figure 6 compares simulated H-pol and V-pol scatter power levels for an omni-reference

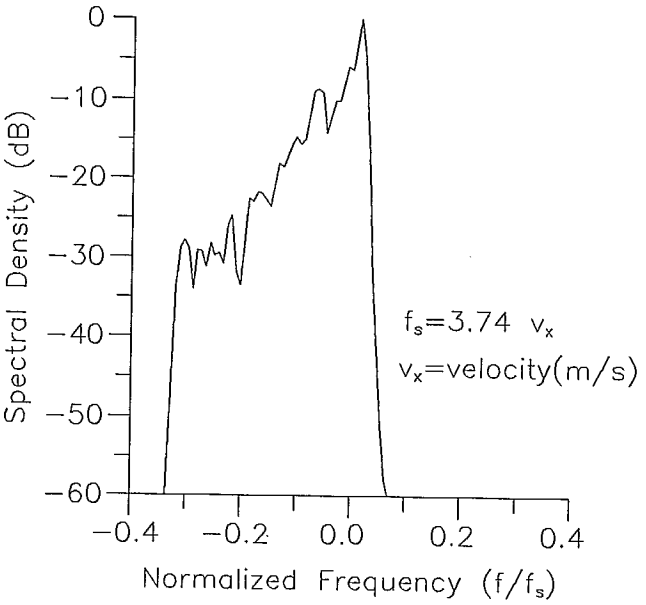


Figure 5 : Doppler spectrum for E=10° and a C-pol omni-antenna.

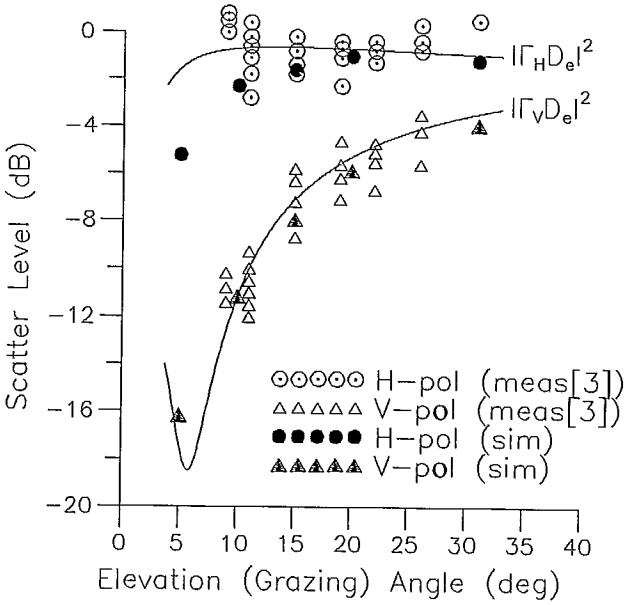


Figure 6 : Measured and simulated scatter power levels for an omni-antenna.

antenna against measurement results from [3] and a simple approximation given by the magnitude squared value of the product of surface reflection coefficient and the divergence factor of the earth [3]. The agreement for elevation angles above 10° is quite good.

APPENDIX A
SURFACE MODEL

The wave profile used in this study is described by a sum of random phase sinusoidal components, with a white bandpass spatial spectrum,

$$z_w(x) = \frac{\sqrt{2}\sigma_z}{\sqrt{N_T}} \sum_{k=-N}^{N} \cos(2\pi u_k x + \theta_k)$$

where $N_T=2N+1$ is the number of sinusoidal components, θ_k is a uniformily distributed random phase, $u_k=u_o(1+\frac{k\gamma}{N_T})$, $\gamma = 1.344$, and $u_o = (1.0726\ \Lambda)^{-1}$. The dependence of the rms waveheight σ_z and average sea wavelength Λ on sea state is provided in Table A.1 for a few representative cases.

For the simulation runs, N_T was chosen to be 21. With the above parameter values, the wave profile generated by this model is approximately Gaussian with a correlation distance of $\Lambda/4.44$. According to [4], this is a realistic model for a 1-dimensional cross-section of the ocean surface.

Sea State Index	RMS Waveheight (Meters)	Average Sea Wavelength (Meters)
3	0.35	22.3
4(i)	0.525	30.75
4(ii)	0.61	34.3
5(i)	0.76	41.9
5(ii)	0.92	48.8

Table A.1: Representative sea state table

A nice feature of this surface model is that derivatives, which are required in the geometrical search for specular points and radius of curvature computations, are easily determined.

REFERENCES

[1] P. Beckmann and A. Spizzichino, "The Scattering of Electromagnetic Waves From Rough Surfaces", The MacMillan Company, New York, 1963.

[2] K.W. Moreland, "Ocean Scatter Propagation Models For Aeronautical and Maritime Satellite Communication Applications", Master's Thesis, Faculty of Engineering, Dept. of Systems and Computer Engineering, Carleton University, February 1987.

[3] R.W. Sutton et.al., "Satellite-Aircraft Multipath and Ranging Experiments at L-band", IEEE Trans.Comm, May 1973.

[4] P. Horn et.al., "Theoretical Study of Multipath Effects In An Aeronautical Satellite System", Messerschmitt-Bolkow-Blohm, Contract Report 1064/70CG (RFQ/737) for European Space Research and Technology Centre, Noordwijk, Holland, November 1970, p.127-128.

Session 7
Mobile Terminal Equipment

Session Chairman - *Hans Haugli,* Inmarsat, UK
Session Organizer - *Tom Jedrey,* JPL

A Satellite Data Terminal for Land Mobile Use

Colin A. Sutherland
Canadian Astronautics Limited
1050 Morrison Drive
Ottawa, Ontario, K2H 8K7
Phone: 613-820-8280
FAX: 613-820-8314

ABSTRACT

Telesat Mobile Incorporated (TMI) has recently introduced the Mobile Data Service (MDS) into Canada. This paper outlines the system design and some key aspects of the detailed design of the Mobile Earth Terminal (MET) developed by Canadian Astronautics Limited (CAL) for use with the MDS. The technical requirements for the MET are outlined and the equipment architecture is described. The major design considerations for each functional module are then addressed. Environmental conditions unique to the land mobile service are highlighted, along with the measures taken to ensure satisfactory operation and survival of the MET. Finally, the probable direction of future developments is indicated.

REQUIREMENTS

The requirements for the MDS are to provide two-way digital messaging together with position reporting for long-haul vehicles and their associated dispatch centres. A central Hub facility provides communication between the dispatch centres and the satellite, and the METs communicate directly with the satellite. Protocols are an extension of Inmarsat's Standard 'C'[1] optimized for land mobile use. Message formats include pre-formatted, coded and free form text, each with two levels of priority, and broadcast messages in the outbound direction. Both scheduled and solicited position reports are provided. The Loran-C 100 KHz navigational system was selected for the positioning system in the initial version of the equipment.

Table 1 shows the primary performance requirements for the MET. In addition, there are stringent phase noise requirements, the need to reject strong adjacent signals, and a complex system of modulation and coding. These must be considered in the context of a low-cost terminal that will be constantly exposed to the rigours of the land mobile environment. Totally automatic operation and a simple user interface are dictated by the intended application.

ARCHITECTURE

The MET comprises three subsystems; Antenna, Transceiver and User Terminal. Figure 2 shows the block diagram of the arrangement adopted. The philosophy was to place in the transceiver all functions not located of necessity at either the antenna or the user terminal. This minimizes both the size and weight of the user terminal, and the replacement cost of the antenna.

Table I. Primary Performance Requirements for the MDS MET

Transmit frequency range:	1626.5 to 1660.5 MHz
Receive frequency range:	1530 to 1559 MHz
Channel spacing:	5 KHz
G/T:	-22 dB/°K min.
EIRP:	+15 dBW min.
Transmit duty cycle:	2.5%
Continuous transmission:	620 ms max.
Transmit/receive switching time:	<150 ms
Initial acquisition time:	5 minutes
Elevation coverage:	15° to 35°

Transceiver

The Transceiver contains the following components: up/down converter (Converter), baseband processor (BBP), Loran-C receiver, power amplifier (PA), and power supply (PS). The functions of each unit are as follows:

Converter. Provides reference frequency generation, frequency synthesis, down-conversion from L-band to baseband, upconversion of data input to BPSK L-band output, and separation of the received Loran-C signal.

BBP. Performs all digital processing, including unique-word (UW) detection, de-interleaving, demodulation, Viterbi decoding, descrambling, and the corresponding inverse operations. Also provides the satellite protocol processing and system control for the entire MET.

Loran-C Rx. Receives and demodulates 100 KHz Loran-C pulses, performs automatic selection of chains, measures time differences, compensates for propagation path characteristics, and computes latitude and longitude.

Power Amplifier. Class-C amplifier providing 35 W minimum output power over the full uplink band of 1626.5 to 1660.5 MHz. Also provides control bias for the transmit/receive (T/R) switch in the antenna subsystem.

Power Supply. Converts nominal 12 V vehicle power into +5 V, +15 V, and +28 V regulated supplies for the rest of the MET.

User Terminal

The user terminal was developed for CAL by Gandalf Technologies Incorporated. It has a full "Qwerty" keyboard with a numeric keypad, cursor control keys and four special function keys, a 40 character by four line liquid crystal display, two LEDs and an annunciator.

Antenna

The antenna uses a single element for both L-band communication and Loran-C reception. The single circuit board provides the following functions:

Low noise amplifier (LNA). Provides over 30 dB of low noise gain over the full receive band of 1530 to 1559 MHz.

Loran-C preamplifier. Provides gain and buffering for 100 KHz Loran-C signals

T/R switch. A pin diode switch to connect the antenna element to either the PA output or the LNA input.

Bias tee. Removes +15 V dc supply from coaxial cable for use in antenna subsystem.

Diplexers. Separate and combine L-band and Loran-C signals.

DESIGN DETAILS

Converter

The converter presented a difficult design challenge. The combination of high frequency, close channel spacing and low phase noise led to the adoption of a three loop configuration for the main synthesizer (LO1) with a fourth loop for the second local oscillator (LO2). The crystal reference oscillator operates at 8 MHz, and an AFC loop performs fine tuning by using the outbound TDM signal from the satellite as a frequency reference. Downconversion is performed in three stages; first to 70 MHz, then to 4 MHz, and finally to a pair of quadrature baseband channels. Two ten bit A/D converters digitize the baseband signals for further processing by the BBP. No upconversion takes place as such. The main synthesizer generates the required L-band frequency directly, and this is then BPSK modulated and fed to the PA.

Baseband Processor

The BBP uses two processors; a TMS320C25 for digital signal processing, and a 16 bit μP for control and satellite protocol functions. A large programmable logic device is used to implement the necessary logic functions, and several serial interfaces are provided. The internal interfaces are for control of the synthesizer, reference oscillator, T/R switch and PA, and data input from the A/D converters and Loran-C receiver. The external interfaces are for the User Terminal and an Auxiliary Port. A test port is also provided.

Loran-C Receiver

The receiver selected is a board-level product intended for land mobile applications. It includes automatically tunable notch filters for interference rejection and provides completely automatic operation.

Power Amplifier

The PA has five silicon bipolar stages; two linear and three grounded base class 'C'. Only the linear stages are keyed since the class 'C' stages draw no current until they are driven. As stated earlier, the PA is located in the transceiver, which results in about 2 dB of loss in the antenna cable and leads to a requirement for a nominal 40 W output from the PA. The output stage uses a pair of devices combined in quadrature since no suitable single device was available at the time of device selection. The greatest challenge in developing this module has been to ensure

unconditional stability over the full temperature range. This has been achieved by a combination of careful layout, extensive decoupling and the use of lossy ferrite beads on bias lines.

Power Supply

The power supply operates at 70 KHz to avoid interference with the Loran-C receiver. It is designed to accommodate the zero to full load transient on the 28 V output imposed by keying the PA. Excellent cross-regulation is critical to avoid frequency chirp on the transmitted burst. Linear regulators are used on two of the outputs to meet this requirement.

User Terminal

The user terminal has a rubber membrane keypad and a "supertwist" LCD display housed in a small, robust injection-moulded enclosure. Internally, a microprocessor with associated RAM and EPROM, a serial I/O interface and a power supply provide the required functionality. A single cable provides both power and a two-way data path from the Transceiver to the User Terminal.

The software uses a simple menu-driven system for selecting and editing messages, which can also be stored for future recall.

Antenna

Figure 1 shows a cross-section of the antenna, which presented one of the most interesting challenges of the whole development program. A gain of +2 dB was required to meet the G/T specification, and an omnidirectional configuration was specified to avoid the difficulties of beam steering. After a number of attempts, a quadfilar helix with a circular ground plane was adopted. Even with the specified gain, the losses in the T/R switch dictated the use of a GaAs FET first stage in the LNA, followed by two bipolar stages. A bandpass filter between the first and second stages protects the LNA from overloading in the presence of an adjacent transmission from another MET. The antenna ground plane is provided by the circuit board, whose active components are mounted on the underside.

The helix functions also as an electrically short monopole for reception of vertically-polarized signals at 100 KHz. The output from the Loran-C preamplifier is combined with the L-band output from the LNA in a diplexer and fed to the transceiver down the receive coaxial cable.

The T/R switch is implemented with a series/shunt PIN diode combination, and control bias is fed from the PA via the transmit coaxial cable.

ENVIRONMENTAL CONSIDERATIONS

The primary application for the MET is in long-haul trucks, and the implications for the design of the equipment must be given serious consideration. Some of the documented[2] physical stresses in these vehicles include 20 g shocks, vibration up to 4 g over 10 Hz to 1 KHz, wide temperature and humidity extremes, temperature shock and cycling, and exposure to oil and chemicals. In addition, the antenna may be exposed to wind loading, solar radiation,

precipitation, salt fog, blowing sand and dust, small flying rocks, rotating brushes, high pressure hoses, and the occasional tree branch. A heavy aluminum extrusion has been selected for the transceiver housing, and this in turn is installed in a shock-mounted tray. The antenna uses rugged polycarbonate mouldings with a clamped O-ring seal. Careful design and selection of materials combined with extensive environmental testing has led to a design which is considered likely to give good service for many years.

Electrical stresses arise from equipment connected to the vehicle supply, and can include transients up to 600 V for 1 ms, and occasionally 150 V for 400 ms. Clearly a power supply designed for a nominal 12 V input needs careful protection to survive such treatment. A combination of fusing, filtering, and transient absorption devices has been adopted to ensure uninterrupted operation during the shorter transients and survival of the longer variety.

FUTURE PRODUCT EVOLUTION

Two trends are clearly apparent in considering the future of the MET as an evolving product:

Cost Reduction

A constant downward pressure on prices exists. This will be met by measures such as a higher level of integration, and the arrival of new devices for frequency synthesis, power amplification, and digital processing. Increased capital investment will also be required to optimize tooling, streamline the production flow, and improve test facilities.

Inter-operability

Future generations of mobile terminal will be expected to operate not only with the Canadian MDS system, but also with AMSC, Inmarsat's Standard 'C', and MSAT.

REFERENCES

1. **INMARSAT**, Standard-C, System Definition Manual, Release 1.3, July 1989.

2. **SAE** 1988. Joint SAE/TMC Recommended Environmental Practices for Electronic Equipment Design (Heavy-Duty Trucks) SAE J1455. Society of Automotive Engineers, Inc.

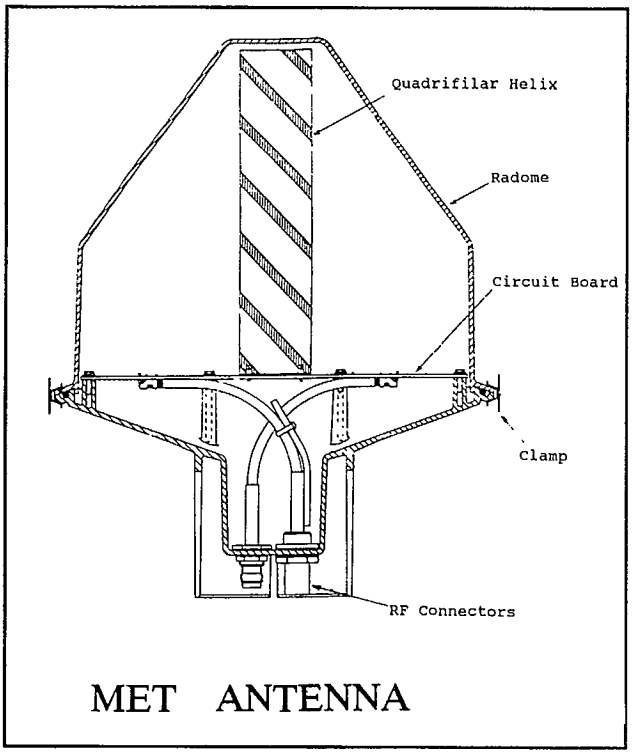


Figure 1. Cross-sectional view of MET Antenna

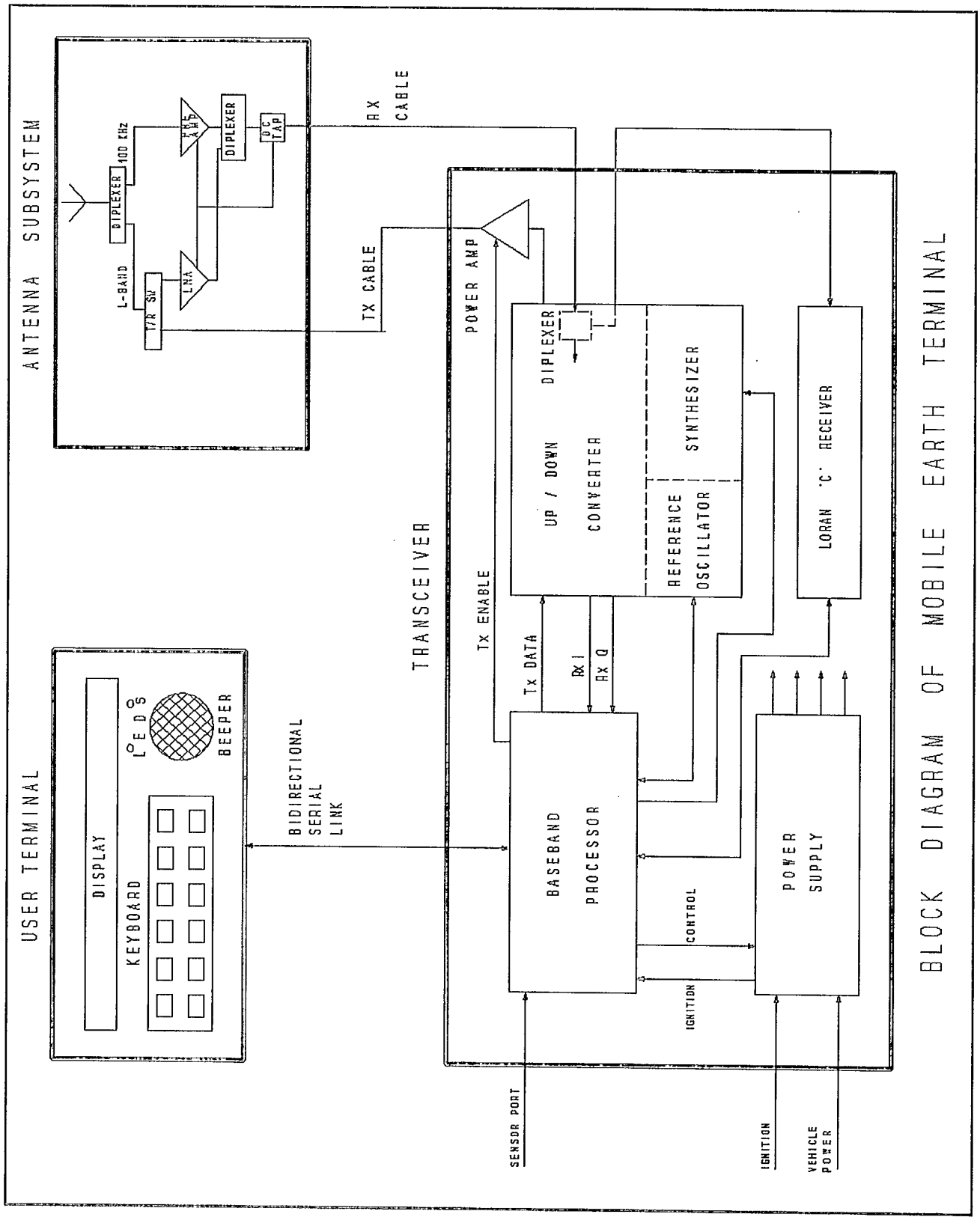


Figure 2. Block Diagram of the MDS MET

A Spread-Spectrum Modem Using Constant Envelope BPSK for a Mobile Satellite Communications Terminal

N. Iizuka, A. Yamashita, and S. Takenaka

FUJITSU LABORATORIES LTD. KAWASAKI
Radio and Satellite Communication Systems Lab.
1015 Kamikodanaka Nakaharaku Kawasaki 211 Japan
Phone: 044-777-1111, Fax: 044-754-2646

E. Morikawa, and T. Ikegami

Kashima Space Research Center
Communications Research Laboratory
Ministry of Posts and Telecommunications
893-1 Hirai Kashima Ibaraki 314 Japan
Phone: 0299-82-1211, Fax: 0299-83-5728

ABSTRACT

This paper describes a 5-kb/s spread-spectrum modem with a 1.275-MHz chip rate for mobile satellite communications. We used a Viterbi decoder with a coding gain of 7.8 dB at a BER of 10^{-5} to decrease the required received power. This reduces the cost of communication services. The spread spectrum technique makes the modem immune to terrestrial radio signals and keeps it from causing interference in terrestrial radio systems.

A class-C power amplifier reduces the modem's power consumption. To avoid nonlinear distortion caused by the amplifier, The envelope of the input signal is kept constant by adding quadrature channel signal to the BPSK signal. To simulate the worst case, we measured the modem's output spectrum using a limiting amplifier instead of the class-C amplifier, and found that 99% of the spectral power was confined to the specified 2.55 MHz bandwidth.

1. INTRODUCTION

Since the lower power spectral density of the spread-spectrum technique reduces interference between satellite and terrestrial communications systems, the technique has the advantage of reduced required received power. This reduces communication cost in satellite communication systems. The coding gain in forward error correction techniques such as in the Viterbi decoder also reduces received power. The service cost of a mobile satellite communication system also depends on the power consumption of the mobile terminal. Although the high efficiency of a class-C amplifier effectively reduces power consumption, the output spectrum of the usual binary phase-shift keyed (BPSK) signal is changed significantly due to nonlinear distortion. To avoid such distortion, the envelope of the input signal must be kept constant. Various constant-envelope BPSK (CE-BPSK) modulation techniques have been proposed [1-3]. When used for mobile satellite communications systems, however, these techniques have the following problems: (1) Remaining envelope variation is about 3 dB due to imperfect suppression, (2) the supplementary signal reducing the envelope variation generates a line spectrum which increases interference with terrestrial communication systems, and (3) interference between the BPSK signal and supplementary signal makes demodulation difficult.

To overcome these problems, we propose an improved CE-BPSK modulation scheme which features: (1) digital signal processing for accurate supplementary signal generation, (2) a polarity inversion switch to remove the line spectrum of the supplementary signal, and (3) a spread-spectrum technique to prevent interference between the supplementary and BPSK signals. We confirmed the improved output spectrum by computer simulation. The measured BER performance and output spectrum of the modem showed little degradation, even though a limiting amplifier with a more nonlinear characteristic was substituted for class-C amplifier.

2. CE-BPSK MODULATION

2.1 Generation of Constant Envelope Signal

Although the high-efficiency of the class-C amplifier reduces power consumption in a mobile satellite communications terminal, this amplifier's highly nonlinear output characteristic requires that the input signal have a constant envelope. We studied modulation in

which a constant-envelope BPSK signal is obtained by adding a supplementary quadrature channel signal. The conventional way of obtaining the constant-envelope signal restricts the supplementary signal to the upper half of the phase-state vector diagram[1], causing an offset in the constellation. Adding a supplementary signal generates a line spectrum which can interfere with terrestrial radio systems. The technique we propose makes positive and negative polarities of the supplementary signal equally probable, so adding the supplementary signal produces no offset or line spectrum.

The conventional circuit for generating the supplementary signal consists of analog circuits which make precise squaring and rooting difficult, requiring the following approximation to be used:

$$Q(t) = \pm(A\text{-}I(t)^2)^{1/2} \simeq A\text{-}|I(t)| \quad (1)$$

A: Constant,
I(t): BPSK signal
Q(t): Supplementary signal

The envelope of the modulated signal may deviate as much as 3 dB. Digital signal processing generates the supplementary signal our method uses, easily satisfying the following equation:

$$Q(t) = \pm(A\text{-}I(t)^2)^{1/2} \quad (2)$$

The block and phase-state vector diagrams of our constant-envelope modulator are shown in Figs. 1 and 2, and the block diagram of our circuit for generating the supplementary signal in Fig. 3. When the spread-spectrum signal is demodulated, the signal is remapped by correlation with the PN code. The supplementary signal can thus be removed in the remapping circuit by generating the supplementary signal with a small cross-correlation to the PN code at the modulator. This makes the CE-BPSK demodulator the same as the conventional BPSK demodulator, simplifying demodulation.

2.2 Polarity Switching

The conventional way of obtaining the constant-envelope signal produces an offset in the constellation and the addition of the supplementary signal generates a line spectrum. To remove the offset, we ensured that positive and negative polarities of the supplementary signal were equally probable. This required polarity switching. Switching the polarity when the supplementary signal amplitude is large varies the transient amplitude from positive to negative, or vice versa. Transient variation of the supplementary signal amplitude causes out-of-band radiation. To avoid such variation, the polarity is switched when the supplementary signal amplitude is smaller than threshold Th.

3. SIMULATION

We used computer simulation to evaluate the performance of our proposed technique. To raise the transmitter power efficiency , we must reduce the power of the supplementary signal. If the rolloff factor decreases, the maximum amplitude of the BPSK signal increases. In our CE-BPSK modulator, the quadrature signal supplements the difference between the maximum amplitude and minimum amplitude of the BPSK signal. Thus, if the rolloff factor decreases, the supplementary signal power increases. The 100% rolloff filter we used for spectrum shaping raised the transmitter's power efficiency, as shown in simulation (Fig. 4). The transmitted power increases 1.7 dB when the 100% rolloff filter is used. Since a class-C amplifier reduces power consumption at least 3 dB, CE-BPSK modulation reduces power consumption by at least 1.3 dB. Using computer simulation we calculated the power spectrum of the CE-BPSK signal after the limiting amplifier (Fig. 5). 99% of the spectral power was confined to the specified 2.55 MHz bandwidth using CE-BPSK modulation. 91% of the spectral power was confined to the same bandwidth using BPSK modulation. Out-of-band radiation performance is improved markedly by adding the supplementary signal to the BPSK signal through a nonlinear amplifier.

In our spread-spectrum modulator, one cycle of the PN code sequence is not inverted when the information is "one," and one cycle of the sequence is inverted when the information is "zero." The PN code sequence modulated by the information is transmitted. At the demodulator, when the cross correlation between the received signal and the PN code is positive, "one" is received and, when it is negative, "zero" is received. Interference signal having no cross correlation to the PN code are removed[4]. However, An interference signals similar to the PN code cannot be removed. In the CE-BPSK demodulator, the received supplementary signal becomes the BPSK interference signal. To estimate degradation due to the interference signal, we calculated the cross correlation between supplementary signal and the PN code (Fig. 6). Because the cross correlation is negligible, we

assume that the degradation in BER performance due to interference is negligible.

4. MODEM

Table 1 lists the system parameters of the CE-BPSK spread-spectrum modem, which operates at a 1.275-MHz chip rate and a 70-MHz IF frequency. The information rate is 4.8 kb/s without forward error correction and 2.4 kb/s with it. Adding unique words increases the data rate of the modem to 5 kb/s. We used direct sequence modulation to spread the spectrum. Because the PN code length is 255, the process gain is 24 dB. The lower spectral power density of the spread-spectrum technique reduces interference with terrestrial radio systems, and the spread-spectrum demodulator reduces interference from terrestrial radio systems[4].

4.1 Modulator

The modulator (Fig. 7) encodes the forward error correction code, inserts unique words into the transmitted data stream, and performs spread spectrum and CE-BPSK modulation. We used digital signal processing to generate the supplementary signal and to shape the BPSK signal spectrum. The CE-BPSK signal has a small envelope variation.

4.2 Demodulator

The demodulator (Fig. 8) remaps the spread spectrum signal, removes the supplementary signal, demodulates the BPSK signal, and decodes the forward error correction code. The BPSK signal can be demodulated after analog-to-digital conversion by digital signal processing. A Costas loop was implemented in software using a digital signal processor.

4.3 Forward Error Correction

The high-coding gain of the Viterbi decoder enables forward error correction convolutional encoding to be used. We selected a code rate of 1/2. The Viterbi decoder uses an 8-level soft decision and a 7-bit constraint length. Its path memory circuit uses path tracing, so the single LSI decoder is compact[5].

5. RESULTS OF EXPERIMENT

5.1 Output Spectrum

To verify the performance of CE-BPSK modulation, we fed the CE-BPSK signal through a limiting nonlinear amplifier instead of a class-C amplifier, (Figs. 9 and 10). Signal spectra at the modulator output are shown Figs. 11 and 12 for comparison. When BPSK modulation was used with the limiting amplifier, the sidelobe signal power increased. When CE-BPSK modulation was used with the limiting amplifier, the percentage of the bandwidth signal power (2.55 MHz) was 99% of the total signal power, the same as the value obtained with the linear amplifier. These results show that out-of-band radiation is markedly improved by adding the supplementary signal to the BPSK-modulated signal when a nonlinear amplifier is used.

5.2 BER Performance

The BER versus carrier-to-noise-density ratio (C/N_0) was measured with IF back-to-back connection for both the linear and limiting amplifiers (Fig. 13). The BERs shown are with and without forward error correction and supplementary signal. When the linear amplifier was used, the difference in BER performance between CE-BPSK and conventional BPSK was 1.8 dB, almost the same as the contribution of the supplementary signal (1.7 dB). Thus, interference between the BPSK and supplementary signals is negligible. The coding gain with forward error correction was 7.8 dB at a BER of 10^{-5}, almost the same as the 8-dB theoretical value. When the limiting amplifier was used, the modem has almost the same BER performance as that using the linear amplifier.

6. CONCLUSION

We developed a spread-spectrum modem using CE-BPSK modulation for mobile satellite communications. The constant envelope signal enables a class-C amplifier to be used to ensure decreased power consumption and increased transmission power. The spread-spectrum modem reduces interference with terrestrial radio systems. We verified the effectiveness of our technique for spread spectrum communication by measuring the modem's output spectrum and BER performance. Our results show this modem to be promising for use in mobile satellite communications.

REFERENCES

1) C. Andren, "PSK Sidebands Reduced by

Premodulation Filtering," Microwave Journal, Vol. 21 January 1978.
2) **H. Yazdani** *et al.*, "Constant Envelope Band-Limited BPSK Signal," IEEE Trans., COM-28, no. 6, June 1980.
3) **E. Dickerson** *et al.*, "Generation of Constant Envelope Signals," IEEE Trans., COM-30, no. 12, December 1982.
4) **R.C. Dixon**, "Spread Spectrum Systems", New York, John Wiley, 1976.
5) **A. Yamashita** *et al.*, "A New Path Memory for Viterbi Decoders," Globecom'87, Tokyo, Japan, 1987, pp. 2076-2079.

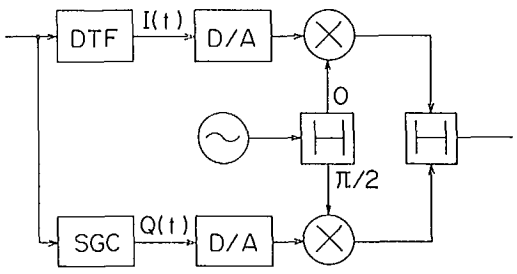


Fig. 1 Modified constant-envelope modulator

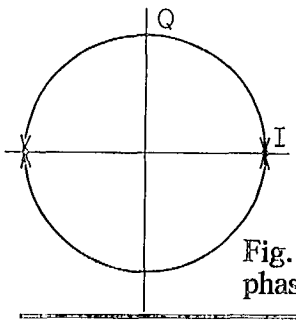


Fig. 2 Modified CE modulator phase-state vector diagram

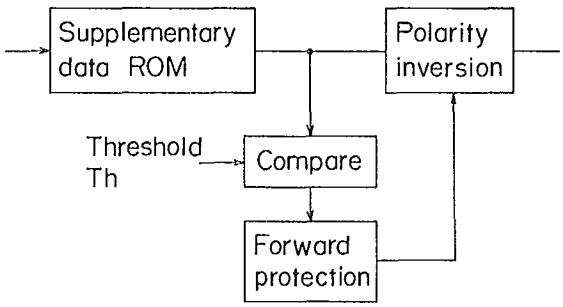


Fig. 3 Supplementary signal generation circuit

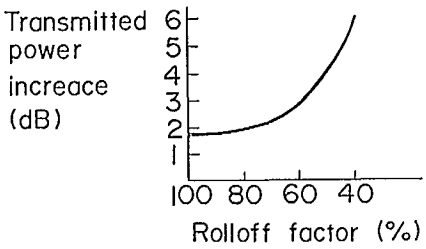


Fig. 4 Relation between increased transmitted power and rolloff factor

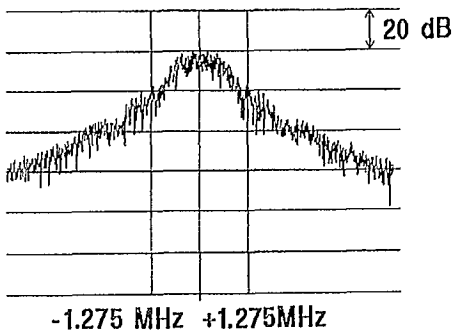


Fig. 5 CE-BPSK signal spectrum (simulated)

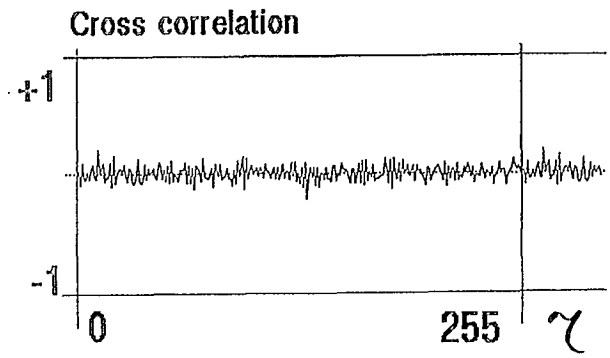
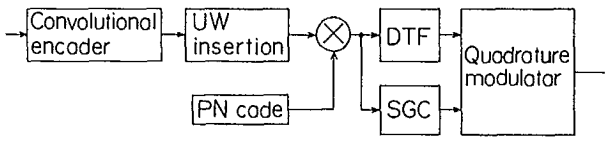


Fig. 6 Cross correlation between supplementary signal and PN code

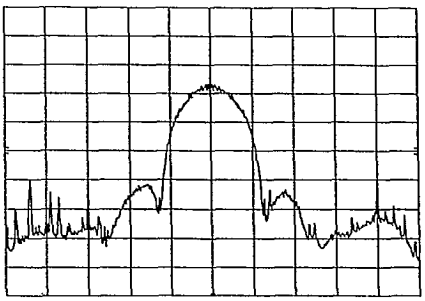
Table 1 Modem parameters

Modulation:	Spread-spectrum with CE-BPSK
Initial synchronization:	Digital matched filter
Tracking:	Delay locked loop
IF frequency:	70 MHz
Data rate:	4.8 kb/s without FEC
	2.4 kb/s with FEC
Modem bit rate:	5 kb/s
Chip rate:	1.275 Mb/s
Process gain:	24 dB
Forward error correction:	R = 1/2, K = 7 soft-decision Viterbi decoder
Carrier-to-noise density ratio:	40 dB·Hz to 60 dB·Hz



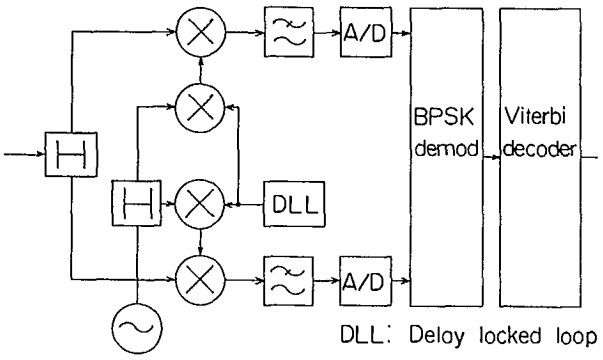
UW: Unique words
DTF: Digital transversal filter
SGC: Supplementory signal generation circuit

Fig. 7 Modulator



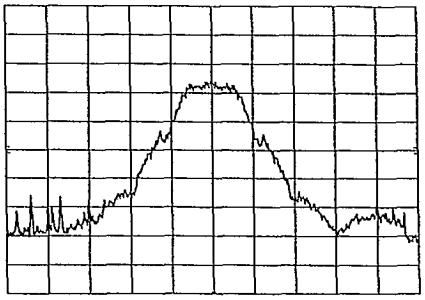
10 dB/div
1 MHz/div
Resolution
BW: 30 kHz

Fig. 11 BPSK signal spectrum at modulator output



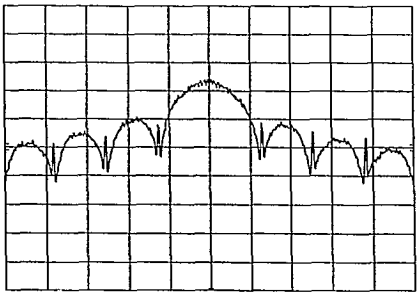
DLL: Delay locked loop

Fig. 8 Demodulator



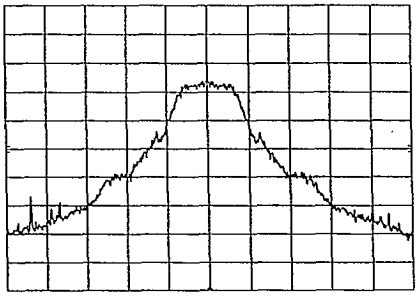
10 dB/div
1 MHz/div
Resolution
BW: 30 kHz

Fig. 12 CE-BPSK signal spectrum at modulator output



10 dB/div
1 MHz/div
Resolution
BW: 30 kHz

Fig. 9 BPSK signal spectrum with limiting amplifier



10 dB/div
1 MHz/div
Resolution
BW: 30 kHz

Fig. 10 CE-BPSK signal spectrum with limiting amplifier

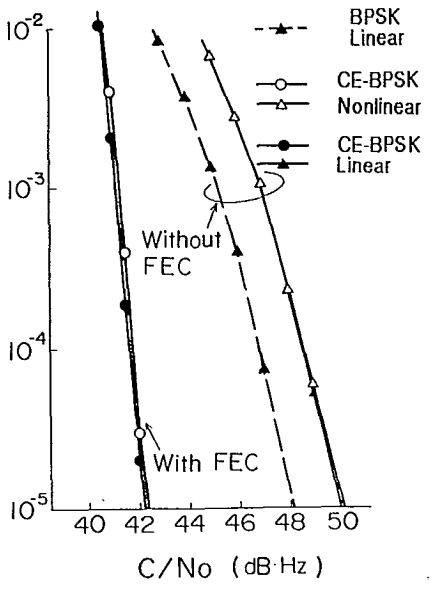


Fig. 13 BER performance

Description and Performance of a Digital Mobile Satellite Terminal

N. Lay, T. Jedrey, J. Parkyn, D. Divsalar
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109 USA
FAX: (818) 354-6825

ABSTRACT

A major goal of the Mobile Satellite Experiment (MSAT-X) program at JPL is the development of an advanced digital terminal for use in land-mobile satellite communications. The terminal has been developed to minimize the risk of applying advanced technologies to future commercial mobile satellite systems (MSS). Testing with existing L-band satellites was performed in fixed, land-mobile and aeronautical-mobile environments. JPL's development and tests of its mobile terminal have demonstrated the viability of narrowband digital voice communications in a land-mobile environment through geostationary satellites. This paper provides a consolidated description of the terminal architecture and the performance of its individual elements.

INTRODUCTION

A key objective of the MSAT-X program has been the development and demonstration of critical technologies required to enable a commercial land-mobile satellite service (LMSS). These technologies include medium gain tracking antennas, power and bandwidth efficient modulation techniques, and high quality low bit rate voice codecs. The demonstration of these technologies required the development of an integrated terminal to evaluate their individual and combined performance. Several different platforms were employed to field the terminal ranging from a demonstration automobile utilizing a conformal phased array antenna to fully instrumented terminals located within an experimental van and within an experimental aircraft. Figure 1 indicates one configuration of the terminal elements within the trunk of the demonstration automobile.

TERMINAL ARCHITECTURE

The mobile terminal is composed of five different subsystems: the speech codec, the terminal processor, the modem, the transceiver, and the antenna subsystem. A functional block diagram of the terminal is shown in Figure 2. The speech codec provides good quality speech at 4800 bits per second [1,2]. The terminal processor provides a terminal user interface and implements all networking and control functions [3]. The modem implements a power and bandwidth efficient modulation scheme designed for the demanding LMSS channel [4,5]. The transceiver converts the modem output to the L-band transmit frequency, and converts the received L-band signal to the modem input IF, as well as performs pilot tracking [6]. The antenna subsystem tracks the satellite and transmits and receives the transceiver output [7].

Speech Codec

The codecs developed for MSAT-X by UC Santa Barbara (UCSB) and the Georgia Institute of Technology (GIT) were required to be robust in a burst error channel at an average bit error rate (BER) of 10^{-3}. Integral to any speech coding scheme that is to operate in this environment is the ability to detect error events and take actions to mitigate the effects of these errors as well as the ability to reacquire fairly rapidly after a long error event such as shadowing. Critical in the implementation of this scheme is that it does not require a large overhead in terms of the 4800 bps data rate.

The actual voice algorithms employed in the UCSB/GIT codecs have been well described [1,2] and this general class of vector quantizing codecs

has seen a great deal of recent algorithmic development. In short, the compression algorithms operate by initially removing redundancy in the input speech corresponding to vocal tract and pitch information through the use of short and long term predictors. The residual error sequences are vector quantized through the use of codebooks in the case of Vector Adaptive Predictive Coding and Pulse Vector Excitation Coding (UCSB) or are regenerated as delayed versions of the excitation sequences used at the receiver for Self Excited Vocoders (GIT). Typically error sequences are chosen by minimizing a perceptually weighted mean squared error for the entire sequence.

Terminal Processor

All elements of the terminal are tied together under the control of a central device, the terminal processor (TP). The TP implements a variety of functions in response to user input via a keyboard. Transmissions in the form of continuous or packetized data tests, voice conversations or text messages are initiated by the TP through a channel connect protocol consisting of an initial connect packet query and an acknowledgement from the far end terminal. Transmissions are similarly terminated through the use of this protocol. In addition to protocol implementation, the TP selects data transmit and pilot and data receive frequencies by directly controlling the transceiver synthesizers. Pilot acquisition by the antenna subsystem is initially performed manually under TP control. After initial acquisition, pilot reacquisition during loss of signal (from long deep fades) may be performed either manually or automatically with the TP reading pilot signal strength from the transceiver.

Modem

The key design goals for the MSAT-X modem required the transmission of 4.8 kbps data over a Rician K=10 channel at an average E_b/N_0 of 11 dB while achieving a 10^{-3} BER. In addition, the modem was required to be robust in the presence of shadowing and time varying Doppler [4].

The modem accepts as input a serial 4.8 kbps data stream. This data is encoded with a 16 state rate 2/3 trellis code which achieves a coding gain of 1 dB at a BER of 10^{-3} relative to uncoded DQPSK in AWGN. This gain increases to 2 dB when considering a Rician LMSS channel. Three bit symbols from the trellis encoder are mapped to one of eight phases and then block interleaved (16×8) prior to differential encoding. Baseband inphase and quadrature signals are pulse shaped prior to analog output to the transmitter. The modulator employs a square root raised cosine pulse shape with 100% excess bandwidth to achieve spectral efficiency.

At the demodulator, the signal is quadrature demodulated by sampling at four times the IF and decimating by two the inphase and quadrature signals. The signals are time aligned by averaging one stream and lowpass filtered to the signal bandwidth plus the maximum expected Doppler (or frequency offset) of 200 Hz. Both signals are interpolated for greater timing resolution and then subsampled. The symbol timing recovery mechanism incorporates a random walk filter to smooth jitter in the recovery process and to provide an extremely robust estimate of symbol timing during loss of signal due to shadowing. After subsampling the two intersymbol interference free points per symbol period of the pulse shape, an instantaneous Doppler estimate is formed by differential detection. These estimates are smoothed by a lowpass filter whose bandwidth is on the order of the maximum expected Doppler rate. This smoothed estimate is utilized in the matched filter implementation and residual Doppler correction after differential detection. Sixteen bit soft decision outputs are passed to the de-interleaver and Viterbi decoder. A 4.8 kbps serial data stream is then output to the TP.

Transceiver

The MSAT-X transceiver provides a flexible RF package to interface the antenna and modem in a mobile environment. The transceiver provides pointing information for the antenna controller, frequency translation for the modem and test points for propagation data gathering. Input data lines are provided for frequency selection by the TP and output lines provide a sampled version of noncoherent received pilot power.

The front end provides transmitter to receiver isolation while sharing a common antenna, RF filtering and low noise receive amplification. The transmitter and receiver are isolated through the use of an RF diplexer. This unit accommodates the broad frequency range required for the vari-

ous field experiments as well as provides low insertion loss and high isolation. The front end exhibits an overall measured noise figure of 1.8 dB. The transmitter accepts inphase and quadrature baseband output from the modem and upconverts these signals to L-band in a controlled manner to preserve spectral efficiency.

In the receiver, the RF input from the front end is mixed down to an IF of 28 MHz and bandlimited to 4 MHz. Pilot and data channels can be selected independently anywhere within this band to a resolution of 5 kHz. Pilot and data signals are downconverted through separate IF chains where final bandlimiting is performed on each signal. The data signal is output to the modem at an IF of 28.8 kHz. A narrower bandpass filter rejects noise from the pilot signal prior to its downconversion to a tracking loop. All synthesizers within the receiver can then be locked to the received pilot. This provides some measure of Doppler tracking for the data signal and allows the derivation of amplitude and phase information for use by the antenna pointing circuitry.

Antennas

Antenna development for MSAT-X has concentrated on medium gain (8-12 dB) steerable antennas for power efficiency and to combat multipath fading and intersatellite interference. The antenna development has followed a dual design approach consisting of two phased array development efforts under contract to JPL and an in house mechanically steered tilted array design [7]. Average antenna gains for these units range from 10 to 13 dB. Elevation beamwidths range from 28^{o} to 55^{o}, while azimuth beamwidths range from 38^{o} to 55^{o}. Three of these antennas are shown in Figure 3 -- the Teledyne Ryan Electronics (TRE) phased array and two versions of the JPL mechanically steered tilted array antenna, an initial prototype and a mechanically re-engineered lower profile version. Both of the JPL antennas as well as the Teledyne array were employed in land-mobile satellite field tests along the eastern Australian coast during July, 1989.

Both antennas use inertial rate sensors to maintain point during periods of signal loss due to shadowing or fading when the closed loop tracking methods are inoperative. Closed loop pointing information is derived from a coherently demodulated pilot signal provided by the transceiver. The closed loop tracking technique utilized in the phased array is based on a sequential lobing technique, while the mechanical antenna employs a single line pseudo-monopulse technique.

In addition to these antennas, omnidirectional drooping dipoles, and more recently a planar Yagi array [8] have been developed for experimental purposes.

TERMINAL PERFORMANCE

Subsystem specifications, and subsystem and overall mobile terminal performance gathered from laboratory and field tests are presented in this section.

Speech Codec

The performance of the speech codecs developed for MSAT-X can be characterized in two ways: (1) the actual quality of the compression algorithms in accurately reproducing the input audio signal, and (2) the performance of the codec in an error environment typical of the land-mobile satellite channel. The first type of performance evaluation has been reported in a variety of sources [2,9]. These performance evaluations have been performed primarily for the UCSB VAPC algorithm. To summarize these results, the Diagnostic Rhyme Test and Diagnostic Acceptability Measure scores for this codec in the quiet background environment are 91.9 and 65.5, respectively.

The second type of performance characterization has been performed by evaluating the use of the codec (the UCSB codec) in the actual land-mobile satellite environment. In this channel, burst errors are expected as a natural part of the channel characteristics and techniques need to be adopted to partially compensate for these errors. These include frame synchronization and error detection/correction and mitigation strategies [1,2]. The net effect of the above strategies was an effective muting of the received digital speech in a high error rate environment. This threshold was achieved very quickly due predominantly to the steepness of the modem performance curve. As a result, subjectively unpleasant audio artifacts were kept to a minimum. The approximate threshold that corresponded to this muting response occurred at an approximate BER of 5×10^{-3} and above. Below this threshold, the speech quality

was acceptable, and error induced audio artifacts were negligible.

Modem

Modem performance has been previously reported for both AWGN and Rician K=5,10 dB channels [4,10]. In this section, new results are presented for the BER performance in the presence of static and time varying frequency offsets. In addition, BER degradation results due to phase noise in the received signal are presented.

In Figure 4, modem performance curves are shown for a variety of frequency offsets ranging from 0 Hz up to 400 Hz. The 0 Hz curve corresponds to the laboratory measured AWGN performance. Note that in the land-mobile arena, the maximum offset due to Doppler can be upper bounded by 200 Hz, and that performance degradation above this offset can be primarily attributed to loss of signal energy outside the passband of the initial demodulator lowpass filters. The tracking performance is also shown for a linearly varying frequency offset ranging from 0 Hz to 200 Hz at a rate of 100 Hz/s -- a rate substantially higher than any expected land-mobile Doppler rates. There is less than 0.3 dB of degradation for static and tracking performance up to a 200 Hz offset. At offsets of 300 and 400 Hz the degradation increases to 0.75 and 1.5 dB respectively. This resultant loss could be decreased by increasing the front end lowpass filtering bandwidth to completely allow the received signal to pass undistorted.

Another area impacting modem performance is degradation induced by oscillator phase noise. The sources of this phase noise are distributed throughout the communication link, however, the satellite transponder and the terminal receiver (due to pilot tracking at low levels) have been identified as the two primary sources of phase noise. A software simulation of a transponder phase noise specification was developed, and the effects of the phase noise on the modem simulated [11]. The phase noise model was based on data for the one sided power spectral density of the INMARSAT II satellite, provided by Hughes Aircraft. Simulation results indicate that a model with roughly 10 dB more phase noise power than Inmarsat II was shown to introduce less than 0.3 dB of degradation in both AWGN and K=10 fading, indicating robustness in the presence of satellite transponder phase noise. Tests involving the terminal receiver are discussed in the following section.

Transceiver

The mobile terminal transceiver fulfills two important roles. It provides a phase locked pilot tracking reference for antenna pointing, and as a local frequency reference. Typical receive signal levels range from -130 to -120 dBm, while transmit levels can range up to +46 dBm. The performance specifications which enable these levels of operation are summarized in Table 1.

Table 1 Transceiver Specifications

Parameter	*Value*
Noise Figure	1.8 dB
Receiver Sensitivity	-130 dBm (42 dB-Hz)
Tx/Rx Isolation	>100 dB
Receive Frequency Range	1539-1556 MHz
Transmit Frequency Range	1639-1656 MHz
Step Size	5 kHz
Data Passband Delay Variation	2 μs
Data Passband Ripple	0.5 dB
Data IF Architecture	Triple Conversion
Pilot Loop Bandwidth	300 Hz
Pilot IF Architecture	Double Conversion
Long Term Stability	10^{-6}
Short Term Stability	3×10^{-8}

For transmit signal amplification a variety of units were used. In the JPL Colorado tower experiments, the transmitter's integral driver amplifier's low power output was sufficient. In the aeronautical experiments, a TWT amplifier with 200 Watt saturated capability was used in it's linear region to provide output levels ranging from 5 to 10 Watts. In the AUSSAT land-mobile experiment a solid state linear amplifier that could produce 40 Watts at the 1 dB gain compression point was loaned to JPL by AUSSAT. An equivalent solid state amplifier is currently in development at JPL for use in land-mobile experiments in North America.

As indicated in the previous section, another potential source of phase noise is the remodulation of phase noise present in the pilot tracking

loop on to pilot locked local oscillators in the mobile receiver. Although the receiver local oscillators are not required to be phase locked due to the Doppler/frequency correction scheme in the modem, this investigation does yield useful insights into the required quality of free running oscillators in any future receiver design, and into the tradeoff between opening up the modem filter bandwidths versus employing a pilot tracking scheme. Performance tests have included parametrizing BER vs. pilot C/N_0, indicating degradation ranging from 0.25-1.1 dB correpsonding to C/N_0 values from 51.5-42.2 dB-Hz [12]. Current investigations include parametrizing phase noise power spectral densities versus C/N_0 and generating a corresponding model for use in simulations.

Antennas

Table 2 details the tracking and acquisition performance and G/T values for each of the MSAT-X antennas [14]. The tabulated values correspond to the right hand circularly polarized version of the Ball antenna and the left hand polarized versions of the Teledyne and low profile JPL mechanical antenna. The JPL antenna includes the highest level of RF circuitry integration of all the developed mechanical antennas to reduce insertion loss and the resultant antenna temperature. Both the left hand Teledyne and JPL antennas were employed in field testing in Australia.

Table 2 Antenna Performance

Parameter	*Mechanical*	*TRE*	*Ball*
Peak Gain	11.3 dB	10.8 dB	11.5 dB
G/T	-14.0 dB	-15.4 dB	NA
Elevation Beamwidth	55^o	28^o	32^o
Azimuth Beamwidth	38^o	55^o	55^o
Acquisition Time	<15 sec.	<3 sec.	<3 sec.
Tracking Error	$\pm 2^o$	$\pm 5^o$	$\pm 7^o$

Overall Terminal Performance

A set of performance curves detailing the performance of the terminal in an aeronautical environment [14] is presented in Figure 5. The experimental results are compared to laboratory AWGN data. The curve GROUND.RHS.BER shows the terminal return link performance for the aircraft mounted terminal while stationary on the ground. Less than 0.5 dB of degradation is observed. The remaining two curves represent the flight performance on different dates. The difference in performance is attributed to light or no turbulence for the curve AIR.RHS.329, and very heavy turbulence for the curve AIR.RHS.331. In all cases, the experiment data was observed to be within 1.0 dB of laboratory performance.

A representative curve of the overall terminal BER performance data for a stationary data test taken during the JPL/AUSSAT land-mobile experiment [15], is presented in Figure 6. The mobile terminal was located near Taree, north of Sydney on the eastern Australian coast, with a clear line of sight link to the Japanese ETS-V satellite, and the fixed terminal was located in downtown Sydney. The antenna utilized for this test was the low profile JPL mechanical antenna. Also plotted is the laboratory performance of the modem for the AWGN channel. Field results are within 0.5-0.7 dB of laboratory performance. The return link (mobile-to-fixed) performance for the mobile travelling around a twenty minute test loop is presented in Figure 6. The test results are from both the JPL and TRE antennas. These tests were performed just south of Brisbane, Australia. During the tests, the van velocity ranged from 0 mph up to 40 mph. The terrain can be characterized as predominantly clear with sporadic shadowing. For all data points, the performance relative to laboratory AWGN results is within 1.2 dB.

CONCLUSIONS

The MSAT-X Mobile Terminal represents the culmination of many technology development efforts necessary for a narrowband digital LMSS. The terminal integrates the approaches in each of the technology areas into a communications system suitable for demonstrating and further evaluating system operation in the digital mobile satellite arena. To date, the terminal has demonstrated robustness in the mobile environment and performance approximately within 1 dB of theory/simulation.

ACKNOWLEDGEMENT

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and

Space Administration.

REFERENCES

1. Gersho, A. 1988. High Quality Speech Compression at 4800 bps Using Vector Quantization. *Phase Three Final Report JPL Contract 957063;* Department of Electrical & Computer Engineering, University of California Santa Barbara

2. McGrath, S., Barnwell, T. 1988. Development, Design, Fabrication, and Evaluation of a Breadboard Speech Compression System at 4800 bps. *Phase Four Final Report JPL Contract 957074;* Georgia Institute of Technology

3. Cheetham, C. 1988. The Terminal Processor: The Heart of the Mobile Terminal. *MSAT-X Quarterly.* No. 12.

4. Jedrey, T., Lay, N., Rafferty, W. 1988. An All Digital 8-DPSK TCM Modem for Land Mobile Satellite Communications. *Proceedings ICASSP-88.*

5. Simon, M., Divsalar, D. 1988. The Performance of Trellis Coded Multilevel DPSK on a Fading Mobile Satellite Channel. *IEEE Transactions on Vehicular Technology.* Vol. 37, No.2:78-91

6. Parkyn, J. 1988. L-Band Receiver for MSAT-X. *MSAT-X Quarterly.* No. 13

7. Woo, K. 1988. Vehicle Antenna Development for Mobile Satellite Applications. *Proceedings IEE Fourth International Conference on Satellite Systems for Mobile Communications and Navigation.*

8. Huang, J. 1989. Microstrip Yagi Array for Low-Profile Mechanically Steered Antenna Application. *MSAT-X Quarterly.* No. 20.

9. Kemp, D., Sueda, R., Tremain, T. 1989. An Evaluation of 4800-bps Voice Coders. *Proceedings ICASSP-89.*

10. Jedrey, T., Lay, N., Rafferty, W. 1988. An 8-DPSK TCM Modem for MSAT-X. *Proceedings Mobile Satellite Conference 1988.*

11. Divsalar, D. 1989. *The Effects of Phase Noise of the Satellite Transponder on the Performance of MSAT-X Modem.* Interoffice Memorandum 331-89.2-009 (Pasadena, California; Jet Propulsion Laboratory Internal Document).

12. Dessouky, K., Lay, N., Parkyn, J. *The Bench 1 (B1) Tests.* Interoffice Memorandum 3392-88-72 (Pasadena, California; Jet Propulsion Laboratory Internal Document).

13. Estabrook, P., Oliver, G., Jamnejad, V., Huang, J., Thomas, B. 1990. The Impact of External Noise Sources on Antenna Noise Temperature and the Characterization of the Antennas Used in the Joint MSAT-X/AUSSAT Experiment. *MSAT-X Quarterly.* No. 22.

14. Jedrey, T. 1990. The MARECS-B2 Satellite Experiment: Flight Segment Results. *MSAT-X Quarterly.* No. 22.

15. Jedrey, T., Rafferty, W. 1990. The MSAT-X/AUSSAT Land-Mobile Satellite Experiment: An Overview. *MSAT-X Quarterly.* No. 22.

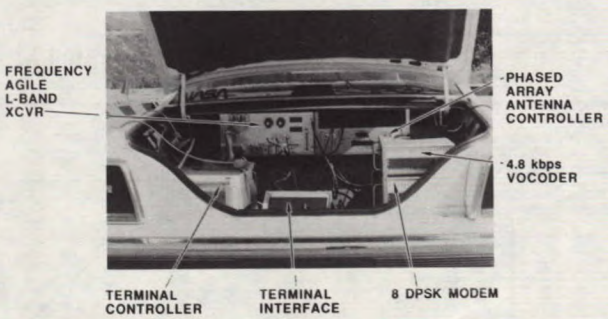


Figure 1 Terminal Configuration in Automobile

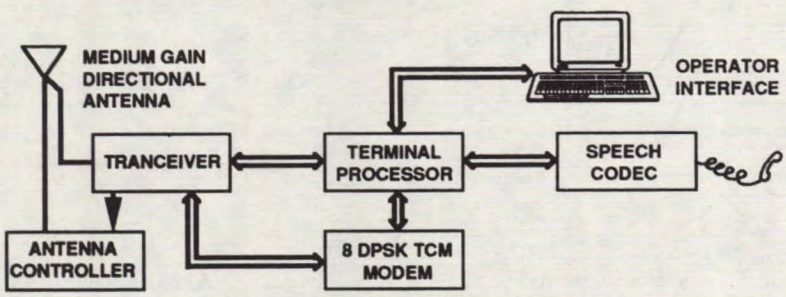


Figure 2 Terminal Block Diagram

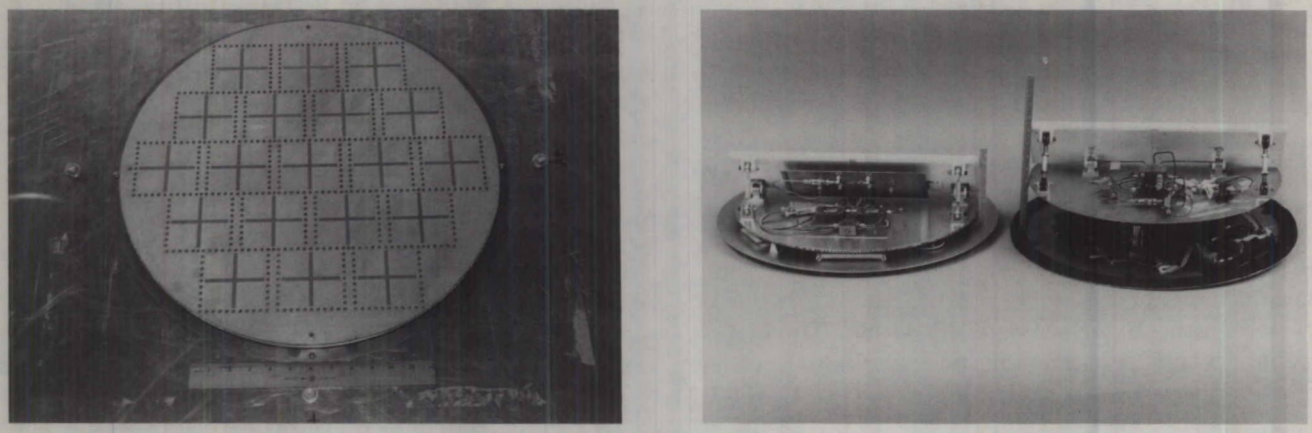


Figure 3 MSAT-X Antennas

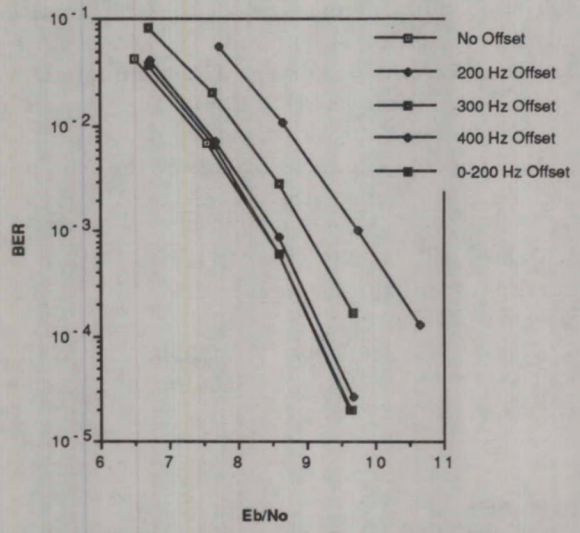


Figure 4 Performance with Frequency Offsets

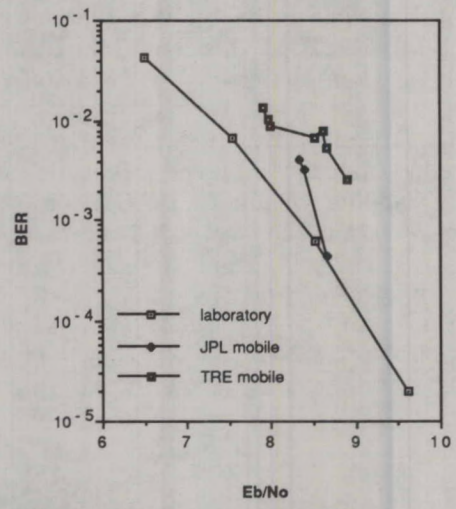


Figure 6 Terminal Mobile Performance

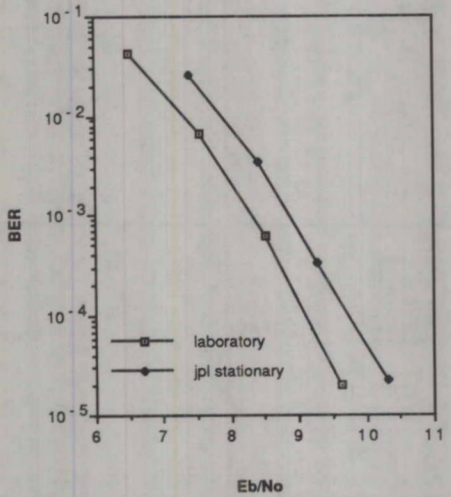


Figure 5 Terminal Stationary Performance

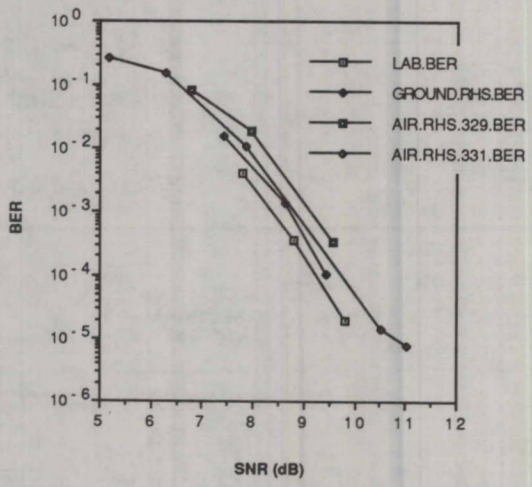


Figure 7 Aeronautical Terminal Performance

L-Band Briefcase Terminal Network Operation

P. Rossiter
D. Reveler
L. Tibbo

SkyWave Electronics Ltd.
300 March Road
Suite 304
Kanata, Ontario
Canada

Phone: (613) 592-0908
Fax: (613) 592-2104

ABSTRACT

During 1989, SkyWave Electronics Ltd. developed a lightweight, battery-powered, L-band Briefcase Satellite Terminal (LBT) which is capable of providing truly portable voice and data communications on a global basis.

The LBT is designed to operate through Inmarsat's Atlantic Region Satellite MARECS B2, and Teleglobe Canada's 18-meter Laurentides Earth Station at Weir, Quebec.

A small operating network, consisting of up to 20 mobile terminals and homing on the Laurentides Earth Station, was set up in the spring of 1990 to provide commercial service to LBT users, both domestic and international.

This paper describes the briefcase terminal and the operation of the network.

BACKGROUND

The motivation for developing the LBT was two-fold.

Our first objective was to develop and evaluate new communications technology being proposed for future use in the North American MSAT system planned for 1993. This technology included new voice "modulation" techniques such as Amplitude Companded Single Sideband (ACSSB) which permits power-efficient voice communication within a 5 kHz RF channel. To support ACSSB operation, it was necessary to devise effective approaches to frequency control, signaling, linear solid state power amplifiers, as well as innovative antenna design. Development of the ACSSB itself, including the digital signal processing (DSP) implementation, had been performed earlier by Lodge[1] and others at the Canadian Department of Communications.

The second objective was to design a terminal which would be extremely portable, lightweight, and battery-powered and could be used to provide voice and data communication service, through Teleglobe Canada, within the MARECS B2 footprint. Typically users might include:

- News crews
- Disaster relief crews
- Diplomatic users
- Expeditions
- Construction project teams

From the outset it was known that there was a severe shortage of communication capacity for

new services on MARECS B2 and that only a single 25 kHz, 21 dBW (C to L-band EIRP) channel could be made available by Inmarsat for the briefcase terminal network. Furthermore, this channel had to be shared between the briefcase terminals and two existing Ontario Air Ambulance Service (OAAS) aeronautical mobile terminals.

In order to permit effective time and power sharing of the allocated channel by up to 20 users (18 LBT plus 2 OAAS), it was necessary to design and implement a small network control system. The Network Controller is a PC-based system which manages power, channel allocation, access priority, and call detail recording. It is installed at the network hub in the Laurentides earth station.

NETWORK OVERVIEW

The briefcase terminal network consists of the five main elements shown if Figure 1.

The MARECS B2 satellite, stationed at 26° W covers the Atlantic Region including Europe, Africa, South America, and eastern North America. It uses L-band (1.5/1.6 GHz) for the satellite-to-mobile link and C-band (4/6 GHz) for the satellite-to-hub link.

The LBT (up to 18 units) provides voice and data communications to the user in the field. It is operated in a "stop and talk" mode.

The OAAS terminals (2 units) are used for emergency communication to ambulance aircraft used by the Ontario Ministry of Health.

The hub of the network is Teleglobe's 18-meter C-band Earth Station at Weir, Quebec, just north of Montreal.

The Network Controller, hub channel units, and PSTN interconnects are also situated at the Laurentides station.

BRIEFCASE TERMINAL DESIGN

The briefcase terminal was designed to meet the following overall requirements:

- Full duplex voice operation
- Full duplex data (2400 bps)
- Dial-up operation to the PSTN
- Weight not to exceed 32 lbs
- Deployment time 1 - 2 minutes
- 1 hour operation on internal batteries

The design approach is shown in Figure 2.

The antenna consists of two hybrid-combined, right hand circularly polarized, 16-element, microstrip patch array antennas, each with a nominal gain of 15 dBiC. The overall antenna gain is about 17 dBiC after accounting for combining and feeder losses Each panel is 17 3/4" x 14" x 5/8" and weighs under 2 lbs. The antennas attach to the LBT lid during operation and stow in the LBT carry case for transportation.

The diplexer is a low-loss (typ 0.5 dB) cavity filter design.

The 20 W SSPA uses a linear Ga As FET final device and bipolar transistor drivers.

The upconverter is synthesized in 5 kHz steps for channel selection and is of conventional design. The downconverter is synthesized in 50 Hz steps and provides both channel selection and AFC tuning capability for the demodulators.

The heart of the terminal is the ACSSB/DMSK modem. ACSSB is used for voice communications and the co-resident DMSK (differential minimum shift keying) is used for both data communications and signaling purposes. The modulator and demodulator are implemented on two TMS 320C25 digital signal processors operating at 40 MHz clock rate.

Call setup, mode control and the user interface are managed by a CMOS V40 control processor.

The user interface consists of a sonalert, a cellular radio style handset which includes a 2 line LCD display, status indicators, and a keypad for control and call set up. The handset also contains the microphone and earpiece.

The power system includes a power conditioner, sequencer, charging circuitry, and 20 'D' size NiCad cells.

Table 1 shows a brief specification of the LBT terminal:

PARAMETER	SPECIFICATION
Voice Mode	ACSSB
Data Mode	DMSK 2400 bps, sync
G/T	-9 dB/K
EIRP	30 dBW
Transmit Band	1636.5 - 1645.0 MHz
Receive Band	1535.0 - 1543.5 MHz
Occupied BW	5 kHz
Frequency Stability	±1.6 kHz
Operating Time	1 hour
Power	12 V @ 7A 115/230 V ac
Batteries	NiCad (20 'D' cells)
Size (briefcase)	18" x 13" x 4.5"
(packed)	19" x 14" x 8.5"
Weight (batteries, case, antennas)	32 lbs

TABLE 1 : LBT Specifications

Operation

Operation of the LBT involves three steps:

- Set-up
- Peaking on the satellite
- Call establishment

First the LBT is unpacked and placed on a surface with a clear line of sight to the satellite. The two antenna panels are attached to the lid with velcro fasteners and the RF cables are connected.

The unit is then powered on and the briefcase is oriented in the general direction of the satellite. In a special antenna-pointing mode the demodulator measures the amplitude of the satellite pilot tone and modulates an audio tone in the earpiece. Peaking the antenna towards the satellite is simply a matter of adjusting the briefcase lid, in azimuth and elevation, for the highest pitch audio tone. At this point, the lid is locked in position and a call may be placed.

Placing a voice call involves entering the called party's PSTN number and pressing the call originating buttom.

For data calls, the LBT appears like a "Hayes" modem. Calls can be placed manually using the data port or with a commercial software package.

Calls can also be placed in the reverse direction (i.e., from the PSTN to the briefcase). To do this, the calling party first dials the Laurentides Earth Station's designated line and then overdials to the desired called LBT party.

Performance

Table 2 shows the ACSSB link budget.

PARAMETER	LBT TO HUB	HUB TO LBT
Uplink		
EIRP up (dBW)	24.0	59.1
FSL (dB)	189.0	200.9
Misc. Loss (dB)	0.5	1.5
G/T (dB/K)	-10.9	-14.0
C/Io (dB-Hz)	60.0	67.1
C/No th (dB-Hz)	52.2	71.3
C/No uptot (dB-Hz)	51.5	65.7
Downlink		
EIRPdn (dBW)	-16.5	18.0
G/T (dB/K)	36.3	-9.9
FSL (dB)	197.1	188.4
Misc. Loss (dB)	1.7	0.5
Boltz Const	-228.6	-228.6
C/No th (dB-Hz)	49.6	47.7
C/No dn (dB-Hz)	49.6	47.7
C/No link (dB-Hz)	47.4	47.6

Table 2 : ACSSB Link Budget

The EIRP values shown are average values and, in the LBT, correspond to between 5 and 6 dB output back off from saturation.

ACSSB voice quality is excellent at 50 dB-Hz, good at 47 dB-Hz, and intelligible down to 38 - 40 dB-Hz.

Data communications performance is excellent (BER $\leq 10^{-4}$) for C/N_0 values above 44 dB-Hz.

NETWORK OPERATION

The Briefcase Terminal Network was authorized to operate with the following constraints

- Maximum L-band EIRP of 21 dBW
- OAAS carriers require the full 21 dBW
- LBT carriers require 18 dBW each
- Occupied bandwidth 25 kHz
- Channel time/power utilization records
- Carrier down when idle operation

Under these design constraints a three-channel system and a PC-based network controller emerged. Figure 3 shows the frequency plan written for allocated 25 kHz bandwidth. The network control computer manages the hub channel units to ensure the above constraints are always met.

Channel Assignment

Each of the three channels was assigned a specific purpose. Channel 1 was assigned for OAAS-1 calls, OAAS-2 call requests and potentially LBT preemptive requests. No LBT calls are allowed on this channel, keeping it free for OAAS call setups. OAAS-2 shares its call channel with one of the LBT groups so that it does not interfere with the OAAS-1 air unit.

Channels 2 and 3 remain for use by the LBTs. A circuit sharing algorithm was developed that allowed a hub channel unit to uniquely address a LBT within its own group. This algorithm allows a number of LBTs to share a single hub channel unit. It supports calls both into and out of the hub. Idle LBTs can detect when their hub channel unit is busy. When its hub channel unit is busy the LBT will block any further call setup attempts until the channel becomes free.

The system design assumes that OAAS is inactive a large percentage of the time since the LBT user is generally not informed of OAAS call activity and if an LBT user makes a call request while OAAS is active, the request will not be answered.

Call Authorization/Billing

The network controller continually monitors the call state of each hub channel unit. It uses state transitions and their associated parameters to build call detail records. At the end of a call the call detail record is sent to disk and optionally to a printer. The PC clock is used for logging all call start and end times. Each call detail record contains the remote user identification number and the dialled PSTN number if it was an inbound call.

All outgoing and incoming call requests block the hub channel unit until it gets authorization from the network controller to transmit. Depending on the network's current status and the call request's status, the network controller either authorizes or terminates the call request. Before a call is authorized the network controller will allocate any network resources the pending call will require.

Resource Management

The network configuration is defined in a database created by a separate setup program. Three parameters are stored for each hub in the data base

- Group number
- Priority
- EIRP allocation

All hub channel units with the same call channel share a common group. In no case will the network controller authorize two hub channel units within the same group. Groups allow different services and different hub channel units to share a common frequency. In the system, for example, one group will have a hub channel unit connected to the PSTN while another hub channel unit will connect to a

dedicated line.

Each hub channel unit is assigned a priority level. The network controller will terminate any existing calls when a higher priority call request arrives.

EIRP allocation presently is quite simple but adequate for the three channel system. Only two EIRP settings are possible: either the hub requires all the EIRP or it shares the EIRP. If it requires the entire EIRP it must be the only hub transmitting. The network controller assumes that total EIRP of all groups with hubs sharing EIRP does not exceed the total allocated EIRP. In this network all LBT hub channel units must be set for a maximum EIRP of 18 dBW.

Preemption

The network controller supports preemptive LBT calls. A preemptive call allows a LBT to terminate the call current within its group and subsequently place its own call. In order to provide a preemptive capability, the network requires a dedicated hub channel unit in the LBT's call group that listens to a separate channel. Like all hub channel units, the preemptive hub channel unit must first get authorization before it completes the call. The call setup always proceeds on the group call channel.

A "type of service" parameter is included in the configuration database. This parameter allows the network controller to move preemptive hub channel units to a second channel to listen for call requests.

Table 3 summarizes the configuration parameters for the hub channel units in the network.

HUB	GROUP	PRIORITY	EIRP	SERVICE
OAAS-1	1	1	All	Normal
OAAS-2	3	1	All	Normal
PSTN	2	3	Share	Normal
PSTN	3	3	Share	Normal
DATA	3	3	Share	Normal
Preempt	3	2	Share	Preempt

Table 3 : Hub Configuration Parameters

CONCLUSION

The paper has presented details of the L-band Briefcase Terminal and its associated Network Control System.

SkyWave wishes to thank DOC, Teleglobe, and Inmarsat for their support in this activity.

REFERENCES

(1) Lodge, J and Boudreau, D. "The Implementation and Performance of Narrow Band Modulation Techniques for Mobile Satellite Application". IEEE International Conference on Communications, Toronto 1986, Conference Record.

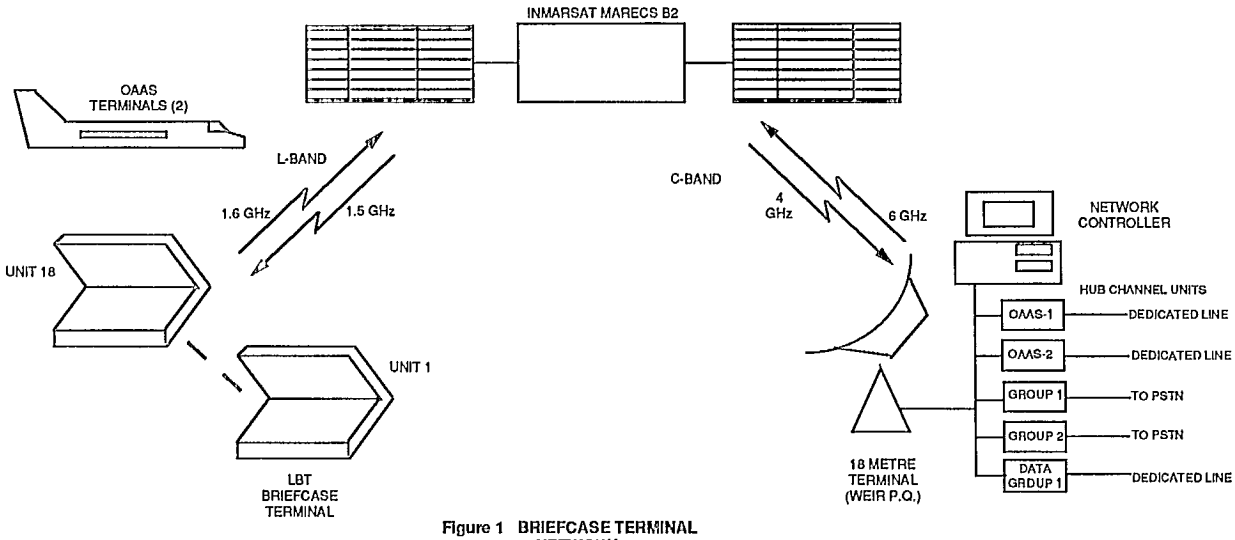


Figure 1 BRIEFCASE TERMINAL NETWORK

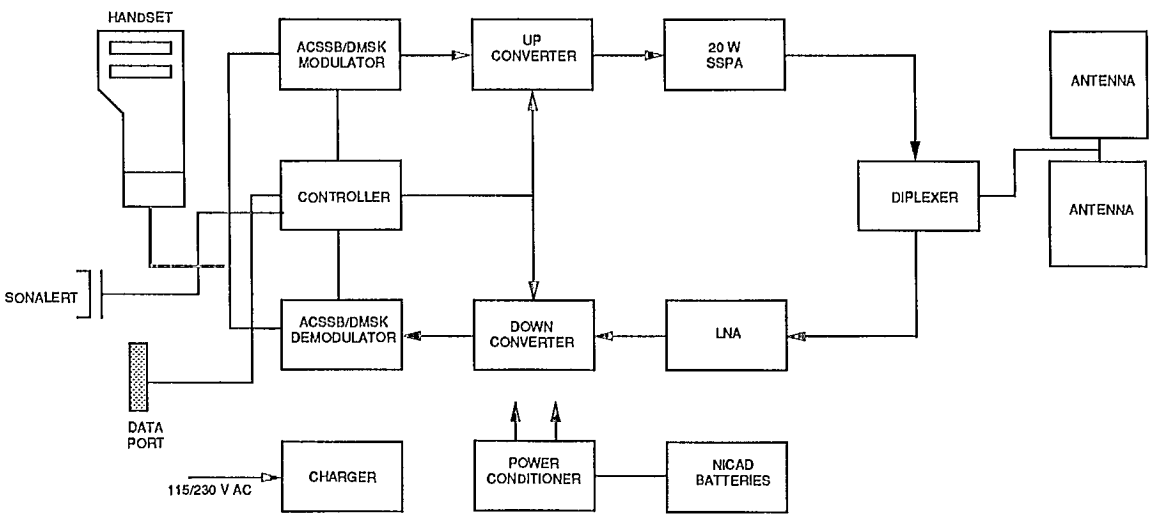


Figure 2 BRIEFCASE TERMINAL DESIGN APPROACH

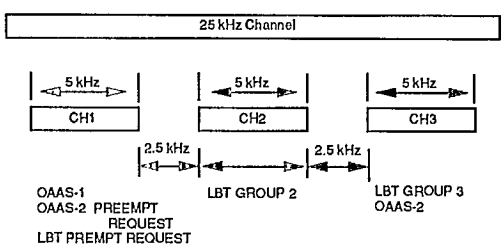


Figure 3 FREQUENCY ALLOCATIONS

A Description of QUALCOMM Automatic Satellite Position Reporting (QASPR®) for Mobile Communications

William G. Ames
QUALCOMM, Inc.
10555 Sorrento Valley Road
San Diego, CA 92121
Phone: (619) 587-1121
Fax: (619) 452-9096

ABSTRACT

Two-Satellite position reporting for mobile communications has been introduced into the OmniTRACS mobile satellite communication system. The first engineering demonstration in the USA occurred on 7-January-1990. This system significantly improves position reporting reliability and accuracy while simplifying the terminal's hardware. The positioning technique uses the original OmniTRACS TDMA timing signal formats in the forward and return link directions plus an auxiliary, low power forward link signal through a second satellite (Ranger) to derive distance values. The distance values are then converted into the mobile terminal's latitude and longitude in real time as messages or acknowledgments are received at the Network Management Center (NMC). A minor augmentation of the spread spectrum profile of the return link allowed the resolution of periodic ambiguities. The system also locates the two satellites in real time with fixed platforms in known locations using identical mobile terminal hardware. Specialized hardware in the satellite was not required, and the reference oscillator in the mobile hardware was held to the original specifications, 10 ppm long term, 100 ppb short term. New mobile terminal software provides for pointing the directional OmniTRACS antenna on a scheduled basis plus acquiring and tracking of the Ranger's Forward Link signal. Initial accuracies on the order of 1/4 mile have been realized uniformly throughout the USA using a satellite separation of 22 degrees and there are no dead zones, skywaves, or cycle slips as found in terrestrial systems like LORAN-C. An altitude map provides distance from the center of the earth. As of 24-January-1990, OmniTRACS in Europe is hosting the first operational implementation of the positioning system where a 9 degree satellite separation is exhibiting 2/3 mile random errors as expected. These accuracies are first results and will be improved in the near term by system tuning, wider baselines in the fixed-sites, a more precise topographic model as needed and finer timing in the forward and return signals.

INTRODUCTION

QASPR is the acronym for QUALCOMM's Automatic Satellite Position Reporting service in the OmniTRACS mobile communication system. This is QUALCOMM's preferred method of position reporting over LORAN-C because of its consistency, reliability, economy and accuracy. QASPR is controlled at the Hub by the Network Management Center (NMC) computer. The system uses a small amount of transponder capacity in a second satellite, called the Ranger, in a leased agreement similar to the messaging satellite's lease arrangements. The full time lease with back-up, allows protection against satellite failures, planned outages or government policies beyond QUALCOMM's control, contrary to our experience in working with external positioning systems like LORAN-C and GPS or highly specialized satellite systems. The OmniTRACS mission is to provide a complete two-way mobile communication service under one roof, with consistent, reliable messaging and positioning 24 hours a day, 365 days a year. QASPR now provides the internal position reporting component to that mission. QASPR also performs real-time satellite location tasks, further eliminating complexity and dependence on external sources for ephemeris data. Figure 2 shows the standard system block diagram of the OmniTRACS communication system.

PERFORMANCE

Since the first engineering demonstration in early January 1990, QASPR continually demonstrated reliable and accurate performance even while vulnerabilities in the system operation were identified and protections built in. QASPR has also been a integral part live OmniTRACS operations in Europe since January 24, 1990. QASPR will be available commercially in the USA beginning in May 1990.

Table 1 summarizes QASPR's positioning performance in CONUS using 12° satellite spacing. This is expected performance from the composite total of range errors, including 'one-shot' timing errors to the mobile unit and current satellite location errors. The actual system is currently performing at one-half this error due to a satellite separation of approximately 24°. Altitude model error accounts for an additional absolute positioning bias, typically less than 1/4 mile. Additional positioning faults result from matching solved positions to dated information in the map database. This has introduced a "position war" element into the system, which is usually decided in favor of the satellite-based vehicle solution. This was confirmed with independent GPS fixes. OmniTRACS easily supports reporting of external system data such as GPS and comparisons can be made simultaneously.

Table 1. Expected Vehicle Position Performance (position error in miles)

User Location	Vehicle Positioning Error at Earth Surface
San Diego	0.36
Pacific Northwest	0.35
North Midwest	0.32
New England	0.36
South Florida	0.37
South Texas	0.31
Satellite Separation = 12° (103°W and 91°W	

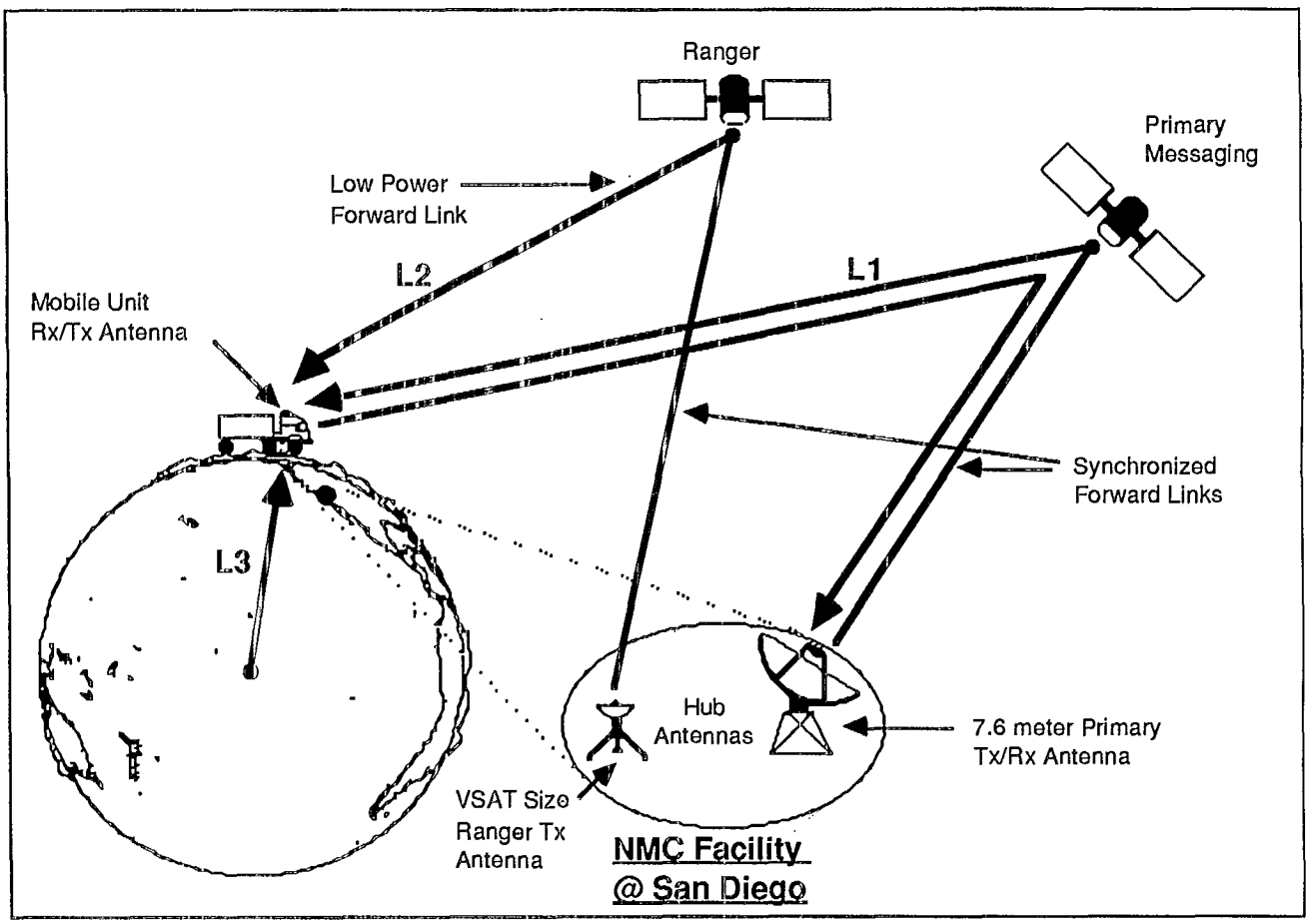


Figure 1. QASPR System Components

SYSTEM MECHANICS AND VEHICLE POSITIONING

The QASPR method uses two satellites, separated by 12° in orbital arc, to obtain positioning accuracy of 1/2 mile or better on the earth surface. Submicrosecond timing in the round-trip-delay, and a delta-time measured at the MCT (mobile communications terminal) provide all the inputs to compute the MCT position. The OmniTRACS Return Link Receivers (RLR) measure round trip delay during receipt of any return packet or acknowledgement. Delta-time is a derived value from the MCT comparing the reception of two synchronized Forward-Links. Vehicle position is achieved by multilateration from three known points in space.

Figure 1 highlights the system components involved in QASPR. Normal messaging is performed through the Primary satellite with Forward and Return links. Measuring and tracking round-trip-delay is a necessity for receiving return messages and therefore was already in place for the QASPR system, excepting finer resolution of the measurement. The Ranger uplink antenna provides a one-way, low power, Forward-Link signal (similar to the primary with no phase modulation) through the Ranger satellite. The period of frequency modulation in this identical copy of the primary waveform is long enough in time so that positional ambiguities do not arise when MCTs across the network compare the two signals. The MCT acquires the Ranger signal as necessary (see Figure 3) and reports the derived time difference with any return message or acknowledgements of forward messages. The NMC then processes the round-trip-delay and delta-time pair immediately as the return message is compiled for delivery to the customer. The two pieces of timing information define distances to the two satellites, L1 and L2 Figure 1, while the third distance, L3, is derived from a model of the earth's shape. A topographic model improves accuracy beyond what is achievable using only a smooth earth model.

Since QASPR incorporates and supports all of the previous methods of position reporting, a brief review of the position reporting mechanism using a sensor card, such as LORAN-C, will help clarify the QASPR technique. With a sensor card, MCT position is defined by an external navigation system. The sensor card communicates the position information directly to the main processor card in the MCT's Indoor Unit. Acknowledgements, position polls or mobile initiated messages use reserved packets to report the position information. The OmniTRACS system delivers the position and messages to the customer following their arrival through the satellite link. A single satellite transponder mix (two transponders) is used for communication to and from the MCT, even though there may be several transponders/satellites supporting a large community of users (10,000 or more). There is only one Forward-Link to the MCT and a Return-Link through the same satellite. Hence, in the original position reporting mode, the mobile's antenna seeks to stay focused toward one direction in the sky.

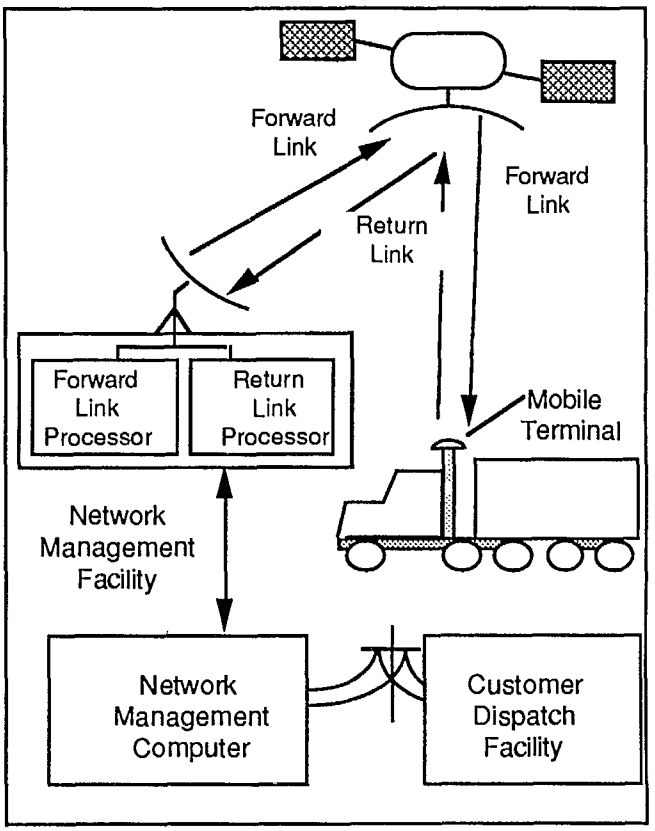


Figure 2. OmniTRACS Block Diagram

In the new QASPR technique, the same MCT antenna acquires and tracks two separate Forward Link signals from two different antenna pointing directions. Figure 3 shows the sequence of events. The additional Forward Link arrives from the Ranger satellite, 12° to 24° away from the primary messaging satellite depending on

vehicle location in CONUS. Hence, the MCT must locate the Ranger at start up to expedite Ranger acquisitions when it's time to perform positioning tasks. Typically, the MCT performs Ranger acquisitions only infrequently as commanded by NMC control messages, or as required for scheduled position reporting intervals. Unsolicited position reports (mobile initiated) or specific events at the MCT, other than messaging, may trigger position reports, e.g., trailer disconnects or vehicle sensor data exceeds tolerance values. The position report packet now contains only delta-time information rather than a true position as in the previous case.

Fine time determination in round trip delay required a minor adjustment of the return link waveform to increase the code length in the spread spectrum component of the wave form. Round trip delay prediction is already necessary for return link messaging otherwise detection and demodulation of the message will not happen reliably. For QASPR, forcing exact alignment of the TDM frame boundaries with a longer spread spectrum code provided the precision required. It was also necessary to make the long code work with the original shorter codes received in the same radio hardware. Hence this avoids a major retrofit of software in the field.

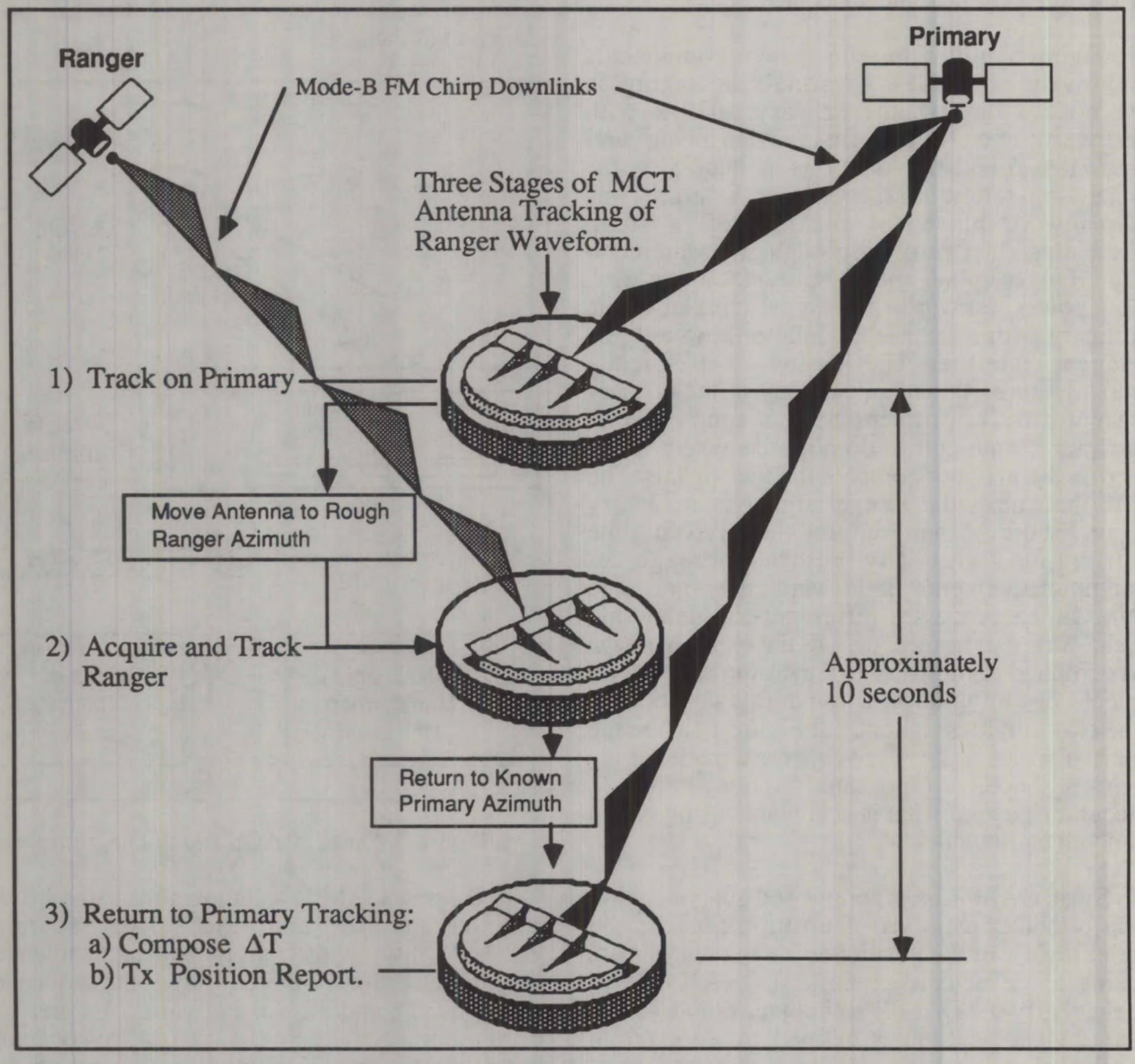


Figure 3. MCT Antenna and Signal Tracking Functions

RANGER SATELLITE FUNCTIONS

The major new MCT function is the determination of the delta-time (hop time offset) between two arriving waveforms: one from the Primary satellite and one from the Ranger satellite. The current MCT hardware cannot track these two signals in parallel because the MCT antenna must point at each satellite individually to provide sufficient signal gain. Therefore, the MCT timeshares receiving the waveform from both satellites to obtain the hop time offset.

As shown in Figure 3, the MCT antenna stops tracking the Primary's downlink, acquires and tracks Ranger hop timing, then returns to the Primary. Because of the importance of a precisely controlled time synch for Return Link message transmissions, the delta-time is measured in the last stage of the operations when the MCT antenna returns to the Primary. This allows the most precise determination of the delta-time because returning to Primary tracking from Ranger tracking is much shorter than the initial changeover to the Ranger signal acquisition and tracking. This includes time for antenna azimuth rotations for a total of about 10 seconds. So, upon return to the Primary, the delta-time is measured then prepared for the return message.

The mobile unit does not require new performance levels in mobile terminal clock accuracy. The mobile terminal clock has life time stability of better than 10 ppm however a short term stability of better than 100 ppb is the necessary specification for the QASPR technique to perform to the positioning accuracies mentioned earlier. This requirement again allowed the introduction of the QASPR into the OmniTRACS system without major hardware redesign. In total, the QASPR method has allowed for the elimination of the LORAN sensor card and its associated antenna hardware plus cabling and installation, and the addition of a small chip to implement a specialized long spread spectrum code for the return link.

SATELLITE POSITIONING TECHNIQUE

Although the primary function of QASPR is locating the MCTs, a reliable means of supporting that function with current and accurate satellite positions is necessary. Rather than obtain ephemerides from satellite controllers, satellite position is obtained through the reverse process of multilaterating the satellite from Fixed MCTs.

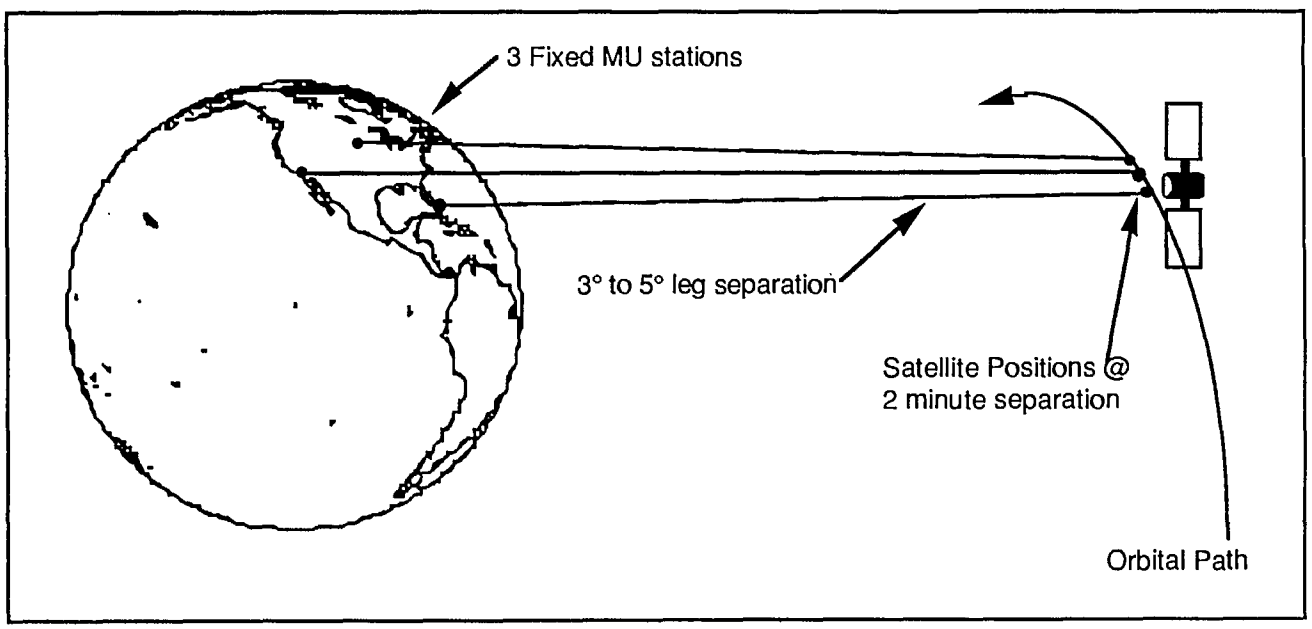


Figure 4. Satellite Positioning From Fixed Mobile Units

The alternative of obtaining a satellite ephemeris report from satellite owners implies a delayed version of what the satellite was doing perhaps weeks ago, and therefore will lack the required accuracy at current time. This approach also has the undesirable complexity of setting up and organizing an efficient interface with every possible satellite owner who could provide the transponder capacity. The ability to pinpoint the satellite in real time provides a robust positioning system which immediately follows any station keeping maneuvers and quickly adjusts to backup satellite configurations in the unlikely event of a satellite failure.

As shown in Figure 4, three or more fixed observing sites are used to determine the satellite positions in real-time. One is located with the Hub station. At least two others are placed in geometrically diverse locations to provide the most favorable conditions for the satellite position solutions. The station locations are defined from survey data and the lengths to the satellite are defined through the same process as vehicle positioning. Hence, Fixed Units are identical in form and function to the mobile equipment. The NMC makes the distinction regarding Fixed Units status.

FIXED UNIT ALLOCATION

Fixed MCTs are used to track and determine the positions of the two satellites in real time. There is complete identity in hardware and software used by the Fixed Units and the mobiles. This provides for a robust system that doesn't require two separate code versions and each group (mobiles and fixed) can act as continuous test beds for the other. The main difference within the system between them is how the NMC views them. Geographical location is important from the standpoint of positioning accuracy both for the satellites and the MCT location.

Figure 3 (near to scale for satellite distance to the earth) shows that the angular separation of directions from example Fixed Unit locations toward the satellite are very narrow even for the wide baselines provided by the CONUS geography. Therefore, Fixed Units must be spread out as far as possible, yet have antenna gains high enough to satisfy link budget requirements. For reliability, the five Fixed Unit locations shown in Figure 5 have been installed. Any three (including San Diego) of the five can provide primary location data; the spares provide backup if a primary fails or loses its link. The five fixed unit sites are currently operating near the following cities:

San Diego, CA
Minot, ND
Seattle, WA
Bangor, ME
Melbourne, FL

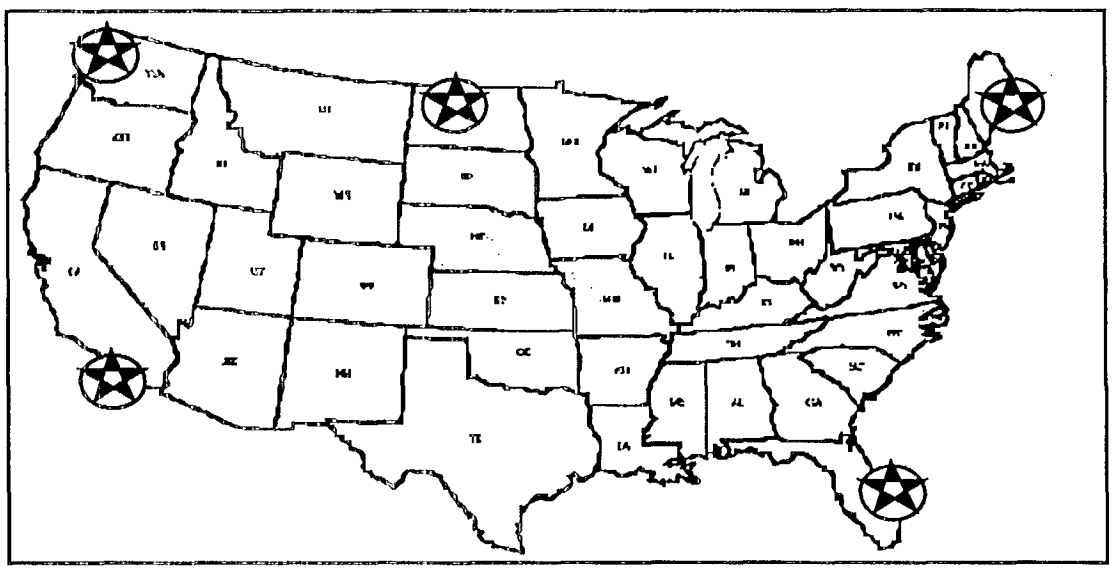


Figure 5. Fixed Unit Locations for CONUS

Frequency Stabilization for Mobile Satellite Terminals via LORAN

Gregory J. Ernst, Steven M. Kee,
Robert C. Marquart
Hughes Network Systems
11717 Exploration Lane
Germantown, MD-20874
United States
Tel: (301) 428-5500 Fax: (301) 428-1868

SYNOPSIS

Digital satellite communication systems require careful management of frequency stability. Historically, this has been accomplished by continuously powered, high cost, high performance reference oscillators. Today's low cost mobile satellite communication equipment must operate under wide ranging environmental conditions, stabilize quickly at power up, and provide adequate performance margin to overcome RF link impairment unique to the mobile environment. Methods for frequency stabilization in mobile applications must meet these objectives without incurring excessive performance degradation. A frequency stabilization scheme utilizing the LORAN (LOng RAnge Navigation) system is presented.

Position determination and reporting are requirements in many mobile satellite terminal applications. The LORAN system is widely available and a low cost means of providing position determination. Many mobile satellite terminals include an integrated LORAN Receiver for this purpose. A high stability frequency reference can be derived from the received LORAN signal. Application of this technique to transmit-only mobile satellite terminals is particularly useful.

The LORAN Receiver compares the frequency error of the mobile terminal's crystal oscillator reference to the received LORAN signal. Periodically, the terminal's processor system interrogates the LORAN Receiver to obtain this frequency error. After digital processing, a frequency adjustment is fed to a digital to analog converter whose output is used to tune the crystal oscillator for minimal frequency error. This negative feedback servo loop removes the frequency offset effects due to temperature and time.

The implementation of the servo loop incorporates features to ensure reliable frequency stabilization. The LORAN Receiver computes the frequency error as well as a received signal quality indicator. A frequency error accompanied by a poor signal quality indication is ignored by the microprocessor. Otherwise, the frequency error is processed along with the existing frequency offset to establish an updated frequency adjustment. The most recent frequency adjustment is always stored in nonvolatile memory so that upon power up the previously stabilized frequency can be obtained promptly. During periods of extended LORAN outages, the stability is maintained because the crystal oscillator is temperature compensated. In addition, the crystal oscillator provides the short term frequency stability performance of the terminal.

Design evaluation tests of the Hughes Network Systems SkyRider™ Mobile Satellite Terminal demonstrate freqency stability of better than 3×10^{-8} over temperature and time under continuous LORAN coverage. The long term aging effects can easily be maintained to less than 10^{-7} with only monthly LORAN availability. Temperature stability without LORAN is a function of the design of the crystal oscillator, but can be less than 10^{-7} with a state of the art implementation.

The full text of the paper is in the Appendix ..... page A-9

Mobile Terminal Equipment Design Utilising Split-loop Phase-lock Techniques

P.B. Kenington* J.P. McGeehan* D.J. Edwards†
University of Bristol
Centre for Communications Research
Queens Building, University Walk
Bristol BS8 1TR, United Kingdom
Tel: +44 272 303726, Fax: +44 272 255265

Abstract

The design and resultant performance of the terminal equipment in a mobile satellite system is vitally important in respect of the overall cost/performance compromise of the whole system. Improvements in system performance which also result in a reduction of the equipment cost are rare. However, this paper details a significant advance in terminal design, utilising a novel form of 'Split-loop' phase-locked receiver/downconverter system to enable an accurate, stable and wide coverage terminal to be realised at a reduced cost. The system has the capability of automatically locking onto any carrier within a complete transponder, and can cope with severe amplitude modulation and fading effects.

1 Introduction

The split-loop phase-lock configuration has replaced the conventional long-loop as the state of the art in PLL receiver technology. It has been shown to possess superior qualities in both acquisition and in the tracking of difficult signals and has recently been applied to the field of mobile satellite receiver design in conjunction with the Electronic Beam-Squint tracking system [1], [2]. The superior stability of the design can be exploited to permit phase-locked control of the complete receive/downconversion chain, thus giving far greater versatility to a mobile terminal.

The split-loop phase-locked loop (PLL) receiver configuration was originally conceived as a method of eliminating the problem of false-lock inherent in the long-loop PLL configuration [3]. The problem of false-lock manifests itself by the cessation of the natural pull-in effect of the long-loop on a frequency which bears no obvious relationship to the wanted signal or its modulation (if present). The effect is caused by the delay present in the receiver's IF stages, and particularly the crystal filter. The split-loop is able to eliminate this effect as it effectively removes the IF delay from the time critical parts of the phase control loop. It is this elimination of the effects of the IF delay which also leads to a number of other advantages of the technique, and these will be discussed below.

One of the candidate schemes for mobile terminal tracking is the Electronic Beam-Squint (EBS) tracking system. The EBS system is a new form of satellite tracking system combining an accuracy approaching that of monopulse with a cost of the same order as that of a step-track system. The tracking information in this system is derived by switching the main lobe direction of the antenna into four positions around boresight using PIN diodes. This results in a very rapid method of determining the tracking error and leads to the possibility of performing averaging on the received tracking information in order to remove noise and scintilation effects. The result of this is that the antenna need only move to actually follow the satellite, and not also to determine the tracking information. This leads to a further advantage in the reduced wear of the antenna positioning mechanics.

The rapid switching operations described above lead to a switched amplitude modulation on the received beacon signal as seen by the tracking receiver. The tracking receiver must be able to accurately detect and reproduce this modulation in order for the tracking controller to make correct decisions. The current long-loop design of receiver is unable to perform this task satisfactorily due to the presence of the IF delay within the loop.

*University of Bristol, UK
†University of Oxford, UK

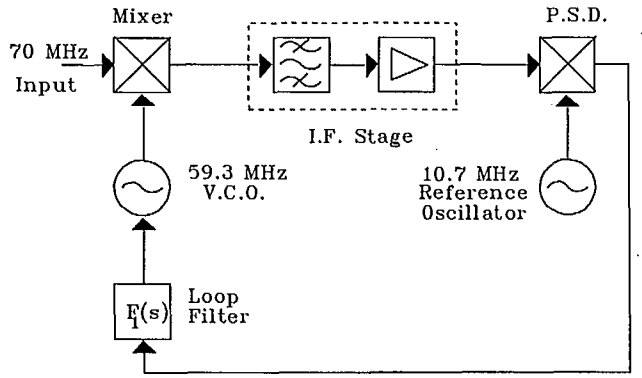


Figure 1: Configuration of the long-loop receiver

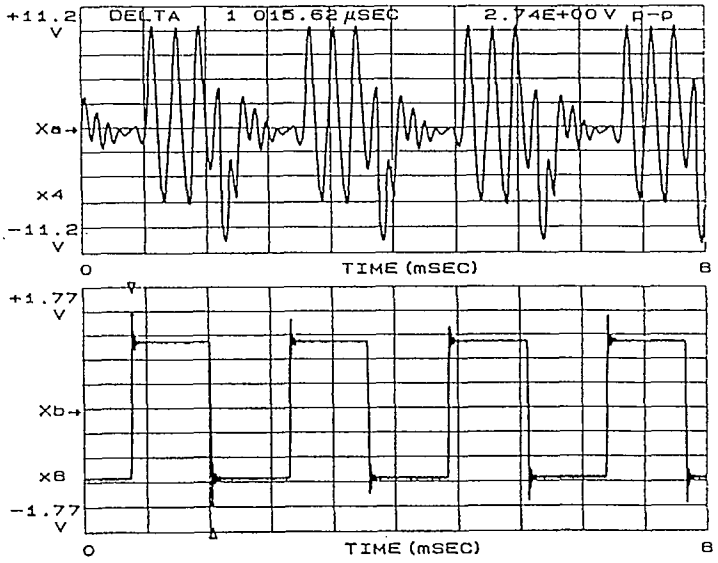


Figure 2: Upper trace: Experimental results showing the coherently detected output of a long-loop receiver, when receiving a squarewave AM signal with a modulation index of 0.95 and using an IF filter bandwidth of 15 kHz.
Lower trace: Squarewave modulating signal.

2 The Long-Loop Receiver

In general, current tracking receivers are based around the 'Long-Loop' phase-lock technique. The basic configuration is shown in block diagram form in Figure 1 [4]. This is intended to be a simplified diagram, and does not include any of the automatic gain control or coherent detection elements which are assumed.

Enhancements of the conventional long-loop receiver have been developed for use in satellite systems. One example is described in [5], utilising two phase-locked loops to detect both co-polar and cross-polar beacon polarization states. Level detection in this case is provided coherently as mentioned above.

There are a number of disadvantages inherent in this type of receiver. Biswas *et al* [6] noted the problem of false-locking, and has suggested a method for its elimination. However, this method is less than straightforward to implement, and its inclusion greatly complicates the loop analysis under certain circumstances.

A second disadvantage is inherent in the loop filter utilised in the long-loop configuration. This filter requires a proportional term as part of its characteristic (to obtain a non-zero damping factor [4]), and as a result, any noise on the beat-note will phase modulate the down-converting local oscillator. Modulation components, such as those described below, which appear at the output of the IF filter, will be mixed with the reference oscillator in Figure 1. The resulting sum and difference outputs of the phase sensitive detector give rise to sidebands at the first local oscillator.

Further to these inherent problems, E.B.S. causes the beacon signal to be corrupted by severe, discontinuous amplitude modulation, and this creates a number of problems of its own. This is not true of, say, step-track or conical scan systems, as they produce a continuous amplitude modulation which is generally of quite a low level. It is essential to construct a receiver which has the ability to supply a valid output to the tracking controller and one which is as stable as possible when faced with this type of signal.

If we now examine, in detail, two of the problems which EBS causes to a long-loop tracking receiver, it will become apparent why the long-loop is inappropriate for this system. Figure 2 (top trace) shows the coherently detected output of a long-loop receiver, when receiving a carrier which is amplitude modulated by the squarewave shown in the lower trace. It is evident from this figure that the long-loop is slipping cycles when faced with the severe amplitude modulation representative of the signals produced by the EBS system. This effect is caused by the inclusion of the IF stage delay within the control loop of the long-loop receiver. Thus a long-loop receiver is unable to correctly interpret the EBS tracking information present on the beacon signal.

Figure 3 shows the first local oscillator spectrum of a long-loop receiver when receiving the signals mentioned above. It can be seen that this oscillator is being

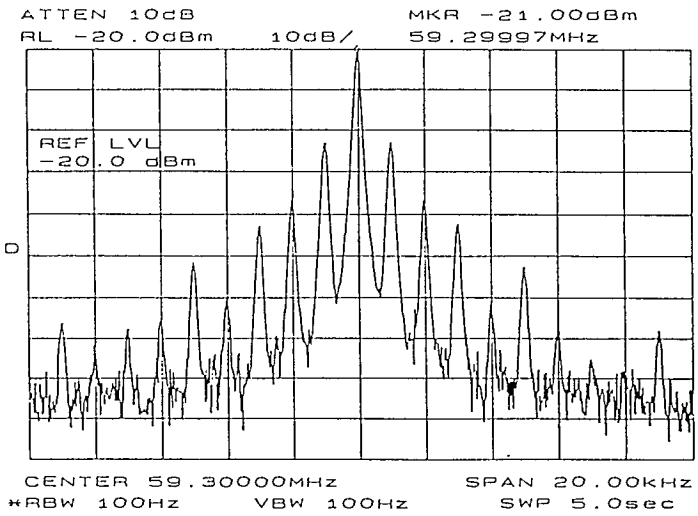


Figure 3: Experimental result showing the first local oscillator spectrum of a long-loop receiver, when receiving a squarewave AM signal with a modulation index of 0.95 and using an IF filter bandwidth of 15 kHz.

phase-modulated by spurious signals present on the receiver's beat-note, resulting in undesirable sidebands. These sidebands can downconvert adjacent channels, which will in turn cause interference to the wanted channel and degrade the performance of the receiver. This problem, known as *reciprocal mixing*, has previously been reported in the literature [7].

3 The Split-Loop Receiver

The split-loop derives its name from the technique of 'splitting' the loop filter into proportional and integral components, with each feeding its own local oscillator. The first local oscillator of Figure 4 is fed by an integrator, with the second being fed by a directly-coupled (DC) broadband amplifier.

Thus, it can be seen that the integrator and first local oscillator are responsible for tracking long-term drift of the carrier, with the proportional stage and second local oscillator passing the beat-note signal and tracking rapid frequency changes, such as doppler shifts. Since the loop formed around the PSD, the proportional stage and the second local oscillator is effectively first-order, the beat frequency bypasses the IF stage, and all other primary sources of delay, and

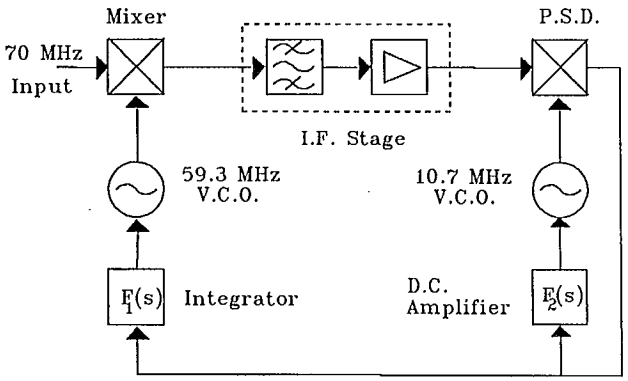


Figure 4: Configuration of the split-loop receiver

as a result the possibility of false-lock is eliminated [8].

Furthermore, since the action of correcting for rapid changes is independent of the IF delay, the split-loop is better able to correct for the rapid changes in loop bandwidth caused by the tracking information present on the beacon signal in the Electronic Beam-Squint tracking system. This is demonstrated by Figure 5 which shows the split-loop response under the same conditions as those of the long-loop above (Figure 2).

Finally, if we again examine the spectrum of the first local oscillator (Figure 6), a notable improvement in the level of the unwanted sidebands can be observed. This is due both to the improvement in overall stability of the loop, and also to the fact that the first local oscillator is fed by an integrator, which greatly attenuates any residual modulation on the beat-note. This modulation is thus effectively isolated from the oscillator input, resulting in its improved spectral purity.

4 Incorporation of the Down-converter in the Split-loop Mobile Terminal

The superior performance of the split-loop configuration makes it ideal for use in a mobile terminal, particularly if the downconversion stages are included within the loop. Figure 7 outlines a new form of satellite receiver which permits locked frequency control to extend to include the front-end downconversion system. Again, only the main loop components are shown, with coherent detection elements being assumed.

This new technique enables the whole of the down-

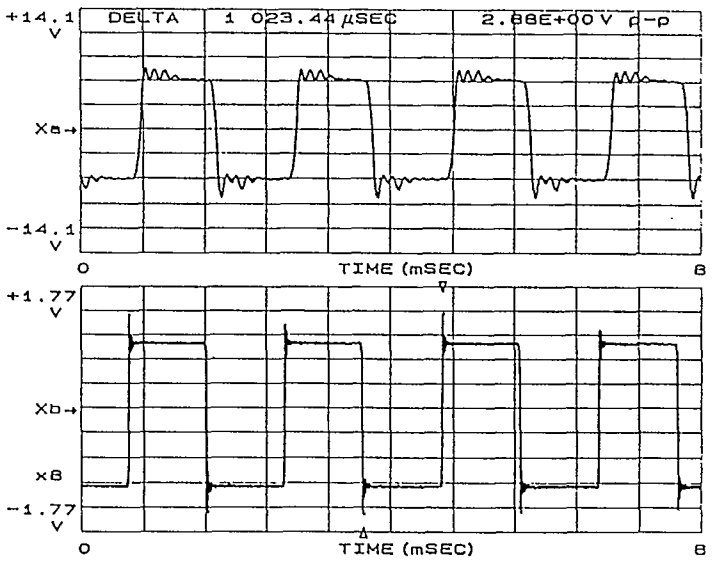


Figure 5: Upper trace: Experimental results showing the coherently detected output of a split-loop receiver, when receiving a squarewave AM signal with a modulation index of 0.95 and using an IF filter bandwidth of 15 kHz.
Lower trace: Squarewave modulating signal.

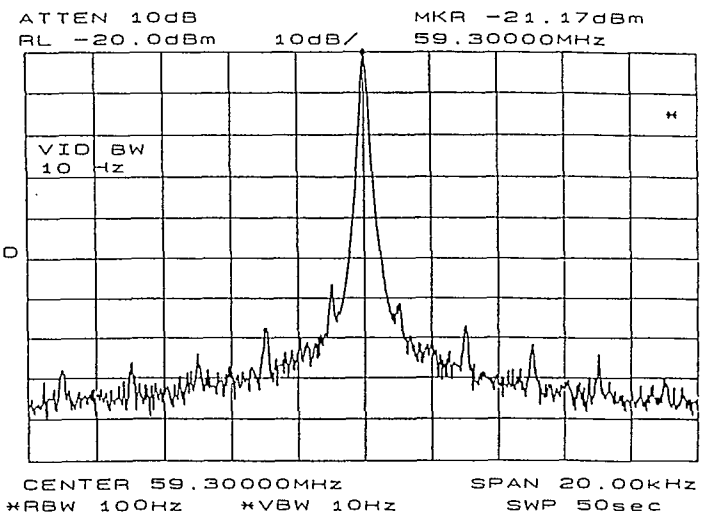


Figure 6: Experimental result showing the first local oscillator of a split-loop receiver, when receiving a squarewave AM signal with a modulation index of 0.95 and using an IF filter bandwidth of 15 kHz.

conversion, tracking and reception chain to be phase-locked using a single loop. This in turn results in the frequency coverage of the system being limited only by the downconverter system employed, thus permitting the whole of a transponder or even a whole satellite allocation to be accessed automatically. Thus, a system can be conceived in which the operator requires only to aid in the initial satellite positional acquisition, with all further operation in both tracking (utilising the 'Electronic Beam-Squint' tracking system [1], [2]) and frequency acquisition being automatic.

A principle advantage of this configuration is the ability for automatic acquisition and tracking over a wide frequency bandwidth. This is extremely advantageous in a military scenario, as it provides almost total immunity from jamming of the beacon signal, since automatic tracking and wide-bandwidth locking would permit operation anywhere within the band, and not only on the beacon signal.

In a civil context, terminal equipment designed using this technique would be both cheaper and easier to operate. The specification of the intermediate local oscillators is relaxed, resulting in a more cost-effective system, and this is combined with the capability of fully automatic operation.

5 Conclusions

This paper has outlined the advantages of the split-loop receiver configuration, particularly with reference to its remarkable ability to cope with severe forms of modulation and fading. The system as described has the ability to automatically control the operation of the downconverter to enable tracking to be performed on any carrier within a transponder. This removes the need for manual tuning of the downconverter or an expensive microwave synthesizer.

The use of the split-loop in a new form of mobile satellite receiver results in considerable improvements in many areas, together with a reduction in cost due to a relaxation of the specification of some of the system components. Although this receiver configuration was originally envisaged for use with the Electronic Beam-Squint tracking system, its numerous advantages make it ideal for use in most phase-locked mobile satellite receiver applications.

6 Acknowledgements

The first author is grateful to the U.K. Science and Engineering Research Council for the award of a research studentship. All of the authors are grateful to British

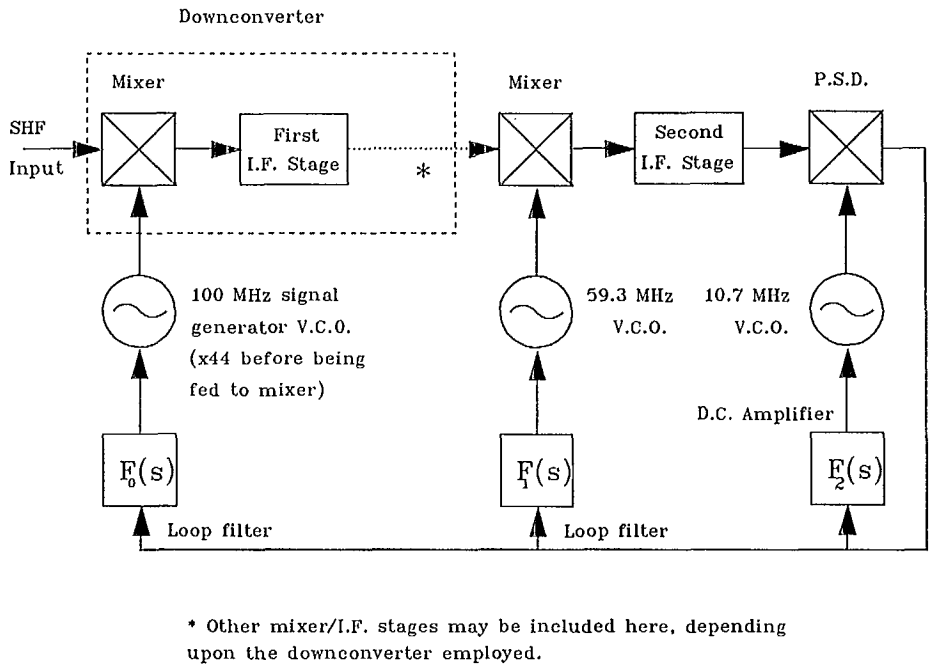


Figure 7: Configuration of the split-loop receiver/downconverter mobile terminal

Telecom International for the use of research facilities at the Goonhilly Down Satellite Earth Station.

References

[1] P.B. Kenington, D.J. Edwards and J.P. McGeehan, "Receivers for an electronic beam-squint tracking system," *IEE Proc. I*, vol. 136, no. 4, pp. 305–311, August 1989.

[2] P.B. Kenington and D.J. Edwards, "Tracking receiver design for the electronic 'beam-squint' tracking system in the mobile environment," in *Proceedings of the 39th IEEE Vehicular Technology Conference*, pp. 742–748, San Fransisco, April/May 1989.

[3] J.P. McGeehan and J.P.H. Sladen, "Elimination of false-locking in long-loop phase-locked receivers," *IEEE Transactions on Communications*, vol. COM–30, no. 10, pp. 2391–2397, 1982.

[4] F.M. Gardner, *Phaselock Techniques*, New York: Wiley, 1979.

[5] F.T. Tseng and L.S. Lee, "Accurate detection of satellite beacon polarization states using cascaded heterodyne phase-locked loops," *IEEE Transactions on Communications*, vol. COM–30, no. 1, pp. 283–293, January 1982.

[6] B.N. Biswas, P. Banerjee and A.K. Bhattacharya, "Heterodyne phase locked loops – revisited," *IEEE Trans. Comm.*, vol. COM-25, no. 10, pp. 1164–1170, 1977.

[7] J.P. McGeehan and A. Lymer, "The problem of speech pulling and its implications for the design of phase-locked SSB radio systems," *IEE Proc. F, Commun., Radar & Signal Process.*, vol. 128, part 6, pp. 361–369, 1981.

[8] J.P.H. Sladen, "The performance of phase-locked loops for frequency control in single sideband land mobile radio receivers," *University of Bath PhD Thesis*, chapter 3, 1983.

Linear Transmitter Design for MSAT Terminals

Ross Wilkinson, John MacLeod*, Mark Beach, and Andrew Bateman,
Centre for Communications Research,
University of Bristol,
Queen's Building, University Walk,
Bristol, BS8 1TR, England.
Tel: +44 272 303726, Fax: +44 272 255265

ABSTRACT

One of the factors that will undoubtedly influence the choice of modulation format for mobile satellite, is the availability of cheap, power-efficient, linear amplifiers for mobile terminal equipment operating in the 1.5-1.7GHz band. Transmitter linearity is not easily achieved at these frequencies, although high power (20W) class A/AB devices are becoming available. However, these components are expensive and require careful design to achieve a modest degree of linearity. In this paper an alternative approach to RF power amplifier design for MSAT terminals using readily-available, power-efficient, and cheap class C devices in a feedback amplifier architecture is presented.

INTRODUCTION

Global mobile communication services such as maritime, aeronautical and land mobile are most effectively provided by using satellite transponders. The long range and limited radiated power of the satellites, together with the modest allocation of radio spectrum[1] given to these services necessitates the efficient use of both bandwidth and power, if a service is to be provided for a large user population. The modulation format selected must therefore be tolerant to the restrictive propagation characteristics of the mobile radio channel, if intelligible and reliable communications are to be provided.

Significant advances in both analogue and digital modulation techniques have been made for both this and the closely related field of terrestrial mobile communications. The majority of modulation techniques currently used for analogue and digital radio are in existence not because of their spectral efficiency, but rather because of their compatibility with non-linear RF amplifier technology. Examples are analogue frequency modulation and constant-envelope data modulation formats (MSK, GMSK, TFM, etc.). However, techniques which combine both amplitude and phase modulation, such as analogue linear modulation (single sideband)[2] and M-ary QAM[3], have significant advantages in terms of power drain and spectral efficiency, and thus are well suited to mobile satellite applications[4].

Numerous proposals have been made regarding the choice of modulation format for future generation mobile satellite systems. These include the use of amplitude-companded single sideband with a transparent tone-in-band (ACSSB-TTIB) by both the Australians[5] and Canadians[6], and heavily filtered Quadrature Phase Shift Keying by Inmarsat[7] for their Standard-M service. In order to obtain the necessary adjacent channel performance in these systems, it is essential that a linear channel is employed. This requirement dictates the use of either linear, or linearised RF power amplifiers.

This paper describes a simple technique for the linearisation of high efficiency class C amplifier stages using feedback compensation. New results obtained for a linear transmitter system operating at 1.7GHz are presented here, and compared with results previously obtained at 900MHz[8],[9].

*On sabbatical leave from Footscray Institute of Technology, PO Box 64, Footscray, Victoria, Australia.
Tel: +61 3 6884422, Fax: +61 3 6884806.

LINEAR POWER AMPLIFIERS

The requirements of efficient spectrum utilisation, good power-efficiency, and compatibility with any modulation scheme can only be met by a completely *linear* transmitter. To date, all linear transmitters have utilised devices operating in classes A or AB, where their characteristics are approximately linear. These transmitters, with their poor efficiency, and large heatsinks, have not been attractive in small battery-powered equipment.

The use of a high-gain, narrow-band feedback loop allows amplifiers whose open-loop characteristics are far from linear to be employed[10]. A class C RF amplifier, for instance, would normally only be suited to constant-envelope signals, but the prototype amplifiers here described have achieved excellent performance with a two-tone test signal, the most rigourous test of linearity. The advantages of class C amplifiers over their more linear counterparts are increased efficiency (a most important factor in battery-powered systems), and reduced size (since they require a smaller heatsink). In addition, they do not suffer from the gradual deterioration in performance due to drift in the bias point, etc.

IMPLEMENTATION OF THE CARTESIAN LOOP TRANSMITTER

Essentially, the transmitter is a direct conversion design: the baseband message information, in quadrature components, is modulated directly onto the local oscillator, without the use of any intermediate frequency translation. The quadrature components of the modulation are generated by digital signal processing, and hence the channel filtering, and modulation scheme are controlled in software, and may easily be reconfigured.

Any modulation which can be represented in the form of two quadrature channels, whose bandwidth does not exceed the transmitter's linearised bandwidth, may thus be upconverted and amplified without distortion.

A schematic of the transmitter is shown in figure 1. The I and Q components of the modulation are supplied by the DSP, and modulated onto the LO via a phasing/Weaver method upconverter. A series of amplifiers

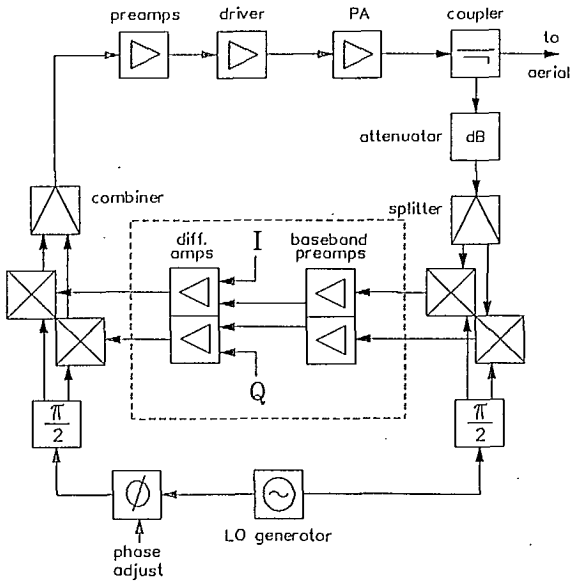


Fig.1 Schematic of the cartesian loop transmitter.

bring the low-level output from the mixers up to the required output power. The amplifiers in this chain are progressively more non-linear, with the final stage being in class C. After the amplifiers, a directional coupler samples the output signal, which is then attenuated and fed into a quadrature demodulator, akin to the upconverter. A pair of identical differential amplifiers, one in each quadrature arm, derive error signals between the required modulation (from the DSP) and the downconverted output. It is these error signals which drive the upconverter. Figure 2 shows an ideal two-tone test signal, and the error signal generated in the feedback loop around a class C amplifier: the distortion in this signal is exactly complementary to that in the amplifiers and upconverter, and thus the output is undistorted. The plot shows the predistortion to be most evident near the minimum of the modulation envelope, where the class C stage is in its most non-linear region.

The feedback loop also includes phase compensation, here shown in the local oscillator driving the upconverter. This is required to cancel the phase shift through the amplifiers, coupler and attenuator, so that the downconverted signals are in phase with their respective upconverted counterparts. If this phase was not equalised, then positive feedback around the loop could occur, resulting in wideband spurious outputs. Since the phase shift through the transmitter will vary with operating frequency, this phase equalisation must be altered as the channel is changed.

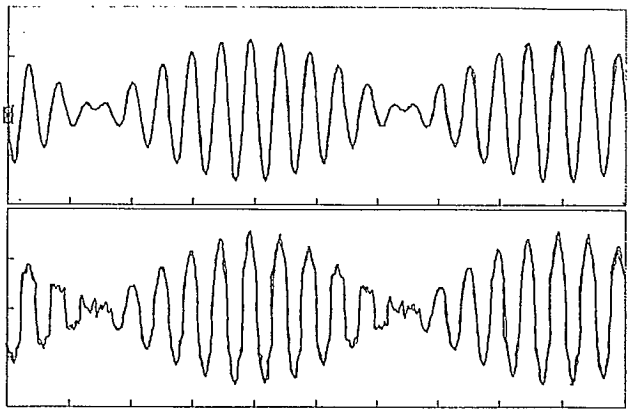


Fig.2 A two-tone test signal and predistorted error signal.

The loop dynamics are controlled by the differential amplifiers: the aim is the greatest suppression of distortion, whilst maintaining overall stability. The gain and bandwidth of the feedback loop must be designed accordingly.

DC offsets in the differential amplifiers and the mixers in the up- and down-converters will result in an unwanted RF output at the local oscillator frequency. This will change with operating frequency, time, temperature, etc, and so steps must be taken to null out this residual DC. The use of a TTIB-based system will relax considerably the need for this suppression, since the unwanted DC will be insignificant compared with the level of the pilot tone.

RESULTS FROM PROTOTYPE SYSTEMS AT 1.7GHZ AND 900MHZ

Two prototype systems have been constructed, both employing diode-ring mixers in the up- and down-converters, and using the same differential amplifier configuration, consisting of a first-order loop with a bandwidth of 10kHz (note that this is effectively 10kHz on either side of the local oscillator, i.e. a 20kHz correction bandwidth).

The 1.7GHz system included an ACR 2001 transistor as the final class C amplifier, and figure 3 shows the results achieved with feedback, compared with the open-loop performance. The tones are spaced at 4.2kHz, and the output power is 400mW PEP. Although the device is capable of more output, initial tests showed that it was not possible to achieve stable closed-loop operation at higher power levels. Examination of the gain/phase characteristics of the amplifier chain over the top 10dB of the output range, figure 4, shows an abrupt change of over 35dB and 120°: much more than can be accommodated by the gain and phase margin of the loop.

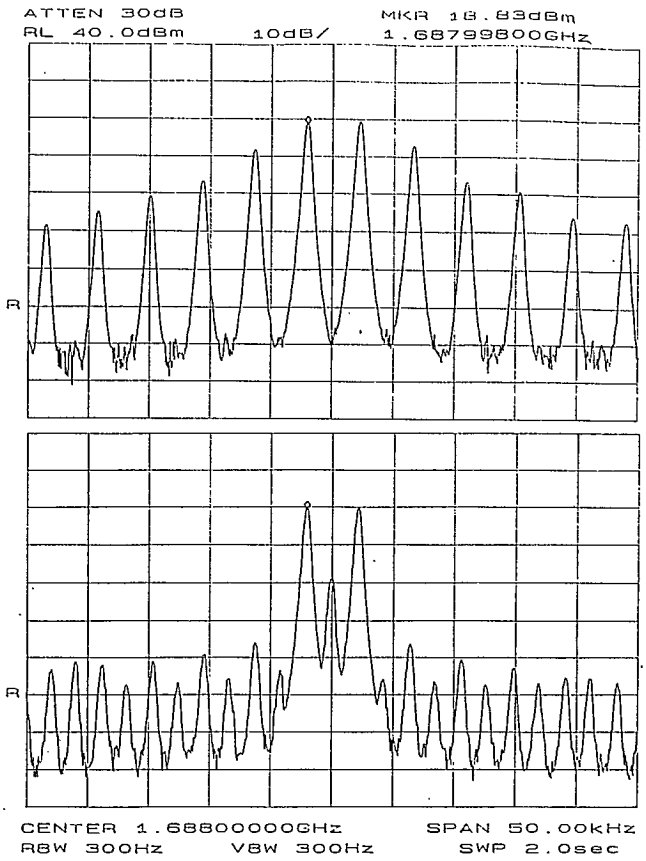


Fig.3 Transmitter for 1.7GHz open- and closed-loop.

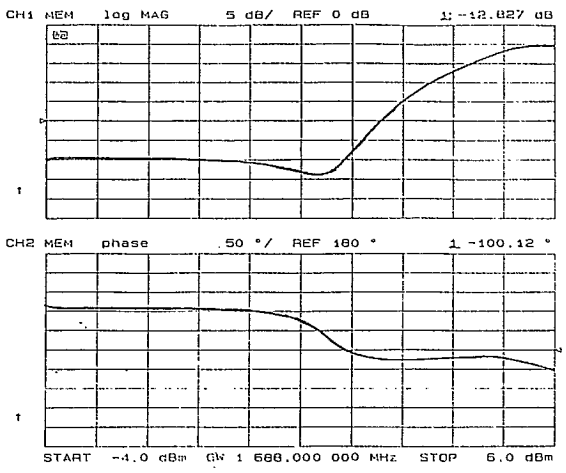


Fig.4 Gain/phase characteristic for 1.7GHz amplifier.

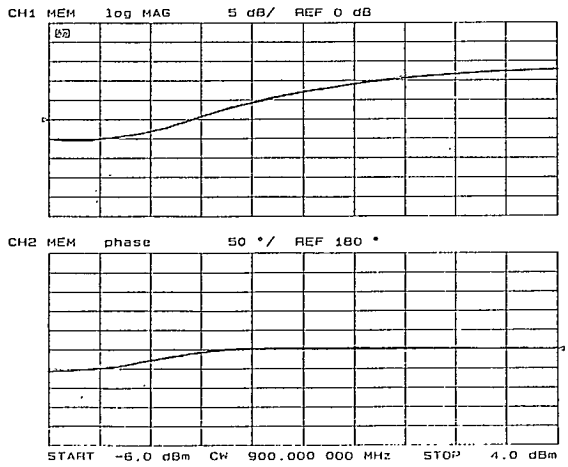


Fig.5 Gain/phase characteristic for 900MHz amplifier.

For comparison, figure 5 shows the same measurement for the 900MHz system: here the change is more gradual, 15dB and 60°, but this is still a class C amplifier, (here an NE080191 transistor). The open- and closed-loop plots at an output of 1W PEP (full power) are shown in figure 6. This illustrates what may be achieved with an efficient linearised class C transmitter.

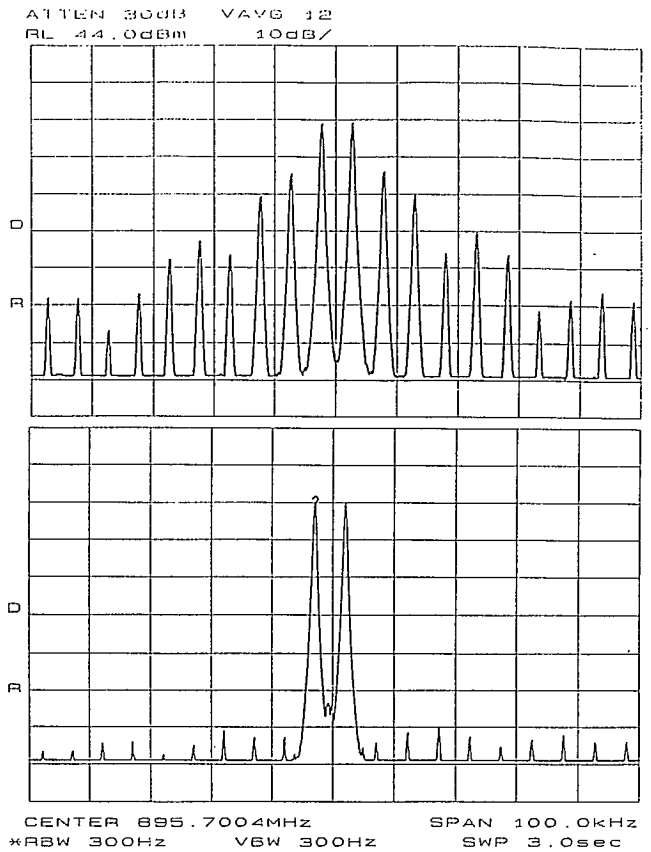


Fig.6 Transmitter for 900MHz open- and closed-loop.

CONCLUSIONS

It has been shown, using the cartesian feedback technique at 900MHz, that excellent intermodulation suppression (>60dB for a two-tone test signal) and good power efficiency (55-65%) can be obtained. Initial results at 1.7GHz show that although the feedback loop performs just as well at this frequency, the amplifier itself requires careful design. The use of a RF non-linear simulation software package would facilitate the design of such an amplifier, and work in this field should soon be underway at Bristol. Some advantage may also be gained by using higher-order feedback loops, to more carefully tailor the gain and phase margin of the system.

Hence, linear power amplifiers for future generation mobile satellite terminal equipment can be designed using a linearised class C architecture, which has better linearity and power efficiency performance then its class A counterpart. Besides mobile satellite applications, another major driving force behind the development of linear transmitter architectures at 1.7GHz is the personal communication networks (PCN's), which is also being addressed at Bristol. In addition, other techniques for linearising amplifiers, offering wider channel bandwidths and power efficiency are also being considered. These include adaptive predistortion and LINC[11].

ACKNOWLEDGEMENT

The authors wish to thank David Bennet for his invaluable help with the fabrication of the up- and down-converters and the loop phase shifter for the 1.7GHz system.

REFERENCES

[1] **M.Goddard**, 'Frequency Allocations for Mobile Satellite Systems and Satellite Navigation,' IEE 4th Int. Conf. on Satellite Systems for Mobile Communications and Navigation, CP294, Oct 1988, pp10-14.

[2] **A.Bateman & J.P.McGeehan**, 'A Survey of Potential Narrowband Modulation Techniques for use in Satellite and Cellular System,' Second Nordic Seminar on Digital Land Mobile Radio Communications, Sweden, 1986.

[3] **K.Feher**, 'Digital Communications,' Prentice-Hall, 1983.

[4] **W.J.Whitmarsh, A.Bateman & J.P.McGeehan**, 'Analogue and Digital Linear Modulation Techniques for Mobile Satellite,' Second International Mobile Satellite Conference, Ottawa, Canada, June 1990.

[5] 'Trials Report on the Subjective Performance of Modulation Techniques for Mobile Satellite,' Australian Telecom Internal Report, 1989.

[6] **J.S.Butterworth**, 'Satellite Communications Experiment for the Ontario Air Ambulance Service,' ibid [1], pp83-86.

[7] 'On the Subjective Voice Quality Test of ACSSB Modulators and Digital Vocoders,' Inmarsat Internal Document, 1989.

[8] **A.Bateman & R.J.Wilkinson**, 'Linearisation of Class C Amplifiers using Cartesian Feedback,' IEEE Workshop on Mobile and Cordless Telephone Commuications, London, Sept 1989, pp 62-65.

[9] **A.Bateman, D.M.Haines & R.J.Wilkinson**, 'Direct Conversion Linear Transceiver Design,' IEE 5th International Conf. Mobile Radio & Personal Communications, Warwick, Dec. 1989, pp 53-56.

[10] **V.Petrovic**, 'Reduction of Spurious Emission from Radio Transmission by means of Modulation Feedback,' IEE Conf. Radio Spectrum Conservation Techniques, Sept. 1983, pp 44-49.

[11] **A.Bateman, R.J.Wilkinson & J.D.Marvill**, 'The Application of DSP to Transmitter Linearisation,' 8th Euro. Conf. Electrotechnics, Stockholm, June 1988.

Integrated DSP/RF Design for an MSAT Transmitter

S.P.Stapleton, J.K. Cavers
Communications Science Laboratory
Simon Fraser University
Burnaby, British Columbia
V5A 1S6, Canada
Phone: 604-291-4371
Fax: 604-291-4951

ABSTRACT

In mobile radio systems, the relatively inefficient use of the spectrum by existing FM modulation techniques is limiting the available channels. Linear modulation methods are thought to provide the necessary push into more efficient useage of the spectrum. The limitation, to date, of the linear modulation techniques is the fact that the nonlinear power amplifier tends to spread the spectrum and thus offset any spectrum efficiency advantage. Linearization techniques are presently being investigated, in the literature, in order to recover some of the lost spectrum, due to spreading. To this end, DSP techniques are thought to provide real time predistortion methods. Some of the complex modulation blocks, normally confined to analog circuitry, are being implemented using DSP circuitry. This paper will address the linearization techniques, for MSAT tranceivers, that use an integrated DSP/RF design.

INTRODUCTION

To date, mobile radio systems use FM modulation schemes like MSK,TFM, and others. Even though FM modulation is inefficient in spectrum useage, it is still widely used because of its constant envelope property. The constant envelope property provides the use of power efficient nonlinear amplifiers, at the same time as achieving low out-of-band radiation (-60 dB to -70 dB).

The next generation of digital mobile radio systems will use Quadrature Amplitude Modulation (QAM) patterns as a means of increasing the spectrum efficiency beyond that of FM modulation schemes. However, since QAM presents a variant envelope, it will be necessary to use more linear less-efficient power amplifiers. Linear amplifiers have the disadvantage of poor power efficiency to the extent that they cannot satisfy the required -60 to -70 dB adjacent channel interference power level. While, this severe constraint on adjacent channel interference may be relaxed for certain systems such as cellular mobile telephone systems, it remains as a requirement in general mobile radio systems.

The extremely high degrees of transmitter linearity required for operation with an acceptable overall system bit-error-rate (BER) performance is difficult to obtain. Because of the large number of possible states in the transmitted phase-plane, as well as the high peak-to-average signal power inherent in QAM modulations, any degradations due to transmitter nonlinearity can become quite severe. Transmitter intermodulation distortion (IMD), which can be observed on a QAM phase-plane as a combination of AM-to-AM and AM-to-PM, moves the transmitted states closer to decision thresholds, degrading the operating margins and hence attainable system BER. In addition to the increase in BER the distortion will cause spectrum spreading, which will place a limit on the maximum system data rate allowable to maintain within the FCC frequency mask limitations. In order to achieve the optimum tradeoff between spectrum spreading and efficiency, techniques for linearizing the power amplifier must be adopted.

Linearization techniques can be categorized into one of the following:
1) feed-forward, 2) feedback, 3) LINC and 4)

predistortion. These techniques have conventionally been implemented using analog circuitry. However, the advent of Digital Signal Processing (DSP) devices permits complicated operations to be handled in a simple manner. This paper demonstrates how DSP devices can be integrated into mobile radios.

DSP/RF INTEGRATED ARCHITECTURE

We have developed a digital signal processing (DSP) technique to apply compensating predistortion to the signal, at baseband, before modulation (Figure 1). The complex vector predistortion will track the inverse function of the MSAT power amplifier's AM-to-AM and AM-to-PM distortion. A recent paper[1] discussed this method and demonstrated significant improvement. Central to such baseband approaches is a quadrature modulator to translate the complex signals to an IF. The level of precision required in meeting emmission constraints for mobile applications makes the use of analog technology for quadrature modulation quite difficult. The problems, which include amplitude and phase imbalance, and component drift, can be circumvented by digital implementations.

The benefits to the designer of using the digital quadrature modulator are as follows:

1) It is a precision digital solution to a common design problem. Unlike the analog equivalents, amplitude matching and phase separation are essentially perfect, there is no residual carrier or DC offset, and there is no performance drift. The digital quadrature modulator requires one digital to analog converter (DAC), whereas an analog quadrature modulator driven by a DSP chip would require two DAC's.

2) It simplifies the RF analog design. Upconversion from a 480 KHz IF, for example, simply requires the second IF stage to remove images 960 KHz away, which is a relatively easy filter implementation.

3) The digital quadrature modulator, together with the DSP chip driving it, can act as an universal modulator, for a multimode tranceiver. FM, AM, SSB, ACSB, TTIB, MSK, QPSK, 16QAM, etc, can all be resident in the DSP chip's ROM. and selected as appropriate during operation.

Some of the potential problems with using digital quadrature modulators in mobile transmitter architectures are: the significant DC power consumption; the delay introduced by internal filtering operations on both the digital quadrature modulator and DSP chips; and the reduced spectral purity.

DIGITAL QUADRATURE MODULATOR

General Description

Figure 2 is a functional view of an ideal quadrature modulator (quad mod). The complex baseband signal $v_2(t)$, represented by real and imaginary components $v_{2I}(t)$ and $v_{2Q}(t)$, is modulated to form the real bandpass output $\tilde{v}_3(t)$ with complex envelope $v_3(t)$, at a center frequency f_{c3} :

$$\tilde{v}_3(t) = v_{3I}(t) \cdot \cos(2\pi f_{c3} t) - v_{3Q}(t) \cdot \sin(2\pi f_{c3} t)$$
$$= Re[v_3(t) \cdot \exp(j2\pi f_{c3} t)] \qquad (1)$$

Ideally, the complex envelope $v_3(t)$ equals the input $v_2(t)$. The frequency domain sketches show the input $V_2(f)$ as asymmetric to emphasize the fact that the input is complex. Its bandwidth is W_2 and the corresponding RF bandwidth is $2W_2$. Figure 3 illustrates the operation of the digital quad mod. The inputs are digital, and enter at a relatively low sampling rate f_{s2} . The output is analog, from a DAC driven at a sampling rate f_{s3} , which is precisely 4 times the center frequency f_{c3} . The cos/sin carriers in (1) degenerate to:

cseq: 1,0,-1,0,1,0,-1,0,1,0,-1,0,1 ...
sseq: 0,1,0,-1,0,1,0,-1,0,1,0,-1,0 ... (2)

which eliminates both trigonometric generation and multiplication. Since f_{s3} is significantly greater than f_{s2} , we need more samples of $v_2(t)$ than are available from the input. The intermediate samples are created by interpolation (i.e. lowpass filtering). The periodic zero in (2) means that the I and Q channels supply the output alternately, which cuts computation in half. In a further simplification, we can combine the cseq and sseq sequences with the interpolation coefficients.

One consequence of interpolation is delay,

which increases with the length of the interpolator. The delay becomes important when the quad mod is embedded in a feedback loop, as in the case of adaptive amplifier linearization.

Interpolation and Computation

Performing interpolations at the rate f_{s3} is less demanding than it appears, as an example will make clear. Suppose that a 16 point interpolating lowpass filter is sufficient. Figure 4 shows the impulse response and a typical output sequence generated by convolution if the step up ratio $f_{s3}/f_{s2} = 8$. Only two input samples affect each output sample. Since I and Q inputs alternate, only 2 multiply/adds are required per output sample.

More generally, for an N-point lowpass filter, we require:

Nf_{s2}/f_{s3} multiply-adds per output sample, and
Nf_{s2} multiply-adds per second. (3)

For a fixed output sampling rate, therefore, reducing f_{s2} (provided it is above the Nyquist rate $2W_2$) allows a proportional increase in the number of points.

Interpolator Design

This section develops frequency domain design guidelines for the quad mod interpolator. Although the quad mod algorithm accommodates a wide range of sampling and center frequencies, we will assume specific values, in order to make the discussion and illustration simpler:

$W_2 = 17.5$ kHz (5 kHz MSAT channels, 7th order predistortion)
$f_{s2} = 480$ kHz (to simplify the illustration only)
$f_{s3} = 1920$ kHz, $f_{c3} = 480$ kHz (4)

Figure 5a shows that the input spectrum consists of images with a repetition interval of 480 kHz. The interpolator suppresses images other than the one centered at 0 Hz (Figure 5b), thereby changing the effective repetition interval to 1920 kHz. The greatest attenuation is at the folding frequency of 960 kHz. The complex carrier (1) is also sampled at 1920 kHz, and therefore has the spectrum shown in Figure 5c. Figure 5d is the result of convolving the spectra of the sampled input and the sampled complex carrier, to represent multiplication of the corresponding time domain signals. Finally, the operation of taking the real part of the time domain signal in (1) has a spectral equivalent of superimposing the conjugate reflection, as shown in Figure 5e.

The desired components in Figure 5e are at +/- 480 kHz. The rest may require further suppression by an analog bandpass filter, depending on the application. However, the frequency reversed image superimposed on the desired one at 480 kHz cannot be suppressed by filtering. It derives originally from the input image at 960 kHz.

The most important design objective for the lowpass filter, therefore, is attenuation of at least 60 dB in a 35 kHz stop band centered at 960 kHz. A secondary objective is to maximize the attenuation of other input images, so that the design of any following analog filters is simplified. The attenuation at other frequencies can take any convenient value, since there is no signal component there to require attenuation.

To continue the example, assume that the DSP hardware is capable of 11.52 million multiply-adds per second (6 multiply/adds per output sample at an output sampling rate f_{s3} = 1920 kHz). Since the step-up ratio is 1920/480 = 4, we can have N = 24 points in the lowpass filter. Figure 6 shows a multiband design using the Parks- McClellan algorithm with only 16 points, which places 35 kHz stopbands at 480 kHz and 960 kHz. Clearly, the interpolation more than meets the requirements. An analog bandpass is necessary only to suppress harmonics at multiples of 960 kHz from channel center, which is not difficult.

Lower Input Sampling Rates

The 480 kHz input sampling rate used in the example above is unnecessarily high for MSAT applications. However, it can be reduced with no change to the output center frequency or computational load. Suppose the input sampling rate is $f_{s2} = 60$ kHz,

sufficient for 7th order predistorted MSAT channels. The images to be removed are 35 kHz wide, centered at multiples of 60 kHz. If we use the same DSP hardware as above, then from (3) the filter length can be N = 192, without change to the computation rate. Figure 7 shows a Kaiser window design which suppresses even the closest images by almost 50 dB.

Number of Bits

The DAC adds quantization noise. Its level relative to the desired signal is easily computed. Assume the peak DAC voltages are +/- 1. Then the quantization noise power per sample is:

$$\sigma_q^2 = (2 \cdot 2^{-n})^2/12 \qquad (5)$$

where n is the number of DAC bits. Assume a typical data signal with a peak to average power ratio of 6 dB. Then the signal power per sample is

$$\sigma_v^2 = 1/4 \qquad (6)$$

The ratio of signal and quantization noise power spectral densities (PSDs) is more important than per-sample variances. Assume for simplicity that the signal spectrum is approximately rectangular with bandwidth W_2, and that successive quantization noise samples are independent (i.e. white). Then the ratio of PSDs of quantization noise and signal is given by:

$$\frac{PSD_q}{PSD_v} = \frac{\sigma_q^2}{f_{s3}} \cdot \frac{W_2}{\sigma_v^2} = \frac{W_2}{f_{s3}} \cdot \frac{4}{3} \cdot 2^{-2n} \qquad (7)$$

For $W_2 = 5$ kHz, and $f_{s3} = 1920$ kHz, we can achieve a noise floor of -60 dBc with only 6 DAC bits. However, this figure is valid only for random data. Short period repetitive sequences produce quantization noise line spectra. In this case, 10 to 12 bit DACs are advisable.

PREDISTORTION

The general transmitter configuration is shown in Figure 8. The modem, predistorter, quad mod. and quad demod. are implemented digitally at complex baseband. The RF power amplifier, and up and down converters, which translate between IF and RF, are analog. Since the power amp exhibits both AM-to-AM and AM-to-PM distortion, any modulation with nonconstant envelope will generate adjacent channel interference because of intermodulation products. The function of the predistorter is to apply an inverse nonlinearity, so that the power amplifier output is free of intermodulation products.

The technology under development is a digitally implemented predistorter which operates at complex baseband[2]. Currently, the predistorter can operate at a sampling rate of 240 KHz, which is sufficient for 7th order nonlinearities generated from a standard 25KHz mobile channel.

In a nonadaptive configuration, there is no feedback path. The predistorter coefficients are determined for a given amplifier at the factory and are fixed. Aging and temperature changes will cause the amplifier characteristics to drift, of course, so the predistorter and amplifier are no longer perfectly matched. It is possible, though, that the combination will still meet the relatively lenient MSAT specifications on out-of-band power, and it is much cheaper than an adaptive configuration.

Adaptivity is achieved by feeding back the demodulated amplifier output and comparing it to the desired output, then updating the predistorter coefficients. Note that adaptation can be very slow if all it has to track is drift in the amplifier characteristics; however, a channel switch can introduce larger changes, and readaptation must be rapid to avoid dropouts. The requirement of a receiver operating in parallel with the transmitter means that it may be an expensive technology.

CONCLUSIONS

This paper has described the integration of DSP hardware into a mobile communications transmitter. We have addressed the significance of implementing the quadrature modulator digitally. The design considerations for the digital quadrature modulator have been discussed. In order to improve the spectral efficiency of mobile radio systems, QAM patterns will be required. However, the variant

envelope present in QAM will necessitate the use of predistortion techniques. DSP chips can be used for complex baseband predistortion. Whether the system is dynamic or static will depend on the system FCC limitations.

ACKNOWLEDGEMENTS

This research was supported by the British Columbia Science and Technology Development Fund. We are pleased to acknowledge assistance from J. Wang, N. Fried, K. Douglas and C. Anderson.

REFERENCES

1 **Bateman, A.,Haines, D.M., and Wilkinson, R.J.** 1988. Linear Tranceiver Architectures. Proc. IEEE Vehicular Tech. Conf., pp. 478-484.

2 **Cavers, J.K.** 1990. A Linearizing Predistorter With Fast Adaptation. Proc. IEEE Vehicular Tech. Conf.

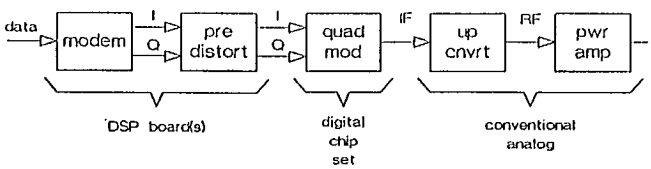


Fig. 1. MSAT Terminal with Static Predistortion

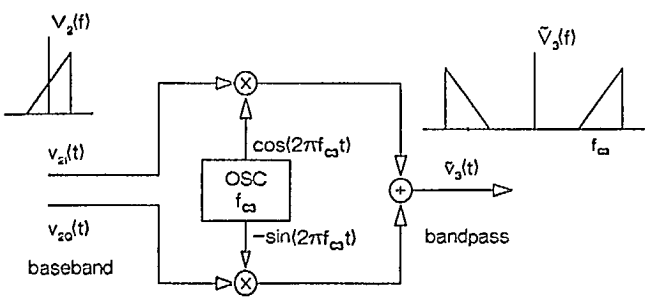


Fig. 2. Analog Quadrature Modulator

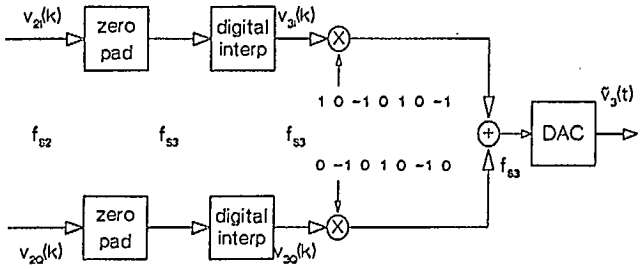


Fig. 3. Digital Quadrature Modulator

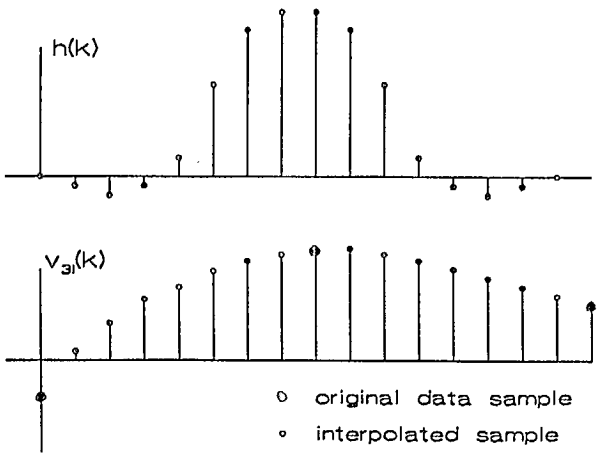


Fig. 4. Impulse Response for a 16 Point Interpolating Lowpass Filter

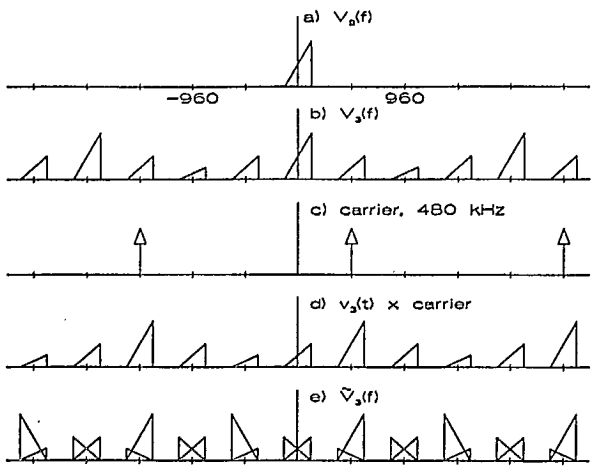


Fig. 5. Frequency Domain Spectrum for the Digital Quad. Mod.

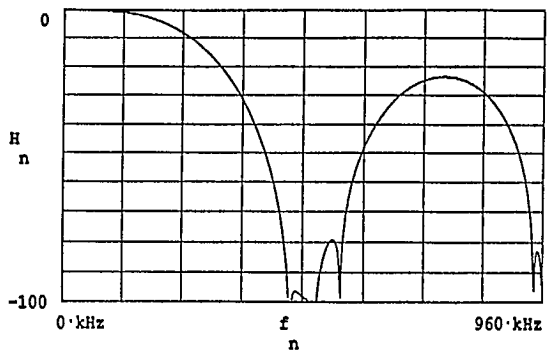


Fig. 6. 16-Point Filter for 480 KHz Samples

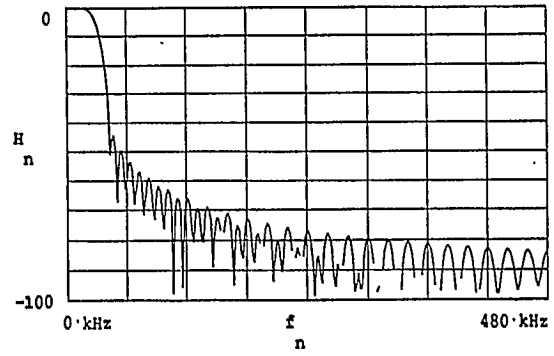


Fig. 7. 192-Point Filter for 60 KHz Samples

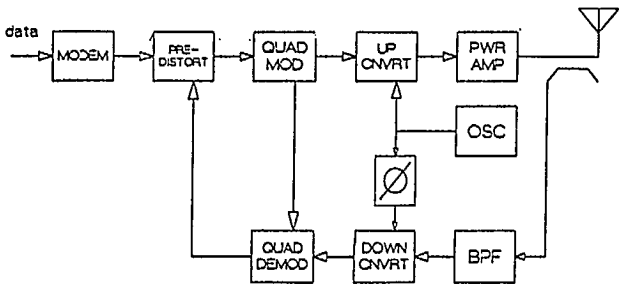


Fig. 8. Transmitter Functional Block Diagram

Session 8
Network Architecture and Control

Session Chairman - *William W. Wu,* Stanford Telecom, USA
Session Organizer - *Tsun-Yee Yan,* JPL

MSAT Network Architecture

N.G. Davies and B. Skerry
Telesat Mobile Inc.
P.O. Box 7800 Ottawa, Ontario
K1L 8E4, Canada
Phone: (613) 746-5601
Fax: (613) 746-2277

ABSTRACT

Telesat Mobile Inc. will provide Mobile Satellite Services (MSS) in Canada starting in 1993 using a geostationary MSAT satellite. The communications system, which is being developed in close cooperation with the American Mobile Satellite Corporation, will support mobile voice and data services using circuit switched and packet switched facilities with interconnection to the public switched telephone network and private networks.

Control of the satellite network will reside in a Network Control System (NCS) which is being designed to be extremely flexible to provide for the operation of the system initially with one multi-beam satellite, but with the capability to add additional satellites which may have other beam configurations. The architecture of the NCS is described. The signalling system must be capable of supporting the protocols for the assignment of circuits for mobile public telephone and private network calls as well as identifying packet data networks. The structure of a straw-man signalling system is discussed.

INTRODUCTION

Telesat Mobile Inc. (TMI), which is authorized to provide mobile satellite services (MSS) in Canada, will launch a dedicated MSAT satellite in late 1993 to provide mobile telephone, mobile radio and mobile data services to customers on the move in any part of Canada. The American Mobile Satellite Corp. (AMSC), the licensed MSS operator in the U.S.A., will procure a similar satellite. The satellites of the two organizations will provide mutual back-up. TMI is cooperating closely with AMSC to define and to procure the necessary satellites and ground segment equipment. Contract negotiations are under way with satellite suppliers with the objective to enter into a contract in the second quarter of 1990. The design and procurement of ground segment equipment will proceed in parallel. The satellites and the ground system are to be ready for full operation by 1994.

This paper describes the baseline planning of the infrastructure for the ground segment. Through close cooperation with AMSC, it is planned that a common communications infrastructure be designed, such that customers will be able use their mobile terminals anywhere in North America and will have a choice of suppliers for mobile terminals.

MSAT SERVICE REQUIREMENTS

TMI estimates the potential demand for MSS at 300,000 to 450,000 mobile terminals in Canada by the year 2000. Studies of the requirements for the mobile data services indicate that the requirements for mobile data terminals could reach 50% to 60% of the total demand. Assumptions regarding this market penetration rate indicate that TMI's system should support 75,000 to 130,000 mobile voice terminals by the year 2000. In comparison, it is anticipated that the market in the U.S.A. could be three to ten times greater.

The penetration of the market will depend critically upon the cost of the mobile terminals and the charge for air time, which should not be too much greater than the equivalent costs and charges for cellular services. The MSAT

communications network must be designed to provide as much satellite circuit capacity as possible with a given satellite using mobile terminals that are cheap, very reliable and are appropriate to the land, sea and air mobile environments. The network design should permit many services to be provided through mobile transceivers which are manufactured to a common standard.

To aid the design of the communications network, the service requirements in Canada have been aggregated into the following four categories.

Mobile Radio Service (MRS) - a circuit switched voice and data private network service.

Mobile Telephone Service (MTS) - a circuit switched voice and data service with interconnection to the public switched telephone network.

Mobile Data Service (MDS) - a packet switched data service with interconnection to private and public data networks.

Aeronautical Services - services which are compatible with international standards for mobile satellite communications with aircraft.

The first three categories of service will support land and marine mobile services as well as aeronautical services to private and small commercial aircraft. A separate category has been defined for the provision of aeronautical mobile satellite services to large commercial aircraft which will most likely adhere to international standards which are defined by the International Civil Aviation Organization (ICAO). Other categories may be added as experience is gained.

THE BASELINE SYSTEM

The MSAT system will use the L-band frequencies 1530.0 - 1559.0 MHz (downlink) and 1631.5 - 1660.5 MHz (uplink) for communication with mobile earth terminals (METs). The service area is the whole of North America and the islands of Puerto Rico and Hawaii with an option for coverage of Mexico and the Canadian and U.S.A. controlled international flight information regions. This coverage has been defined for the satellites of both systems in order that in-orbit back-up of each system can be provided. The baseline design for the first generation system [1] has envisaged five to nine beams with re-use of the limited frequency spectrum up to 1.3 times. Ku-band frequencies in the 11/13 GHz allotment bands, in a single continental beam, will be used for feeder-links to fixed feeder-link earth stations (FESs).

The frequency translating satellite communications transponder will relay signals between Ku-band feeder-link earth stations and L-band mobile terminals. The satellite transponders will not constrain the choice of communications signals. The baseline system design is for FDMA/SCPC multiple access narrow-bandwidth SCPC channels, nominally 5 kHz in bandwidth. The conceptual design of the satellite transponder has considerable flexibility to reallocate power and bandwidth to the different beams or service areas of the system.

A wide variety of mobile terminal types will be able to operate with the MSAT satellite. While the detailed parameters of many of these terminals are not yet firmly established, the parameters for voice communications terminals have been extensively explored. It is anticipated that these terminals will have an antenna gain of 10 to 13 dBic (circularly polarized) and a corresponding G/T of -15 to -12 dB/K. MET EIRP will be below 16 dBW.

Both analog and digital forms of speech modulation are being considered. Amplitude companded single sideband (ACSSB) modulation is assessed to be a very rugged form of modulation for use under the severe multipath and blocking anticipated in the land mobile environment. 4.8 kbps digital speech with trellis coded modulation is also becoming very attractive. The latter form of modulation may permit satisfactory operation at lower values of C/No, permitting more channels to be assigned through a given satellite, but has a sharper threshold and hence may not be so satisfactory in a highly shadowed environment.

The range of communications capacity of the baseline L-band satellite transponder and the two types of speech modulation is some 1,000 to 2,000 assignable voice channels which can support 100,000 to 200,000 METs for voice or the equivalent in data traffic [2].

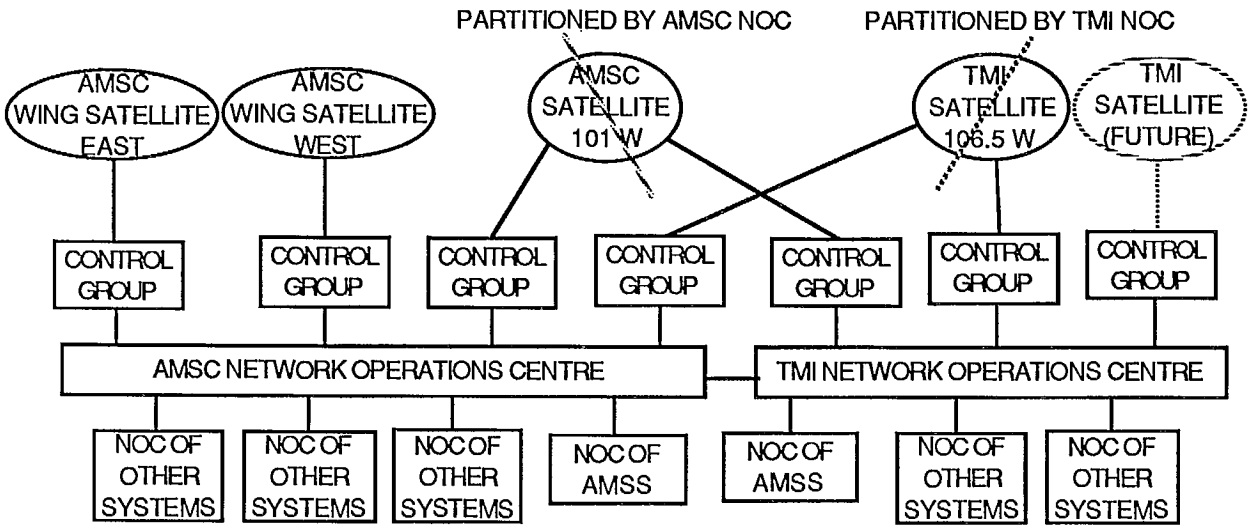


Figure 1. General Architecture of a System to Provide MSS in North America

The bandwidth and power available in a satellite transponder will be subdivided into one-way channels of various bandwidths and power levels. Generally, forward and return link channels will be paired to support two-way communication. A symmetrical channel pair forms a circuit. Pairings may be asymmetric for provisioning of asymmetric services such as packet switched data and to support signalling. Circuits may be used for DEMAND PERIOD or for FULL PERIOD circuit switched service. DEMAND PERIOD circuits are used on a call-by-call, as needed, basis; FULL PERIOD circuits are devoted to a long term particular use.

NETWORK ARCHITECTURE

The general architecture of the systems to provide MSS in North America is illustrated in Figure 1. The satellites owned respectively by TMI and AMSC are located at several different orbital positions. Communications services provided through any of these satellites, or a portion of a satellite, is managed within a Control Group, to be described in more detail later. For business reasons or for network restoral in the event of satellite failure, capacity on the two central satellites may be allocated to both a TMI and an AMSC Control Group.

A TMI and a AMSC Network Operations Centre will manage and control all the resources of the respective systems. A NOC will coordinate the frequencies to be used in each beam and will allocate blocks of satellite bandwidth and power within beams as satellite circuits to a Control Group. The NOC also allocates blocks of satellite bandwidth and power to customers such as the commercial aviation community, who may wish to operate their own mobile satellite system. The NOC receives and acts upon requests for pre-emption of frequencies if required for aeronautical safety services. In responding to the flight safety requirements, the NOC may need to reconfigure the frequency bands to be used by each satellite in each beam and so advise the GCs.

The NOCs will carry out the administrative functions associated with management of the total MSS system. This includes the registering of METs with their attributes, setting up conditions so that properly authorized METs may access the system, the recording of system usage for billing purposes and the accumulation of network performance records to aid in long-term system planning.

This architecture clearly has provision to accommodate growth through the addition of new satellites (which may have a different transponder and beam configuration), and the addition of Control Groups. It also provides needed flexibility in the allocation of resources to different Control Groups.

Within this overall architecture, the provision of DEMAND PERIOD circuits to and from METs is managed by a Network Control System (NCS). The structure of the NCS is shown diagrammatically in Figure 2. The elements of a NCS are a Network Operations Centre (NOC), a Network Control Centre (NCC) and the portions of the FESs and METs that are associated with the signalling channels through which the assignment of circuits and other network control functions are exercised. A Group controller (GC) within the NCC assigns satellite circuits to meet customer requirements. All interfaces within the NCS will be in accordance with a specified network standard. The operator interface at the MET and the terrestrial network interface at the FES will not be standardized in order to maintain flexibility to meet the needs of specific end-user customers or service provider organizations.

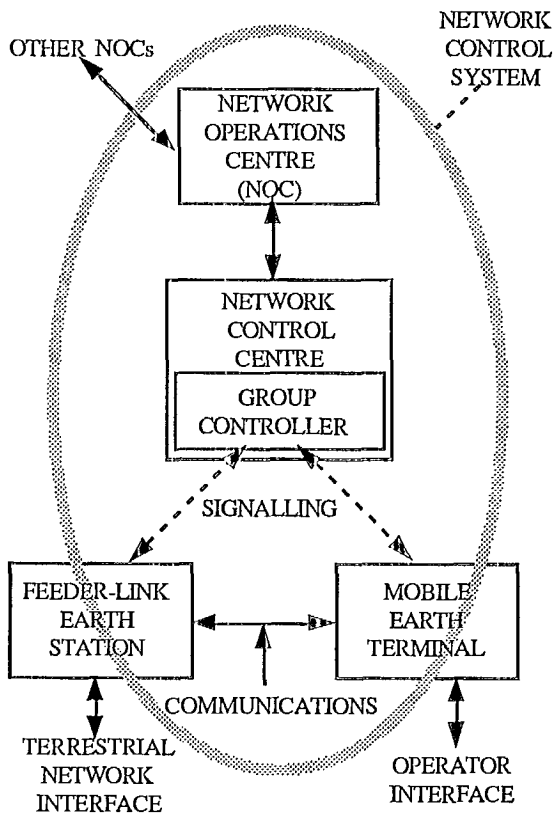


Figure 2. Architecture of the NCS

Because bandwidth and power are valuable and limited, the available satellite circuits are formed into pools large enough to ensure high circuit usage efficiency. The various customer networks within the NCS will use circuits from circuit pools on an as-needed basis.

METs located within the various beams of the system access the GC through L-band signalling channels to request a satellite circuit for communication involving requirements for an interconnection through a FES to the PSTN or to a private network. The necessary instructions from a GC to a FES to set up a communications channel with a MET are provided through Ku-band signalling channels. Calls between METs, which are not expected to be a large part of the total service, are established by double hopping through a FES which is suitably equipped to support such calls.

Two examples of virtual networks will illustrate the application of these features of the NCS. For the first example, a single FES serves a group of METs to form a private network. Call connections are made between the METs and a PBX attached to the FES. The FES is provisioned with several voice channel units of a single type. The interface between the PBX and channel units provides the necessary protocol conversion. Virtual network managers within the GC set up the pre-assigned routing option for this closed user group. The second example involves a virtual network configured for providing PSTN access by subscribing METs. FESs are strategically located throughout the country and interconnected to the PSTN. Each FES is equipped with a number of channel units and PSTN connections based on traffic engineering forecasts and the routing method selected. A virtual network manager within the GC sets up call routing through a FES which is closest to the called party in the PSTN. Appropriate routing is also provided for calls to METs originating from the PSTN.

The network architecture provides for the incorporation of packet switched services, as a separate network, which will likely operate using FULL PERIOD channels administered by a Data Hub [3].

SIGNALLING

The signalling system provides a method for mobile terminals to access the NCS for the exchange of channel assignment and network

management information among all the elements of the NCS. The signalling system must serve the anticipated number of mobile terminals and provide the required response time for setting up calls. One or more signalling channels will be needed in each beam of each satellite.

The signalling system must provide for positive and continuous control of all METs in the system. For this reason and because of the need for a continuous signal for MET antenna tracking and the benefits of a common frequency reference, consideration is being given to requiring all METs to receive a signalling channel at all times, even when a call is in progress.

The strawman concept for the signalling system described here supports the circuit switched services for MSAT. The packet switched data services, the MDS, operate on separate channels. This signalling will be used to set up calls between an FES and a MET. Call take-down will be managed by the FES via some form of signalling on the traffic channel with appropriate notification to the GC.

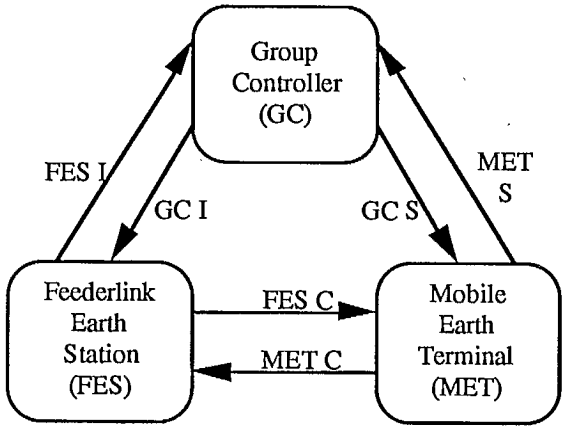


Figure 3. Signalling Channel Configuration

Figure 3 shows the channel configuration between a GC, an FES, and a MET. Six different channel types have been defined as shown in Table 1 [4].

The interstation signalling channels (GC-I and FES-I) will be a simplified form of many existing VSAT systems and will not be described further here.

The communications channels (FES-C and MET-C) are demand-assigned circuit switched channels. The type of traffic to be carried on the channel will be negotiated during the call set-up and may include the type of modulation to be used. Thus the same signalling infrastructure will be able to set up calls for ACSSB, 4.8 kbps CELP, stream data, etc.

Channel	Description	Characteristics
GC-I	Interstation sig. ch. from GC to an FES	TDM
FES-I	Interstation sig. ch. from an FES to GC	TDMA
GC-S	Outbound sig. ch. from GC to a MET	TDM
MET-S	Inbound random access sig. ch. from a MET to the GC	Sloppy Slotted Aloha
FES-C	Outbound com. ch. from an FES to a MET	Traffic
MET-C	Inbound com. ch. from a MET to an FES	Traffic

Table 1. Channel Descriptions

Due to the nature of the signalling channels, it is necessary to have a fixed packet size for the exchange of information. These packets are referred to as signalling units (SU) and are 96 bits long. This size was deemed to be adequate for most applications and is also the size that other MSS operators have chosen [5,6].

The signalling channels may use A-BPSK transmitted at 2400 sps. Rate 1/2 coding may be used thus giving an information rate of 1200 bps. The use of higher rate block codes is being investigated, especially for the inbound channel (MET-S) in order to improve throughput.

The outbound GS-S channel uses time division multiplexing and channel could be interleaved over a one second frame. Each frame could handle 12 SUs. Periodic bulletin boards would be transmitted on this channel in order to relay vital network information (i.e. congestion information, alternate signalling channels, etc.).

The inbound signalling channel (MET-S) could use Sloppy Slotted Aloha (SSA) [7]. This scheme permits the use of small guard times.

However METs do not necessarily have to have correct timing on their first access. The MET-S channel will be able to support 10 slots per second.

In addition to the SSA approach, various reservation schemes are being investigated which may lead to greater throughput. The ratio of inbound to outbound channels will depend upon which reservation scheme is selected.

The connect time is defined as the time from which the MET or FES initiates a call to the time an end-to-end connection is established between a MET and FES. The performance objectives are as follows:

	Connect Time (seconds)			
Performance Objective	Degree of Shadowing None	Light	Moderate	Heavy
Mean	2.5	3.0	4.0	5.0
90%	5.0	6.0	8.0	10.0

Given one GC-S channel as described above and as many inbound channels as are necessary, it is desired to estimate how many users could be supported. The following assumptions are made: three outbound packets needed to set up a call; an average busy hour usage of 0.01 erlangs per user; and an average call duration of 60 seconds.

Working with these averages, and ignoring queuing problems, this leads to one GC-S channel being able to support or 240 calls in progress (approximately 24,000 users). This number would be lower in actual practice, so that the queues could be minimized. The number of inbound channels required per outbound channel will depend upon the reservation scheme adopted, but will be in the range of one to five.

CONCLUSION

This paper has described the concept for the ground segment infrastructure that will be needed to support the mobile voice and data services that will be provided by MSAT. The design of the ground segment will be developed further during 1990 and procurement of hardware will commence at the beginning of 1991.

The design process will be pursued in cooperation with the AMSC with the intention that systems to be operated by the two companies should provide a fully compatible service.

The detailed design and implementation of the systems will be a major and exciting challenge.

REFERENCES

1. **Bertenyi, E.** The MSAT satellite. In *Proceedings of the 40th Congress of the International Astronautical Federation*, Spain 9-13 October 1989.

2. **Davies, N.G. and Roscoe, O.S.** The provision of mobile satellite services in Canada. To be published in *Canadian Journal of Electrical and Computer Engineering*, 1990.

3. **Kerr, R.W. and Skerry, B.** An Architecture for the MSAT Mobile Data System. To be published in these proceedings *Second International Mobile Satellite Conference*. June 1990.

4. **CCIR.** A Network Control and Signalling System for the North American MSS. *CCIR Report 8/531 Annex 5*. 1989.

5. **Aussat.** *Mobilesat System Description.* Aussat Mobilesat, September 1989.

6. **Inmarsat.** *Standard-M System Definition Manual.* Inmarsat, September 1989.

7. **Crozier, S.** Sloppy Slotted Aloha. To be published in these proceedings *Second International Mobile Satellite Conference*. June 1990.

The AMSC Network Control System

W. B. Garner
American Mobile Satellite Corporation
1233 20th Street, N.W., Suite 301
Washington, DC 20036
Phone: 202-331-5858
FAX: 202-331-5861

AMSC will be constructing, launching and operating a satellite system in order to provide mobile satellite services to the United States. AMSC will build, own and operate a Network Control System for managing the communications usage of the satellites, and to control circuit switched access between mobile earth terminals and feeder-link earth stations. This paper provides an overview of the major Network Control System functional and performance requirements, the control system physical architecture and the logical architecture.

FUNCTIONAL REQUIREMENTS

The Network Control System (NCS) will provide centralized control, management and administration of all services utilizing the AMSC satellites. From the NCS perspective, satellite use is divided into two major categories, Full Period and Demand Period. Full Period services are those services which make use of dedicated satellite bandwidth and power that is administered by a control system that is essentially independent of the NCS. Examples are aviation safety service networks that operate to FAA standards, packet data networks that do not make use of circuit switched type access, or networks that are operated by independent service providers. Demand Period services are those services which use satellite power and bandwidth on demand and which are directly controlled by the NCS. Demand Period services are provided through circuit switched, full duplex, two-point circuits set up between Mobile Earth Terminals (METs) and Feeder-link Earth Stations (FESs) by the NCS. All satellite use, whether Demand Period or Full Period, will be under the ultimate control of the NCS.

Demand Period Services

Groups of METs and FESs using Demand Period Services can be set-up into separate networks with customized routing, communications channel types and customized calling services. Networks may be configured for interconnection to the public circuit switched networks or to private circuit switched networks. A universal numbering plan will be used that will permit universal calling.

MET Access Management

All METs will be subject to a number of requirements designed to insure that they are capable of operating correctly, that they are authorized for access to the system and that they continue to operate properly once operational. In addition, the control system will be able to exert positive control to manage congestion, abnormal operation or to invoke preemption on behalf of safety services.

METs must be type approved prior to being allowed to be registered in the system. Type approval will be used to verify that the particular model of MET meets minimum performance and operational specifications required to operate successfully within the network.

METs must undergo commissioning before being allowed to use the network

for the first time or after being decommissioned. The commissioning process will verify that the MET is type approved and capable of operating in a satisfactory manner through measurement of key performance parameters and the exercising of appropriate control and signalling sequences.

Each time a request is placed by a MET to enter or use the network, it will be subject to an authentication and authorization procedure. This procedure will verify that the MET ID is authentic and that the MET is authorized to use the system. The procedure will verify that the MET is registered to a legitimate owner and that the owner has current administrative approval to use the system.

Once commissioned, METS will be subject to performance verification testing under the control of the NCS. These tests will be designed to verify that an individual MET operates properly and to minimum performance standards.

All METs using the services of the NCS will be under positive control while operating in the system. Positive control means that the NCS can issue commands at any time directed to one or groups of METs (and associated FESs) that must be acted upon immediately. A primary use for this mechanism is to be able to preempt satellite channels already in use for reassignment to Aeronautical Mobile Satellite (Route) Services, AMS(R)S, upon demand. This is expected to be a relative rare event. Another use is to shutdown METs that have been identified as malfunctioning and that might cause disruption to other users of the satellite. Another use is to implement source control of access when congestion or abnormal operation require limitation on how and when access is permitted.

Network Management and Administration

Management and administration of the network will be centralized with appropriate backup. The NCS will allocate system resources such as satellite spectrum and power to various uses, respond to preemption demands by safety services and insure that they are carried out, manage service restoral and failure recovery, and establish and maintain system databases and configuration control processes. The NCS will provide status monitoring and control support for maintenance and network operations functions. It will collect call records and usage data, generate service status information and reports, provide network performance analysis data and support customer service activities related to status and performance. It will interface to the network control systems of other (non-AMSC) systems for the purposes of intersystem coordination and operation. Examples of other systems are those of Telesat Mobile, Inc. and INMARSAT.

System Sizing

The AMSC system will eventually be comprised of several satellites in several orbital locations. The satellites will each have multiple beams on the service-link (mobile) side. Initially there will be single beams on the feeder-link side although long term there may be multiple beams as well. The first generation system is expected to be able to service at least 0.5 million METs of various types, with long term projections into the several millions. Feeder-link Earth Stations are expected to number into the hundreds minimum, possibly into the thousands. Each orbit slot and it's associated satellites may support several thousand individual communications channels. The NCS must be able to administer and manage the total resources of all these assets.

Availability and Reliability

The NCS must have availability and reliability commensurate with its central and essential role in managing the system.

METs and FESs will be assigned to Control Groups. There may be several Control Groups, each managed separately within the NCS. The availability of a Control Group is specified as 0.999 in the worst month, while the availability of at least one Control Group operative in the NCS is specified at 0.9999 in the worst month. The reliability of an individual Control Group has been provisionally selected as no more than four outages in excess of 30 seconds in any 24 hour period.

The NCS must be designed so that there is no single point failure, since a single point failure might cause unacceptably long outages.

Circuit Switched Connection Performance

It is expected that Grade of Service requirements will vary from one group of users to another. Accordingly, the NCS must be able to efficiently service call blocking rates ranging from 1% to 15% in a nonqueued mode, and optionally be able to provide call queueing with a variable wait delay before blocking. Connection holding times are expected to average from about 20 seconds for dispatch type services to several minutes for general telephone use. The NCS must be able to handle connection requests at an average rate of at least one per 20 seconds per circuit and to service large groups of channels within a Control Group at that rate. The average connection time under clear path conditions should be no more than 2.5 seconds, with 90% of connections completed in less than 5 seconds. The connection time may be longer under poor propagation conditions. Disconnect time should average no more than two seconds.

Signalling and the Radio Link Environment

The signalling between the NCS and the METs and FESs must be robust and reliable to insure a high probability of successful control communication. The service link side will be subject to multipath, shadowing and path blockage as the METs move. The motion will also introduce doppler, and doppler spreading if multipath is present. The feeder-link side will be subject to occasional flat fading caused by hydrometers (rain, snow, ice) that must be compensated for either by adequate link margins, power control, or transfer between feeder-link earth stations. The NCS feeder-link path must be especially robust as it may serve large numbers of users.

PHYSICAL ARCHITECTURE

The Network Control System makes use of several physical facilities as illustrated in Figure 1. The AMSC owned facilities are the Network Operations Center, Network Control Centers and the satellites (not shown). Also, logically part of the NCS are the controllers in the METs and FESs that interwork with the NCS. The Network Operations Center will consist of computer facilities, related peripherals and network operator consoles that support overall network administrative, management and operational needs. It may be colocated with one of the Network Control Centers. The Network Operations Center will be backed up by another suitably equipped facility. There will be two geographically separate Network Control Centers. Geographic redundancy will be used to guard against catastrophic failure of a facility, to provide a means for continuing operations when maintenance or upgrades or modifications require taking down an Network Control Center, or to provide access diversity when the other Network Control Center is subject to propagation induced outages. A Network Control Center will consist of an RF terminal to work with the satellite on the

feeder-link side, signalling channel units for working through the satellite to the METs and FESs, computers and related peripherals for using the signalling channels and interfacing to the Network Operations Center, and such other equipment as is typically required to run an earth station complex. The Network Control Centers do the real-time management of satellite access.

Each Network Control Center will provide at least one set of bidirectional signalling channels for use in each beam. The METs will use these signalling channels for obtaining access to the system. Forward channels will be continuously broadcast with signalling information packet multiplexed into them. Included in these channels will be periodic system configuration and status information required by METs to enter and use the network, or to inhibit access or use if necessary to maintain the integrity of the control system. The return signalling channels will permit random access by the METs to the control system.

Each Network Control Center will have bi-directional signalling channels through the satellite for use by the FESs. There will be at least two channel sets for redundancy. The forward channel will operate in a continuous broadcast packet multiplexed mode. The return channels will operate in a TDMA mode, with possibly a random access mode for low usage FESs. The communications protocol will be based on a well established standard. An alternate interconnection path using either leased line or telephone dial backup facilities will be supported at the Network Control Centers.

The MET signalling channels and FES signalling channels will be used to assign a specific circuit to an MET-FES pair. Once assigned, the MET and FES will establish contact with one another through the communications channel. Present thinking is to require METs to continuously receive an NCS control channel even when using a communications channel, and to use this channel for control purposes. This approach would permit the NCS to directly control certain aspects of MET operation required for system protection (improper operation, system failure and recovery, preemption for AMS(R)S). It would also support the provisioning of custom calling services such as call or message waiting.

The Network Operations Center will provide interconnections to the control centers of Full Period service providers or to other systems such as TMI and INMARSAT. These interconnections will use either X.75 or X.25 Gateway communications protocols.

LOGICAL ARCHITECTURE

The logical architecture is hierarchical in nature. At the top of the hierarchy is the Network Operations Center. The NOC is the ultimate authority in the system. It partitions and assigns resources to Control Groups, manages priority and preemption, generates databases for use of the Control Groups in the carrying out of their functions, provides the interface to external systems, and generally provides the overall management of the system.

Reporting to the NOC are one or more Control Groups. There will be at least one Control Group per satellite. Control Groups are groups of METs and FESs that share in some fashion a collection of space resources (circuits, geographical access), and a Group Controller that manages the Control Group. Group Controllers reside in Network Control Centers. The top level logical relationships within a Control Group are illustrated in Figure 2.

Group Controller

A Group Controller logically consists of a Group Resource Manager, Network Access Processor and one or more Virtual Network Managers.

The Network Access Processor handles all those functions related to the transmission and reception of access and signalling messages between the Group Controller and the METs and FESs. It services all signalling channel sets in all beams served by the Control Group, operates the access protocols over the signalling channels, passes service requests and responses between the METs or FESs and the Virtual Network Manager or Group Resource Controller as appropriate.

The Virtual Network Manager serves a specific set of METs and FESs forming a Virtual Network. Normal requests for service by a member of a Virtual Network are handled by the associated Virtual Network Manager. The VNM contains a database listing all of the attributes, routing algorithms and permitted services associated with each member of the Virtual Network and of the Virtual Network as a set. Using this information, the VNM determines the type of connection to be made (if permitted), which MET/FES pair to connect and the priority to be given. It requests an appropriate circuit assignment from the Group Resource Manager which returns a specific circuit assignment, if available, to the VNM. The VNM passes the assignment on to the MET and FES involved via the Network Access Processor. At the completion of a call, the Virtual Network Manager returns the circuit to the Group Resource Manager.

The Group Resource Manager manages the Control Group as a whole. It maintains the circuit assignment pool used by the Virtual Network Managers, assigning and deassigning circuits as needed. It obtains the members of the circuit pool from the NOC as the NOC dictates. It manages priority and preemption within the Control Group. It conducts commissioning and performance verification testing of METs. It collects call records and status information for periodic transfer to the NOC. It provides status monitoring, alarming and diagnostics for the Group Controller, and supports network operator access and intervention.

Because the Virtual Networks are likely to support a variety of Demand Period services and capabilities, the NOC and its associated Group Controllers must be able to support a wide range of configuration options from which Virtual Networks may be constructed. The categories of configuration options are circuit configuration options, connection routing options and address screening options.

Circuit configuration options define a specific set of standard communication circuit types that will be assignable by the NCS. Each type will have a specified signal format, modulation, channel spacing and power settings associated with a class of METs or FESs.

Connection routing options define the specific methods available for selecting MET/FES connections for each call. The NCS will be capable of establishing connections for either MET originated or FES originated calls. There will be a general purpose routing structure available that can be tailored to specific Virtual Network needs. The structure will support either public or private network inter-working. It will support preassigned routing, alternate routing, unrestricted routing and addressed based restricted routing. The ability to form routing algorithms which can do "least cost routing" is also highly desirable. Least cost routing will be applicable to Virtual Networks that interconnect into public networks where tariffs vary depending on routing.

Address screening options define how an individual call is handled based on source and destination addresses associated with the call. Address screening will be used for routing purposes, but it will also be used to restrict or permit different types of connections based on preestablished custom calling subscriptions. Examples are one-way calling in which calls can be placed in one direction but not the other, restricted destination calling in which a particular source can call only a selected set of destination addresses, and closed user group calling in which calls can be placed between, and only between, a limited set of addresses.

SUMMARY

The functional requirements, physical architecture and logical architecture of the AMSC Network Control System have been described. The system will be designed, developed, built and installed in time for operations to begin through the satellites in late 1993.

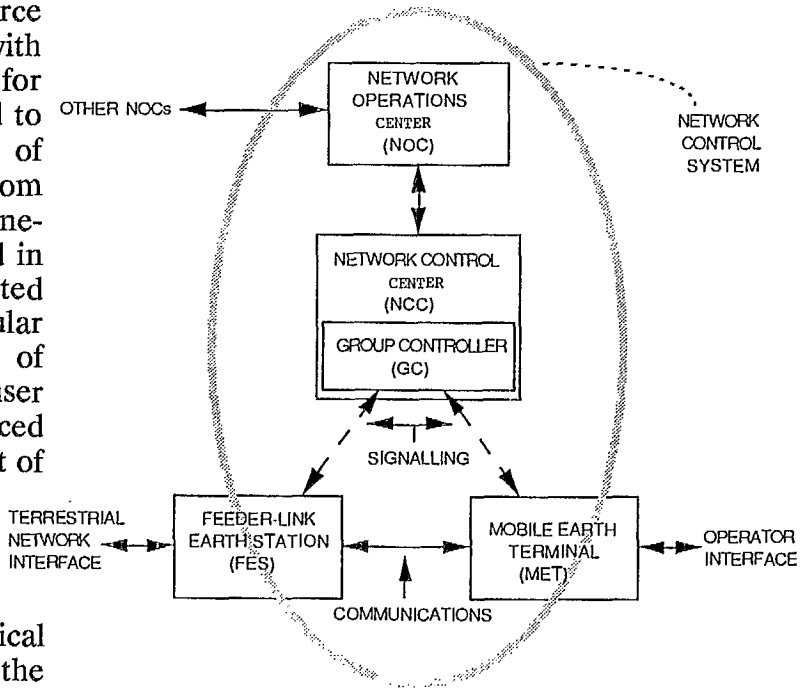


Figure 1. The Network Control System

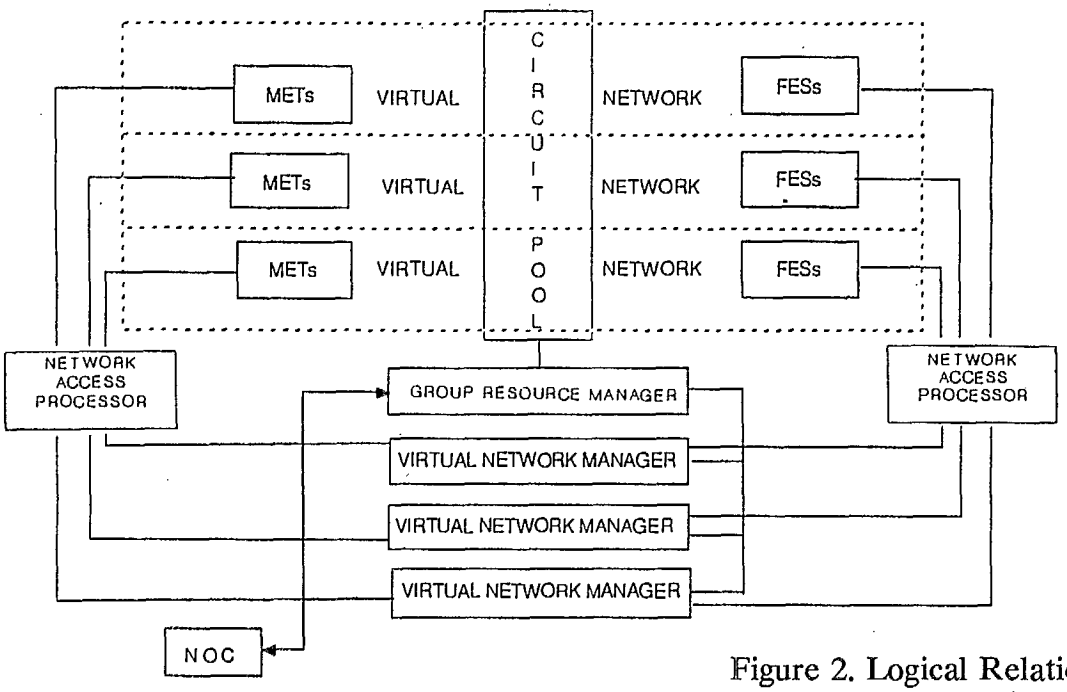


Figure 2. Logical Relationships Within a Control Group

An Architecture for the MSAT Mobile Data System

R.W. Kerr and B. Skerry
Telesat Mobile Inc.
P.O. Box 7800 Ottawa, Ontario
K1L 8E4, Canada
Phone: (613) 746-5601
Fax: (613) 746-2277

ABSTRACT

The MSAT MDS will offer a wide range of packet switched data services. The characteristics and requirements of the services are briefly examined. A proposed architecture to implement these services is presented along with its connectivity requirements. A description of the inbound and outbound channels is provided which are based upon the signalling for the circuit switched services. Also, the duties of the Network Management System are examined.

INTRODUCTION

The Mobile Data System (MDS) for MSAT will provide packet switched data to mobile users anywhere in North America. Due to the harsh nature of the land mobile environment, the system will have to be very robust in order to provide a virtual 'error-free' connection between the Mobile Earth Terminal (MET) and the Data Hub (DH). The Data Hub will be the central point at which all METs interconnect to other public and private data networks. Smaller Regional Data Hubs will provide service to METs in limited areas (one or two beams) and will possibly allow for a reduction in backhaul charges.

Central to the design philosophy behind the METs is the concept of an integrated voice and data terminal. All voice and data services will be available on the same MET, although the MDS will operate independently of the Mobile Telephone System (MTS) and the Mobile Radio System (MRS). However, it is anticipated that with the advent of ISDN in the Public Switched Telephone Network some of the services offered on the MDS could operate within the framework of the MTS.

Given the dual nature of the basic MET, it is foreseen that significant cost savings can be achieved by adopting as much as is feasible for the MDS from the signalling for the circuit switched services (MRS/MTS). Thus the same modulation, coding, interleaving etc. that is used for MRS/MTS signalling will be used as the basis for the packet switched data services [1]. As the MRS/MTS signalling packet sizes have not been optimized for data throughput, some changes will obviously have to be made.

Unlike some currently proposed mobile satellite systems [2,3] the MSAT MDS will not operate on the same inbound and outbound signalling channels used for the circuit switched services. This makes it possible to guarantee access to the system for MRS/MTS regardless of any congestion that would result from the MDS traffic. Also, this allows the data service to expand and change in a modular fashion independent of the MRS/MTS. Different service providers can decide to offer their own packet or circuit switched services autonomously.

Service Offerings

The MDS will provide packet switched data services to both integrated voice/data METs as well as dedicated data terminals. Many of these services are currently being offered [4], but the MSAT MDS will allow for faster, cheaper, and more flexible service. Some of these services are described below, although the list is not exhaustive.

Vehicle Location. Automatic Vehicle Location (AVL) will allow a central dispatcher to

monitor the position of a fleet of vehicles, with periodic updates. This implies some sort of navigation device on the vehicle, such as Loran-C or GPS. This can be coupled with navigation services to allow for optimum routing of a vehicle.

Two-way general messaging. Short messages can be sent between a MET and public or private data networks. An electronic mail box would be a value added service to store messages if a MET is unavailable.

File transfer. This would provide for the MET to exchange relatively long files with the DH. There would be a tradeoff as to whether a file should go as packets on the MDS or demand a circuit via the MRS/MTS.

Interactive data. In order to allow for interactive data sessions, the MDS will have the capability of guaranteeing a specified response time. This could be done either on a priority basis or by assigning a TDMA slot for the MET.

Monitoring and control. Supervisory Control and Data Acquisition (SCADA) will allow for METs to monitor and control METs on a real-time basis.

Virtual circuits. A MET will be able to request a virtual circuit of variable throughput in order to establish a 'virtual circuit'. This would give a user a TDMA slot into which they can put whatever data they wish. From the DH to a MET, the user can be allocated one of the TDM slots on a demand-assigned basis.

ARCHITECTURE

The main Data Hub will be the central point in a double star network configuration. It will manage communications between a large number of MSAT data terminals and various public or private data networks. In the overall MSAT network, the Data Hub will communicate with a Group Controller (GC) to request channels for the operation of the packet data network (see Figure 1). The GC is responsible for managing power and bandwidth resources. Note that this architecture has many similarities with existing VSAT networks [5].

It is possible as that traffic in certain areas or beams will require the addition of a regional data hub. A regional DH may only need to operate in one or two beams and would handle a subset of the MDS customers (see Figure 2). The main DH will communicate in all beams and can assign METs to a regional data hub. Once assigned to the regional DH, that DH will control a MET.

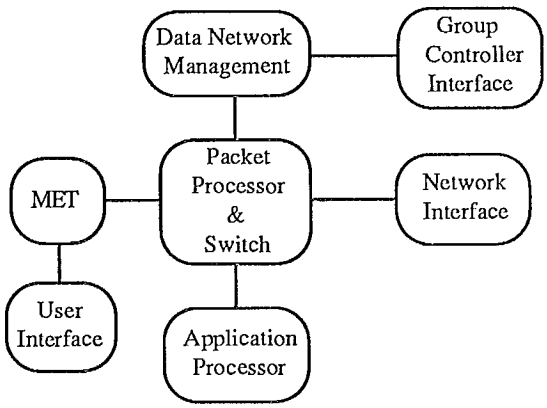


Figure 1. Data Hub Architecture

Network operation will be managed on the assigned channels by the Network Management System (NMS) at the Data Hub. The NMS carries out all of the functions which are required for the efficient operation of the network. A description of the NMS capabilities are discussed below. This structure reduces the GC processing load and the need for combined packet data and signalling channels for the GC, as the DH has the responsibility of managing the data channel implementation and user assignments.

The NMS has direct control over the Packet Processor, which is responsible for the operation of the communication links and all of the operations required to convert satellite packets to a form of data usable by the destination network and routing it to that network (and vice versa). The protocols used on the other networks are well defined standards. In the initial implementation, the protocol that will be used will be X.25. Other protocols may be implemented at a later date, depending on customer demand.

On the satellite side of the Packet Processor, the communication link protocols between the MET and the DH can not be well established protocols. This is because the propagation environment for Mobile Satellite communications can result in bursts of errors . The proposed proto-

cols for the MDS satellite communication links are described in a later section.

The MSAT MET will convert the data for the User terminal from the satellite protocol to a standard protocol to allow for an asynchronous connection between the MSAT MET and User terminal. This terminal could take the form of a hand-held keyboard/display, or possibly a personal computer.

The MDS will be able to offer a wide variety of additional features through the addition of an application processor. This processor handles value added services, which enhance the service seen by the customer and are not a function of a communications network. For example, these features could include a mailbox, a paging service, or broadcast services where a message is periodically repeated.

Channel Descriptions

As indicated previously the MDS is based upon the signalling for the MRS/MTS [1] which in turn draws heavily upon other proposed systems [2,3,6]. However, the MDS has been designed to allow greater data throughput and more flexibility while at the same time trying to reduce the complexity of both the METs and the network infrastructure.

Signalling unit structure. The basic signalling unit (SU) structure is 96 bits long. Note that although the term signalling unit is used, the packet may contain data and/or signalling (but of course only for the MDS). The 96 bit SU includes a CCITT recommended 16 bit Cyclic Redundancy Check (CRC) for error checking.

In order to gain greater data throughput, SUs can be 'chained' together either in the same frame or in subsequent frames. This allows an initial SU (ISU) to be followed by one or more subsequent SUs (SSU) as in other systems [6]. Thus the addressing and other overhead needs to be sent only in the ISU.

Modulation and coding. The proposed initial form of modulation to be used on the MDS is Aviation Binary Phase Shift Keying (A-BPSK) at a transmission rate of 2400 sps (the same as the MRS/MTS signalling). Rate 1/2 coding (k=7) will be used which gives an information rate of 1200 bps.

Future expansion of the MDS could lead to the inclusion of trellis coded modulation (TCM) at 4800 bps [7]. This may be fairly easy to implement because TCM may be used as one of the modulations for the voice channel. Others could be added as well. Thus a fully developed MDS could have a wide range of modulation and coding techniques as in the aeronautical system [3]. This would provide service to a variety of METs with different characteristics and under various propagation environments.

DH-P channel. This is the outbound TDM channel from the Data Hub to the METs. Its frame structure is very similar to the signalling outbound channel for the GC (see figure 3). Data packets may be grouped in 3 or 4 SUs (ISU and 2 or 3 SSUs) in one frame to increase the data throughput to a MET. The DH configuration will be broadcast periodically to ensure every has the current parameters and channel assignments.

MET-R channel. This inbound random access channel will be similar to the signalling access channels. The difference is that short messages will be allowed on these channels (see Figure 4). Messages of up to 3 packets will be able to be sent. Initial SU with the address and header and up to two Subsequent SUs which contain a reference to the ISU, data and CRC, hence increasing the amount of data that can be sent in a packet.

The MET-R channel will use Sloppy Slotted Aloha [8] in order to minimize the guard times. Sloppy Slotted Aloha is a form of Slotted Aloha in which METs are allowed to 'spill over' into adjacent slots on their first access, but will have their timing corrected for subsequent accesses. The retransmission interval of the MET-R channel will vary dynamically in accordance with the traffic.

MET-T channel. The inbound TDMA channel will have very short guard times like the MET-R channel because any MET using the MET-T channel is assumed to have its timing corrected. It may have larger packets that MET-R channel, and/or the capability to assign multiple adjacent slots in every frame to increase the data throughput.

Network Management System. One of the duties of the NMS is scheduling long-term accesses on MET-T channels. A few of the expected applications require allocation of capacity for periodic reporting. The reservation may be needed in one of two ways, either by a customer selected reporting interval or at a specific time. For a given interval, the NMS assigns the MET to a slot that is a given time period from the first transmission where the MET requested the service. The specific time reservation the NMS assigns capacity within a minute of the customer selected time. The NMS re-assigns METs' accesses when it is necessary.

It is expected that short term assignments to the MET-T channel, such as reservations for multi-packet transmissions will be sent on a separate MET-T channel to avoid conflicts between long and short term reservations.

The NMS also exercises control on all of the METs accessing the DH. This is required to prevent malfunctioning METs from interfering with other communications, by either corrective actions, such as adjustments to the MET's frequency or timing, or a complete shutdown of the MET. Control of the MET is also a requirement of the network to accommodate pre-emption requirements for AMSS.

SUMMARY

The Mobile Data System on MSAT will provide for a wide variety of services for users across North America. By adapting as much as possible from the circuit switched services, the cost of both the network infrastructure and the METs can be significantly reduced. At the same time, by offering the MDS on separate channels from MRS and MTS signalling, independence of the services is assured, which enhances the system's flexibility and modularity.

The concept of a regional Data Hub offers further possibilities for gradual system expansion to meet specific market needs. By offering interconnectivity with public and private data networks as well as many value added services, the MDS will be able to provide a comprehensive set of data services.

REFERENCES

[1] "A Network Control and Signalling System for the North American MSS", CCIR SG8 Final Meeting, November 1989.

[2] Mobilesat System Description, Aussat Mobilesat, September 1989.

[3] Aviation Satellite Communications (SatCom) System, Project Paper 741, AEEC, March 1988.

[4] M. Shariatmadar, K. Gordon, B. Skerry, H. El-Damhougy, D. Bossler, "System Architecture for the Canadian Interim Mobile Satellite System", Proceedings of the Mobile Satellite Conference, pp 163-169, May 1988.

[5] J. Stratigos, M. Rakesh, "Packet Switch Architectures and User Protocol Interfaces for VSAT Networks", I E E E Communications Magazine, Vol. 26, No.7, pp 39-47, July 1988.

[6] Standard-M System Definition Manual, INMARSAT, September 1989.

[7] R. Kerr, P.J. McLane, "Coherent Detection of Interleaved Trellis Encoded CPFSK on Shadowed Mobile Satellite Channels", IEEE GlobeCom '89, November 1989.

[8] S. Crozier, "Sloppy-Slotted Aloha", Second International Mobile Satellite Conference, published in these proceedings, June 1990.

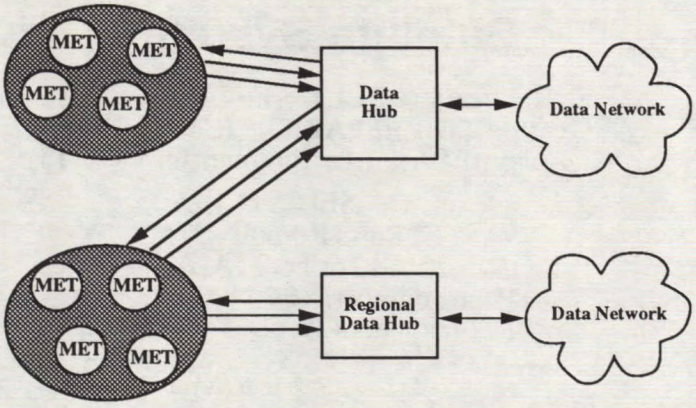


Figure 2. Network Topology

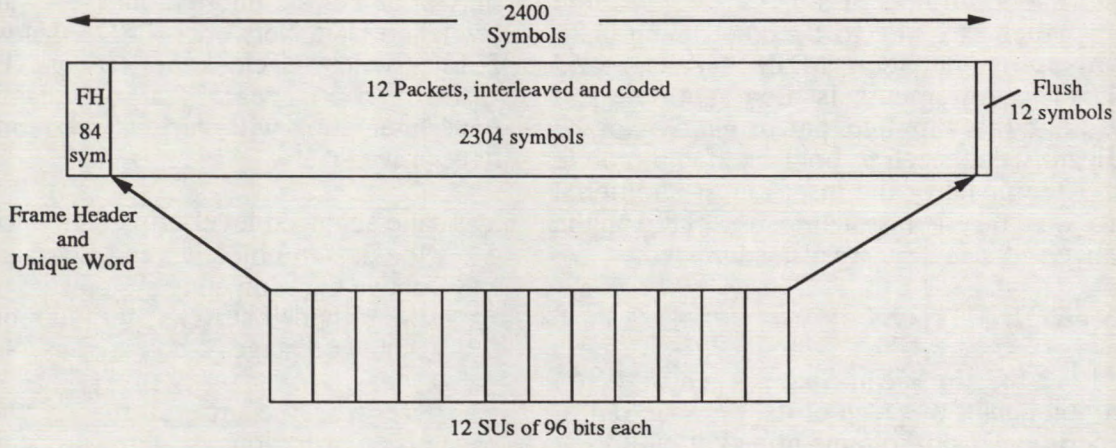


Figure 3. DH-P Channel Format

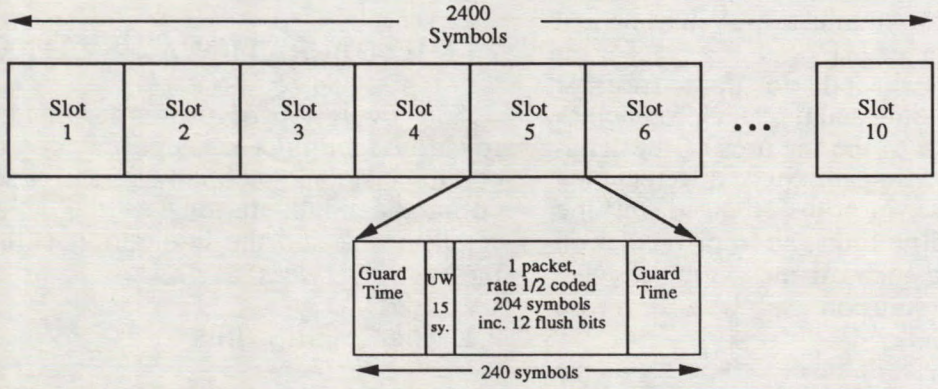


Figure 4. MET-R Channel Burst Format

Message Handling System Concepts and Services in a Land Mobile Satellite System

S. Barberis (CSELT), F. Settimo (=);
A. Giralda (TELESPAZIO), I. Mistretta (=);
C. Loisy (ESA); J.L. Parmentier (SAIT)

CSELT
Via G. Reiss Romoli, 274
I - 10148 Torino, ITALY
Phone number: +39 11 21691
Fax number: +39 11 2169909

ABSTRACT

A network architecture containing the capabilities offered by the Message Handling System (MHS) to the PRODAT Land Mobile Satellite System (LMSS) is described taking into account the constraints of a preexisting satellite system which is going to become operational. The mapping between MHS services and PRODAT requirements is also reported and shows that the supplied performance can be significantly enhanced to both fixed and mobile users. The impact of the insertion of additional features on the system structure, especially on the centralized control unit, are also addressed.

1. INTRODUCTION

In 1982 the European Space Agency (ESA) started, on behalf of seven of its Member States, an experimental programme aimed at exploring the feasibility of providing satellite communications to mobiles equipped with light and inexpensive terminals. After a two year phase spent to characterize the radio propagation environment between satellite and mobiles, a low data rate satellite communication system named PRODAT was designed [1].

The system is based on data transfer protocols, data encoding and modulation schemes specifically adapted to the features of the land-mobile to satellite radio path. Since different data transfer protocols were to be used on the terrestrial and satellite links, an implementation based on store-and-forward messaging service seemed a natural solution and was therefore selected for PRODAT.

In order to verify both the design performance and the suitability of this telecommunication service to satisfy the potential users needs, a prototype system was built to provide an experimental service between some fifty mobile equipments and a few fixed positions. The connections between fixed users and the Earth Stations, having access to the satellite, were established by means of the available public networks, i.e. the international telex and Packet Switched Data Networks (PSDN) as well as the Public Switched Telephone Network (PSTN).

Over two years of successful field experimentation with various user communities have proved that:

- the achieved level of performance had met the design objective and was significantly above the competing systems;
- the provided service met the requirements of a wide range of potential users.

Therefore, the upgrading of the design towards an operational system was decided.

The following chapters briefly present the upgrading of the satellite link performance and the requirements of an operational service; a typical PRODAT Center and the integration of MHS and PRODAT are then described.

2. THE OPERATIONAL SYSTEM

The evolution of the experimental system towards a product susceptible of providing a commercial and potentially attractive service has commanded adaptations in the areas of the satellite link and the integration with standard messaging services.

2.1 The satellite link

Minimising the impact of the space segment utilization charges to the operating costs is of

primary importance in this competitive environment [2]. It is therefore essential that the satellite link design provides an optimal usage of the satellite communication capacity, also when the radio path is interrupted due to obstacles.

For the experimental system, a fixed apportionment of the total communication capacity among all mobiles in session at the same time had been chosen. The operational system, on the contrary, has a dynamically variable channel assignment scheme which limits the real time capacity sharing only to those mobiles which experience good link conditions. Those who are temporarily obstructed are not allowed to transfer data until more favourable conditions are restored. This is achievable in PRODAT owing to the mode of operations of the mobile which, unlike other systems, is full duplex and allows to continuously monitor the link status.

2.2 Integration with standard messaging services

Although the connections with the fixed users in the experimental system use standard protocols on public networks, no attempt could be made at the time of design to comply with the new messaging standards. Today, however, there is a significant incentive in seeking compliance with the new standards, in particular the X.400 CCITT Recs. This because:

- it allows PRODAT to become a ready made product, susceptible of being integrated as an add-on to a messaging network, public or private, when service extension to mobiles, using a satellite connection, is desired;
- the compliance with well accepted standards provides the scope for more economical system implementations; in particular, it is possible to take advantage of the commercial availability of products developed for a wider market than the application considered. A significant economy of scale, when compared with customized developments, is thus achieved.

3. OPERATIONAL REQUIREMENTS

The PRODAT system has the difficult task of facing a market where other systems (Standard C, Qualcomm) provide a similar service. The obtainable performances will thus be compared by both the network operator and the users.

In this competition context, the success of a system will depend not only on the system design quality but also on the fulfilment of several operational requirements, necessary to achieve the target performance degree. Within such requirements, those related to Operations and Maintenance pertain to the System manager, whose choices impact on the achievement of, for example, a high reliability level. Other requirements, described in the following, are more peculiar to the system; among these:

- Service definition
- Network architecture definition
- User requirements.

3.1 Service definition

The services supported by the PRODAT System are divided into four main categories:

- basic services
- supplementary services
- value added services
- emergency messages.

Basic services ensure a message exchange between fixed and mobile users in a store and forward mode. Broadcast and multidestination transmission, as well as request/reply, are particular applications of the first category.

To cope with specific requests coming from the Users community, other optional services like notification delivery, deferred delivery, waiting messages cancellation are also foreseen. More details about the introduction of optional service elements are presented in sect. 5.

3.2 Network architecture definition

The design objectives of the PRODAT network architecture must comply with adequacy and flexibility requirements and can be classified according to the following list:

- to fit into the existing technical standards
- to comply with European regulations
- to allow for future technological changes
- to follow the demand increase with a modular expansion of the radio channels capacity.

The consequent guidelines for the architecture implementation are the following:

- isolation of the satellite access, terrestrial access and message switching functions
- functions location as much as possible in proximity of users and/or network operator
- choice of advanced structures like the CCITT X.400 framework.

3.3 User requirements

In addition to the user interface implementation, some care has been devoted to the security requirements in the areas of:

- access control to the radio link
- privacy of the user related information.

The first problem has been solved with subscriber identity authentication functions which protect the network from the misuse of resources by non authorized people using manipulated or stolen mobile terminals. For the second problem, the network has been equipped with protection functions of the subscriber identity.

4. THE PRODAT CENTER

In the described satellite system configuration, a key role is played by a centralized management unit, the PRODAT Center, which must:

- ensure the connection between fixed and mobile environment linking the satellite segment, accessed by the mobiles, with several terrestrial networks
- to provide an efficient overall network management.

The first requirement consists in ensuring a reliable message transfer between originator and recipient harmonizing the two different communication media involved in a transaction. The satellite subnetwork is the most critical and the PRODAT Center has been equipped with those functions (bidimensional Reed-Solomon coding, CDMA access technique) capable of protecting the transmitted messages at a satisfactory level.

On the fixed network side, the X.400 standard offers a variety of services by means of existing or emerging products provided that the PRODAT Center is interfaced appropriately. In this way, the advantages of an advanced message handling system would be available not only to the users of telematic services such as TELEX but also to OSI Personal Computers.

Two main functional subunits can be identified on the radio side, namely:

- the Satellite Earth Terminal, with the antenna, the frequency converters and the associated facilities;
- the Satellite Access Unit which handles the satellite protocol and interfaces with the store-and-forward part.

More details can be found in the scheme of fig.1 which includes:

- the IF subsystem performing the BPSK/SCPC modulation and, in the opposite way, despreading, QPSK demodulation and vertical decoding. The functions for controlling carrier level and frequency are located here;
- the Satellite Control Unit, responsible not only for the satellite link protocol but also for the other high level protocols. Hence, it provides a reliable transfer service to the Message Transfer Agent (MTA), devoted to message handling;
- the Terrestrial Access Unit, responsible for handling the terrestrial network protocol and the interface with the MTA.

The PRODAT Center also performs the management functions necessary for housekeeping, communication monitoring and customer data handling (configuration, security, billing, accounting, planning etc.). Automatic procedures and powerful User Interfaces are provided as a support to the operator.

A modular concept has been adopted for implementation so that fault recovery, maintainability and further improvements are easier. The needed functionalities are implemented around a communication subsystem based on Ethernet and a Data Base subsystem.

5. MHS AND PRODAT

5.1 An overview of MHS

The Message Handling System (MHS) is a standard electronic mail store-and-forward system which allows the message exchange between users, humans or computers. Several functional components, i.e. the Message Transfer Agent (MTA), the User Agent (UA), and the Access Unit (AU) interwork with each other to provide MHS services.

The UA allows Users to access the MHS services. The MTAs cooperate with each other to supply a reliable relay and delivery: this feature makes them the fundamental MHS entities. The AU allows the Users of other telematic services to communicate with MHS Users.

In the MHS environment two basic services have been standardized, namely:

- Message Transfer Service (MTS). It is a general store-and-forward and application independent service which involves only MTAs;
- InterPersonal Messaging (IPM). It allows a message exchange between Users in an electronic mail fashion.

The structure of an MHS message has been standardized in two components, an "envelope" and a "content". On the former, all information needed to route the message is written; the latter is the IPM message, i.e. the useful data.

In the following, a mapping between MHS services and PRODAT requirements is attempted for both MTS and IPM, commenting the service elements deemed significant to an LMSS.

5.2 Services and architectures

5.2.1 The MTS service

The MTS is accomplished by means of a set of MTA centres which communicate with each other via a standardized protocol named P1. Every MTA performs the following basic actions:

- **submission:** it allows an originating UA to request the transfer of a message to the MTA directly connected with the UA;
- **delivery:** it allows an MTA to deliver a message to the UA recipient;
- **transfer:** it involves two or more MTA centres. Every MTA routes the message through the network according to the informations written on the envelope;
- **notification:** it informs the originating UA of the message delivery/non delivery to the UA recipients.

The analogy between the basic interactions of an MTA and the main functions of a LMSS is fairly evident and, furthermore, it can be shown that several MTS service elements are already somehow provided by PRODAT.

The entire set of standardized MTS service elements is contained in [3], with a subdivision into basic and optional ones. The former class includes, in particular: Submission and delivery, Non delivery notification, Message identification, Submission and delivery time stamp Indication which do not need further comments.

Among the latter, Probe, Deferred delivery and Deferred Delivery Cancellation seem to be tailored for an environment where the recipient is sometimes unreachable and the information contained in a message is useful only for a limited period. The Probe service element, in particular, consists in sending a short message and reports about the mobile reachability; in this way, useless occupancy of resources can be saved, especially when long messages are transmitted.

Also the Grade of Delivery Selection allows PRODAT users to specify how urgent the relay through the MTS and delivery to the PRODAT Center must be. The chosen urgency level will be afterwards translated into the appropriate transmission priority before accessing the radio link; if necessary, also the retransmission frequency could be increased in order to maximize the probability of reaching the destination mobile in time.

Finally, the Hold for Delivery could be used by an overloaded PRODAT Center as a flow control mechanism when the number of queueing messages exceeds a predefined threshold. The responsibility for a temporary message storage lies in the MTS until the recipient is again ready to accept delivery.

The study of the correspondence between MTS and PRODAT services has been accompanied by the definition of an access network architecture (fig.2). Here, the fixed PRODAT Users are not necessarily regular MHS customers and this allows a simplified terminal equipment. The AU must be designed on purpose: it behaves essentially as a concentration point where all software needed to ensure compatibility between the (PRODAT-like) User generated messages and MTS resides.

Between AU and MHS/PRODAT Interface, the message exchange is fully standardized to the application layer (protocol P1). This is represented by means of the familiar OSI stack without layer 6, not standardized by MHS. The Reliable Transfer Server is a standard layer 7 service element, responsible for the reliable relay between MTAs.

The functional simmetry of the MTS-based access architecture points out that the PRODAT Center is logically equivalent to a fixed User. Hence, it will receive only the content of the P1-messages delivered to the MTA recipient. A suitable MHS/PRODAT Interface provides the necessary adaptation so that changes to the

PRODAT Center internal structure are minimized. The received messages are then stored according to the queuing modality and processed according to the design of the satellite protocol.

5.2.2 The IPM service

In the access scheme of fig 1, Mobiles are not direct MHS users and the advantages of MHS practically involve only the terrestrial link, i.e. the less critical element of the PRODAT network. A higher service efficiency can be achieved with an access network architecture, based on InterPersonal Messaging, which includes Mobiles in the MHS community.

IPM provides to the individual Users all the communications services typical of electronic mail. Being a higher level application than MTS, it supplies to IPM subscribers several basic and optional additional services. The former allow to:

- send and/or receive messages of the electronic mail system;
- use the capabilities of MTS;
- identify, at user level, the messages submitted to the MTS for transfer;
- recognize the information types contained in the message body. This is obviously not possible to MTS, which only sees the message "envelope".

Some optional services are attractive for PRODAT applications. For example, the Reply Request Indication and the Replying IP-message Indication represent a ready made upgrade to an existing PRODAT service and correlate automatically Original and Reply messages. Non receipt and Receipt Notification Requests are a much more interesting piece of information than the Non Delivery and Delivery Notification supplied by MTS. The former, in fact, indicate if the mobile has been successfully reached; the latter only indicate if the PRODAT Center has been reached via the terrestrial link but nothing can be said about the final destination.

Expiry Date Indication and Obsoleting Indication help to avoid transmitting obsolete messages to a mobile which has been unreachable for some time, when the radio link improves. The saving of radio resources usage is evident.

Fig 3 shows a proposal for integrating IPM into the PRODAT environment. It is evident, first of all, that Fixed and Mobile PRODAT customers are individually considered and that each one has access to a dedicated storage area (UM) of the computers hosting the UA software. This makes it possible to address users in an IPM fashion by means of the standard application protocol P2 (based on protocol P1 for transfer). User Memories are not yet provided in the PRODAT Center design.

Another interesting feature of this second access architecture is the symmetry between the Fixed Users side and the PRODAT Center side: this should allow replicating, or at least reusing with minor modifications, a part of the involved software. It is finally noted that the proposed architecture is particularly favourable to those Users who are already MHS subscribers and want to becom PRODAT customers too: these ones, in fact, need no additional equipment to access PRODAT with the benefits of MHS.

6. CONCLUSIONS

The achieved results allow to state that the disadvantages of a mobile environment (typically on the radio link) can be mitigated not only by a robust access protocol to the space segment but also with reliable message transfer procedures on the terrestrial link. The paper contains some implementation proposals which validate the basic idea showing how the theoretical service elements of CCITT Recs can be integrated in a real system.

Acknowledgement

The authors are grateful to Mr. C. Berrino (CSELT) for his valued suggestions in the access architecture definitions.

REFERENCES

1 **Rogard, R.; Jongejans, A.; Loisy, C.** 1989. Mobile Communications by Satellite - Results of Field Trials conducted in Europe with the PRODAT System. In: *ESA Journal 1989, Vol. 13 pp. 1-12*

2 **Bartholome, P; Rogard, R.** 1988. A Satellite System for Land-Mobile Communications in Europe. In: *Proceedings of the Mobile Satellite Conference, Pasadena Ca. May 3-5 1988 pp. 37-42*

3 **CCITT Recommendation X.400** *Blue Book, Melbourne 1988*

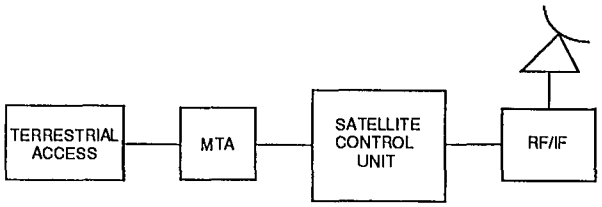


Fig. 1 - PRODAT Centre Architecture

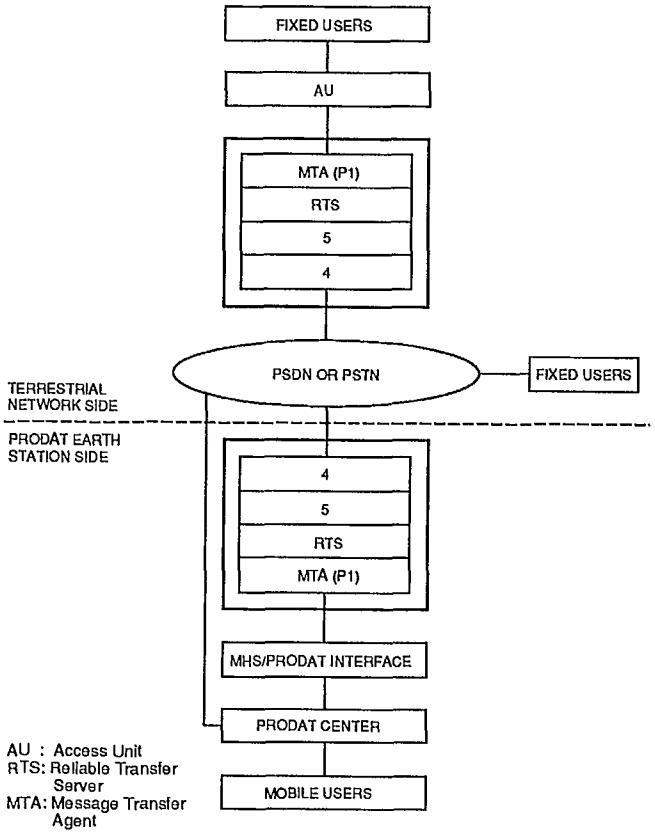


Fig. 2 - MTS based architecture

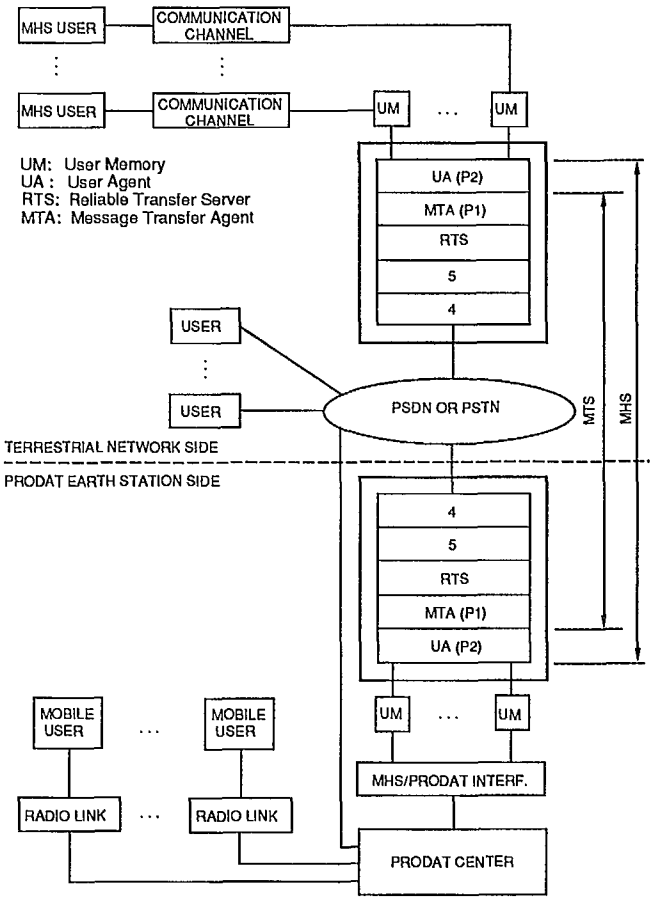


Fig.3 - IPM based architecture

An Alternative Resource Sharing Scheme for Land Mobile Satellite Services

Tsun-Yee Yan and Miles K. Sue
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

Abstract

This paper presents a preliminary comparison between the two competing channelization concepts for the Land Mobile Satellite Services (LMSS), namely frequency division (FD) and code division (CD). Both random access and demand-assigned approaches are considered under these concepts. The CD concept is compared with the traditional FD concept based on the system consideration and a projected traffic model. It can be seen that CD is not particularly attractive for the first generation Mobile Satellite Services because of the spectral occupancy of the network bandwidth. However, the CD concept is a viable alternative for future systems such as the personal access satellite system (PASS) in the Ka-band spectrum where spectral efficiency is not of prime concern. The effects of power robbing and voice activity factor are incorporated. It has been shown that the traditional rule of thumb of dividing the number of raw channels by the voice activity factor to obtain the effective number of channels is only valid asymptotically as the aggregated traffic approaches infinity.

1 Introduction

In 1983 the Jet Propulsion Laboratory (JPL), as lead center for the National Aeronautics and Space Administration (NASA)'s Mobile Satellite Program, selected to follow the frequency division (FD) concept for the Mobile Satellite Experiment (MSAT-X) project. One of the main system considerations was to offer low cost terminals to the general subscribers. The concept of code division (CD) was discussed but was not pursued because of the cost and maturity of the technology for mobile applications.

1.1 Resource Sharing Approaches

There are two basic approaches to utilize these two channelization concepts. One is the random access approach whereby each subscriber is allowed to transmit without making explicit requests to a Network Control Center (NCC). The other one is the demand-assigned approach. Each subscriber must make a connection request known to the NCC before a transmission is granted. In MSAT-X, an Integrated-Adaptive Mobile Access Protocol (I-AMAP) governs the operations of subscribers accessing the network. It is an example of the demand-assigned approach using the FD concept. The protocol consists of the channel access protocol and the connection protocol. The channel access protocol is identical to the random access approach without the NCC. The prime objective of using I-AMAP was to maximize the number of subscribers for the first generation Mobile Satellite Services (MSS).

On May 31, 1989, The Federal Communications Commission (FCC) granted the MSS license to the American Mobile Satellite Corporation (AMSC). It is very likely that AMSC will adopt the FD concept for the first generation MSS. Advanced technologies developed under MSAT-X might be transferred to AMSC.

Ideally, I-AMAP which is a generic protocol is independent of the channelization technique. Results derived for the MSAT-X therefore can be directly applied to predict the performance of the CD network using the I-AMAP under certain conditions. The differences lie in the hardware implementation between these two concepts and the throughput of the random access technique in making connection requests.

This paper provides a preliminary analysis between these two competing concepts for the MSS from a different perspective. Previous discussions compare the

performance of these two concepts based on the satellite operating characteristics [1, 2]. Recently the joint Qualcomm/PacTel experiment demonstrated the feasibility of adopting CD concept for the cellular radio environment. The concept of CD emerges as a promising candidate for designing future MSS.

1.2 Impacts of Voice Activity

During a conversation, the gaps between talk spurs can be exploited to maximize the utility of satellite resources. The voice activity factor increases the effective number of voice channels. Because the voice activity is statistical in nature, the possibility of satellite overloading exists. Overloading occurs when the number of voice packets arriving at the satellite from different active subscribers exceeds the threshold value, beyond which the satellite will no longer be able to provide adequate power amplification for each channel. Overloading can result in performance degradation, i.e., increased bit error rate (BER) and packet errors, and lower voice quality. In case of mixed voice and data traffic, overloading can result in increased data packet errors and retransmissions. The extent of degradation depends on the link characteristics and is severer for downlink limited channels. This paper examines the probability of overloading and the resulting degradation for voice only traffic. Typical parameter values are based on the current personal access satellite system (PASS) design. Results will be used to establish the margin (TWT backoff) needed to meet the reliability requirement.

For systems following the CD concept, the statistical nature of voice activity also causes the self-noise level to vary accordingly. This factor should be considered in system design. In this analysis, this effect is ignored.

Section 2 describes the impact on system performance due to voice activity factor, or commonly known as the VOX factor, for FD under the demand-assigned approach. Section 3 describes the effect of VOX on the CD concept. Section 4 relates the overloading probability with the satellite transponder margin to meet the reliability requirement. Section 5 presents a framework to compute the throughput of a CD network under the random access approach. Section 6 provides a summary.

2 Effect of VOX on FD Architecture

Let B be the available bandwidth and W be the channel spacing for a FD architecture under the demand-assigned approach. For voice application, the effect of the channel access protocol operated on request channels can be ignored. Without loss of generality, the total number of available channels, N_t, can be written as $N_t = \frac{B}{W}$. In practice, the available satellite power may not be able to support all N_t channels. Let P be the total RF power and P_f be the power required to achieve the bit error rate of γ_f per FD channel. We made a simplified assumption that the channel will be error-free whenever the power exceeds P_f. The number of channels available under the power constraint is then given by $N_p^{(f)} = \frac{P}{P_f}$. In general $N_p^{(f)} < N_t$ represents a power limited satellite, and $N_t < N_p^{(f)}$ represents a bandwidth limited satellite.

Let λ be the aggregated average call arrival rate and $(\frac{1}{\mu})$ be the average call duration. (μ is the average service rate per call.) If the satellite is bandwidth limited, for any given blocking probability α_f and N_t, the scenario can be modeled by a $M/M/N_t/N_t$ queuing model under the Poisson assumption. The aggregated traffic that can be supported is given by the well known Erlang-B formula,

$$\alpha_f = \frac{\rho^{N_t} \frac{1}{N_t!}}{\sum_{k=0}^{N_t} \rho^k \frac{1}{k!}} \qquad (1)$$

where $\rho = \frac{\lambda}{\mu}$.

Equation (1) does not consider that at least half of the time either party is listening rather than talking. If the satellite is bandwidth limited, we can not take advantage of the gaps between talk spurs. However, in a power limited satellite, since $N_p^{(f)} < N_t$, the maximum number of channels that carries the conversation is smaller than the available number of channels N_t. During the silence period of a conversation, the available satellite power P_f can support additional channels. In this paper, we consider the case of only power limited satellites.

Consider the same queuing model that the call service rate is increased by a factor of ν. This corresponds to a VOX factor of $(\frac{1}{\nu})$. For a power limited satellite, the probability that an incoming call will be blocked for a given $N_p^{(f)}$ is the probability that a call will overload the satellite.

Let N_r be the number of channels required for any

given ρ and δ_0 be the desired overloading probability. Then

$$\delta_0 = \frac{(\frac{\rho}{\nu})^{N_r} \frac{1}{N_r!}}{\sum_{k=0}^{N_r} (\frac{\rho}{\nu})^k \frac{1}{k!}}. \quad (2)$$

Note that the total satellite power required is $N_r P_f$. If the resulting $N_r < N_p^{(f)}$, the satellite has sufficient power to support all N_r channels without even considering the VOX factor. A more interesting case is when $N_r > N_p^{(f)}$.

Let

$$\beta = 1 - \frac{N_r}{N_s} \quad (3)$$

where N_s is the number of channels without the VOX factor (i.e. $\nu = 1$). Table 1 shows the percentage of improvement due to VOX factor of 50% for an overloading probability of 5%. Tables 2 and 3 show the corresponding improvement for 1% and 0.1%. It is interesting to observe that as the traffic increases, the improvement due to VOX approaches a limit. Specifically, let

$$\delta_0 = \frac{(\frac{\rho}{\nu})^{N_r} \frac{1}{N_r!}}{\sum_{k=0}^{N_r} (\frac{\rho}{\nu})^k \frac{1}{k!}} = \frac{\rho^{N_s} \frac{1}{N_s!}}{\sum_{k=0}^{N_s} \rho^k \frac{1}{k!}}. \quad (4)$$

As ρ increases, the summations of both denominators are dominated by the last two terms. From (4), for large enough ρ,

$$\frac{(\frac{\rho}{\nu})^{N_r} \frac{1}{N_r!}}{(\frac{\rho}{\nu})^{N_r} \frac{1}{N_r!} + (\frac{\rho}{\nu})^{N_r-1} \frac{1}{(N_r-1)!}} = \frac{\rho^{N_s} \frac{1}{N_s!}}{\rho^{N_s} \frac{1}{N_s!} + \rho^{N_s-1} \frac{1}{(N_s-1)!}}. \quad (5)$$

Rearranging terms in (5), we have

$$\lim_{\rho \to \infty} \left(\frac{N_s}{N_r} \right) = \nu \quad (6)$$

Figure 1 shows the percentage of improvement due to the VOX factor for various traffic conditions. From Figure 1, it can be seen that the improvement due to the VOX factor approaches a limit as predicted by (6). The usual rule of thumb of dividing the number of raw channels by the VOX factor to obtain the effective number of channels is only asymptotically true for heavy traffic conditions. The fluctuation at low traffic levels is due to the integral values of the number of channels.

Traffic (Erlangs)	Number of Channels		(β)% of Improvement
	w/o VOX	w/ VOX	
1.0	4	3	25.0
11.0	16	10	37.5
21.0	27	15	44.4
31.0	37	21	43.2
41.0	47	26	44.7
51.0	57	31	45.6
61.0	67	36	46.3
71.0	77	41	46.8
81.0	87	46	47.1
91.0	96	51	46.9
101.0	106	56	47.2
111.0	116	61	47.4
121.0	126	66	47.6
131.0	135	71	47.4

Table 1: Performance Improvement due to VOX of 50% and $\delta_0 = 5\%$

Traffic (Erlangs)	Number of Channels		(β) % of Improvement
	w/o VOX	w/ VOX	
1.0	5	4	20.0
11.0	19	12	36.8
21.0	31	19	38.7
31.0	43	25	41.9
41.0	54	31	42.6
51.0	65	36	44.6
61.0	76	42	44.7
71.0	87	48	44.8
81.0	97	53	45.4
91.0	108	59	45.4
101.0	118	64	45.8
111.0	129	70	45.8
121.0	139	75	45.7
131.0	150	81	46.0

Table 2: Performance Improvement due to VOX of 50% and $\delta_0 = 1\%$

Traffic (Erlangs)	Number of Channels		(β) % of Improvement
	w/o VOX	w/ VOX	
1.0	6	5	16.7
11.0	23	15	34.8
21.0	36	22	38.9
31.0	49	29	40.8
41.0	61	35	42.6
51.0	73	42	42.5
61.0	84	48	42.9
71.0	96	54	43.8
81.0	107	60	43.9
91.0	118	66	44.1
101.0	129	72	44.2
111.0	141	78	44.7
121.0	152	84	44.7
131.0	163	89	45.4

Table 3: Performance Improvement due to VOX of 50% and $\delta_0 = 0.1\%$

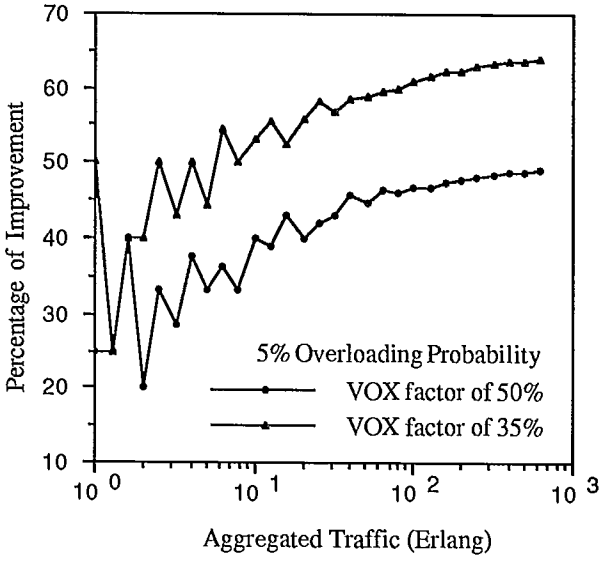


Figure 1: Performance Improvement due to VOX factor for various traffic loading

3 Effect of VOX on CD Architecture

Now, consider the case of a CD architecture. Let P_c be the power required to achieve the bit error rate of γ_c per CD channel. We made a simplified assumption that the channel will be error-free whenever the power exceeds P_c. The number of channels available under the power constraint is then given by $N_p^{(c)} = \frac{P}{P_c}$. Unlike the FD concept, the maximum number of communications channels is solely determined by the quantity $N_p^{(c)}$. Under the demand-assigned approach, we can make a simplified assumption that the NCC controls the number of requests. In this case, the scenario operated under CD concept is similar to the FD architecture. Results developed in the previous section are valid for the CD concept under the demand-assigned approach.

4 Performance Degradation

This section relates the overloading probability to the required link margin (or degradation) for a given satellite transponder design point. If the link is limited by uplink thermal noise, overloading effect can be negligible. On the other hand, if the link is severely limited by downlink thermal noise, overloading can cause significant degradation. Figure 2 shows the degradation of $\frac{C}{N_0}$, $\Delta_{\frac{C}{N_0}}$, in dB as a function of per channel downlink power normalized to the design point value. Results in Figure 2 are parameterized by η, where η is the ratio of downlink thermal noise to uplink thermal noise at design point. It should be noted that the satellite traveling wave tube (TWT) is assumed to be operating at constant output power in order to examine the impacts of voice activities. Effects such as interference and other channel impairments are ignored.

Table 4 relates the overloading probability with the degradation $\Delta_{\frac{C}{N_0}}$ for a fixed design point as a function of η for two traffic scenarios: high traffic (101 erlangs) and low traffic (21 erlangs). Results in Table 4 are computed for the case with VOX and the satellite TWT sized for a fixed operating point corresponding to 5% overloading probability. Degradations of $\frac{C}{N_0}$ are calculated as the overloading probability changes from 5% to 1% to 0.1%. (In other words, system availability increases from 95% to 99% to 99.9% as shown in Table 4.) Entries in Table 4 for system availabilities of 99% and 99.9% are based on Tables 2

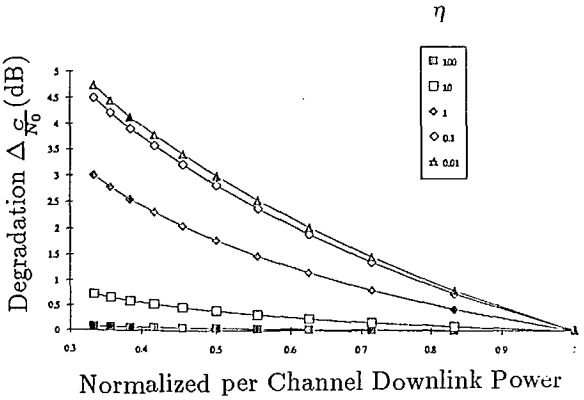


Figure 2: Degradation on the total $\frac{C}{N_0}$

and 3. As expected, the degradation is more severe for low traffic systems. Based on the table, in order to achieve a high system availability (99.9%), a margin of about 2 dB is necessary for low traffic and downlink limited systems. For uplink limited systems, the degradation is negligible. For the case without VOX, lowering the overloading probability from 5% to 1% does not, in practice, result in performance degradation.

5 Throughput of Random Access CD Network

Both CD and FD concepts can be utilized under the random access approach. This section provides a framework to evaluate the throughput of the CD network under the random access approach. It has been suggested [3] that a CD network should be operated at about 10% of the total network bandwidth for reasonable error performances without specific coding techniques. This observation was supported by Xiang [4] in his detailed computation of the maximum number of subscribers for non-coherent-phase CDMA at BER of 10^{-3} for non-fading channels. In [5], the maximum throughput for CDMA was derived for convolutional coded systems. Later papers concentrated on calculating the bit error and packet error performance for different combinations of modulation and spreading techniques on various transmission conditions [6, 7]. In [8], the throughput of a slotted random access CDMA system was derived under the Poisson arrival model. This paper includes the voice activity in the model by incorporating a parameter ν as discussed in Section 2.

We assume a slotted system where the arrival process is Poisson with an average aggregated arrival rate λ_c. Each subscriber transmits a packet of fixed length L. Let $P_c(m)$ be the probability of correctly receiving a packet from a designated subscriber, given that there are $(m-1)$ other simultaneous transmissions. Let $\mathcal{S}$ be the throughput (packets/slot) of the random access CDMA system. Using the voice activity model described in Section 2, the throughput is given by

$$\mathcal{S} = \rho_e e^{-\rho_e} \sum_{m=0}^{\infty} \frac{\rho_e^m}{m!} P_c(m+1), \qquad (7)$$

where $\rho_e = \frac{\rho_c}{\nu}$ is the effective traffic and $\frac{1}{\nu}$ is defined in Section 2. The quantity $\rho_c = \lambda_c T$ represents the aggregated traffic per slot, where T is the slot size. The average number of retransmissions is given by

$$r_{av} = \frac{\rho_e}{\mathcal{S}} - 1.$$

Equation (7) represents a generic expression which relates the throughput of a CD network with the modulation and spreading technique. Once the probabilities of $P_c(m)$ are specified, the throughput of the CDMA system can be determined. The computation of $P_c(m)$ for a specific CDMA system is quite involved and is beyond the scope of this paper. In the following, we demonstrate the computation using (7) by three examples. Example 1 and example 2 illustrate two bounding conditions for a random access CD system. Example 3 shows the throughput of a sample CD system under the random access approach.

Example 1. Let $P_c(1) = 1$, $P_c(m) = 0$ for $m \geq 2$, and $\nu = 1$. This corresponds to a slotted ALOHA system without considering the advantage of the VOX factor under the CD concept. From (7), the throughput is given by

$$\mathcal{S} = \rho_c e^{-\rho_c}.$$

Example 2. Let $P_c(m) = 1$ for all m, and $\nu = 1$. This corresponds to a perfect system with arbitrary number of CD channels without the VOX factor. The throughput is therefore ρ_c.

Example 3. Consider a binary direct-sequence spread spectrum system with random code sequences. Lehnert and Pursley [6] have provided the upper and lower bounds for the bit error probability. They also compared the performance with the Gaussian approx-

η	Degradation $\Delta_{\frac{C}{N_0}}$(dB) for selected system availability and traffic				Comments
	21 Erlangs		101 Erlangs		
	99%	99.9%	99%	99.9%	
$\frac{1}{100}$	1.0	1.7	0.6	1.1	extremely downlink limited case
$\frac{1}{10}$	0.9	1.5	0.5	1.0	downlink limited case
1	0.5	0.9	0.3	0.6	equal uplink and downlink noise
10	0.1	0.2	0.1	0.1	uplink limited case
100	0	0	0	0	extremely uplink limited case

Table 4: Degradation of Total $\frac{C}{N_0}$ based on 5% Overloading Probability Design Point

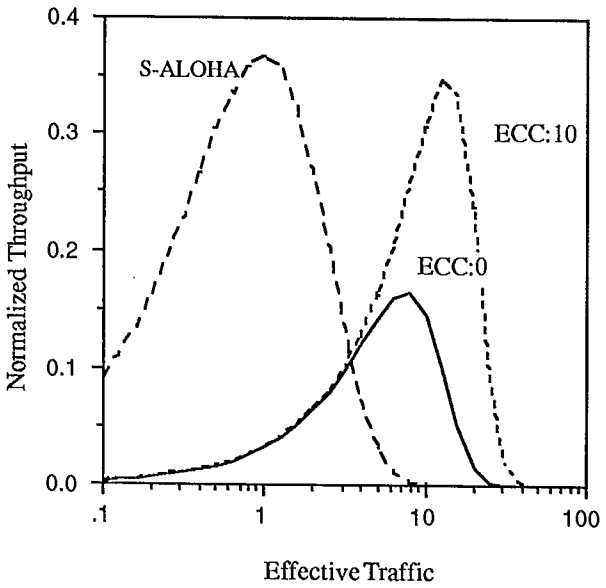


Figure 3: Normalized Throughput of the sample CD network in Example 3

imation. In [7], it has been shown that the Gaussian approximation is accurate only when there are a large number of simultaneous subscribers on the channel; otherwise the approximation can be optimistic by several orders of magnitude. They computed the average probability of packet error versus the number of subscribers. Table 5 gives the $P_e(m)$ versus m. In Table 5, number of chips per data bit is 31, packet length is 1000 including the error correction bits. The error correction capability of the code is assumed to be 10. In Figure 3, the normalized throughput of this CDMA system is denoted by ECC:10. Also in Figure 3, we include the case of no error correction capability for comparison. Notice that the vertical axis of Figure 3 is normalized by the bandwidth expansion of the sample CD network.

Figure 3 shows that, for this sample CD network, the maximum throughput is smaller than that of the classical S-ALOHA network. In case of no error correction capability (ECC:0 in Figure 3), the maximum throughput is less than half of that of the S-ALOHA network. The maximum throughput of 0.35 packets per channel per slot of the sample CD network is attainable at 13.65 packets per slot offered traffic. In contrast, the S-ALOHA network achieves the maximum throughput of 0.37 at 1 packet per channel per slot offered traffic. However, the 13.65 times increase in the offered traffic is in the expense of the 31 times increase in the network bandwidth. If we can utilize the same increase of bandwidth for the S-ALOHA system, the effective traffic that can be supported is 31 packets per slot. The efficiency of the sample CD network is therefore $\frac{13.65}{31} = 44\%$ of the S-ALOHA network.

It should be noted that the maximum throughput is not the only measure to conclude the effectiveness of a CD network. Other system parameters such as delay, stability, achievable throughput, complexity of the terminal design, bandwidth availability, and the incorporation of the random access technique into an operational protocol must also be considered. Furthermore, the analysis presented in this paper is based on a set of simplified assumptions and a sample CD network. The incorporation of the noise activity in the random access model may not reflect realistic conditions. Different conclusions may be drawn if the scenario changes. Detail studies must be conducted to determine the advantages and disadvantages of each channelization concept.

6 Summary

The throughput of CD under the random access scheme and the effects of voice activity factor have

m	$P_c(m)$	m	$P_c(m)$	m	$P_c(m)$
1	1.0	11	1.0	21	0.28
2	1.0	12	0.98	22	0.20
3	1.0	13	0.96	23	0.15
4	1.0	14	0.92	24	0.10
5	1.0	15	0.85	25	0.07
6	1.0	16	0.78	26	0.05
7	1.0	17	0.70		
8	1.0	18	0.60		
9	1.0	19	0.48		
10	1.0	20	0.37		

Table 5: Probability of Success for various m of Example 3

been investigated using a set of simplifying assumptions. It has been shown that the traditional rule of thumb of dividing the number of raw channels by the voice activity factor to obtain the effective number of channels is only valid asymptotically as the aggregated traffic approaches infinity under the demand-assigned approach for both CD and FD networks. The CD concept is compared with the traditional FD concept under the random access approach and a projected traffic model. It can be seen that CD is not particularly attractive for the first generation MSS with limited bandwidth. However, the CD concept is a viable alternative for future systems such as the personal access satellite system in the Ka-band spectrum with ample available bandwidth. It should be noted that there are many other factors affecting the choice of CD versus FD. A number of studies have been performed in search for a proper resource sharing scheme for PASS from different perspectives. Some of the comparisons can be found in [9, 10].

Acknowledgement

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

References

[1] B. Levitt, et al., "A Unified Analysis of the Effects of Uplink Noise and Multi-User Interference on the End-to-End Performance of a DS/CDMA Bent -Pipe Satellite Relay System," JPL Internal report, December 19, 1989.

[2] K. Dessouky and M. Motamedi, "FDMA and CDMA PASS Design Tradeoffs, Comparisons of Results, and Conclusion," JPL IOM-3392-90-001, January 3, 1990. (Internal document)

[3] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of Spread Spectrum Communications – A Tutorial," IEEE Trans. on Communications, COM-30, May, 1982. pp.855-884

[4] H. Xiang, "Binary Code-Division Multiple-Access Systems Operating in Multipath Fading, Noisy Channels," IEEE Trans. on Communications, COM-33, No. 8, August 1985. pp.775-784

[5] J. Hui, "Throughput Analysis for Code Division Multiple Accessing of the Spread Spectrum Channel," IEEE Journal on Selected Areas in Communications, SAC-2, July, 1984. pp.482-486

[6] J. Lehnert and M. Pursley, "Error Probabilities for Binary Direct-Sequence Spread-Spectrum Communications with Random Signature Sequences," IEEE Trans. on Communications, COM-35, No. 1, January 1987. pp.87-98

[7] R. Morrow and J. Lehnert, "Bit-to-Bit Error Dependence in Slotted DS/SSMA Packet Systems with Random Signature Sequences," IEEE Trans. on Communications, COM-37, No. 10, October 1989. pp.1052-1061

[8] D. Raychaudhuri, "Performance Analysis of Random Access Packet-Switched Code Division Multiple Access System," IEEE Trans. on Communications, COM-29, No. 6, June, 1981. pp.895-901

[9] M. Motamedi and M. Sue, "A CDMA Architecture for a Ka-Band Personal Access Satellite System: Complexity and Capacity," Proceeding 13th AIAA International Communication Satellite Systems Conference, Los Angeles, CA, March 11-15, 1990.

[10] K. Dessouky and M. Motamedi, "Multiple Access Capacity Trade-Offs for A Ka-Band Personal Access Satellite System," Proceeding International Mobile Satellite Conference, Ottawa, Ontario, Canada, June 18-20, 1990.

Advanced Multiple Access Concepts in Mobile Satellite Systems

Fulvio Ananasso
Electronics Engineering Dept. (DIE) - University of Roma/Tor Vergata
Via O.Raimondo 8 - 00173 ROMA (Italy)
Phone: + 39-6-24990-462(or -451)
FAX: + 39-6-2490-519

ABSTRACT

The paper expands upon some multiple access strategies for Mobile Satellite Systems (MSS), as investigated in the frame of three separate studies for the International Maritime Satellite Organisation (INMARSAT) and European Space Agency (ESA). SS-FDMA, CDMA and Frequency-Addressable Beam architectures are addressed, discussing both system and technology aspects and outlining advantages and drawbacks of either solution with associated relevant hardware issues. An attempt is made to compare the considered options from the standpoint of user terminal/space segment complexity, synchronization requirements, spectral efficiency and interference rejection.

The opinions expressed in the paper, although related to the mentioned study contracts awarded to Telespazio, are not necessarily those of INMARSAT and ESA.

1. INTRODUCTION

Next generations of multibeam Mobile Satellite Systems (MSS) are expected to extensively utilize on-board routing and processing functions, to cope with the expansion of user communities demanding MSS services and the possibility of efficiently routing the signals under variable traffic pattern conditions. The latter substantially calls for proper *flexibility* of the system to adapt itself to different scenarios, at the same time minimizing the on-board/on-ground hardware complexity to keep the service rendering cheap and user friendly.

The selection of the better suited multiple access technique is a crucial issue in that respect, associated to the selection of modulation/coding schemes and (possible) on-board regeneration, decoding and switching at channel level. Incidentally, this was indeed the rationale behind an ESA Request for Proposal ("Applicability of different on-board routing and processing techniques to a Mobile Satellite System"), issued in March 1990.

Three possible access schemes are addressed in this paper, on the basis of some studies carried out by the author on behalf of Telespazio. The first one concerns Satellite-Switched Frequency Division Multiple Access (SS-FDMA), and is treated in section 2. The second concerns Code Division Multiple Access (CDMA), addressed in section 3 - although Direct Sequence (DS) spread spectrum is generally envisaged in such systems, we report some investigations on Frequency Hopping (FH) techniques, popular in military applications but not really in satellite CDMA systems -. Section 4 briefly considers the Frequency-Addressable Beam (FAB) concept as applicable to the present context, although only some qualitative comments can be made at this stage since the bulk of investigations performed so far concerned Data Relay Satellites (DRS), where the extremely small number of envisaged users suggested a mixed frequency/phase address scheme. Finally, section 5 attempts to make some comparison among the mentioned access schemes, and some conclusions are drawn.

2. SS-FDMA

The viability of SS-FDMA techniques for MSSs has been investigated in several studies carried out in USA, Canada, Japan and Europe.

An INMARSAT study ("Study of Channelization and Routing for Inmarsat 3rd generation Satellite Transponder", Contract INM/218) was awarded in 1987 to Telespazio, aiming at identifying critical technologies for an SS-FDMA on-board router ("IF Processor") able to be utilised for either oceanic coverage (Atlantic, Pacific or Indian Ocean Region, AOR, POR, IOR) with the same repeater architecture.

The study provided an L-band baseline system handling up to about 500 channels 20 kHz spaced, derived by 18 kbps signal information rate + rate 3/4 FEC encoding and Offset-QPSK modulation. The $\pm$ 10° Field Of View (FOV) in the space-to-mobile user link was covered by 18 contoured/spot beams, whose shape is independent of the covered oceanic region and is determined by an RF/IF Beam Forming Network (BFN). The

L-band frequency plan is split into several, different bandwidth frequency slots for interbeam connectivity. 24 slots are allocated to AOR , 19 to IOR and 21 to POR. In total, (24+19+21=) 64 slots are envisaged; their widths span from 20 kHz (1 single channel) to 2180 kHz (109 channels) [1].

Fig.1 shows the principle architecture of the considered MSS repeater. In the C-band forward (outbound) link - a specular configuration characterizes the return (inbound) link -, the (global coverage) feeder link signal is frequency-converted to a convenient IF prior to feed the SS-FDMA router. This consists of an 18-output (=L-band beams) IF Processor, made up by a Filter Bank (FB) and a Channel-to-Beam Allocation (CBA) matrix. An IF/RF BFN - Beam Interpolation (hardwired) matrix + RF Combining Network - synthesizes 18 L-band beams by feeding a 19-horn 3.4-m reflector antenna, High Power Amplification (HPA) being provided by a smaller number of amplifiers.

Reconfigurability requirements of matching different traffic scenarios - i.e. oceanic regions by the same repeater architecture were met utilizing a "switched path filter" arrangement for the IF Processor [3], whereby a 26-way - instead of the theorethical (24+19+21=) 64-way - *fixed* FB is cascaded to a 26x18 CBA matrix. Interbeam connectivity is rearranged by properly activating the matrix cross-points.

The IF Processor specifications (detailed in Table 1) originated a survey of possible technologies, primarily for what concerns FB, which was afterwards designed utilizing a bank of 26 Surface Acoustic Wave (SAW) filters @ IF=30 MHz - 60 kHz minimum 0.5 dB bandwidth instead of the nominally required 20 kHz, definitely beyond the state-of-the-art with the required shape factor, 2.18 MHz maximum bandwidth - having less than 30 kHz 0.5/40 dB transition band.

TABLE 1. IF PROCESSOR SPECIFICATIONS

FILTERS	- 60-to-2180 kHz bandwidth required - 1 dB peak-to-peak (+-0.5 dB) maximum in-band ripple - 4 degrees peak-to-peak (+-2 degs) maximum in-band deviation from phase linearity - 40 dB minimum out-of-band (sidelobe) rejection - 30 kHz maximum (0.5-40 dB) transition bandwidth, at least for narrowband filters (up to some hundred kHz) - 60 kHz minimum (0.5 dB) bandwidth
SWITCH MATRIX	- substantial phase linearity - 50 dB minimum ON-OFF isolation per switch
PROGRAMMABLE MIXERS	- capability of mixing with any LO frequency within the comb of 1335 (10 kHz-spaced) components from 6.25 MHz to 19.6 MHz (each beam requires one such CW frequency-per-covered ocean region) - low phase noise

The very nice frequency response achieved with SAW filters is counterbalanced by the large number of taps (around 5,000) required to assure the steepness of the passband-to-stopband transition (in order to maximize the system spectral efficiency), dictating their physical size to be unusually large, 360 x 38 x 17 mm each packaged filter i.e. 14.2 x 1.5 x 0.7 inches. The overall IF Processor was designed by utilizing current 1 μm IC bipolar technology for the CBA switching matrix. The mass and size was 8.3 kg and 40 x 70 x 4 cm, respectively, whereas the power consumption was 10.9 W.

The study, formally completed in 1988, is being integrated by some follow-on activities at the University of Rome and Telespazio, aiming at investigating the possibility of implementing the SS-FDMA router with Digital Signal Processing (DSP) hardware ("Baseband Processor"), utilizing either "multistage" or "block" methods for channel demultiplexing [4]. This would also ease regeneration and *time* switching of channels, implementing *true* on-board processing (OBP) functions and direct-to-user (DTU) operation. Some promising results applicable to mobile satellite systems will be published shortly.

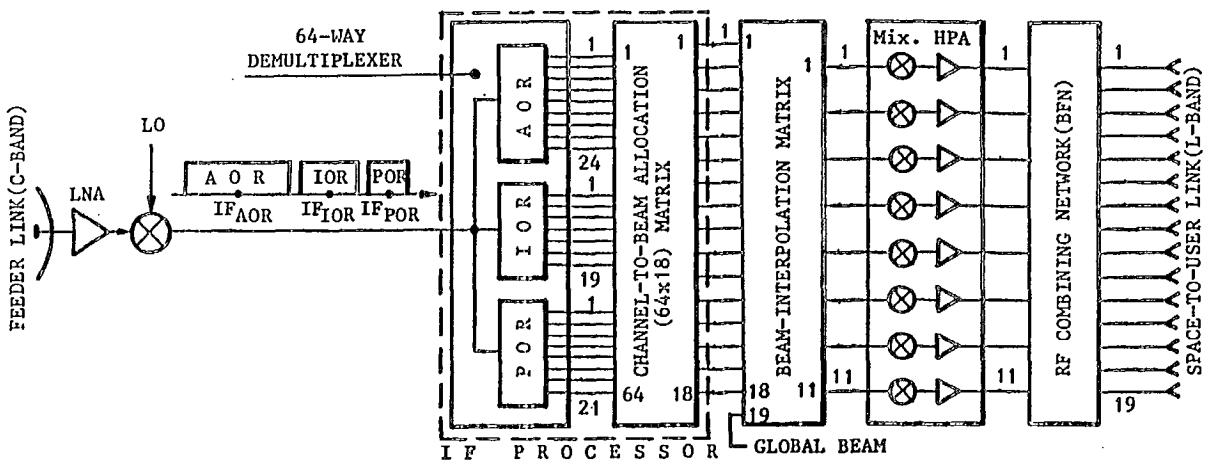


Figure 1. Forward link repeater with 64-way fixed-filter demultiplexing; AOR, IOR, POR are the demux portions selecting the pertinent bands

One interesting alternative to the mentioned architecture is based upon Surface Acoustic Wave (SAW) *chirp* Fourier transform (CFT) processors, that permit to turn the input FDMA channels into output TDM slots [4][5]. Subsequent processing (time-gating and inverse Fourier transform) permits reconfigurable filtering, enabling in principle channel-by-channel routing (not easily achievable by the use of filter banks implemented with any technology).

As a general remark, SS-FDMA is very efficient for what concerns interbeam connectivity, the well known drawbacks caused by non-linear intermodulation effects being at present fairly well balanced by the availability of space-qualified HPA linearizers and (inherently rather linear) Solid State Power Amplifiers (SSPA). Differently from several years ago, a little amount of amplifier back-off is now required (1-2 dB at most), which, together with the off-the-shelf availability of effective FEC codecs for the involved narrowband signals, permit the user terminals to be rather cheap and simple, with no synchronization constraints. On the other hand, although this multiple access technique is inherently "non regenerative" [3], the required on-board technologies are not trivial, and the system performances are substantially interference-limited, differently from the spread-spectrum approach addressed in the next section.

3. SPREAD SPECTRUM (CDMA)

Code Division Multiple Access (CDMA) techniques are inherently better suited in interference environments since they exploit spread spectrum peculiarities. "*Averaging*" techniques (e.g. Direct Sequence (DS) spread spectrum) reject (narrowband) interference by averaging them on a wide frequency range; "*avoidance*" methods (e.g. Frequency Hopping, FH) try to avoid that the frequency slots occupied by the useful signal and interference(s) overlap each other. Table 2 summarizes advantages and drawbacks of both techniques. In general, FH is less sensitive to *near-far* problems, although it is less (spectrum) efficient by roughly a factor of 2 than DS. However, this processing gain limitation is overshadowed by the possibility of spreading the signal(s) over much wider bandwidths, which in turn increases the system efficiency [6]. It must be underlined that, whenever bandwidth efficiency is used as a criterion to compare spread srectrum CDMA and SCPC/FDMA systems, the comparison greatly favours the latter (at least in global coverage systems), unless FEC coding is utilized or any peculiar interference-reduction technique is adopted (spot beam coverage, polarization discrimination, voice activation,...[8]). This is because the CDMA system capacity (M user) is limited by co-channel noise (*self* noise), that is the interference contribution by the other (M-1) users. It is evident that any factor that can reduce the self noise contribution results in an increase of spectral

	ADVANTAGES	DRAWBACKS
DS	* GOOD ANTIJAMMING PERFORMANCE * DETECTION HARD TO UNAUTHORIZED LISTENER (SECURE COMMUNICATION) * GOOD MULTIPATH DISCRIMINATION	* WIDEBAND CHANNEL WITH LITTLE PHASE DISTORTION * LONG ACQUISITION TIME, OR * HIGH PN CHIP RATE * NEAR-FAR PROBLEM
FH	* LARGE SPREADING POSSIBLE * AVOIDING SOME PART(S) OF THE RF SPECTRUM IS POSSIBLE * SHORTER ACQUISITION TIME THAN DS * REDUCED NEAR-FAR SENSITIVITY	* COMPLEX FREQUENCY SYNTHESIZER * ERROR CORRECTION REQUIRED * NON-COHERENT DETECTION GENERALLY TO BE USED

TABLE 2. DIRECT SEQUENCE (DS) VERSUS FREQUENCY HOPPING (FH) SPREAD SPECTRUM.

efficiency, i.e. a higher number of users can be handled by the system, the overall self noise power density being equivalent to that produced by a lower number of users.

CDMA techniques require code synchronization; the possibility of using a kind of network (code) synchronization based on distributing a master code in the forward (outbound) link has demonstrated improved system efficiency, resulting in rather simple and cheap user terminals [9].

An ESA contract to Telespazio ("Space Applications of Frequency Hopping Techniques", ESA/ESTEC Contract 7497/88), started in 1988 and still in progress to build-up a proof-of-concept piece of hardware, aims at assessing the possibility of using CDMA in several environments, including MSSs [10]. The possible utilization of FH as opposed to DS has been investigated and some interesting results derived on the system capacity in the two configurations in comparison with FDMA.

Basically, FH system performances are limited by the possible occurrence of collisions or "hits", resulting whenever the user instantaneous frequency hops on a frequency slot already occupied at that time by another user (packet *erasure*). The data transmitted in that "dwell" interval would then completely lost unless some proper FEC technique is utilised; however, the receiver tolerates a certain amount of (partial) overlapping [11], i.e. is capable of discriminating against partial hits provided that the interfering power (across the received frequency slot) is below a given threshold P (fig.2).

The error probability P_b for a given number of users (M) depends on the probability that, in the event of collision, the foresaid receiver threshold is exceeded. If we assume a spot-beam coverage, the interfering signals coming from different beams have envelopes weighted by the beam pattern, i.e. the related interference (self noise) results lower than in a global beam environment. This means that $P_{b,spot}(M) < P_{b,glob}(M)$, or, equivalently, $P_{b,spot}(M') = P_{b,glob}(M)$, where M'>M: due to the self noise reduction, the system can handle at the same error probability a higher number of users. For the same (self noise reduction) reason, signals transmitted on orthogonal polarizations permit to increase the spectral efficiency; in the case of voice activation

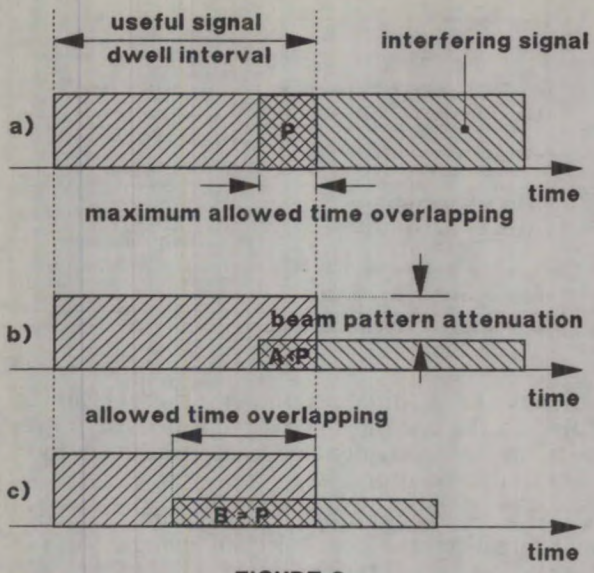


FIGURE 2.
Partial hits between (non-synchronized) users. a)global coverage; b) and c)spot beam coverage.

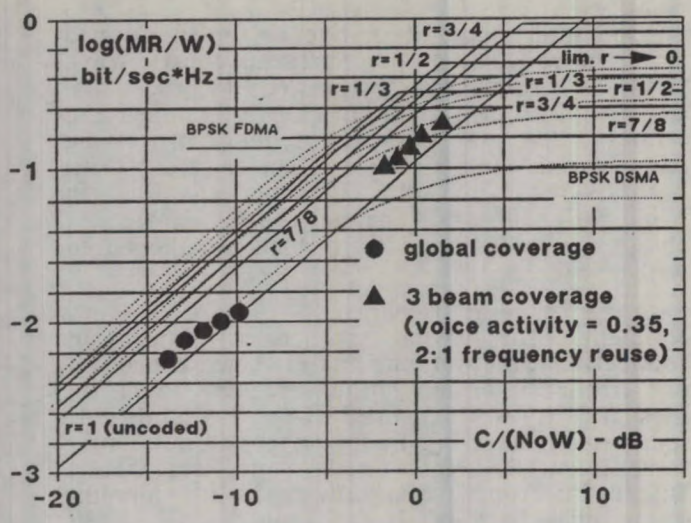


Fig.3. 16-FSK Slow FH Spectral Efficiency @ BER= 10^{-5} versus Carrier-to-Noise Ratio.

- only pilot tones are transmitted during non-active intervals to maintain synchronization -, the average number of simultaneous users is reduced by roughly a factor of 3 : 1 (=1/activity factor), which reduces the hit probability and, in turn, permits to increase the channel capacity.

For an M-user system, each user transmitting R bit/s in the presence of a rate-r FEC coding with N-FSK modulation, resulting in an individual (frequency bin) bandwidth W_b to be hopped across F slots (overall RF bandwidth $W=FW_b$), the spectral efficiency MR/W can be expressed as $MR/(FW_b) = M\log_2(N)/(FNr)$. On the other hand the carrier-to-noise ratio is given by $C/(N_oW) = (E_b/N_o)R/W$. Hence, for a given number of users (M), level of (FSK) modulations (N), frequency bins (F) and code rate (r), the spectral efficiency MR/W is derived, and the corresponding E_b/N_o (and thus $C/(N_oW)$) value to reach the objective error probability level can be calculated.

Some relevant results are plotted in fig.3 for a 16-FSK (*slow* FH) case - the spectral efficiency has been computed at $P_b=10^{-5}$, although 10^{-3} is sufficient in most MSS vocoders -, where some solid and dotted lines are drawn as a reference, indicating the performances of (BPSK) FDMA and DS-CDMA, respectively, in the presence of matched Reed-Solomon codes (rate 1/3-to-7/8). It can be seen that the presence of 3 spot beams, voice activation (gain about 3) and frequency re-use (2:1) by polarization discrimination substantially permits the system to handle about 3x3x2=18 times as many user with respect to absence of those features. However, at least for private user networks, CDMA techniques seem to require extremely large signalling burden to handle a multiple-beam environment, so that the advantages in terms of spectral efficiency should be probably traded-off versus network/system aspects.

FH satellite systems require some (limited) degree of complexity in the user terminals, although at present the development of fast synthesizers made up by Numerically Controlled Oscillators (NCO) permits to employ *fast* FH (several hops-per-symbol), that exhibits attractive system performances. It has to be pointed out, however, that DS spread spectrum techniques have considerably higher maturity than FH in multiple access systems, so that the real applicability of FH-CDMA will depend to a large extent on the technology advances in the years to come.

4. FREQUENCY ADDRESSABLE BEAMS

A separate ESA contract ("Study of an Experimental Payload for Ka Multiple Access Services", ESA/ESTEC Contract 7616/88) awarded to Telespazio in 1988 and still in progress concerning follow-on activities on MSSs, substantially aims at investigating the "frequency-addressable" beams concept [12] as applicable to both Data Relay Satellites (DRSs) and MSSs. According to it, a one-to-one relationship does exist between the channel *frequency* and the mobile terminal *location* (fig.4). This is achieved by means of suitable - and in principle simple - BFNs substantially composed by arrays of tapped delay lines (loaded waveguides). The system is designed such as to continuously track the user(s) throughout the motion, by sweeping the carrier frequency across the allocated bandwidth. This causes the necessary band being much wider than in principle required, programmable VCOs being called (at least) on-ground to count for the instantaneous position(=frequency) of the user within FOV. The user terminal is therefore not too simple, although no synchronization problems exist and standard FDMA equipment can be utilized.

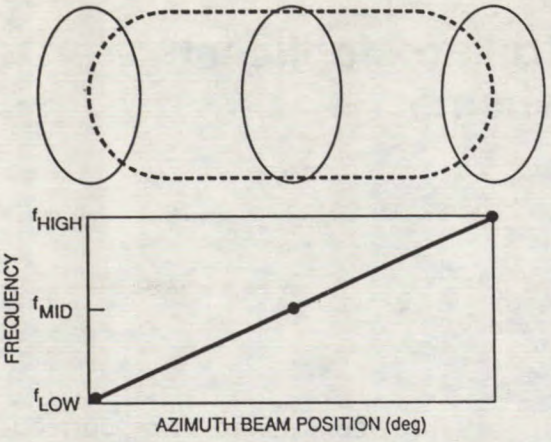


Fig. 4 Frequency-address beam arrangement

PARAMETER	SS-FDMA	FH-CDMA	FAB
Spectral Efficiency	GOOD	GOOD	FAIR
(Co-Channel) Interference Rejection	POOR	GOOD	POOR
Synchronization requirements	NO	YES	NO
Terminal Cost/Size	LOW	LOW-MEDIUM	MEDIUM
Space Segment Complexity	MEDIUM	LOW	MEDIUM
Flexibility/Reconfigurability	GOOD	FAIR	FAIR

TABLE 3. SS-FDMA versus FH-CDMA and Frequency Address Beams

5. CONCLUSIONS

Table 3 attempts to make a comparison among the three multiple access systems outlined in this paper. SS-FDMA represents maybe a good compromise between flexibility and cost, whilst CDMA is more viable in interference-limited environments. Frequency-addressable beam concepts still require some development efforts to optimize performances.

REFERENCES

[1] **F.Ananasso, F.Delli Priscoli**: "Non-Regenerative On-Board Processing for Multi-Beam Mobile Satellite Systems", IEEE/ICC-88, Philadelphia (PA, USA), June 1988, paper 16.7.1

[2] **F.Ananasso, J.M.Deacon**: "Design of SAW Filters in SS-FDMA Routers for Mobile Satellite Systems", International Journal of Satellite Communications, Vol.7, August 1989, p.321

[3] **F.Ananasso, P.de Santis**: "On-board Technologies for User-oriented SS-FDMA Satellite Systems", IEEE/ICC-87, Seattle (WA, USA), June 1987

[4] **F.Ananasso**: "Low Bit Rate on-board Demultiplexing and Demodulation", IEEE/ICC-89, Boston (MA, USA), June 1989

[5] **P.M.Bakken, K.Grythe, A.Ronnekleiv**: "The On-board FROBE SAW/Digital Signal Processor", 8th International Conference on Digital Satellite Communications (ICDSC-8), Pointe-à-Pitre (Guadeloupe), April 1989

[6] **A.J.Viterbi**: "Spread Spectrum Communications - Myths and Realities", IEEE Communications Magazine, Vol.17, No.5, May 1979, pp.11-18

[7] **A.J.Viterbi**: "When Not to Spread Spectrum - a Sequel", IEEE Communications Magazine, Vol.23, No.4, April 1985, pp.12-17

[8] **I.M.Jacobs** et al.: "Comparison of CDMA and FDMA for the MobileStar^{sm} System", 1988 International Mobile Satellite Conference (IMSC-88), Pasadena (CA, USA), May 1988

[9] **R.De Gaudenzi, R.Viola**: "High Efficiency Voice Activated CDMA Mobile Communication System Based on Master Code Synchronization", IEEE/GLOBECOM-89, Dallas (TX, USA), November 1989

[10] **F.Ananasso** et al.: "Satellite Applications of Spread Spectrum Frequency Hopping Techniques", IEEE/GLOBECOM-89, Dallas (TX, USA), November 1989

[11] **J.Wieseltier, A.Ephremides**: "Discrimination against Partial Overlapping Interference", IEEE Transactions on Communications, Vol.COM-34, No.2, February 1987

[12] **J.D.Thompson**: "Frequency Addressable Beams for Mobile Communications", AIAA 12th Communications Satellite Systems Conference, Arlington (VA, USA), March 1988

Simulation of the Australian Mobilesat Signalling Scheme

Mushfiqur Rahman
Australia Telecom Research Laboratories
770 Blackburn Avenue
Clayton, Victoria 3168, Australia
Phone: 03-541-6809
FAX: 03-543-3339

ABSTRACT

The proposed Australian Mobilesat system will provide a range of circuit switched voice/data services using the B series satellites. The reliability of the signalling scheme between the Network Management Station (NMS) and the mobile terminal (MT) is of critical importance to the performance of the overall system. In this paper we present simulation results of the performance of the signalling scheme under various channel conditions and coding schemes.

1. INTRODUCTION

The outbound signalling channel of the proposed Mobilesat system is a TDM channel designed to carry information from the Network Management Station (NMS) to the mobile terminals. All signalling messages,both inbound and outbound , will be formatted into uniform signalling units (SU) of 96 bits (12 bytes). Each SU will include 16 check bits (the last two bytes) for error detection. The 16-bit cyclic redundancy check (CRC) sequence for error detection is based on CCITT recommendation X.25. The details are available in the AUSSAT Mobilesat specifications [1].

The reliability of the outbound signalling scheme is of critical importance to the mobilesat system, since all call set-up procedures between mobile terminals (and a number of other control and network management functions) are accomplished through successful transmission and reception of SU's between the Network Management Station and the Mobile Terminals. A two tier approach has been adopted to ensure high reliability.

(i) The 96-bit SU is encoded into a 128 bit coded signalling unit. The additional 32 bits provide forward error correction capability.

(ii) The coded SU is repeated three times within a superframe structure consisting of 72 coded SU's.

The repetition interval between two signalling units is 24 blocks which ensures that the probability of the repeated SU being affected by the same fade is very small, since they are sufficiently apart from each other. A SU is deemed to be correctly transmitted if one or more of the three transmissions is successful. The overall TDM super frame format is depicted in figure 1 below. Within a frame length of 24 blocks two blocks are used for bulletin board information (BB) and there is a 32 bit unique word (UW) preceding a subframe of 8 blocks.

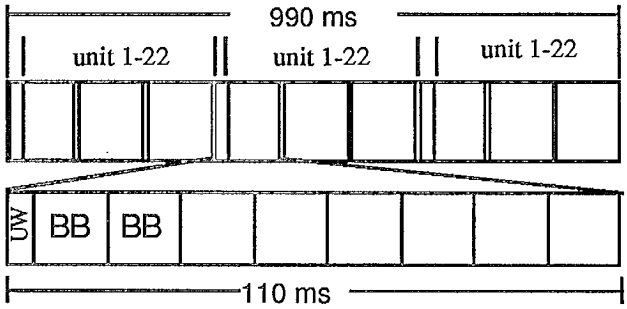


Fig 1: TDM Super Frame Format

In this paper we examine the performance of the outbound signalling scheme under Australian propagation conditions and show that the proposed scheme is viable under those conditions.

2. CHANNEL MODELLING

The channel is characterized by random variations of the received signal power, due to

the motion of the mobile unit. The propagation loss, with respect to unimpaired line of sight reception is a function of the terrain and the vegetation around the mobile unit in the direction of the satellite. Therefore, in order to simulate the output of the demodulator, the nominal signal power with respect to line of sight, has to be modified to take into account the effect of shadowing due to vegetation.

Telecom Research Labs carried out a number of propagation studies [2,3] to examine the effect of shadowing on the received signal power. The study was carried out on different road sections in the suburbs of Melbourne, with 35% and 85% vegetation density around them. These sections of the test route will henceforth be referred to as Channel-1 and Channel-2. Received signals were recorded at 1500 Hz and expressed in dB's with respect to line of sight. Figure 2 shows a plot for Channel 2.

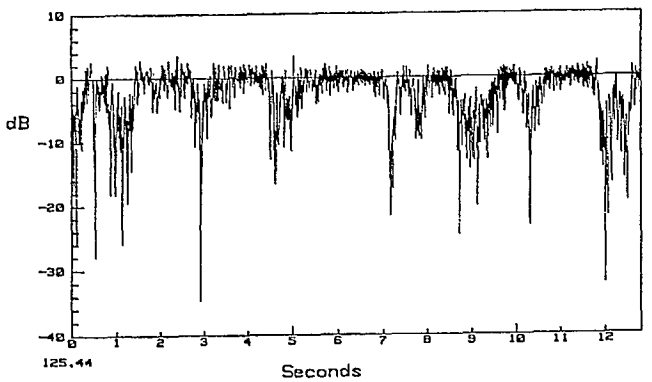


Fig 2. Effect of shadowing on received signal

Modelling of the channel was accomplished by assuming a nominal value of E_S/N_O, E_S = received symbol energy and N_0 = thermal noise power spectral density. The received signal energy was then modified on a bit by bit basis, by the additional loss (or gain) in dB's due to shadowing or multipath. From the resultant modified value of E_S/N_O, the bit error probability was calculated. It should be noted that the propagation data was sampled at 1500 Hz, whereas the signalling rate of the outbound channel is 9600 Hz. A simple fixed interpolation was used, in which each propagation data was held constant for 6 consecutive received symbols, giving an effective rate of 9000 Hz, which is close to the actual signalling rate.

3. CODING SCHEMES

3.1. Reed-Solomon Code

Reed-Solomon Codes are a class of non-binary codes defined over the Galois Field. RS codes when viewed as binary codes, ie. when the q-ary symbol of the code is expressed as a binary vector of the appropriate length, have some burst error correcting capability. The SU's are likely to be subjected to burst errors on account of the shadowing effect of vegetation. and therefore RS codes were logical candidates for providing error correction.Also recalling that there are 96 bits of data in a SU and they are required to be encoded into 128 bits of coded data, the rate of the encoder should be 3/4. Therefore a suitable candidate code is the (16,12) RS code defined over GF(8). The code alphabet has 256 symbols and thus every code symbol can be represented by a unique 8-bit binary vector. Therefore the (16,12) RS code over GF(8) can be viewed as a (128,96) binary code. From an error correcting point of view, the code can correct any two symbol errors. In terms of binary errors, this can be any two random errors, affecting two separate symbols; the code is also capable of correcting larger number of errors if they occur in bursts, the maximum burst length being 16 when they span exactly two symbols.

3.2. Convolutional Code

Simulation was also carried out for rate 3/4 convolutional codes, with constraint length = 7. The rate 3/4 code was obtained from rate 1/2 codes by puncturing (deleting) 2 output bits from every 6 output bits, corresponding to 3 input bits.

The generator polynomial for the rate 1/2 code, constraint length = 7 are:

G1: $1 + X^2 + X^3 + X^5 + X^6$

G2: $1 + X + X^2 + X^3 + X^6$

In the rate 1/2 code three consecutive input bits generate six output bits according to the following input/output relationship:

Input bit time:	1	2	3
Output Sequence:	G1 G2	G1 G2	G1 G2

The rate 3/4 coded data sequence is obtained by deleting 2 bits from each block of 6 output bits from the rate 1/2 convolutional encoder (ie. G2 from input time 2 and G1 from input time 3.

4. RESULTS

4.1. Throughput

Simulation of the outbound signalling scheme was carried out for both of the coding schemes described in the previous section. The channel was modelled with the stored propagation data as discussed in section 1. Figures 3-8 summarize these results and provide a basis for comparison between the two coding schemes.

Figures 3-4 show the performance of the (16,12) RS code for Channel-1 and Channel-2. Without repetition (1-shot), at E_s/N_0 = 12 dB,throughput reaches only 96.1% for the sparsely vegetated Channel-1 and 89.5% for the more heavily vegetated Channel-2. This clearly suggests that the shadowing effect of vegetation is far too severe for the Reed-Solomon code to be effective as a burst error-correcting code. An analysis of the propagation data reveals that most fade durations lie between 10 - 100 ms which implies that if a SU falls within a fade, the number of errors would be more than the maximum error correcting capability of the code. Fig.4 shows the effect of repetition, where each SU is repeated three times, the repetition interval being 24 SU's - conforming to the superframe format of 72 SU's as in AUSSAT specifications. A significant improvement in throughput, as defined by the probability of at least one successful transmission out of three, was achieved with this repetition scheme, as was expected.

Since the RS code functions primarily as a random error-correcting code, unable to cope with the long error bursts, it is logical to use a better random error correcting code with the same level of redundancy. Figures 5-6 show the improvement in throughput when a 3/4 rate convolutional code with constraint length 7 was used. It is quite apparent that the convolutional code offers significant advantage over the Reed-Solomon code. For the case of relatively heavy shadowing (85%) as in Channel-2 this enhanced performance will have important implications

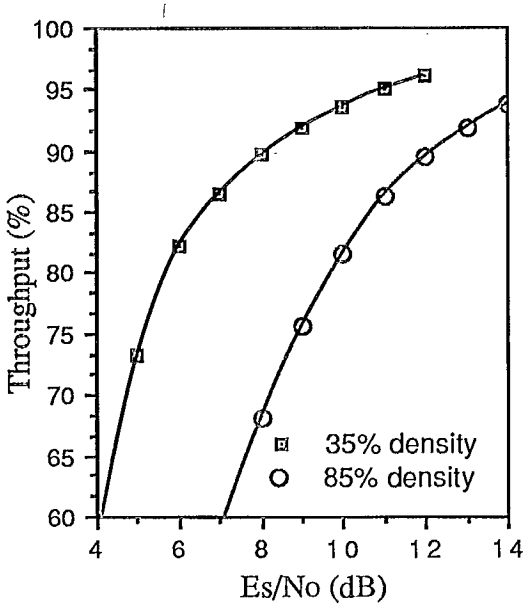


Fig3: Throughput for RS code - 1 shot

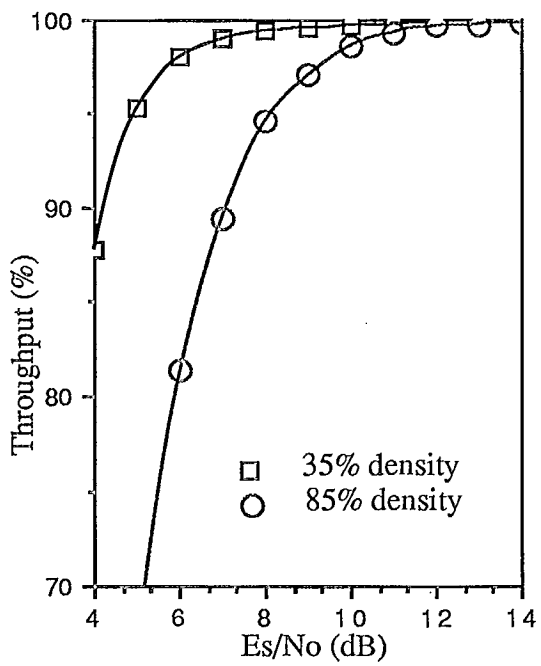


Fig 4. Throughput for RS code - 3 shot

with regard to the quality of service at the customer level. As an example, at E_s/N_0 = 8 dB,

RS coding with 3 repeats achieves a throughput of 94.5%, while convolutional coding achieves a

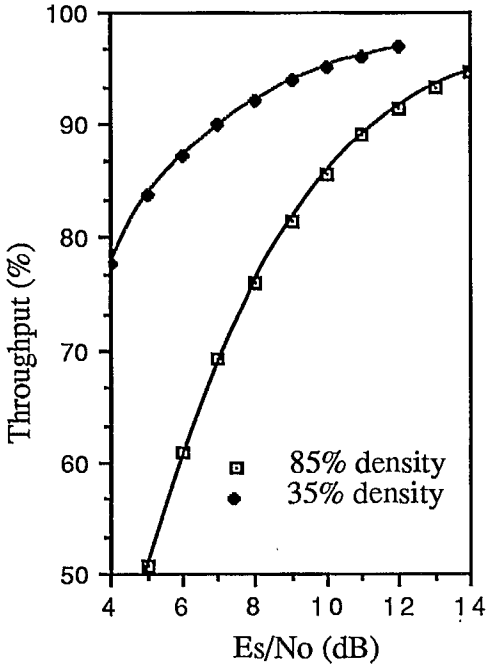


Fig 5: Throughput for Convolutional code 1shot

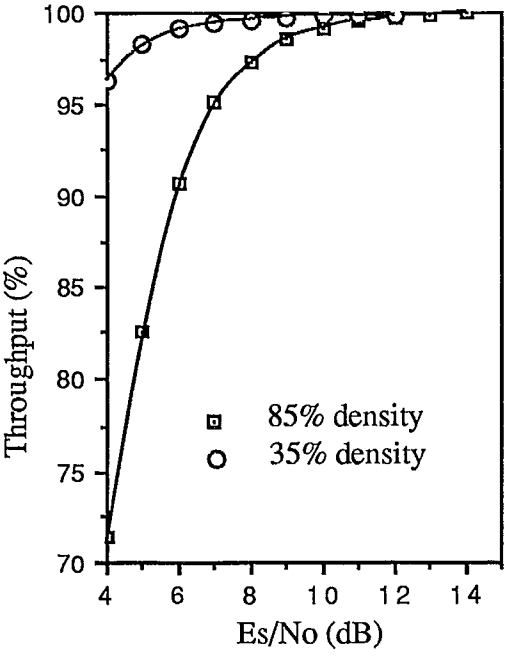


Fig 6: Throughput for convolutional code 3 shot

throughput of 97.3%. The probability of making a successful call will of course depend to a great extent on the reliability of the outbound signalling scheme as measured by these percentage throughputs. It may be observed that even for 85% vegetation density, a fairly high reliability can be achieved at moderate values of E_s/N_0. At 10 dB a reliability of over 99% can be achieved, with 3 repeats.

4.2. Capacity of the outbound channel

An altogether different viewpoint of the performance can be examined in terms of the capacity of the outbound channel, measured in terms of the number of good packets per second. It is obvious that repeating the SU's 3 times increases their reliability and hence increases service quality. On the other hand, repetition decreases the effective number of packets/sec which the outbound channel can handle. This reduced capacity will ultimately affect the probability of blocking, which in turn decreases service quality. Figure 7 shows packets/sec as a function of E_s/N_0 for both schemes, 3-shot and 1-shot, both using 3/4 rate convolutional coding in conjunction with Viterbi decoding. With 3 repeats, the maximum capacity is 21.8 packets/sec while without any repeats the maximum capacity is 65.45 packets/sec. The actual capacity is a little less because in a frame spanning 24 SU's, two slots are used for bulletin board information. As the number of mobile terminals increases with a corresponding increase in total traffic, the lower packet handling capacity of the repeat scheme may in fact cause a deterioration of the overall network performance. However this situation can be remedied by providing an additional outbound channel.

4.3. Performance with 2 repeats:

It is interesting to examine the performance of the signalling scheme with 2 repeats instead of 3. As expected, the reliability measured in terms of percentage throughput decreases as shown in figure 8. However, it should be noted that throughput reaches above 90% at $E_s/N_0 = 8$ dB for Channel-2 (85% vegetation). In terms of capacity, the 2-repeat scheme is approximately 50% better than the 3-repeat scheme.

5. CONCLUSIONS

The outbound signalling scheme as envisaged in the AUSSAT specifications can perform quite

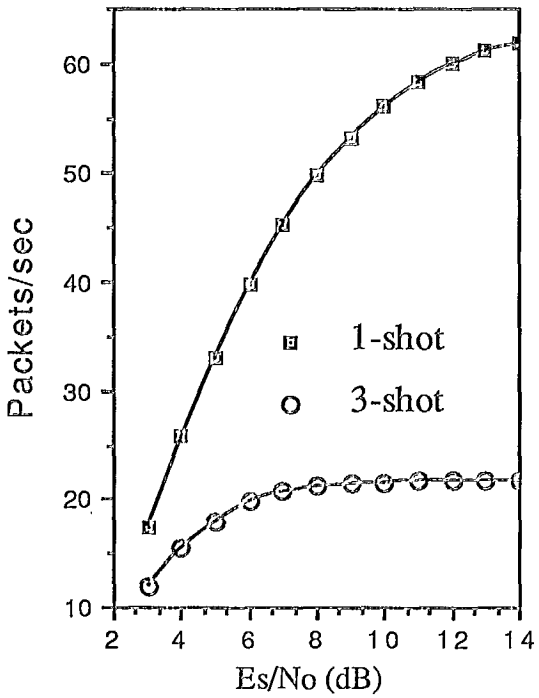


Fig 7: Capacity of the outbound channel

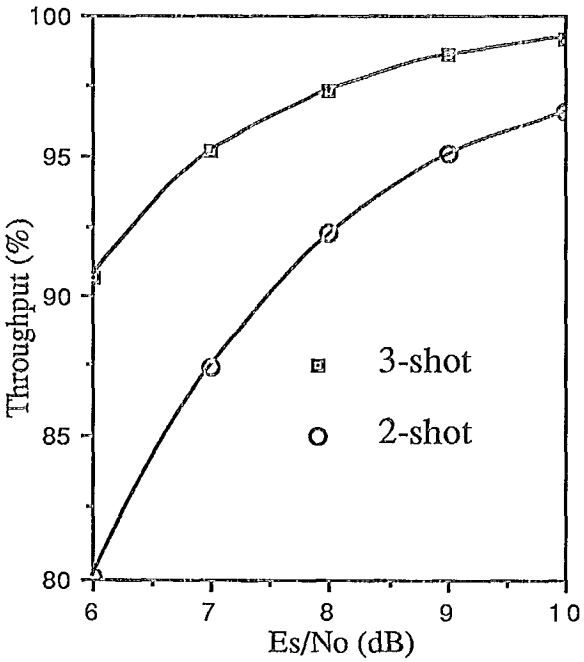


Fig 8. Throughput for a 2 - repeat scheme

satisfactorily in terms of reliability of the SU's. Convolutional coding, in conjunction with 3 repeats can ensure 99% reliability at around 10 dB for 85% road-side vegetation. The AUSSAT specifications calls for a nominal value of 12 dB for the outbound signalling channel and therefore the signalling scheme will be very reliable even in high density vegetation areas.

Convolutional coding outperforms the RS code in all cases. With soft decision decoding, there should be another 1-2 dB improvement. Because of the duration of bursts which are typically greater than 10 ms or more, use of a burst error correcting code is not helpful.

For light shadowing conditions, a 2-repeat scheme can provide an acceptable level of throughput (>99%) and yet provide 50% more capacity. Thus under heavy traffic conditions the NMS may switch to a 2-repeat scheme to cope with the increased traffic.

6. REFERENCES

1. Mobilesat Technical Specifications, AUSSAT Mobilesat, September 12, 1989.
2. A.J. Bundrock and R.A. Harvey, "Propagation Measurements for an Australian Land Mobile-Satellite System", Australian Telecommunication Research, vol 23, Number 1, 1989.
3. I.T. Hawkes, "Land Mobile Satellite Transmission Measurements at 1500 Hz", Section Report, Radio & Satellite Network Section, Australia Telecom Research Laboratories, 1989.

Random Access to Mobile Networks with Advanced Error Correction

Michael Dippold
DLR, German Aerospace Research
Institute for Communications
D-8031 Weßling, F.R.G.
FAX ++8153 28243

ABSTRACT

A random access scheme for unreliable data channels is investigated in conjunction with an adaptive Hybrid-II ARQ scheme using RCPC codes FEC. A simple scheme with fixed frame length and equal slot sizes is chosen and reservation is implicit by the first packet transmitted randomly in a free slot, similar to Reservation Aloha. This allows the further transmission of redundancy if the last decoding attempt failed. Results show that a high channel utilization and superior throughput can be achieved with this scheme that shows a quite low implementation complexity. For the example of an interleaved Rayleigh channel and soft decision utilization and mean delay are calculated. An utilization of 40% may be achieved for a frame with the number of slots being equal to half the station number under high traffic load. The effects of feedback channel errors and some countermeasures are discussed.

I. INTRODUCTION

The efficiency of channel access techniques for mobile services is degraded by channel impairments like multipath propagation and shadowing. More complicated reservation schemes show a faster degradation than the simpler random access methods under these conditions. Many recent publications are devoted to this problem. For example the improved efficiency of Slotted Aloha through the capture effect was studied in [1] for a Rayleigh channel and in [2] for the landmobile satellite channel. In [3] the effect of capture with a tree collision resolution algorithm was analyzed for obstacles between transmitter and receiver. The performance of a reservation scheme was compared to that of Slotted Aloha in [4].

In most of the investigations it was assumed that a packet has to be repeated if the detection failed due to bit errors. This basicly is the mechanism of type-I or type-II hybrid ARQ (Automatic Repeat Request). However, especially for time varying channel quality like in mobile applications it has been shown that more effective schemes exist, e.g. [5]. Essentially they are based on the idea of successive parity transmission and usage of all the information of the previous unsuccessful decoding attempts. In general this increases the throughput between transmitter and receiver. The criterion of point-to-point throughput used in ARQ analysis is sometimes not useful in a multiuser environment in connection with random access. Here channel utilization is an usual measure of system quality.

We investigate the use of such an advanced Hybrid ARQ scheme in the random access multiuser application for mobile data transmission. It is applicable for bursty data transmission or random access reservations for a SCPC voice channel. Although access schemes with fixed slot sizes are better to implement and work more reliable, the adaptive hybrid ARQ methods perform best, if redundancy can be transmitted in very small portions. But also with equal packet sizes their main advantage is the use of all formerly transmitted data for an actual decoding attempt in the receiver. Because the receiver stores all the data received in unsuccessful detection attempts until an ack-

nowledgement can be issued, even a one packet message becomes multiple packets in the case of packet errors.

For this reason we will regard an implicit reservation random access scheme instead of a pure Slotted Aloha technique. Details will be discussed in section 2. The ARQ scheme considered is based on the adaptive ARQ/FEC (Forward Error Correction) with RCPC-Codes and Viterbi decoding presented by Hagenauer in [6]. It is briefly reviewed in section III and a model for the combined performance of this method with the random access is presented. Section IV gives numerical results for the example of the Rayleigh fading channel.

II. RANDOM ACCESS WITH RESERVATION

The most effective way to use adaptive ARQ protocols is to allow arbitrary lengths of the various parity packets. However, variable slot size increases the complexity and hence network overhead and error possibilities of the access scheme. Therefore we consider only a fixed equal slot length. Slots are organized in a fixed size frame for the uplink from the mobile to the center (satellite). On the return link acknowledgements (ACK) and reservation confirmation are issued. The frame length exceeds the signal round trip time. Every station is allowed to transmit messages with variable length and hence variable packet number but may only use one slot per frame. Every packet is coded according to the coding rule and data and redundancy is stored for every packet until an ACK was received. We assume the following access algorithm:

On the return link for every data carrying slot a feedback is received unless a collision took place (it is reasonable to assume that the receiver distinguishes between collision and packet error with high probability). Any station to start transmission tries one not reserved slot. There are the following feedback signals (no feedback in the case of garbled data after a collision):

- Packet error and reservation for a further attempt ("RES")
- Acknowledgement for a successfully decoded packet and reservation for a next packet ("AR")
- Acknowledgement for the last packet and release of this slot ("ACK")

If Collision:	Try again in a free slot according to Slotted Aloha; no feedback.
If Success:	Done (1-packet message); feedback ACK. This slot is reserved in the next frame for the next packet (m-packet message); feedback AR.
If error:	Same slot in next frame reserved for transmission of additional redundancy; feedback RES.

Every packet carries an indicator whether more information packets are to follow or not. If decoding was not yet successful it is clear that more redundancy will be sent and hence the automatic reservation. This scheme is basicly the classical reservation Aloha introduced by [7] and is illustrated in Fig. 1. If all the priority is transmitted and decoding was not yet successful the station will try from the beginning and join the set of new arrivals. We do not consider code combining methods here.

Feedback Errors: We assume that no feedback error occurs as data may be protected by strong codes. However, there is a certain probability for landmobiles in the L-Band or shorter wavelength bands that shadowing occurs and no feedback signal is received.

- A station misses the AR or RES feedback: After its own transmission this leads to a new access attempt and the reserved slot is wasted. But if reservations for other stations are not heard, this may cause a collision, if the station tries a new access in this slot in the next frame. This problem is alleviated if the center transmits a reservation plan for the next frame. Only slots within a round trip time cannot be indicated.
- A station misses an ACK: If it was an ACK for its own transmission it keeps transmitting in a released slot. This also

can cause collision and the station is not informed about its own success. A collision indicator instead of no feedback in this case can be used. Missing ACK's for other stations means that the station in mind considers this slot as reserved although it has already been released.

All these errors result in a somewhat lower throughput and higher delay but should not be a problem if the probability for their appearance can be kept low.

Analysis

For the analysis we assume a stationary system with the following parameters:

F number of slots in every frame
H average number of packets per message including redundancy transmission
R average number of slots, which are reserved for transmission in a frame
M number of users in the system.

A station can only start its transmission by accessing a free slot. We model the complete system by a Markov-chain where the number of reserved slots in a frame is the state variable r. According to [8] we assume that all non-reserved F-r slots may be accessed in a Slotted Aloha manner with a channel utilization U_0. A station can hold only one reserved slot per frame. Any of the M-r stations not transmitting in the current frame starts transmission in the next frame with probability U_0. Any of the r active stations in the actual frame will have finished in the next frame with probability 1/H. The average message length H depends on the channel errors and will be calculated in the next section. In this model we can describe the probability that the number of reserved slots r is changed by i in the next slot

$(-r \le i \le F-r)$ by:

$$p_{r,\,r+i} = \sum_{l=0}^{r} \left[\binom{r}{l} (1/H)^{l} (1-1/H)^{r-l} \cdot \right. \tag{1}$$

$$\left. \cdot \binom{M-r}{l+i} U_0^{\,l+i} (1-U_0)^{M-r-l-i} \right]$$

Solving the linear equation system

$$\begin{aligned} &\ldots \\ &P_r = p_{0,r} P_0 + p_{1,r} P_1 \ldots + p_{F,r} P_F \\ &\ldots \end{aligned} \qquad 0 \le r \le F \tag{2}$$

we get the stationary probability distribution for the number of reserved slots with the mean value $R = E\{r\}$.

The channel utilization is the ratio of the average number of successful information packets per packet time. On the average m/H packets out of H are information, the other are redundancy. Related to the frame length F this results in:

$$U = \frac{R\, m/H}{F} . \tag{3}$$

The delay D for a message from the time a station starts to access the channel until an acknowledgement is received has several contributions:

$$D = D_1 + D_H + T_{RT} + 1.5 \tag{4}$$

D_1 is the delay until the first packet has been successfully placed into a slot,
$D_H = (H-1)\,F$ is the time in slots until the last packet has been transmitted. Additionally after T_{RT}, the roundtrip time, an acknowledge can be received. For D_1 we assume that a station has to wait F/(F-R) slots on the average for a free slot, which is then accessed with a probability of success equal to U_S. As the whole system depends on the user activity an optimization would require a joint analysis of both the Slotted Aloha and the reservation system. It can be shown from the Markov analysis of Slotted Aloha, eg. [2] that the probability of success U_S for a station seeking access to the system is proportional to $(1-\sigma)^{-R}$, with σ being the arrival rate in a station in packets/slot. After K free slots on the average Slotted Aloha allows another attempt after a collision. Under the assumption of independent locations of free and reserved slots in a frame K free slots is equivalent to a period of $T_K = K\,F/(F-R)$ slots in total. Under these assumptions D_1

results in:

$$D_1 = \frac{F}{F-R} U_S + U_S(1-U_S)\, T_K + U_S(1-U_S)^2\, 2T_K + \ldots = \frac{F}{F-R} U_S + T_K \frac{1-U_S}{U_S} \quad (5)$$

U_S indicates the probability that an access in the Slotted Aloha mode is successful (taken approximately equal both in the new arrival and repetition case here).

III. AVERAGE MESSAGE LENGTH

A message may consist of m packets each of length of one slot. Every packet is coded with the Punctured Convolutional Code of rate 1/N and stored in the transmitter. According to the scheme presented in [6] this results in m·N packets that may be transmitted at most. Decoding is performed after every packet and only if not successful the next portion is transmitted in the next frame. Due to the equal length constraint the coderate is decreasing in larger steps than necessary by the RCPC code alone. For example with the puncture period P a 1-packet message (m=1) exhibits the following successive coderates: P/(P+P), P/(P+2P)..., P/(NP). However, for m > 1 a somewhat finer gradation can be allowed by interleaving the different information packets and their redundancy packets. But this must be known in the receiver. The superior performance of this adaptive scheme compared to conventional Hybrid-ARQ was recently shown in [9].

If we imagine the total coded message stored in n·m·P columns, the transmission proceeds in steps of ΔI additional columns after each decoding attempt. In the following we constrain ourselves to the case ΔI=P and m=1, where P is the puncture period. The average number of transmitted columns I_{AV} is then according to [6]:

$$I_{AV} = \sum_{k=1}^{N} kP\,(1- p_e(k\Delta I)) \prod_{i=0}^{k-1} p_e(i\Delta I) + N P \prod_{i=0}^{N} p_e(i\Delta I) \quad (6)$$

With $p_e(k\Delta I)$ we express the probability that a packet cannot be decoded with kΔI columns of bits already received. Of course with p_e=0 in the error free case only P columns have to be transmitted and no additional redundancy is required.

In the transmitter a buffer capacity of m·N·Lp bits is required for a m-packet message. The demand for buffer is more critical in the receiver as in the extreme case (all slots reserved, all stations transmit full code) it increases with the frame length F: N·m·Lp·F. On the average only m·H·R·Lp bits of buffer are needed. An overflow would lead to a loss of already received information and will only effect U and D. Therefore a tradeoff between buffer space and utilization and delay is possible.

IV RESULTS

For the numerical evaluation of eq. 6 a Rayleigh channel with perfect channel state estimation and soft decision is chosen as an example. The packet error probability p_e was calculated using the figures in [6]. With $p_e(P)$=1 (certain packet error after the first transmission) the average message length $H = I_{AV}/P$ approaches H = 2 for large E_s/N_o. The dependence of H on E_s/N_o is shown in fig. 2. However, an inner code with some error correction facility instead of a CRC in the first packet would reduce $p_e(P)$ below 1 and H below 2 for the price of some additional redundancy. For comparison the mean message length H is also drawn in fig. 2 for the case ΔI=2. The advantage of a transmission of redundancy in smaller units than P becomes beneficial for higher E_s/N_o as the coderate may be decreased

in finer steps. This would be possible in point-to-point connections or with unequal slot sizes in a multiaccess scheme. The following examples show the performance of the model from section II for the following figures: user activity per station $\sigma = 0.01$ messages/slot, operation point of Slotted Aloha $U_O = 0.33$, number of stations M = 30, round trip time $T_{RT} = 2.4$ slots, no feedback errors. In Fig. 3 the channel utilization and the delay as function of the mean number of packets H are depicted (fixed frame length F = 20 , number of packets per uncoded message m = 1). The delay D increases about linear with H, while the channel utilization U decreases. For higher m the utilization would be improved for a given H as is typical for reservation schemes.

A sensitive design parameter of the multiple access scheme described is the frame length. It has to be optimized in dependence of the number of stations M, the user activity σ and the channel quality (represented by H). Fig. 4 illustrates the mean delay D and channel utilization U as a function of the number of slots/frame F. For longer frames U decreases as too many slots stay idle. But with higher utilization the delay increases to infinity as all slots are reserved and new arrivals are blocked. With very high F the delay becomes larger again, as it takes H frames until a message is completed. However, long frames in the order of M are not sensible as fixed TDMA would be simpler and free of collisions.

V CONCLUSIONS

It has been shown that adaptive Hybrid ARQ schemes can be utilized in multiaccess environments. An implicit reservation scheme was chosen that is relatively simple to implement. Channel utilization and expected delay were calculated for the hybrid ARQ/FEC error correction with RCPC codes on a Rayleigh fading channel. The effectiveness of the adaptive ARQ is somewhat reduced by the constraint of equal size slots in the multiple access scheme. However, this disadvantage can be alleviated by the use of a high rate error correcting inner code in the hybrid ARQ scheme, or by reducing slot size to achieve multiple packets per uncoded message. The frame length is an important design parameter. Its influence on utilization and delay has been investigated. Further optimization of the combined adaptive ARQ and multiaccess scheme is required with respect to slot size, user activity and different channel characteristics.

REFERENCES

1. **Habbab, I. and Kavehrad, M. and Sundberg, C.E.** 1989. ALOHA with Capture Over Slow and Fast Fading Radio Channels with Coding and Diversity, IEEE Journ. Select. Areas in Commun., SAC 7, 1, pp. 79-88.
2. Dippold, M. 1989. Performance of a Slotted Aloha Access Scheme for Packet Transmission on the Land-Mobile Satellite Channel. Proc. 8th Intern. Conf. on Dig. Satell. Comm. ICDSC-8, pp. 477-481.
3. **Cidon, I.; Kodesh, H. and Sidi, M.** 1988. Erasure, Capture and Random Power Level Selection in Multiple-Access Systems, IEEE Transact. Commun. COM-36, 3, pp. 263-271.
4. **Böttcher, A. and Dippold, M.** 1989. A Comparison of Two Multiple Access Schemes for Satellite Networks with Error-Prone Transmission Channels. Proc. 1st Europ. Conference on Satellite Commun., ECSC-1, pp. 283-293.
5. **Hagenauer, J.** 1987. Hybrid ARQ/FEC Protocols on Fading Channels using Rate Compatible Punctured Convolutional Codes. Proc. ICC 1987, 21.4.
6. **Hagenauer, J.** 1988. Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications. IEEE Transact. Communication COM 36, 4, pp. 389-400
7. **Crowther, W. et al.** 1973. A System for Broadcast Communication: Reservation Aloha. Proc. Hawaii 6th Internat. Conference on System Science.
8. **Lam, S.S.** 1980. Packet Broadcast Networks - A Performance Analysis of the R-Aloha Protocol. IEEE Transact. on Computers 29, 7, pp. 596-603.
9. **Niinomi, T. et al.** 1990. Selective Repeat Type-II Hybrid ARQ/FEC Scheme Using Rate-Compatible Punctured Convolutional Code. To be published in Proc. ICC, Atlanta, 1990.

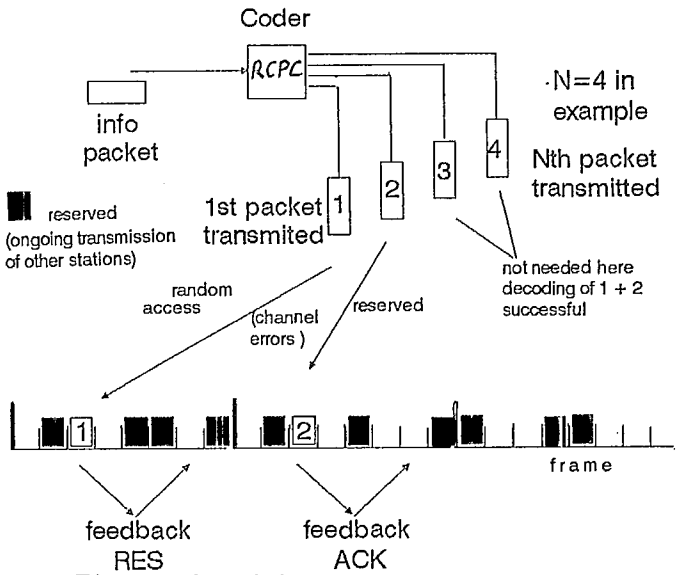


Fig. 1: Implicit reservation scheme

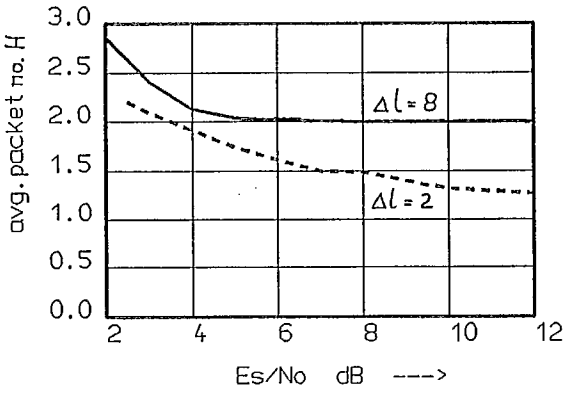


Fig. 2: Mean message length in packets (m=1, Rayleigh channel, soft decision, RCPC-Code N=4, P=8).

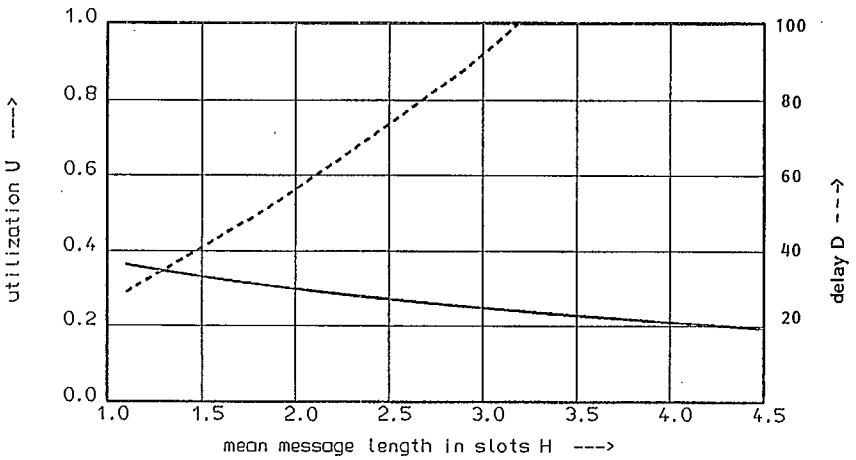


Fig. 3: Channel utilization U and delay D vs. mean message length H (M=30, F=20).

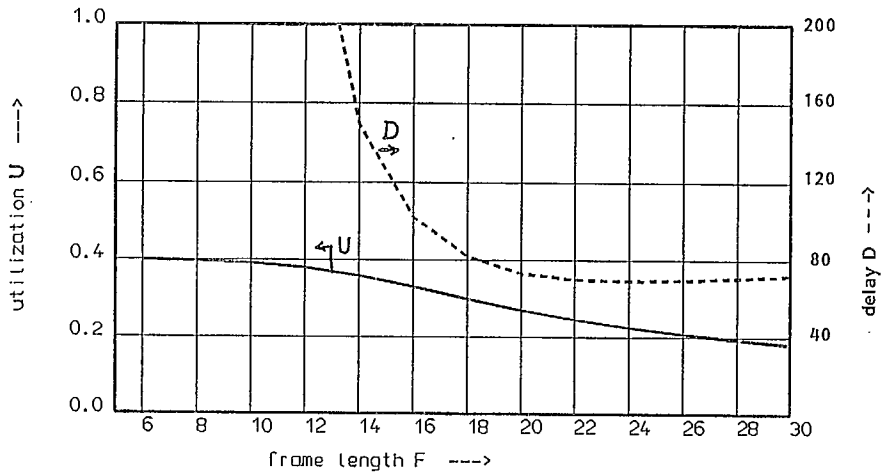


Fig. 4: Channel utilization U and delay D vs. frame length F (M=30, H=2.5).

Sloppy-Slotted ALOHA

Stewart N. Crozier

Department of Communications, Communications Research Centre, DSAT,
PO Box 11490, Station H, 3701 Carling Ave., Ottawa, Canada,
K2H 8S2, Ph: (613)-998-9262, Fax: 990-0316

ABSTRACT

Random access signaling, which allows slotted packets to spill over into adjacent slots, is investigated. It is shown that sloppy-slotted ALOHA can always provide higher throughput than conventional slotted ALOHA. The degree of improvement depends on the timing error distribution. Throughput performance is presented for Gaussian timing error distributions, modified to include timing error corrections. A general channel capacity lower bound, independent of the specific timing error distribution, is also presented.

1.0 INTRODUCTION

The random access channels for the North American MSAT system are likely to involve some form of slotted ALOHA signaling. A potential problem is the large guard times which may be required to ensure that packets stay correctly slotted. A variation in distance of 6000 km from mobile terminal to satellite, results in a two way propagation delay variation on the order of 40 msec. As an example, with 192 transmission bits per packet, a guard time of 40 msec corresponds to half a packet length for 2400 bps, and a full packet length for 4800 bps. The throughput reduction resulting directly from the use of non-zero guard times is well understood. If the guard time is narrower than the width of the timing error distribution, then packets transmitted in adjacent slots will occasionally collide. The throughput reduction, caused by adjacent packet collisions, is usually assumed to be small, or forced to be negligible by choosing a sufficiently large guard time. It is shown that this is not the best strategy, and that the optimum guard time is often much narrower than the width of the timing error distribution. Random access signaling, which allows slotted packets to spill over into adjacent slots, is denoted as Sloppy-Slotted ALOHA.

The throughput and channel capacity performance of classical unslotted ALOHA [1], and classical slotted ALOHA, including non-zero guard times [2, 3, 4], is reviewed. The corresponding performance measures for sloppy-slotted ALOHA are derived. Performance results are presented for Gaussian timing error distributions, modified to include a fraction of traffic with corrected timing.

A convenient and general channel capacity lower bound, independent of the specific timing error distribution, is also presented. This lower bound is particularly useful for designing signaling systems where most users are expected to have accurate timing, but a few users could have very large timing errors, and the type and width of the timing error distribution is unknown.

2.0 CLASSICAL ALOHA PERFORMANCE

2.1 Unslotted ALOHA

The throughput performance of classical unslotted ALOHA is given by the well known formula [3, 4]

$$S = G\, e^{-2G} \qquad (1)$$

where S is the normalized channel throughput in packets per packet length, and G is the normalized channel traffic, or offered traffic, also in packets per packet length. The channel capacity, C, is defined as the maximum channel throughput achievable, and is found by differentiating (1) with respect to G and equating to zero. The result is

$$C = \max[S] = \frac{1}{2e} \simeq 18.4\,\% \qquad (2)$$

and occurs for G=0.5. Unslotted ALOHA does not require guard times since there are no slots to guard.

2.2 Slotted ALOHA

The throughput performance for classical slotted ALOHA, which assumes that the required guard time is negligible, is given by [3, 4]

$$S = G\,e^{-G} \tag{3}$$

where S and G are as defined above. The channel capacity for this case is

$$C = \max[S] = \frac{1}{e} \simeq 36.8\,\% \tag{4}$$

and occurs for G=1.0. The channel capacity is twice that of unslotted ALOHA. The above result holds only if it is assumed that packets transmitted in adjacent slots never collide, and that the necessary guard time is negligible. This is not the case in practice, and non-zero guard times will be required. This is especially true for the MSAT system, which will involve time-varying propagaton delays with large delay differences.

2.3 Slotted ALOHA with Non-Zero Guard Times

The analysis of classical slotted ALOHA with non-zero guard times is identical to that with zero guard times, provided the traffic statistics are presented in terms of packets per slot, instead of packets per packet length. The result is

$$S' = G'\,e^{-G'} \tag{5}$$

where S' is the channel throughput in packets per slot, and G' is the offered traffic, also in packets per slot. When the packet and slot lengths are the same, corresponding to zero guard time, S' and G' are equal to the normalized traffic parameters S and G, respectively, and equations (5) and (3) are equivalent. If the guard time is not zero, then the slot length is given by

$$\tau_s = \tau_p + \tau_g = (1 + g)\,\tau_p \tag{6}$$

where τ_p is the packet length in units of time, τ_g is the guard time, and

$$g = \frac{\tau_g}{\tau_p} \tag{7}$$

is the normalized guard time, measured in packet lengths. The normalized traffic parameters, S and G, are given by

$$S = S'\,\frac{\tau_p}{\tau_s} = \frac{S'}{1 + g} \tag{8}$$

$$G = G'\,\frac{\tau_p}{\tau_s} = \frac{G'}{1 + g} \tag{9}$$

Combining (5), (8), and (9), gives the result

$$S = G\,e^{-(1+g)G} \tag{10}$$

The channel capacity is given by

$$C = \max[S] = \frac{1}{e(1+g)} \tag{11}$$

and occurs for $G=1/(1+g)$. Figure 1 shows the channel capacity as a function of the normalized guard time, g. As the guard time approaches one full packet length, the capacity degrades to that of unslotted ALOHA.

3.0 SLOPPY-SLOTTED ALOHA

3.1 Throughput and Capacity

The previous conventional slotted results are based on the assumption that packets always fall within their intended slots. It was shown that the performance of slotted ALOHA is poor, approaching that of unslotted ALOHA, if the required guard time is on the order of a packet length. Reducing the guard time results in the following two effects: (a) It increases the potential channel capacity by increasing the number of slots available per unit time, and (b) It introduces the possibility of adjacent packet collisions, which in turn will reduce the channel capacity. Finding the optimum guard time obviously involves a trade-off between these two effects. The approximate throughput performance, with non-zero guard times, and with the possibility

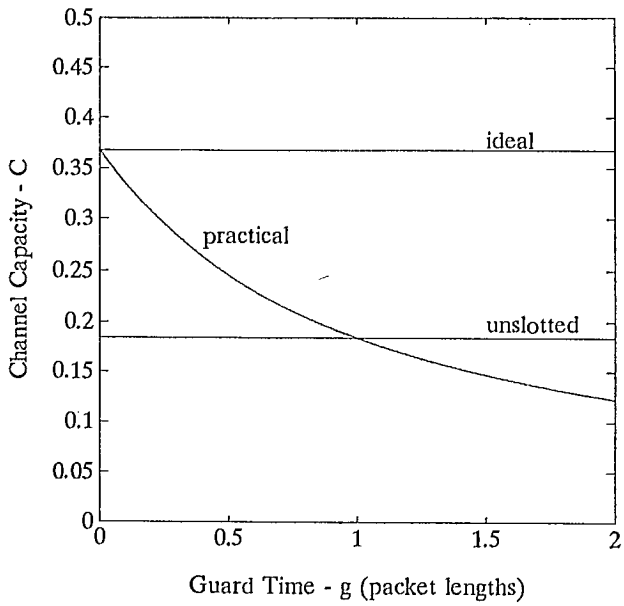


Fig. 1: Channel capacity of conventional slotted ALOHA with non-zero guard times.

of adjacent packet collisions, is derived in Appendix A. The result is

$$S \simeq G\, e^{-(1+g)(1+2p)G} \qquad (12)$$

where, S, G, and g are as defined previously, and p is the probability that 2 packets transmitted in adjacent slots collide. The channel capacity evaluates to

$$C \simeq \frac{1}{e(1+g)(1+2p)} \qquad (13)$$

and occurs for $G = 1/[(1+g)(1+2p)]$. Comparing (2) and (13), it is seen that the capacity of sloppy-slotted ALOHA is higher than the capacity of unslotted ALOHA only if $(1+g)(1+2p) < 2$. Ideally one would like to keep both g and p as close to zero as possible. One must be traded off against the other, however, since p is a function of g and the timing error distribution.

3.2 Gaussian Timing Error Distribution

Consider a Gaussian timing error distribution with a standard deviation of d packet lengths. The derivation in Appendix A is accurate for $d \leq 0.25$. With $d = 0.25$, a timing error of half a packet length or more, in either direction, will occur with a probability of 4.5%. The channel capacity is evaluated in Appendix B, and is given by

$$C \simeq \frac{1}{e(1+g)\left[1 + 2Q\left[\frac{g}{\sqrt{2}\,d}\right]\right]} \qquad (14)$$

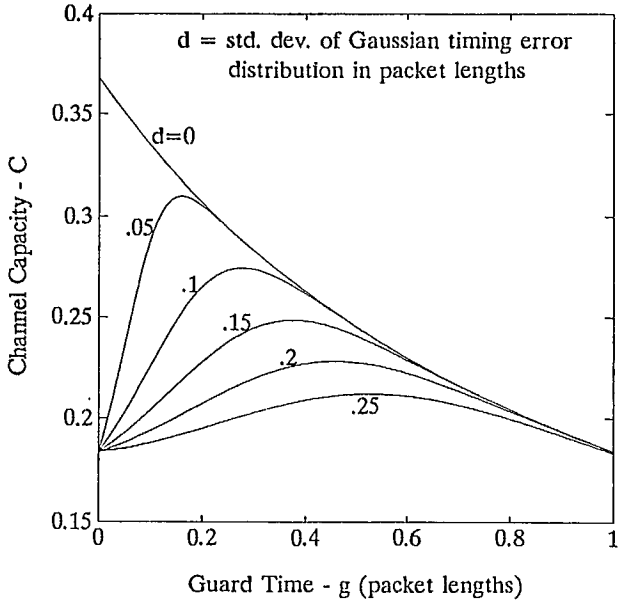


Fig. 2: Channel capacity of sloppy-slotted ALOHA versus guard time, with Gaussian timing error distribution.

where Q[x] is the area under the tail of the normal distribution from x to infinity. Figure 2 shows the channel capacity versus guard time, with the standard deviation of the timing error distribution as a parameter. Typically, one might choose g to be many standard deviations, to keep the number of miss-slotted packets small. For example, one would expect 4.5% of all packets to be miss-slotted with $g = 4d$. The optimum guard time is defined as the guard time which maximizes the channel capacity. For very wide timing error distributions, the optimum guard time is seen to be closer to $g = 2d$, which corresponds to over 30% of all packets being miss-slotted. The optimum guard times, and corresponding optimum channel capacities, are presented with the results in the next section.

3.3 Gaussian Distribution with Corrections

A Gaussian timing error distribution, modified to include timing corrections, is now considered. Fraction q of all transmitted packets are assumed to have uncorrected, Gaussianly distributed timing errors, with a standard deviation of d packet lengths. The other transmitted packets, fraction (1-q), are assumed to have perfect timing. The channel capacity is evaluated in Appendix B. The result is

$$C \simeq \frac{1}{e(1+g)\left[1+4q(1-q)\,Q\left[\frac{g}{d}\right]+2q^2\,Q\left[\frac{g}{\sqrt{2}\,d}\right]\right]} \qquad (15)$$

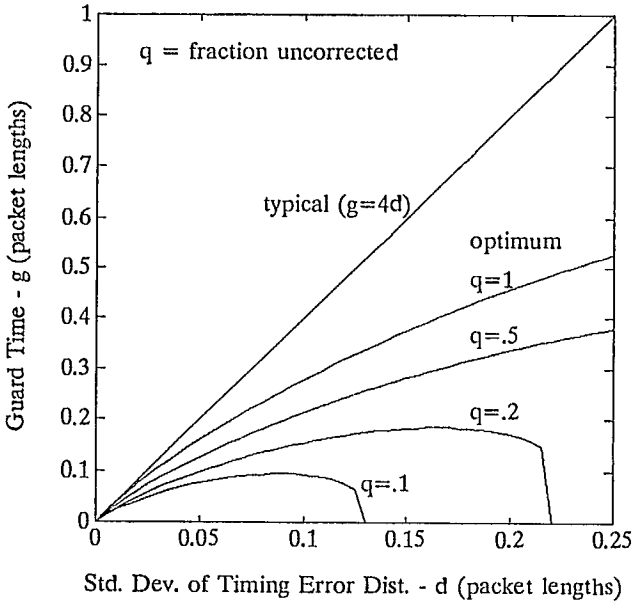


Fig. 3: Optimum guard time versus standard deviation of Gaussian timing error distribution, modified to include timing corrections.

Figure 3 shows the optimum guard times versus standard deviation, d, with fraction uncorrected, q, as a parameter. This figure was obtained using numerical methods on (15), to find the guard time which maximized the capacity for each d and q. Note that, with $q < 0.5$, as the width of the timing error distribution becomes very large, the optimum guard time jumps back to zero. Figure 4 shows the corresponding optimum channel capacities.

3.4 Channel Capacity Lower Bound

A lower bound on channel capacity is derived in Appendix C. The lower bound is general in that it is independent of the specific type and width of the timing error distribution. All that is required is the probability of being miss-slotted, m, given a specific guard time, g. The result is

$$C \geq \frac{1}{e(1+g)[1+2m-m^2]} \qquad (16)$$

This lower bound is plotted versus guard time in Figure 5, with m as a parameter. As one might expect, the lower bound predicts ideal slotted performance with g=0 and m=0, and ideal unslotted performance with g=0 and m=1. The greater the guard time the poorer the bound.

This lower bound is useful for designing systems where corrected timing is possible, but not all packets will have corrected timing, and the timing error distribution without corrections is not well known, or time-varying. For example, it may be known that, on average, the timing of 95% of all packets can be corrected to within 10% of a packet length, and that the timing for the remaining 5% is very unpredictable. It can be seen, from the lower bound plotted in Figure 5, with g=0.1 and m=0.05, that the channel capacity is at least 30%.

4.0 CONCLUSIONS

A tight approximation for the throughput and channel capacity, with sloppy-slotted ALOHA signaling, was derived. Performance results were presented for Gaussian timing error distributions, modified to include timing corrections. The results show that sloppy-slotted ALOHA can always provide higher throughput than conventional slotted ALOHA. The degree of improvement depends on the specific timing error distribution. The greatest improvement is for wide timing error distributions, with the optimum guard time often being close to zero.

A convenient and general channel capacity lower bound, independent of the specific timing error distribution, was also presented. This lower bound is particularly useful for designing signaling systems where most users are expected to have accurate timing, but a few users could have very large timing

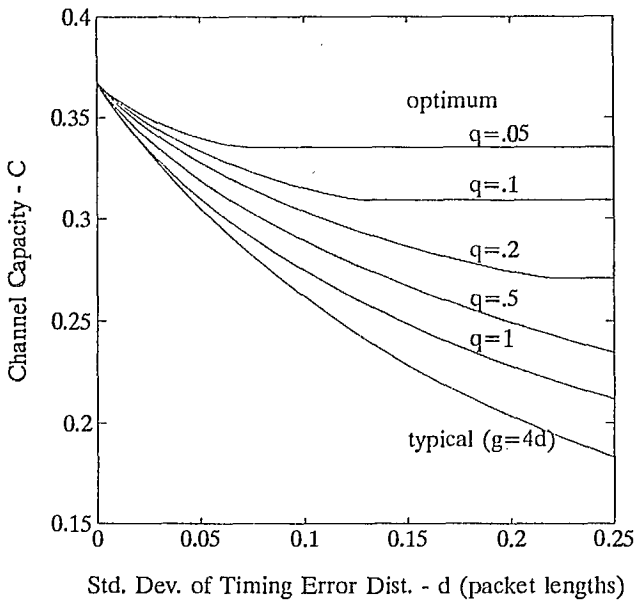


Fig. 4: Optimum channel capacity versus standard deviation of Gaussian timing error distribution, modified to include timing corrections.

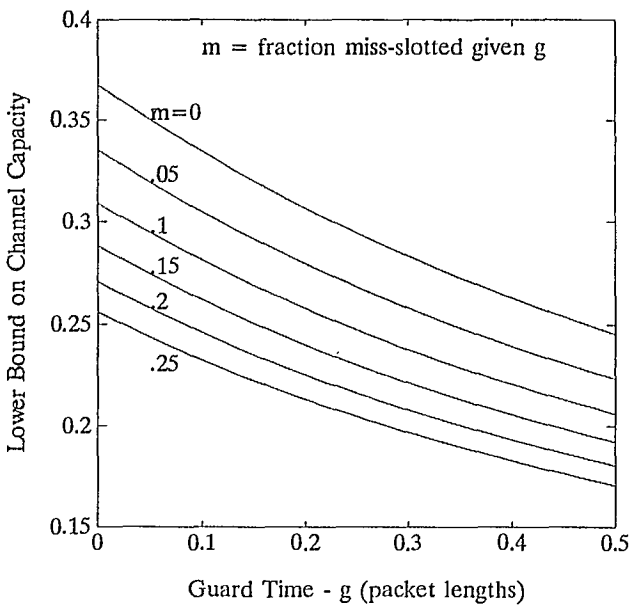


Fig. 5: Channel capacity lower bound versus guard time, with fraction of miss-slotted packets as a parameter.

errors, and the type and width of the timing error distribution is unknown.

Appendix A: Analysis of Sloppy-Slotted ALOHA

The following definitions are used:
N = number of users
M = middle (or co-slot) collision
L = left collision
R = right collision
$\bar{x}$ = no x, or not x (eg. $\bar{M}$ = no middle collision)
P(x) = probability of x
$P(x \cap y)$ = probability of x and y
P(x|y) = probability of x given y

Then S'/N = average throughput in packets per slot per user = probability of a successful packet per slot per user, and G'/N = average offered traffic in packets per slot per user = probability of a transmission per slot per user. It follows that

$$\frac{S'}{N} = \frac{G'}{N} P(\bar{M} \cap \bar{L} \cap \bar{R}) \qquad (A.1)$$

Multiplying through by N, and using the fact that M is independent of L and R, gives

$$S' = G' P(\bar{M}) P(\bar{L} \cap \bar{R}) \qquad (A.2)$$

M is independent of L and R because P(M) does not depend on the timing error distribution. This is because a second co-slot transmission is always assumed to cause a middle collision. L and R are not independent. Knowing that a left collision has occurred reduces the probability of a right collision, since it is more likely that the packet of interest has shifted left than right. Expressed mathematically,

$$P(R|L) \leq P(R) \qquad (A.3)$$

Note that

$$\begin{aligned} P(\bar{L} \cap \bar{R}) &= 1 - P(L) - P(R) + P(L \cap R) \\ &\geq 1 - P(L) - P(R) \\ &= [1-P(L)]\,[1-P(R)] - P(L)\,P(R) \\ &= P(\bar{L})\,P(\bar{R}) - P(L)\,P(R) \end{aligned} \qquad (A.4)$$

Also, it follows from (A.3), that

$$\begin{aligned} P(\bar{L} \cap \bar{R}) &= P(\bar{L})\,P(\bar{R}|\bar{L}) \\ &\leq P(\bar{L})\,P(\bar{R}) \end{aligned} \qquad (A.5)$$

With equal traffic and timing error statistics for all users, the symmetry of the problem forces P(L) = P(R), even if the timing error distribution is not symmetric. From (A.4) and (A.5), the following lower and upper bounds on $P(\bar{L} \cap \bar{R})$ are obtained

$$P(\bar{L})^2 - P(L)^2 \leq P(\bar{L} \cap \bar{R}) \leq P(\bar{L})^2 \qquad (A.6)$$

These bounds differ only by $P(L)^2$. Typically, in the operating region of interest, P(L) is fairly small, so that $P(L)^2$ is a very small second order effect. Even for P(L) as high as 10%, the bounds are only 1% apart. Thus, $\bar{L}$ and $\bar{R}$ are approximately independent in the operating region of interest. Using the upper bound of (A.6) as an approximation, (A.2) simplifies to

$$S' \simeq G' P(\bar{M}) P(\bar{L})^2 \qquad (A.7)$$

For N users, $P(\bar{M})$ is given by

$$P(\bar{M}) = \left[1 - \frac{G'}{N}\right]^{N-1} \qquad (A.8)$$

Taking the limit as N approaches infinity gives

$$P(\bar{M}) = e^{-G'} \qquad (A.9)$$

For N users, $P(\bar{L})$ can be approximated by

$$\begin{aligned} P(\bar{L}) \simeq\ & P(0 \text{ left adjacent packets}) \\ & + P(1 \text{ left adjacent packet}) \times P(\text{ no overlap}) \\ & + P(2 \text{ left adjacent packets}) \times P(\text{ no overlap})^2 \\ & + \ldots \end{aligned} \qquad (A.10)$$

$$\begin{aligned} &= \left[1 - \frac{G'}{N}\right]^N \\ &+ \binom{N}{1}\left[\frac{G'}{N}\right]^1\left[1 - \frac{G'}{N}\right]^{N-1}(1-p) \\ &+ \binom{N}{2}\left[\frac{G'}{N}\right]^2\left[1 - \frac{G'}{N}\right]^{N-2}(1-p)^2 \\ &+ \ldots \end{aligned} \qquad (A.11)$$

where

$$\binom{N}{n} = \frac{N!}{(N-n)!\,n!} \qquad (A.12)$$

is the number of "N choose n" combinations, and p is the probability that 2 packets transmitted in adjacent slots overlap (collide). The approximation is only for the high order terms, and is due to the independence assumption for the probability of no overlap when 2 or more packets are transmitted in the same adjacent slot. The independence assumption provides a very tight lower bound in this case, provided p is small (e.g. less than 10%). Taking the limit as N approaches infinity gives

$$\begin{aligned} P(\bar{L}) &\simeq e^{-G'}\left[1 + G'(1-p) + \frac{1}{2!}\left[G'(1-p)\right]^2 + \ldots\right] \\ &= e^{-G'}\,e^{G'(1-p)} \\ &= e^{-pG'} \end{aligned} \qquad (A.13)$$

Substituting (A.9) and (A.13) into (A.7) yields (as N approaches infinity)

$$S' \simeq G' e^{-(1+2p)G'} \qquad (A.14)$$

Performance can be presented in terms of the normalized throughput, S, and offered traffic, G, measured in packets per packet length, by accounting for the non-zero guard times, as in Section 2.3. The result is equation (12).

Appendix B: Gaussian Distribution with Corrections

Let u and v represent the timing errors, measured in packet lengths, for the first and second of two adjacent packets, respectively. With probability (1-q), the timing is correct, and with probability q, the timing is Gaussianly distributed with standard deviation d. The probability that the two adjacent packets collide is given by

$$p = P(u > v+g) \qquad (B.1)$$

$$= \iint_A f_u(x)\, f_v(y)\, dx\, dy \qquad (B.2)$$

$$= \int_{L_x} (1-q) f_u(x)\, dx + \int_{L_y} (1-q) f_v(y)\, dy + \iint_{A-L_x-L_y} f_u(x) f_v(y)\, dx\, dy \qquad (B.3)$$

$$= 2q(1-q)Q\left[\frac{g}{d}\right] + q^2 Q\left[\frac{g}{\sqrt{2}\,d}\right] \qquad (B.4)$$

where $f_u(x)$ and $f_v(y)$ represent the probability density functions for u and v respectively, L_x and L_y are the infinite half-lines shown on the x and y axes in Figure 6, and A is the shaded half-plane. Substituting (B.4) into (13) gives (15), and letting q=1 gives (14).

Appendix C: Channel Capacity Lower Bound

Using definitions similar to those in Appendix B, the probability that the two adjacent packets collide can be upper bounded as follows:

$$p = P(u > v+g) \qquad (C.1)$$

$$= \iint_A f_u(x)\, f_v(y)\, dx\, dy \qquad (C.2)$$

$$\leq \iint_{A+B+C} f_u(x)\, f_v(y)\, dx\, dy \qquad (C.3)$$

$$= \tfrac{1}{2} - \iint_D f_u(x)\, f_v(y)\, dx\, dy \qquad (C.4)$$

$$= \tfrac{1}{2} - \tfrac{1}{2}(1-m)^2 \qquad (C.5)$$

$$= m - \tfrac{1}{2}m^2 \qquad (C.6)$$

where A, B, C, and D represent the regions shown in Figure 7, and the probability of u or v being outside the box is m. Substituting (C.6) into (13) gives (16).

References

[1] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications", Proceedings, Fall Joint Computer Conference, 1970.

[2] L. Roberts, "ALOHA Packet System With and Without Slots and Capture", Computer Communications Review, April, 1975.

[3] R. Rosner, *Distributed Telecommunications Networks, via Satellites and Packet Switching*, Wadsworth, Inc., 1982.

[4] W. Stallings, *Data and Computer Communications*, Macmillan Publishing Company, New York, 1985.

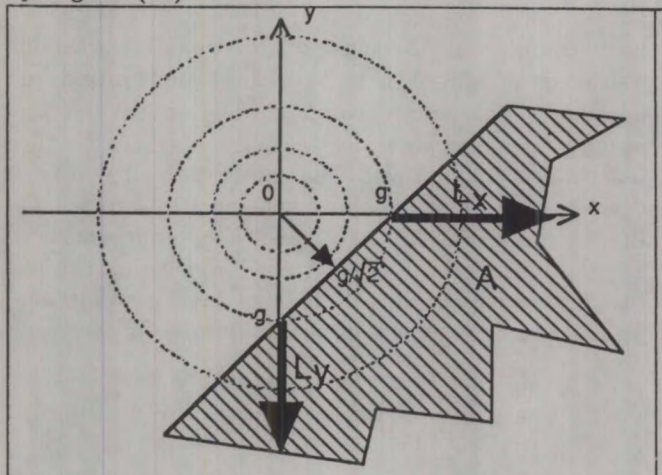


Fig. 6: Joint timing error distribution (Gaussian).

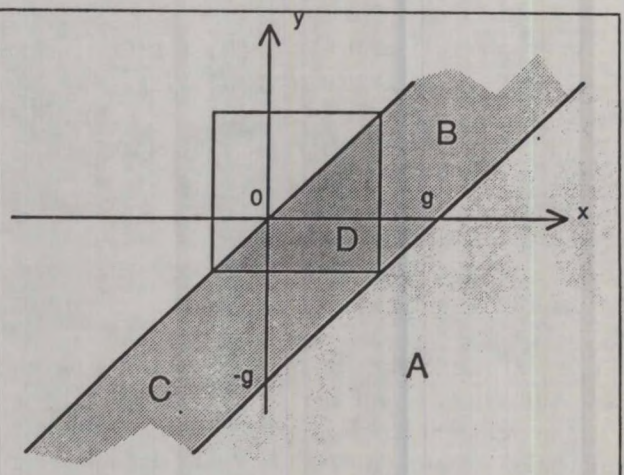


Fig. 7: Joint timing error distribution (General).

Performance of Random Multiple Access Transmission System

N. Phinainitisart
Chulalongkorn University
Bangkok, Thailand

W. W. Wu
Stanford Telecom
1761 Business Center Drive
Reston, VA 22090
Phone: 703-438-8110
FAX: 703-438-8112

ABSTRACT

This paper determines the performance of the RMA technique applied to direct terminal-to-terminal link with large number of potential users. The average signal-to-noise (SNR) is derived. Under Gaussian assumption, the approximation of the probability of error is given. The analysis shows that the system performance is affected by the sequence length, the number of simultaneous users, and the number of cochannel symbols, but is not sensitive to the thermal noise. The performance of using very small aperture antenna for both transmitting and receiving without a hub station is given.

INTRODUCTION

Most satellite networks, at present, use earth stations as shared nodes or involve a number of small terminals in a star network interacting with a central master station, or hub station. Random Multiple Access (RMA) technique will allow very small aperture terminals (VSAT) to be installed directly on end-user premises (fixed or mobile terminals) and have a mesh configuration. The reason for mesh configuration is to avoid double-hop of messages through the gateway or hub stations, thus, allowing small terminal to small terminal interconnectivity.

Within a satellite antenna coverage area, all terminals potentially can directly communicate with any other terminal. For n terminals, the number of possible terminal-to-terminal link is n(n-1)/2. If the system consists of five thousand terminals or more, then the potential number of connection is over 12,497,500 different links. Under such conditions, frequency division multiple access (FDMA), or time division multiple access (TDMA), or any other orthogonal access system are not likely to be feasible. The direct sequence spread spectrum (DS-SS) is also not shown for applications of such large number of user links. Therefore, it appears that RMA is an attractive solution to provide access for large number of VSAT.

The basic transmission scheme of a RMA system is the transformation of each binary bit into a set of time-frequency (TF) symbols. These symbols form a code sequence which is uniquely assigned to a particular user. The receiver of the system, then, performs the inverse transformation from a sequence of symbols back to the corresponding binary bit.

The procedures for RMA code sequence generation, which utilizing Euclidean Geometry (EG) Difference Set {De} and Projective Geometry (PG) Difference Set {Dp}, were presented in [1] [2] [4]. In this paper only the system performance results based on the derivation from {De} are presented.

Section Two describes the system model and defines the notations. Section Three addresses the issue of average signal to noise ratio. Section Four relates signal to noise ratio to the probability of error with sequence parameters. Numerical examples are given in Section Five.

SYSTEM MODEL

Our model for a RMA transmission system presented in Figure 1 is taken from [1]. The L active users simultaneously transmit through the satellite transponder to L receivers. Each RMA transmitter contains a set of frequency generators (FG), a set of delay units, biphase modulators, and assembler. Assume that each transmitter has the same power and there is only one overlapping symbol between any two sequences.

The #k user's data signal b_k (t) is a sequence of unit amplitude, positive and negative, rectangular pulse of duration T. The #k user is assigned a RMA code sequence, which contains M symbols, to represent each of it's binary information bit. Each FG generates a carrier frequency with proper delay by the delay unit to represent a symbol in the code sequence. The biphase modulator phase modulates the data signal b_k (t) of the #k user onto the carrier. The number of such FG and delay units depends on the number of symbols/sequence (M). After passing through the assembler, the corresponding code sequence S_k (t) is formed for transmission. The remaining (L-1) transmitters are identical to the #k transmitter except that different FG and delay units are used to generate different code sequence for each of the transmitters.

A the receiving end, the corresponding FG and delay units, and correlation detections are used to reverse the encoding process from the received signal W(t). The total output from all M correlation receivers of any #k user (C_k) will include the desired signal b_k (t), channel noise n(t), and multiple-access interference (MAI) signals. The recovery of the data bit b_k is done by the threshold comparator which provides at T interval a +1 or -1 output state, depending on whether C_k is larger or smaller than zero.

The analysis of the interaction between RMA signals in a multiple access transmission environment is presented in [1] [3]. The analysis shows that SNR depends on the number of simultaneous users (c) the number of energy per data bit, and the design of TF matrix.

Because the statistic of the simultaneous users is time varying, the exact evaluations of SNR and error rate Pe are difficult. Instead, an approximation will be used, based on the assumption that all noise is additive white Gaussian noise and large number of users are involved. The system is operating in the environment where there are strong interfering signals, hence, the system performance is not sensitive to the variation in the thermal noise. Consequently, allowing the use of very large number of small aperture terminals for both transmission and reception.

APPROXIMATION OF THE AVERAGE SIGNAL-TO-NOISE RATIO

We have shown in [1] [3] that the output from each correlation receiver includes the desired signal, channel noise, and the multiple-access interference (MAI). This MAI is caused by all simultaneous users whose sequence symbols are represented by the same carrier frequencies as the desired symbols. The total number of cochannel symbols for each carrier is simply the number of time-delay units in the TF matrix, or T_d.

The precise value of MAI can be calculated only when the statistic of the simultaneous users are known at a particular time for a particular system determined by the corresponding TF matrix. Therefore, the average value of multiple access interference and consequently the signal-to-noise ratio, which is useful for preliminary system design where the statistic of each user is not available but the sequence parameters and TF matrix are known, will be determined. The

SNR at the input of the threshold comparator can be expressed as [1].

$$\overline{SNR} = \left[\frac{M}{\frac{\bar{U}}{3M} + \frac{N_o}{2E_b}} \right]^{1/2} \quad (1)$$

$$\text{where } \bar{U} = \bar{a} + c\bar{d} = \frac{\bar{U}}{3M} \quad (2)$$

The parameters ā and d̄ are the part of MAI of any user, which is caused by his own transmitter, and any other simultaneous user accordingly.

ā = average number of designated cochannel symbols ($0 \le a \le M$)
d̄ = average number of cochannel symbols between designated user and any other user ($0 \le d \le M$)
c = total number of simultaneous users ($0 \le c < L$)

Both ā and d̄ can be expressed in term of sequence length, as followed:

$$\bar{a} = yM \quad \text{where } 0 \le y \le 1 \quad (3)$$

$$\bar{d} = \frac{\text{max. } (U) - yM}{\text{max. } (c)} \quad (4)$$

Parameter y is a user's cochannel symbols-to-sequence symbols ratio, or

$$\bar{y} = \frac{\left[\text{average number of cochannel symbols in a sequence}\right]}{\left[\text{total number of symbols in that sequence}\right]}$$

When all users are activated simultaneously, we have

$$\text{max. } (c) = L-1 \quad (5)$$

$$\text{where } L = L_e = M(M+1) \quad (6)$$

The maximum value of parameter U are given by [1]

$$\text{max. } (U) = M[(M+1)T_d - 1] \quad (7)$$

Substituting (7) and (6) into (4)

$$d = \frac{M[(M+1)T_d - 1] - yM}{M(M+1) - 1} \quad (8)$$

Since M>>1, (8) reduce to

$$d \propto T_d \quad (9)$$

From (2), (3) and (9), we have

$$\bar{U} = yM + cT_d \quad (10)$$

Finally substituting $\bar{U}$ into (1), we have

$$\overline{SNR} = \left[\frac{M}{\frac{1}{3M}[\bar{y}M + cT_d] + \frac{N_o}{2E_b}} \right]^{1/2} \quad (11)$$

In terms of energy per data bit, (11) can be expressed as

$$\overline{SNR} = \left[\frac{M}{\frac{1}{(E_b/N_o)_{MAI}} + \frac{1}{(E_b/N_o)_{th.}}} \right]^{1/2} \quad (12)$$

where $\left(\frac{E_b}{N_o}\right)_{th.}$ = energy -per-bit/channel thermal noise density ratio

$$\left(\frac{E_b}{N_o}\right)_{MAI} = \frac{3M}{\bar{y}M + cT_d} \quad (13)$$

or $\left(\frac{E_b}{N_o}\right)_{MAI}$ = energy-per-bit/multiple-access-interference density ratio

PROBABILITY OF ERROR

In this case of a communication system with additive white Gaussion noise (AWGN), which includes multiple access interference noise and channel noise, a data bit error occurs at the threshold comparator output when the integrated amplitude of the noise is larger than the integrated amplitude of the desired signal in the opposite direction.

The threshold comparator provides, at T intervals, a +1 or -1 output state, depending on whether C is larger or smaller than zero, respectively. Thus, the probability of error is given by

$$Pe = P[C_i>0/b_i = -1]\ P\ [b_i= -1] + P[C_i<0/b_i = +1\ P[b_i= +1] \quad (14)$$

where $P[C_i>0/b_i= -1]$ = probability of having sampled C_i larger than zero, given that a -1 (b_i) is being transmitted

For equa-probable signaling, (14) becomes

$$Pe = 0.5\{P[C_i>0/b_i = -1] + P[C_i<0/b_i = +1]\}$$

or Pe = 0.5 P[(total noise at the sampling instant) > (total desired signals at the sampling instant)]

The probability of error can be approximated by

$$Pe \approx Q[\overline{SNR}] \quad (15)$$

From (11) and (15), we can conclude that the probability of error of the RMA system depends on the following parameters:

- sequence length (M)
- user's cochannel symbols-to-sequence symbols ratio (y)
- number of simultaneous users (c)
- number of time-delay in the TF matrix or number of cochannel symbols (T_d)
- energy-per-bit/channel-thermal-noise density ratio $(E_b/N_o)_{th.}$

Note that the parameters M, y and T_d depend on the structure of the sequence and the size of the TF matrix, but the $(E_b/N_o)_{th.}$ can be calculated from the link budget parameters, such as transmitted carrier power, bit rate, etc. It is of interest to know how the system will perform when the number of simultaneous users are varying, as this is the only parameter which can not be predetermined or fixed due to the random access nature of the system. Thus, the graphs in Figures 2 - 5 are of Pe versus C for different values of M, y, T_d and $(E_b/N_o)_{th.}$

Our objective is to keep the value of $(E_b/N_o)_{th.}$ to a minimum; as the higher $(E_b/N_o)_{th.}$ may result in higher uplink carrier power, or larger antenna at the received station or a reduced data rate. We assume that all stations have the same power, thus, any increase in energy per data bit from the desired user also means an increase in the MAI. Therefore, $(E_b/N_o)_{th.}$ has little effect on the Pe as we can observe in Figure 3. The Pe is also not sensitive to the change of parameter y as shown in the plot in Figure 4. On the contrary Figures 2 and 5 demonstrate that the sequence length and the number of cochannel symbols in the TF matrix greatly affect the probability of error, especially when the number of simultaneous users are relatively very small compared to the total number of users.

Plots in all figures are for sequences of length 59, 97, 107, 127, and 169 which can be constructed from EG sets. The corresponding number of users are calculated from (6) as followed:

Table 1. Total Number of Users For Each Sequence

Sequence Length (M)	Total No. of Users (L_e)
59	3,450
97	9,560
107	11,556
127	16,256
169	28,730
289	83,810

For preliminary system design it is useful to be able to carry out a tradeoff between the parameters M, T_d, and c as they all have significant effect on Pe. Such a tradeoff can be obtained from (11) and (15) as:

$$Pe = Q\left[\frac{M}{\frac{1}{3}\left[y + \frac{cT_d}{M}\right] + \frac{N_o}{2E_b}}\right]^{1/2} \quad (16)$$

Given y = 0.2, E_b/N_o = 0.2 dB, (16) becomes

$$Pe = Q\left[\frac{M}{\frac{1}{3}\left[0.2 + \frac{cT_d}{M}\right] + 0.477}\right]^{1/2} \quad (17)$$

The effects of these three parameters can be summarized as followed:

1) For a fixed number of simultaneous users, we have to increase the sequence length in order to improve the system performance. An example is shown in Table 2

Table 2. Pe for Different Sequence Length When Number of Simultaneous Users is Fixed (c = 25), and T_d = M

M	Pe
289	5.95E-09
169	6.71E-06
127	8.00E-05
107	2.65E-04
97	4.72E-04
59	4.90E-03

2) If a certain percentage of users is expected to transmit simultaneously, we'll find that smaller size system will have a better performance than the bigger one. This can be observed from the Table 3.

Table 3 Pe for Different System Size When c is a Fixed Percentage of Total Users (c = 0.001% of L), and T_d = M

Total users	c = 0.001% of Total users	Pe
83,810	84	7.4E-04
28,730	29	2.5E-05
16,256	16	1.7E-06
11,556	12	6.3E-07
9,506	10	2.9E-07
3,540	4	1.0E-08

3) Since the noise in this RMA system is dominated by cochannel interference, we can improve the system performance by reducing the number of cochannel symbols. In other words,

we have to reduce T_d while increasing F accordingly, so that the total number of symbols ($V = T_d*F$) can remain constant.

Table 4 and Figure 5 show how we can significantly reduce the probability of error by increasing the carrier frequency. In this specific example, the sequences are derived from a EG set, hence, $T_d*F=M^2$. When the TF matrix is a square matrix, we have $F = T_d = M$. If we reduce the cochannel symbols by half, the TF matrix becomes a rectangular matrix and $F = 2M$.

Table 4. Pe for Different Cochannel Symbols (T_d) (M = 169)

c (users)	$T_d = M$	$T_d = 0.5M$
14	4.48E-09	3.38E-15
16	3.45E-08	1.03E-13
18	1.64E-07	1.45E-12
20	5.84E-07	1.29E-11
22	1.68E-06	8.12E-11
24	4.12E-06	3.89E-10
26	8.87E-06	1.50E-09
28	1.72E-05	4.88E-09
30	3.0E-05	1.38E-08

4) If the system is designed to provide a fixed nominal bit error rate performance, the number of simultaneous users can not exceed a certain limit. An example is shown in Table 5.

Table 5. Maximum Number of Simultaneous Users for Different Sequence Length When $Pe = 10^{-7}$ and $T_d = M$

M	Maximum c	Max. c as % of Total Users
289	31	0.037%
169	16	0.056%
127	13	0.08%
107	11	0.095%
97	10	0.1%
59	5	0.14%

EXAMPLES OF LINK BUDGET FOR VERY SMALL APERTURE ANTENNAS

In this section some link budgets are presented to illustrate that the proposed RMA technique is of use to a system with very small aperture terminals (VSAT).

We have demonstrated in the previous section that the $(E_b/N_o)_{th.}$ has some contribution to the system performance (Pe) but it will require a big increase in $(E_b/N_o)_{th.}$ to have any noticeable effect on Pe. In addition, the carrier power varies with $(E_b/N_o)_{th.}$, thus, it is important to keep the value of this parameter at some minimum level. Since the lowest level of $(E_b/N_o)_{th.}$ assumed in this section is 2.0 dB, we will use this same value for conveniently relating the link budgets to the PE already calculated. The link budgets are designed for the system which the transmit and receive stations are the same size.

	14/11 GHz			6/4
Antenna Diameter (m).	1.2	1	0.8	1
Antenna Gain (dB)	42.3	40.8	38.8	35.0
HPA (dB/carrier)	-24.1	-21.6	-16.9	-1.4
EIRP/carrier (dBW)	18.2	19.2	21.9	33.6
Satellite G/T (dB/K)	1.7	1.7	1.7	-2.0
Up Link (C/T) thermal (dBW/K)	-188.1	-186.4	-184.4	-169.1
Antenna Gain 1 m^2 (dB-m^2)	44.0	44.0	44.0	37.0
Saturation Flux Density (dBW/m^2)	-90.5	-90.5	-90.5	-81.0
Input Backoff/ carrier (dB)	-53.6	-51.9	-49.9	-51.1
Output Backoff/ carrier (dB)	-48.1	-46.4	-44.4	-46.6
EIRP Downlink (dBW)	-7.1	-5.4	-3.4	-10.6
Earth Station G/T (dB/K)	15.5	13.8	11.8	10.0

At Ku-band the EIRP is assumed as 41.0 dB,

downlink (C/T) is -197.6 dBW/K which is almost the same as the total (C/T). At C-band the EIRP is 36.0 dB and the total (C/T) or (C/No) is the same as that of Ku-band.

In order to increase the speed of information bit, we have to increase the transmitted EIRP or increase G/T of the earth terminal. Table 6 shows bit rate versus (E_b/N_o)th..

Table 6. $(C/N_o)_{th.}$ Required for Different Bit Rate

R_b bit/sec.	dB	Total $(C/N_o)_{th.}$ dBW/Hz
1200	30.8	31
2400	33.8	34
4800	36.8	37
9600	39.8	40

As an example, if we increase the information bit rate of 1200 bit/sec. to 4800 bit/sec., we have to increase the EIRP proportionally. That is to raise the uplink EIRP level up to 6 dB. Another alternative is to introduce error coding in order to increase the G/T by 6 dB.

CONCLUSION

The expressions for approximation of the signal-to-noise ratio and the probability of error of the RMA system has been obtained, based on the Gaussian interference assumption. It is assumed that the system consists of a large number of users, thus, long sequence are required.

The evaluations of the error probability are presented. The results show that the sequence length, the number of cochannel symbols in the TF matrix and the number of simultaneous users, which all contribute to the multiple access interference (MAI) parameters, affect the bit error rate. On the contrary, the system is not sensitive to the changes in the energy-per-bit/channel thermal noise density ratio $(E_b/N_o)_{th.}$, thus, allowing the use of very small aperture terminals (VSAT) for both transmission and reception without a larger hub station. The examples of link budget for VSAT are given for illustration.

However, the system performance will degrade rapidly as the number of simultaneous users is increasing. Therefore, this access technique is suitable for systems with large number of users and light traffic. This technique can be applied to emergency network, military network, thin route commercial network, and mobile satellite communication for aeronautical, land or maritime.

REFERENCES

1. **N. Phinainitisart.** An access technique for very small aperture antenna. *Thesis for the Degree of Doctor of Engineering*, Chulalongkorn University, Bangkok, Thailand (May, 1990).

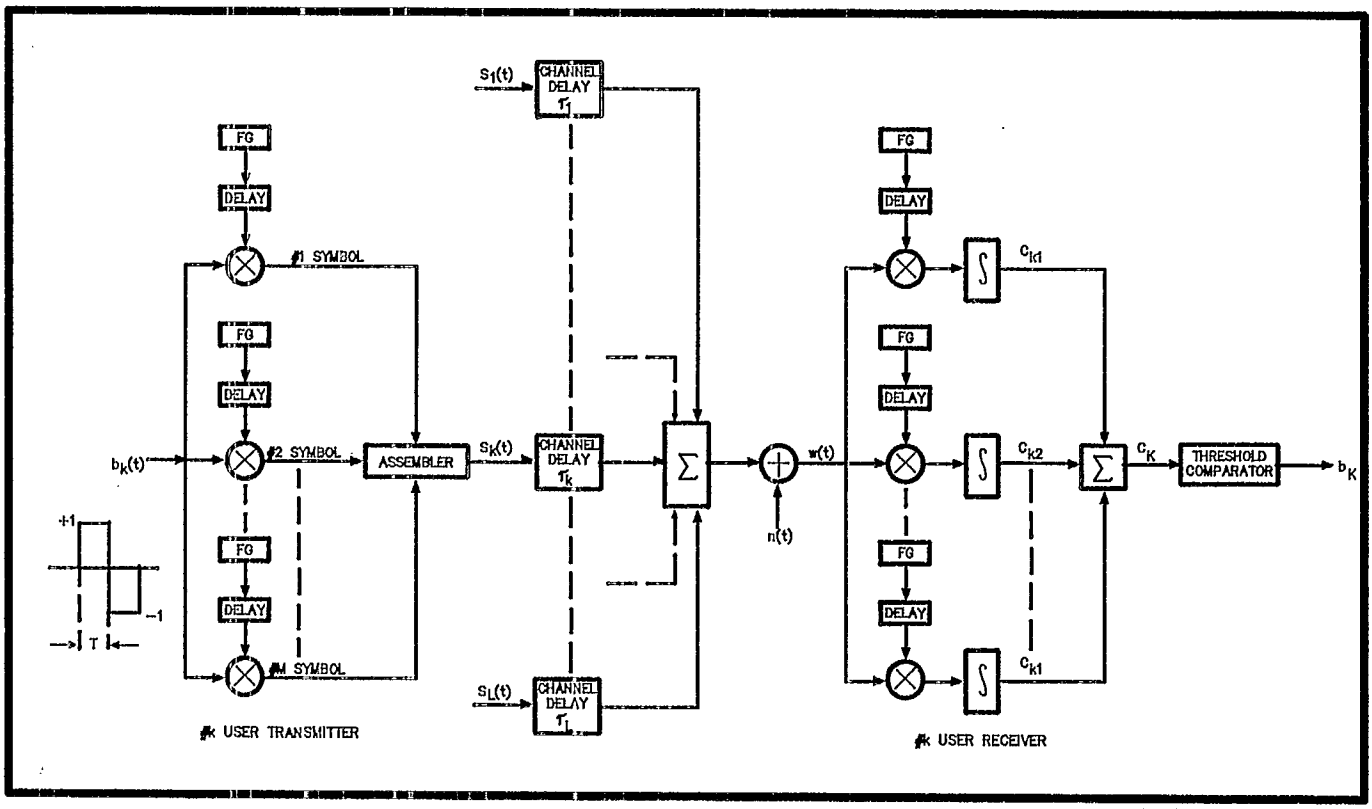
2. **N. Phinainitisart.** W.W. Wu. Code sequences for gatewayless transmission. *Proceeding IEEE GLOBECOM* (Nov., 1989), Vol. 3.

3. **N. Phinainitisart.** W.W. Wu. Analysis of random multiple access transmission system. *International Journal of Satellite Commun.* to be published in 1990.

4. **W.W. Wu.** *Elements of Digital Satellite Communication.* 2 Vols. Computer Science Press, 1985.

5. **K. Feher.** *Digital communications: satellite/earth station engineering.* Prentice Hall Inc., 1983.

6. **T.T. Ha.** *Digital satellite communications.* 2nd ed. Mcgraw-Hill, 1990.



4/10/90 MIS88T\BW5773

FIGURE 1: RMA SYSTEM MODEL

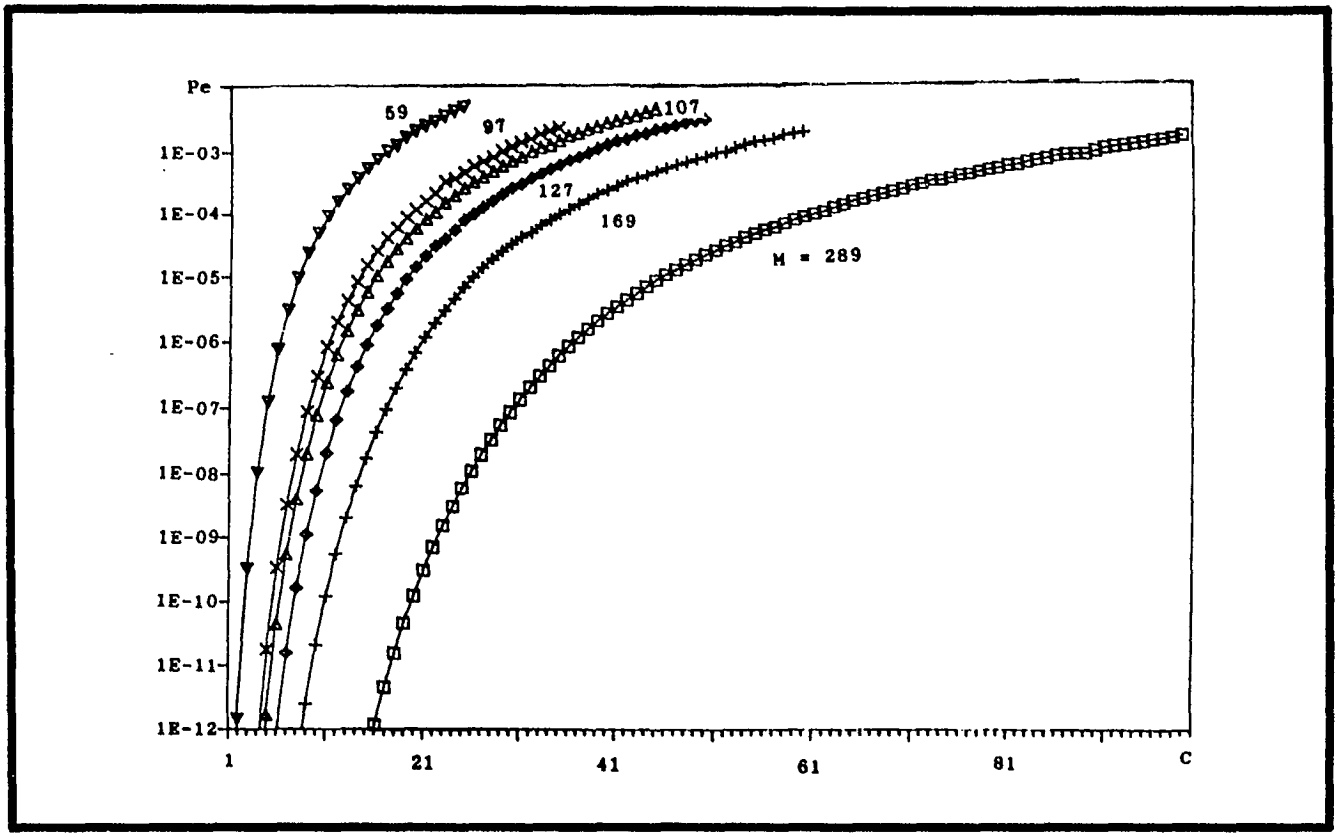


FIGURE 2: Pe AS A FUNCTION OF c FOR DIFFERENT M
($(E_b/N_o)_{th.}$ =0.2 dB, y=0.2, T_d=M)

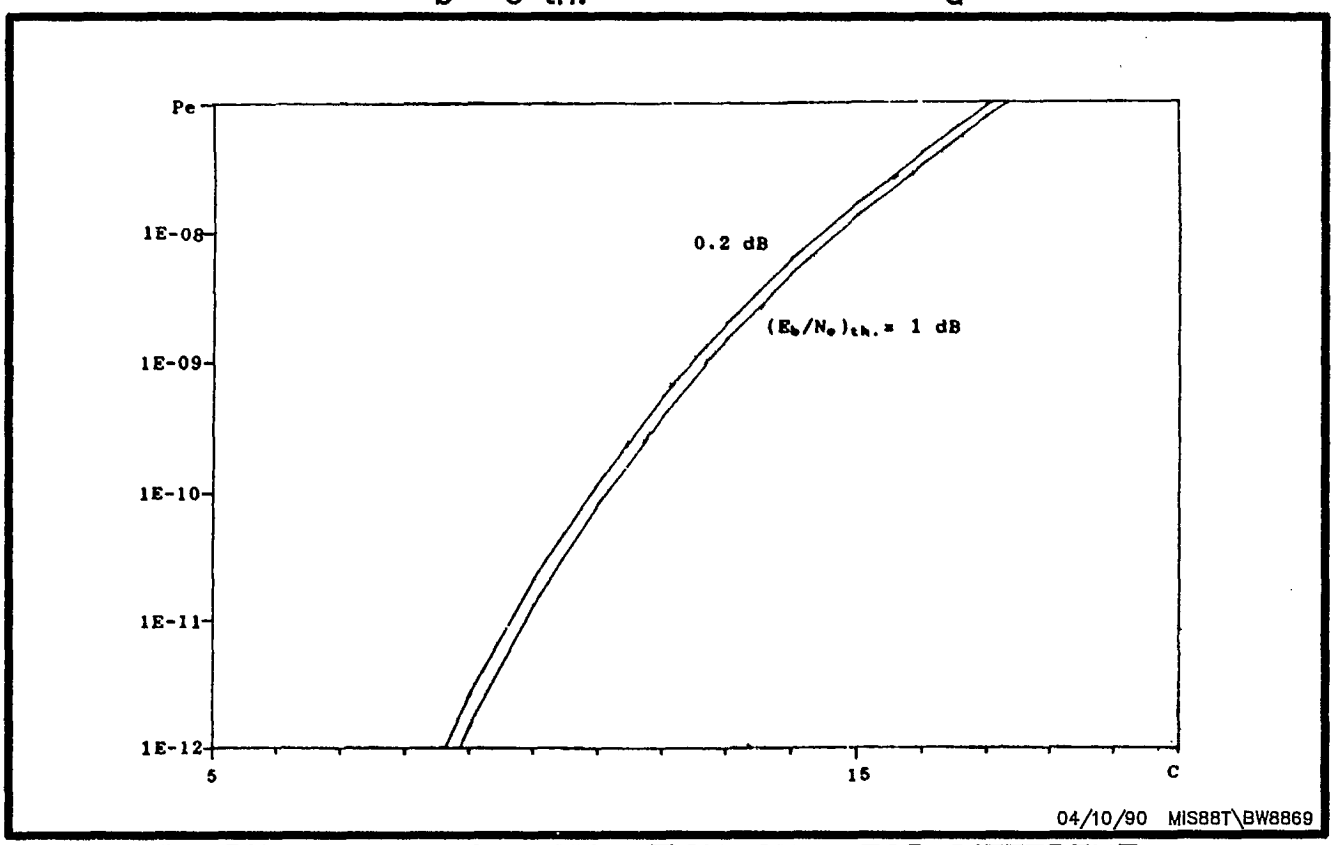


FIGURE 3: Pe AS A FUNCTION OF c FOR DIFFERENT
$(E_b/N_o)_{th.}$ (y=0.2, M=169, T_d=M)

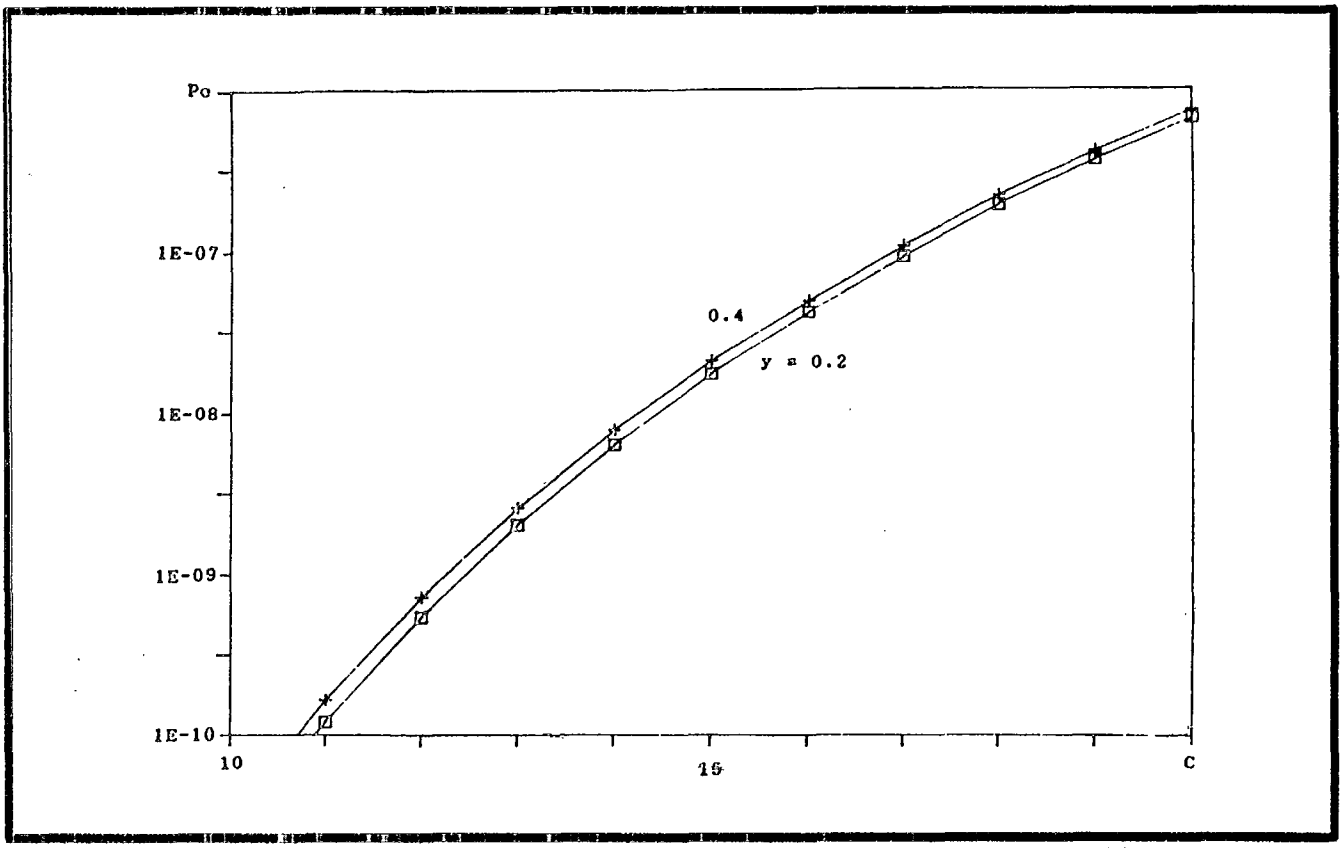


FIGURE 4: Pe AS A FUNCTION OF c FOR DIFFERENT y
$(E_b/N_o)_{th.}$ =0.2 dB, T_d=M, M=169

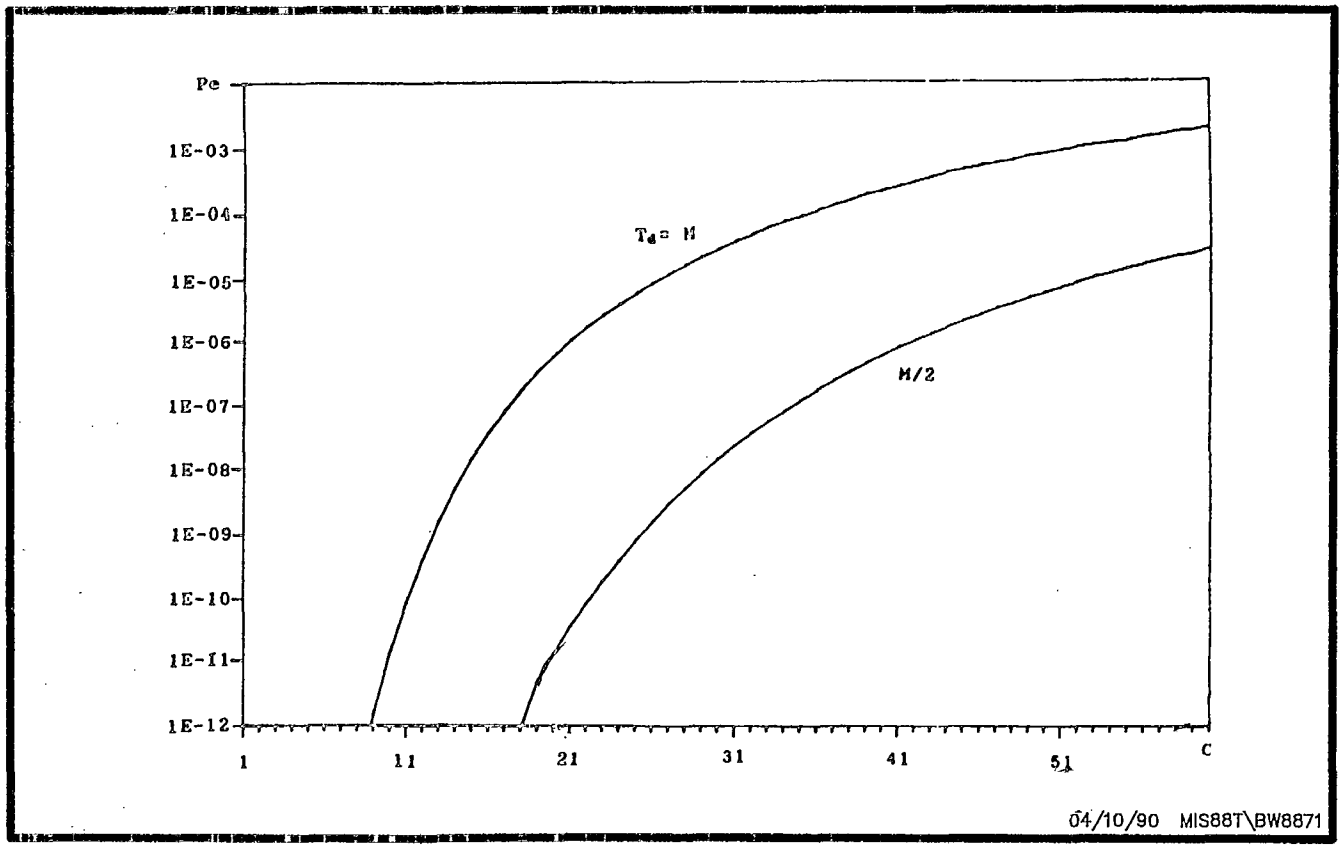


FIGURE 5: Pe AS A FUNCTION OF c FOR DIFFERENT T_d
y=0.2, M=169, $(E_b/N_o)_{th.}$ =1 dB

Session 9

Spacecraft Technology

9

Session Chairman - *F.J.F. Osborne,* Spar Aerospace, Canada
Session Organizer - *Joe McNally,* DOC

Space Segment Configuration of the Canadian Mobile Satellite System

Elemer Bertenyi
Telesat Canada
1601 Telesat Court
Gloucester, Ontario K1B 5P4
Tel: (613) 748-0123 Fax: (613) 748-8712

SYNOPSIS

Implementation of the Canadian mobile satellite system has passed another important milestone with the issuing in 1989 of the MSAT space segment Request for Proposal by Telesat Mobile Inc., and negotiations with spacecraft vendors in the early part of 1990. The performance of the spacecraft, which is a principal component and a major cost element of the MSAT system, has been specified to satisfy the requirements of cost efficient two way voice and data communications to mobile and transportable terminals within Canada, and also to provide in-orbit backup services to the United States of America mobile satellite operator.

This paper reviews the performance requirements, the main parameters and the configuration of the MSAT spacecraft. The major features of the communications subsystem are discussed in some detail.

Key technology items include the L-band RF power amplifier which must operate with a high DC to RF power efficiency and generate low intermodulation when loaded with multi-carrier signals; and the large diameter deployable L-band antenna. The development status and expected performance of these spacecraft components is examined.

The full text of this paper was not available at press time

LLM: An L-Band Multibeam Land Mobile Payload for Europe

J.Benedicto , E.Rammos
G.Oppenhaeuser , A.Roederer
European Space Agency
P.O.box 299, 2202 AG Noordwijk, The Netherlands
Phone:31-01719-83944 Fax:31-01719-84596

1 Abstract

The European Space Agency is developing, in the frame of the ARTEMIS programme, a multibeam reconfigurable mobile payload to provide pre-operational land-mobile satellite services at L-band over Europe .

The LLM payload features high capacity at L-band, efficient use of the L-band spectrum resources and flexibility in reconfiguring the allocation of bandwidth and RF power resources to the different beams. Additionally, a number of features have been added to the payload purely for experimental purposes, like the provision of one steerable spot beam which can be repositioned anywhere within the coverage area, and the possibility to reuse L-band frequencies by spatial discrimination between non-adjacent beams, or via orthogonal polarisations.

This paper describes the architecture of the payload, and the hardware implementation of the most critical subsystems.

2 Introduction

The expansion of the Mobile Satellite Service to new emerging communities of users (in particular land-based mobiles and aircrafts) is very much dependent upon the additional capacity and flexibility that the space segment can bring into the system. This leads to specific payload requirements, like the need to generate multiple high-gain spot beams over the service area, to provide adequate beam-to-beam interconnectivity with efficient means of reconfiguring the allocation of on-board power and spectrum resources between the different beams to minimise the system blocking probability. Moreover, the limitation of the L-band spectrum resources suggests the need for frequency reuse which can be provided by spatial or polarisation discrimination.

An L-band Land-Mobile (LLM) payload has been designed to meet the above system requirements and to establish a payload architecture concept with growth capability, such that it could be expanded, in a second generation LLM, to increase further the L-band EIRP and the frequency reuse capability.

3 Payload performance

The main payload performance parameters are summarised in this section.

FREQUENCY PLAN: The signals to and from the mobiles are relayed by the satellite at L-band in the 1.6/1.5 GHz frequency bands (Mobile Link), and the feeder link with the Base and Gateway fixed stations is established at Ku-band, in the 14/12 GHz frequency bands.

The L-band frequency plan follows the WARC-MOB'87 allocation (Fig-1). The LLM payload subdivides the L-band spectrum into three functional bands:

- L1: 1530-1533 MHz for telephony/data traffic to land-mobiles through spot beams (3 MHz)
- L2: 1555-1559 MHz for telephony/data traffic to land-mobiles through spot beams (4 MHz)
- G: a slot of 4 MHz contained anywhere between 1533-1555 MHz for low data rate traffic to land-mobiles through a global Eurobeam

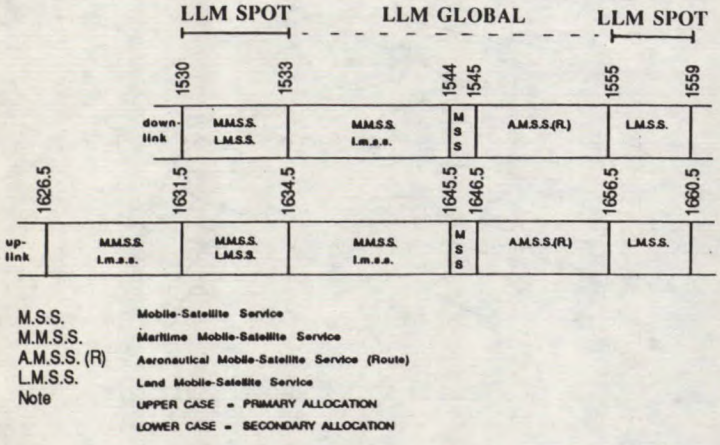


Fig.1: MSS WARC-MOB'87 frequency allocation

The bands L1 and L2 are combined together in the Feeder Link into the "M" type of channels. Channels M1, M2 and M3 reproduce the same part of the L-band spectrum in order to allow frequency reuse by spatial isolation and/or polarisation discrimination, and to cater for redundancy within the "M" type of filter banks. In summary the LLM payload covers 7 MHz (3+4) allocated on a primary basis to the Land-Mobile Satellite Service (LMSS), and 4 MHz from the band allocated to LMSS on a secondary basis and to AMSS (aeronautical) on a primary basis.

At Ku-band four non-adjacent channels are used. Three of them, M1, M2 and M3 are 8.5 MHz wide whereas the fourth, G, covers 4 MHz. The "M" channels provide 7 MHz of useful bandwidth in 1 MHz slots, and additional 250 KHz slots to cater for the guardbands of the IF Processor SAW filters (Fig-2). The channels are distributed across the 14.0-14.5 GHz band in the up-link and 11.45-11.70 GHz and 12.50-12.75 GHz in the down-link. The exact location of the channels is not defined yet, and will be coordinated with Ku-band european operators, in particular EUTELSAT.

SPECTRUM HANDLING EFFICIENCY: One of the most important features of the IF Processor presented in section 4 is its high efficiency in handling the L-band spectrum. The use of offset frequency conversions for adjacent frequency slots (Fig-2) eliminates the loss of spectrum due to filter guardbands, eases at the same time the out-of-band rejection requirements for the SAW filters, and reduces the coherent multipath component through the repeater.

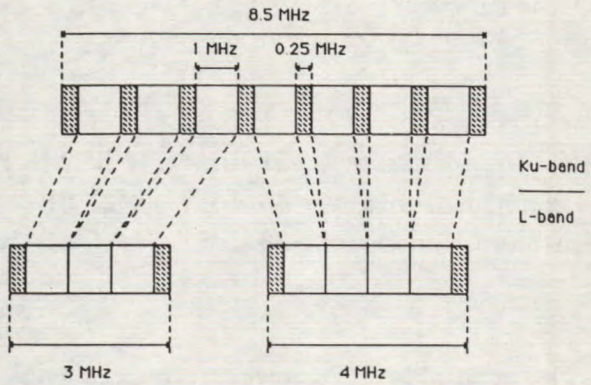


Fig.2: LLM channelisation plan

COVERAGE: The LLM payload covers all Western Europe, the upper North Africa and part of the Middle East. The coverage is provided, at L-band, in three complementary ways:

- with a single Eurobeam (Fig-3)
- with a set of six fixed spot beams (Fig-4)
- with a steerable beam which can be repositioned anywhere within the coverage area with a resolution of one degree.

One additional fixed spot beam is generated in opposite circular polarisation (RHCP), overlapped with one of the LHCP beams, in order to carry out experiments of frequency reuse by polarisation discrimination in CDMA and FDMA systems.

The total number of beams that the payload can generate at any one time is 9. Two additional steerable beams are implemented onboard for redundancy purposes.

The same coverage is implemented in the Forward and the Return links.

The coverage at Ku-band is equivalent to the L-band Eurobeam.

EIRP AND G/T: The payload exhibits the following EIRP and G/T performance:

coverage	EIRP	G/T
L-band spot	51 dBW	+2.5 dB/K
L-band Eurobeam	45 dBW	-2.0 dB/K
Ku-band Eurobeam	38 dBW	-2.0 dB/K

The available RF power resources can be shared in any ratio among the Eurobeam and any of the fixed or steerable spot beams at L-band, in order to match the non-uniform distribution of the traffic across the coverage area.

The L-band EIRP requirements have been sized with the intention to provide a good balance between RF power and spectrum availability. The capacity of the payload in terms of communication channels is approximately 1000 voice channels of 4.8 Kbps data rate (R3/4 Viterbi encoding), QPSK modulated with a channel spacing of 10 KHz.

BEAM-TO-BEAM ISOLATION: The L-band spot beam coverage of Europe (Fig-4) is such that 15 to 20 dB of isolation are provided between the NW and SE beams and between the SW and the two most SE beams. This feature will be used to perform experiments of frequency reuse by spatial discrimination in FDMA and CDMA systems.

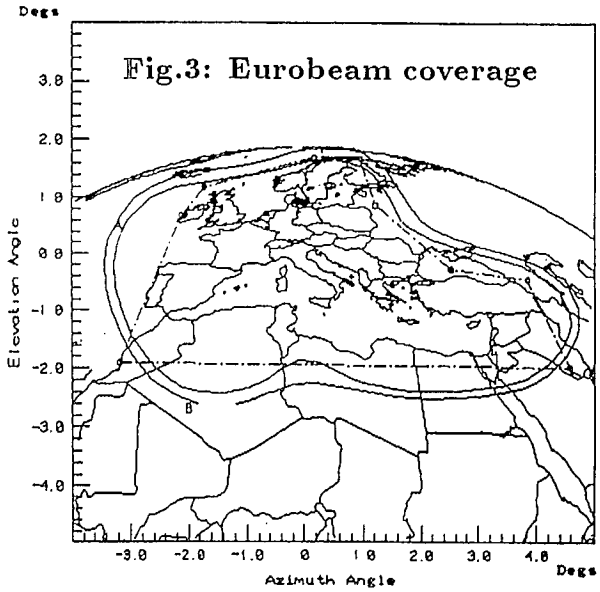


Fig.3: Eurobeam coverage

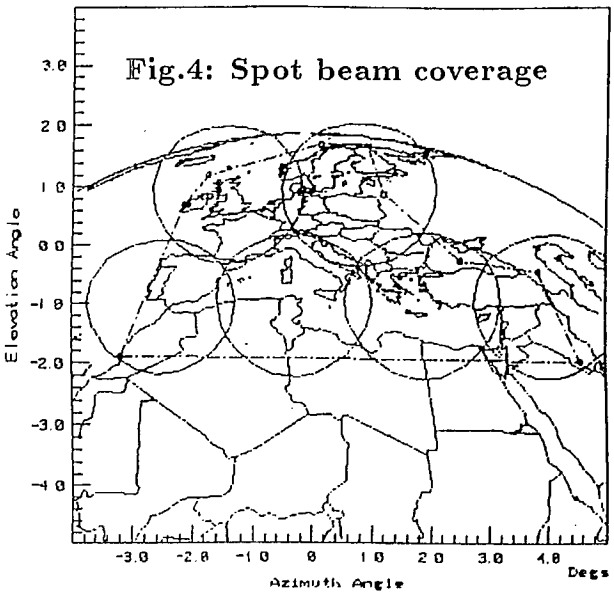


Fig.4: Spot beam coverage

IN-ORBIT RECONFIGURABILITY: The LLM payload exhibits a high degree of in-orbit reconfigurability to ease the operation of the spacecraft, and to optimise, at any time, the use of the on-board resources to maximise the system capacity. The reconfiguration of the payload can be implemented at different levels:

- flexible allocation of the L-band spectrum resources to any of the spot beams or the Eurobeam.
- flexible share of the generated L-band RF power (120 Watts) among the different beams, to match the variable assignation of bandwidth per beam. This variable power sharing capability is implemented with no loss in TPA power efficiency.
- the provision of a steerable spot beam to cope with seasonal or diurnal peak traffic distributions without the need to reconfigure the capacity allocated to any of the remaining fixed beams.
- the flexibility in the positioning of the G channel, intended to ease the tasks of coordination with other systems and to make a provision for an eventual future increase (in WARC-MOB'92) of the bandwidth allocated to the LMSS service.

4 Payload description

GENERAL: A block diagram of the LLM payload is given for illustration in Fig-5. The payload consists of the Forward and Return transponders. The Forward (FWD) transponder receives the signals from the fixed earth stations at Ku-band and relays them to the mobiles at L-band. The Return (RTN) transpouder does the reverse operation.

In the Forward transponder, the Ku-band receiver amplifies and down-converts the signals from Ku-band to L-band (1950-1979 MHz) and feeds them into the IF Processor. The FWD IF Processor (IFP) operates with L-band input and output interface frequencies. It is responsible for the routing of signals through the payload. Channel filtering is implemented using SAW filter technology , and channel-to-beam switching is realised with L-band MMIC multithrow switches. The output of the IFP is a set of 11 beam drive signals (nine active beams and two redundancy ones) at L-band (1530-1559 MHz). The L-band TX front end uses a novel multimatrix antenna feed system with overlapped beam feed clusters [1] which consists of a low power Beam Forming Network (BFN) , a set of twelve cold-redundant 10 Watt transistor power amplifiers (operating at 16 dB of noise power ratio), and a set of three 4x4 Butler-like hybrid matrices, which steer the amplified beam drive signals to the corresponding antenna feed elements. The L-band antenna is common for TX and RX, and consists of a large unfurlable reflector (5x5.6 meter projected aperture) fed from its focus by an array of twelve feeds. Each of the twelve antenna feed elements is equipped with an L-band diplexer, and in addition three of them with an orthomode transducer to provide dual circular polarisation for one beam.

The signals received by the L-band antenna are amplified by 15 LNAs which are connected to the L-band diplexers. The RX beams are configured by a passive RTN Combiner which performs the opposite of the combined function of the FWD low-power BFN and Butler-like matrices . The RTN IFP performs a function equivalent to the FWD IFP. The only difference is in the degree of channelisation of the G channel, which is increased to 1 MHz (as opposed to 4 MHz in the FWD) to protect the transponder against interference from systems using the same frequency band. The L-band output signals of the RTN IFP are up-converted to Ku-band by the RTN Up-Converter which drives a

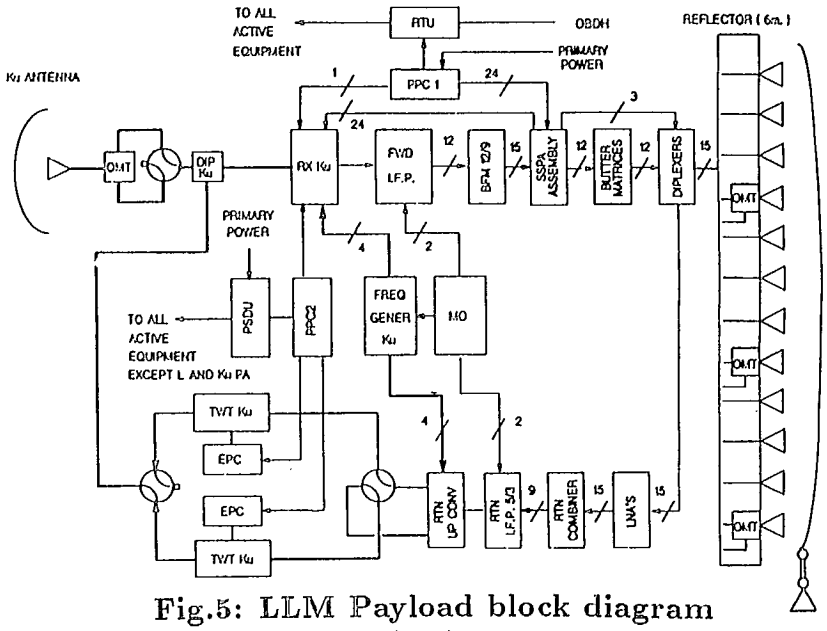


Fig.5: LLM Payload block diagram

cold-redundant Ku-band High Power Amplifier providing 20 Watts of useful output power at 21 dB of noise power ratio. The Ku-band antenna is a center-fed reflector antenna of 450 mm diameter, which is common for TX and RX.

The most characteristic subsytems of the LLM payload are the IF Processor (including Frequency Generation equipment), and the L-band Transmit and Antenna subsystems. Those are described in detail herebelow:

IF PROCESSOR: The IFP performs the onboard routing of the signals to and from the different beams. It is therefore responsible for providing an adequate interface between the Feeder Link and Mobile Link frequency plans, to handle efficiently the spectrum resources, and to provide the beam-to-beam interconnectivity paths through the payload.

The IFP contains a bank of SAW filters operating in the range 140 to 169 MHz which cover the channels M1,M2, M3 and G . The filters provide more than 37 dB of out-of-band rejection at 250 KHz away from the passband. The filter bandwidths are of 1 MHz for the filters contained in channels M1, M2 and M3 and 4 MHz for channel G. All filters are manufactured on quartz substrate in order to achieve good temperature stability. Each of the filter outputs, upconverted to L-band, can be routed to any of the 11 input ports of the BFN by the use of miniaturised MMIC multithrow switches which provide over 40 dB isolation.

In order to maximise the efficiency in handling the L-band spectrum, a novel technique of filter guardband reduction has been implemented. It consists of compressing (or expanding in the RTN transponder) the seven 1 MHz slots contained in a 8.5 MHz-wide "M" channel down to contiguous 3 and 4 MHz bands (L1 and L2) as shown in Fig-2. Each 1 MHz slot at IF is up-converted by a different local oscillator signal. Adjacent local oscillator signals are offset by an amount equal to the filter transition band (250 MHz). In this way 100 per cent spectrum efficiency is obtained at L-band.

The frequency overlapping at L-band of the channels M1, M2 and M3 would allow reusing the same frequencies in three beams at the same time. Moreover, since each of the 1 MHz frequency slots is controlled in frequency by a different frequency synthesizer of 250 KHz resolution, up to six slots can be overlapped in

frequency to provide six-fold frequency reuse, or the use of the same band in all spot beams, for a multibeam CDMA system.

The implementation of such an IF Processor is largely dependent upon the use of miniaturised frequency synthesizers. Seven such synthesizers are needed for the LLM IF Processor, the outputs of which are reused by the FWD and RTN transponders. Each synthesizer features 250 KHz resolution and 25 MHz bandwidth in the L-band frequency range (1500 MHz), with outstanding phase noise and spurious performance. The mass of each unit is 150 grams, and the power consumption 2.5 Watts.

L-BAND TRANSMIT AND ANTENNA SUBSYSTEMS: The L-band antenna has to provide simultaneously one Eurobeam, a set of six fixed spot beams and one steerable spot beam, as indicated before. In order to meet the EIRP requirements, the Eurobeam edge-of-coverage directivity has to be around 27.5 dBi (Fig-3) and the spot beam directivity around 33.5 dBi (fig-4). In addition to that, RHCP has to be provided in one of the spot beams for experimental purposes. However, the requirements which have driven to a large extent the design of the L-band antenna are the beam-to-beam isolation for frequency reuse and the need to have flexibility in RF power to beam allocation in order to cope with fluctuations of the traffic across the coverage area.

In view of the above requirements, an extensive trade-off among various candidate antenna configurations has been performed [2].

The specified coverage/gain/isolation requirements impose an antenna size of about five meter projected diameter.

Active foldable multibeam arrays can provide the required flexibility but they are complex, suffer from scan loss and the required taper for low sidelobe operation implies usually lower overall efficiency. This is also the case for hybrid array-fed reflectors for which, in addition, reflector oversizing is required, unless complex amplitude and phase control is included.

Focus-fed reflectors suffer from excessive spillover or beam cross-over losses unless overlapping feed clusters are used .

Distributed amplification using a Multiport Amplifier can be used but, in the case of the LLM payload, the required matrices would be large and complex. Also, feeding multiple feeds for each beam from the same matrix generally leads to non-uniform amplitude excitation of the amplifiers with a consequent reduction in efficiency.

In order to reduce the above problems, a novel semi-active antenna configuration has been selected for the LLM payload, the "Multimatrix-fed reflector" (ESA patent pending) [1]. In the baseline LLM configuration a deployable reflector of 5x5.6 m projected aperture is used with a feed array of 12 elements in its focus, fed by a multimatrix transmit section. Three feeds are used for each beam, with one feed shared between adjacent beams.

In a Multimatrix-fed reflector antenna, each element in a beam feed cluster is fed from a different, smaller, Butler-like matrix or gen-

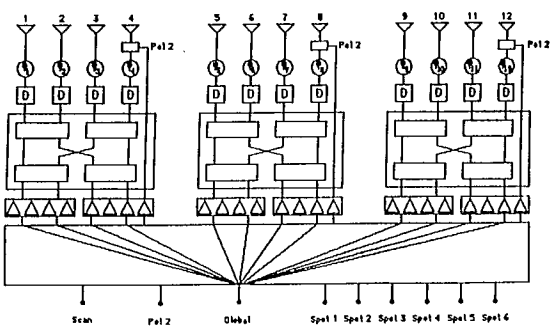


Fig.6: "Multimatrix" TX subsystem

eralised coupler, as shown in Fig-6. The total number of these matrices is equal to the number of feeeds used for each beam. Their order depends on the total number of feeds , and therefore on the total number of required beams. In the LLM payload three 4x4 output matrices are used to feed twelve LHCP feed ports. The RHCP feed ports are feed directly from the TPA section. Each matrix is fed by identical amplifiers. For each beam, the power is directed towards one matrix output port by proper relative phasing of the inputs (0,90,180 and 270 degrees phases are only required). All solid-state amplifiers operate at the same input drive level, with independence of the actual beam-to-beam traffic distribution. This feature maximises the DC to RF conversion efficiency. Power division and phasing for each beam at the amplifier inputs is obtained by a low level beam former (BFN). A fixed BFN is used for the fixed beams, whereas a variable one (5 bit phase control, no amplitude control) is used for the steerable beam and the redundancy beams.

The computed directivity countours for the six fixed spot beams generated by this antenna are presented in Fig-4, showing a crossover of about -3 dB. The steerable beam can also be "scanned" at the nominal or intermediate positions between fixed beams by sharing one or two feeds between adjacent positions. The Eurobeam is generated as shown in Fig-3. A fixed non-uniform phase excitation for the global beam is achieved by adjusting the cable length at the feed level. A compensation for this fixed taper is performed at the low level BFN for the remaining spot beams.

A failure tolerance analysis has been performed for the baseline configuration and the main conclusion is that both the spot and global coverages are seriously degraded by one amplifier failure, and consequently two-for-one cold redundancy has been provided at TPA level. Also a sensitivity analysis versus the BFN errors, the TPA amplitude and phase tracking errors and the hybrids phase and amplitude imbalance has been performed. The results show that although the spot and eurobeam directivity degradation is negligible (assuming realistic hardware implementation), the isolation, and therefore the frequency reuse capability is seriously degraded for aggregate BFN and TPA tracking errors exceeding 0.8 dB and 8 degrees.

L-BAND ANTENNA TECHNOLOGY: In view of the large reflector size required for the LLM payload, an unfurlable reflector is required. Two different reflector technologies have been adressed as candidates [3]:

- Inflatable Space Rigidized reflector
- Foldable radial rib reflector

The Inflatable Space Rigidized reflector is the present baseline. This reflector is manufactured out of prepreg gores precut and joined together. It is composed of three basic elements: A torus used to stabilise and stretch the structure, the reflector/radome membranes to seal the structure and support RF reflecting/transparent surfaces, and the structural interfaces to join the reflector to the spacecraft and to the inflation system. The wall of the reflector consists of a thin fiber reinforced composite lamina, plus a Kapton foil metallised on one side. After inflation in space (duration about 10 minutes) the solar radiation hardens the matrix by a chemical reaction and rigidises the structure from which the nitrogen is evacuated 48 hours after deployment.

Several models of this reflector have been developed and tested including a 6m centered model and a 3.5m offset model. A 10m model is presently under development. A design has been performed for the LLM mission. The reflector is folded for launch in a container providing mechanical packaging and thermal protection. In orbit, after deployment, the reflector

is attached to the spacecraft via its container serving as a supporting arm as shown in Fig-7. The total mass of the reflector subsystem including container/arm, pneumatic system and locking mechanism is estimated to be 40.0 Kg.

THE FEED ARRAY: The 12-radiators feed array is mounted on the East spacecraft wall and connected with coaxial cables to the diplexers and the active components. Candidate technologies for the feed array include printed type radiators such as patches or slots, and disk-dipoles or cup-dipoles radiators.

The technology has to provide dual CP capability and passive intermodulation and multipaction free operation up to 40 watts of CW power per feed element.

5 Spacecraft accomodation

The LLM payload is one of the three advanced communication payloads of the ARTEMIS satellite, a pictorial view of which is shown for illustration in Fig-7. The LLM L-band unfurlable reflector is attached by its arm/launch container to the East panel of the spacecraft. The L-band feed array is held in its inclined position from the Earth-facing panel .The Ku-band TX/RX reflector is located on the Earth-facing platform antenna tower. The L and Ku-band power amplifiers are located on the North panel in order to radiate the dissipated power efficiently. The L-band LNAs, L-band diplexers and Butler matrices are distributed between the East and Earth-facing panels, as close as possible to the L-band TX/RX feed array in order to minimise the RF losses. The rest of the equipments is distributed on different panels according to the accomodation constraints imposed by the other payloads.

The mass of the LLM payload is typically 150 Kg and the DC power consumption is 600 Watts. The reliability of the payload has been calculated to be 0.856 for a lifetime of 10 years.

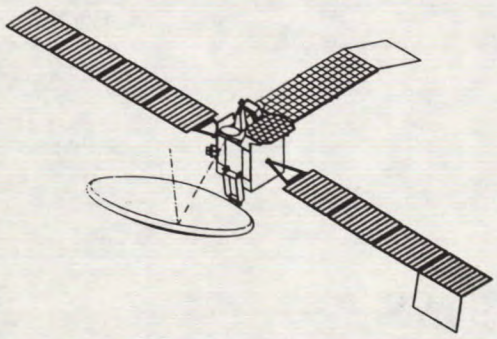


Fig.7: ARTEMIS satellite pictorial view

Acknowledgements: The authors would like to associate with this article R.Rogard, A.Siswick and M.Sabbadini from ESA and the companies Selenia Spazio, Marconi Space Systems and Space Engineering for their valuable contribution to the payload design.

References

1. A.Roederer, M.Sabbadini. A novel semi-active multibeam antenna concept. IEEE AP-S, 1990.

2. L-band Land Mobile payload study. ESTEC contract no.8531/89 Phase B2-1 Final Report, March 1990 (Selenia Spazio).

3. A.Roederer, E.Rammos. Large satellite antenna developments at ESA. ISAP 1989, Tokio.

The Design of a Linear L-Band High Power Amplifier for Mobile Communication Satellites

N. Whittaker, G. Brassard, E. Li, P. Goux,
V. Reginato, M. Coté
SPAR AEROSPACE LTD
21025 Transcanada Hwy
Ste-Anne-de-Bellevue, Québec, Canada
Phone: (514) 457-2150
Fax: (514) 457-2724

ABSTRACT

This paper describes a linear L-band solid state high power amplifier designed for the space segment of the M-SAT mobile communication system. This amplifier has been developed by SPAR AEROSPACE LTD with support from the Canadian Government. It is capable of producing 35 Watts of RF power with multitone signal at an efficiency of 25 % and with intermodulation products better than 16 dB below carrier.

INTRODUCTION

The high power amplifiers that provide the L-band downlink power are the most critical elements of the mobile satellite payload. Since 1980 SPAR has gained a solid understanding of the design implications for such amplifiers.

A number of linear HPA's have been designed and built, reflecting the evolution of the M-SAT regulatory issues, one rated at 100 Watts operating at 866 to 870 MHz[1], one at 50 Watts operating between 1.545 and 1.555 GHz [2] and one at 35 Watts operating between 1.530 and 1.559 GHz. The first two designs utilize bipolar power transistors, the third GaAs MESFETs (of which a total of six have been built). In addition to this work, various high efficiency techniques have been evaluated.

The subject of linear high power amplification is complex and requires a high degree of analytical capability particularly in the area of the evaluation of the multicarrier nature of the signal and its impact on the design. Sophisticated test set-ups are required to accurately simulate this multicarrier signal.

Key design parameters in addition to efficiency and linearity are overdrive and mismatch protection, phase and gain tracking between the amplifiers and multipaction elimination.

This paper presents design details of the GaAs FET amplifier together with measured performance.

DESIGN CONSIDERATIONS AND REQUIREMENTS

Linearity and Efficiency

Linearity must be defined in the context of the specific signal being amplified. The signal to be amplified will consist of many individual carriers spread randomly across a 29 MHz band centered at 1.5445 GHz. The resulting signal envelope will have a noise-like amplitude profile and contain frequencies from DC to 29MHz. The amplifier must be capable of processing this signal without distortion of its amplitude and frequency make-up.

In practice, some degree of distortion is acceptable allowing an increase in efficiency.

Overdrive and Mismatch protection

Because the amplifier will have to operate below saturation for linearity reasons, there is a potential risk of an above nominal demand for DC power. Clearly this cannot be allowed to happen where more than 75% of the total payload DC power is used by the HPA's. The amplifier must therefore contain a means of accurately controlling power demand, without degrading linearity.

The amplifier should be capable of surviving an open or short circuit at its output. (Note that this is usually achieved with an output isolator which also assure a good output impedance, but because of uncertain phase characteristics and a relatively high insertion loss, its use is considered undesirable.)

Phase and gain tracking

The M-SAT satellite employs a multi-beam concept where groups of up to 8 HPA's are embedded between hybrid matrices. For each amplifier's signal to split and combine efficiently, the dynamic gain and phase characteristics of each amplifier must be closely matched under all conditions.

Multipaction

The likelihood of multipaction breakdown within the unit is high considering the power levels involved. The design must be such that this phenomena cannot occur.

Power Converter

An Electronic Power Converter (EPC) is required to provide the amplifier's specific supply voltages from the spacecraft's power bus, to receive telecommands and to provide the interface for telemetry signals.

DESIGN APPROACH

The first L-Band HPA used 35 Watt bipolar transistors employing dynamic bias [2]. The major problem with this approach was the degradation of linearity as envelope bandwidth increased beyond 10MHz. In addition its complexity, hence reliability projections were disturbing.

The decision to undertake a new design in late 1987 (in reaction to the WARC-MOB'87 allocation of 29MHz) coincided with the introduction of a 20 Watt GaAs MESFET. Despite having a significantly lower power output than bipolar, the FET has intrinsic linearity, higher gain, lower thermal resistance, is relatively simple to bias, hence is an attractive alternate. Fundamental differences between the two devices result in the comparison of a complex and sensitive design (bipolar) with a simple stable and reliable one (FET).

Past experience, which included knowledge of efficiency enhancing and linearization techniques, enabled a sensible evaluation of the options. It was concluded that the best approach for the next version of HPA would be to use GaAs FETs in a simple Class A-B configuration.

The design activities were initiated by a thorough evaluation of 20 Watt L-band GaAs MESFETs. Devices from two manufacturers were tested, with similar results. The evaluation process involved the determination of quiescent (drain) current and the gate and drain wideband impedance characteristics commensurate with linearity and efficiency. This was carried out through extensive measurement characterization processes with the help of automated test systems in conjunction with non-linear computer analyses. This evaluation process considered the inherent idiosyncratic behavior of GaAs FETs, particularly when biased in Class B, such as low frequency oscillations, memory of saturation, etc.. From these results, the optimum operating

conditions were determined and an output amplifier module was designed and tested. Results from this module indicated that it would be possible to meet the basic requirements by combining four such modules in parallel.

UNIT ELECTRICAL DESIGN

The block diagram of the complete amplifier is shown in Figure 1. All the active devices of the RF chain are GaAs MESFETs sharing two common supply lines of -10 and +10 Volts.

The first module is a wideband L-Band amplifier (A1) designed with two cascaded GaAs MESFET devices. It has 20 dB of gain with a 1 dB compression point of +19 dBm.

Module A2 is a voltage controlled attenuator which not only limits the drive signal when necessary, but also provides the means to compensate unit gain change with temperature, and to accurately set the initial gain. It is followed by a voltage controlled phase shifter (A3) that allows initial setting of the unit's phase and provides compensation for phase changes with temperature.

Module A4 is a medium power amplifier with a GaAs MESFET biased in Class A. It provides a gain of 14 dB with a saturated output power above +27 dBm. It is followed by a 10 Watt GaAs MESFET amplifier (A5) biased in Class A-B that provides 13.5 dB of gain.

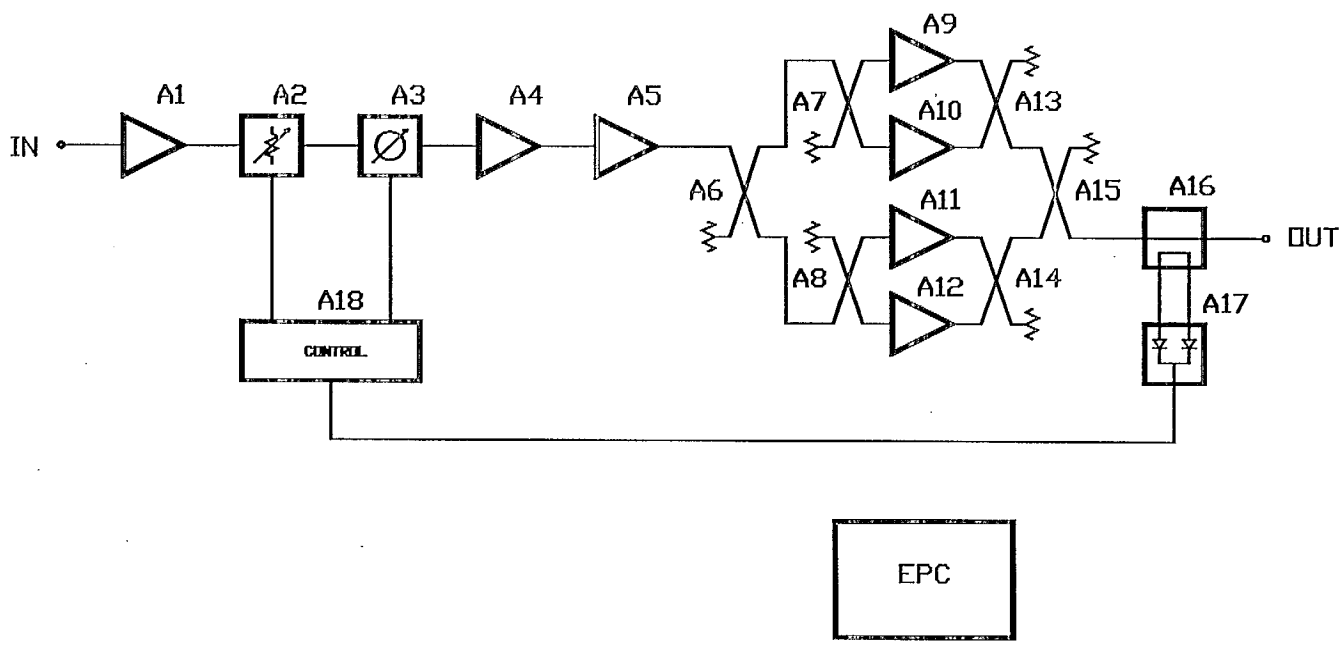


Figure 1: Block diagram of the amplifier

Four identical output modules (A9 to A12) as described above, are then combined in parallel using Lange couplers for power splitting and combining. The efficiency was optimized by careful selection of the operating point and bias impedances, along with an output matching circuit that reflects the second and third harmonics back to the device with the appropriate phases [3].

A bi-directional low loss coupler (A16) having 25 dB directivity, allows accurate signal detection for both incident and reflected power. Detection is achieved using two pairs of Schottky barrier diodes biased in their square law region (module A17). In each pair, one diode provides temperature compensation. Processing of the detected signals is accomplished in module A18, which provides output power telemetry and the control signal to the attenuator A2. A commandable override of the control signal is included. Module A18 also provides the control signal to the phase shifter A3.

The design of the Electronic Power Converter reflects the latest development in power conversion techniques. The main features of the EPC are, a high frequency of operation to minimize weight and size, excellent output load and line regulation, high efficiency (above 87% at full load), high reliability, low conducted and radiated emission levels, short circuit protection, sequential start-up to protect the HPA, and input/output monitoring.

MECHANICAL DESIGN

The HPAs' compartmentalized housing is of cast aluminium. The four output modules A9-A12 are mounted onto a common plate which is fitted into the unit from the bottom, forming the baseplate of the unit. This approach is used as the power transistors are soldered to the baseplate to obtain a well defined thermal resistance. By having a removable section, the assembly and test tasks are simplified.

RF modules are built with MIC hybrid techniques using alumina substrates mounted on kovar carriers. Packaged surface mount components are used in conjunction with circuit functions realized in microstrip transmission lines on alumina substrates.

The EPC consists of a printed circuit board assembly mounted on top of the RF chassis, the floor of the EPC forming the cover of the RF section. Connections between the two sections are achieved by hard wiring.

The dimensions of the overall HPA/EPC unit shown in Figure 2 are 28.2 cm long by 16cm wide by 6.9 cm high. The mass of the breadboard unit before weight relief is 2.9 Kg.

MEASURED PERFORMANCE

The unit development program has resulted in the production of two breadboard units and four 'quasi-engineering models'. Test results are summarized in Table 1.

All HPA's meet the basic requirements of output power, efficiency, linearity and gain. Temperature tests performed on the four QEM's show that the gain and phase tracking between units over temperature and power level are within 0.6 dBpp and 20^{o}pp respectively.

Note that the HPA's meet both input and output return loss requirements without using an isolator. Output return loss is ensured by the balanced structure of the four output modules. It has been measured with two different approaches, one by injecting a signal at the output at a different in-band frequency, and the other by performing a load pull test. Both methods confirmed a 20 dB return loss across the frequency band and over the dynamic range.

Table 1
L-Band HPA Requirements and Measured Performance

Parameter	Target	Measured
Operating Frequency Range	1.530-1.559 GHz	1.530-1.559 GHz
Nominal Output Power	35 Watts	35 Watts
Peak Output Power		80 Watts
Gain	50 dB	50 dB
Efficiency (with EPC)*	25 %	25 %
Intermodulation distortion *	16 dB NPR	16 dB NPR
Gain vs Temperature (per 10°C) (-5oC to 55°C)	0.3 dB_{pp} 0.5 dB_{pp}	0.2 dB_{pp} 0.5 dB_{pp}
Gain Tracking between units	0.6 dB_{pp}	0.6 dB_{pp}
Phase Tracking between units	28^{o}_{pp}	20^{o}_{pp}
Mass (Breadboard) (Flight Model)	 2.6 Kg	2.9 Kg
Input / Output Return Loss	20dB min.	> 20 dB
Noise Figure	10 dB max.	6 dB at 60°C
Spurious Modulation	-57 dBc max.	-68 dBc
Overdrive		
Level above nominal	up to 21 dB	up to 25 dB
Duration	indefinite	indefinite
Maximum RF Output	40 Watts	40 Watts
Maximum DC Power	155 Watts	< 155 Watts

* Measured at Nominal output power with a noise signal of 29 MHz bandwidth

The overdrive and output mismatch protection features of the HPA were tested successfully. The HPA showed no degradation in linearity (noise intermodulation) when overdriven by 25 dB. Under this condition, the output power is automatically limited to 40 Watts. The amplifier also proved that it can withstand an infinite VSWR on its output port for an indefinite time without any degradation. When submitted to such a condition, the protection loop attenuates the signal such that the amplifier is working almost at quiescent power.

Areas within the unit where multipaction is most likely to occur are the microstrip to non-microstrip signal interfaces. Representative sections of the unit were built and tested in vacuum at levels at least 6dB above the highest expected power. In addition, the Lange coupler and the directional coupler were both tested at 180 Watts. No evidence of breakdown multipaction was detected.

CONCLUSIONS

This paper has presented the key design considerations for a spaceborne linear high power amplifier.

Using GaAs FET's as the key device, not only is the resulting amplifier's performance satisfactory, but the design is conceptually simple. In conjunction with the inherent reliability of the FET, this results in a unit that is optimum for space application.

REFERENCES

1. **Whittaker, N. Brassard, G. Butterworth, J.S.** 1986. The design, evaluation & modelling of a UHF power amplifier for a mobile satellite transponder. IEEE International Conference on Communications 1986.

2. **Brassard, G.** 1988. Development of a linear L-band High Power Amplifier for satellite application. RF EXPO EAST 1988.

3. **Tyler, V.J.** 1958. A new High Efficiency High Power Amplifier. Marconi Review 21 pp96-109.

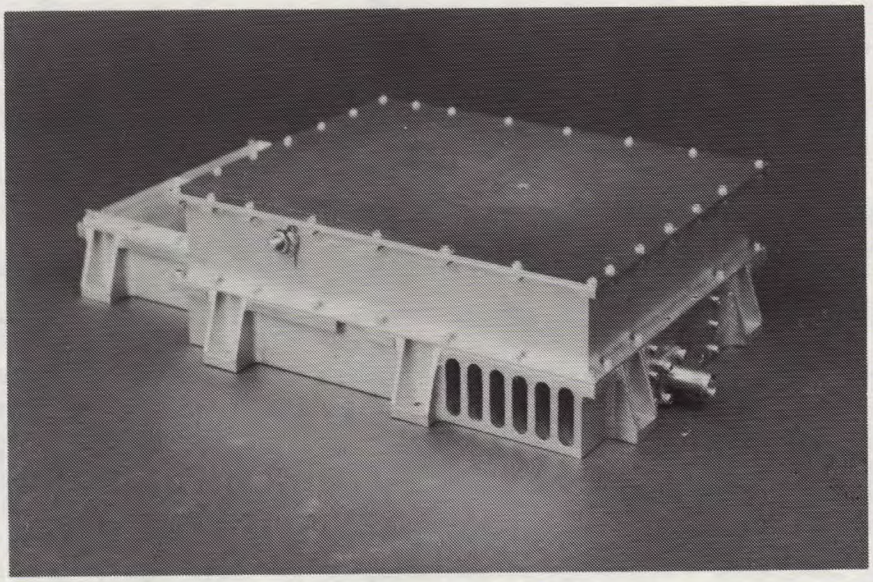


Figure 2: Overall HPA/EPC unit

A Method for Modelling Peak Signal Statistics on a Mobile Satellite Transponder

André Bilodeau, Michel Lecours
Marcel Pelletier, Gilles Y. Delisle
Laval University
Quebec City, Canada G1K 7P4

ABSTRACT

This paper proposes a simulation method to model the peak duration and energy content of signal peaks in a mobile communication satellite operating in an FDMA mode and presents an estimate of those power peaks for a system where the channels are modeled as band limited gaussian noise, which is taken as a reasonable representation for ACSSB, MSK or PSK modulated signals. The simulation results show that, under this hypothesis, the level of the signal power peaks for 10%, 1% and 0.1% of the time are well described by a Rayleigh law and that their duration is extremely short and inversely proportional to the total FDM system bandwidth.

INTRODUCTION

Assuming that the transmitting power in a mobile communication satellite is to be provided by power amplifiers whose instantaneous power consumption is related to the instanteneous T_X power output, it is of interest to specify the peak transmit power levels which will be required and the duration of those peaks. More specifically if the spacecraft power subsystem is sized to provide power up to the 99% point of the average power distribution, it is important to estimate this value, and to estimate the statistical properties of voltage addition peaks at higher probability levels (i.e.: 99.9%) in order to optimize the short term characteristics of the power amplifier response.

If the channels on the satellite transponder are to be ACSSB, MSK or PSK modulated, this question may be viewed somewhat differently than in the case where the satellite system is required to transmit a large number of narrowband modulated carriers. The individual channels may then be considered to some extent as band-limited gaussian noise.

This paper studies then this question of peak signal statistics on a transponder where the individual channels can be modeled as band-limited gaussian noise. This is analysed for a general transmission system configuration of one thousand voice-activated 5 kHz bandwidth channels with an SCPC-FDMA configuration, which are modeled as 400 fully activated channels randomly chosen inside the FDM band.

THEORETICAL FOUNDATIONS

The theoretical foundations for such an analysis have been developed in the classical papers by S.O. Rice on "Mathematical Analysis of Random Noise", [1,2] In particular, Part III of this ensemble of two papers discusses the statistics of random noise currents: *Probability distributions associated with the maxima of the currents and the maxima of its envelope are developed. Formulas for the expected number of zeros and maxima per second are given...* The results have been presented and expanded in a number of classical textbooks, such as the one by J.S. Bendat [3] , where one finds in chapter 3 a discussion of the applications of probability theory to random noise analysis: expected number of maxima per unit of time, distribution of the envelope.

Random noise can be regarded as the sum of a large number of independent events and hence can be related to a gaussian statistical law. The case of noise characterized by particular spectral properties (such as band-limiting) can be studied through the use of its power spectrum and correlation function, which results in gaussian-distributed signal models. The probability of these signals exceeding a certain peak can be easily estimated, and the relationships for signal power peaks can be derived. It is also possible to show that the signal envelope will follow a Rayleigh distribution.

These results permit then to compute the probability that a signal exceeds a certain power level, as is the concern here.

EXPECTED NUMBER OF LEVEL CROSSINGS

Let us consider the random noise y(t) and its time derivative $y'(t)$. Their joint probability expresses the probability that y(t) lies in the interval (y,y+dy) when its derivative is in the interval $y', y'+dy$. To find the expected number of crossings of y(t) through the interval (y,y+dy), the amount of time that y(t) is in the interval must be divided by the time required to cross the interval, which is related $y'(t)$ to the derivative of y(t). Hence, the expected number of passages per unit of time of y (t) through the interval $y'(t) = y'$ for all values of is given by.

$$N(y(t) = y) = \int_{-\infty}^{+\infty} |y'| f(y,y') dy' \qquad (1)$$

where the absolute value of y' is used since crossing time must be a positive quantity. So, if y = 0, one will obtain the expected number of zeros. For other values of y, one will obtain the expected number of passages through given levels.

This is directly y applicable to estimating the probability of a signal having peaks over a certain power level as is the concern here. One field of application of this concept has been in the statistical analysis of fades in the mobile radio-channel. Lee [4] (chapters 2 and 5) gives results for the probability density for normal-distributed signal levels in dB.

EXPECTED DURATIONS OF MAXIMA AND MINIMA

Rice [1] had developed the basis on which to estimate the durations of maxima and minima (peaks and fades). This question appears, however, to have been developed further by researchers in the mobile radio field, who have applied it to the estimation of fade durations. It can be as easily applied to the estimation of peak durations.

According to Lee [4] (chapters 2 and 5), for a signal , the relation between the average duration of the signal fade (or of the signal peak) and the expected number of passages per second at a particular level y is given as the ratio between the expected amount of time per second where y(t) is below (or above) the given level y and the average duration of fades under (or of peaks above) this given level.

Results are given by Lee [4] (page 73) for average durations of fades under a given level in the case of a Rayleigh distributed variable, whose time derivative is gaussian distributed. These results could be adapted to the estimation of peak durations, as is the concern here.

ADVANTAGES OF A SIMULATION APPROACH

It has been shown in the previous section that a good theoretical foundation has been established for the understanding and the analysis of the question under concern here, namely the modeling of the peak duration statistics for band-limited noise channels in an FDM configuration.

From this theoretical foundation, one expects the satellite channel signal to be normal-distributed. One expects its envelope to be Rayleigh-distributed. One expects then the probability of the power peaks to follow

those of a Rayleigh distributed signal. As illustrated, for example in Lee [4], page 175, the 90%, 99%, 99.9% power levels would be about 3,7 or 9 dB above the signal rms value 10% 1% or 0.1% of the time.

One would then have to make the required assumptions on the signal statistics and on its power spectra to undertake the analytical derivations or the numerical integrations leading to the estimates of the frequency of occurence of these power peaks and to the estimates of their durations.

It appears then legitimate to ask a number of questions about the quantitative aspects of the model. Let us admit the mobile satellite channel is normal-distributed with a power spectrum and a certain variance. But one needs also to make assumptions about the power spectrum and the variance of the derivative in order to compute the joint probability density function f(y,y') At a certain point, the level of confidence in the quantitative results becomes limited. Quoting from J.S. Bendat [3] , page 128, it can be written that:

"In applying those formulas, many assumptions about the physical nature of y(t) are involved so that it is quite gratifying to learn, when one is through, that good agreement with empirical data has been found."

So, for the completion of this study, and in order to obtain with a high degree of confidence the quantitative results desired on the peak traffic statistics, it was judged preferable to proceed to the simulation of the FDM channel, or rather of a restricted subset of the channel.

SIMULATION

As mentioned earlier the general transmission system configuration was one with one thousand voice-activated 5 kHz bandwidth channels with an SCPC-FDMA configuration, modeled as 400 fully activated channels randomly chosen inside this 5 MHz FDM band.

The simulation was carried out on a SUN workstation. For consideration related to the speed and capacity of the work station, it was found impractical to simulate the FDM system mentioned above (400 channels out of 1000 in a 5 MHz bandwidth). Simulations were then carried out for 4 channels out of 10 in a 50 kHz bandwidth, and for 40 channels out of 100 in a 500 kHz band width: the results were then extrapolated for the case of 400 channels out of 1000 in a 5 MHz bandwith.

To realize the individual 5 kHz channel simulation, a gaussian noise over a 10 kHz bandwidth was generated and convolved with a 3 kHz 3 dB bandwidth baseband filter having a 30 dB attenuation outside the 5 kHz band.

SIMULATION WITH BASEBAND CHANNELS

The simulation procedure was first tested for the case where forty 5 kHz baseband channels are summed. With the signal amplitude normalized for a power of unity, the power probability distribution and the density function were computed. From these, power levels of 1.9, 6.9 and 8.7 db respectively over the rms value for 10%, 1% and 0.1% are measured or calculated directly from the simulated signal: these values are well in agreement with a Rayleigh model (Lee [4], p. 175). The frequency of occurrence (number/second) of these power peaks is 185.6, 31.7 and 2.4 per second.

SIMULATION WITH 5 KHZ CHANNELS IN FDM CONFIGURATION

As stated in section 4, the simulation took into account a maximum number of forty out of one hundred 5 kHz bandpass channels, which was a compromise between the calculation time and the precision of the results. In order to estimate the results for four hundred out of one thousand bandpass channels in an FDM configuration, the relation between the bandwidth and the statistical parameters had to be obtained. The results for four out of ten 5 kHz bandpass channels and for forty out of one hundred 5

kHz bandpass channels were computed, and extrapolated.

To realize this simulation, a gaussian noise over an appropriate bandwidth was generated and convolved with a bandpass filter of 3 kHz bandwidth centered at a given frequency f_c. This operation was repeated for different values of f_c. The next step was to sum these filtered signals to obtain the total signal.

RESULTS AND CONCLUSION

As a result of these simulations, and of the extrapolation to the case of four hundred active 5 kHz channels out of one thousand in an FDM configuration, the following results have been obtained for the signal level distribution, the frequency of occurence and the duration of power peaks.

The signal level distribution is well described by a Rayleigh law and the signal levels in dB, above the signal rms value, for the 90%, 99% and 99.9% probability levels are in the order of:

90%: 1.6 dB above the signal rms value
99% : 6.9 dB above the signal rms value
99.9% : 9.1 dB above the signal rms value

In other words, the instantaneous signal power will be 1.6 dB above the rms value 10% of the time, 6.9 dB above the rms value 1% of the time and 9.1 dB above the rms value 0.1% of the time.

For the frequency of occurence (per second) of the peaks above the 90%, 99% and 99.9% probability level, it is estimated to be in the order of:

90%: (1.6 dB): 400 000 per second
99%: (6.9 dB): 49 000 per second
99.9%: (9.1 dB): 4 900 per second

In other words, the instantaneous power level will rise 1.6 dB above the rms value 400 000 times/second, will rise 6.9 dB above the rms value 49 000 times per second and will rise 9.1 dB above the rms value 4 900 times per second.

As for the duration of those peaks, they are extremely short. 90% of the peaks which exceed the signal rms value by 1.6 dB, 6.9 dB or 9.2 dB will have a duration not exceeding 0.28, usec, 0.18 us or 0.19 usec respectively

As a conclusion, it may be said that simulation is an efficient and quick method to obtain information on the statistical power structure of a signal in a satellite transponder. One basic issue however is to have a valid statistical representation of the signal carried by the individual channels. To that extent, it would be important for future work to include such questions as a more accurate modeling of the user channel signal for specific types of modulation such as ACSSB or QPSK, taking into account the problem of possible addition of discrete spectral components present in the individual signals.

The authors wish to acknowledge the financial and scientific support of SPAR Aerospace Montreal, and in particular the advices of MM. K. Hing, S. Irani, R. Bélanger and of Dr. H.J. Moody.

REFERENCES

1- **Rice S.O.** "Mathematical Analysis of Random Noise" (Parts I, II) Bell System Tech. J., vol 23, pp 282-332, july 1944

2- **Rice S.O.** "Mathematical Analysis of Random Noise" (Parts III, IV), Bell System. Tech. J., vol 24, pp 46-156, july 1945

3- **J.S. Bendat** "Principles and Applications of RAndom Noise Theory " (chapter 3), Wiley, 1958

4- **W.C.Y. Lee** "Mobile Communications Engineering" (chapter 6), Mc Graw-Hill, 1982.

Enhanced Spectral Efficiency Using Bandwidth Switchable SAW Filtering for Mobile Satellite Communications Systems

Robert Peach and Alastair Malarky
COM DEV
Cambridge, Ontario, Canada

ABSTRACT

Currently proposed mobile satellite communications systems require a high degree of flexibility in assignment of spectral capacity to different geographic locations. Conventionally this results in poor spectral efficiency which may be overcome by the use of bandwidth switchable filtering.

Surface acoustic wave (SAW) technology makes it possible to provide banks of filters whose responses may be contiguously combined to form variable bandwidth filters with constant amplitude and phase responses across the entire band. The high selectivity possible with SAW filters, combined with the variable bandwidth capability, makes it possible to achieve spectral efficiencies over the allocated bandwidths of greater than 90%, while retaining full system flexibility. Bandwidth switchable SAW filtering (BSSF) achieves these gains with a negligible increase in hardware complexity.

INTRODUCTION

Mobile satellite communications systems must operate with very limited bandwidth allocations and must serve a wide variety of inexpensive terminals. For a given transmitter power, the EIRP and figure of merit (G/T) are enhanced by dividing the spectrum between a number of spot beams, each serving a different geographical region. For maximum operational efficiency, flexibility is needed in the allocation of bandwidth to different beams; this generally requires a large number of filters, with a corresponding loss in spectral efficiency due to the guardbands between adjacent channels. The conflict between flexibility and spectral utilisation can be removed by the use of contiguous filter banks, where adjacent filters may be operated simultaneously to provide a single continuous channel.

The second section of this paper describes the various filtering options, and shows the advantages of bandwidth switchable filtering using contiguous filter banks. The third section considers surface acoustic wave (SAW) transversal filters, describing both currently achievable performance and basic performance tradeoffs, and shows that they are the most suitable components for this application. The fourth section discusses the design of bandwidth switchable SAW filter banks, and the fifth section describes experimental data obtained at COM DEV on a trial three-channel filter bank at 150 MHz.

BANDWIDTH SWITCHABLE FILTERING

Figure 1 shows the general transponder configuration while Figure 2 shows the corresponding frequency channelisation. Bandwidth switchable filtering is a technique for replacing adjacent filters in a particular beam by a single filter covering the same total bandwidth; as illustrated in Figure 3, this avoids wasting the guardband regions between individual filters. If a continuous band of frequencies is always assigned to a particular beam, then the filters within that band may be replaced by a single equivalent filter. With this scheme, guardbands are only required between beams and not between individual filters, and spectral efficiency is greatly increased.

The simplest method of implementing a bandwidth switchable filter system would be to provide separate equivalent filters for every combination of adjacent channels that might be employed. However, for any reasonably flexible system, this would involve a large increase in system cost and complexity, not to

mention the extraordinarily difficult shape factors that would be required for the broadband filters. As discussed later in the paper, SAW technology offers a solution to this problem. SAW filter banks may be designed such that adjacent filters operated in parallel act as a single continuous filter. This provides all the advantages of bandwidth switchable filtering, while retaining essentially the same transponder configuration as that illustrated in Figure 1.

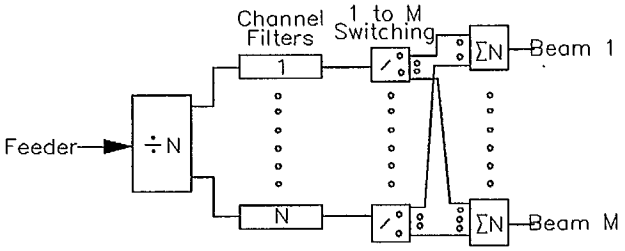


Figure 1. Flexible Transponder Configuration

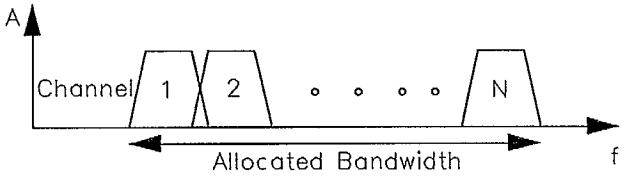


Figure 2. Spectral Channelising

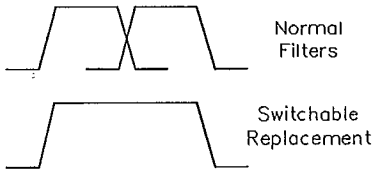


Figure 3. Switchable Filters

Spectral efficiency is defined as the ratio of the usable bandwidth to the allocated channel bandwidth [1]. For mobile communications systems, the available bandwidth (WARC 87) is small. This puts a premium on the efficient utilisation of the bandwidth due to its sensitivity in determining the revenue earning capability of satellites. In addition, spectral efficiency is also of importance on the satellite feed links where there are also bandwidth restrictions. This is of particular importance where the L-band spectrum is reused.

Mobile satellite communication systems also require flexible assignment of communications capacity between beams. Conventional filtering systems can only achieve this at the expense of spectral efficiency or at increased cost and complexity in the satellite.

One method of subdividing the L-band spectrum is by the use of filters which overlap at their passband edge as shown in Figure 4. By assigning adjacent filters to non-overlapping beams, no L-band spectrum is lost. This approach has a number of deficiencies. The constraint of adjacent filter assignment to non-overlapping beams severely restricts the flexibility of the system while the translation to the feeder band, as shown in Figure 4, also restricts the designer in employing bandwidth switchable filtering. It restricts each beam to employing a maximum of 50% of the spectrum separated into non-adjacent channels. Also, the feeder allocation requires at least two times the L-band bandwidth for each frequency use, unless individual local oscillators are used for the filters, which is prohibitively inefficient and costly. Additional problems are also encountered due to the overlapping regions of the filters.

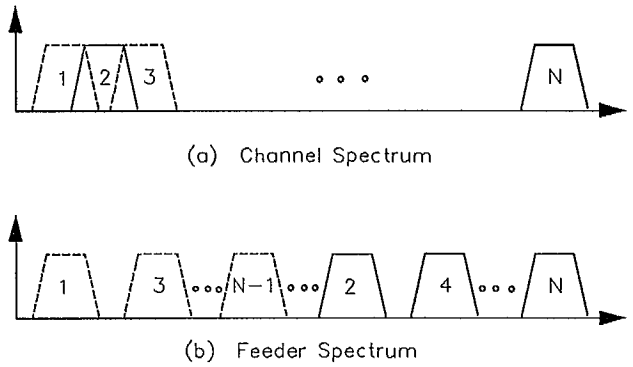


Figure 4. Overlapping Spectrum Channelisation

Non-Overlapping Filters

The preferred method of subdividing the L-band spectrum is by the use of filters which do not overlap within the passband, as shown in Figure 2. This approach allows any beam to access any number of channels up to the full L-band capacity and is not restricted in flexibility. Also, the feeder allocation only requires the L-band bandwidth for each frequency use. However this approach cannot achieve 100% spectral efficiency. The minimum passband separation required is dictated by the filter transition width. The spectral efficiency is thus directly related to the filter characteristics and the number of filters by the following equations:

$$B = (B_T-(N-1)B_t)/N \quad (1)$$

$$E = 100\,N\,B/B_T = 200/(1+S);\ N>>1 \quad (2)$$

B_T total bandwidth available
N number of channels across band
B_t filter transition bandwidth
B bandwidth for each channel
S filter shape factor = $1+2B_t/B$
E effective system L-band spectral efficiency (%)

Bandwidth Switchable Filtering

The bandwidth switchable filtering approach uses non-overlapping filters as shown in Figure 2 but can also connect filters which cover any combination of channels. By only assigning adjacent filters to individual beams, the spectral efficiency becomes dependent only on the number of beams simultaneously sharing the L-band spectrum, the level of frequency reuse and the number of unswitched divisions of the band. These divisions occur where the L-band spectrum is divided into a number of independent sub-bands, each of which is supported by a bandwidth switchable filter bank. The maximum spectral efficiency and usage is given in these cases by the following equations:

$$U = 100(1-(D-1)B_t/B_T) \quad (3)$$

$$E = 100(1-((DM/R)-1)B_t/B_T) \quad (4)$$

M number of beams in the system
D number of sub-bands
U maximum L-band spectral usage per beam (%)
R frequency reuse factor
E effective system L-band spectral efficiency (%), when all M beams are active in each sub-band.

Comparison of Filtering Options

Table 1 shows a relative comparison of the options discussed above. The system assumed uses the WARC forward L-band allocation of 29 MHz, excluding the 1 MHz search and rescue (SAR) band; it has 9 beams and reuses the L-band spectrum 3 times. To achieve system flexibility the spectrum is subdivided into 4 separate sub-bands, each of which is further divided by 8 filters, each with transition bandwidths of 200 kHz.

For mobile satellite communications, where capacity assignment flexibility is a key requirement, bandwidth switchable filtering offers the optimum performance in all respects and achieves very high spectral efficiency.

Table 1 : Comparison of Filter Options

Parameter	Over-lapping	Non-Over-lapping	Switched
L-band spectral efficiency (%)	100	77.9	92.1
Maximum L-band spectrum available per beam (%)	50	77.9	97.9
Capacity assignment flexibility	Poor	Good	Good
Feeder spectral efficiency (%)	50	77.9	92.1

SURFACE ACOUSTIC WAVE FILTERS

Proposed mobile satellite systems typically require filters with amplitude and phase ripples of 0.8 dB and 5° p-p respectively, transition bands of 200 kHz, channel bandwidths from a few hundred kilohertz to a few megahertz, and rejections of 40 dB close-in and 50 dB further out. The bandwidth and shape factor specifications limit the IF frequency to the VHF region, and the only possible technologies are miniature LC filters, crystal filters, or SAW transversal filters. Low inductor Q makes LC technology unsuitable for highly selective filters, and crystal resonators cannot accommodate the required range of bandwidths. SAW filters, however, can meet the proposed specifications, and they have the additional virtues of small size, low weight, and linear phase response; furthermore, they have a very simple structure which makes them highly reliable.

The basic structure of a SAW bandpass filter is shown in Figure 5. It consists of a polished substrate of some piezoelectric material with an aluminum electrode pattern formed photolithographically on the surface. The electrodes form two separate transducers, each consisting of a pair of busbars with numerous parallel electrodes running between them. The surface acoustic wave propagates in a direction parallel to the busbars and perpendicular to the electrodes. The principle of operation is almost identical to that of a finite impulse response (FIR) digital filter. Each transducer acts as a tapped delay line with each electrode functioning as an individual tap; the weighting function is implemented by moving the positions of the electrode breaks in the region between the busbars. The weighting function is closely related to the impulse response, which is sin x/x for a bandpass filter, and this can be clearly seen by examining the break pattern in a SAW transducer.

For classical filters the complexity is usually determined by the required number of poles. This concept is not applicable to SAW filters, and a more appropriate measure is the required impulse response duration. If a bandpass filter has a passband ripple of $20 \log\left(\frac{1+\delta_p}{1-\delta_p}\right)$ dB and a stopband level of $20 \log \delta_s$, then it is known semi-empirically that:

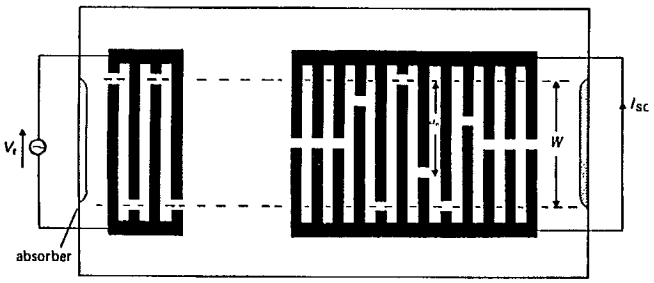


Figure 5. Simplified SAW Bandpass Filter Structure

$$\log (\delta_p \delta_s) \cong -1.05 - 1.45 B_t T, \quad (5)$$

where B is the transition bandwidth and T is the required length of impulse response [2]. The impulse response duration is therefore independent of the absolute filter bandwidth, and for the proposed specification all filters would be of essentially the same size. Shape factor is not directly constrained by impulse response duration; however, low shape factors require sin x/x weighting functions with narrow mainlobes and numerous small sidelobes which are difficult to implement accurately.

SAW devices usually employ one of three materials, lithium niobate, lithium tantalate, or quartz. Lithium niobate exhibits the strongest piezoelectric effect, and is good for broad band devices, but it has a poor frequency/temperature coefficient (-75 to -94 ppm/°C depending on orientation). At the other extreme, quartz is only weakly piezoelectric, but it can have a zero first-order temperature coefficient. The most popular quartz substrate, the ST-cut, has a frequency temperature law:

$$\frac{\Delta F}{F} = -3 \times 10^{-8} (T-T_o)^2. \quad (6)$$

T_o is nominally 25°C, but can be varied by slight adjustments to the substrate orientation. Hence, a 100 MHz quartz filter with a 60°C operating range would move by only 2.7 kHz if T_o were placed in the centre of the range. For satellite applications, size, weight, reliability and power consumption are at a premium, and quartz filters, which avoid the use of ovens, are the preferred choice.

Lithium niobate filters rarely have both transducers in line as in Figure 5; instead, they are linked by another component called a multistrip coupler; this has the advantage of allowing each transducer to be weighted without restriction. For quartz the transducers must be placed directly in line, and although unrestricted weighting, called apodisation, can be applied to one transducer, the electrode breaks in the other must be placed at the extremes of the aperture, adjacent to one busbar or the other; this restricted form of weighting is usually called withdrawal weighting. The performance achievable with quartz filters is reduced by these constraints, but it is still adequate for the purpose.

The following list of parameters represents proven performance limits for quartz. All limits can be achieved simultaneously and some can be exceeded at the expense of others.

Passband amplitude ripple, dB	<0.5
Passband phase ripple, p-p	<3°
Stopband rejection, dB	>50
Centre frequency, MHz	<200
Shape factor	>1.2
Fractional bandwidth, %	<5
Transition bandwidth, kHz	>200

The fractional bandwidth limit is imposed by the low piezoelectric coupling of quartz; it is likely that the limits on passband ripple and shape factor will be improved in the future. It is advisable to restrict centre frequency to 200 MHz; this is not a firm limit, but filter performance does start to degrade beyond this point.

The insertion loss of SAW filters is usually fairly high, typically 20 - 30 dB. This is not due to acoustic attenuation, which is very small, but to the bi-directionality of the transducers and to the electrical mismatch with the source and load. The high mismatch makes the filter characteristics insensitive to variations in the source and load impedance, and reduces unwanted acoustic signals. There are techniques for making true low loss SAW filters, but these increase complexity and reduce filter performance. For the proposed satellite application, the simplest device structure operated with a 25 - 30 dB insertion loss appears to be the most appropriate choice.

The surface wave velocities for all commonly used materials are in the range 3000 - 4000 m/s; a microsecond of impulse response therefore translates into 3 - 4 mm of substrate length. Allowing for various margins, and for factors such as transducer spacing, the overall length of devices conforming to the proposed specification would be 6 - 7 cm.

BANDWIDTH SWITCHABLE SAW FILTERING

The concept of bandwidth switchable SAW filtering (BSSF) is quite simple [3]. In this implementation it consists of a bank of contiguous filters with overlapping transition regions. Each individual filter must satisfy the general channel specification, and when adjacent filters are operated in parallel the responses in the common transition region must add vectorially to produce a continuous, composite passband response.

For a quartz SAW filter most of the frequency selectivity is provided by one transducer, which is consequently much larger than the other. If the tap weights in this transducer are $\{h_n\}$, and the electrode time spacing is τ, then the frequency response can be written:

$$H(\omega) = A(\omega) \sum_n h_n e^{-j\omega n\tau},$$

where $A(\omega)$ includes the response of the second transducer. The design problem therefore involves the approximation of the desired response by a linear function of the tap weights. This problem may be solved by one of two standard methods, the most popular being the Remez exchange algorithm [4]; the alternative approach is to use linear programming, which offers unrivalled flexibility, but is computationally expensive. These optimal methods have now largely displaced window functions as a design tool.

With these methods it is easy to design filters with specified transition region characteristics. Fortunately, the transition response required for BSSF differs only slightly from the unconstrained form, and other parameters are little affected, although the filter is somewhat squarer than it would otherwise have been. Linear phase SAW filters are particularly easy

to design, and two such contiguous filters designed to cross at their 6 dB points are quite a good approximation to the BSSF requirement.

As will be seen in the next section, the key to success in designing individual filters and BSSF systems is in the modelling of second order effects, particularly diffraction and circuit loading. Current simulation techniques are very accurate, but the very high precision required for satellite channeliser applications shows that weaknesses still exist.

EXPERIMENTAL RESULTS

The results described in this section represent the initial data from the first BSSF designs attempted at COM DEV. To simplify the preliminary assessment, the BSSF filters were designed to operate in a 50 Ω system without matching circuits, and this accounts for the high insertion loss. In practice, matching circuits would be used and the effects of these would have to be allowed for in the design. However, the initial purpose was to assess the intrinsic properties of the SAW devices without extraneous complications.

The BSSF trial filter set consisted of three devices, each with design ripples of 0.3 dB and 2°p-p in the passband, and transition widths of 200 kHz. The centre frequencies were 147.8 MHz, 150 MHz, and 152.2 MHz. Practical BSSF systems would probably use filter bandwidths less than the value of 2 MHz chosen here; however, the purpose of this exercise was to define performance limits, and for this reason filters with a low shape factor of 1.2 were selected. Figure 6 shows the superimposed response of the three filters as measured directly on a network analyzer; no matching circuits were used; the amplitude levels were within 1 dB of prediction and so closely matched that no balancing was necessary. For greater clarity, Figures 7 and 8 show the experimental and theoretical responses respectively for the 150 MHz filter. The agreement is very close, except for the stopband region which is degraded by spurious acoustic signals. The stopband in Figure 6 is limited by electromagnetic breakthrough in the device package, and this effect has been gated out in Figure 7 to give a clearer view of the acoustic response.

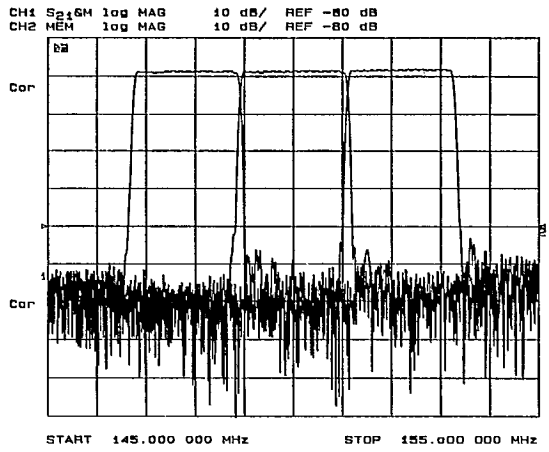


Figure 6. Superposed Filter Responses

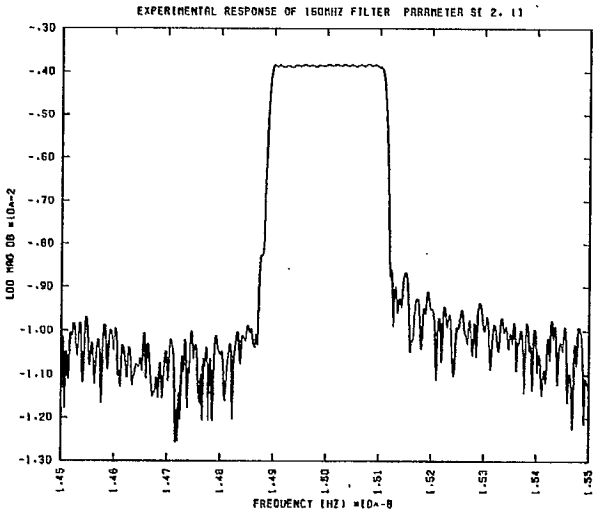


Figure 7. Gated Response of 150 MHz Filter

To assess the properties of the devices when interconnected in BSSF mode, the four S-parameters of each filter were accurately measured, and the various modes of interconnection were then simulated on the computer. This approach was used in the interests of speed and flexibility.

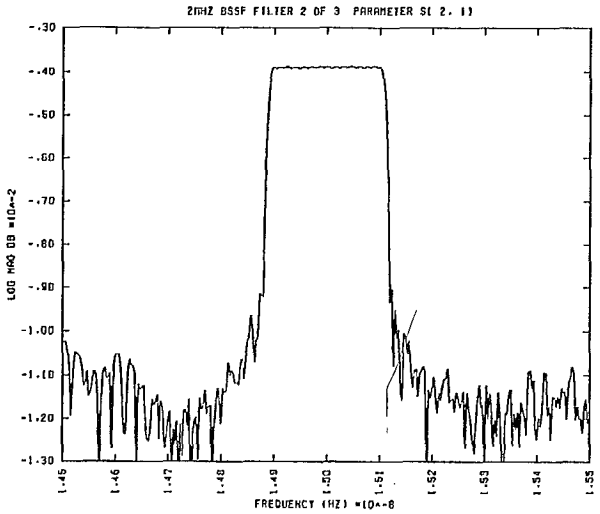


Figure 8. Theoretical Response of 150 MHz Filter

The first method of interconnection employed 50 Ω power dividers to couple the input and output ports. Figure 9 shows the composite response, again using gating to remove electromagnetic breakthrough; Figures 10 and 11 show detailed passband amplitude and phase responses without gating. The experimental filters showed absolute phase offsets of several tens of degrees, though this was as great within devices of the same type as between devices of different types. The responses in Figures 9, 10 and 11 therefore used phase shifters to balance the phase response, but no amplitude adjustment was employed. The degradation of amplitude ripple in the transition regions can be clearly seen, though the target specification of 0.8 dB is almost met. The excess amplitude ripple can be considerably reduced by adjusting the phase shifters, but this degrades the phase response.

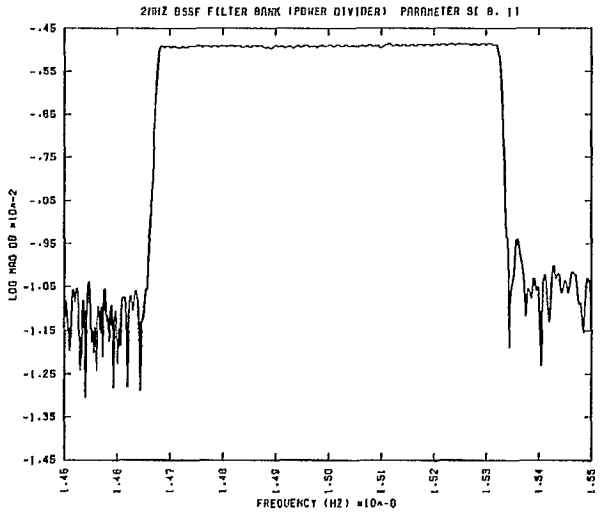


Figure 9. Combined Filter Amplitude Response Using Power Divider Coupling

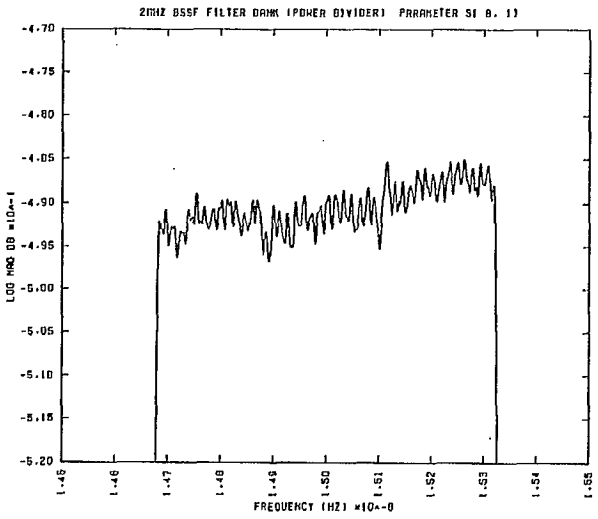


Figure 10. Combined Passband Amplitude Response Using Power Divider Coupling

The slight tilt on the filter passbands is in accord with the design, as they were intended to operate in BSSF mode with the inputs and outputs directly connected. The corresponding responses for this mode of interconnection are shown in Figures 12, 13, and 14. In this case, a small degree of amplitude correction (0.4 dB) was employed. However, the overall ripple was

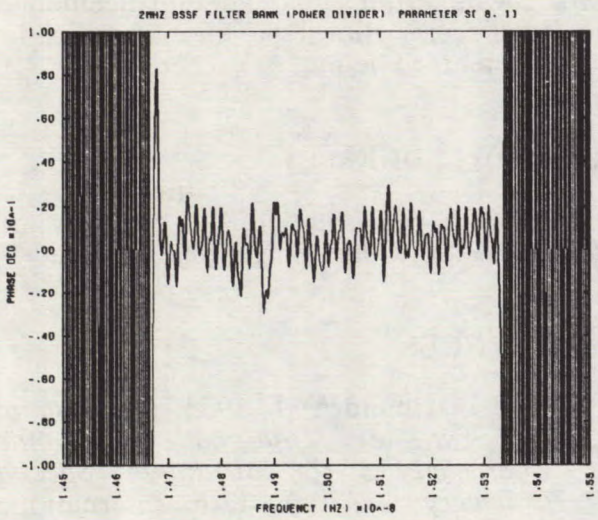


Figure 11. Combined Phase Response Using Power Divider Coupling

not greatly different from that in the power divider case. The use of direct coupling at either the input or the output ports has advantages in reducing component count in the final system, but it may produce difficulties due to reduced isolation.

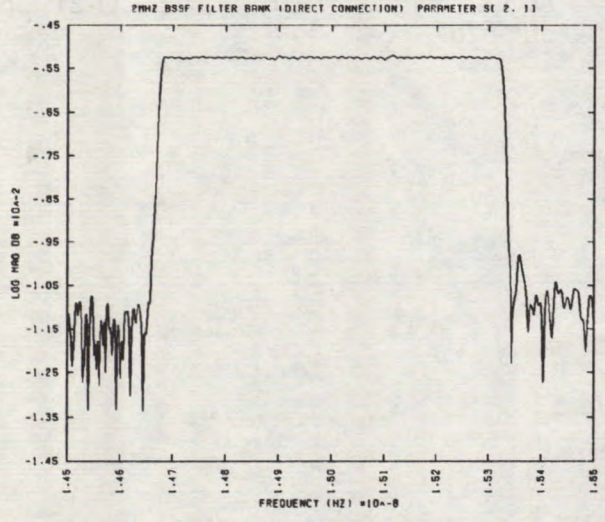


Figure 12. Combined Filter Amplitude Response Using Direct Coupling

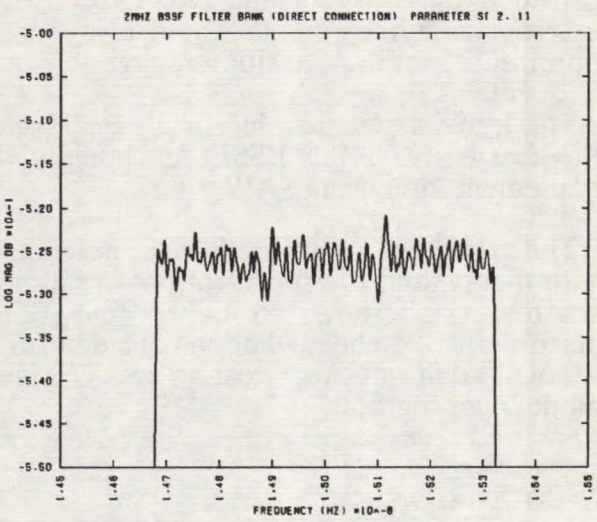


Figure 13. Combined Passband Amplitude Response Using Direct Coupling

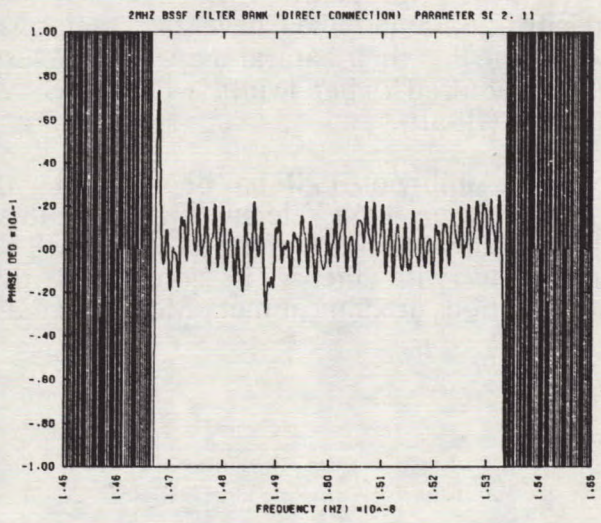


Figure 14. Combined Passband Phase Response Using Direct Coupling

These results show that the SAW devices are intrinsically very well balanced in amplitude. The absolute phase shift through the devices is over 6×10^5 degrees, and some phase imbalance between production filters is expected. However, it appears to be somewhat high at the moment, possibly due to strain in the substrate mounting. Some external adjustment of

amplitude and phase balance will always be necessary, and, in practice, this would be largely provided by the matching circuits. Use of external circuit elements would also permit optimisation of the transition region.

This initial assessment has firmly established the feasibility of the BSSF concept and its implementation using SAW filters.

The measured results are in close agreement with theory, and the specifications are almost satisfied. There is no doubt that BSSF performance can be further enhanced by fairly simple additions to existing SAW filter modelling programs.

CONCLUSIONS

It has been demonstrated that bandwidth switchable filter banks can provide flexibility in the allocation of available channel capacity without sacrificing spectral efficiency. SAW transversal filters are the only devices that can meet the proposed mobile satellite specifications; and they have the additional advantage that their natural shape is very close to that required for bandwidth switchable SAW filtering (BSSF).

The feasibility of BSSF has been tested with a set of three, high selectivity quartz filters chosen to test the limits of the technology. The agreement with theory is good, and the specifications are almost met. Measured data indicates that simple refinements in device modelling and optimisation of matching circuits will permit significant enhancement of performance, allowing all specification requirements to be met or exceeded.

ACKNOWLEDGEMENT

This work was funded, in part, by the Department of Communications, Canada.

REFERENCES

[1] **Kudsia, Chandra M.** 1988. *Optimisation of Satellite Transponder Channel Characteristics for Maximum Spectral Efficiency.* AIAA 12th International Communication Satellite Systems Conference, pp. 581-586.

[2] **Morgan, D. P.** 1985. *Surface-Wave Devices for Signal Processing.* Elsevier, Amsterdam.

[3] **Andreassen, O.** 1990. *A SAW Filter Bank for Telecommunications Satellites.* Microwave and RF Engineering, pp. 43-48, January/February.

[4] **McClellan, J. H., Parks, T. W., and Rabiner, L. R.** 1973. *A Computer Program for Designing Optimum FIR Linear Phase Digital Filters.* IEEE Trans. AU-21, pp. 506-526.

Advanced Communications Payload for Mobile Applications

S. A. Ames and R. K. Kwan
Ford Aerospace Corporation, Space systems Division
3825 Fabian Way, Palo Alto, CA 94303-4697

ABSTRACT

An advanced satellite payload is proposed for single hop linking of mobile terminals of all classes as well as VSAT's. It relies on intensive use of communications on-board processing and beam hopping for efficient link design to maximize capacity and large satellite antenna aperture and high satellite transmitter power to minimize the cost of the ground terminals. Intersatellite links are used to improve the link quality and for high capacity relay. Power budgets are presented for links between the satellite and mobile, VSAT, and hub terminals. Defeating the effects of shadowing and fading requires the use of differentially coherent demodulation, concatenated forward error correction coding, and interleaving, all on a single link basis.

INTRODUCTION

Each class of mobile terminal, land, maritime, and aeronautical, has its own set of system constraints. These stem from differences in traffic demand including data rates, number of users, and the use of scheduled or demand access. Of great importance are the differences in channel characteristics including the carrier frequency, fading, and interference conditions. Thus the design of an optimal link for each of the three mobile classes will be different thereby requiring flexibility and programmability in the satellite. It has been demonstrated that on-board communications signal processing (OBP) can achieve lower cost in mobile terminals and higher efficiency in spectrum utilization. However, the use of OBP has meant that the payload will have to be committed to a specific waveform prior to launch. Because of economic and traffic uncertainties, a commercial satellite owner/operator would most likely be reluctant to build a satellite with a committed OBP payload. To alleviate this limitation, a versatile OBP payload is needed which flexibly meets the communications requirements of all classes of users in the mobile community. The payload flexibility will enable not only connectivity between mobile and fixed ground terminals, but also between mobiles of the same class as well as between mobiles of different classes.

SATELLITE PAYLOAD DESIGN PHILOSOPHY

The satellite payload described here uses OBP, intersatellite links (ISL), and scanning spot beam techniques to achieve unprecedented flexibility in waveform processing and connectivity for mobile communications. The on-board modems, codecs, interleavers and deinterleavers are programmable by ground command which allows the processing described above to be achieved in "uncommitted" hardware which can be adapted to traffic demand. The OBP also allows for the implementation of hopping spot beams by virtue of digital buffering of data between hops. For a mobile satellite (MSAT) the hopping spots can be ultra-narrow beams to achieve super high antenna gain to service low capability terminals including very small personal terminals. The ISL permits intercontinental circuit connections on a single hop basis. It also enables a higher elevation angle line of sight from a terminal to the satellite thereby allowing the terminal to employ a narrower beam antenna for mitigation of multipath fading, shadowing and interference.

SYSTEM OPERATION AND REQUIREMENTS

A number of design features will facilitate reducing the cost and complexity of the earth terminals for which typical design parameters are given in Figure 1. These advantages can be illustrated by describing the procedures and links required to setup and implement voice calls between users. Calls can be made either between mobiles and telephone network users or directly between mobiles. It is assumed that all mobiles carry LORAN or GPS receivers so that they know their location and, therefore, can have their calls forwarded to them through the proper beam. The MSAT will broadcast a beacon signal modulated by the network symbol clock through an earth coverage beam. When transmitting to the satellite, it will be required that the mobile terminal synchronize its reference frequency to the beacon carrier, its symbol rate to the beacon clock, and adjust its baud phase, through the use of GPS

timing or a closed loop protocol, so that its baud arrival time at the satellite is synchronous with all other uplink baud arrival times. Having all users synchronize to align their baud transition times at the satellite is not too stringent a requirement at the low baud rates of coded voice and it greatly simplifies and enhances the reliability of the satellite multichannel demodulator (MCD). Requiring all users to synchronize their frequency synthesizer reference oscillators to the beacon carrier will ensure that the uplink FDMA signals can have a minimum of guard band between channels without mandating expensive ultrastable carrier oscillators in the mobiles. The differentially coherent 8PSK demodulation requires a very small ratio of Doppler shift to symbol rate, less than .05 assuming an L-band carrier. Estimates of the maximum value of this ratio for an automobile is about 0.1 and about 1.0 for a subsonic aircraft which shows that the modems on the mobile platforms will need to use active Doppler phase compensation to avoid excessive BER degradation.

The call setup procedure for random and demand access will by its very nature be outside of the reservation type of multiple access system; for example, an ALOHA type system operating at very low data rate through an earth coverage beam. The setup establishes a set of reserved uplink and downlink frequency channels, time slot reservations, and a pair of beams between the terminals and the satellite. In the case of the nonhopping beams which cover the metropolitan areas, only frequency and beam reservations are required, while the users in a hopping beam will also need a time slot assignment.

Since the satellite is a fully regenerating repeater, the uplink and downlink are effectively decoupled and the link calculations can be performed independently. Then the overall link BER is at worst twice the BER of either for balanced links and, for highly unbalanced links, approaches the BER of the weakest link. Public switched network users are connected through a central ground station called a hub. All satellite traffic to and from hubs will be transmitted on a single carrier in TDM through a Ku-band CONUS coverage beam.

Traffic between mobiles and the satellite will be transmitted through two types of L-band spot beams, fixed and hopping. Spot beams will be about one degree wide with fixed ones over urban areas having a large average volume of use. Mobile users communicating through fixed spot beams will use the minimum transmission data rate 4.8 Kbps and a very bandwidth efficient modulation, differentially coherent 8PSK (8DPSK). When coding of rate 1/2 is used and if it is filtered with a 40 % excess bandwidth raised cosine filter, this signal occupies 4.5 KHz and can be confined to channels on 5 KHz centers including ± 250 Hz guard bands which is a standard for mobile channels [1]. The block diagram for a mobile terminal is shown in Figure 2.

PAYLOAD DESCRIPTION

The advanced payload concept for an MSAT baseband processor shown in Figure 3 is representative of the functions in the payload and not the quantities of each type of block. The interfaces to and from the baseband processor are shown to be at 4 GHz, a convenient and typical satellite IF frequency, and the transponder bandwidth shown as a standard 72 MHz which could accommodate a maximum of 14,400 5 KHz channels. On reception, the FDMA and FDMA/TDMA 8PSK signals will first encounter a multichannel demultiplexer and demodulator (MCD). The MCD will likely be a SAW demultiplexer followed by a time-shared digital demodulator [2]. The demultiplexer uses the chirp Fourier transform principle to transmultiplex the FDMA signals to a TDMA signal which is then sampled, digitized and operated on by the digital demodulator which could employ standard digital signal processors or ASIC's. More recently even more compact MCD's have been described using integrated optics [3]. The 8DPSK demodulator is a Doppler compensating type and requires neither a carrier nor clock recovery phase-locked loop (PLL) [4]. However, for use with nonfading VSAT links, to achieve more power and/or bandwidth efficiency, the modem could be flexibly programmable by command to accommodate both coherent and differentially coherent demodulation as well as be adaptable to different modulation formats and symbol rates. Contracts for flexible satellite modems and complementary flexible codecs have been issued by NASA in the past several years and are currently under development [5], [6].

Following the demodulator is the FEC decoder and deinterleaver. To mitigate the channel fades, a concatenated pair of coders and deep interleavers at both the transmission ends of the links

and at the receiving ends are used. The outer codec typically uses a Reed-Solomon (RS) code while the inner code is convolutional using Viterbi Algorithm decoding. Interleaving may be used between the inner and outer encoders and decoders as well as between the inner codec and modem. The purpose of the interleaver is to randomize the channel errors to enable the codes to perform their error correction by distributing error bursts over the shuffled code words such that the number of errors in a single code word does not exceed its maximum correction capability. Because of the very deep and long fades, occasionally the error correction capability of the code will be exceeded and an error will be committed. At these times, and for data transmission only, the capability of the RS decoder to detect the presence of errors (beyond its capability to correct) will be exploited to request a repeated transmission from the transmitter (ARQ). Because of the presence of the decoder in the payload, ARQ's can be generated in the payload thereby avoiding the up and downlink delay which would result from placing the decoder only on the ground. Throughput is degraded because of the ARQ but for low data rate high reliability transmissions, such as paging, this is probably acceptable. Because of its greater error tolerance and more severe timeliness requirement, no ARQ can be used for digitized voice.

Using concatenated convolutional and RS codes on a Rayleigh fading channel with interleaving and MSK modulation, recent simulation results showed more than 40 dB coding and interleaving improvement [7]. Interleaving used in these results ranged from about 500 to more than 4000 bits resulting in interleaving delays of from about 100 ms to more than 800 ms for each interleaving and deinterleaving pair of operations For mobile to mobile the interleaving delay for an uplink and downlink would be twice these values. Clearly the upper limit is excessive for voice conversations thereby restricting the maximum interleaving depth to about 500 bits for voice but possibly much higher for services such as one-way paging and data.

The decoded data will undergo descrambling and will be entered into an input buffer memory followed by a routing switch and output buffer memory. It is in these blocks that the data is routed to its intended destination port, formatted, and repacked with other data for its downlink transmission. Data destined for FDMA downlinks to mobiles will be rescrambled, FEC encoded, modulated, and multiplexed for transmission through a common L-band power amplifier on a specific beam. The payload concept allows other modes of transmission to be accommodated such as wideband TDMA and wideband continuous data streams. The latter may be those arriving and leaving the payload via the ISL or to and from Ku-band VSAT's and hubs.

The baseband switch, buffer memory, beam hopping, as well as all traffic timing is under the control of the timing and control processor which is programmed through the order wire transmission from a master control station. The program controls access authorization, priorities, and logs traffic for billing. The master control station commands also will upload programs to control the payload configuration and schedules for the high capacity, non-demand access links such as hub links, video relay, and ISL.

TYPICAL LINK PERFORMANCE

The power budgets between a mobile and satellite for both voice and voiceband compatible data is given for no coding in Table 1. A modest 39 dBm e.i.r.p. is assumed for the ground transmitter while a large multibeam satellite antenna with a 47 foot diameter is assumed. Unfurlable parabolic antennas for L-band with this diameter are feasible. A circularly polarized ground antenna has been recently reported with autotrack capability and very reasonable mounting requirements atop a ground vehicle [8]. The FDMA downlink is for a metropolitan area spot beam with 1000 FDMA carriers. This requires a 200 watt transmitter radiating through the large 47 foot diameter satellite aperture.

The power budgets for the links to VSATs and hub terminals are given in Tables 2 and 3 respectively. These links operate in Ku-band using CONUS coverage beams. The links to the hubs use TDM transmission at 4.8 Mbps to be received by hub stations with 14.5 foot diameter antennas. The 1000 uplink signals at 4.8 Kbps on a single beam are not the same signals that are aggregated for downlink transmission to a single hub. The TDM transmissions to the hubs are aggregated according to the hub destination of the uplink signals arriving at the OBP payload on many different beams.

If possible, it would be desirable to tile the entire field of view of the satellite with contiguous one degree spot beams, but those beams not falling on metropolitan areas would only have a fraction of their receive and transmit capacity utilized. More efficient use of the satellite resource can be achieved by having, in addition to the fixed metropolitan area beams, a number of agile beams which are time shared between low density user locations. The algorithm of the time sharing might be strictly periodic, according to preprogrammed estimate of traffic demand, or it might even be adaptable to traffic demand. Assuming the case where a single beam is equally time shared among 100 spots, the uplink and downlink data rates would need to be increased by a factor of 100. The ground terminal peak e.i.r.p. would then need to be increased by 100 fold (although the average would still be 1 W) and the satellite downlink e.i.r.p. per carrier also increased by 100 over that of a stationary beam to maintain the same revisit time. The latter is only possible since the number of carriers per hopping spot is much fewer than the 1000 assumed for each metro beam. Thus, if the hopping spot only needs to service ten users, then the 200 W satellite transmitter would be sufficient to close the link even for 100 times speedup in the data rate. The main impact on the ground user of this speedup is that the synchronization precision to align the uplink bauds at the satellite must be 100 times better.

There are at least three classes of mobile platforms, land, sea, and air. Seagoing platforms which have been operating for a number of years in the INMARSAT system do not experience severe fades and some test results have recently been reported to indicate that the propagation environment for the aeronautical link suffers only minor fading [9]. The most difficult environment for satellite communications is land mobile which, unlike ships at sea or airborne terminals, are subjected to a hostile propagation environment caused by obstructions and reflecting objects in the terrain which produce partial or complete blocking of the line of sight path (flat fades and shadowing), mutipath fading (frequency selective fades), and combinations of these which will dictate the following important characteristics of the land mobile signal:

◦ Because of periodic deep fades, differentially coherent detection is preferred to coherent detection to avoid the frequent cycle slips in the coherent demodulator carrier PLL. (The ground terminal PLL for the pilot tone will have a very narrow tracking bandwidth and will not break lock unless the fade persists for an inordinately long time; besides its only function is to track out the long term frequency drifts between the satellite and reference oscillators.) The cost of 8DPSK versus coherent 8PSK detection is about 3 dB in power efficiency although it is possible to use multiple symbol differential detection to achieve less degradation [10].

◦ In the fading channel, a great deal of additional link margin will be required. This margin will be achieved through a combination of forward error correction (FEC) coding and interleaving of symbols to mitigate the effects of error bursts.

The links given in the tables are for additive white Gaussian noise channels which are applicable to the ocean and airborne environments of mobile platforms and the satellite to VSAT and hub links. They are not applicable to the land mobile links except under unusually good conditions. Measurements on the L-band channel have shown the probability density function (PDF) of the received power to exhibit Rician fading in the non-shadowed environment and, when shadowing is present, to have a conditional Rayleigh amplitude PDF, (i.e., a conditional exponential PDF of power), conditioned on the short term mean received power which fits the lognormal PDF [11]. To properly evaluate the overall link fading requires the summing of the PDF's from the shadowing and nonshadowing cases where each is weighted by the frequency of its occurrence. Evaluation of coding and interleaving with this complex fading model will need to be achieved through simulation for candidate modulation formats such as 8DPSK.

CONCLUSIONS

◦ Air, sea, land mobile, and VSAT communications will use the same satellites with flexible OBP functions to reduce the cost of the overall communications system and the ground terminals in particular by trading off for more complex satellite payloads

◦ Hopping beams between areas of low demand will optimize cost of service

◦ To close the land mobile links in a fading and shadowing environment, differentially coherent

demodulation and concatenated coding and interleaving will be employed

• MSAT systems will use a multi-satellite constellation connected by ISL's to maintain the highest feasible elevation angles to the satellites in order to minimize fading and interference

REFERENCES

1. "The MSAT-X Project: NASA Explores Advanced Communications Technologies", NASA Tech Briefs, Vol. 14, No. 2, pp. 12-14, Feb. 1990
2. Saggese, Enrico and G. Chiassarini, "SAW and Digital Technologies in Multicarrier Demodulators", Int'l. J. of Sat. Com., VOL. 7, 295-306, Oct.-Nov. 1989
3. Ananasso, Fulvio, and I. Bennion, "Optical Technologies for Signal Processing in Satellite Repeaters", IEEE Communications Magazine, pp. 55-64, VOL. 28, No.2, Feb. 1990
4. "Fast Correction for Doppler in MDPSK Signals", NASA Tech Briefs, Vol. 13, No. 5, Item # 136, May 1989
5. RFP3-319317Q, "Programmable Digital Modem", NASA Lewis Research Center, Sept. 21, 1988
6. RFP3-310480, "Flexible High Speed Codec", NASA Lewis Research Center, Jan 15, 1988
7. Wong, K. H. H., et al., "Channel Coding for a Satellite Mobile Channel", Int'l. J. of Sat. Com. , VOL. 7, 143-163, July-Sept. 1989
8. "Mechanically-Steered, Mobile Satellite-Tracking Antenna", NASA Tech Briefs, Vol. 14, No. 2, Item # 42, Feb. 1990
9. Hase, Yoshihiro, et al., "ETS-V/EMSS Experiments on Aeronautical Communications", paper 7.1, ICC Conference Record, June 11–14, 1989
10. Divsalar, Dariush and M. K. Simon, "Multiple-Symbol Differential Detection of MPSK", IEEE Trans. on Com., VOL. 38, No. 3, pp. 300-308, March 1990
11. Lutz, Erich, et al., "Land Mobile Satellite Channel - Model and Error Control", Satellite Integrated Communications Networks, E. del Re, et al. (eds.), Elsevier Science Pub. B. V., 1988

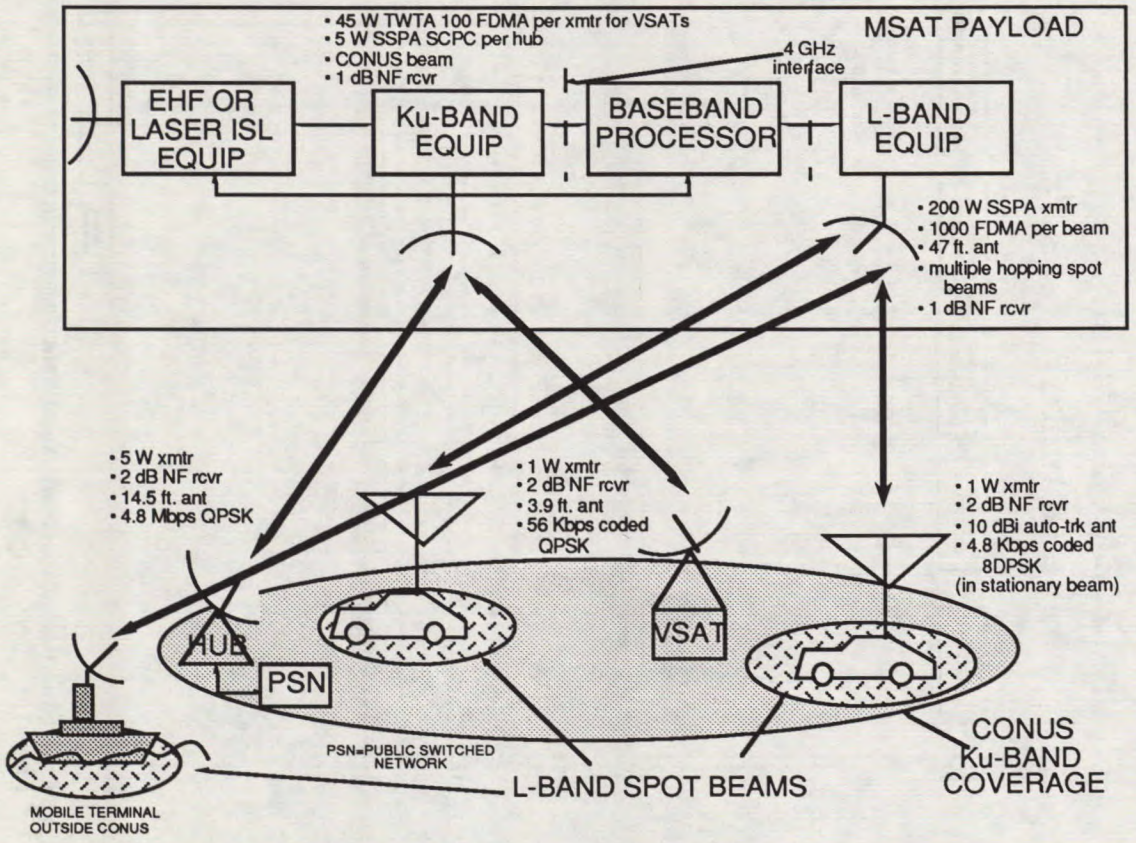


Figure 1 MSAT system showing salient terminal characteristics

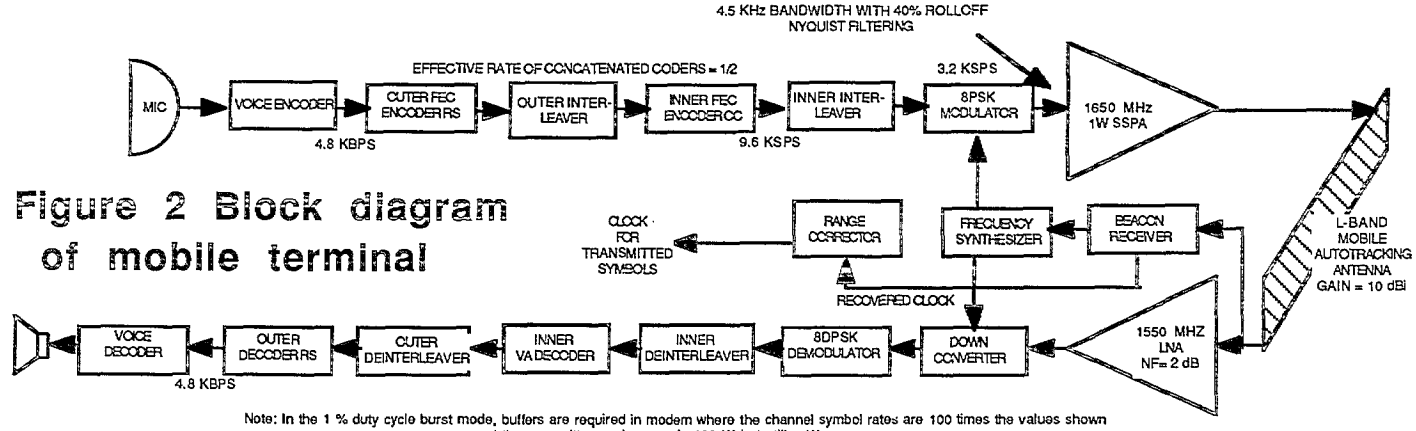


Figure 2 Block diagram of mobile terminal

Note: In the 1 % duty cycle burst mode, buffers are required in modem where the channel symbol rates are 100 times the values shown and the transmitter peak power is 100 W but still 1 W average.

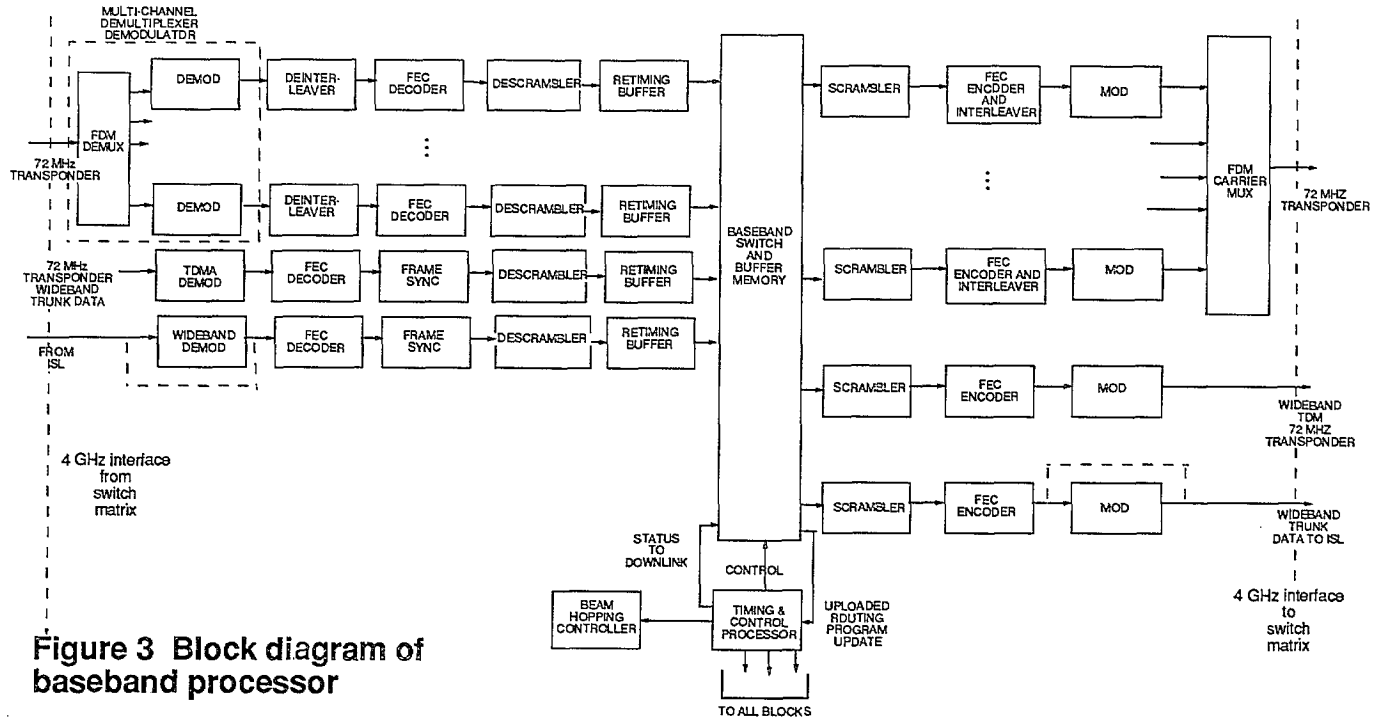


Figure 3 Block diagram of baseband processor

MSAT Mobile Links									
				COMMENTS	DOWNLINK 1550 MHz			COMMENTS	
Ground Transmitter Power	30.0	dBm	1	W SSPA.	Satellite Transmitter Power	53.0	dBm	200	W SSPA
Ground Transmitter Losses	1.0	dB			Satellite Transmitter Backoff	6.0	dB	FDM	
Ground Transmitter Antenna Gain	10.0	dBi		JPL auto-tracking	Satellite Power per Carrier	17.0	dBm	1000	carriers
Ground e.i.r.p.	39.0	dBm			Satellite Transmitter Ant. Gain	44.5	dBi	47' diam, 55% effic., 1degree b.w.	
Free Space Loss	189.0	dB			Satellite Transmitter Losses	1.0	dB		
Satellite Antenna Gain	45.0	dB		47' diam., 55% illumination efficiency	Satellite e.i.r.p	60.5	dBm	per carrier	
Sat. Rec. System Noise Temp.	25.6	db-deg K		290 deg earth, 1dB noise figure	Free Space Loss	188.5	dB		
Satellite G/T	19.4	db/deg K			Ground Terminal Ant. Gain	10.0	dBi	JPL auto-tracking	
Eb/No	29.7	dB		4.8 Kbps data, modem impl. loss 1.5 dB	Ground Receiver Noise Temp.	22.3	dB-deg K	2 dB noise figure	
Required Eb/No	14.0	dB		BER=10e-3 for 8DPSK	Ground Receiver G/T	-12.3	dB/deg K		
Margin for BER=10e-3	15.7	dB		AWGN without coding	Eb/No	20.0	dB	4.8 Kbps, 1.5 dB modem impl. loss	
Required Eb/No	17.0	dB		BER=10e-6 for 8DPSK	Required Eb/No	14.0	dB	D8PSK, BER=10e-3	
Margin for BER=10e-6	12.7	dB		AWGN without coding	Margin	6.0	dB	AWGN without coding	

Table 1 MSAT Mobile Links

MSAT VSAT Links									
UPLINK 14.2 GHz				COMMENTS	DOWNLINK 11.9 GHz			COMMENTS	
Ground Transmitter Power	30.0	dBm	1	W SSPA	Satellite Transmitter Power	46.5	dBm	45	W TWTA
Ground Transmitter Losses	1.0	dB			Satellite Transmitter Backoff	6.0	dB	FDM	
Ground Transmitter Antenna Gain	42.5	dBi		3.9' ant.	Satellite Power per Carrier	20.5	dBm	100	carriers
Ground e.i.r.p.	71.5	dBm			Satellite Transmitter Ant. Gain	29.0	dBi	1' diam., CONUS coverage	
Free Space Loss	207.5	dB			Satellite Transmitter Losses	1.0	dB		
Satellite Antenna Gain	30.5	dB		1' diam., CONUS coverage	Satellite e.i.r.p	48.5	dBm	per carrier	
Sat. Rec. System Noise Temp.	25.6	db-deg K		290 deg earth, 1dB noise figure	Free Space Loss	206.0	dB		
Satellite G/T	4.9	db/deg K			Ground Terminal Ant. Gain	40.5	dBi	3.9' ant.	
Eb/No	18.5	dB		56 Kbps data, modem impl. loss 1.5 dB	Ground Receiver Noise Temp.	22.3	dB-deg K	2 dB noise figure	
Required Eb/No	10.6	dB		BER=10e-6 for QPSK, uncoded	Ground Receiver G/T	18.2	dB/deg K		
Margin for BER=10e-6	7.9	dB		AWGN	Eb/No	10.3	dB	56 Kbps data, modem impl. loss 1.5 dB	
Required Eb/No	11.3	dB		BER=10e-7 for QPSK, uncoded	Required Eb/No	5.5	dB	BER=10e-6 for QPSK, R=1/2 K=7, VA decoder	
Margin for BER=10e-7	7.2	dB		AWGN	Margin	4.8	dB		

Table 2 MSAT VSAT Links

MSAT Hub Links									
UPLINK 14.2 GHz				COMMENTS	DOWNLINK 11.9 GHz			COMMENT	
Ground Transmitter Power	37.0	dBm	5	W SSPA	Satellite Transmitter Power	37.0	dBm	5	W SSPA
Ground Transmitter Losses	1.0	dB			Satellite Transmitter Backoff	0.0	dB	TDM	
Ground Transmitter Antenna Gain	52.4	dBi		14.5' ant.	Satellite Power per Carrier	37.0	dBm	1	carrier
Ground e.i.r.p.	88.4	dBm			Satellite Transmitter Ant. Gain	29.0	dBi	1' diam., CONUS coverage	
Free Space Loss	207.5	dB			Satellite Transmitter Losses	1.0	dB		
Satellite Antenna Gain	30.5	dB		1' diam., CONUS coverage	Satellite e.i.r.p	65.0	dBm		
Sat. Rec. System Noise Temp.	25.6	db-deg K		290 deg earth, 1dB noise figure	Free Space Loss	206.0	dB		
Satellite G/T	4.9	db/deg K			Ground Terminal Ant. Gain	52.0	dBi	14.5 ' diam. hub terminal	
Eb/No	16.1	dB		4.8 Mbps data, modem impl. loss 1.5 dB	Ground Receiver Noise Temp.	22.3	dB-deg K	2 dB noise fig.	
Required Eb/No	10.6	dB		BER=10e-6 for QPSK, uncoded	Ground Receiver G/T	29.7	dB/deg K		
Margin for BER=10e-6	5.5	dB		AWGN	Eb/No	19.0	dB	4.8 Mbps, 1.5 dB implem. loss	
Required Eb/No	11.3	dB		BER=10e-7 for QPSK, uncoded	Required Eb/No	10.6	dB	QPSK, BER=10e-6	
Margin for BER=10e-7	4.8	dB		AWGN	Margin	8.4	dB	AWGN, uncoded	

Table 3 MSAT Hub Links

Payloads Development for European Land Mobile Satellites: A Technical and Economical Assessment

G. PERROTTA, F. RISPOLI, T. SASSOROSSI
SELENIA SPAZIO - Via Saccomuro, 24
00131 - ROME (ITALY)
Phone: (06)-43681
Fax: (06)-4090675
(06)-4090773

ABSTRACT

In the frame of the European Space Agency activities, Selenia Spazio has defined two payloads for Mobile Communication, one for pre-operational use (European Land Mobile System, EMS) and one for promoting the development of technologies for future mobile communication systems (L-band Land Mobile Payload, LLM).

The paper provides a summary description of the two payloads, their main performances and, in conclusion, an economical assessment for the potential mobile communication market in Europe.

1. INTRODUCTION

The land mobile satellite communication market is expected to grow considerably in Europe during the next decade. Several initiatives are taking place trying to meet the user needs offering a number of services which range from the simplest radiolocalization and messaging to more sophisticated data and voice channels.

The European Space Agency (ESA) has been widely active in promoting the technology for Land Mobile Communica tions since many years and its PRODAT low data rate terminal is currently conducting field trials using the MARECS satellite.

The next step is the deployment of an European Land Mobile System (EMS) oriented for closed users groups (i.e. trucks) and allowing data and voice communications with an initial capacity of 300 to 500 channels. The technical and economical viability of the EMS project has been amply assessed by several studies performed by different organizations. The payload can be embarked as an additional passenger of the ITALSAT F2 or EUTELSAT II communication satellites by the end of 1992.

These opportunities would allow to start the mobile service by early 1993 in response to the urgent users needs and exploit a lucrative market segment which predicts up to one million of users which can be economically served via satellite by 1995 throughout the european territory.

According to this scenario there exists an amply justification for the deployment of an operational system for a fully developed service with a 2000 channels dedicated satellite.

Therefore ESA has planned to fly an experimental payload on board the Technological Satellite Mission (ARTEMIS) by 1993 for which Selenia Spazio has been selected as Prime Contractor.

This is to develop and qualify all critical elements which will be required for a future high capacity operational mission. The main characteristics of the L-band Land Mobile Payload (LLM), which is presently in the phase B, i.e. design definition, are the use of multiple beams and of an efficient technique for the utilization of spacecraft resources through flexible adaptation to traffic demand.

This paper outlines the importance of the Land Mobile Services in Europe where a number of different systems and operators are making plans for the near future. The two payloads which have been recently defined (EMS, LLM for ARTEMIS) will be described underlying

the major technical issues. In particular the design philosophy and the necessary technologies is discussed taking into account recent developments in the areas of antennas, beam forming networks and power amplifiers. Also an economical assessment in terms of potential revenues for the mobile services is given under different hypothesis of traffic capacity.

2. THE EMS MISSION AND ITS PAYLOAD

2.1 Mission and Payload Requirements

The European Mobile System (E.M.S.) will provide two ways voice and data communications between Fixed Earth Stations (FES) and land mobiles located anywhere in Europe and, possibly, North Africa. Links between satellite and mobiles will be at L-band. Feederlink to/from FES will be at Ku-band.

Table 2.1 excerpts the main EMS payload characteristics.

The definition of a first generation L-band payload addressed properly the technological drivers. Many ESA contracts are infact underway, across the european industry, spurring advanced technologies development and test.

In line with this approach, Selenia Spazio performed several studies for ESA [1], [2], [3] aiming to define a payload architecture making optimum use of the on going development in Europe.

2.2 The Payload Architecture

Fig. 2.1 shows the payload block diagram: it consists of two antennas and a multichannel transparent forward and return repeater.

A forward channellized repeater links the FES to the land mobiles. On the other side, a return channellized repeater links, the mobile terminals to the FES.

The 2.5 x 1.5 m^2 reflector antenna provides an elliptical beam at L-band covering Europe and North Africa (Fig. 2.2). The elliptical reflector is fed by a cluster of four self-diplexed patches radiating in left hand circular polarization at the transmit and receive bands. The feederlink coverage is provided by an indipendent Ku-band antenna of the Dragone type [4], chosen

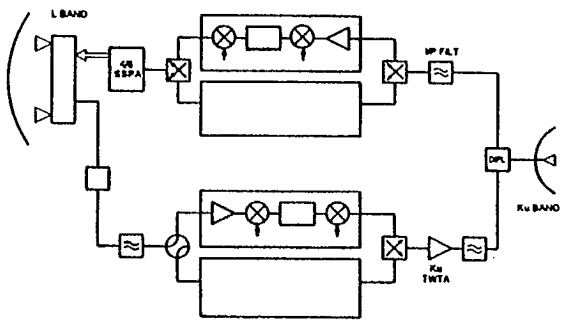


Fig. 2.1 EMS Payload Block Diagram

on the grounds of its inherently good crosspolar performance.

In the receive side, a passive combiner sums the four signals received by the four antenna patches before LNA amplification. In the transmit side, four power amplifier modules are directly connected to the four patch elements. A single high power beam is synthesized by the four RF signals.

The maximum power level, through the transmit chain, does not exceed that of each power module. This considerably reduces multipaction and passive intermodulation problems.

This is an important factor since, in the EMS mission, the frequency plan is such that the 7th order intermodulation products fall in the receive band. This and the very high EIRP require an unprecedented overall strategy for PIMP avoidance and control.

PIMP sources are all passive components interested to the flow of high power density RF multicarrier signal such as:

- transmit filters and/or diplexers
- passive RF circuitry
- antenna feeds
- reflectors
- RF cables

A distributed power amplifier concept will help in reducing the PIMP generation. Besides, the inherent transmit to receive isolation of self-diplexed patch elements (yet unproven in terms of PIMP generation and

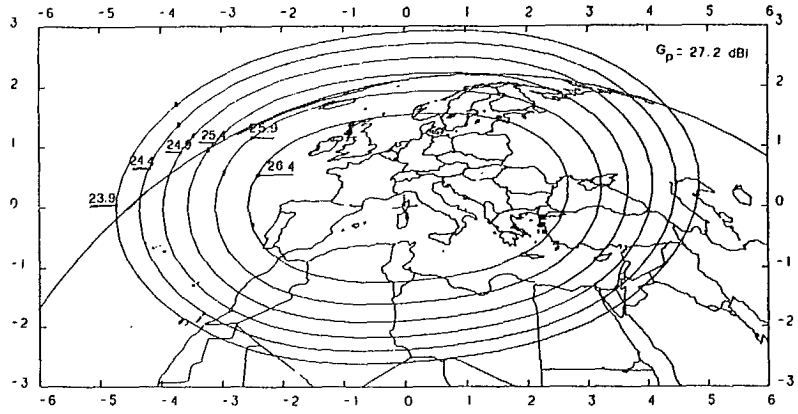


Fig. 2.2 EMS Antenna Isolevel Gain Curves

multipacting behaviour) is expected to be a determining factor for the feasibility of such advanced high power mobile systems.

Solid state L-band power amplifiers have been assessed, compared to TWTAs in terms of mass, efficiency, reliability, and finally chosen for the EMS transponder. At present, two L-band SSPA concepts are under developments in Europe: the DEBS, by MSS, and the PAMELA, by MBB. Both types have efficiencies of about 30%, at a C/I of about 18 dB under multicarrier operation (10 tone test), and low mass.

The forward and return channellized sections serve to multiplex and demultiplex the non-adjacent frequency slots available for the land mobile services. The forward repeater frequency plan shown in Fig. 2.3, consists of three non contiguous bands about 4 MHz wide.

The use of SAW filters will provide the necessary selectivity. A 120 MHz intermediate frequency is used.

Frequency coordination, with existing Ku-band Satellite Systems, requires, however, a more flexible return repeater multiplexing scheme. We slice each of the three non-adjacent frequency slots of Fig. 2.3 into three virtual channels (Fig. 2.4). The assignement of the return link channels may cope with any future spectrum utili zation restraint. Nevertheless, fixed bandwidth filter multiplexers have a poor bandwidth utilization efficiency, due to the guardbands required in proximity of crossovers.

The AME company studied, under ESA contracts, a potentially attractive solution to this problem: the Bandwidth Switchable Saw Filter, or BSSF [5]. The device consists of a bank of phase matched SAW filters and a switchable combining network, with controlled phases. It is, thus, possible to combine any filter subset achieving a variable bandwidth filter. The outstanding advantage of the concept consists in the elimination of the guardbands at filters crossover, when using contiguous filters.

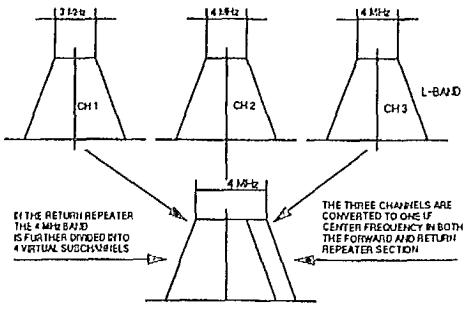


Fig. 2.3 EMS Forward Repeater Frequency Plan

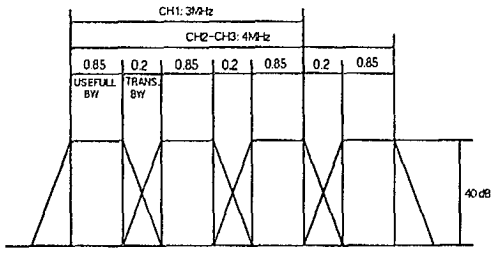


Fig. 2.4 EMS Return Repeater Frequency Plan

Table 2.1 EMS Payload Characteristics

- Coverage	Global European
- L Band EIRP	44.5 dBW
- L Band G/T	- 1.2 dB/K°
- Forward Useful Bandwidth	11 MHz total in three channels 4, 4. and 3 MHz wide
- Return Useful Bandwidth	9.35 MHz, in 11 channels 0.85 MHz wide each. May implement a BSSF combiner
- Ku-Band EIRP	34.9 dBW
- Ku-Band G/T	-1.4 dB/K°
- Payload mass including Ku-Band Feederlink	60 Kg.
- Payload DC Power including Ku-Band Feederlink	400 W

3. THE L-BAND LAND MOBILE PAYLOAD

3.1 Mission Requirements

The L-band Land Mobile Payload must provide the same kind of service of EMS Payload, but with different coverage and EIRP requirements.

A spot beam coverage over Europe will be provided at high EIRP (+ 51 dBW) and G/T (+ 2.5 dB/K°) levels.

3.1.1 LLM Payload Main Characteristics

The main performances of the LLM Payload are summarized in Table 3.1.

Europe will be covered both by a global coverage and by 6 spot beams as depicted in Fig. 3.1 and 3.2.

A steerable beam is available that can be moved over the positions shown in Fig. 3.3

The same frequency can be reused between spot beam no. 1 and 6 and between spot 3 and 5.

The LLM channel G (4 MHz wide) is permanently connected to the Euroglobal beam, whereas each 1 MHz slot in the L1 and L2 frequency bands can be connected to each spot beam of the L-band coverage.

The LLM Payload is able to perform as EMS back-up system.

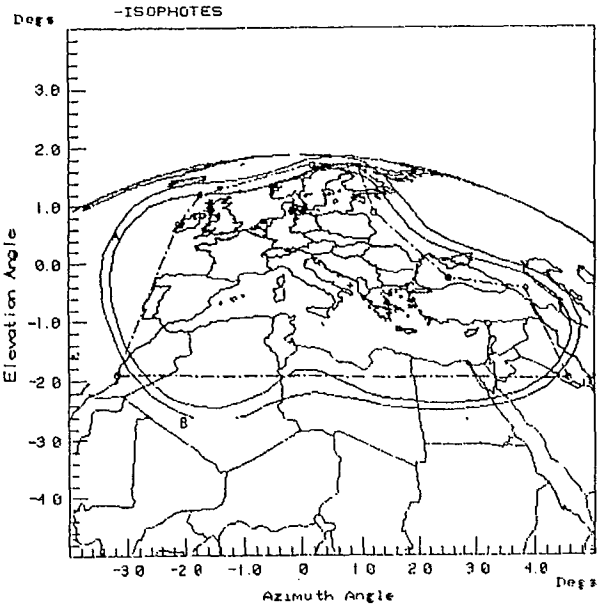


Fig. 3.1 LLM Payload Gain Contours (Euroglobal Coverage)

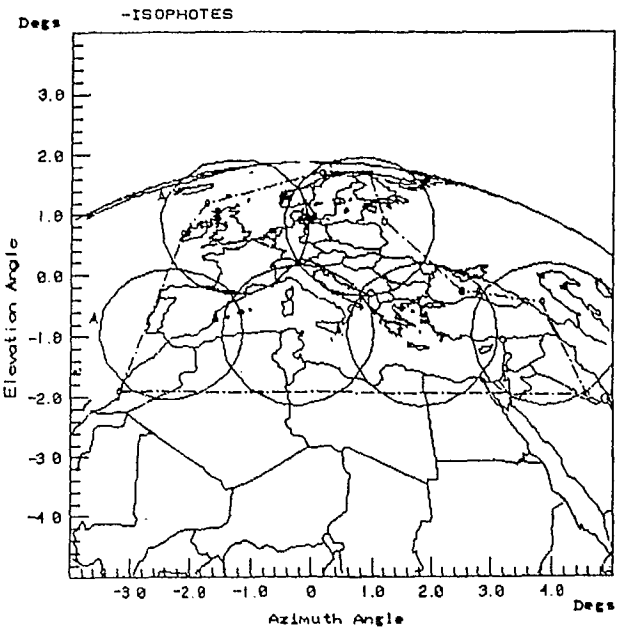


Fig. 3.2 LLM Payload Gain Contours (Beam Coverage)

3.2 LLM Payload Block Diagram Description

The LLM Payload L-band Section is based on a large reflector, offset antenna with the following characteristics:

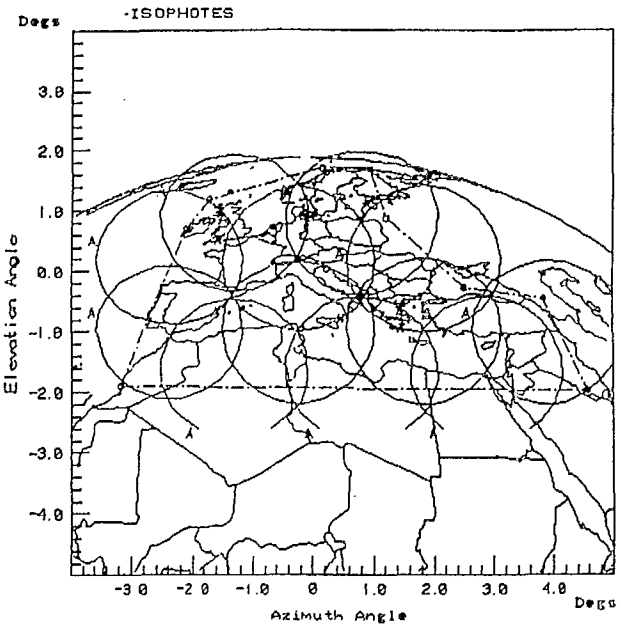


Fig. 3.3 LLM Payload Gain Contours (Steerable Beam)

- Projected Aperture 5 m
- F/D ratio 0.625
- Focal Length 3.125 m
- Clearance 0.71 m

The reflector is implemented by an inflatable structure manufactured by Contraves.

The L-band Feed Array is composed by 12 radiating elements, that are LH circular polarized except three of them that can provide both senses of circular polarization.

The Euroglobal Beam is generated by combining the signals from all of the twelve radiating elements, whereas the spot beams are synthesized by the signals from selected triplets of feed elements.

The correct phase distributions for the beams are generated in the Forward and Return Beam Forming Matrix.

The L-band High Power Section is formed by 12 dual-redundant SSPA's, of 10 W of output power each.
The Power Amplifiers are arranged as 3 Multi port Amplifiers of order 4 x 4. This particular configuration has been defined by ESTEC and is patented as "Multimatrix Antenna".

One of its most important feature is that it allows to share the available RF power among the various beams in whichever ratio according to the traffic demand.

In the Return Link, each line from the Feed Array to the BFM is equipped with a dual-redundant LNA.

There are two IF Processors, one in the Forward and one in the Return Link.

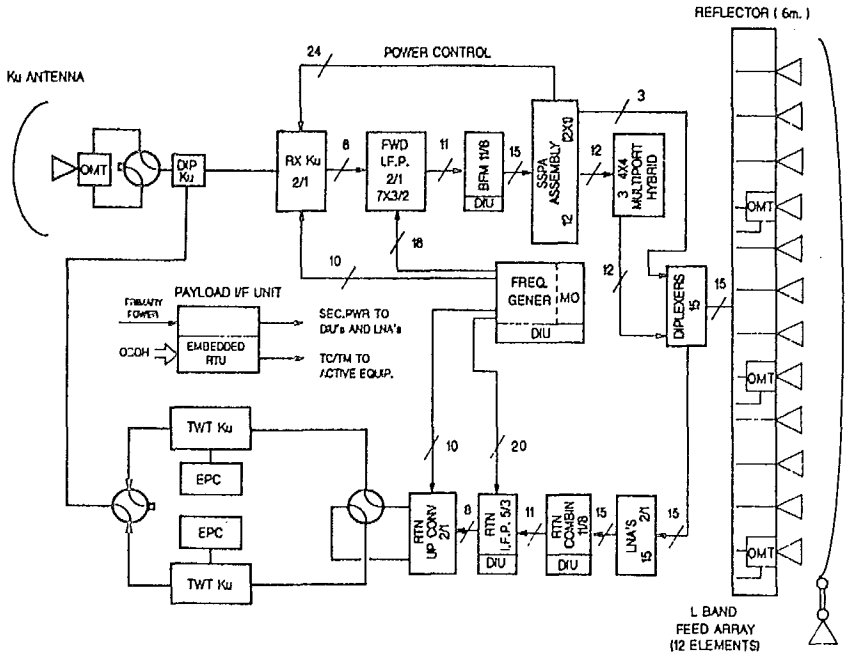


Fig. 3.4 LLM Payload Block Diagram

They perform channel filtering by using SAW filters and provide channel-to-beam switching function.

Each of them has a section dedicated to the EMS mode.

The frequency band of the Feeder Link of this Payload is at Ku-band, but the frequency allocation is not yet been defined. The Ku-band antenna is a 42 cm. onset parabola that can provide signals both in vertical and horizontal polarization.

The actual sense of polarization can be selected by ground command to cope with coordination requirement with other systems.

The received signals are LN amplified and down-converted by the Ku-band Receiver and then provided to the Forward IF Processor.

On the other side, the signals from the Return IF Processor are up-converted in the Return Upconverter.

The RF power at Ku-band is provided by a TWTA of 20W linear power that works at 6 dB back-off to cope with the linearity requirement.

4. ECONOMICAL ASSESSMENT

The growth projections for mobile communications in Europe predict a potential market of 12 million users by 1995. At least an 8% of these can be served economically via satellite. To satisfy this demand the EMS payload should be available in orbit by 1993 providing some 400 channels.

The ITALSAT F2, scheduled for launch in the same time frame, is a convenient opportunity to embark the EMS payload because of the minimum impact on ITS primary telecommunication mission [1].

In this respect the estimated cost of the space segment for the EMS mission is in the order of 10 to 13 million dollars per year while the potential revenue could be up to three times higher. The operator can take advantage of this opportunity considering also the possibility to have in orbit the LLM payload whose equivalent capacity could be leased for operational purposes. Financial risks should be agreed on the basis of the effective traffic carried out while preserving the technological content of the EMS and LLM missions which represent the forerunner of a dedicated commercial system.

CONCLUSIONS

Selenia Spazio has defined two payloads for Mobile Communications. The EMS Payload is a relatively "simple" payload in order to provide a pre-operational mobile communication service over Europe with limited risks within the required time frame (1993). The LLM Payload, through the experimentation of beam coverage at very high EIRP levels, intends to provide a field for development and testing of advanced technologies such as very large antennas, Beam Forming Matrices and Multiport Amplifiers that will be useful for future mobile communication payloads.

Both the systems have been demonstrated to be able to provide an adequate financial revenue.

Table 3.1 LLM Payload Performance Summary

EIRP	L-band Euroglobal	+45.9 dBW
	L-band Spot Beam	+51.4 dBW
	Ku-band Euroglobal	+38.3 dBW
G/T	L-band Euroglobal	+ 4.9 dB/K°
	L-band Spot Beam	- 1.2 dB/K°
	Ku-band Euroglobal	- 1.5 dB/K°
Coverage	See Fig. 3.1 and 3.2	
Linearity	L-band C/I > 16.1 dB	
(10 tone test)	Ku-band C/I > 21.1 dB	
P/L mass	140 kg.	
P/L Power Cons.	611 W	

REFERENCES

1. Selenia Spazio Final Report: "Study of Technological Boundaries for the EMS Payload accomodation on Italsat F2" ESA Contract 8070/88 April 1989

2. G. Perrotta, F. Rispoli: "Characteristics of first generation L-band Payloads for land mobile applications in Europe" 19th European Microwave Conference London Sept. 1988

3. ESA contract 8332/89: "Investigation into a Combined Italsat/PSDE SAT2 Mission" ongoing at Selenia Spazio.

4. A. Saitto, D. Patel, G. Doro, B. Trombetta: "High Efficiency Satellite Antenna" Electronic Letters; April 1983, Vol. 19

5. Andreassen, Viddal: "Bandwidth Switchable SAW Filters (BSSF) on Quartz" Ultrasonic Symposium Proceedings, 1988

6. Selenia Spazio Document: "L-band Land Mobile Payload Detailed Definition and Interface Description", RPT/SAT/0181/SES, March 1990

7. S. Egami and M. Kawai, "An Adaptive Multiple Beam System Concept" IEEE Journal vol. SAC-5, no. 4, May 1987

A Generalized Transmultiplexer and its Application to Mobile Satellite Communications

Osamu Ichiyoshi
NEC Corporation
4035 IKEBE-CHO, MIDORI-KU
YOKOHAMA, 229 Japan
Tel: 045-939-2215
Fax: 045-933-9768

ABSTRACT

This paper presents a generalization of the digital transmultiplexer technology. The proposed method can realize a transmultiplexer (TMUX) and a trans-Demultiplexer (TDUX) filter banks whose element filters can have bandwidths greater than the channel spacing frequency. This feature is useful for many applications in communications fields. As an example, a satellite switched (SS) FDMA system is proposed for spot beams satellite communications in particular for mobile satellite communications.

1. INTRODUCTION

The transmultiplexer based on FFT and Digital polyphase concept was proposed in 1974[1]. Since then extensive applications have been made to FDM/TDM transmultiplexers for General Switched Telephone Networks[2,3]. Another interesting area of its application is to the onboard processing circuit for satellite switched (SS) FDMA systems[4], which will be quite effective for mobile satellite communications. An exhaustive summary and prospect of the technology is given in the literature[5].

The conventional transmultiplexer method has an inherent limitation that the bandwidth of the element filter of the filter bank is bounded by the channel spacing frequency. This limitation is too restrictive for general SS/FDMA systems which must be able to provide transparent transmission paths for any portion of the assigned frequency ranges.

This paper presents a generalization of the transmultiplexer which can provide such filter banks.

2. CONVENTIONAL TRANSMULTIPLEXER

A block diagram of the trans multiplexer is given in Figure 1. In the figure (A) shows a multiplexer (TMUX) and (B) a demultiplexer (TDUX). The TMUX output Y (Z) is given by

$$Y(Z) = \sum_{i=0}^{N-1} Z^{-i} G_i(Z^N) \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}ki} X_k(Z^N) \qquad (2\text{-}1)$$

where

$$Z = e^{j2\pi fT} = e^{j2\pi f/fs} \qquad (2\text{-}2)$$

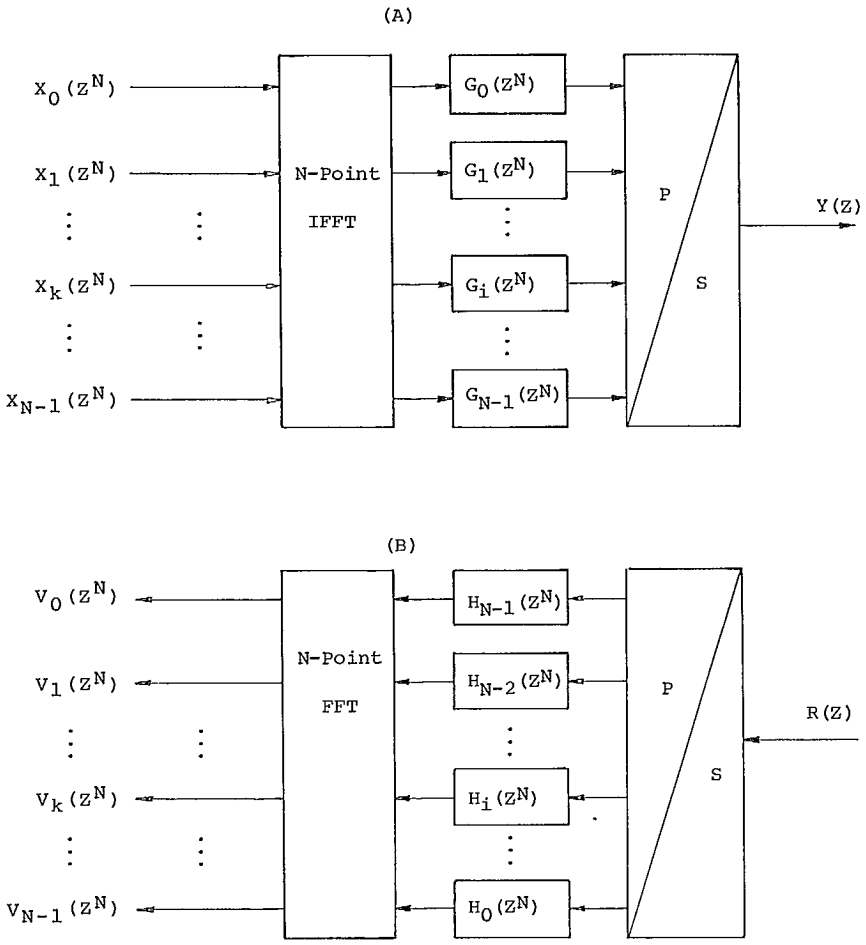
$$fs = 1/T \ ; \text{ Sampling frequency} \qquad (2\text{-}3)$$

$$X_k(Z^N) = \sum_n x_k(n) Z^{-nN} \qquad (2\text{-}4)$$

Those $G_i(Z^N)$ $(i=0,1,2, \ldots, N-1)$ in eq(1) are the subfilters which constitute the basic LPF G(Z) as

$$G(Z) = \sum_{\ell=0}^{L-1} g(\ell) Z^{-\ell} = \sum_{i=0}^{N-1} Z^{-i} G_i(Z^N) \qquad (2\text{-}5)$$

$$G_i(Z^N) = \sum_{\ell=0}^{L/N-1} g(\ell N+i) Z^{-\ell N} \qquad (2\text{-}6)$$



(A) Multiplexer (MUX)
(B) Demultiplexer (DUX)

Figure 1 Conventional Digital FDM Circuit

The output of the k-th channel of TDUX is given by

$$V_k(Z^N) = \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) R_i(Z^N) \qquad (2\text{-}7)$$

where the receive signal R(Z) is decomposed into subsequences $R_i(Z^N)$ as follows

$$R(Z) = \sum_n r(n) Z^{-n} = \sum_{i=0}^{N-1} Z^{-i} R_i(Z^N) \qquad (2\text{-}8)$$

$$R_i(Z^N) = \sum_n r(n \cdot N + i) Z^{-Nn} \qquad (2\text{-}9)$$

It is apparent from Figure 1 that the TMUX and TDUX are mutualy inverse operations.
It should be noted that because sampling frequencies for the signals $X_k(Z^N)$ or $V_k(Z^N)$ are

$$\Delta f = \frac{1}{NT} = \frac{fs}{N} \qquad (2\text{-}10)$$

the bandwidth of each multiplex or demultiplex filter can not exceed Δf.

3. GENERALIZED TRANSMULTIPLEXER

The structure of the generalized TMUX and TDUX is given in Figure 2.

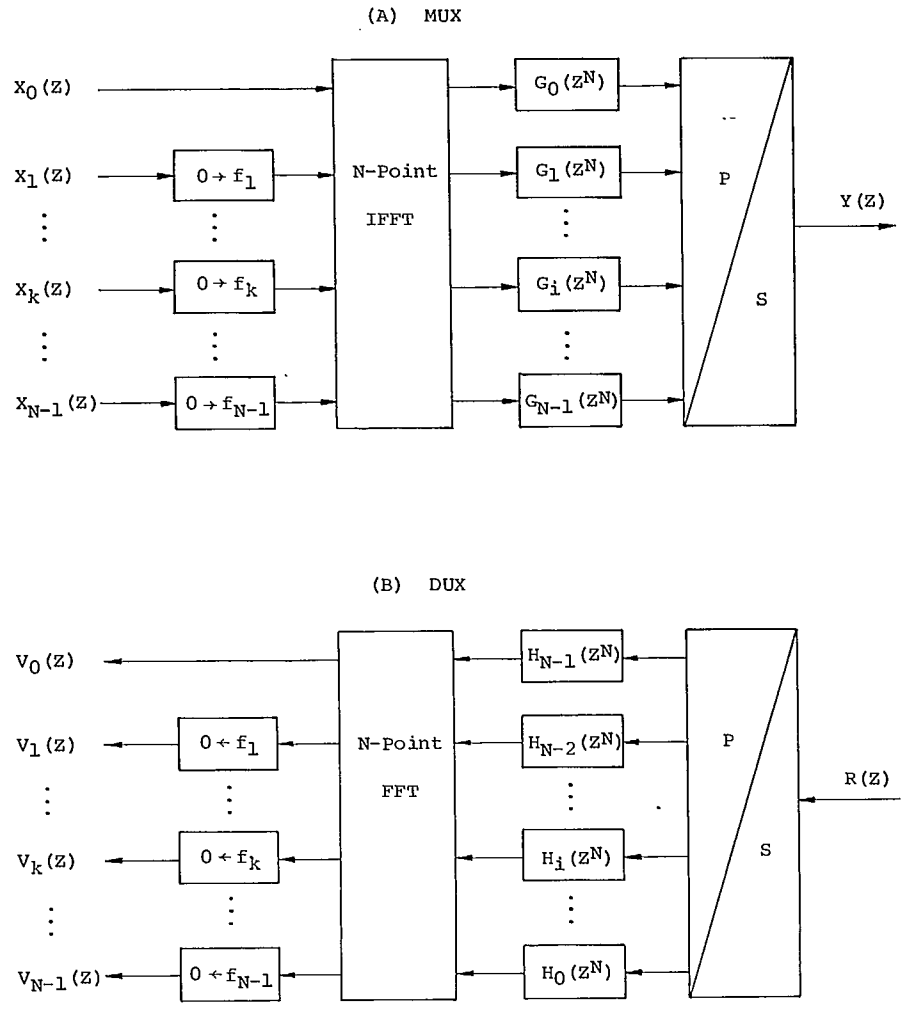


Figure 2 Generalized Digital FDM Circuit

3.1 TMUX

In general the k-th input $X_k(Z)$ is expressed as

$$X_k(Z) = \sum_n x_k(n) Z^{-n} \qquad (3\text{-}1)$$

This signal is to be frequency converted to f_k;

$$f_k = k \cdot \Delta f = k \cdot \frac{fs}{N} \qquad (3\text{-}2)$$

The conversion is made by the variable transform;

$$Z \rightarrow e^{-j2\pi f_k T} Z = e^{-jk\frac{2\pi}{N}} Z \qquad (3\text{-}3)$$

Hence;

$$X_k(Z;k) = \sum_n e^{j\frac{2\pi}{N}kn} x(n) Z^{-n} \qquad (3\text{-}4)$$

In the same manner the original LPF G(Z) is also frequency converted to f_k in order to select $X_k(Z;k)$;

$$G(Z;k) = \sum_{i=0}^{N-1} e^{j\frac{2\pi}{N}ki} Z^{-i} G_i(Z^N) \qquad (3\text{-}5)$$

Thus the output of the TMUX will be

$$\begin{aligned} Y(Z) &= \sum_{k=0}^{N-1} G(Z;k) \cdot X_k(Z;k) \\ &= \sum_{i=0}^{N-1} Z^{-i} G_i(Z^N) \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}ki} \cdot X_k(Z;k) \end{aligned} \qquad (3\text{-}6)$$

Eq (3-6) is depicted in (A) of Figure 2. The figure and eq (3-6) tell that the generalized TMUX is basically of the same structure as the conventional TMUX including it as a special case as shown below. From eq (3-4)

$$X_k(Z;k) = \sum_n e^{jk\frac{2\pi}{N}n} x_k(n) Z^{-n}$$

If the input signal $X_k(Z)$ has data only for every N samples, then

$$X_k(Z;k) = \sum_{n'} x_k(n'N) Z^{-n'N} = X_k(Z^N)$$

Thus the generalized TMUX reduces to the conventional TMUX.
If the input $X_k(Z)$ has data for every N/2 samples, then

$$X_k(Z;k) = \sum_n (-1)^{kn} x_k(n\tfrac{N}{2}) Z^{-n\frac{N}{2}} \qquad (3\text{-}7)$$

Eq (3-7) can be expressed in another form;

$$X_k(Z;k) = X_k(Z^N;e) + (-1)^k Z^{-N/2} X_k(Z^N;O) \qquad (3\text{-}8)$$

where

$$X_k(Z^N;e) = \sum_n x_k(nN) Z^{-nN} \qquad (3\text{-}9)$$

$$X_k(Z^N;O) = \sum_n x_k(nN+\tfrac{N}{2}) Z^{-nN} \qquad (3\text{-}10)$$

Because the sampling frequency is $2\Delta f$, the bandwidth of the input signal can be greater than Δf up to $2\Delta f$.

3.2 TDUX

The receive signal R(Z) is first frequency converted from f_k to 0(Hz), and then selected by the LPF H(Z). The frequency conversion is made by the variable transform

$$Z \to e^{j\frac{2\pi}{N}k} \cdot Z \qquad (3\text{-}11)$$

Hence;

$$R(Z;-k) = \sum_n r(n) e^{-j\frac{2\pi}{N}kn} Z^{-n} \qquad (3\text{-}12)$$

Then the low pass filter output is given by

$$V_k(Z) = H(Z)R(Z;-k)$$

$$= Z^{-(N-1)} \sum_{i=0}^{N-1} \sum_n Z^{(i-n)} H_{N-1-i}(Z^N) \cdot e^{-j\frac{2\pi}{N}kn} r(n)$$

$$= Z^{-(N-1)} \sum_m Z^{-m} e^{-j\frac{2\pi}{N}km} \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) r(m+i) \qquad (3\text{-}13)$$

The equation gives the structure (B) of Figure 2.
If we select one of every N samples, then we put m=nN to obtain

$$V_k(Z^N) = Z^{-(N-1)} \cdot \sum_n Z^{-nN} \cdot \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) r(nN+i)$$

$$= Z^{-(N-1)} \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) \cdot R_i(Z^N) \qquad (3\text{-}14)$$

Thus the generalized TDUX includes the conventional TDUX as a special case.
If we select one of every N/2 samples, then we put m=nN/2 to obtain

$$V_k(Z^{N/2}) = Z^{-(N-1)} \sum_n Z^{-nN/2} (-1)^{kn} \cdot \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) r(nN/2+i)$$

$$= Z^{-(N-1)} \{ V_k(Z^N;e) + Z^{-N/2} (-1)^k \cdot V_k(Z^N;O) \} \qquad (3\text{-}15)$$

where for even samples

$$V_k(Z^N;e) = \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) R_i(Z^N) \qquad (3\text{-}16)$$

and for odd samples

$$V_k(Z^N;O) = \sum_{i=0}^{N-1} e^{-j\frac{2\pi}{N}ki} H_{N-1-i}(Z^N) R'_i(Z^N) \qquad (3\text{-}17)$$

where

$$R'_i(Z^N) = \sum_n r(nN+\tfrac{N}{2}+i) Z^{-nN} \qquad (3\text{-}18)$$

$$(i=0,1,2, \ldots, N-1)$$

The generalized TDUX above gives the output with sample frequency $2\Delta f$. Therefore the bandwidth of the filter $H(Z)$ can be greater than Δf up to $2\Delta f$.

It is apparent from Figure 2 that the TMUX and TDUX are mutually inverse operations. If the output of TDUX is fed to the corresponding input of TMUX, then the total frequency response is given as follows.

$$Y(Z) = \sum_{k=0}^{N-1} G(Z;k)V_k(Z;k)$$

$$= \sum_{k=0}^{N-1} G(Z;k)H(Z;k)R(Z)$$

$$= R(Z) \sum_{k=0}^{N-1} G(Z;k)H(Z;k) \qquad (3\text{-}19)$$

Thus the whole transmission path is transparent if

$$\sum_{k=0}^{N-1} G(Z;k)H(Z;k) = 1 \qquad (3\text{-}20)$$

4. SATELLITE SWITCHED (SS) FDMA SYSTEM

The block diagram of a SS/FDMA system for mobile satellite system is given in Figure 3. The SS/FDMA system is very effective for mobile satellite communications systems with a number of spot beams. For example mobile to mobile links can be easily set up to avoid the double delay which would be incurred by double hops via feeder links. By adoption of spot beams for feeder links Very Small Aperture Stations (VSATs) can be easily set close to customer facilities and can reduce utility cost for terrestrial networks.

The proposed SS/FDMA system uses the generalized TDUX to divide incoming signals from each beam, put them into the Base Band Switch Matrix (BBSM) which executes the inter-beam switching.

The output of the BBSM is frequency multiplexed by the generalized TMUX to be sent to each downlink.

The total frequency response of the proposed system is depicted in Figure 4. The filters for TMUX and TDUX are designed to satisfy eq (3-20). Thus the proposed system can provide adaptive bandwidth SS/FDMA system with perfectly transparent transmission paths.

5. ACKNOWLEDGEMENT

The author wishes to thank T. Furukawa and other colleagues of him for guidances and useful discussions.

REFERENCES

1. M.Bellanger and J.L.Daguet, 1974 "TDM-FDM transmultiplexer; Digital Polyphase and FFT" IEEE trans, vol. COM-22, pp.1199-1204, Sept. 1974.
2. R.Maruta and A.Tomozawa, 1978 "An improved method for digital SSB-FDM modulation and demodulation IEEE trans, vol. COM-26, pp.720-725, May, 1978.
3. F.Takahata, Y.Hirata, A.Ogawa and K.Inagaki, 1978 "Development of a TDM/FDM transmultiplexer" IEEE trans, vol. COM-26, pp.726-733, May, 1978
4. W.H.YIM, C.C.D.KWAN, F.P.COAKLEY and B.G.EVANS, 1988 "Multi-Carrier Demodulators For ON-BOARD Processing Satellites" International Journal of Satellite Communications Vol. 6, pp.243-251 (1988).
5. H.Scheuermann and H.Göckler, 1981 "A Comprehensive Survey of Digital Transmultiplexing Method" Proceedings of the IEEE, Vol. 69, No.11, November, 1981.

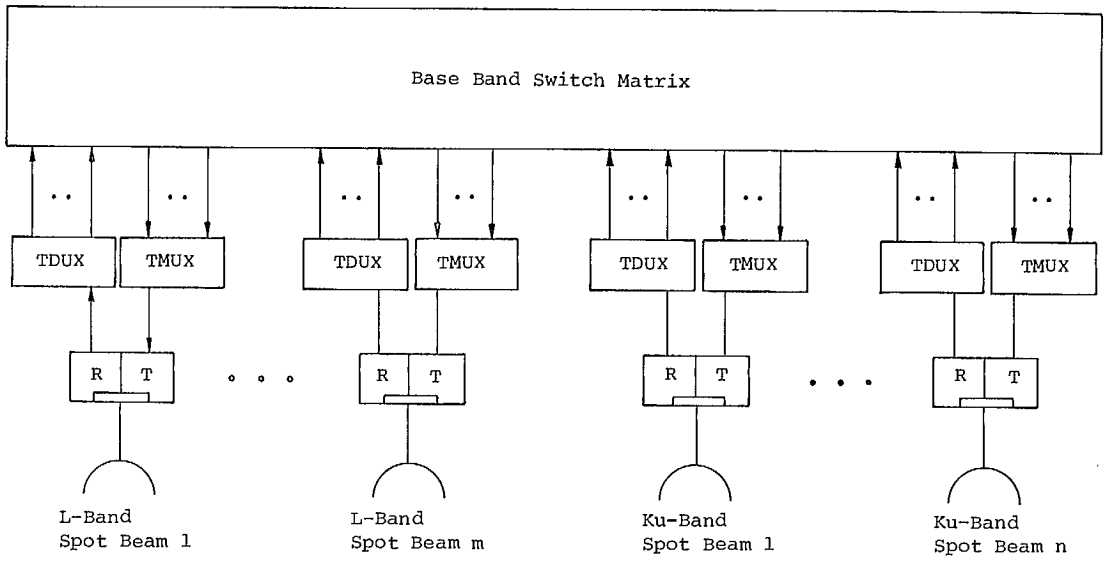
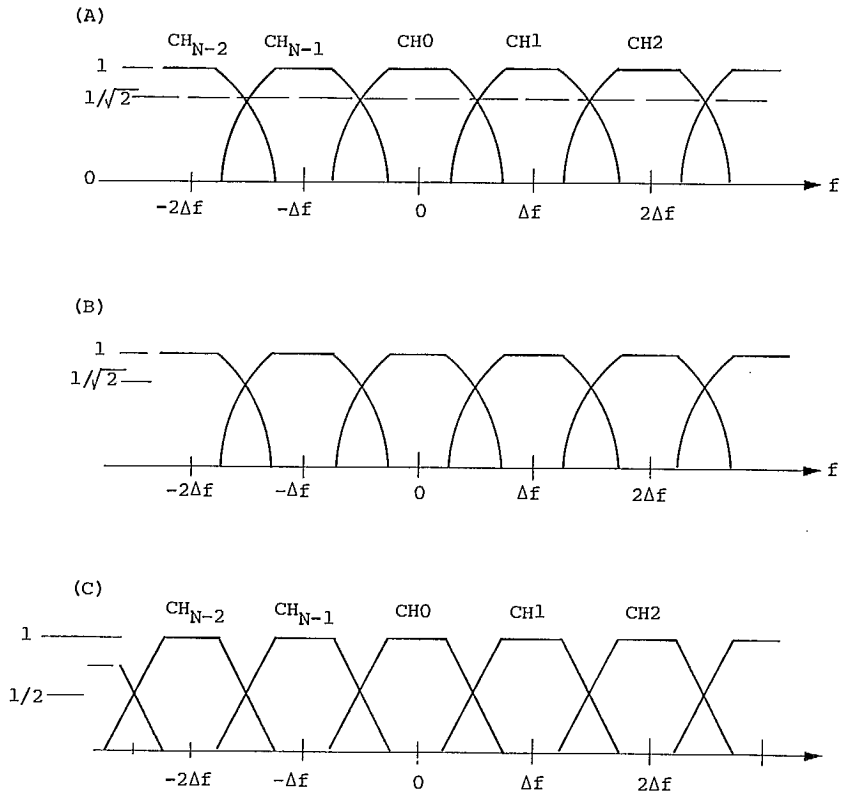


Figure 3 Satellite Switched FDMA System



(A) TDUX Filter Bank
(B) TMUX Filter Bank
(C) TDUX-TMUX Filter Bank

Figure 4 Satellite Switch Frequency Characteristics

Payload System Tradeoffs for Mobile Communications Satellites

H.J. Moody
Spar Aerospace Limited
21025 TransCanada Highway, Ste-Anne-de-Bellevue, Quebec
H9X 3R2 Canada
Phone: (514) 457-2150
FAX: (514) 457-2224

ABSTRACT

This paper describes system level trade-offs carried out during M-SAT design activities. These trade-offs relate to the use of low level beam forming, flexible power and spectrum distribution and selection of the number of beams to cover the service area. It is shown that antenna performance can be improved by sharing horns between beams using a low level BFN and that greatly increased power utilization is possible using a hybrid matrix concept to share power between beams.

INTRODUCTION

Since 1980, Spar Aerospace Limited has been engaged, under contract to the Department of Communication and to the Department of Industry, Science and Technology Canada, in a study and design of a payload for a mobile communication satellite. This paper describes some of the trade-offs carried out during the design of the payuload.

REQUIREMENTS

The design requires the complete coverage of Canada and the United States including Alaska, Hawaii and Puerto Rico. The minimum aggregate EIRP (AEIRP) is to exceed 54 dBW. Frequency reuse capability is required between the east and the west areas demanding a minimum of four beams across the continent. Capability of adapting to changing traffic patterns is also a requirement. As a minimum the capability should allow power to be moved between the two eastern beams and between the two western beams though not between the east and the west.

SYSTEM TRADE-OFFS

The system trade-offs described here include the following:

1. Number of beams covering the continent.
2. Frequency reuse capability.
3. Low level beam forming for improved antenna performance.
4. Capability of changing the allocation of power and spectrum between beams to match the evolving traffic pattern.

NUMBER OF BEAMS

Configurations have been evaluated for 10 beams, 8 beams and 4 beams covering Canada and the United States not counting Alaska, Hawaii and Puerto Rico, which require additional beam(s) in all cases. The 10 beam and 8 beam configurations consisted of 5 and 4 circular beams over Canada and an equal number of slightly elongated beams over the United States. The 4 beam configuration has elongated beams each of which covers part of Canada and part of the United States.

The advantage of fewer beams is the reduction of spacecraft hardware associated with each beam and the possibility of using a smaller L-Band reflector. The advantage of more beams is the increased antenna gain and resulting increased EIRP and in the case of the 10 beam configuration, the increased frequency reuse capability.

FREQUENCY REUSE

Frequency reuse patterns are well known and are characterized by the number of frequency segments required (one per beam) so that the beams are sufficiently far apart and the frequency segments can be reused. For the triangular beam format, the frequency segments can number 4, 7, 9 or higher for a symmetric reuse pattern. For a square beam format, the number can be 4, 9 or higher. Since the North American continent favoured a square beam format, the choice was between 4 and 9 segments of frequency. The 4 and 9 segment plans require every other beam and every third beam to use the same frequency respectively.

To have frequency reuse in every other beam requires that, not only must the first sidelobe be suppressed but also that the main lobe of the nearest alternate beam must not extend into the designated reuse region. This demands very narrow beams formed using large reflectors and results in a large gain difference between the peak of the beam and the edge of the coverage area. This causes a large difference in carrier power through the reverse transponder and an increased demand for feeder link power.

These problems are eased by reusing the spectrum only in every third beam. This allows use of a smaller antenna diameter and reduced feeder link transmitter power. There is a small impact on the antenna gain with the smaller antenna as the system is sized on the antenna gain at the cross over point between beams rather than the peak antenna gain.

FLEXIBLE POWER DISTRIBUTION

In a multibeam environment, the traffic in each beam may not be known at the time the payload is designed, and in addition some changes in traffic distribution can be expected during the life of the spacecraft. If fixed power is assigned to each beam, then it should be matched to the expected traffic in each beam. Preferably, flexibility of power distribution should be built into the payload so that the capability can be matched to the traffic as the traffic evolves. A simplified block diagram of a transponder with fixed power per beam is shown in Figure 1.

The disadvantage of fixed power assignment can be seen from Figure 2. The maximum power utilization factor is shown as a function of the traffic fraction serviced by a fully powered beam for three cases; two beams with two equal power pools, four beams with four equal power pools, and ten beams with ten equal power pools. The power utilization factor is defined as the total power radiated by all beams as a fraction of the total power available in all beams. This can be determined by looking only at the performance of one amplifier in one beam. That is, the power utilization factor is given by the power radiated by one beam as a fraction of the total power available in all beams divided by the fraction of the traffic carried by that beam. The same utilization factor is obtained by considering beams which are not fully loaded. Also, power pools may be unequal as is the case if the beam powers are taylored to the expected traffic distribution. The utilization faction will be on the appropriate line in the case of a fully loaded amplifier and below the line for one not fully loaded but the same utilization factor will be obtained no matter which amplifier is considered.

As the traffic builds up, there will be a point at which even the lightly loaded beams reach full power and the power utilization factor will reach unity. However, in the meantime, traffic has been rejected in the heavily loaded beams.

A design approach is to interconnect the amplifiers for a number of beams by means of a hybrid matrix network[1] which combines the separate power pools into a single power pool and allows any beam to dip into the pool for the number of carriers present at each instant of time. The hybrid matrix consists of a network of 3 dB hybrids connected so that a binary number of input and output ports exist. The configuration for a 4 X 4 (4 input and 4 output ports) hybrid matrix is shown in Figure 3. An input hybrid matrix is placed before the amplifiers which takes the signal at each input port and divides it equally between all amplifiers. An output hybrid matrix is placed after the amplifiers which collects the signal from all amplifiers and directs it to the corresponding output port. In this way, every signal extracts an equal amount of power from all amplifiers in the hybrid matrix power amplifier. A simplified block diagram of the transponder is given in Figure 4 showing the location of the input and output hybrid matrix. In this way for example, four beams with four power pools can be converted to a single power pool by connecting all beams to the hybrid matrix network.

FLEXIBLE FREQUENCY DISTRIBUTION

In order to make use of the power flexibility, as provided by the hybrid matrix, it is necessary to have a corresponding amount of spectrum flexibility.

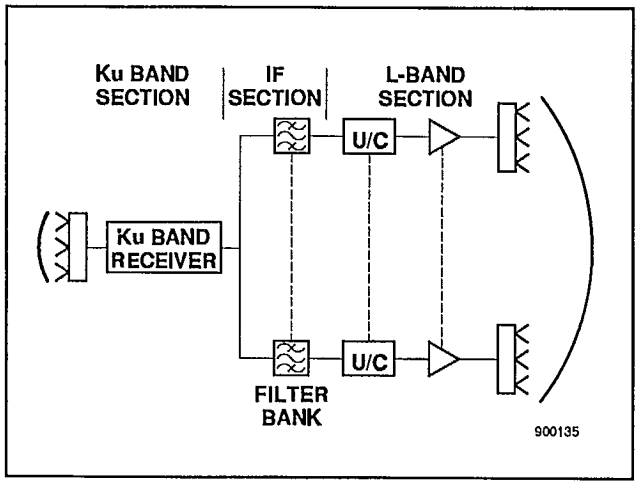


Fig. 1. Forward repeater with fixed power per beam

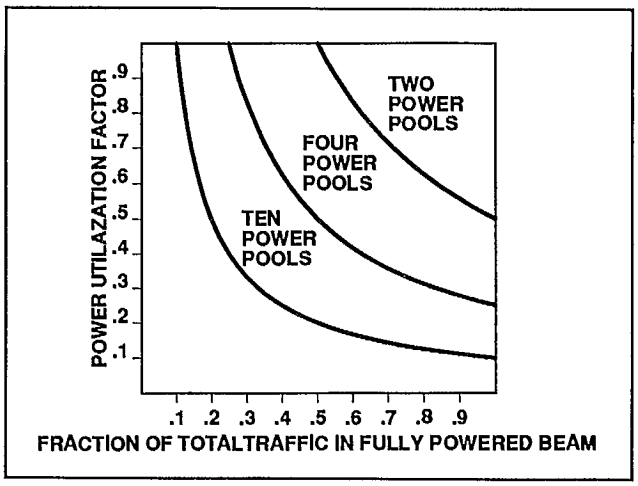


Fig. 2. Maximum power utilization curves for 2, 4 and 10 pools of power

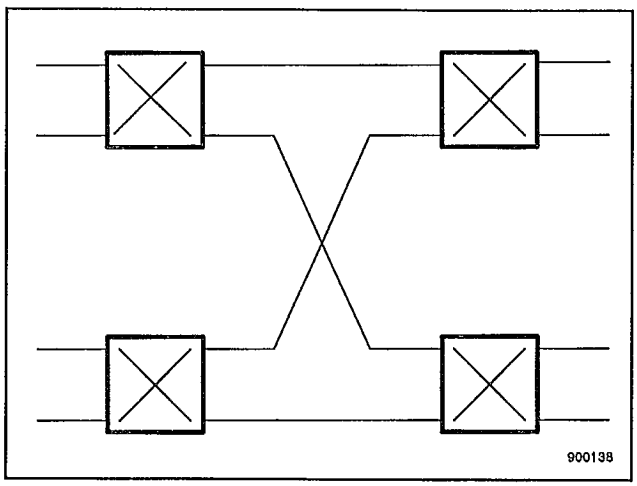


Fig. 3. Block diagram of a 4X4 hybrid matrix

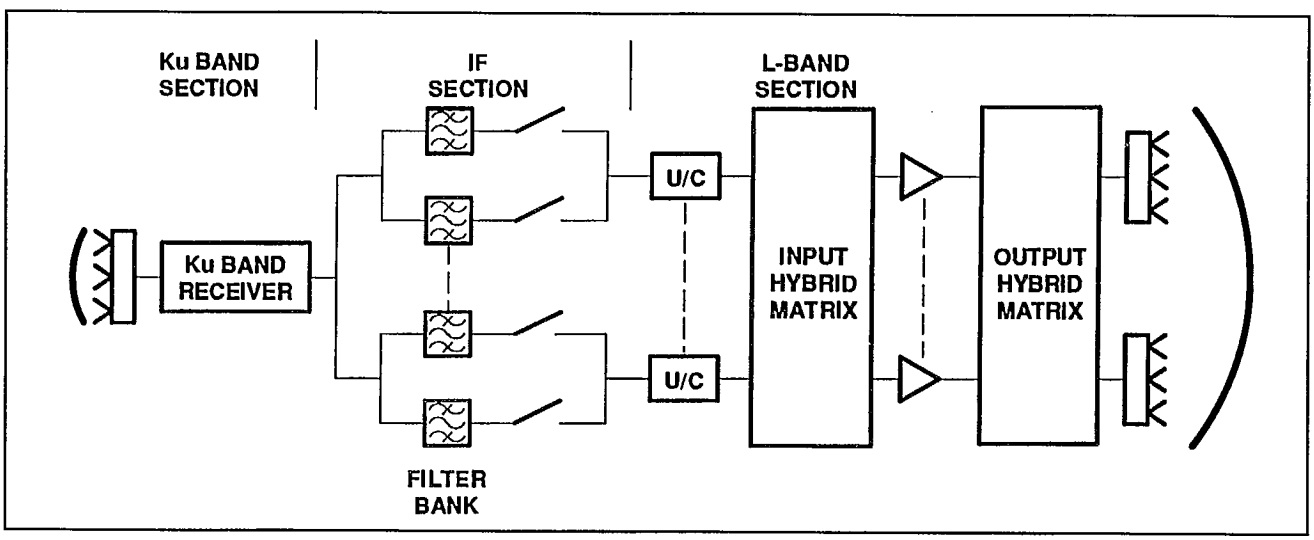


Fig. 4. Forward repeater with flexible power and spectrum distribution

Table 1. Comparison between filter bank and filter switch matrix approach to spectrum flexibility

	Filter switch matrix	Filter bank
Total bandwidth (MHz)	29	29
Filter bandwidth (MHz)	0.5 - 2.0	3.5
Switches per filter	3 to 5	one
Several filters connectable to one beam	yes	yes
Several beams connectable to one filter	yes	yes
Number of filters	40 per use	8 per beam

This is provided in a stepwise fashion by switching filters rather than in a continuous fashion as provided for power with the hybrid matrix. There are two ways of implementing the filter switching as shown in Figures 5and 6, namely a filter switch matrix approach and a filter bank approach.

The filter switch matrix approach connects the filters permanently to the feeder link and connects them to the desired beam by means of switches. A full 29 MHz of filters is provided for each frequency use and connected to the feeder link for transmission to the ground.

The filter bank approach provides a full 29 MHz of filters permanently connected to each beam and these are connected or disconnected as desired.

The filter bank approach tends to have more filtering and less switching whereas the filter switch matrix approach tends to have less filtering and more switching. However, to minimize the number of filters in the filter bank approach, the filter bandwidth is increased. A summary comparison between the two approachs is given in Table 1.

Table 2. Comparison of shared and un-shared horn approaches

	Un-shared horns	Shared horns
Reflector size (m)	5X6	5X5
Edge of coverage gain	Ref.	+1 dB
Peak to edge gain delta	5-6 dB	2.5-3 dB
Reuse isolation	Ref.	+2-3 dB

LOW LEVEL BEAM FORMING

To provide good overlap between beams, it is necesary to share horns between beams, that is to use the same horn as part of the horn cluster for adjacent beams. Because of the loss in the beam forming network, the network is placed before the final power amplifiers (or after the initial LNA's on the receive side) with one amplifier used for each radiating element. This is shown in simplified form in Figure 7. The advantage of sharing horns in this way is a higher antenna gain at the cross-over point between beams and a lower gain variation across the beam compared to un-shared horns. An additional advantage is the reduction of sidelobes in the reuse region. The disadvantage with shared horns is that the antenna and the transponder are intermingled and the performance of each can not be separately evaluated. Table 2 summarizes the performance obtained for M-SAT with both shared and un-shared horn approachs.

REFERENCES

1. **Egami and Kawai.** May 1987. An Adaptive Multiple Beam System Concept. IEEE Journal on Selected Areas in Communications, SAC5, pp. 630-636.

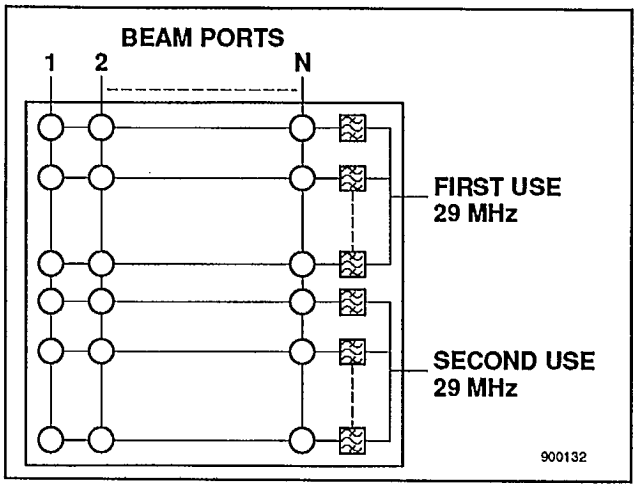


Fig. 5. Filter switch matrix configuration

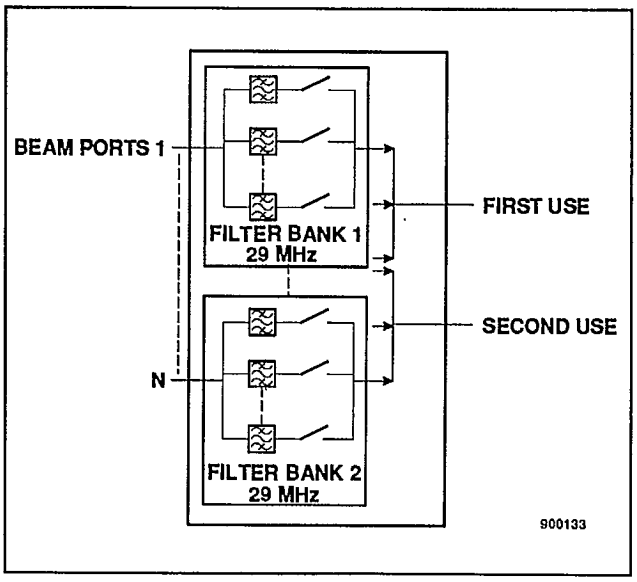


Fig. 6. Filter bank configuration

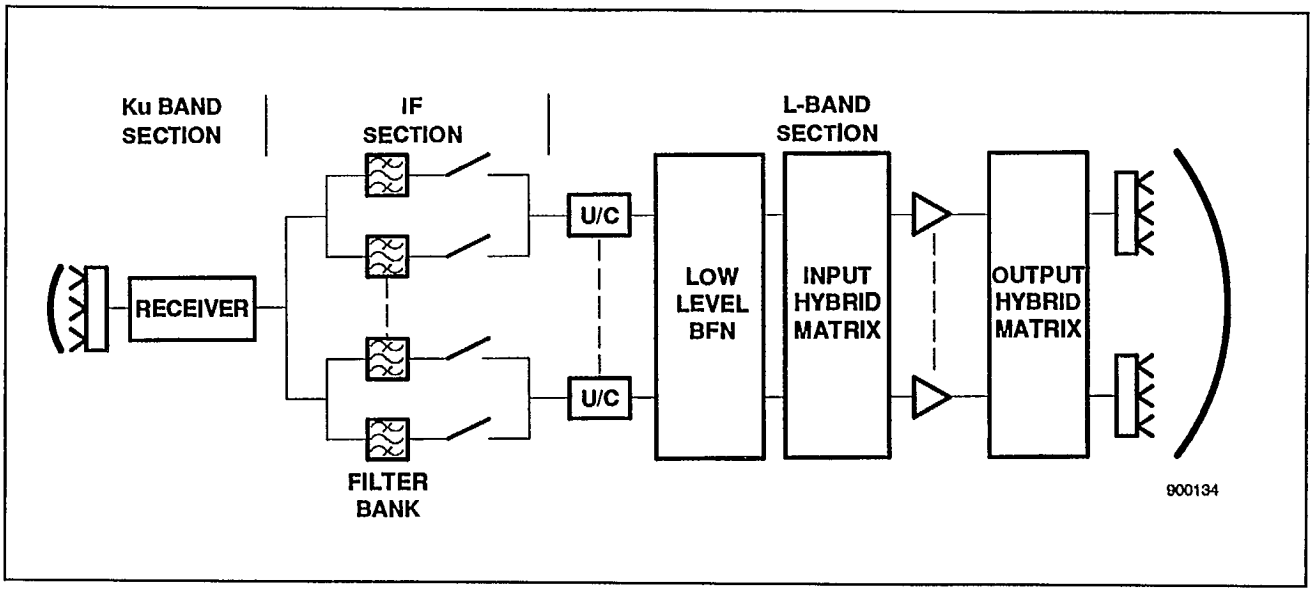


Fig. 7. Forward repeater with low level beam forming and flexible power and spectrum distribution

Session 10
Regulatory and Policy Considerations

Session Chairman - *Maurice K. Nunas,*
Department of Communications, Canada
Session Organizer - *Robert Bowen,* DOC

10

U.S. Development and Commercialization of a North American Mobile Satellite Service

Ray J. Arnold
Director, Communications and Information Systems Division
National Aeronautics and Space Administration
600 Independence Avenue S.W.
Washington, D.C. 20546
Phone: 202-453-1510
FAX: 202-755-9234

Valerie Gray
Manager, MSAT-X Government/Industry Interface Subtask
Jet Propulsion Laboratory/California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109
Phone: 818-354-1260
FAX: 818-393-9876

Jerry Freibaum
Formerly: Chief, Policy and Regulatory Studies, NASA Communications Division
Senior Advisor to the Public Service Satellite Consortium
600 Maryland Avenue S.W., Suite 220
Washington, D.C. 20024
Phone: 202-863-0890
FAX: 202-863-0897

ABSTRACT

U.S. policies promoting applications and commercialization of space technology for the "benefit of mankind," and emphasis on international competitiveness, formed the basis of NASA's Mobile Satellite (MSAT) R&D and user experiments program to develop a commercial U.S. Mobile Satellite Service. Exemplifying this philosophy, the MSAT program targets the reduction of technical, regulatory, market, and financial risks that inhibit commercialization. The program strategy includes industry and user involvement in developing and demonstrating advanced technologies, regulatory advocacy, and financial incentives to industry. Approximately 2 decades of NASA's satellite communications development and demonstrations have contributed to the emergence of a new multi-billion dollar industry for land, aeronautical, and maritime mobile communications via satellite. NASA's R&D efforts are now evolving from the development of "enabling" ground technologies for VHF, UHF, and L-Band mobile terminals, to Ka-Band terminals offering additional mobility and user convenience. Many aspects of NASA's MSAT Program are described in detail.

INTRODUCTION

U.S. Commercial MSS Context

The remarkable growth in recent years of a diverse mobile communications market clearly demonstrates the concept of people emancipating themselves from wall plugs for a variety of communications applications. Mobile satellite communications systems will comprise an important segment of this market, staking out, and likely increasing, their market share for specific applications such as remote/"thin-route" communications in which the Mobile Satellite Service (MSS) is considered the optimum choice.[1] User needs such as interoperability, wide-area coverage, and reliability combine to boost the mobile satellite's utility for public safety and disaster communications as well as numerous other applications.

NASA's commercial space policy encourages private sector involvement in commercial space endeavors. Commercially provided mobile satellite communications are considered a viable alternative to direct public expenditures for acquiring and maintaining Government communications platforms to meet Government research and communications needs. In addition, U.S. technology policies in the 1980s emphasized and encouraged programs that would enhance U.S. competitiveness in international markets. Finally, the 1958 National Aeronautics and Space Act and other U.S. technology policies mandated "socially beneficial" as well as commercial applications of space technology. The combination of all of these factors formed the backdrop for NASA's highly cooperative Government/industry MSAT Program.[2]

During almost 2 decades of leadership in developing communications satellites, NASA and Goddard Space Flight Center conducted over 100 land, aeronautical, and maritime mobile satellite experiments and studies, using

NASA experimental satellites. Among these were the ATS Series satellites (1, 3, 6) launched in the 1960s and 1970s, and the Communications Technology Satellite (CTS), launched in 1976. The CTS was a cooperative Department of Communications/Canada (DOC)/NASA direct broadcast satellite technology development program begun in the late 1960s. The ATS and CTS series were precursors to today's Mobile and Broadcast Satellite Services, paving the way for the current small ground terminal market, and opening up many new frequency bands for satellite services, including L-Band.

Through this experience, NASA determined by the late 1970s that the necessary technology development and regulatory actions appeared feasible for the commercialization of MSS. Through extensive market studies conducted in the early 1980s, NASA concluded that a vast rural communications market existed that was, as yet, untapped. Also, international interest in mobile communications markets was growing, prompting NASA to develop a strategy for the U.S. commercialization of MSS.

NASA's approach to accelerating MSS commercialization in the United States revolved around a basic requirement: the need to reduce the financial and market risks to commercializing the MSS industry.[3] These risks were associated in part with the technology, namely market aggregation for this specific service and the viability and cost of the technology. A significant risk was also presented by the regulatory process. Both of these problems gave rise to a NASA R&D risk-reduction strategy that incorporated technical, regulatory, financial, and institutional elements. All of these elements are critical aspects of the equation needed to commercialize the MSS in the U.S. context, in essence helping to make private commercial space ventures competitive to alternative investments.[4] Inherent limitations of spectrum and orbital slots and other key challenges associated with financial and market risks to the MSS formed the cornerstones of the NASA Mobile Program. These were: (1) development of efficient ground technologies and techniques to conserve power, spectrum, and orbit at the lowest possible cost to users; (2) development of a user base for the technology; (3) allocation of frequencies for a viable MSS; and (4) development of financial incentives to attract industry involvement.

Scope of the Paper

This paper describes the evolution of NASA's MSAT Program and how it is structured to address technical, institutional, financial, and regulatory risks involved in MSS development. Discussion includes historical and current MSAT-X portions of the program, as well as the evolving Personal Access Satellite System (PASS) and Advanced Communications Technology Satellite (ACTS) Mobile Terminal (AMT) Experiment program thrusts (Figure 1).

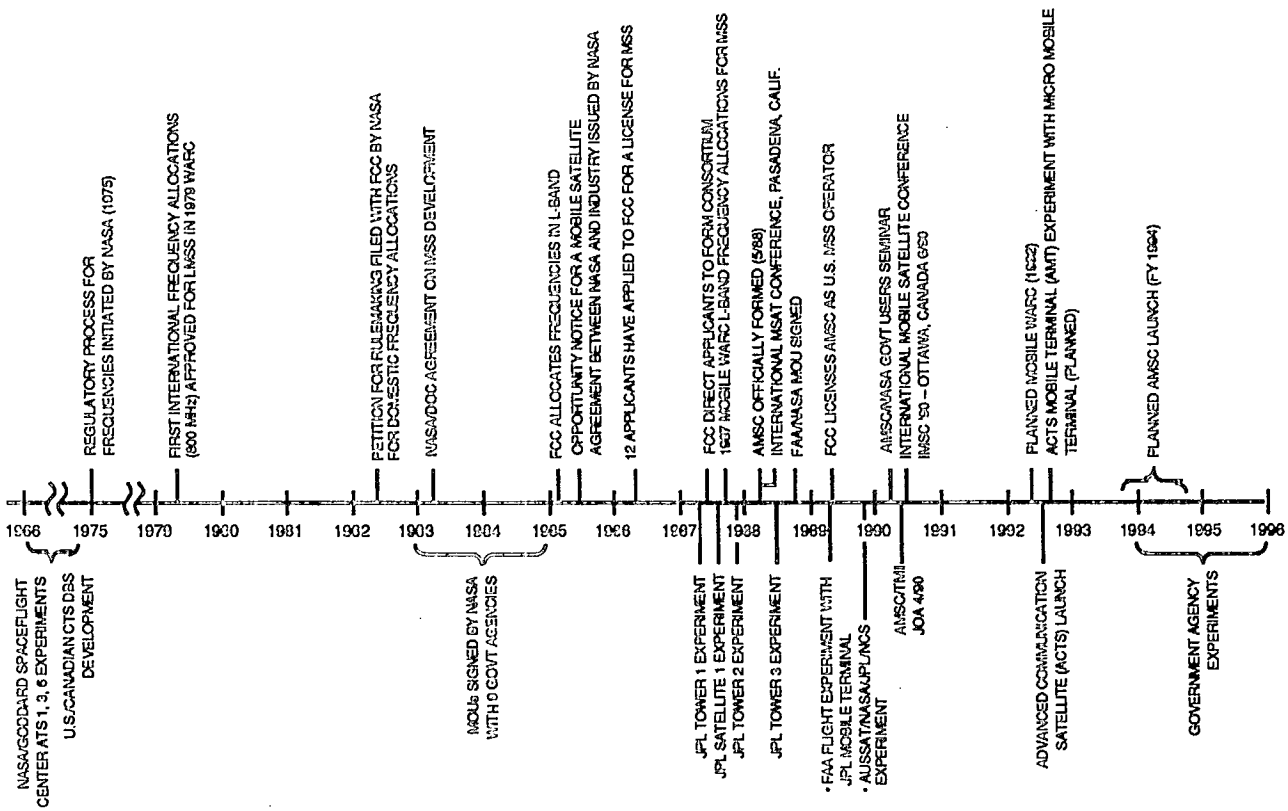


Fig. 1. Key dates in NASA MSAT development

MSAT PROGRAM UPDATE

Program Objectives

The primary objectives of the MSAT Program are to: (1) accelerate the introduction of the first U.S. commercial MSS and (2) enable and enhance future generations of mobile satellite systems.[5] The program was designed to be a highly cooperative Government/industry effort, focused on industry and user needs, and in particular on reducing the risks to industry of commercial MSS development.

MSAT and MSAT-X

NASA's MSAT Program (including the MSAT-X Task) consists of the following elements: (1) high-risk enabling technology R&D, (2) regulatory aspects of MSS, (3) a NASA launch offer to industry, (4) Government agency user experiments, and (5) institutional relationships to support MSS commercialization. In 1982, the Jet Propulsion Laboratory (JPL) was designated the NASA MSAT Program lead center for MSS studies and critical technology identification and development. JPL's MSAT-X Task is focused on mobile satellite communications ground segment development and satellite experimentation; however, MSAT-X also includes the Program's Government/industry interface subtask.

Technology Development. The MSS is a satellite-based communication system, providing voice and data communications to mobile users. Satellites, in geosynchronous orbit, relay two-way digital transmissions over a wide geographical coverage area, making this type of mobile communications ideal for rural (thin route) areas, or as an augmentation of terrestrial systems.

There are three constituent portions of the system. The first is the land-based portion of the network, including the Network Management Center which oversees the operations of the network; gateways, which provide interfaces to other networks such as the Public Switched Telephone Network (PSTN); and base stations, which are normally centers for dispatch operations, not necessarily connected to other networks. The second is the satellite (space segment). The third is the mobile terminal, which is located within the mobile vehicle and comprises the radio, antenna, and user interface. A principal challenge is to have the mobile terminal be low-cost and small in size to ensure a wide customer base and economic viability.[6]

By the early 1980s, the technology critically needed for a first-generation system was considered feasible. A large (10-meter) spacecraft antenna and other vital technologies had already been developed and demonstrated by NASA and Goddard Space Flight Center. Because of the limited bandwidth expected to be available and geostationary orbit placement considerations, the emphasis was on developing or improving the spot beam technology; spectrum-efficient modulation and coding techniques; small, low-cost, low-powered mobile ground terminal hardware; high- and medium-gain directional antennas; and frequency-sharing techniques.

The JPL MSAT-X Task consists of five aspects of ground segment/mobile terminal technology development: (1) steerable, medium-gain vehicle antennas supporting orbit reuse and power conservation; (2) near-toll-quality digital speech at 4800 bps; (3) efficient modems (4800-9600 bps) for 5-kHz channel transmission; (4) network architecture and multiple access techniques for optimal system throughput; and (5) propagation studies to reduce design uncertainties.[7]

First- and second-generation designs are complete, and a prototype terminal has been developed and demonstrated by JPL, with substantial industry involvement and user experimentation. A pre-prototype, mechanically steered, tilted array antenna and an 8 DPSK-TCM modem were developed in-house by JPL. However, in support of the program's technology transfer goals, industry and universities were significantly involved in system development through R&D contracts. Two contractors were involved in the development of antennas: Teledyne-Ryan Electronics and Ball Aerospace developed and delivered breadboards of the phased-array antennas. Speech coders were developed through R&D contracts, one with the University of California at Santa Barbara and another with the Georgia Institute of Technology.

Numerous MSAT-X experiments have been conducted using tower-mounted and satellite transponders. A series of pilot field experiments (PiFEX) was planned and implemented to demonstrate the ability of the moving vehicle to acquire and track the signal and to test the end-to-end link of all subsystem technologies. Three of the PiFEX tests were conducted with a transponder atop a 1000-ft tower to emulate the satellite (Towers 1, 2, and 3). Other tests were conducted using actual satellites in 1987 and 1989. Satellite testing with the mobile terminal was conducted in a flight environment aboard a Boeing 727 in 1989, in cooperation with the Federal Aviation Administration's (FAA's) Technical Center, using INMARSAT's MARECS-B2 Satellite. An L-Band satellite experiment, with JPL- and industry-developed mobile terminal hardware, was also conducted by NASA/JPL in cooperation with AUSSAT, using the Japanese ETS V Satellite's southern beam. Conducted in Australia in 1989, this experiment was made possible through an agreement between AUSSAT and Japan's Communications Research Laboratory (CRL). The National Communications System (NCS), one of the Federal Government agencies having a Memorandum of Understanding (MOU) with NASA, participated in this experiment to validate some of its requirements for a mobile terminal.

Many systems and market studies have been conducted by JPL, including some newly developed network ing techniques and protocols that are being evaluated in JPL's research test bed. Propagation effects studies conducted in JPL's mobile terminal laboratory housed in the

propagation measurement van continue to be invaluable in validating MSS technology. A recently completed cost study of the mobile terminal equipment, conceptually upgraded to incorporate state-of-the-art technologies, indicates that manufacturer's costs are within a reasonable range of estimates.[8] The overall result of these efforts has been the development of a significant body of knowledge in L-Band communications. This information has been documented and disseminated through technical publications, conferences, and workshops designed to promote technology transfer, and facilitate increased user/industry involvement.

Regulatory Aspects. Spectrum allocations and licensing are crucial to being able to obtain investment capital. NASA's effort in the area of regulatory risk reduction was focused on obtaining primary frequency allocations and adequate bandwidth in the 800-MHz band or L-Band or both. This was initiated by NASA in 1975, as part of the U.S. preparations for the 1979 Space World Administrative Radio Conference (WARC). After more than 4 years of domestic proceedings, the United States included in its international positions recommendations for frequency allocations in the 800-MHz band. However, during subsequent Federal Communications Commission (FCC) domestic proceedings on terrestrial mobile allocations in the early 1980s, opposition to an MSS allocation became evident, especially in the 800-MHz band. To preserve the commercial viability of the MSS, and to facilitate the commercialization of its technology, NASA filed a Petition for Rulemaking in 1982 for frequency allocations for the MSS.[9] The Petition argued for an 800-MHz allocation and received formal support from more than 80% of the 92 filings on the Petition.

The number and substance of the filings were clear evidence of the broad user constituency formed as a result of more than 20 years of public and private sector user experiments. This same user community serves, in effect, as a partially aggregated market for QUALCOMM, GEOSTAR, and INMARSAT. The American Mobile Satellite Consortium (AMSC) should also benefit from this process, which has served to reduce the market risk to all.

The FCC Rulemaking Proceeding, initiated by NASA, resulted in frequency allocations in 1985 for the service in L-Band rather than the 800-MHz band, due to economic, political, and regulatory considerations.[10] L-Band allocations were then extended worldwide as an outcome of the 1987 Mobile WARC, culminating 12 years of domestic and international regulatory efforts to achieve this goal.

Unresolved issues remain concerning licensing, transborder operations, interoperability, standards, inter-system interference, coordination, service areas, and additional spectrum provisions for growth of domestic and international MSS.[11] These are serious problems since the numbers of planned systems are increasing substantially. International and national working groups and regulatory bodies are working to resolve most of these issues so that many, if not most, can be resolved before or during the planned 1992 WARC.

Industry/Launch Offer. Reducing financial risk was the objective of NASA issuing an Opportunity Notice for a Mobile Satellite Agreement to industry in 1985 to provide standard launch services for the first U.S. licensed MSS provider, in exchange for satellite capacity for Government experimentation.[12] It was anticipated that combining experimental and operational modes on the same satellite in a Government/industry joint venture or partnership would reduce costs to both parties and significantly expedite the commercialization process.[13] Twelve companies had filed applications for licenses with the FCC by 1986; however, they were later directed by the FCC in 1987 to form a consortium. Eight of these companies eventually formed the AMSC by May 1988, and the AMSC finally received its license in May 1989.[14] The AMSC (now "Corporation") and Telesat Mobile Inc. (TMI) Canada signed a Memorandum of Understanding in the summer of 1989, and, pursuant to this MOU, have recently signed a Joint Operating Agreement (JOA) for providing roaming capability and mutual satellite capacity backup (April 25, 1990). AMSC and TMI are each currently involved in procuring their satellites. AMSC and NASA negotiations are under way for the exchange of a launch and satellite capacity, based upon the 1985 Opportunity Notice. The launch is currently carried on NASA's mixed fleet manifest, with a planned launch date in Fiscal Year 1994. AMSC plans to initiate early service (data only) prior to this launch date via leased space segment capacity.

Further indication of the market viability of this service is the expressed interest that other companies have in marketing related services. On an international basis, competing MSS providers are considering the use of portions of the same spectrum allocated in the 1987 Mobile WARC. Annual projections for the MSS business in the 1990s are placed in the multi-billion dollar range.[1]

Government Agency/User Development Experiments. A key component of the approach to address market risk relates to fostering market aggregation. Between 1983 and 1988, NASA signed MOUs with 10 Government agencies to be experimenters on the satellite for a 2-year period or more, using bartered capacity. These agencies include eight federal and two state agencies. Additional agencies have recently expressed interest in participating in the experiments, prompting the development of a plan for soliciting experiment proposals to allocate experimental capacity, based upon proposal merit and experiment value. Implementing these experiments will be linked to the outcome of AMSC/NASA negotiations for the launch; however, the substantial interest of Government agencies in experimentation,

indicated in a recent Government users' seminar, clearly illustrates the potential Government demand for the service.

Throughout the development of the MSAT mobile terminals, the MOU agencies were encouraged to interact with the NASA/JPL research staff. Two agencies, the FAA and the NCS, also participated in satellite experiments with the system, relevant to their agencies' operational concerns. These experiments were successful, and served to validate the usefulness of the system in meeting the operational requirements important to the users. They also afforded the users an opportunity to become familiar with the system in a "hands on" setting.

Institutional Relationships. Both international and domestic institutional relationships are included in NASA's program structure (Figure 2). In 1983, NASA and Canada's DOC extended their earlier cooperation from the 1970s CTS Program, through signing an agreement to foster development of a North American MSS. Domestically, the institutional strategy has been aimed at developing a strong Government user base for the technology as well as disseminating relevant technical information to industry (the intermediate user) in a timely fashion. NASA's formalized relationships with 10 Government agencies that have signed MOUs to be experimenters with the satellite capacity, like Canada's Government trials program, support the objective of user base development. JPL's use of industry and university R&D contracts for developing the mobile terminals was designed to promote closer Government/industry ties and cooperation. A Government/industry interface subtask at JPL supports the integration of NASA's launch offer activities with Government experiment planning and oversees the dissemination of MSAT-X technical information to industry and user groups.

A whole family of institutional relationship issues pertinent to a viable domestic and international MSS is evolving to include: ownership, tariff arrangements, service backup agreements, interoperability, service area overlaps, priority of traffic, network control, and conflicts with international carriers. The recent establishment of a JOA by AMSC and TMI, described previously, should address one or more of these issues.

Advanced MSAT Research in Ka-Band

NASA's MSAT Program is now evolving to research and development in Ka-Band technologies, based upon projected use of, and constraints upon the L-, C-, and

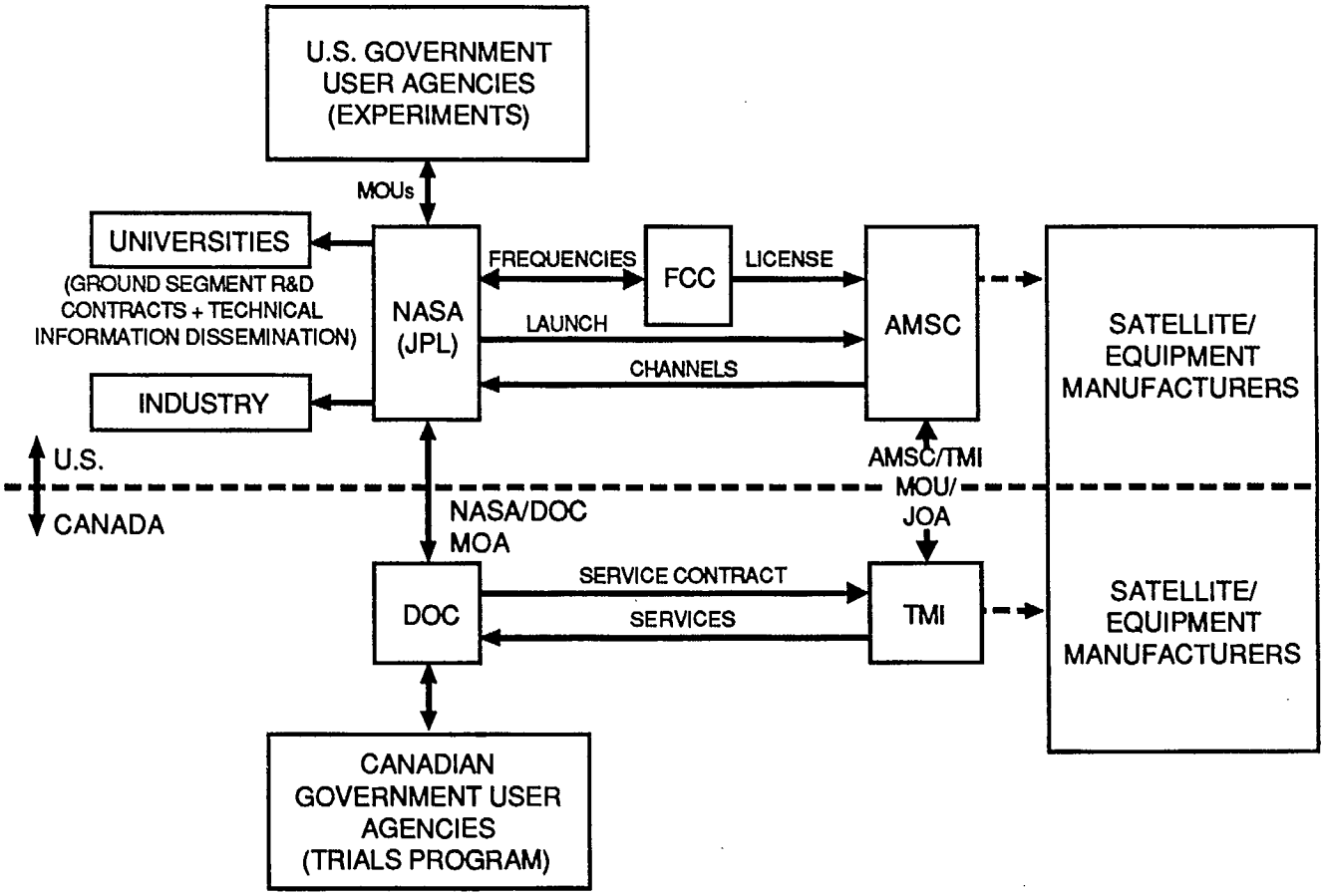


Fig. 2. Mobile Satellite Communications Program: Institutional Arrangements (NASA/DOC/Industry)

Ku-Bands. Although technical and economic challenges exist, the tradeoff is the availability of spectrum and orbital parking slots, making Ka-Band an important follow-on of L-Band research for mobile satellite applications.[1] Two research thrusts are under way at JPL: (1) PASS technology development and studies to evaluate a number of terminal options for personal communications, including stationary, transportable and ambulatory (personal) options; and (2) the development of a terminal for an ACTS satellite experiment planned for 1992. The Acts Mobile Terminal experiment is linked to the PASS task by providing an early demonstration of a mobile micro terminal application of the ACTS satellite as well as a "satellite of opportunity" for validating other PASS technologies.

LESSONS LEARNED

Value of User Experiments

Emphasis on user involvement in mobile satellite applications helps to validate system usefulness and refine its design. It also provides users and industry with an early awareness of and exposure to evolving technology, thus addressing the "push" and the "pull" of technology transfer. In addition to supporting market aggregation, this also serves to help support regulatory positions, when needed. An excellent example of user involvement was the AUSSAT experiment, which gave the NCS an opportunity to conduct an early experiment over an L-Band satellite. This experiment served as an excellent opportunity to give visibility to U.S. technology developments in a highly cooperative international setting.

Value of Information Dissemination

Because technology transfer is dependent in part upon technical know-how, good quality information and effective information dissemination methods are essential. Technical publications, conferences, workshops, and technical interchanges are all useful mechanisms that support technology transfer.

Importance of Addressing Technical, Regulatory, Institutional, and Financial Aspects of the MSS

Accelerating commercialization requires that the complex interactions of technology readiness, institutional, market, and policy factors impacting a technology's commercialization pathway be molded into an integrated program strategy. Integrating program elements associated with a variety of potential barriers has helped to keep this pathway open.

The Case for Lowering Risk

Reducing the risk to industry is only partly addressed through Government R&D and demonstrations of long-term, high-risk advanced technologies. To attract private investment, it is essential that industry be assured of direct benefits (i.e., profits of sufficient magnitude) that a specific investment (e.g., MSS) is attractive compared to other alternatives.[2] Expectations of profits are associated with risk, especially in areas such as defined markets, regulatory delays, and system costs. This constitutes a strong case for a multi-faceted, risk-reduction strategy.

Viability of MSS

Market viability can be seen in cost studies, market projections, and increasing public and private sector applications and demand for this service. Viability and growth of the system should continue when mobile satellite is viewed as a complementary as well as a competing system.[1] Its ultimate viability, however, is likely to depend upon the allocation of sufficient spectrum to assure its financial viability in a commercial world.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of Dr. John Dickman, NASA MSAT Program Manager, Dr. William Rafferty, JPL MSAT-X Manager, and Dr. Byron Jackson in reviewing the paper, to Ms. Laura Steele for contributing to early planning of the paper; and to Ms. Ermine van der Wyk for editing the paper.

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. **NASA.** 1990. NASA *Communications Division Strategic Plan*, Draft, pp. 5–11.

2. **Office of Technology Assessment.** 1982. *Civilian Space Policy and Applications*, Library of Congress 82-600564, p. 3-48, 267, 272, 283.

3. **Citibank N.A.** 1981. Financial study for a Satellite Land Mobile Communications System. JPL Contract for NASA No. BP-736564, p. 16.

4. **Freibaum, J.** 1986. Commercial space policy: Theory and practice, in *Tracing New Orbits: Competition and Cooperation in Global Satellite Development*, Section 11, ed. D.A. Demac (Columbia University Press), pp. 157–159.

5. **NASA.** 1985. NASA's Mobile Satellite Communications Program (Technical Briefing Package). JPL 400-257 (Pasadena, California; The Jet Propulsion Laboratory), pp. 2–6.

6. **Rafferty, W.** 1988. NASA's Mobile Satellite Development Program. *Mobile Satellite Conference*

Proceedings, JPL Publication 88-9 (Pasadena, California; The Jet Propulsion Laboratory), pp. 9, 12–15.

7. **Dickman, J. E.** 1989. NASA Mobile Satellite Communications Program Briefing Chart, p. 14.

8. **Jackson, B.** 1990. *Cost Study for MSAT-X Vehicle Communications System*. JPL D-7210 (Pasadena, California; The Jet Propulsion Laboratory), p. 32.

9. **NASA.** 1982. Petition for Rulemaking for the Establishment of a Mobile Satellite Service, FCC RM-4247 (NASA Headquarters, Code EC).

10. **Newman, D. B., Jr.** 1986. Land Mobile Satellite Service – Part II: L-Band. *IEEE Communications Magazine*. Communications and the Law. Vol. 24, No. 6, pp. 50–52.

11. **Freibaum, J.** 1988. International and Domestic Mobile Satellite Regulatory Proceedings: A Comparison of Outcomes and Discussion of Implications. MSAT Conference Paper, JPL Publication 88-9 (Pasadena, California; The Jet Propulsion Laboratory), pp. 71a–71f.

12. **Office of Space Science and Applications (OSSA).** 1985. NASA Headquarters, *Opportunity Notice for a Mobile Satellite Agreement*, Section 1, p. 1.

13. **Freibaum, J.** 1980. Land Mobile Satellite Service: Concept, policies, and regulatory issues, *INTELCOM 80*, Los Angeles (Dedham, Massachusetts; Horizon House International), p. 5.

14. **MSAT-X Quarterly.** 1989. FCC Grants MSS License to AMSC. Vol. 20, (Pasadena, California; The Jet Propulsion Laboratory), p. 1.

Canadian Development and Commercialization of a North American Mobile Satellite Service

Demetre Athanassiadis
Chief, Satellite Applications
Department of Communications
300 Slater Street
Ottawa, Canada K1A 0C8
Phone: 613-990-4111
FAX: 613-952-1231

ABSTRACT

Canada recognized early the value of mobile satellite communications, originally through the planning of a military system (MUSAT) and subsequently through the development of the Canadian MSAT system. Acting on behalf of the government, the Department of Communications (DOC) defined and carried out a complete plan for the implementation of Mobile Satellite Services (MSS). Based on an extensive dialogue between government, industry and users and encompassing all technical, economic regulatory and institutional aspects, this plan resulted in the completion by 1986 of a comprehensive business plan and a decision for commercial MSS delivery. The Canadian lead for a commercial system was quickly followed by others, and in particular the U.S., giving rise to the concept of North American MSS. In moving rapidly towards commercialization, we must learn from the past and plan for the future.

INTRODUCTION

Canada has been in the forefront of satellite communications for almost 30 years. In 1962, Canada became the third country in the world to launch a satellite and to pioneer satellite communications through Alouette 1, an upper atmospheric research satellite.

Although in intervening years surging international demands resulted in the proliferation of communications satellites, Canada continued to play a leading role in the development of new applications communications satellites. The DOC teamed up with the U.S. National Aeronautics and Space Administration (NASA) in the development and launch in 1976 of the Communications Technology Satellite (CTS), a direct broadcast satellite. Again, in the 1970's, the DOC in cooperation with the Canadian Department of National Defence (DND) and other government departments, developed the concept of a Canadian multi-purpose UHF satellite communications system called MUSAT. This system was conceived as the most cost-effective means of satisfying government needs for voice and data communications with ships, aircraft, vehicles and transportable stations operating in the Canadian north and other areas where conventional terrestrial systems could not provide services for technical and economic reasons. The MUSAT system would have operated in the 240-400 MHz band which is used extensively in both Canada and the United States for military communications and was the forerunner of MSAT.

At the 1979 WARC, the decision was taken to permit shared satellite and terrestrial communications services to mobile radio and telephone users in the 806-890 MHz frequency band in Region 2 (North and South America). This new allocation allowed the benefits of mobile satellite communications to be extended to the general public and non-military government applications in such areas as oil exploration and exploitation, mining, trucking, shipping, business, personal communications, law enforcement, forest fire fighting, ambulance communications and resource management in remote areas. The availability of mobile satellite communications to the public at large was of distinct interest in view of the Canadian topography. The DOC re-directed its efforts to conduct preliminary studies in cooperation with NASA in early 1980 on the use of satellite technology in the provision of public mobile communications services in Canada and the U.S. These studies explored system concepts and gave strong indications that market demand would be sufficient to ensure the future commercial viability of satellite systems providing services to mobile terminals if the required technology and services were developed.

Thus the main emphasis shifted to minimizing or eliminating the technical, commercial and administrative risks involved. In effect, the government invested approximately $20M, during 1982 and 1983, in order to fully justify the $200M-$300M implementation expenditure involved.

The government absorbed the complete cost of these studies since, under the original program direction, government was to be the main investor and owner of a demonstration MSAT system. The private sector's subsequent expression of interest in eliminating the demonstration system and moving directly to a commercially funded and owned system was in line with the strategic objectives defined for the Program at the outset. In fact, the private-sector interest in a commercial system was proof that the original targets of the Program had not only been met, but were exceeded. In 1984,

Telesat Canada assumed the role of the satellite owner and operator, and later transferred this role to a subsidiary Telesat Mobile Inc. (TMI).

This change in program direction did not significantly affect or alter the original objectives, although it did increase the strategic importance of the market and economic viability studies as well as the expedient implementation of the appropriate regulatory, institutional and international arrangements.

CANADIAN MSAT PROGRAM DEVELOPMENT

Program Objectives

There are two primary objectives governing the Canadian MSAT Program:

1) The development and availability of efficient and diverse mobile communications services in underserved areas of Canada; and

2) The creation of the necessary conditions under which Canadian domestic industry can develop and market effectively MSAT services and products domestically and internationally.

Although preliminary studies were encouraging, there still prevailed considerable skepticism as to whether or not the state of the art was sufficiently advanced as to permit the introduction of such services at realistic cost and risk. In this environment, it was essential to the achievement of objectives that the key collaborators be assured of reasonable risk-benefit tradeoffs. In the case of government it was important to show that the overall socio-economic benefits justified the expenditure of the public funds required for technology development and risk-sharing programs. Service providers had to be reasonably assured of sufficiently attractive returns on investment that would outweigh the higher than normal risks involved. Finally, manufacturers had to be reasonably assured that near term research and development outlays would translate into long-term portfolio and revenue gains.

Program Structure

Given the environment, it was evident that the DOC would have to provide much more than technical, financial and institutional support. What was required was an all-encompassing BUSINESS PLAN for the Development and Commercialization of Mobile Satellite Services (MSS), to be developed and managed by DOC on behalf of one of the main players (government) in consultation and cooperation with the other two (service and manufacturing industries). The resulting structure entailed numerous elements, with the main ones being:

1) system and service definition
2) user indoctrination and interaction
3) market definition and development
4) development of enabling technologies
5) overall commercial viability and socio-economic benefit evaluation
6) domestic regulatory and institutional arrangements
7) government risk-sharing support
8) international coordination and cooperation

The intent was for this BUSINESS PLAN to be fully executed in the five-year interval 1982-1987, with the ultimate success criterion of handing Program leadership over to the private sector at the end of that period.

System and Service Definition: Early in the program evolution, it was recognized that MSAT was unlike any other telecommunications system previously implemented which catered to point-to-point telecommunications. This was, indeed, a Telecommunications Distribution Network of end-user mobile satellite services similar, in many respects, to the telephone network.

Thus, two early documents were prepared, one dealing with a proposed system concept,[1] and the other defining a strawman services and applications scenario.[2] The system concept served as the main source document for the numerous technical system studies carried out in-house and through contracts let to the private sector. The strawman service description formed the basis of an unprecedented program of end-user indoctrination and interaction. It is interesting to note that the main service concepts contained in the strawman document have withstood the test of time and form the basis of presently proposed services and applications.

User Indoctrination and Interaction: In the 1982-1984 timeframe, user and service provider awareness of the potential of MSS was lagging considerably. The user audience to which the services were addressed was primarily familiar with private network radio communications and the majority of prospective MSS service providers were primarily familiar with point-to-point satellite communications. DOC initiated an intensive indoctrination program reaching hundreds of potential user organizations across Canada. Some two hundred of these became sufficiently interested as to enter into a continuing dialogue with DOC. This dialogue resulted in a broad base of educated users on one hand and on the other provided invaluable feedback on potential service refinements and improvements that were instrumental in finalizing the repertory of MSS services presently advocated. The user interaction resulted in a further originally unexpected benefit. Due to the broad user base involved and the close interaction, extremely detailed market demand and cost-sensitivity projections were obtained. In many respects, these projections proved to be more accurate and meaningful than the information compiled through statistical market and cost-sensitivity studies. For instance, the first indication of the now proven large demand for message oriented and other data services first became apparent through the user interaction process. Early user involvement has proven an invaluable tool in system and service evolution.

Market Definition and Development: Apart from the market and cost sensitivity obtained through the user interaction, a major statistical study was carried out by Woods Gordon Management Consultants.[3] This study was primarily based on feedback received from an extensive mailed questionnaire and a limited number of interviews. Although the questionnaire format failed to reveal certain characteristics and aspects of the market, such as the issue of data services previously mentioned, it constitutes, nonetheless, the most

extensive independent market analysis available and its findings, modified through subsequent experiences, formed the basis for the commercial viability and socio-economic benefit assessments.

In addition to the overall market potential the Woods Gordon study provided inputs on the cost sensitivity of the MSAT services indicating the percentage of the potential market to be captured if the services were to be priced according to a number of pricing alternatives. Based on the above and after applying certain correction factors, to account for information made available at a later time or received from other sources, TMI developed the "baseline" market scenario used in the business plan.[6]

While market definition and sensitivity studies provide a good indication of potential commercial viability, the actual rate of market penetration is highly dependent on the early development of a broad user base. This is especially true in the case of MSS where the lag between system and service definition and actual implementation is very large. Recognizing this, the DOC defined and sponsored a program of trials and experiments aimed at generating user awareness and allowing potential end-user service providers to develop specialized services and products. This program is still ongoing and is dependent on the procurement of airtime on satellites of opportunity. While extremely well received this program has not reached its full potential due to uncertainties associated with the procurement of adequate airtime and overall program delays that have impacted user confidence.

Development of Enabling Technologies: For all the emphasis paid on other elements of the Program that helped define and justify the overall business plan, technology development consumed the bulk of the budget and resources of the program. Over twenty major studies were commissioned on space and ground segment development.[4] Major attention was paid to payload, user terminal and UHF antenna technology considered critical in the formulation of the overall business plan. In parallel, extensive in-house research has been and continues to be carried out at the Communications Research Centre (CRC) of the DOC on both digital and analogue vocoder development. A major portion of the products of this development has been already passed on to industry.

By virtue of catering to the expedient development of a Canadian commercial system, technology development has focused primarily on near term development targets. As a result, second generation requirements have only just started to receive serious attention.

The most significant results, to date, have been accomplished in the area of payload development with SPAR Aerospace holding a prominent place amongst international satellite manufacturers.

Overall Commercial Viability and Socio-Economic Benefit Evaluation: The effect of the Program on a number of sectors had to be assessed in order to arrive at an overall evaluation of the impact of MSS introduction. Key areas evaluated were the impact on the communications industry as a whole,[5,6] the commercial viability for the implementors of MSS,[7] the impact on the manufacturing industry,[8] and the end-user benefits.[9] It being crucial that these evaluations were realistic and unbiased, the DOC chose to task the principals, wherever possible, in the conduct of the required studies. Using these studies as an input, overall socio-economic assessment[10] was carried out. The results were interesting, although predictable. Users of MSS stood to gain the most and the fastest; Communications Carriers could expect very good long-term returns after a somewhat bumpy ride initially; the potential markets generated for equipment suppliers were extremely large but subject to keen competition; and finally, the overall benefits to the country promised to be very substantial thereby fully justifying government financial support and risk-sharing.

Regulatory and Institutional Arrangements: Given the extremely encouraging results from other elements of the program, the DOC moved quickly to resolve the regulatory and institutional issues involved. Within the framework of existing domestic communications policy and regulations the following actions were taken: TMI was designated as the sole satellite owner and operator; a framework for end-user service distribution was established; and guidelines for the appropriate regulatory policies were issued. Suddenly the issue of frequency allocations became the only remaining unresolved piece of the puzzle.

Ever since the inception of the Program, a domestic frequency allocation at 800 MHz was taken as a given. In the euphoria created by the allocations obtained at the 1979 WARC, certain major issues pertaining to the adequacy of spectrum and the difficulties of bilateral coordination were partially overlooked. Suddenly it became apparent that the planned Canadian use of the very limited UHF spectrum available at 800 MHz would be subject to long and arduous negotiations between the Canadian and American administrations. As a result, the search for additional or alternate spectrum began again in both Canada and the U.S. with attention directed at the 1.5 GHz band. But the results of this search could not be known prior to WARC-MOB 1987. With all other aspects of the domestic MSAT business plan in place, government and industry agreed to finalize and publicly announce, in early 1986, the implementation decision for the domestic MSAT system, subject to adequate spectrum allocations in the 1987 time frame.

Government Risk-Sharing Support: Key to the final decision for MSAT implementation were the official endorsement and submission by TMI of a comprehensive Business Proposal,[11] and an extensive Cabinet Document prepared by DOC. In April 1986, the government approved a support package containing the following major items:

1) approximately $30M of 50¢ dollars for the development of MSAT space and ground segment products;
2) up to $20M for the sponsoring of user trials; and
3) a $126M guaranteed lease for MSAT services to be procured from TMI over the life of the first generation system or a period of 15 years, whichever is longer.

Assessment of the MSAT Business Plan Development: By virtue of its charter to be the

chosen government instrument for both communications technology development and regulatory and institutional policies, DOC was in a position to define, develop, and justify a complete Business Development Plan.

Under this Plan, and having proved technical feasibility, commercial viability and overall socio-economic benefits, the DOC succeeded in putting in place the following commercialization enabling arrangements.

- a satellite owner and operator was selected and assured of a licence as early as 1983;
- an original domestic spectrum decision was taken as early as 1982 and modified in 1986 as a result of the associated FCC rulemaking;
- key players for the space and ground segment were identified as early as 1983;
- the types and levels of government support and incentives required to reduce risk were developed; and
- Program implementation received final approval by April 1986

All of the above helped define the set of key players and the form of the playing field, thus allowing the cooperative evolution of detailed systems and service parameters, a proposed service distribution scenario and a unified government and private sector front in international negotiations and coordination matters. Technical-economic tradeoffs were considered early. For instance, even though the technical feasibility of large spacecraft antennas was already demonstrated and the DOC was proposing antennas as large as 10 meters in diameter, it became apparent that both manufacturers and service providers were not prepared to adopt them for the first generation on the grounds of cost and risk. Other countries have since adopted a similar cautious approach on their proposed and, about to be, implemented systems.

Effective April 1986, the Canadian MSAT Program has been turned over to the private sector with DOC maintaining a support and oversight role. This is one year prior to the originally anticipated date and could have occurred as early as 1985 in the absence of spectrum availability complications.

Finally, commercialization of the Canadian Program acted as a catalyst for the acceleration of commercial plans in the U.S.

International Cooperation and Coordination: Throughout the MSAT Program development the DOC engaged in numerous cooperative arrangements involving domestic and international collaboration. Major domestic cooperative arrangements were executed with Telesat Canada first through a Memorandum of Understanding (MOU) subsequently replaced by a Joint Endeavour Agreement (JEA). Similarly, an MOU was executed between DOC and Spar Aerospace. Internationally, by far the most notable cooperation has been that with NASA, officially governed by the conditions contained in the NASA/DOC MOA executed in 1983. However the NASA/DOC cooperation is primarily a long standing association which has extended over a period of 20 years and is expected to continue in the 1990's at least as far as the present and future generations of MSS are concerned.

DOC's mandate further extends in the areas of international coordination. In this capacity DOC represents Canada in international fora such as the ITU and CCIR and in bilateral and multilateral negotiations such as frequency coordination, reciprocal operating agreements, etc. In this capacity, DOC has maintained a continuous interface with the FCC on MSAT Program matters, holds frequent meetings on MSAT spectrum coordination with INMARSAT, the U.S.S.R. and the U.S. and speaks on behalf of Canada in the ITU. These meetings which are de facto intense and partisan, have once again demonstrated that even the 1.5-1.7 GHz allocations established in 1987 are not sufficient to meet all reasonable domestic and international demands. Clearly the WARC-MOB 1992 will be faced with renewed demands for additional MSS allocations.

NORTH AMERICAN MSAT DEVELOPMENT

The difficulties encountered with the frequency allocations in the 800 MHz band were a sobering experience for Canada and possibly NASA. These difficulties did act, however, as a silver lining in that they generated wide awareness of the limited resource environment and of the need to think in terms of a North American MSS as opposed to individual national programs. This concept had certainly been considered by DOC and NASA in the past, but might have not been easy to sell to the regulators and the entrepreneurs but for the spectrum crisis that developed. As a result, there developed strong synergies between private sector counterparts in Canada and the U.S. as well as between the two administrations.

Finding themselves in a clearly underdog position prior to WARC-MOB 87, the two administrations aligned their positions and embarked on a worldwide campaign to generate international support for their proposal to generalize the existing mobile satellite allocations in the 1.5 - 1.7 GHz bands. Although not all of the Canada/U.S. proposals were accepted in the WARC-MOB 87, the resulting modifications contained in the final acts of the conference must be regarded as a victory, given the strong opposition provided by most European administrations and others.

The would be MSS providers in both Canada and the U.S. also moved quickly to endorse the North American service concept. Spectrum shortages were amongst the leading factors but the process was also assisted by the fact that U.S. MSS hopefuls found it expedient to align their efforts with those of TMI who was already the designated Canadian satellite owner and operator. Finally, both sides had a lot to gain by minimizing their initial capital expenditures. For whatever reasons, the present common approach towards MSS development makes immanent sense. The recent signing of a Joint Operating Agreement between TMI and the American Mobile Satellite Corporation (AMSC) is a major step towards the long overdue implementation of MSS in North America. While there are still outstanding issues remaining, it is reasonable to assume that these will be resolved in an expeditious manner; there is just too much competition domestically and internationally to afford the luxury of further major delays.

The DOC, in its continuing role of support, oversight and policy implementation, will continue to fully support and encourage the quick implementation of MSS in North America.

LESSONS LEARNED

Successes and Failures

There have been numerous accounts of the Canadian MSS development and, for his part, this author vows that this is his last attempt at a post mortum. MSS is well on the road to maturity and should be treated accordingly. For these reasons, a moment of reflection is well in order.

It is perhaps an overlooked fact that the MSS development has signalled the end of infancy for satellite communications services, for prior to MSAT and for all the mystique of space, satellite systems remained a glorified form of carrier or cable systems. For the first time, a satellite-based system had to cope with end-user service and all the complexities involved in the interconnection, distribution and control of a large communications network. The trend is irreversible! More and more future satellite service offerings will involve total network approaches.

The realization followed naturally but not without considerable strain that implementation decisions should not be left solely in the hands of satellite designers but instead should be strongly influenced and guided by service, network and marketing experts. The total network and service design concept was born.

Upon reflection, the single most important outcome of the approach adopted was the early involvement of the end-users and the end-user service providers in an iterative process that proved invaluable. The message was very clear: "MSS is from the people and for the people". Such mass consultations usually tend to become cumbersome and ineffective unless properly orchestrated. For this reason, the formation of committees was limited to a minimum, most of the dialogue being on a one to one basis. Ultimate control remained with the program planners.

DOC's wide jurisdictional powers on telecommunications coupled with the prevailing Canadian communications policies were instrumental in the development and implementation of an integrated Business Development Plan that resulted in the resolution of all Canadian domestic issues in the short period of 3 years. In a contrast in styles, the MSAT development in the U.S. depended upon the respective jurisdictions of NASA, FCC and to a certain degree NTIA with a resulting much longer interval of completion. The inherent merits of the two styles not being the issue, it is undoubtedly clear that prompt Program implementation was the result of the ability to embrace all Program activities within a single entity.

While recalling the many bright points of the Program, one cannot but temper them with the recognition of things not so well accomplished. Domestically the area that is lagging the most is the emergence of a strong manufacturing base for ground segment equipment. For all the intensive research carried out by the DOC, for all the incentives and in contrast to the satellite manufacturers, this sector has been reluctant to commit to a long range portfolio development choosing to focus instead on short-term product development. As a result, the area most lagging today is that of network and user equipment development.

Finally, the handling of spectrum issues leaves a lot to be desired. Undoubtedly, the most frustrating and damaging to the Program development have been the years 1985-1987 during which spectrum availability was in serious jeopardy. Not only did user confidence sink to a low but in addition, this period was marked by the emergence of a number of competing systems and services.

Paradoxically, the total violation of the old and proven saying of not counting one's chickens before they are hatched may have proven to be the Program's silver lining! The development of MSS was based on the perceived availability of mobile spectrum at 800 MHz. But for this perception, it is highly unlikely that NASA and DOC would have embarked so vigorously on the MSAT Program development. In the absence of such commitment, it is questionable if the North American private sector would have been ready to commit to any form of cooperative program development, or if the Canadian and U.S. administrations would have pushed as hard as they did at WARC-MOB 87 or if the rest of the world would have been willing to take seriously any attempts to modify the existing mobile satellite allocations at 1.5 - 1.7 GHz. What could have been the greatest blunder in North American MSS development turned out to be a blessing in disguise.

But the spectrum lesson must remain with us because the future success and survival of MSS will undoubtedly depend on it. A lot was made in WARC 87 of the "great compromise" praised by so many of us that were there. Was it really a great compromise or was it that INMARSAT suddenly realized, in mid conference, that it had just as much to gain and perhaps more by agreeing to selective change? We may never really know! One thing is certain; we must never again allow MSS viability to be subject to uncertainty and last-minute compromises.

The Moment of Truth

For all the things that went right or wrong North American MSS implementation is only a few years away and yet several major issues need still to be addressed and resolved. Bilateral issues such as transborder service, licensing and type approval must be addressed; Difficult multilateral spectrum coordination deliberations must be completed; and above all the perception of a viable and expanding service offering must be maintained and enhanced through further mobile satellite allocations in 1992.

ACKNOWLEDGEMENTS

The author acknowledges the contribution of D.C. Buchanan, Director, MSAT Program, in suggesting and reviewing the paper and the numerous contributors to the success of the MSAT Program, without whom this paper would not have been possible.

REFERENCES

[1] DOC: March 1983 - MSAT Communications System Concept Document.

[2] DOC: May 1983 - MSAT Service Description.

[3] Woods Gordon: June 1984 - The Market for MSAT Services: Study to Determine Future Market for Mobile Satellite Services in Canada and the Benefits accruing to Users.

[4] DOC: 1987 - MSAT Phase B Final Report.

[5] Telecom Canada: December 1984 - MSAT Telco Opportunity Assessment.

[6] KVA Communications and Electronics: April 1985 - Study to Assess the Impact of MSAT on Radio Common Carriers.

[7] Telesat Canada: January 1985 - MSAT Phase B Commercial Viability Study.

[8] Woods Gordon: March 1985 - The Manufacturing Impact Study.

[9] Wescom Communications: April 1985 - Study to Evaluate the Quantitative Social Impacts and Additional User Benefits of MSAT.

[10] Econanalysis: May 1985 - The Overall Socio-Economic Analysis of the MSAT Program.

[11] Telesat Mobile Inc.: December 1985 - Revised MSAT Business Proposal.

Developments in Land Mobile Satellite Service in Europe

D A R Jayasuriya
Radiocommunications Agency
Waterloo Bridge House
Waterloo Road
London SE1 8UA
United Kingdom

ABSTRACT

The evolution of land mobile radio has reached a stage to benefit from satellite communications. The provision of a service on a pan-European basis makes the use of satellites a viable proposition. The paper describes the European position on both system and space segment aspects of the land mobile satellite service. Also, some of the functions of the European institutions, such as ETSI, CEPT and CEC, in establishing these services are identified in the paper.

INTRODUCTION

The need to develop land mobile satellite service (LMSS) in Europe has long been recognised by many European organisations. The studies and discussions have been continuing in various fora, notably within the European Space Agency (ESA). These studies have included market aspects, technical matters (including standards and frequency availability) and regulatory aspects concerning the operation of LMSS terminals on a pan-European basis.

The pan-European digital cellular radio system, GSM, is expected to begin operation in 1992. The other land mobile systems which operate in many countries on a national basis will continue to provide services. In addition to these the market needs identified for pan-European satellite services, for both data and voice, will be fulfilled by systems proposed by ESA, Inmarsat and Eutelsat. The low speed data systems include: ESA PRODAT, Inmarsat Standard-C and Eutelsat EUTELTRACS. The former two are L band systems whilst the latter operates at Ku band. Further plans are also afoot to provide voice services. To this end, in addition to Inmarsat Standard-M, Eutelsat has been examining the possibility of providing a pan-European (regional) service with the embarkation of an L band payload, currently being developed by ESA, on one of its Eutelsat II satellites.

In all cases mentioned above the utilisation of the frequency spectrum necessitates detailed study. The studies concerning the L band systems indicate the difficulties associated with frequency coordination. This with future traffic/market projections have shown the need for additional frequency spectrum for mobile satellite services and discussions are underway within CEPT on future frequency requirements for mobile

satellite services (MSS) in preparation for WARC 92.

The provision of a mobile service on a pan-European basis necessitates the establishment of a harmonised regulatory regime to enable the operation of mobile terminals throughout Europe. These issues are being considered by various European organisations, such as European Telecommunications Standards Institute (ETSI), Conference of European Postal and Telecommunications Administrations (CEPT) and the Commission of the European Communities (CEC).

EUROPEAN MARKET

Market studies conducted by various organisations suggest that in spite of the developments taking place to establish GSM system in Europe there is a sufficient demand to justify the establishment of mobile services, for both voice and data, by satellite. These studies also recognise the need for the proposed low speed data systems to take advantage of the "window of opportunity" that exists prior to the implementation of the GSM system. Nevertheless some level of system integration between European LMSS and GSM will provide a further advantage to the LMSS market by establishing it as a supplementary service to GSM especially in the areas where GSM service is patchy.

LOW SPEED DATA SYSTEMS

Low speed data systems have been identified as a key LMSS market for pan-European applications in various market studies. Over the last few years ESA and Inmarsat have carried out trials of their systems known as PRODAT and Standard-C respectively. Recently Eutelsat has entered the market with a system called EUTELTRACS which operates through Eutelsat's Ku band regional beams in Europe. It is also understood that the LOCSTAR radiodetermination satellite system is expected to offer a messaging service using its low speed data transmission capability.

Prodat

ESA, after carrying out a comprehensive study of system aspects, has successfully demonstrated the PRODAT bi-directional data transmission system to the European mobile market. The PRODAT system, designed to operate at L band, with omnidirectional antennas, at present utilises the Inmarsat space segment. The design of this system has been preceded by investigations into the problems associated with the propagation of signals to all three mobile environments, namely land, maritime and aeronautical. The system has benefited from ESA studies considerably, in that, it has demonstrated its ability to operate reliably in a highly variable land mobile radio channel where shadowing is a frequent event.

Inmarsat Standard-C

This system was initially designed for applications in the maritime service. Since then it has adapted for land mobile applications and successful demonstrations have been performed throughout Europe and beyond using Inmarsat global satellite network. This system also employs omnidirectional mobile antennas.

Eutelsat EUTELTRACS

This system has recently been

introduced into Europe and operates via Eutelsat space segment at Ku band primarily assigned to the Fixed Satellite Service (FSS). Following initial trials of the system carried out by European PTTs to ascertain the system performance, Eutelsat has now entered into the pre-operational demonstration phase for a six month period from the beginning of January 1990. Initial indications are that the system performs satisfactorily in a LMSS environment. This system employs a relatively narrow beamwidth mobile antenna coupled with spread spectrum transmissions to enable it to co-exist with other Ku band FSSs without causing harmful interference. The latter issue, the question of frequency compatibility, is currently under study in some European administrations.

Locstar

The LOCSTAR organisation in Europe has proposed the implementation of an RDSS system to offer, in addition to the radiodetermination service, a messaging service which utilises the system's low speed data transmission capability. The proposed system is to operate at L band where primary allocations exist in some of the European countries. It is clear, therefore, that successful frequency coordination on a pan-European basis is required for the implementation of the service.

VOICE SYSTEMS

Inmarsat has already initiated discussion throughout the world on voice systems in order to facilitate the development of the Standard-M system. The proposed system is at L band and uses directional antennas. Further discussions on European requirements for voice systems are expected to take place in the recently formed Mobile Satellite Group of ETSI, known as SES5.

SPACE SEGMENT

The space segment has an essential role to play in determining the type of service to be provided and the type of market to be addressed. The current discussions in Europe concern three types of space segment, namely Inmarsat, ESA/Eutelsat EMS and Eutelsat Ku band. The LOCSTAR system employs a dedicated satellite system hence it is not subject to further discussion in this section.

The type of LMSS services required for various markets can succinctly be described as "public" or "private" services. The public service entails the provision of communications with a connection to various public networks, such as PSTN or PSDN. On the other hand some markets rely on private services where a small earth station situated at customers' premises could address directly a closed user community.

Inmarsat

At present the only L band space segment available within Europe is that provided by Inmarsat. The frequencies of this space segment lie within the primary maritime mobile satellite service (MMSS) where low speed data transmissions can be provided for LMSS applications on a secondary basis. Once Inmarsat II generation satellites are in orbit the availability of frequencies will improve by nearly three fold and frequencies within primary LMSS bands become available. The proposed Inmarsat III satellites with spot

beam capability will allow the provision of enhanced services in the long term.

European PTTs have elected to provide their initial LMSS public low speed data service (Standard-C) using the Inmarsat space segment. The current generation of satellites and also the II generation satellites are ideally suited to enable European PTTs to offer Standard-C as a public service since existing or planned Coast Earth Stations (CES) infrastructure for MMSS Standard-C, which by definition have connections to public networks, need only minor modifications to adapt for land Standard-C. Inmarsat system can also be configured to provide broadcasts to a closed user group, as exemplified by Enhanced Group Call (EGC) service, transmitted via a CES.

ESA/Eutelsat EMS

The ESA, under its mobile satellite programme, is developing a specialised L band payload, called EMS, for European mobile satellite applications. EMS is expected to provide services to a wider European region initially concentrating on LMSS. These services will be able to take advantage of higher eirp provided by EMS payload to offer enhanced services. Discussions are currently underway between Eutelsat and ESA on the embarkation of this payload on one of Eutelsat's II series satellites.

The feeder links to the EMS payload are expected to be at Ku band. Therefore, EMS would be an ideal vehicle to offer private closed user services based on VSAT up link facilities, which can be located at customers' premises. In addition, of course, public services can be provided by establishing necessary interfaces with the public networks.

Eutelsat Ku band

Subject to the satisfactory outcome of the pre-operational trials and frequency compatibility studies, this service is expected to be available on a pan-European basis using Eutelsat's "Eurobeam" coverage and with connections to the public data networks provided at the Central Hub Station.

FREQUENCY SPECTRUM

With the development of many global and regional satellite services operating at L band the frequency coordination has become almost an insurmountable task. This coupled with certain traffic forecasts for MSS for the next decade has led to the suggestion that more spectrum is needed to cater for the predicted demand.

The inputs to the studies underway in CEPT, in preparation for the WARC 1992 conference, have identified this need for some additional spectrum allocations to MSS in the L band. This immediately raises the problem of identifying the part of the spectrum which could be allocated to these services. In addition, it is necessary to examine whether the MSS requirements, in particular those for LMSS, can be satisfied with regional allocations in contrast to the global allocations generally made for satellite services. Frequency sharing also plays a crucial role. If studies show that MSS systems require a near exclusive allocation then ways to see how this can be achieved need to be investigated. This in turn leads to the suggestion for a "staged allocation" where the nominated spectrum becomes available at a specified time in

the future, but in time to meet the demand, once the exisisting services have been relocated.

TRANSBORDER OPERATION

The provision of a land mobile service on a pan-European basis necessitates the establishment of a harmonised regulatory regime to enable the carriage and operation of mobile terminals throughout Europe. The authorisations for the operation of the mobile terminals are generally granted by the administrations who consider, amongst other things, the general type approval requirements of the terminals. These issues are currently under discussion in CEPT, ETSI and CEC.

CEPT and Circulation Card

CEPT has adopted the necessary provisions, in the form of a Recommendation, to facilitate the transborder operation of LMSS systems. This method, initially developed for Inmarsat Standard-C trials, consists of the issue of a "circulation card" which authorises the users to carry and operate the mobile terminals within the countries identified in it. However, the issue of the circulation card is conditional upon the administrations authorising/licensing the specific satellite service and the use of the associated mobile terminals. The latter issue, of course, is related to the type approval of equipment.

Recently Inmarsat, having noted the rapid development of portable mobile terminals, has initiated discussions among its member states for the accommodation of transborder use of Inmarsat terminals in member and non-member countries.

ETSI and Type Approval

As a part of the preparations for the Single Market in 1992 the European Community is developing a policy of testing, certification and inspection which aims to eliminate barriers to trade by encouraging mutual recognition of test results and certificates. The issues relating to telecommunications have been taken up by ETSI which plans to achieve this objective by setting European Telecommunications Standards (ETSs). The afore mentioned SES5 Committee of ETSI has the function of developing such harmonised standards for the mobile satellite equipment. The adoption of ETSs and the mutual recognition which results from that, would enable service operators to satisfy a condition of the circulation card.

European Commission

It is the policy of the CEC to assist the harmonisation effort in Europe to provide pan-European services to the benefit of manufacturers, operators and users. In this regard the CEC has been examining the question of MSS as a part of its studies into the role of communications satellites in Europe. The relevant issues are to be raised in a discussion document (Green Paper) which the CEC intends to publish later this year. It is expected that amongst the issues to be raised there will be policy initiatives on encouraging the development of European standards and achieving some level of system integration between the GSM and LMSS services.

CONCLUSIONS

LMSS has been identified as a viable market in Europe.

Several initiatives are under way to provide a number of systems operating both at L and Ku bands. These will fulfill a vital segment of the pan-European communication market.

Initial studies have indicated the need to allocate additional spectrum to all MSS to meet the expected demand. This coupled with the establishment of a harmonised regulatory regime should facilitate the operation of LMSS voice and data services on a pan-European basis. The work underway in ETSI, CEPT and CEC would enact procedures to achieve these goals.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the permission of the Director of Radio Technology, Radiocommunications Agency, DTI, to publish this paper and assistance provided by his colleagues.

The Radiocommunications Agency is an Executive Agency of the Department of Trade and Industry (DTI)

Implementation of Mobile Satellite Services in Developing Countries - The Mexican Experience

Alexis Reimers and Jorge Weitzner
Corporación Nacional de Radiodeterminación
Dante 26 BIS, México D.F. 11590, México
Phone: (905) 533-2784
FAX: (905) 533-2904

ABSTRACT

In the present paper, we present an analysis of the differences between Developing Countries (DCs) and Industrialized Countries (ICs) that concern Mobile Satellite Services (MSS) providers and regulators, and make a series of recommendations that may improve the odds for a successful implementation of MSS in DCs.

MSS MARKETS

At a time when Mobile Satellite Services (MSS) gain momentum all over the world, industry firms and regulators find themselves busy developing the terms and standards in which the services will operate in the future. Since the technology was, and is being originated in the Industrialized Countries (ICs), it is natural that their efforts focus on addressing the needs and requirements that ICs have for MSS.

As this process takes place, the industry participants should bear in mind that Developing Countries (DCs) represent an important potential market from which they could reap generous benefits. For this to happen, MSS providers must first understand the peculiar requirements and needs that DCs have from MSS, which differ significantly from the requirements of ICs.

At the DC's end, policy makers and regulators must appraise the far reaching effects that the adoption of MSS may have in their countries' economies. For this reason, they must formulate sensible regulations capable of promoting the investment required to provide MSS within a framework that considers the country's needs and priorities.

DIFFERENCES BETWEEN DCs AND ICs

For the introduction of MSS in DCs, industry firms and regulators must consider contextual and implementation factors that differ significantly from the ones they encounter in ICs.

Contextual Factors

i) Needs

Among country blocks, their geography and economic specialization determine the needs and applications for telecommunications. While ICs tend to specialize economically in services and the transfer of information, DCs specialize in manufacturing and in the exploitation of their natural resources across extensive geographical and deficiently communicated regions. While the former demand intensive communications able to transfer information at high speeds in order to enhance the quality of services, the later require links that support their industrial activities and the distribution of goods. One suitable service to fulfill these needs is the object of our discussions at this conference.

ii) Infrastructure

The infrastructure of DCs is characterized by a limited availability of the most elementary resources needed to provide MSS, such as satellite coverage, terrestrial links and trained engineers. It is also distinguished by big imbalances between private and public services, where on one hand we find private users with sophisticated networks for the transmission of data and voice while on the other we find deficiencies in public services such as the telephone and telegraph networks.

iii) Objectives and Strategies: Telecommunications as a Strategic Sector

One of the most significant factors that distinguish DCs are the objectives and strategies they follow, shaped in the form of their particular regulations. These are defined by the national priorities as perceived by the decision makers. While in the ICs the main objective has been the economic development of the nation, the goal in DCs has been to maintain socio-political harmony within a framework of economic development. This has been reached through a strong, centralized government and the designation of strategic enterprises or sectors, as is the case of telecommunications.

In the past, the reasons to designate telecommunications a strategic sector are mainly three: to have control over the basic infrastructure which is vital for the security of the country and the economy; to have a public service provided to all the population uniformly and equitably; and to provide the service with the lowest possible cost by achieving economies of scale in the investment and its operation. For the future, we foresee a trend towards decentralization as reflected in current global political and economical events.

iv) Legal Restrictions

The legal restrictions and regulations to foreign investment in strategic sectors in DCs are originated mainly for reasons of sovereignty, control and the creation of a national technological base. The restrictions vary considerably from country to country and deal with issues such as whether foreign firms can participate, in what degree and what kind of commitments are required from them -which include amount of investment, degree of local integration, technology transfer and reduced charge service to government institutions among others.

v) Underlying Economic Conditions

It is evident that DCs and more particularly Latinamerican countries, find themselves in a far reaching process of political, economical and social changes. At the same time, their economies are characterized by high inflation, unemployment, recession, a decreasing standard of living and a growing external debt. These problems are aggravated by a shortage of foreign currency due to credit restrictions imposed by lenders while many of their exports have been affected by lower prices in the international markets. Because of these factors, the relative cost of introducing MSS to DCs is much higher than the cost of introduction to ICs, as manifested by the comparison of the price paid for a monthly rental fee for the service in terms of the minimum wages paid in each country.

vi) Applications and Benefits

As in ICs, the primary application (for the foreseeable future) of MSS in DCs is in the area of logistics management for medium and long haul trucks. But in DCs, this sector of transportation plays a much greater role in the national economy given the lack of adequate, alternative transportation means as those employed in

ICs (such as railroads, airplanes and fluvial means). Unfortunately for DCs, the demand for transportation is not being fulfilled due to the lack of resources to provide and maintain the service while at the same time, a great part of their vehicle fleet works at low utilization levels because, among other reasons, the absence of proper communications required for the management of logistics. In México and until recently, these problems were compounded because of regulations that restricted the transit of transport vehicles to fixed routes. The de-regulation that was introduced in the past year pretends to correct this problem and encourage greater utilization of vehicles. The introduction of MSS could complement this effort by providing the communications infrastructure needed to coordinate logistics and improve productivity.

vii) Technological Culture

The technological culture of potential users of MSS in DC's is rather unsophisticated, considering that telephone communications and computers are marginally integrated into their logistics. Because of this, MSS represent a "quantum leap" to users in DC's, both in terms of the level of sophistication of the technology they are used to handle and in the opening of new opportunities due to the integration of MSS to their operations.

Implementation Factors

i) Process to Obtain Required Permits

The process to obtain the required permits to operate and exploit MSS in DC's is complex, because of the many parties involved in the decision making. Through the process, the authorities consider besides technical issues such as standards, band allocation and interference problems, other issues that concern the political, economical and social viability of the project: who is the service provider, its nationality, what experience does it have, how much power does it gain, what are the effects on different sectors of the economy, what is the cost to the country in terms of monetary reserves and so on.

ii) Configurations of the Systems

The configurations of MSS for DCs must consider that fundamentally, MSS represent a productivity tool to their economies, intensive in manufacturing and extractive industries, while in ICs, besides increasing productivity, the service enhances the quality of services thus playing an important effect in the competitiveness of firms in the ICs' service oriented economy.

iii) Understanding the Market

Understanding of the market, its needs, segmentation, leaders, attitudes towards technology, business culture, the degree of customization and integration of equipment, expected rates of failure that determine service and maintenance expenses, among others, are issues crucial to the success of the implementation, and singular to the idiosyncracy of DCs.

iv) Objections by Third Parties

The objections to the implementation of MSS in DC's can come from unexpected directions, mainly because of the threat that a new, far reaching technology such as MSS signifies to the market position of traditional leaders. Influential, well established firms involved in communications, transportation, truck production, transportation brokerage, etc. may oppose the introduction of a new productivity tool such as MSS that may adversely affect their hold on the market.

RECOMMENDATIONS TO MSS PROVIDERS

From the former analysis, it is clear that industry firms considering participation in

DCs must develop very focused and well planned strategies, much different from the ones devised for ICs. As these firms plan for significant investments to develop technology and infrastructure to provide the service, they must bear in mind that in order to achieve a successful penetration in DCs, they must provide technologies and standards that fulfill the markets' specific needs and are in accordance with the country's national objectives.

Specific recommendations concerning the entrance of MSS providers to DCs are the following:

i) Find a Strong Local Associate

Given the complexities and the restrictions that MSS providers face in their effort to enter DC's, the association with a strong local firm with a good record in dealing with the authorities is essential. MSS providers must also check the potential associate's capabilities in case technology transfer and local manufacturing is required. Most important, the firm must assess the associate's honesty and reputation because, in case of any dispute, local laws will naturally favor a local firm against a foreign one. There are several types of agreements that can be signed with a local associate: Commercialization Agreements, Royalty Agreements, Joint Ventures and Direct Investments. Each alternative has its pros and cons, and their analysis must consider the following issues:

- Type of association wanted or required by law.
- Percent of participation in venture, if allowed by regulations.
- Investment in terms of technology and cash.
- Investment in local infrastructure to provide service and to service and maintain equipment.
- Local integration of equipment.
- Return in terms of profits or royalties.
- Legal Restrictions.

ii) Understand Laws

Given that restrictions and regulations vary among industries and some may be negotiable, a good understanding of the laws and regulations, as well as of the idiosyncracies of the regulators is needed in order to reach the best possible conditions allowed by the law.

iii) Plan for Added Costs

MSS firms must plan for the added costs involved in providing the infrastructure needed for MSS in DCs, such as building terrestrial links and recruiting and training technical people.

iv) Plan for Technology Transfer

Given that DCs policies and laws condition the introduction of new technologies, such as MSS, to the provider transferring technology to the DC in varying degrees, MSS providers need to adopt a long term strategy that considers technology transfer along with proper protection of patents and processes.

v) Configure to the Market's Needs

When developing hardware to provide the service, industry firms should consider an open architecture flexible enough to accommodate operation using the locally available infrastructure. Thought should also be given to the fact that in DCs, the MSS short and medium term market comprehends mainly industrial users which have more need to data communications, as opposed to the smaller consumer segment which has better use for voice communications.

vi) Customize

Although the current globalization craze favors the standardization of products, it is clear that industrial products and services, such as MSS in DC's, require a high

degree of customization in order to provide a solution to the specific needs of each market and user. In this respect, the most important aspect to be considered is that, to succeed with the users, MSS must be designed to provide total solutions to specific problems rather than be presented as a sophisticated communications link.

vii) Educate the Market

In order to fulfill the expectations from MSS, a big effort must be spent in the education of a market which is usually reluctant to adopt new technologies. The training programs to operate the hardware should consider the language and the idiosyncracies of users.

RECOMMENDATIONS TO POLICY MAKERS AND REGULATORS IN DCs

For years, in ICs as well as in DCs, norms that regulate public services have lagged significantly behind the technological innovations that drive those services. The consequence of this has been a high cost in terms of the benefits that society could have derived from the services if the regulations were actual. This problem is more acute in DCs, where regulations must consider a much greater number of factors, as explained in previous paragraphs.

Given their need to improve the condition of their economies, DCs can not afford the luxury of the burdensome cost mentioned before. Some suggestions that may help formulate regulations to reap the expected benefits from MSS are the following:

i) Re-evaluate Strategic Sectors

It is unnecessary and quite costly for DCs to have all the value added telecommunications services centralized. Government should allow that these services be provided by national, private firms and focus instead in formulating and enforcing proper regulation to guide the exploitation of such services according to the national objectives and needs. The implementation of this policy will not jeopardize the country's sovereignty if proper regulation is formulated to control and supervise such service providers.

ii) Nurture a Technological Base

The nurturing of a technological base is one of the national priorities of DCs. It is required, among other reasons, in order to modernize and develop technology to better serve the particular needs of the country and to improve the terms of international exchange for the DC.

Regulators and policy makers must provide a solid ground to foster a real international technological cooperation. At the same time, policy makers must give impulse to the country's scientific and technological programs with the goal of reaching the levels where technology is not only adopted but also generated.

iii) Use Standards that Maximize Utilization of National Resources

The regulation of standards must contemplate the utilization of infrastructure already available in the country in order to maximize its utilization, minimize the outflow of precious foreign exchange and provide needed experience to qualified personnel to operate and develop technology in the future. Given that most of the market operates within the limits of the country, primary consideration should be given to standards that employ national links. For the market segment that requires greater coverage, as in the case of international flights for example, hybrid standards that allow international coordination of links should be employed.

iv) Define Regulation According to the Market's Needs

Regulators must research the needs and

requirements of the sectors that may adopt MSS. As a result, they will be able to formulate the proper regulations that can guarantee the greatest benefits from the service. The combination of properly regulated MSS together with the de-regulation of transportation, as in the case of México, will provide the solutions that DCs require in order to increase transportation productivity and efficiency, achieving thus the desired effects in the overall economy.

CONCLUSIONS

DCs represent an attractive potential market for MSS participants while at the same time, MSS technology could render great benefits to the DCs economies. For this to happen, MSS technologies must be flexible enough to adapt to the specific needs of the countries and employ their available infrastructure.

In order to succeed in these markets, MSS providers must understand the underlying contextual and implementation issues that affect them, and choose a suitable local firm through which the main hurdles of implementation can be overcome.

Regulators at DCs should keep in mind that a new technology such as MSS offers the opportunity to formulate new regulations to develop a local technological base through technology transfer.

Quoting Carlo de Benedetti who said "technological innovation has progressed like a series of waves clearly defined. Each wave has introduced new series of technologies that in turn, have fused with their applications causing a cascade of innovations". DCs have in MSS a good opportunity to create their own cascade of innovations whose effects will benefit not only the DCs, but the technology providers as well.

REFERENCES

1. Bennedeti, C. 1987. The new role of Europe in the World Market. New York, USA.

2. Capital, Mercados Financieros. 1990. Monthly, February, numb. 28. México D.F., México.

3. COSPAR/COSTED/UN. 1982. Conclusions from COSPAR/COSTED/UN symposium on the role and impact of space research in Developing Countries, COSPAR/IAF/UNISPACE '82 forum, Space Resolutions Vol 3 No 7. Viena, Austria.

4. Fondo de Cultura Economica. 1986. Las actividades espaciales en México, una revisión crítica. México D.F., México.

5. Fondo de Cultura Económica. 1986. Memoria del Primer Seminario Latinoamericano de Conversión Industrial, Exposición, Síntesis y Perspectivas. México D.F., México.

6. Luisignan B. B. 1981. Technology Transfer. Document for the Communication Satellites Planning Center, Stanford University. Stanford CA, USA.

7. Melody, William H. 1986. Telecommunications - policy directions for the technology and information services. Economic and Social Research Council. London, England.

8. Mobile Satellite Conference, NASA-JPL. May 3-5, 1988. Proceedings. USA.

9. Onward and Upward Conference organized by Phillips Publishing. 1989. Proceedings. Washington D.C., USA

10. UIT. 1988. Benefits of Telecommunications for the Transportation sector in Developing Countries. Geneve, Switzerland.

An Upward Compatible Spectrum Sharing Architecture for Existing, Actively Planned and Emerging Mobile Satellite Systems

Bahman Azarbar
Telesat Canada
1601 Telesat Court, Gloucester, Ontario, K1B 5P4 Canada
Phone: 613-748-0123

ABSTRACT

Existing and actively planned mobile satellite systems are competing for a viable share of the spectrum allocated by the ITU to the satellite based mobile services in the 1.5/1.6 GHz range. The very limited amount of spectrum available worldwide and the sheer number of existing and planned mobile satellite systems as known to date dictate adoption of an architecture which will maximize sharing possibilities. A viable sharing architecture must recognize the operational needs and limitations of the existing systems while accommodating the reasonable demands for spectrum of the planned systems. Furthermore, recognizing the right of access of the future systems as they will emerge in time, the adopted architecture must allow for additional growth and be amenable to orderly introduction of future systems. This paper describes an attempt toward devising such a sharing architecture. A specific example of the application of the basic concept to the existing and planned mobile satellite systems as known to date is also discussed.

INTRODUCTION

The 1987 World Administrative Radio Conference on Mobile Services (WARC-MOB-87) redefined the spectrum allocation to mobile satellite systems by explicitly recognizing the need for introduction of land mobile service. Noting the limited extent of raw available spectrum in the 1.5/1.6 GHz range, the conference further requested the ITU to convene a limited Allocation Conference in 1992 to seek additional spectrum.

While it is comforting for mobile service interest groups to know that the ITU will do its utmost to find ways of providing additional growth capacity for future years, it is also well recognized that irrespective of the degree of success of WARC-92 in securing a sizable growth capacity, the most suitable spectrum for truly mobile applications is likely to stay below 3 GHz. This, in conjunction with what we know to date on the nature and extent of worldwide utilization indicate that the spectrum in this range will continue to remain relatively in short supply and by necessity *must therefore be utilized prudently and shared extensively.*

CHARACTERIZING FEATURES OF MOBILE SATELLITE SYSTEMS

Unlike their Fixed Satellite Service counterparts, the emerging mobile satellite systems are characterized by user terminals which by the very nature of the envisaged mobile services cannot afford the ability to discriminate between satellites at small to medium spacings in the geostationary orbit. Consequently, spectrum reuse through orbit separation is not likely to become a widespread and practical sharing tool in the entry period of these systems. As a result, the most effective means of reusing the limited raw spectrum over the next decade or so will most likely remain to be through isolation by geographical separation of the service areas.

There are two practical considerations, however, which tend to inhibit the full potential of a satellite cellular structure in reusing the spectrum. Analogous to its terrestrial counterpart, generation of small cells (satellite beams) requires a larger overall infrastructure which in satellite terms translates into a larger antenna assembly and repeater complex. This in turn requires a minimum user density per beam area to justify the associated cost of a given size beam. Furthermore, assuming that the potential market is sufficient enough in the long

term to support a given beam size, the economic realities of the first generation systems in terms of the speed with which they can build up a subscriber base are likely to further limit the ability of the operators to economically justify such a beam size. In view of this, the elemental cell size which defines the relative ability of the spacecraft antenna in concentrating its transmit capability and receive susceptibility is not likely to be appreciably smaller than 2.5 degrees in satellite coordinates for mobile satellite systems to be implemented over the next decade. This practical consideration is demonstrated[1] to approximately define the upper bound achievable reuse factor over the land mass to be:

REGION	REUSE FACTOR
North America	1.7
South America	2.0
Africa/Europe	4.0
Asia/Australia	4.0
TOTAL	**11.7**

The total reuse factor stated above merely suggests that, if 30 MHz of raw spectrum were made available, 350 MHz would be close to the upper bound worldwide aggregate capacity that can be expected to be realized for sharing amongst the emerging systems which are still at their infancy stage.

To better appreciate this potential capacity, it is worthwhile to note that for mobile satellite systems, the main applications are thin route voice and data messaging which by nature are not spectrum thirsty when spectrally efficient analogue and digital modulations are employed. In fact with any stretch of imagination, it is hard to conceive a realistic scenario in which power constraints of the largest available spacecraft bus today for deployment of commercial payload (3-4 kW DC range) will mandate a single satellite to demand more than 10 MHz of usable spectrum. This 10 MHz is considered to be enough to support a subscriber base per satellite in excess of 200,000 voice users, an order of magnitude larger data and messaging users or a proportion of both. To summarize, , given a 30 MHz of raw spectrum, the first consideration imposed by economic realities of early entry systems defines the upper bound to the capacity to be 350 MHz which is considered to be sufficient to respond to the realistic needs of up to 40 large sized start-up satellites, or alternately, 70 medium sized satellites of Delta class to serve the worldwide market. While this potential capacity by no means is trivial for start-up systems over the next decade or so, it could conceivably be further increased if coverage over the oceans were also added to the assumed aggregate service area.

The second practical consideration relates to the non uniformity of the elemental beam size from system to system in a real life situation. To better appreciate the nature of the problem, Figure 1 is intended to depict the beam sizes of the planned mobile satellite systems as have been filed officially with the ITU by various administration and international organizations. The smallest beam dimension which defines how fast the beam rolls off from the intended service area varies from 2.5 degrees for the Canadian MSAT on one extreme to approximately 18 degrees at the other extreme characterizing a global beam.

This gross disparity, does significantly reduce the potential capacities from the figures derived earlier for the first generation systems. The larger the degree of non uniformity, the smaller becomes the potential capacity. This reduction, however, is partially compensated by the fact that, in general, the larger the beam size, the smaller is the resulting aggregate EIRP for a given power level which in turn implies that less spectrum is needed to fully utilize the satellite's on board power.

While non uniform distribution of beam size reduces the ability of a satellite cellular structure to form an efficient reuse pattern, it is the existence of the global beams which drastically affect the worldwide sharing situation. Global beam satellites by their very nature offer no reuse capability through beam isolation. This means that unless the use of large

mobile antennas in conjunction with large orbit spacings become the prevalent

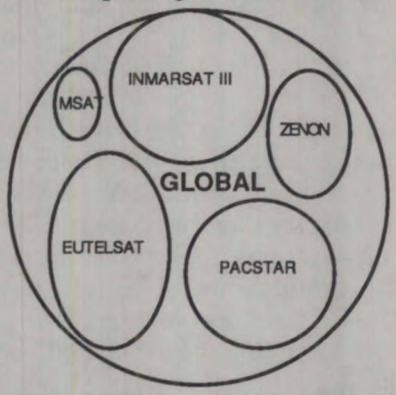


Fig 1. RELATIVE BEAM SIZES FOR PLANNED MOBILE SATELLITE SYSTEMS.

operating mode; a scenario which violates the very foundation of the business cases put forward for emerging mobile systems, very little reuse will be feasible in overlapping beam areas of global versus global or global versus spot beam systems. The coordination problem is further aggravated by the large gain difference between spot and global beams which tends to to lead to non homogeneity and incompatibility of traffic on these two classes of mobile satellite systems. While it is conceivable that global beams will eventually reduce in number due to competitive market forces and, as a logical consequence, in demand for spectrum, global systems will remain a fact of life for the next 10-15 for certain applications and as such will have to be accommodated. Indeed, the only existing mobile satellite systems today, although very modest in power capability and spectrum requirements, are solely characterized by global beams.

MEANS OF COEXISTENCE

Frequency coordination problems of spot beam based systems are generally localized and the task is expected to be manageable as long as the beam sizes are not exceptionally large and grossly non uniform. The variation observed in the beam sizes of the planned systems to date may become more subtle in time as they move from conceptual design stage. It is believed that when time for implementation of the actual hardware arrives, economics will most likely favor the national and regional systems to adopt small to moderate size beams; a position primarily driven by the desire to conserve the satellite's expensive power over the intended service area while maximizing reuse capability and minimizing the foreseen international coordination difficulties. Consequently, the most complex issue in coordination of mobile satellite systems will become finding means of coexistence between global and spot beam based systems.

Based on extensive experience accumulated by Canada over the past three years in relation with frequency/orbit coordination of the Canadian MSAT, four technical tools have been identified as the major facilitators of the process. Depending on the characteristics of the systems under coordination and their associated traffic, any or a combination of these options could be resorted to in order to maximize sharing of the limited spectrum. While all of these technical options are graphically described in Figures 2 to 5, by far the most effective tool for spectrum sharing is Option 1 or simply stated; geographical separation of service areas by means of *Global Radio Horizons* and sharing in an East/West complementary fashion with spot beams.

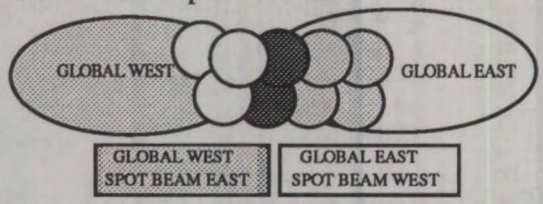


Fig 2. OPTION 1; USE OF RADIO HORIZON.

This option draws upon an organized pattern of spectrum usage to allow sharing and reuse of the same spectrum by both global beam and spot beam systems. The spectrum for spot beams happening to fall in the overlap areas of the global beams indicated by a darker shade in Figure 1 need to be sought by other technical means.

Option 2 as depicted in Figure 3 is intended to usefully employ the intercarrier spectrum which has been left unused by a given operating entity due to such operational considerations as intra-system interference consideration

and/or limitations imposed by the existing ground segment infrastructure.

Global Traffic

|| Spot Beam Traffic

Fig 3. OPTION 2; FREQUENCY INTERLEAVING.

Exploitation of this option, however, mandates careful scrutiny of the uplink inter-system transponder loading and/or power robbing that could result and become excessively large if not attended to. To control this loading, the location of transmiting mobiles, if available to the Network Control Center of the interfering system, could be advantageously used to assign non overlapping uplink carriers for the mobiles operating in the areas close to the spot beams of the victim system.

Option 3 utilizes inter-transponder guard bands built into the hardware design of a given satellite constellation as shown in Figure 4. These guard bands, while normally small in the amount of spectrum, since are not used by a given global network, could be reused several fold by spot beam systems over a relatively large service area.

Global Traffic

||| Spot Beam Traffic

Fig 4. OPTION 3; USAGE OF GUARD BANDS.

Option 4 as depicted in Figure 5, is merely brute force spectrum segmentation approach in the frequency

Global Traffic

||| Spot Beam Traffic

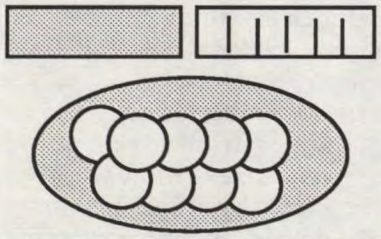


Fig 5. OPTION 4; SPECTRUM SEGMENTATION FOR MUTUALLY INCLUSIVE COVERAGES

coordination process when other options are not viable due to a multitude of technical and operational reasons.

DESIRED SHARING FEATURES

In order to make geographical sharing possible (Option 1), an organized pattern of spectrum usage need to be adopted to maximize sharing and reuse of the same spectrum by both global beam and spot beam based systems. Furthermore, for situations where the only viable solution proves to be brute force spectrum segmentation (Option 4), such an organized pattern of usage is needed even more in order to ensure that coexistence between such grossly incompatible systems will not lead to inefficient sharing scenarios. Moreover, the method should be robust enough not to mandate extensive inter-system traffic coordination on a case by case basis during the early phase that the new mobile satellite systems are moving through their infancy stage and have yet to acquire the operational experience and master the art of coexistence in a shared environment. The assumed sharing architecture should also be amenable to orderly introduction of new entries as they emerge in time; irrespective of whether they have global or spot beam configuration, without disruption of the existing and replacement systems or requiring major reorientation of the actively planned systems. In other words, the adopted strategy should be sound enough not to fall apart like a house of cards when new systems emerge in time.

A PLAUSIBLE USAGE ARCHITECTURE

With the above objectives, we proceed now to define a possible spectrum sharing architecture. Without loss of generality, we concentrate on defining the approach using the 10 MHz spectrum allocated to Aeronautical Mobile Satellite Service (AMSS). The application to the other components of the service such as Land Mobile Satellite Service (LMSS) or Maritime Mobile Satellite Service (MMSS) is similar with differences primarily stemming from the nature and extent of the existing systems or the specifics of the planned systems in each part of the band.

With the exception of ETS V experimental project with limited life expectancy, the operational or existing satellite hardware are those of USSR and INMARSAT Organization. These two entities also happen to have the most extensive plans for satellite based mobile networks around the globe and, with the exception of the planned INMARSAT III, are all characterized solely by global beams. With this preamble, it is logical to start first by devising an architecture which organizes the spectrum usage around these systems. Since INMARSAT and USSR have global networks with extensive concentration in the Atlantic and Pacific regions with a geostationary orbit distribution independent from each other, their usage can be organized over five distinct blocks of spectrum, namely:

o **POR** block to be shared by global beams over Pacific Ocean Region with horizon lines crossing the North America service area. The same spectrum can be used and reused by spot beam based systems in a cellular structure over eastern North America, entire South America, Europe, and Africa. It could further be reused by spot beam systems over western North America and Pacific region in conjunction with a limited quantity of mobile terminal antennas which could provide sufficient isolation towards POR global networks with orbit spacings greater than 30 degrees and for certain traffic combinations which would allow such sharing. To improve sharing possibilities by spot beam systems, POR block can be subdivided into POR1 and POR2W, with INMARSAT using POR1 and USSR using POR2W for global coverage in the Pacific region.

o **AOR** block to be used by global beams over Atlantic Ocean Region with horizon lines limited to mid North America. The same spectrum can be reused by spot beam based systems in a cellular structure over western North America. Like POR case, it could further be used by spot beam systems elsewhere in conjunction with mobile terminal antennas which could provide sufficient isolation towards global networks with orbit spacings greater than 30 degrees for certain traffic combinations. To improve sharing possibilities by spot beam systems, AOR block can be subdivided into AOR1E and AOR2, with INMARSAT using AOR1E and USSR using AOR2 for global coverage in the Atlantic region.

o **AOR W** block is to be used by global beam satellites over the Atlantic having full North/South American visibility. It can also be used by systems having beams over eastern Africa, Asia, Pacific region and Australia. The same spectrum could also be used elsewhere by spot beam systems in conjunction with mobile terminal antennas offering sufficient intersystem isolation.

o **POR E** block is to be used by global beam satellites over the Pacific having full North/South American visibility. It can further be used by systems having beams over eastern Africa, Asia, western Pacific and Australia. The same spectrum could also be used elsewhere by spot beam systems in conjunction with mobile terminal antennas offering sufficient intersystem isolation.

o **IOR** block to be used by global beams over Indian Ocean Region. In order to maximize sharing possibilities for spot beam based systems, this block should further be subdivided into IOR1 and IOR2. Global satellites in the range of 60-85 degrees east longitude should preferably use IOR1 and global satellites within 85-100 degrees east longitude should use IOR2. The same spectrum can be shared by spot beam based systems with service areas primarily in western Europe, western Africa, Atlantic region, North and South Americas and east Pacific. As before, application of mobile terminals with directive antennas is permitted wherever sharing is feasible

o **SP** block is to be solely used by spot beam based systems around the world to effectively deal with their localized coordination problems. These interface problems arise from the non uniformity of spot beam sizes of various systems or are prompted by spectrum requirement for spot beams used as entry port by international systems connecting the continents. Furthermore, this spectrum in

needed to feed the spot beams which fall in the beam cross-over area of two complementary global beams configured in an east/west fashion.

In line with these guidelines, Figure 6 depicts how this basic architecture could be applied towards accommodation of mobile satellite systems as known to date. The systems above the top horizontal line are the global satellites that need to share the spectrum within the block by whatever means seen feasible. The systems immediately below this line are spot beam systems which, transparent to the global systems cited in each block, can use and reuse the same spectrum with minimal coordination requirement amongst themselves in interface areas. The regions identified below the bottom horizontal line represent additional growth capability of the proposed architecture. In the case of INMARSAT II, AOR1W is also labeled as IOR as a transitional step necessitated by the degree of accessibility of AMSS band built into the hardware design of INM II satellites.

The compatible usage in North America is illustrated by means of beam patterns arranged under the relevant blocks of frequencies. These beams approximately demonstrate the North American coverage pattern of interest to date over Canada and USA including Alaska, Hawaii and Puerto Rico. The relative utility of blocks SP1-SP4 in North America is largely dependent upon the degree that these blocks of frequencies will be required to accommodate the needs of spot beams of international systems used as entry port in the coastal areas of North America.

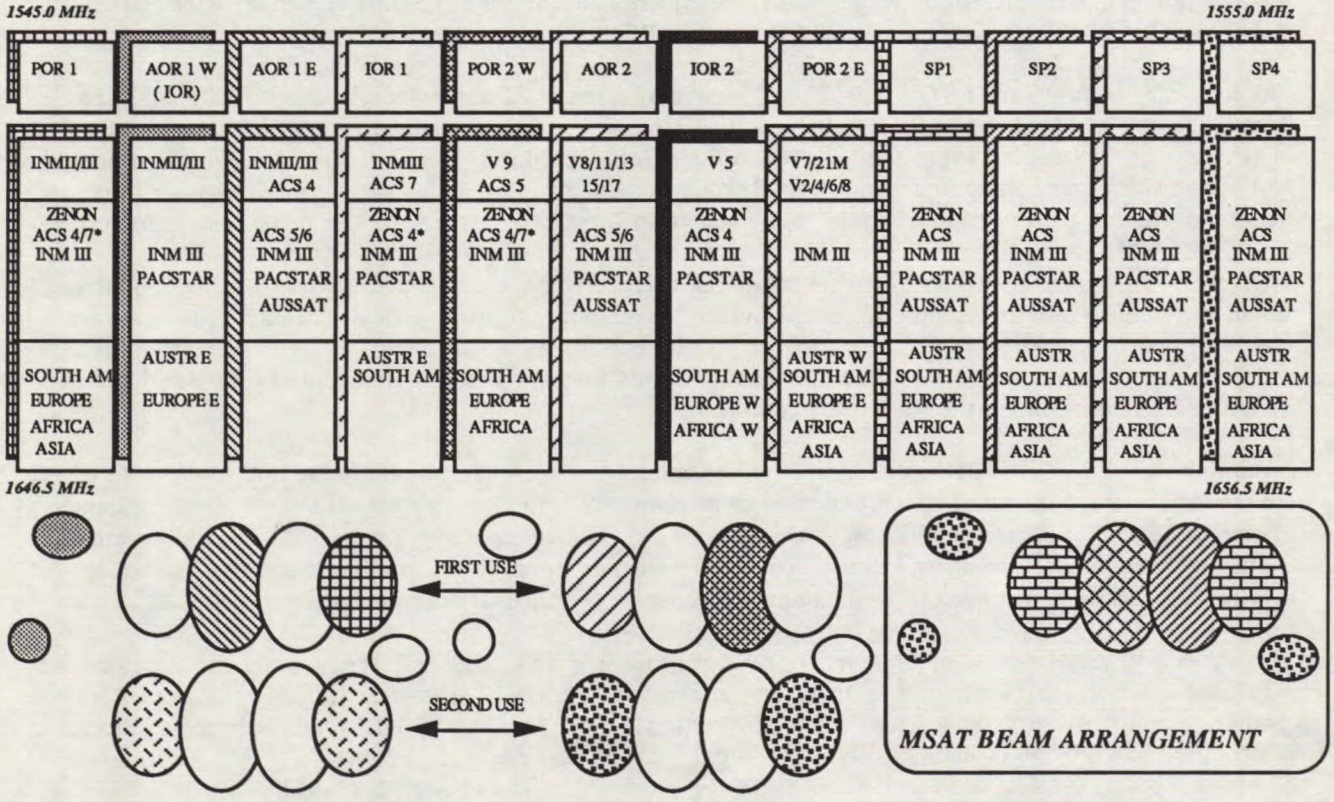


Fig 6. PROPOSED SPECTRUM SHARING ARCHITECTURE TO ACCOMMODATE EXISTING, ACTIVELY PLANNED AND FUTURE MOBILE SATELLITE SYSTEMS AROUND THE WORLD.

REFERENCES

1. Information Paper - Frequency Reuse Considerations in Sharing Common Frequency Allocations by Various Mobile Satellite Services. WARC-MOB-87, ITU, Geneva; Sept-Oct 1987, Doc. No. 56, Canada.
2. Frequency Reuse Using Fixed Spot Beams. Annex I to Report AF/8 on Intersystem Sharing... CCIR Doc. 8/256-E, July 30, 1988.

Mobile Satellite Users and the ITU 1992 World Administrative Radio Conference - Planning for the Future

Lawrence M. Palmer
U.S. Dept. of Commerce
Room 4701
Washington DC 20230
United States
Tel: (202) 377-1304 Fax: (202) 377-1635

SYNOPSIS

The ITU has scheduled a World Administrative Radio Conference to be held in the first quarter of 1992 in Spain to deal with allocation matters. While that conference will address only those items that are part of its mandate, it is anticipated that a review of the mobile-satellite frequency allocations at 1.5/1.6 Ghz will be made, particularly in light of Resolution No. 208 of the 1987 Mobile WARC.

The frequency bands at 1.5/1.6 Ghz have been reviewed by several ITU conferences in the past 20 years. Changes were made in the allocations as the bands lay fallow. Changes were also made at the 1987 WARC, not because of non-use, but because of expressed need by an operational satellite system, because of the applications that had been filed in the United States and Canada, and because of the planned activity underway on three continents to proceed with development and implementation of satellite systems to service mobile customers.

The situation has now changed as systems have been licensed in North America. Testing is being conducted in Japan and Europe. Interim systems are being planned. In certain instances, specific operational systems are underway. Full scale operational systems are just around the corner. As we all know, however, domestic allocation decisions have been made in the United States and Canada that are different from the ITU international table of frequency allocations in the Radio Regulations.

Assuming that the 1992 WARC will address the allocations at 1.5/1.6 Ghz, what are the options open to users? Do the international spectrum allocations need to be changed? And if so, in what fashion? Will the proponents of change be able to influence their respective national positions to reflect such a need? What options are open to the licensees? To the operational users? What about the new applications that have been filed? What are the regulatory issues that will need to be examined in order to permit international systems?

This paper will give a review of the allocations history leading up to the 1987 WARC, and the decisions that were made there. It will examine the different frequency allocations that now exist. Operational uses will be examined. Emphasis will be given showing the options, both from a regulatory and frequency allocation perspective, that may be available to the various parties. The paper will then conclude with a presentation of how well the options may succeed if those views are presented to the ITU 1992 WARC.

The full text of this paper was not available at press time

Mobile Satellite Regulation in the United States

Lon C. Levin, Walter H. Sonnenfeldt
Gurman, Kurtis, Blask & Freedman Chartered
1400 16th Street, N.W., Suite 500
Washington, D.C. 20036 USA
Phone: 202-328-8200
FAX: 202-462-1784

ABSTRACT

During the last decade, the U.S. Federal Communications Commission ("FCC") has developed the regulatory structure for the provision of mobile services via satellite. In May 1989, the FCC awarded American Mobile Satellite Corporation ("AMSC") a license to provide the full range of domestic mobile satellite services in the United States. At that time, the FCC reaffirmed the U.S. mobile satellite industry structure and spectrum allocations that had been adopted previously. Also in May 1989, the FCC authorized the Communications Satellite Corporation ("COMSAT"), the U.S. Signatory to Inmarsat, to provide international aeronautical satellite service via the Inmarsat system. Earlier in 1989, the FCC permitted the use of Ku-band satellites to provide messaging and tracking services. In the mid 80's, the FCC established the Radiodetermination Satellite Service and awarded licenses. Among the mobile satellite matters currently facing the FCC are whether additional spectrum should be allocated for domestic "generic" mobile satellite services, the regulatory structure for the provision of mobile satellite service on an interim basis before AMSC launches its dedicated satellites, and whether to authorize a low-earth orbit satellite system to provide mobile data service.

INTRODUCTION

The past decade has seen an explosion in the use of mobile communications. New and improved technologies have led to rapid growth in paging, cellular telephone, and private mobile radio services. Additional new technologies are continually being proposed. One of the most promising means by which to provide mobile communications is through the use of satellites. Recognizing the potential of satellite technology, the FCC has devoted considerable resources toward the development of a regulatory structure to assure the efficient and effective introduction of mobile satellite services in the U.S. This paper provides an overview of the current status of mobile satellite regulation in the United States.

U.S. DOMESTIC MOBILE SATELLITE SERVICE

In May 1989, culminating a proceeding that began with a Petition for Rulemaking filed by the National Aeronautics and Space Administration ("NASA") in November 1982, the FCC authorized AMSC to construct, launch, and operate the U.S. domestic mobile satellite system consisting of

three satellites using L-band frequencies for mobile links and Ku-band frequencies for feeder links. The first satellite is scheduled to be launched in 1993. The orbital locations assigned to AMSC are 101° W.L. for the central satellite, 62° W.L. for the eastern satellite, and 139° W.L. for the western satellite. At the same time, the FCC reaffirmed its domestic L-band mobile satellite allocations and industry structure regulations. These orders are currently under review before the U.S. Court of Appeals for the District of Columbia Circuit.

Mobile Satellite Spectrum Allocations

The FCC allocated 28 MHz of L-band spectrum in the bands 1545.0-1559.0 and 1646.5-1660.5 MHz for use by the U.S. mobile satellite service system. Rather than adopting a rigid spectrum segmentation plan, the Commission devised an allocation structure that permits all mobile satellite services to be provided across 27 MHz of the allocation, while assuring that aeronautical safety ("AMSS(R)") traffic can enjoy additional protection relative to other services. Due to sharing constraints with Radio Astronomy, the remaining 1 MHz is limited to aviation safety and certain one-way services.

The U.S. allocation combines all non-AMSS(R) mobile satellite services (land, maritime, and non-safety related aeronautical communications) under the designation 'MSS'. Eighteen MHz of the allocation (two 9 MHz segments at 1549.5-1558.5 MHz and 1651.0-1660.0 MHz) are shared between MSS and AMSS(R) on a co-primary basis, with a footnote providing for priority of AMSS(R) over MSS. In 9 MHz (two 4.5 MHz segments at 1545.0-1549.5 MHz and 1646.5-1651.0 MHz), MSS is secondary to AMSS(R). With regard to the remaining 1 MHz, one 0.5 MHz segment is allocated to AMSS(R) on a primary basis (1558.5-1559.0 MHz) and one 0.5 MHz segment is allocated to AMSS(R) and Radio Astronomy on a co-primary basis (1660.0-1660.5 MHz). The 0.5 MHz not used by Radio Astronomy may be used by AMSC for one-way dispatch MSS on a non-interference basis relative to all other users of that band. The entire 1 MHz may be used for AMSS(R) communications provided that Radio Astronomy does not suffer harmful interference.

The U.S. domestic allocation is consistent with international allocations. In pertinent part, both allocations provide the same status for AMSS(R) and land mobile satellite service.

The Commission allocated 200 MHz of Ku-band for feeder link use to each of the three satellites. The central satellite at 101° W.L. is allocated 200 MHz of the 11/13 GHz band. The satellites located at 62° W.L. and 139° W.L. are allocated 200 MHz of the 12/14 GHz band.

MSS Expansion Band Rulemaking. In February 1990, the Commission adopted a Notice of Proposed Rulemaking proposing to reallocate the bands 1530.0-1544.0 MHz and 1626.5-1645.5 MHz for domestic generic mobile satellite services. These bands are currently allocated to the maritime mobile satellite service. The Commission proposes that maritime safety services be afforded real time preemptive priority in the MSS expansion band. Applications seeking permanent authority to use these bands for domestic service

are being held in abeyance pending completion of the rulemaking.

MSS Industry Structure

Due to technical and economic considerations, the FCC decided to license one system (AMSC) at this time to provide the full range of land, maritime, and aeronautical services. The license is based on the FCC's finding that a single satellite system providing all services is currently the best means by which to "ensure efficient use of spectrum, promote safety, and introduce new services to the public in a timely manner." In particular, the FCC concluded that the efficiencies inherent in a single system will help assure that aviation safety services will be made available soon. Accordingly, the FCC authorized AMSC to be the MSS and AMSS(R) licensee.

AMSC will provide space segment on a common carrier basis. AMSC will operate as a carrier's carrier, providing open access to carriers and end users. AMSC's ground segment will be authorized separately. Earth stations accessing the system will be licensed individually. Mobile units will be authorized under blanket licenses, with the exception of aeronautical mobile earth terminals, which may be licensed individually.

Services & Coverage Area

AMSC is licensed to provide the full range of land, maritime and aeronautical services including two-way voice, two-way voice dispatch, two-way mobile data, air traffic control, airline operational and management communications, and aeronautical passenger correspondence. Fixed and transportable services may be provided on a non-interference basis to segments of the population where few alternatives exist. AMSC is required to provide coverage of the entire U.S. domestic market including all 50 states, Puerto Rico, the Virgin Islands, and U.S. coastal areas up to 200 miles offshore, except for those waters that are part of the territory of another country. In recognition of the relationship between the AMSC and Telesat Mobile, Inc., systems for mutual back-up and restoration capability, the FCC authorized AMSC to construct its satellites to cover Canada. AMSC may also construct its satellites to cover Mexico. Authority to operate in Canada and Mexico must be obtained by separate application.

Aeronautical Matters. By its authorization, AMSC is required to accord priority and real-time preemptive access to AMSS(R) communications throughout the entire assigned bandwidth. AMSC also must develop arrangements to "hand-off" aeronautical traffic to other MSS systems, such as the Canadian and Inmarsat systems.

Additionally, the FCC expects that aeronautical feeder link stations and mobile terminals will have certain unique characteristics to meet aviation safety operational requirements. The FCC may require special licensing procedures or application requirements for these facilities. All mobile terminals to be used on aircraft, including those for non-AMSS(R) communications, must be type accepted and licensed under the FCC's Rules governing aviation communications. The FCC expects the Federal Aviation Administration ("FAA"), which is the U.S. entity responsible for aviation safety, to be involved

directly in the development of standards and practices in order to assure that aviation safety satellite services will be of the highest integrity.

Interim Service

The FCC is currently considering the regulatory structure under which interim mobile satellite service can be provided in the U.S. until AMSC dedicated facilities are in operation. AMSC and others have proposed the use of Inmarsat space segment for the interim period. Inmarsat space segment is attractive because it operates on spectrum adjacent to the spectrum assigned to AMSC, thereby enabling users to transition smoothly from the Inmarsat system to the AMSC system.

USE OF INMARSAT IN THE UNITED STATES

Also in May 1989, the FCC established policies for the provision of international aeronautical services in the U.S. via the Inmarsat system. The Commission determined that COMSAT, the U.S. Signatory to Inmarsat, will be the sole U.S. provider of Inmarsat space segment for aeronautical services. In addition, the FCC decided that aeronautical services provided via Inmarsat to aircraft over the U.S., its territories, and adjacent coastal ocean areas could be offered on a permanent basis only to aircraft in international flight. The FCC defines international flights as those between the U.S. and foreign points (and vice versa), and those flying over the U.S. between two foreign points. COMSAT and AMSC are required to develop hand-off procedures for transferring aeronautical traffic between the Inmarsat and AMSC systems. The FCC has received two filings requesting reconsideration of the geographical scope of Inmarsat service in the U.S.

OTHER SATELLITE-BASED MOBILE SERVICES

The FCC has licensed other satellite-based mobile communications service providers. In February 1989, Qualcomm, Inc., was authorized to operate mobile terminals to provide messaging and tracking services using existing fixed satellite service space segment at Ku-band. In August 1986, Geostar Corporation was authorized to construct, launch and operate a dedicated Radiodetermination Satellite Service ("RDSS") system. Geostar has also obtained authority to provide interim RDSS service until its dedicated RDSS satellites become fully operational.

Orbital Communications Corporation ("Orbcomm") has filed a Petition for Rulemaking seeking to establish an allocation in the VHF/UHF bands for a low speed mobile data and tracking service using low-earth orbit satellites. Orbcomm has also filed an application to construct, launch, and operate a twenty satellite low-earth orbit system in the proposed allocation. The FCC has invited public comment on these filings.

REFERENCES

1. **AMSC Authorization Order**, 4 FCC Rcd 6041 (1989).

2. **Spectrum Allocation Reconsideration Order**, 4 FCC Rcd 6016 (1989).

3. **MSS Structure Reconsideration Order**, 4 FCC Rcd 6029 (1989).

4. **MSS Expansion Band Notice of Proposed Rulemaking**, General Docket No. 90-56, FCC 90-63 (released March 5, 1990).

5. **Inmarsat International Aeronautical Services Order**, 4 FCC Rcd 6072 (1989).

6. **Qualcomm Authorization Order**, 4 FCC Rcd 1543 (1989).

7. **RDSS Allocation Order**, 50 Fed. Reg. 39101 (September 27, 1985).

8. **RDSS Structure Order**, 104 FCC 2d 650 (1986).

9. **Geostar Space Segment Authorization Order**, 60 RR 2d 1725 (1986); FCC Mimeo 6144 (1986).

Review of Canadian Mobile Satellite Systems Institutional Arrangements Policy

David Gilvary
Telecommunications Policy Branch
Department of Communications
300 Slater St., Ottawa, Ontario
Canada, K1A 0C8
Phone: 613-998-4328
FAX: 613-952-0567

ABSTRACT

Development of institutional arrangements policy for maritime, land and aeronautical MSS is an integral part of the Canadian telecommunications policy process. An ongoing activity in that process is fitting of MSS institutional arrangements policy within the confines of the 1987 Canadian Telecom Policy Framework. Making sure the fit is correct is a major task at present because technology seems to be driving service demand at rapid growth rates, particularly in the case of land MSS. This growth is stimulating policy and regulatory development efforts to keep pace. In Canada, this is happening in four planned MSS aplications areas: Canada-U.S. transborder (immediate), Teleglobe Canada Inc. international aeronautical MSS (1990/91), Telesat Mobile Inc. EMDS via INMARSAT (1990) and MSAT (1993/94). The need for an up-to-date MSS policy in these areas is emphasized by related developments in the U.S. and elsewhere. It arises because of the growing number of market initiatives proposing North American rather than Canada-only or U.S.-only coverage, such as INMARSAT, Geostar, OmniTRACS and Starlink.

PREAMBLE

1. Mobile communications has benefitted enormously from the many important technology advances which have increased its worldwide public acceptance. Phenomenal market growth is being observed in the cellular mobile industry and its spin-offs, for example, in personal radio services. Riding this wave of popularity are mobile satellite services (MSS), principally in vehicle positioning and cargo status monitoring for the trucking industry.

PURPOSE

2. This paper reviews the current status of Canadian MSS institutional arrangements policy within the context of the more general 1987 Canadian Telecom Policy Framework and makes proposals for MSS institutional arrangements.

BACKGROUND

3. The events leading to commercially viable MSS in Canada have all occurred within the last thirty years. For example, Canada's Space Program was initiated in the early 1960's and continued on through the 1970's. In 1969, Telesat Canada was founded as Canada's commercial fixed satellite service provider. In 1972, the Department of Communications developed the concept of the Canadian multi-purpose mobile communications satellite called Musat.

4. In 1971, the World Administrative Radio Conference (Space WARC-71) first allocated spectrum (in the 1500 MHz L-band) to maritime and aeronautical MSS. In 1975, the Departments of Communications, Transport and External Affairs, (with the support from the then Crown Corporation, Teleglobe Canada) participated in the establishment of the first maritime MSS by the International Maritime Satellite Organization (INMARSAT). In 1980 to 1984, working with the U.S. National Aeronautics and Space Administration (NASA), the Department developed its generic MSS concept known as MSAT.

5. In 1984, the Department transferred development and implementation of the MSAT technology to Telesat Canada which, in April 1987, set up Telesat Mobile Inc. (TMI) for the purpose. TMI was not, however, activated until December 1988.

6. In 1987, the Mobile WARC allocated 7MHz of spectrum to MSS in the L-band. Also in 1987, the Department introduced its Telecommunications Policy Framework.

7. In 1988, INMARSAT moved to provide aeronautical and land MSS and the Government of Canada signed an international agreement permitting the use of INMARSAT in territorial seas and ports. Also in 1988, the Minister of Communications declared TMI the sole provider of MSS capacity in Canada.

8. In 1989, a number of entities worldwide, including Teleglobe Canada, signed aeronautical MSS agreements with the Société internationale de télécommunications aéronautiques (SITA) to provide air traffic control (ATC), air operations (AOC), air administration (AAC) and passenger communications services (APC). Also in 1989, two U.S. companies emerged to provide dedicated radio determination satellite service (RDSS) and land MSS for the North American trucking industry. Subsequently, Canadian companies, wishing to extend these services into Canada, set up commercial arrangements with the U.S. operators. They have been licensed in Canada to provide services on an interim basis until Canadian facilities are available. This is of particular concern to Canada in that transborder services are also involved. The two main companies implementing this RDSS in the U.S. and southern Canada have already moved to provide similar services in Europe; and one of the companies is planning to expand to North Africa.

REVIEW OF ISSUES AND POLICIES AFFECTING CANADIAN MSS

Past MSS Policy Decisions

9. MSS background in Canada evolved in the context of a variety of policy and legal intruments. These ranged from the 1972 and 1982 Canada-United States Transborder Agreements through the various long-standing international conventions, operating agreements and accords to the 1984 agreement between the Department and Telesat on the terms and conditions of transfer of the MSAT technology.

10. The 1987 Telecommunications Policy Framework distinguished between carriers which owned and operated facilities (Type I) and those providing services by leasing that capacity (Type II). In 1988, in declaring TMI to be the sole provider of MSS satellite capacity within Canada, the Department excluded other Type I carriers from owning and operating the MSS satellite facility.

11. Similarly, in the provision of Canada-overseas space segment facilities, Teleglobe Canada is the relevant Type I MSS carrier. It is the sole provider of overseas MSS satellite capacity. This currently applies to INMARSAT. Teleglobe's monopoly on Canada-overseas service is subject to review in 1992 by the Government of Canada under the terms of its 1987 privatization.

12. The policy objectives of TMI and Teleglobe Canada, on entry into the MSS market in Canada, are consistent with the DOC Telecom Policy Framework. Hence, if either were to enter each other's market as a Type II carrier, that is, in the provision of domestic or international MSS, it would have to lease capacity from the other as the sole Type I domestic or international facilities provider.

13. Under authority of the Radiocommunications Act, network control is the responsibility of the Type I carrier.

14. The Department has authorized TMI to develop a generic MSS system at L-band with the proviso that any air traffic control (ATC) requirements are fully accomodated within the System. This is in order to make most effective use of the valuable spectrum-orbit resource allocated to MSS at the Mobile WARC in 1987. It will also allow this new type of service to develop in a way responsive to Canadian requirements in the 1990's.

15. Under long-established terminal attachment and system interconnection rules, land, aeronautical and maritime MSS subscribers would be allowed to own or lease mobile earth stations. Authorization of these terminals under the Radiocommunications Act would be done by DOC in accordance with existing radio regulations. In the aeronautical and maritime sectors, additional safety regulations might also have an impact on the relevant radio equipment standardization.

16. The approach to MSS policy development in Canada recognizes the trend towards integrated North American MSS. This trend is being dictated by market pressures, particularly in the case of the Canada-U.S. transborder trucking industry.

17. TMI is not yet in a position to provide its own MSS satellite facilities. Until then, it will provide interim MSS using alternate capacity which it must lease from another carrier. For example, for its "Road KIT" Early Mobile Data Service (EMDS) it plans to lease capacity from Teleglobe Canada on an INMARSAT satellite. TMI is permitted to offer interim MSS in this way on condition that its subscribers migrate to the Canadian MSAT when it becomes available.

CURRENT MSS ISSUES

18. Different in some respects from WARC Mobile '87 spectrum allocations, the Canadian spectrum policy document, SP1530MHz, allocates separate portions of the L-band to generic MSS with an exclusive aeronautical MSS allocation (sandwiched between them). The air traffic control portion of the aeronautical MSS has priority access to the upper portion of the generic MSS alocation. TMI must satisfy the Departments of Communications and Transport that the mechanisms exist to provide priority ATC, if and when required. Aeronautical public correspondence does not have the same priority and can be accomodated anywhere in the generic MSS bands.

19. In accordance with the above and the 1987 Telecom Policy Framework, Canadian-registered aircraft normally operating only on domestic routes would use Canadian MSS facilities. Foreign-registered aircraft normally operating on international routes would use either Canadian or foreign MSS networks; the question arises in this case as to what carriers and systems these aircraft would be required to use while in transit through Canadian airspace or while terminating flights in Canada. The Proposals at the end of this paper expand on the above notions and deal with how to handle such situations.

20. American MSS systems (such as the Geostar and Qualcomm systems) are currently authorized by the FCC to provide position determination and one/two way short duration messaging to, for example, the trucking industry in the continental U.S.. While these systems do not have the capacity to provide full two-way voice communications, they do have the capability of covering southern Canada. The Canadian company SaTel is the sales agent for prospective Canadian users of the Geostar Positioning Corp's System. Qualcomm's Canadian subsidiary, Qualcomm Communications Canada (QCC), is to offer service through Telesat Canada via a U.S. fixed satellite in eastern Canada and via Anik C in western Canada. After launch of Telesat's Anik E, all-Canada coverage will shift to these facilities.

21. Licensing measures exist to prevent potential Canadian subscribers to U.S. MSS dealing directly with the U.S. carriers involved. These licensing measures are based on the principles of Canadian carriers not being bypassed and MSS being provided in accordance with the Telecom Policy Framework.

22. There are agreements on transborder <u>fixed</u> satellite services with the U.S. Government that can be used as models to provide for transborder MSS. No reason is foreseen why the spirit of these "fixed satellite" agreements should not also apply to transborder MSS. Action to develop transborder MSS agreements are currently underway.

23. Two Canadian policy issues of importance at present are "Compatibility between different MSS systems" and "the extent of equipment standardization". In regard to the compatibility issue, there are some

basic limitations in the extent to which compatibility may exist in the future. One limitation relates to the use by terrestrial land mobile systems of VHF/UHF spectrum as opposed to the cellular/MSS employment of higher frequency bands; this makes it extremely improbable that compatibility standardization would ever be attempted between these two types of technology.

24. While the notion is accepted today that migration from foreign service facilities to Canadian facilities is required, the cost of implementation of such a policy, when Canadian facilities become available, is an important issue. This could be of particular concern to fleet users such as in the trucking industry.

25. In addition to the "cost" issue in migration to Canadian MSS capacity, there is the question of satisfactorily having to define "equivalence" between the new Canadian facilities and the U.S. (or other) facilities from which migration must be effected. The terms and conditions of evaluating "equivalence" would also have to be clear and unequivocal. This could prove particularly difficult because of differences in features offered on one system relative to another and the value judgements required in comparing one feature with another. For example the potential of current U.S. MSS to provide more accurate position/location information compared to the proposed 1993/94 MSAT system is one such feature. This would have to be looked at to determine whether equivalence exists and whether migration in such a case might be reasonable at the time.

26. A marketing question is the need for either TMI or Teleglobe Canada to become a Type II carrier when it wants to enter the market of the other acting as the Type I carrier of record for the particular type of MSS. For example, if Teleglobe wishes to provide domestic land MSS via INMARSAT, according to the 1987 Telecom Policy Framework, it must lease the INMARSAT capacity to TMI (the Type I carrier) and then lease it back as the end-user service provider (a Type II carrier). Further, if Teleglobe wishes to provide domestic land MSS via U.S. satellites, TMI would have to lease the required capacity from the U.S. carrier and release it to Teleglobe Canada.

27. Another facet of the same question lies in the scenario where Teleglobe is permitted to provide domestic land MSS (as a Type II carrier) using, on an interim basis, the INMARSAT System. The question arises as to whether it follows that TMI should be permitted on a quid pro quo basis (as a Type II carrier) to provide domestic aeronautical MSS using the INMARSAT system. This issue is made more complex because of Teleglobe contractual arrangements with France, Australia and SITA in this regard. It raises the further question of whether TMI could provide such a service without the permission of ICAO or the other parties to the Teleglobe accord, even where there is no international MSS involved. It is apparent that, in the event that TMI would be permitted to enter this market, the required INMARSAT capacity would have to be leased from Teleglobe Canada for the purpose.

MSS INSTITUTIONAL ARRANGEMENTS UNDER CONSIDERATION

28. On the basis of the many complex considerations impacting on provision of MSS in Canada as enumerated above, the following institutional arrangements are under consideration by the Department of Communications:

a) General:

- Authorized carriers in Canada (TMI and Teleglobe as Type I carriers) to provide satellite capacity to end users and end-user service providers

  End-users and Type II carriers required to obtain access to MSS capacity only from Type I carriers

- In the absence of Canadian facilities, use of non-Canadian MSS facilities to be permitted. Migration from non-Canadian facilities to the equivalent Canadian facility to be required when the latter becomes available

- The MSS satellite transmitter/receiver and associated control transmit/receive earth stations to be subject to licensing

  Mobile earth station licensing likely to be put in place similar to that currently applicable to terrestrial mobile stations

b) Transborder MSS:

- Extension of the 1982 Canada-U.S. transborder agreements to cover MSS and associated satellite services

- Appropriate arrangements to cover roaming of mobile earth/stations across the Canada-U.S. border in either direction.

c) Aeronautical MSS

- Canadian aircraft normally flying:

  i) on domestic routes to use the TMI MSAT MSS System

  ii) Canada-U.S. routes to use the proposed TMI or AMSC MSAT MSS System with detailed arrangements to be worked out between the two companies

  iii) on international (overseas) routes to use the TMI MSAT or alternate (INMARSAT) with the choice of facilities to be on a normal competitive business basis

- Foreign aircraft normally flying between Canadian and overseas points to use the TMI MSAT or alternate (INMARSAT) System with choice of facility to be on a normal competitive business basis

d) Compatibility Standardization:

- Possible establishment of national/regional MSS standards

- Possible TMI MSAT/AMSC standardization

- Current Canadian MSS SaTel subscribers to be made aware of the possibility of compatibility problems (including cost and inconvenience) required to migrate to MSAT at the time it becomes available.

  Similarly, for TMI's EMDS service versus other MSS via INMARSAT or via U.S. MSS systems

ACKNOWLEDGEMENTS

29. The author would like to thank experts in various fields in the Department of Communications for these valuable contribution in preparing this paper. In particular, appreciation is sincerely expressed for the efforts of Parke Davis, Dr. Robert Bowen and Murray Fyfe in institutional arrangements and spectrum policy, Demetre Athanassiadis in communications satellite technology, Pierre Gagné, Bob Tritt and Janis Doran in international coordination and, last but not least, Dr. Vena Rawat, Ron Amero, and Garry Rolston in spectrum engineering and management.

Licensing of Future Mobile Satellite Systems

Ronald J. Lepkowski
Geostar Corporation
1001 22nd Street, N.W.
Washington, D.C. 20037 U.S.A.
Phone: (202) 778-6008
FAX: (292) 223-6155

ABSTRACT

The regulatory process for licensing mobile satellite systems is complex and can require many years to complete. This process involves frequency allocations, national licensing, and frequency coordination. This paper describes the regulatory process that resulted in the establishment of the radiodetermination satellite service (RDSS) between 1983 and 1987. In contrast, each of these steps in the licensing of the mobile satellite service (MSS) is taking a significantly longer period of time to complete.

INTRODUCTION

Since the early 1970s, a major regulatory objective of the Federal Communications Commission (FCC) has been the establishment of competition in the provision of satellite services. The licensing of multiple systems has been a central feature of the domestic satellite industry in the United States. Moreover, the FCC has continued to adapt its licensing standards to minimize regulatory delay in the authorization of additional satellites as demand for service grows. The resulting competition has stimulated technical and service innovation, and reduced the price of equipment and service to the public.

RDSS has been established in the United States on a similar multiple entry basis. The establishment of the regulatory scheme for RDSS began in 1983, and was completed with the 1987 World Administrative Radio Conference (WARC) for the Mobile Services. Initial service began within the United States during 1988.

Establishing a regulatory framework for MSS has proven to be much more difficult and time consuming. This is due in large part to the different characteristics and service objectives of MSS systems. Nevertheless, competition between multiple MSS systems can produce the same types of innovation and economies that have been experienced in other areas of satellite communications.

THE REGULATORY PROCESS

The licensing of MSS systems involves three broad areas: the frequency allocation process, the licensing process, and the frequency coordination process. The frequency allocation and coordination processes can have both national and international components.

Frequency Allocations

Both the FCC and the International Telecommunication Union (ITU) establish tables which allocate specific bands of frequencies to various radiocommunication services. In each of these "tables of frequency allocations," separate bands of frequencies are allocated to RDSS and MSS, and each of these services is defined differently. RDSS is a satellite service for the "determination of the position, velocity and/or other characteristics of an object, or the obtaining of information relating to these parameters, by means of the propagation properties of radio waves." MSS is a radiocommunication service to earth stations "intended to be used while in motion or during halts at unspecified points." In addition, the mobile satellite service is subdivided into separately defined land, maritime and aeronautical mobile satellite services.[1]

Frequency allocations can be made on a primary basis or on a secondary basis not to cause interference to primary services. Bands may be allocated to one service on an exclusive basis, or bands can be shared between two or more services.

The allocations for RDSS are essentially the same in both the United States and the international tables. In the Western Hemisphere, RDSS shares its allocated frequencies with other radio services on a primary basis.

The situation regarding the MSS L-band frequency allocations is different, however. Prior to the 1987 Mobile WARC, the lower half of the MSS bands (i.e. 1530-1544 MHz downlink and 1626.5-1645.5 MHz uplink) was allocated only for maritime MSS. The upper half (i.e. 1545-1559 MHz downlink and 1646.5-1660.5 MHz uplink) was allocated only for aeronautical MSS, and there were no L-band allocations for land MSS. The 1987 Mobile WARC changed the international MSS allocations by allocating 3 MHz of the lower half of L-band for land MSS on a co-equal primary basis with maritime MSS, and re-allocating 4 MHz of the upper half to land MSS from aeronautical MSS. The United States domestic allocation of these bands does not, as a general matter, distinguish between land, maritime and aeronautical uses of MSS. However, 4.5 MHz in the upper half of L-band has been allocated domestically only for aeronautical safety MSS on a primary basis.

Licenses

Each operator of a satellite system must receive a formal license from a national telecommunications authority to construct and operate the system. Except for systems to be operated by the federal government, such licenses are issued by the FCC in the United States.

The FCC usually establishes a deadline, or "cut-off period", during which applications for licenses to operate satellites in the same frequency bands must be filed if they are to be considered at the same time. If more applications are filed by the cut-off date than can be granted because of mutual interference, the FCC must establish a policy or procedure to select which of the applications to grant.

Frequency Coordination

Since several satellite systems use the same spectrum, it is necessary to insure that their operation does not cause unacceptable interference between them. International frequency coordination procedures have been established by the ITU in Articles 11 and 13 of its international Radio Regulations.

Technical guidance is also available in the reports and recommendations of the International Radio Consultative Committee (CCIR). New reports will be adopted by the 1990 CCIR Plenary Assembly dealing with the derivation of MSS interference and sharing criteria, MSS inter-system frequency sharing and reuse, and technical aspects of MSS frequency coordination[2].

The technical aspects of coordination initially focus on the amount of discrimination or isolation that can be provided between transmissions carried over two satellite systems on the same frequency. A significant amount of isolation can be provided if the satellite antennas cover different geographical areas, or if directive earth station antennas with rapid sidelobe roll-off are used to discriminate between satellites at different orbital locations. For spread spectrum systems, isolation is provided by choosing pseudo random noise codes which have good cross-correlation properties. If sufficient isolation is not available from satellite and/or earth station antenna directivity, then frequency plans assigning different frequencies to each system can be used to eliminate interference.

RDSS LICENSING

The RDSS proceeding began in March 1983 when the Geostar Corporation filed a petition for rulemaking to allocate frequencies and establish licensing procedures for RDSS, together with an application for an RDSS system license. In its rulemaking Dockets 84-689 and 84-690, the FCC allocated domestic frequency allocations for RDSS[3] and an open entry licensing policy for this satellite service[4]. The feasibility of operating multiple RDSS systems in the same band relies on the use of different pseudo random noise spreading codes and coordination of power density levels.

The RDSS licensing procedures have been codified into Section 25.392 of the FCC's rules and regulations. These regulations include the following provisions:

(a) application content
(b) application procedures, including an automatic 60-day cut-off period
(c) blanket licensing of mobile units
(d) authorization of ancillary services
(e) frequency assignment techniques
(f) domestic inter-system coordination
(g) compliance with Docket 84-689 and 84-690 policies.

Initially, four companies were granted licenses to construct, launch and operate RDSS systems. However, only Geostar proceeded and has already begun initial operations.

Following these national decisions, frequency allocations were made for RDSS on a world wide basis at the 1987 Mobile WARC[5]. Moreover, the international Radio Regulations specify sharing criteria, such as EIRP density limits on mobile units and power flux density limits on RDSS satellites, as well as specified coordination distances for mobile RDSS earth stations, so that the conventional coordination procedures of Articles 11 and 13 can be directly applied to RDSS.

MSS LICENSING

The MSS licensing proceeding began with the filing of a petition for rule making by the National Aeronautics and Space Administration in November 1982. The FCC began its Docket 84-1234 rulemaking proceeding in late 1984 to allocate frequencies and issue the initial licenses for domestic MSS.

One of the major delays in the initial MSS licensing involved the controversy over whether to use UHF or L-band frequencies. Once the FCC allocated the upper half of the MSS L-band for domestic use[6], it then had to grant one or more of the twelve pending applications for initial MSS system licenses. The FCC had several options, such as comparative hearings to determine the best applicant, division of the available spectrum among each of the qualified applicants, a rigorous examination of the qualifications of each applicant, a lottery to randomly select a license, or an auction.

The FCC found various flaws with each of these approaches. In addition, the FCC recognized that virtually every applicant advocated a high capacity, multiple spot beam satellite design and that assigning frequencies to multiple satellites would be complicated by the division of the upper MSS L-band frequencies between aeronautical safety MSS and other non-safety MSS. As a result, the FCC established a consortium of the pending applicants to hold the initial domestic MSS system license in the upper half of the MSS L-band. The FCC also required each of the consortium members to make a $5 million cash contribution to fund the initial operations of the consortium.[7]

In selecting this approach, the Commission did not award the consortium a monopoly franchise for domestic MSS service within the United States, nor did it guarantee the economic viability of the system. The FCC stated that additional systems may be licensed in the future if the need arose, if additional allocations were made, or if technological developments made it feasible to divide the available spectrum.[7] The FCC recently began a proceeding to allocate the lower MSS L-band for domestic MSS systems.[8] The FCC will therefore have an opportunity to license one or more new MSS systems in the near future to compete with the initial domestic MSS system it licensed in 1989.

MSS COORDINATION

The initial United States MSS system of three satellites may be required to coordinate with up to 30 other domestic and international MSS satellites under the ITU procedures. The technical aspects of this coordination can be based on inter-system isolation and/or detailed frequency plans.

Inter-System Isolation

Coordination between domestic MSS systems with non-overlapping coverages should be relatively easy to achieve since the satellite antennas in each of the systems can, by themselves, provide enough inter-system isolation to reduce interference levels between co-channel transmissions to acceptable levels.

The more difficult cases involve coordination of domestic MSS systems with international MSS systems using global coverage antenna beams. Such systems provide no discrimination by the spacecraft antenna in the overlapping coverage areas. For spot beam domestic MSS systems, however, coordination of some of the spot beams can be coordinated on the basis that they do not overlap the global beam. For example, there is no geographical overlap between the western spot beams of a United States MSS system and the global beam of an Atlantic Ocean Region INMARSAT satellite.

Another method of achieving coordination is by means of mobile earth station directivity. Although omnidirectional antennas are desirable to avoid the reliability concerns associated with mechanically steered antennas or to reduce the high costs associated with electronically steered antennas, they provide virtually no antenna discrimination. Studies indicate that directive mobile earth station antennas may provide sufficient discrimination to allow co-channel transmissions on satellites spaced widely apart in orbit.[2] For example, the initial United States MSS system will employ three satellites separated by approximately 35° in longitude, and the same frequencies can be used on all three satellites by mobile earth stations with directive antennas. Initially, such directive antennas may also be employed to conserve limited satellite power.

Frequency Plans

In those cases where there is not enough inter-system isolation to permit co-channel transmissions, detailed frequency plans for each system can be developed to insure that the interfering transmissions do not use the same frequencies in both systems.

The impact of such frequency plans on current systems may not be exorbitant, since currently operational and planned systems appear to be power limited. In these MSS systems, the aggregate r.f. power available at L-band for transmission from the satellite is not sufficient to fully utilize the allocated band in an effective manner. Table 1 illustrates this condition.

Satellite Coverage	Global Beam (39 dBW Total EIRP)				Spot Beam (54 dBW Total EIRP)				
Channel Designation	Standard-A	Standard-B	Standard-M	Aero Voice	Omni Voice	Directional Voice	Fixed Voice	Aero Voice	Aero Data
Channel Bandwidth	28	20	8	17.5	5	5	20	17.5	2.5
Satellite EIRP Per Channel (dBW)	18	18	17	22	30.4	24.5	26.8	25	25
Number of Channels (includes VOX)	250	250	320	100	460	1785	1050	1590	795 no VOX
Frequency Re-Use	None	None	None	None	2X	2X	2X	2X	2X
Total Usable Allocated Bandwidth (MHZ)	7.0	5.0	2.5	1.8	1.1	4.5	10.5	13.9	1.0

Table 1. Total Allocated Downlink Bandwidth Usable with Aggregate Satellite EIRP Available

In Table 1, the amount of downlink bandwidth that can be actually used for the assumed aggregate available satellite power is calculated for several representative types of r.f. carriers. For ease of presentation, these calculations assume that the satellite is filled with

only one type of r.f. carrier, rather than the more typical case where a mixture of different r.f. carriers are carried by an MSS satellite. The global beam parameters are representative of the INMARSAT II satellite, and the spot beam parameters are representative of a domestic, multiple spot beam satellite including multi-carrier back-off. It should be noted that several, lower EIRP satellites can be operated by different companies using the same bandwidth used by a single, high EIRP satellite.

As can be seen from Table 1, currently planned MSS systems have sufficient power to utilize only a small portion of 28 MHz of L-band spectrum currently allocated for downlink MSS transmissions, except for the case where directive antennas are used with a spot beam satellite. However, directive earth station antennas can permit the same frequencies to be used by both global and spot beams covering the same geographical area in the usual case where there is a large orbital separation between such satellites. These example calculations demonstrate that coordination of multiple domestic and global beam MSS satellites with overlapping coverage areas is feasible, if spectrum efficient modulation techniques are used in both types of systems.

CONCLUSION

Lengthy licensing processes create economic disincentives and conflict with the national commitment to increased competition and innovation. With the rapid development of new terrestrial services, it is necessary to improve the functioning of these regulatory processes to maintain the competitiveness of the satellite industry. Additional spectrum in the lower half of L-band is being allocated for domestic MSS, and multiple domestic MSS systems are feasible now at L-band.

The full cooperation of existing operators of global beam satellites will be required to accommodate domestic MSS systems. Moreover, both international and domestic MSS systems will have to be configured to minimize the amount of spectrum actually occupied in their systems.

Regulatory bodies will also have to adjust their standards and procedures to recognize the feasibility and desirability of licensing competing domestic MSS systems, just as was done earlier in the case of RDSS. Regulatory bodies will also have to address the long-term spectrum needs of a growing competitive MSS industry that requires allocations beyond those made at the 1987 Mobile WARC. This can be done at the upcoming 1992 World Administrative Radio Conference, which will be able to make additional allocations to MSS in the 1 to 3 GHz portion of the spectrum.

REFERENCES

1. **Federal Communications Commission.** Rules and Regulations. *Code of Federal Regulations*, Title 47, Section 2.1.

2. **International Radio Consultative Committee**, 1989. *Methodology for the Derivation of Interference and Sharing Criteria for the Mobile-Satellite Service*, Document 8/1079; *Intersystem Frequency Sharing and Reuse in the Mobile-Satellite Services Operating at Mid to High Portions of Band 9*, Document 8/1080; *Technical Aspects of Coordination Among Mobile Satellite Systems Using the Geostationary Satellite Orbit*, Document 8/1081.

3. **Federal Communications Commission**, 1985. Report and Order. In Dockets 84-689 and 84-690, *Federal Register*, Volume 50, p. 3901.

4. **Federal Communications Commission**, 1986. Second Report and Order. In Dockets 84-689 and 84-690, *Federal Register*, Volume 51, p. 18444.

5. **International Telecommunication Union**, 1987. *Final Acts of the World Administrative Radio Conference for the Mobile Services*. (Geneva).

6. **Federal Communications Commission**, 1986. Report and Order. In Docket 84-1234, *FCC Record*, Volume 2, p. 1825.

7. **Federal Communications Commission**, 1987. Second Report and Order. In Docket 84-1234, *FCC Record*, Volume 2, p. 485.

8. **Federal Communications Commission**, 1990. Notice of Proposed Rule Making. In Docket 90-56, *Document FCC 90-63*.

A New Licensing Strategy for Canadian Mobile Earth Stations

Ronald G. Amero
Manager, Space Services Frequency/
Orbit Management Division,
Spectrum Management Operations,
Radio Regulatory Branch, Dept. of Communications
300 Slater Street
Ottawa, Ontario K1A 0C8
Tel: (613) 998-3759 Fax: (613) 952-1231

SYNOPSIS

This paper addresses initiatives presently under consideration by the Department of Communications which will greatly simplify the licensing process for mobile earth stations. The current licensing approach is reviewed and the limitations of that system identified. A new approach is needed to respond to the needs of mobile satellite users in the 90s and beyond, one which will have a minimal impact on the user while still respecting the legislative responsibilities of the Department under the new Radiocommunication Act and the General Radio Regulations. The objective of this new strategy is to lead to a transparent licensing scheme for the end user, which in turn could lead to licence exemption of blanket licences as soon as possible.

The full text of this paper was not available at press time

Session 11
Modulation and Coding - II

Session Chairman - *Erich Lutz,*
DLR - German Aerospace Research, West Germany
Session Organizer - *John Lodge,* DOC

11

New 16-PSK Trellis Codes for Fading Channels

Jun Du, Branka Vucetic and Lin Zhang

The School of Electrical Engineering
The University of Sydney, NSW 2006, Australia
Phone: 62-2-692-2154 Fax: 62-2-692-3847

Abstract

Growth in satellite mobile communications leads to increasing requirements for high data rate transmission that can be met by more efficient modulation schemes (M > 8). The 16-PSK trellis coded modulation technique is a very promising solution. In this paper, a class of new 16-PSK trellis codes with improved error rate are designed based on the criteria on fading channels.

1 Introduction

The expansion of satellite communications implies the use of higher radio frequencies (SHF and EHF regions where a major disadvantage associated with such systems is the signal attenuation due to fading. Most of the TCM codes constructed for Gaussian channels suffer significantly degradation on fading channels. It has been pointed out[2][3][4] that the codes for fading channels should be designed by using the criteria different from those for Gaussian channels.

Some efforts have been made to constructing new 8-PSK TCM codes for fading channels[3]. The growth in satellite communications, however, leads to increasing requirements for high data rate transmission that can be met by more efficient modulation schemes (M > 8). So far, the efforts on construction of 16-PSK TCM codes were concentrated on Gaussian channels [1][5][6][7]. The problem of designing 16-PSK TCM codes for fading channels is studied in this paper. A class of new 16-PSK trellis codes with improved performance on fading channels are proposed based on the code design criteria for fading channels. The performances of the 16-PSK TCM systems on Rayleigh fading channels are evaluated both theoretically and by Monte Carlo simulations.

2 System Description

In the system under consideration, the digital source generates data digit stream at a rate of $1/T_b$. Fro the 3-input information sequence, the convolutional encoder generate 4-output coded sequence. The coded sequences are then interleaved and used to modulate 16-PSK waveform at a rate $1/3T_b$ symbols/s according to symbol mapping rules. The channel is described by a slowly varying frequency non-selective Rayleigh fading. The receiver consists of a coherent demodulator followed by the deinterleaver and a Viterbi algorithm.

The channel is modelled by non-selective Rayleigh fading. We shall restrict attention to channels with slow amplitude variations relative to an elementary signaling interval and assume perfect phase tracking of the phase perturbation process. The interleaving depth is chosen to achieve ideal interleaving for a given fade rate in the simulations and assumed infinite in the analysis. In this model, the amplitude of the envelope of the received signal is multiplied by a random variable a, normalized so that $E[a^2] = 1$ to ensure that the received signal energy is equal to the transmitted signal energy E_s. The probability distribution function of a is

$$p(a) = 2ae^{-a^2} \tag{1}$$

The mean of the random variable a is

$$m_a = \frac{\sqrt{\pi}}{2} \tag{2}$$

while the variance is given by

$$\sigma_a^2 = 1 - m_a^2 \tag{3}$$

3 Error Probability and Code Design Criteria

Figure 1 shows a 16-PSK signal constellation and the notation where Δ_i denotes the Euclidean distance between channel symbols. Let C denote the set of all possible coded signal sequences $\{\mathbf{z}_N\}$, where $\mathbf{z}_N$ is a coded symbol sequence of length N:

$$\mathbf{z}_N = (z_1, z_2, \ldots, z_i, \ldots, z_N)$$

Let us suppose that $\mathbf{v}_N$:

$$\mathbf{v}_N = (v_1, v_2, \ldots, v_i, \ldots, v_N)$$

an element of C, is the transmitted sequence. Then v_i represents the MPSK symbol transmitted at time i. The corresponding received sequence at the input of the decoder is:

$$\mathbf{r}_N = (r_1, r_2, \ldots, r_i, \ldots, r_N)$$

At time i, the received symbol r_i is related to the transmitted symbol v_i by:

$$r_i = a_i v_i + n_i$$

where a_i is a complex random variable equal to the envelope of the fading channel normalized to the transmitted signal and n_i is a sample of a zero mean complex additive white Gaussian noise (AWGN) with variance σ^2 in each signal space coordinate, The corresponding received sequence $\mathbf{r}_N$ can be represented as

$$\mathbf{r}_N = \mathbf{a}_N \times \mathbf{v}_N + \mathbf{n}_N$$

where $\times$ denotes vector multiplication.

In the Viterbi algorithm, the path metric $m(\mathbf{r}_N, \mathbf{z}_N)$ is the sum of the branch metrics:

$$m(\mathbf{r}_N, \mathbf{z}_N) = \sum_{i=1}^{N} m(r_i, z_i)$$

The branch metric used in the algorithm is equal to the squared Euclidean distance:

$$m_i = |r_i - z_i|^2 \qquad (4)$$

Definitions

In the context of fading communications, besides the minimum Euclidean distance, d_e, and the number of paths at distance d_e from the correct path, $N(d_e)$, the following parameters play important roles:

1) *Effective Code Length* (ECL): the minimum effective length taken over all error paths in C where the *Effective Length*, denoted by δ, of an error path is the number of erroneous symbols on the error path, and can be expressed as

$$\delta = \sum_{i=0}^{M/2-1} w_i \qquad (5)$$

where M is the number of signal points in the signal constellation and w_i are the weights associated with the corresponding nonzero symbol distances Δ_i along the error path. The ECL of a TCM code is denoted by δ_H in this paper.

2) *Minimum Product Distance*: the minimum product distance taken over all paths with the ECL where *Product Distance* , denoted by Δ_P, is the product of the corresponding nonzero squared symbol distances along the error path and is given by

$$\Delta_P = \prod_{i=0}^{M/2-1} \Delta_i^{2w_i} \qquad (6)$$

The minimum PD is denoted by d_p in this paper.

Similar to $N(d_e)$, we have $N(d_p)$ associated with d_p, which is the number of the paths at product distance d_p from the correct path, In this paper both $N(d_e)$ and $N(d_p)$ are treated as average numbers in the same way as in [7] and [10] where the average is taken over all paths in the trellis.

Pairwise Error Probability

The pairwise error probability of choosing the path $\mathbf{z}_N$ instead of $\mathbf{v}_N$ (first error event probability) is equal to

$$P(\mathbf{v}_N \to \mathbf{z}_N) = \Pr\{m(\mathbf{r}_N, \mathbf{z}_N) < m(\mathbf{r}_N, \mathbf{v}_N)\}$$

or equivalently

$$P(\mathbf{v}_N \to \mathbf{z}_N) =$$

$$\Pr\{2\mathrm{Re}[(\mathbf{z}_N - \mathbf{v}_N)^* \cdot (\mathbf{r}_N - \mathbf{v}_N)] \geq m(\mathbf{z}_N, \mathbf{v}_N)\} \qquad (7)$$

where * denotes the complex conjugate and $\cdot$ denotes the scalar product of two vectors. The

squared Euclidean distance between the transmitted and error path is given by

$$d^2(z_N, v_N) = \sum_{i=1}^{N} |z_i - v_i|^2 = \sum_{i=1}^{N} d_i^2$$

Since d_i is equal to one of the Δ_j, $0 \leq j < M/2$, (Fig. 1) $d^2(r_N, z_N)$ can can be expressed as

$$d^2(r_N, z_N) = \sum_{i=0}^{M/2-1} w_i \Delta_i^2 \qquad (8)$$

With infinite interleaving, the sequence a_N of fading amplitudes constitutes an independent and identically distributed sequence. Equation (7) can be rewritten as

$$P(v_N \rightarrow z_N) = Pr\{2Re[(z_N - v_N)^* \cdot n_N] \geq$$

$$d^2(z_N, v_N) - 2Re[(z_N - v_N)^* \cdot (a_N \times v_N - v_N)]\} \qquad (9)$$

Let us suppose for simplicity that a_N is real and the modulation scheme is M-PSK. Then the right hand side of Eq.(9) can be simplified by noting that:

$$-2Re[(z_N - v_N)^* \cdot v_N] = d^2(z_N, v_N)$$

and

$$-2Re[(z_N - v_N)^* \cdot (a_N \times v_N)] = \sum_{i=1}^{N} a_i d_i^2$$

Then Eq.(9) can be rewritten as

$$P(v_N \rightarrow z_N) =$$

$$Pr\{2Re[(z_N - v_N)^* \cdot n_N] \geq \sum_{i=1}^{N} a_i d_i^2\} \qquad (10)$$

For a given realization of the fading variable a_N, the term on the left hand side of (10) is a zero-mean Gaussian random variable with variance $4\sigma^2 d^2$ and the term on the right hand side is a constant which depends on a_N. The conditional pairwise error probability for a given realization of the fading random variable a_N is given by

$$P(v_N \rightarrow z_N | a_N) = Q\left(\frac{\sum_{i=1}^{N} a_i d_i^2}{2\sigma d}\right) \qquad (11)$$

where $Q(x)$ is the complementary error function defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \qquad (12)$$

In this analysis, we shall distinguish two cases, depending on the value of the ECL δ_H. For large δ_H, using the approximation $Q(x) \leq 1/2\exp(-x^2/2)$ for $x \gg 1$, $P(v_N \rightarrow z_N)$ can be bounded by or equivalently

$$P(v_N \rightarrow z_N) =$$

$$P_{d,D} \leq \frac{1}{2} \frac{2\sigma d}{\sqrt{4\sigma^2 d^2 + \sigma_a^2 D^4}} \exp\left(-\frac{1}{2} \frac{m_a^2 d^4}{4\sigma^2 d^2 + \sigma_a^2 D}\right) \qquad (13)$$

with

$$D = \sum_{i=1}^{N} d_i^4 \qquad (14)$$

The formula (13) can be used for codes with large ECL for any SNR or for codes with medium ECL for small SNR where the influence of error paths with high effective lengths is significant. The bit error probability is upperbounded by

$$P_b \leq \frac{1}{k} \sum_{d,D} b_{d,D} P_{d,D} \qquad (15)$$

where $b_{d,D}$ is the total number of erroneous information bits on all paths at the Euclidean distance d from the transmitted path characterized by the distance D and k is the number of information bits in a symbol.

In the second case, when δ_H is small, the Gaussian assumption is no longer valid and the probability has to be expressed as or equivalently

$$P(v_N \rightarrow z_N) =$$

$$\int_{a_N} P(v_N \rightarrow z_N | a_N) p(a_1) \ldots p(a_N) da_1 \ldots da_N \qquad (16)$$

For large signal-to-noise ratios, we can write

$$P(v_N \rightarrow z_N) \leq \frac{1}{2} \prod_{i=1}^{N} \frac{8\sigma^2 d^2}{d_i^4} \qquad (17)$$

or equivalently

$$P(v_N \rightarrow z_N) = P_{d,w} \leq \frac{1}{2} \frac{(8\sigma d)^{2\delta_H}}{\prod_{i=0}^{M/2-1} \Delta_i^{4w_i}} \qquad (18)$$

It should be pointed out that the formula (18) is valid for short ECL codes and large SNR only.

The bit error probability is obtained by using the union bound technique as

$$P_b \leq \frac{1}{k} \sum_{d,w} b_{d,w} P_{d,w} \qquad (19)$$

where $b_{d,w}$ is the total number of erroneous information bits on all paths characterized by the Euclidean distance d and the set w.

Code Design Criteria

The analytical results indicate that the design criteria for fading channels will depend on ECL values. For the codes with large ECL, the error probability (Eqs. (13)) is dominated by the minimum Euclidean distance, while for the codes with small ECL, Eq.(18) and Eq.(19 show that the error performance is mostly dependent on the ECL, the minimum PD and the number of erroneous bits on error paths, b_d.

To construct good codes for fading channel, it is very important to avoid having parallel transitions in the trellis. All the previously known 16-PSK and 16-QAM codes[6][5][7], however, have parallel transitions. An upper bound on the maximum ECL has been obtained[8]. Table 1 shows the upper bounds on ECL, L_m, for 16-ary trellis codes with constraint length ν.

According to the possible values of ECL, to minimize error probability on fading channels one should attempt to

- when the possible ECL is small
  (I-a) design the codes with maximum ECL and
  (I-b) among the codes obtained form (I-a) maximize d_p first, and then maximize d_e
  (I-c) minimize $N(d_p)$ and $N(d_e)$ among the found codes
- when the possible ECL is large
  (II-a) maximize d_e
  (II-b) minimize $N(d_e)$ among the found codes

The criterion set I can be used to construct the codes with small possible ECL. It is possible to construct new 16-PSK TCM codes with no parallel transitions, though the minimum Euclidean distances are inferior to the known codes.

4 Code Construction

A computer search based on the defined criteria has been carried out in order to find out the best codes for fading channels. To minimize bit error rate, the Gray mapping (i.e., the neighboring symbols in a constellation differ only by one bit) is adopted in code construction.

A stack algorithm for calculating the minimum free Hamming distance of convolutional codes has been used for calculating the minimum free Euclidean distance of trellis codes by replacing the Hamming weight of channel symbols by the Euclidean weight[1]. This algorithm is referred to as Algorithm 1 in this paper.

A modified version of this algorithm, which enables to compute not only d_e but also $N(d_e)$, has been developed[9]. In the modified algorithm, additional information on the multiplicity of the merging paths to a state with the lowest Euclidean weight is added into the stack memory. The search does not terminate d_e is reached, but continues till all the paths at d_e are found. We refer to this modified algorithm as Algorithm 2 here.

In order to compute d_p, another modification of the previous algorithms has been made. The new version can compute the minimum PD among the paths of any specified effective length. We refer to the new algorithm which computes only d_p as Algorithm 3. The other new algorithm which computes d_p as well as $N(d_p)$ is referred to as Algorithm 4.

From Table 1, one can see that the maximum possible ECLs for 16-ary trellis codes are not large. According to the code design criteria in Section 3, the computer search should be carried out following the criteria set I.

To make the search efficient, Algorithms 3, 1, 4 and 2 are used successively for searching the best codes in terms of maximum ECL, d_p, d_e, minimum $N(d_p)$ and $N(d_e)$ among the codes obtained at each previous step.

Table 2 lists the parameters of some new 16-PSK codes, where Code Fi is 2^i-state code proposed for fading channel. As references, the parameters of the corresponding Ungerboeck 16-PSK TCM codes given in [6] are also listed as Code Ui. All of the other known 16-PSK codes[5][7] have parallel transitions and therefore they behave similarly on fading channels. In the table, the parity-check polynomials are given in the octal form, and N and G denote the natural mapping and the Gary mapping respectively.

It can be seen from the table that the proposed codes for fading channels have the maximum ECL and maximum d_p, but smaller d_e than the corresponding Ui type codes. It should be noted that

the search for codes was only based on the first term of d_p and d_e. It is possible to achieve a slight improvement in error rate performance by optimizing the distance spectrum though it requires more complicated search algorithms.

5 Performance Evaluation

The performance evaluation in terms of bit error rate (BER) has been conducted. To compute the analytical bounds for the BER performance, a modified Viterbi algorithm has been implemented to compute distance spectrum. Concurrently, extensive Monte Carlo computer simulations of the schemes above were carried out.

Fig. 2 shows the analytical and simulation BER results for the 16-state code F4 compared with code U4, on Rayleigh channel. Fig. 3 illustrates the results for the 128-state code F7 compared with code U7. A summary of the performance of all proposed Fi type codes is presented in Fig. 4.

The performance results show that all the Ungerboeck codes have error floors above BER=10^{-4} on fading channel due to parallel transitions. The proposed codes achieve significant improvement. As the simulation results show, at the BER = 10^{-3} (which is an acceptable rate for digital voice transmission), the difference between the two classes of codes varies from 1.5 dB for the 256-state codes to 6 dB for the 128-state codes on Rayleigh fading channels. This difference is even larger at lower error rates.

6 Conclusions

The design criteria based on ECL and PD are applied to construction of new 16-PSK trellis codes with improved error rate performance on fading channels. The code performance is evaluated both analytically and by simulation. The comparison with the previously known codes of the same complexity shows that the proposed codes can achieve significant improvement in bit error rate. The new 16-PSK trellis codes will meet the requirements for high data rate transmission in future satellite communications.

Acknowledgements

This work was carried out under the IR&D Commonwealth Department Grant in Satellite Mobile Communications.

References

[1] **Ungerboeck, G.** 1982. "Channel Coding with Multilevel/Phase signals", *IEEE Trans. Inform. Theory*, pp. 55-67.

[2] **Divsalar, D. D.** and **Simon, M. D.** 1988. "The Design of Trellis Coded MPSK for Fading Channels: Performance Criteria", *IEEE Trans. Commun.*, pp. 1004-1011.

[3] **Schlegel, C.** and **Costello, D. J.** 1989. "Bandwidth Efficient Coding for Fading Channels: Code Construction and Performance Analysis", *IEEE J.S.A.C.*, pp. 1356-1368.

[4] **Vucetic, B.** and **Nicolas, J.** 1990. "Construction of M-PSK Trellis Codes and Performance Analysis over Fading Channels" *ICC'90*, Atlanta, USA

[5] **Wilson, S. G. et al**, 1984. "Rate 3/4 Convolutional Coding of 16-PSK: Code Design and Performance Study", *IEEE Trans. Commu.*, pp. 1308-1315.

[6] **Ungerboeck, G.** 1987. "Trellis Coded Modulation with Redundant Signal Sets", *IEEE Commun. Mag.*, pp. 6-21.

[7] **Benedetto, S., et al.** 1988 "Combined Coding and Modulation: Theory and Applications" *IEEE Trans. Inform. Theory*, pp. 223-236.

[8] **Schlegel, C.** 1988. "Bandwidth Efficient Coding for Non-Gaussian Channels", Ph.D. Dissertation, University of Notre Dame.

[9] **Du, J.** and **Kasahara, M.** 1989. "Improvements in Information-Bit Error Rate of Trellis Coded Modulation Systems", *Trans. of IEICE of Japan*, pp. 609-614.

[10] **Rouanne, M.** and **Costello, D. J.** 1989. "An Algorithm for Computing the Distance Spectrum of Trellis Codes" *IEEE J.S.A.C.*, pp. 929-940.

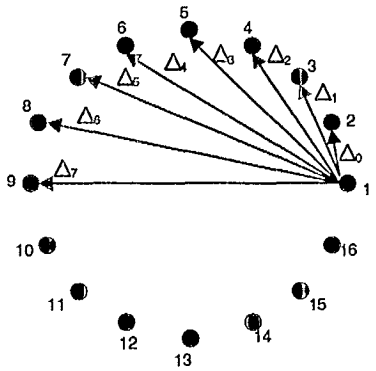


Figure 1: Notation for 16-PSK Modulation

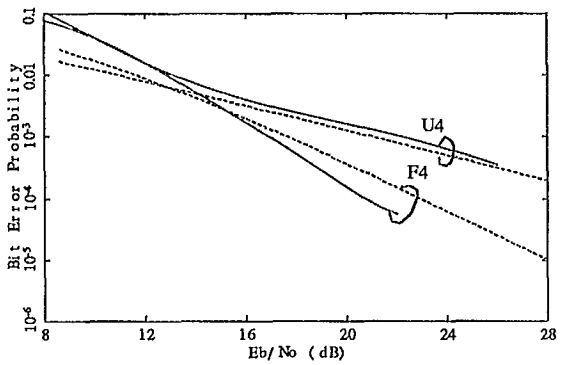


Figure 2: BER performance of the 16-state 16PSK TCM codes on Rayleigh channel
Dash: Analysis; Solid: Monte Carlo simulation

Table 1 Upper Bound on Maximum ECL .

ν	2	3	4	5	6	7	8	9	10	11	...
L_m	1	2	2	2	3	3	3	4	4	4	...

Table 2 16-PSK TCM codes .

i.d.	h^3	h^2	h^1	h^0	δ_H	d_e	$N(d_e)$	d_p	$N(d_p)$	Map.
U3	-	-	04	13	1	1.47	4.00	2.00	2.00	N
F3	15	17	05	13	2	0.89	0.19	0.61	0.25	G
U4	-	-	04	23	1	1.62	8.00	2.00	2.00	N
F4	37	21	15	33	2	0.89	0.34	2.47	0.50	G
U5	-	-	10	45	1	1.91	8.00	2.00	2.00	N
F5	75	73	47	57	2	1.48	0.77	13.7	1.00	G
U6	-	-	024	103	1	2.00	2.00	2.00	2.00	N
F6	165	175	115	143	3	1.48	1.90	0.68	1.00	G
U7	-	-	024	203	1	2.00	2.00	2.00	2.00	N
F7	337	343	307	211	3	1.48	0.03	4.69	2.00	G
U8	-	374	176	427	1	2.08	7.88	2.00	1.00	N
F8*	771	513	463	405	3	1.78	1.39	27.3	1.00	G

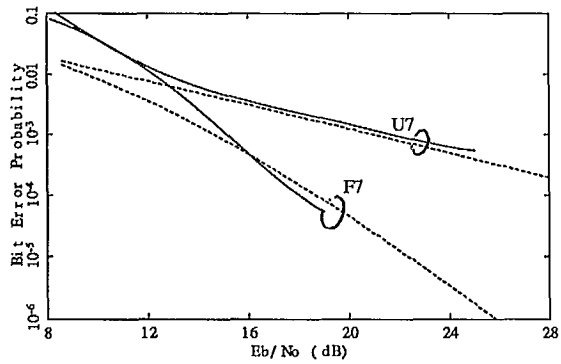


Figure 3: BER performance of the 128-state 16PSK TCM codes on Rayleigh channel
Dash: Analysis; Solid: Monte Carlo simulation

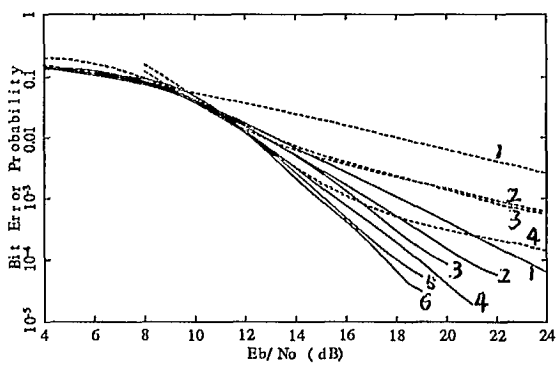


Figure 4: Summary of BER performance of 16PSK TCM codes on Rayleigh channel
Dash 1: Uncoded 8-PSK; Dash 2 to 4: Codes U3,U7 and U8; Solid 1 to 6: Codes F3,F4,F5,F6,F7 and F8; All results are obtained by Monte Carlo simulations

Power and Spectrally Efficient M-ARY QAM Schemes for Future Mobile Satellite Communications

K. Sreenath
Electrical Engineering
University of Ottawa
Ottawa, Ontario, K1N 6N5
Tel: (613) 564–3252
Fax: (613) 564–6882

K. Feher
Elec. Engg. & Comp. Sci.
University of California
Davis, CA 95616

Abstract

An effective method to compensate nonlinear phase distortion caused by the mobile amplifier is proposed. As a first step towards the future use of spectrally efficient modulation schemes for mobile satellite applications, we have investigated effects of nonlinearities and the phase compensation method on 16-QAM. The new method provides about 2 dB savings in power for 16-QAM operation with cost effective amplifiers near saturation and thereby promising use of spectrally efficient linear modulation schemes for future mobile satellite applications.

1 INTRODUCTION

Mobile Satellite (MSAT) communications represent one of the most vital and key areas of the mobile communications field. As a result of extensive research, MSAT terminals accommodating speeds of 4.8 kb/s in 5 kHz bandwidth to provide good quality voice, have been developed. The need to use cost effective non-linear amplification, as well as robust operation to overcome the strong fading behaviour of the MSAT channel has forced the use of MCPSK type signal format (8PSK signal constellation are presently used). Also, to provide better power efficiency, the use of coding (in the form of trellis coded modulation (TCM)) has been applied [1].

Although the present spectral efficiency of 1 b/s/Hz (achieved by 4.8 kb/s in 5 kHz) may be sufficient for today's demand, it is quite certain that higher efficiency is required in the near future to provide new services and accommodate more users in the already congested radio frequency band. The natural direction is to consider the development of MSAT schemes based on the use of M-ary QAM signal formats. However, a few problems related to the mobile satellite channel has to be addressed. They include: Amplifier nonlinearities, fading, doppler effects, phase noise, etc. In [2] feasibility of multilevel signaling scheme (16-QAM) is demonstrated in land-mobile fading channels by employing a Rayleigh fading compensator. It is also shown that fade compensated 16-QAM does not suffer from error floors as compared to conventional GMSK. Use of 16-QAM modulation scheme for satellite channels has also been investigated in [3], [4].

The effect of nonlinearities in the form of AM/AM and AM/PM distortions have to

be considered seriously as they can produce high levels of signal distortion and interference. This problem has to be addressed in the MSAT terminal design to employ cost effective amplifiers. For operation near saturation, the performance of QAM schemes is dictated mainly by the outermost states. The degradations are caused by the phase rotation and amplitude compression of the various states with respect to their nominal positions. Most of the amplifiers exhibit fairly linear AM/AM characteristic if the output is backed slightly from saturation. However, their AM/PM characteristics are significantly nonlinear. As the system performance is mainly controlled by the phase change between the operating point (where the average power of the signal is present) and the point at which the outermost state lies, it is necessary to equalize the nonlinear phase to operate near saturation. The immunity of the QAM schemes can be expressed by their phase margins, where phase margin is defined as the phase rotation required from the nominal value to create the irreducible error floor. For 16-QAM, the phase margin is 16.87 degrees [7].

To overcome nonlinear distortion several methods are possible. For linear modulation schemes, the popular methods employed are, predistortion [5] and new signal constellations [6]. For smaller constellations (< 64 states), predistortion method either in analog or digital form is employed and for larger constellations (> 64 states), both methods are used.

In this paper, we propose the use of 16-QAM for future MSAT terminals. Towards this direction, we first analyze 16-QAM transmission in a Gaussian channel employing a cost effective amplifier in a mobile terminal. A novel phase compensation method to overcome the AM/PM distortion will be introduced. Next, the performance of this scheme is evaluated in the mobile satellite channel environment by using Rician fading model. The performance is analyzed using computer simulation. Organization of the paper is as follows: In section 2, the communication system model is described. In section 3, the performance results are discussed and finally, in section 4, conclusions are presented.

2 SYSTEM MODEL

Block diagram of the communication system is shown in Fig. 1. The transmitter consists of a signal mapper, a spectral shaping filter and a modulator. A square root ($\alpha = 0.5$) Nyquist I filter is equally apportioned between transmitter and receiver. A class-AB amplifier [5] is used as a model to simulate the mobile terminal nonlinear amplifier characteristic. Computer simulation is conducted using an equivalent quadrature baseband model. In order to account for the level compression due to the nonlinear AM/AM characteristic of the amplifier while operating near saturation, an adaptive threshold detector is employed at the receiver [7].

The transmitted signal $x_T(t)$ can be expressed as:

$$x_T(t) = Re\{b(t)\}e^{j\omega_c t} \qquad (1)$$

where $b(t) = b_I(t) + jb_Q(t)$ is the filtered complex baseband signal and ω_c is the carrier frequency. The received signal $x_r(t)$ can be written as:

$$x_r(t) = Re\{c(t)x_T(t)\}e^{j\omega_c t} + n(t) \qquad (2)$$

where $c(t)$ represents the fading process and $n(t)$ is the additive Gaussian noise. In the present case, the fading process is Rician in nature.

The received complex baseband signal expressed by Eq.(2) can be obtained by demodulating with a local oscillator whose frequency is ω_c. It is assumed for the analysis, that good estimates of the carrier phase and symbol timing are available at the receiver. The demodulated in-phase and quadrature signals are sampled and passed on to the threshold comparators in the receiver. For a specified power of white Gaussian noise at the threshold detector input, the error probability of the i^{th} symbol with respect to the in-phase channel is computed as follows:

$$p^I_{e_i} = \begin{cases} \frac{1}{2} erfc \frac{Abs(I_i)-S^I_{i_1}}{\sigma\sqrt{2}}, & i = \pm(2N-1) \\ \frac{1}{2}[erfc \frac{Abs(I_i)-S^I_{i_1}}{\sigma\sqrt{2}} + erfc \frac{S^I_{i_2}-Abs(I_i)}{\sigma\sqrt{2}}], & \\ i = \pm 1, \pm 3, ..., \pm(2N-3) & \end{cases} \tag{3}$$

where I_i is the magnitude of the i^{th} received sample and $S^I_{i_1}$ and $S^I_{i_2}$ are the lower and upper thresholds. $S^I_{i_1}$ and $S^I_{i_2}$ are optimized to lie in the middle of the each eye to minimize the effect of signal distortion. The total symbol error probability for a sequence of N symbols is calculated as

$$P^s_{e_g} = \frac{1}{N} \sum_{i=1}^{N} (p^I_{e_i} + p^Q_{e_i} - p^I_{e_i} p^Q_{e_i}) \tag{4}$$

The error performance in a Rician fading channel can be computed as follows:

$$P_e = \int P^s(_{e_g}/r) f(r) dr \tag{5}$$

where $f(x)$ is the probability density function for Rician channel and r is the gain attenuation applied to the signal by the fading. $f(x)$ for $x > 0$ is given by [8]

$$f(x) = 2x(1+K)e^{-K} e^{-(1+K)x^2} I_o(2x\sqrt{K(1+K)}) \tag{6}$$

where K is the ratio of specular to diffuse power and $I_o(x)$ is the zero order modified Bessel function of the first kind. For MSAT channels, K=10 dB has been considered [8].

In this paper, we propose a simple and effective method to cancel the effect of AM/PM distortion due to the amplifier, thereby enabling the operation of the amplifier closer to saturation. A block diagram of the proposed method is shown in Fig. 2. By using look-up tables, where the data regarding the relationship between the sampled amplifier input envelope and amplifier phase output are stored,, we can predistort the carrier phase at the amplifier input and thereby effectively canceling the AM/PM distortion.This can be effected by providing a phase correction to the transmitter local oscillator as shown in Fig. 2.

3 PERFORMANCE RESULTS

Error performance results in a Gaussian channel is shown in Fig. 3 for phase uncompensated and phase compensated operation. The transmitter is operated at 3 dB output backoff from saturation. It can be observed that by compensating the AM/PM distortion, it is possible to operate with about 2 dB less CNR which is a significant saving in power. The effect of AM/PM compensation can be seen in the signal state space diagrams shown in Figs. 4 & 5. It can be seen that in both cases the outermost states are compressed due to operation near saturation. However, in Fig. 4 the outer states are distorted due to nonlinear phase rotation as only the average phase at the operating point is compensated, while in Fig. 5 the all states are in their nominal positions thereby providing greater decision distance and improved error

performance. The BER performance of 16-QAM in Rician channel (K=10 dB) is also shown in Fig. 3. There is modest penalty in terms of power, but no error floor is observed over the range of observation.

The effect of Doppler shift due to the relative motion between the transmitter and the receiver has not been covered in this paper, as it has been demonstrated in [2] that it can be easily compensated using compensation techniques without significant performance degradations.

4 CONCLUSIONS

A simple and effective method to compensate the nonlinear phase distortion due to the amplifier is proposed. As a first step towards the future use of spectrally efficient modulation schemes for mobile satellite applications effects of nonlinearities on 16-QAM has been investigated. It is shown that about 2 dB power savings in transmitter power can be achieved by compensating phase nonlinearities for 16-QAM. We are confident that this method could be very effective in future mobile satellite systems while employing spectrally efficient modualtion schemes.

References

[1] T.C. Jedrey, et.al. 1988, "An All Digital 8-DPSK TCM Modem for Land Mobile Satellite Communications", Proceedings of ICASSP 88, pp. 1722-1725.

[2] S. Sampei, et.al. 1989,"Rayleigh Fading Compensation Method for 16-QAM in Land Mobile Radio Channels", Proceedings of VTC'89, pp. 640-646.

[3] M.T. Lyons, et.al. 1985,"16-QAM & Trellis Coded 16-QAM on Nonlinear Channels", Proceedings of ICC'85, pp. 493-498.

[4] K. Feher 1989,"A new Generation of 90 Mb/s SATCOM Systems: Bandwidth Efficient, Field Tested 16-QAM ", Proceedings of ICC'89, pp. 552-557.

[5] Y. Akaiwa, et.al. 1987, "Highly Efficient Digital Mobile Communications with a Linear Modulation Method", IEEE JSAC, Vol. SAC-5, No. 5, pp. 890-895, June 1987.

[6] T. Ryu, et.al. 1986,"A Stepped Square 256-QAM for Digital Radio System", Proceedings of ICC'86, pp.1477-1481.

[7] K.Sreenath, et.al. 1989,"QAM and QPRS Digital Broadband Cable Systems", Int. Journal of Digital and Analog Cabled Systems, Vol. 2, pp. 139-148, 1989.

[8] D. Divsalar, et.al. 1987,"Trellis Coded Modulation for 4800-9600 bits/s Transmission over a Fading Mobile Satellite Channel", IEEE JSAC, Vol. SAC-5, No. 2, Feb 1987, pp. 162-175.

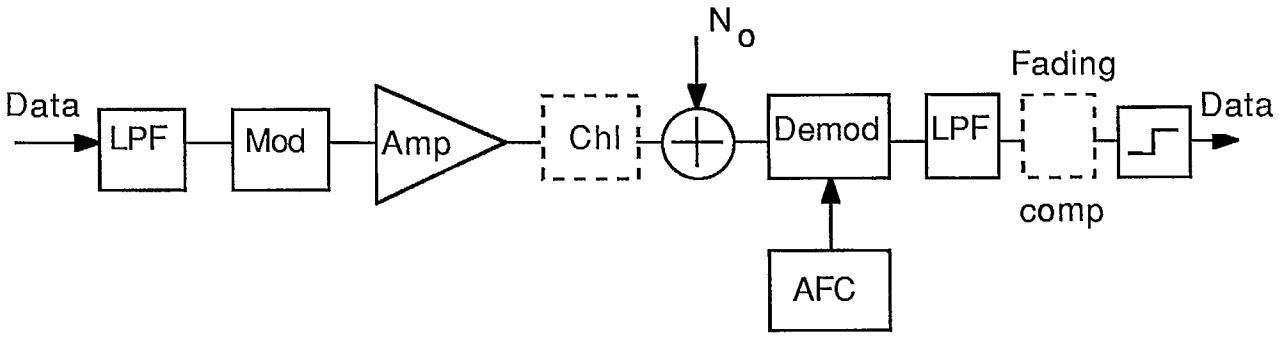


Figure 1. System Model with Linearized Amplifier & Coherent Demodulator

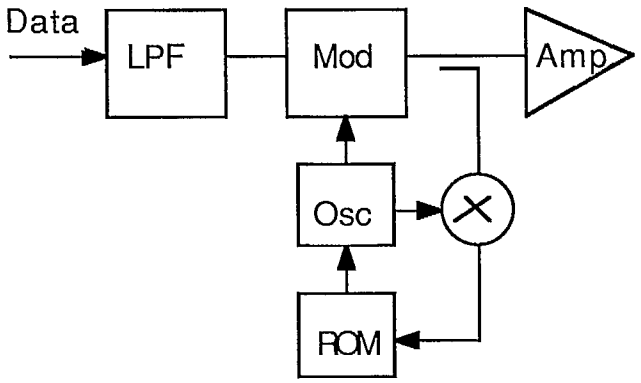


Figure 2. Nonlinear Phase Compensation Method

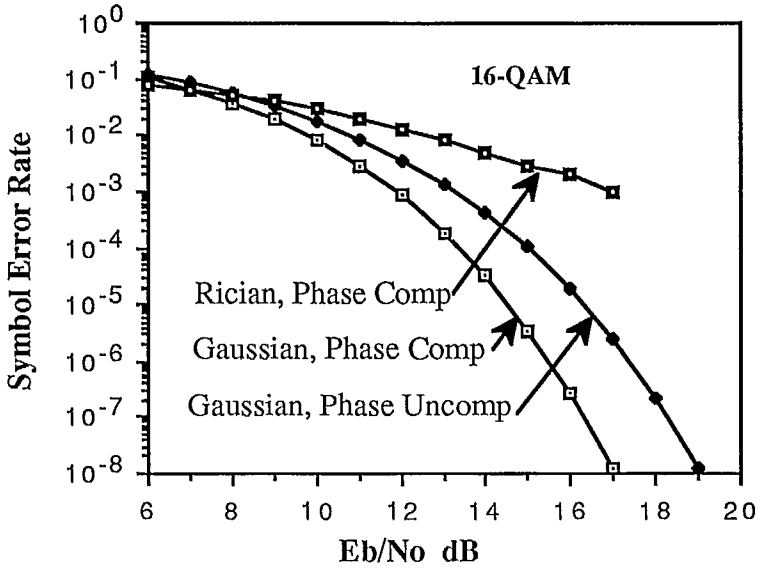


Figure 3. Error Performance of 16-QAM

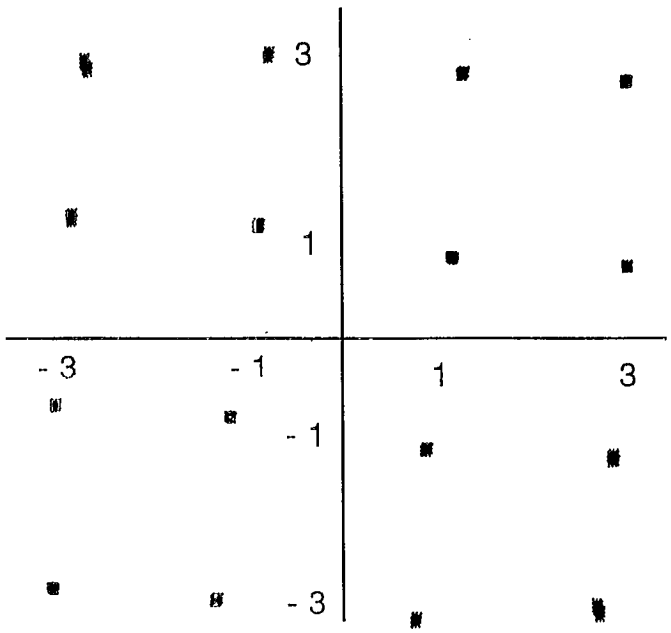


Figure 4. Signal State Space Diagram of 16-QAM Without Phase Compensation

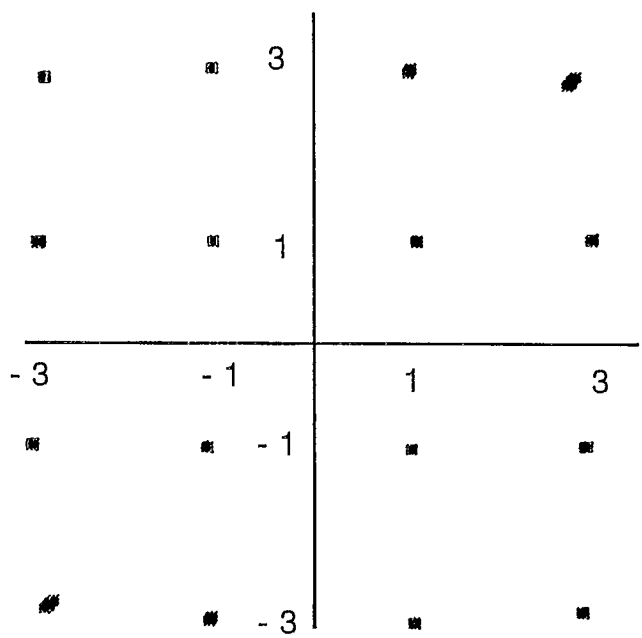


Figure 5. Signal State Space Diagram of 16-QAM with Phase Compensation

A Tone-Aided/Dual Vestigial Sideband (TA/DVSB) System for Mobile Satellite Channels

Gary J. Saulnier, Gilbert M. Millar and Anthony D. de Paolo
Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180-3590
Tel:(518)276-2762
Fax:(518)276-6261

ABSTRACT

Tone-aided modulation is one way of combating the effects of multipath fading and Doppler frequency shifts. This paper discusses a new tone-aided modulation format for M-ary phase-shift keyed signals (MPSK). A spectral null for the placement of the tone is created in the center of the MPSK signal by translating the upper sideband upwards in frequency by one-half the null width and the lower sideband downwards in frequency by the same amount. The key element of the system is the algorithm for recombining the data sidebands in the receiver, a function that is performed by a specialized phase-locked loop (PLL). The system structure is discussed and simulation results showing the PLL acquisition performance are presented.

INTRODUCTION

Mobile-satellite channels are characterized as having Rician multipath fading and Doppler frequency shifts. Due to the narrowband nature of the channels and, in particular, the magnitude of the expected Doppler shifts with respect to the channel bandwidth, the receiver must perform some type of adaptation to the channel in order to maintain acceptable performance. One approach to performing this adaptation uses a tone contained within the signal to calibrate the receiver to the channel characteristics. Such tone-aided modulation has been shown to remove the irreducible error floor that is associated with many modulations in the Rician channel (Rafferty, 1987).

While a number of tone-aided modulation schemes have been proposed (Davarian, 1987, Simon, 1986, McGeehan, 1984) each has some drawback which has detracted from its usefulness. One problem common to all the tone-aided schemes is the undesirable bandwidth expansion of the signal necessary to accommodate the one or more tones. The TA/DVSB system discussed in this paper attempts to minimize the bandwidth expansion by using a single tone placed in a spectral null that is only as wide as is necessary to accommodate the Doppler spread. Results for a binary PSK TA/DVSB system have been reported earlier (Hladik, 1989) and this paper extends this work to MPSK.

As shown in Figure 1, the TA/DVSB signal consists of a tone inserted between the frequency-translated upper and lower vestigial sidebands of a phase modulated carrier. The modulator creates this signal by separating the sidebands and frequency translating the upper sideband upwards in frequency by f_t and the lower sideband downwards by the same amount. Although the TA/DVSB spectrum is similar to that of another tone-aided system, Transparent-Tone-In Band (TTIB) (McGeehan, 1984), the TA/DVSB demodulator uses a different algorithm for recombining the data sidebands. A major consequence of the TA/DVSB recombination algorithm is that

the vestigial sidebands can be kept as small as possible (ideally they would be zero), minimizing the required transmission bandwidth. The translation frequency, f_t, is selected to accommodate the expected Doppler spread while allowing for recovery of the tone in the receiver. Typically, f_t must equal twice the expected Doppler shift so that the pilot recovery filter can extract the tone without any data sideband energy.

MODULATOR

The TA/DVSB modulator accepts M-ary symbols and a produces root-Nyquist MPSK signal whose sidebands have been separated and frequency translated as shown in Figure 1. The tone is also inserted in the modulator. A number of approaches for generating the vestigial sidebands were considered, including Weaver's single sideband (SSB) modulator and the phase-shift SSB modulator (Carlson, 1975). However, a third approach in which SSB signals are directly synthesized was chosen.

Figure 2 is a block diagram of the MPSK TA/DVSB modulator. Consider the processing of the in-phase (I) data first. Impulses representing the I data are input into two transversal filters, one having an impulse response equal to the root-Nyquist pulse shape and a second having an impulse response equal to the Hilbert Transform of the root-Nyquist pulse shape. The outputs from these filters, then, are the pulse shaped I data, $d_I(t)$, and the Hilbert Transform of the I pulse-shaped data, $\hat{d}_I(t)$. These signals are then placed on a carrier of ω_t, where ω_t is the translation frequency, i.e $\omega_t=2\pi f_t$. The resulting signal is the I portion of the baseband TA/DVSB signal,

$$S_I(t)=d_I(t)\cos\omega_t t + \hat{d}_I(t)\sin\omega_t t. \quad (1)$$

The quadrature-phase (Q) data undergoes identical processing producing

$$S_Q(t)=d_Q(t)\cos\omega_t t + \hat{d}_Q(t)\sin\omega_t t. \quad (2)$$

$S_I(t)$ and $S_Q(t)$ are finally multiplied by quadrature sinusoids at the carrier frequency, ω_c, and summed to create the modulator output. When a DC component is added to $S_I(t)$, the tone will appear in the output. The composite output, then, is

$$\begin{aligned} S(t)=&[d_I(t)+\hat{d}_Q(t)]\cos[(\omega_c-\omega_t)t] \\ &+[-\hat{d}_I(t)+d_Q(t)]\sin[(\omega_c-\omega_t)t] \\ &+[d_I(t)-\hat{d}_Q(t)]\cos[(\omega_c+\omega_t)t] \quad (3) \\ &+[\hat{d}_I(t)+d_Q(t)]\sin[(\omega_c+\omega_t)t] \\ &+DC\cos(\omega_c t). \end{aligned}$$

Figure 3 is a transmit spectrum generated using the modulator of Figure 2. The shaping is 65% excess bandwidth root-Nyquist, ω_t is 480Hz and ω_c is 6kHz. The shaping filters use 10 samples per symbol and extend over 8 symbols.

DEMODULATOR

The demodulator consists of a pilot processor, sideband separation processors, a phase-locked loop (PLL) and pulse-shape matched filters. Figure 4 shows this structure.

Pilot Processor

Figure 5 is a block diagram of the pilot processor. The input to the pilot processor is the modulator output (3) after it has been passed through the channel which introduces additive white Gaussian noise (AWGN) along with multipath fading and Doppler frequency shift. The pilot processor assumes that both the tone and the data sidebands have undergone the same phase distortions in the channel and, consequently, uses the pilot as a coherent reference to remove these distortions.

The pilot processor first mixes the input signal to baseband I and Q signals using a local reference, ω_{LO}, which is approximately equal to ω_c.

Since the carrier of the received signal will be offset from ω_c due to Doppler frequency shifts and multipath fading, the I and Q signals will be on a carrier equal to the amount of this frequency offset and the pilot tone frequency will be equal to this offset frequency. Each signal then is broken into two paths, one which uses a lowpass filter to recover the pilot tone, I_P or Q_P, and a second which delays the corresponding I or Q signal by an amount equal to the delay of the lowpass filter.

The pilot tone components, I_P and Q_P, are then subtracted from the corresponding I and Q signals, effectively implementing highpass filters. I_P and Q_P are finally used as coherent references to remove the frequency offset from the I and Q signals. In the absence of noise the pilot processor outputs, I_O and Q_O, will equal the I and Q signals in the modulator as given by (1) and (2).

Sideband Separation Processor

The sideband separation processors act to isolate the upper and lower sidebands from each other, delivering each to the PLL. While this separation can be performed using lowpass and highpass filters if the signal is placed on a carrier, this type of bandpass processing is expensive in a DSP format. Consequently it was decided to perform the sideband separation at baseband using a Hilbert Transform filter. The resulting baseband SSB signals will be complex signals.

Figure 6 is a block diagram of the sideband separation processor. An input signal, x(t), is split into two paths, one which uses a transversal filter to perform a Hilbert Transform and a second which introduces a delay equal to the delay of the Hilbert Transform filter. The Hilbert Transform filter output, $\hat{x}(t)$ is used to create the two SSB signals. The upper sideband (USB) is

$$x_{USB}(t)=x(t) + j\hat{x}(t) \qquad (4a)$$

and the lower sideband (LSB) is

$$x_{LSB}(t)=x(t) - j\hat{x}(t). \qquad (4b)$$

Figure 7 shows recovered USB and LSB signals which have been placed on a carrier of 6kHz for display. The Hilbert Transform filter used was of 101st order.

Phase-Locked Loop

The PLL operates on the outputs of the sideband separation processors and locks a local oscillator to f_t, which is then used to recombine the data sidebands. It is important to note, however, that the PLL does not have to track any channel perturbations since the pilot processor removes these prior to the PLL. The output of the PLL is the recombined data sidebands which are then processed by filters that are matched to the root-Nyquist pulse shape.

As shown in Figure 8, there are four inputs to the PLL, the USB and LSB signals derived from each of the pilot processor outputs, I_O and Q_O. At the heart of the PLL is a voltage controlled oscillator (VCO) which generates a local approximation to ω_t which will be indicated as ω_r. The two VCO outputs are

$$r_1(t)=\exp[j(\omega_r t + \theta_r)] \text{ and} \qquad (5a)$$

$$r_2(t)=\exp[-j(\omega_r t + \theta_r + \pi/2)], \qquad (5b)$$

where θ_r is the phase of ω_r.

Consider the processing of the USB and LSB signals which originate from the I signal. These signals can be expressed as

$$USB(t)=[d_I(t)+j\hat{d}_I(t)]\exp[-j(\omega_t t+\theta_t)]$$

and (6a,b)

$$LSB(t)=[d_I(t)-j\hat{d}_I(t)]\exp[+j(\omega_t t+\theta_t)],$$

where the phase of ω_t is θ_t. USB(t) is multiplied by $r_1(t)$ while LSB(t) is multiplied by $r_2(t)$. The complex conjugate of the LSB(t)$r_2(t)$ product is then multiplied by the USB(t)$r_1(t)$ product to obtain the in-phase error signal

$$e_I(t)=[d_I^2(t)-\hat{d}_I^2(t)]\sin[2(\omega_t-\omega_r)t +2(\theta_t-\theta_r)]+2d_I\hat{d}_I\cos[2(\omega_t-\omega_r)t +2(\theta_t-\theta_r)]+\text{imaginary terms,} \quad (7a)$$

where the imaginary terms are not listed since they are discarded. Processing of the USB and LSB from the Q signal results in the error signal

$$e_Q(t)=[d_Q^2(t)-\hat{d}_Q^2(t)]\sin[2(\omega_t-\omega_r)t +2(\theta_t-\theta_r)]+2d_Q\hat{d}_Q\cos[2(\omega_t-\omega_r)t +2(\theta_t-\theta_r)]+\text{imaginary terms.} \quad (7b)$$

These error signals contain the desired $\sin[2(\omega_t-\omega_r)t-2(\theta_t-\theta_r)]$ term which, when minimized, would drive ω_r and ω_t to the same frequency and phase. The problem, however, is that this term is multiplied by a coefficient that is dependent on the data and its Hilbert Transform. In addition, there is a $\cos[2(\omega_t-\omega_r)t-2(\theta_t-\theta_r)]$ term which, if minimized, would drive the VCO to a quadrature lock.

Simulation of the PLL using either $e_I(t)$ or $e_Q(t)$ as the error signal indicated that the loop could not acquire lock. Actually, for small loop gains, the PLL would tend to maintain its initial phase value, indicating that $e_I(t)$ and $e_Q(t)$ have zero mean. A later analysis showed that the coefficients for both the sine and cosine terms are zero mean.

Through simulation it was determined that hardlimiting the error signal makes the PLL acquire and maintain phase lock. This hardlimiter is indicated in Figure 8, where the error signals from the I and Q channels, $e_I(t)$ and $e_Q(t)$ are first summed and then hardlimited before being sent to the VCO. As will be shown in the next section, this PLL has a considerable amount of self noise which restricts the maximum PLL bandwidth and thereby limits its acquisition performance.

Finally, examination of (7) indicates that the error signal will not suffer from self noise if the data signals, $d_I(t)$ and $d_Q(t)$, do not make any transitions, since the Hilbert Transform terms will all be zero. Consequently, transmitting the same data symbol several times as a preamble is viable approach for obtaining fast acquisition. This property will be demonstrated in the next section.

SIMULATION RESULTS

This section illustrates the performance of the PLL by showing a received eye diagram and a number of PLL acquisition curves. The channel was not included in these simulations and, consequently, the pilot processor was also omitted. The modulator output, at baseband, was fed directly to the sideband separation processors which, in turn, fed the PLL. 8-PSK signalling with root-Nyquist pulse shaping (65% excess bandwidth) was used in the simulations.

Figure 9 is a received eye diagram obtained after the PLL has locked. The loop gain for this diagram was 10^{-4}, a value that produces a tight lock point. This eye diagram was generated after the signal was match filtered. The inter-symbol interference evident in the eye diagram is most likely due to distortion of the low frequency components caused by the sideband separation operation in the modulator.

Figure 10 is a set of PLL acquisition curves for a random data input using loop gain as a parameter. The self noise caused by the presence of the Hilbert Transform components in the error signal is evident. Figure 11

is a set of PLL acquisition curves for a constant data symbol input, i.e. no data transitions. These acquisition curves are linear due to the hardlimiter in the error signal path. Clearly, fast acquisition of the f_t signal is possible with a preamble consisting of a constant data symbol.

CONCLUSIONS

This paper shows that the TA/DVSB system is a viable approach to tone-aided modulation. The baseband implementation of the modulator and demodulator is of relatively low complexity. For burst-mode communications, where fast acquistion of f_t is required, a preamble consisting of the same symbol repeated a number of times is required due to the self-noise of the PLL.

REFERENCES

1. **Carlson, A.B.** 1975. *Communication Systems*. McGraw-Hill.
2. **Davarian, F.** 1987. Mobile digital communications via tone calibration. *IEEE Transactions on Vehicular Technology*. VT-36, pp.55-62.
3. **Hladik, S.M., Saulnier, G.J. and Rafferty, W.** 1989. A tone-aided dual vestigial sideband system for digital communications on fading channels. IEEE Military Communications Conference. pp.709-714.
4. **McGeehan, J.P., Bateman, A.J.** 1984. Phase-locked transparent tone-in-band (TTIB):A new spectrum configuration particularly suited to the transmission of data over SSB mobile radio networks. *IEEE Transactions on Communications*. COM-32, pp.81-87.
5. **Rafferty, W., Anderson, J.B., Saulnier, G.J.**, and Holm, J.R. 1987 Measurements and a theoretical analysis of the TCT fading channel radio system. *IEEE Transactions on Communications*. COM-35, pp. 172-180.
6. **Simon, M.K.** 1986. Dual-pilot tone calibration technique. *IEEE Transactions on Vehicular Technology*. VT-35, pp.63-70.

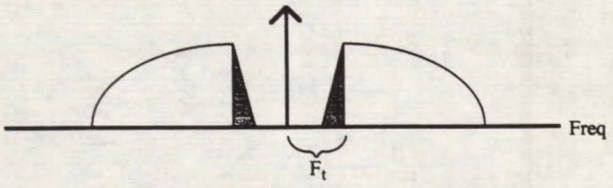


Fig. 1. TA/DVSB frequency spectrum

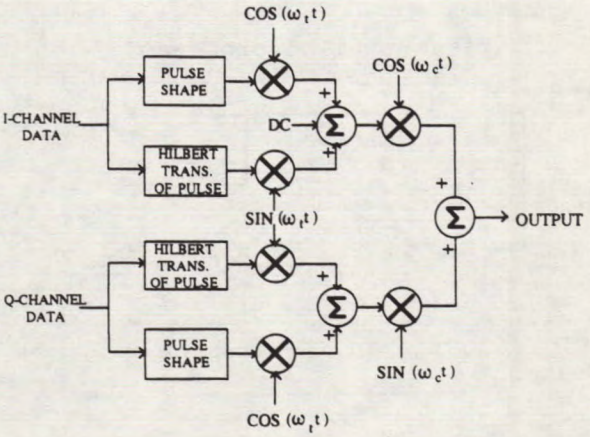


Fig. 2. TA/DVSB modulator

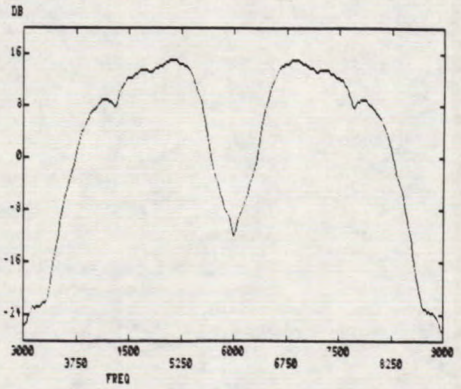


Fig. 3. Modulator output spectrum

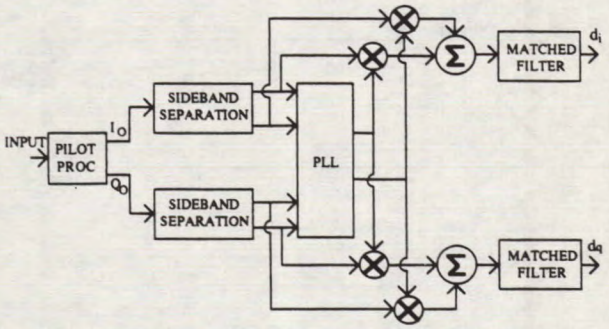


Fig. 4. TA/DVSB demodulator

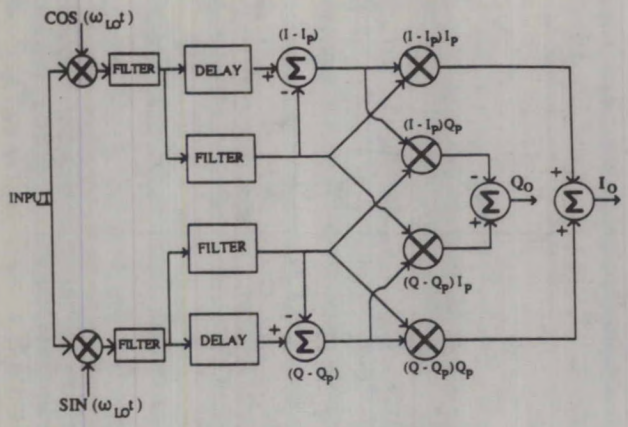


Fig. 5. Pilot processor

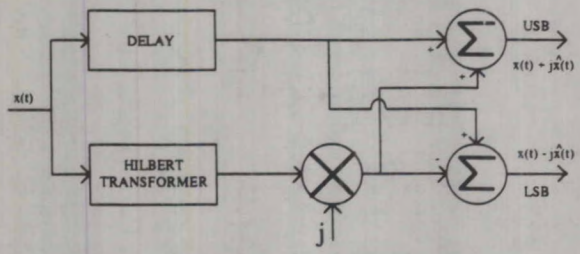


Fig. 6. Sideband separation processor

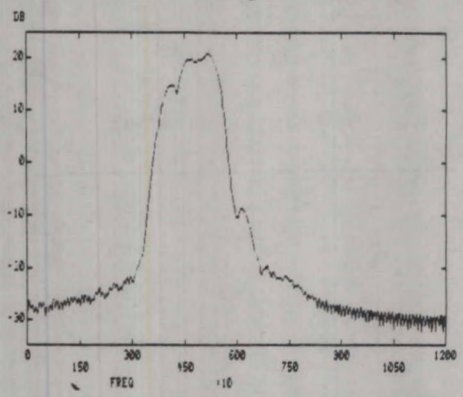


Fig. 7. Recovered sideband

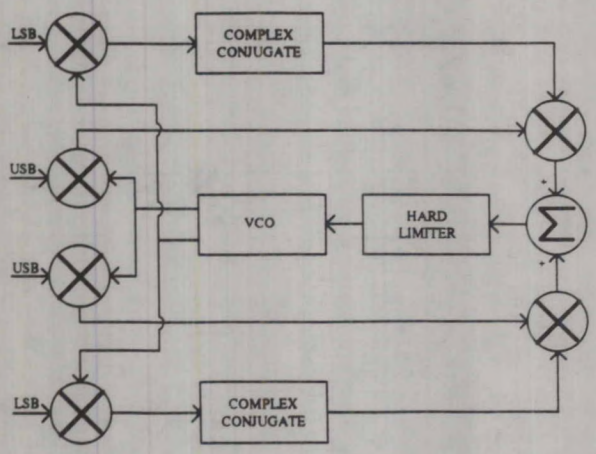


Fig. 8. TA/DVSB phase-locked loop

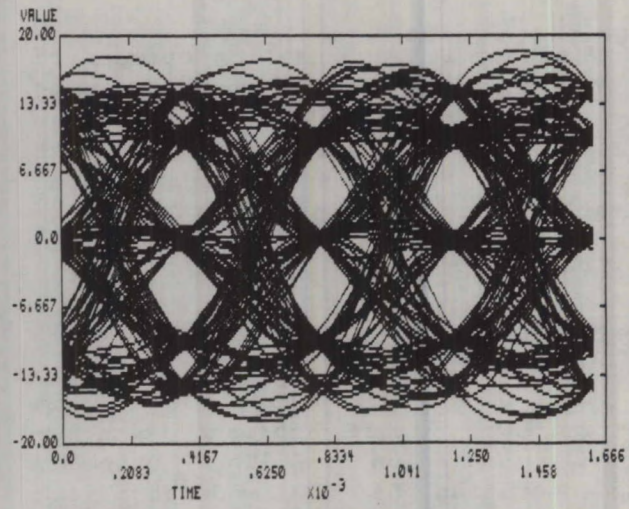


Fig. 9. Recovered eye diagram

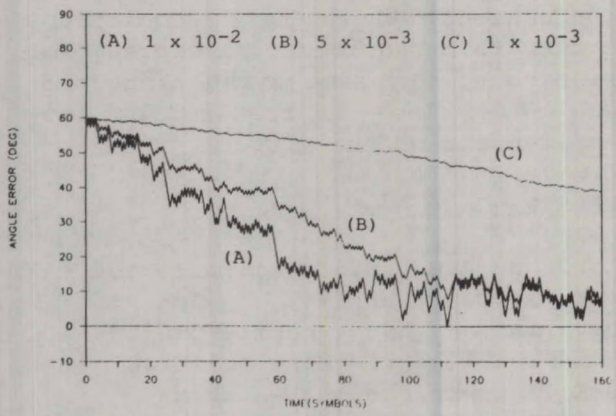


Fig. 10. Acquistion curves for random data

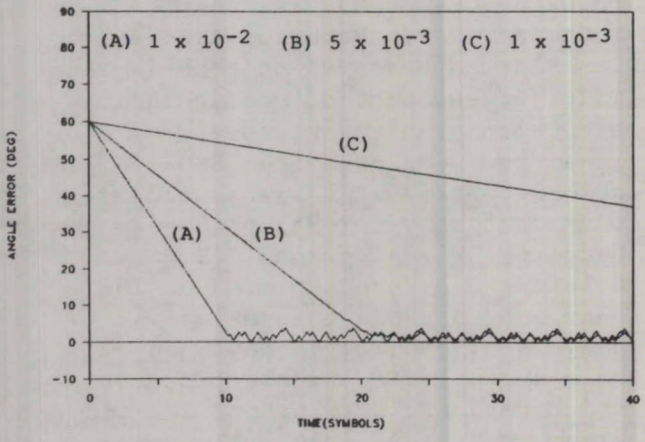


Fig. 11. Acquistion curves for no data transitions

Synchronization Techniques for All Digital 16-ary QAM Receivers Operating over Land Mobile Satellite Links

P. Fines, A.H. Aghvami
Communications Research Group
Department of Electrical and Electronic Engineering
King's College, University of London
Strand, London WC2R LS, U.K.
Phone: 44 71 873 2898
FAX: 44 71 836 4781

ABSTRACT

In this paper, the performance of a low bit rate (64Kb/s) all digital 16-ary Differentially Encoded Quadrature Amplitude Modulation (16-DEQAM) demodulator operating over a mobile satellite channel, is considered. The synchronization and detection techniques employed to overcome the Rician channel impairments, are described. The acquisition and steady state performances of this modem, are evaluated by computer simulation over AWGN and Rician channels. The results verify the suitability of the 16-DEQAM transmission over slowly faded and/or mildly faded channels.

INTRODUCTION

Because of the limited bandwidth availability for digital satellite communications, the current trend has been towards improving the spectrum efficiency of satellite systems by employing high spectrally efficient modulation schemes. Recently 16-ary QAM has received much attention as a candidate for future satellite communications. However, because of the non-constant nature of their amplitudes, 16-ary QAM signals are very sensitive to the non-linearities of satellite channels. That is why 16-ary QAM has not been used in the existing satellite communications systems. With the recent developments in solid-state power amplifier designs and using predistortion techniques, potential application of 16-ary QAM for future satellite systems are now becoming more recognizable.

To satisfy the power limitations and achieve optimum performance, coherent detection at the receiver, must be used. Coherent demodulation is a difficult task in a land mobile satellite environment. This is because multipath fading and Doppler frequency shifts prevent reliable derivation of a carrier reference signal. The carrier amplitude and phase vary considerably and rapidly, depending on the vehicle speed and the severity of the fading. In m-ary QAM transmissions, successful carrier amplitude tracking is equally important as carrier phase tracking. The 16-DEQAM system presented in this study employs coherent detection by means of decision directed carrier amplitude gain control and phase synchronization. Furthermore,large carrier frequency offsets and symbol timing errors have been tackled effectively.

16-DEQAM MODULATOR

The system model under investigation is shown in Fig.1. This model is the baseband equivalent to a typical land mobile satellite link which includes the modulator, the physical channel and the demodulator.

Because of the 16-QAM signal symmetry, there is a fourfold ambiguity in the data recovery. Differential encoding (DE) is used to resolve this phase ambiguity. In order to minimise the BER the assignment of bits to the symbols are such that its nearest neighbours differ in as few bits as possible (exact Gray coding is unachievable). The two most significant

bits (MSB) are differential encoded to represent the change in quadrant. The remain two bits represent the symbol points in each quadrant.

The generated 16-ary DE impulses are converted to bandlimited waveforms by filtering. The transfer function of the filters is square-root raised cosine with excess bandwidth of 50%.

CHANNEL MODEL

In this study no assumptions on carrier amplitude, frequency and phase or symbol timing synchronization are made. The channel model of Fig.1 demonstrates how synchronization errors are introduced as well as the fading and AWGN. The symbol timing errors are introduced by passing the transmitted signal through a variable delay line. Frequency offset errors represent the total local oscillators instability and inaccuracy between transmitter and receiver, as well as large Doppler shifts due to the receiver motion relative to the satellite.

The land mobile satellite channel is assumed non frequency selective for the data rates of interest, and its statistical characteristics are Rician. This fading model [1] considers two signal paths from the satellite to the mobile receiver: a line of sight (LOS) path and a scatter (multipath) path. As a result, the received signal can be modeled as the sum of two components, a direct and a diffuse one. Propagation experiments suggest that for land mobile satellite communications, if the direct path is clear the diffuse component has an rms value 10-20 dB below the direct one. In this study shadowing is not considered, and therefore, the LOS component power is constant.

The diffuse component is modeled with Rayleigh distributed envelope and uniform distributed phase. The time behavior of the fading is mainly characterized by the half of the Doppler spread:

$$f_D = f_c \frac{V}{C} \qquad (1)$$

where fc is the carrier frequency, V is the vehicle speed and C is the speed of light.

The power spectral density of the fading process is assumed according to:

$$S(f) = \begin{cases} \{1-(f/f_D)^2\}^{-1/2}, & |f| \le f_D \\ 0, & |f| > f_D \end{cases} \qquad (2)$$

Two independent white Gaussian sources with zero mean and one sided power spectral density S_o, are filtered by the shaping filters $S(f)$ to give a complex zero mean Gaussian random process with power spectrum $S(f)$ and total power: $2S_oB_x\sigma^2$, where B_x is equivalent noise bandwidth of $S(f)$. For simplicity, the values of S_o and B_x are such that: $S_oB_x=1$.

The random process $y(t)$ results from adding the constant μ. The amplitude of the complex fading distortion process $h(t)=x(t)+jy(t)$ is the Rician distributed carrier attenuation:

$$a(t) = \{x(t)^2+y(t)^2\}^{1/2} \qquad (3)$$

The phase of the fading process $h(t)$ is:

$$b(t) = tan^{-1}\{y(t)/x(t)\} \qquad (4)$$

The severity of the fading is determined by the direct to multipath signal power ratio (Rice factor):

$$c = \mu^2/2\sigma^2 \qquad (5)$$

By selecting $c = 0$ a severely fading Rayleigh channel results. In the limit $c \rightarrow \infty$ the Rician channel reduces to the nonfading AWGN channel.

The coefficients σ and μ are scaled such that:

$$E\{a(t)^2\} = 2\sigma^2 + \mu^2 = 1 \qquad (6)$$

in order to the average transmitted bit energy (Eb) remain unaffected by the channel fading.

16-DEQAM DEMODULATOR

The demodulator incorporates all the necessary synchronization and detection processes. Synchronization is obtained at four levels: carrier amplitude frequency and phase, symbol timing.

The received signal can be expressed as:

$$r(t)=a(t)e^{j[2\pi f_e t+b(t)]}\sum_k s(k)g(t-kT-\tau)+n(t) \quad (7)$$

where, $s(k)$ is the k-th transmitted 16-ary complex symbol, $g(t)$ is symbol pulse, T is the symbol time period, τ is the symbol timing error, f_e is the carrier frequency offset, $a(t)$ and $b(t)$ are the carrier attenuation and phase error introduced by the fading as is given by (3) and (4), and $n(t)$ is complex AWGN with one sided spectral density N_o.

In (7) the effects of fading and synchronization errors are shown clearly with simple mathematical terms. The mathematical model of the demodulator required in order to retrieve the transmitted bits from $r(t)$ is illustrated in Fig.2. The input signal is a sampled version of $r(t)$. To reduce the computing burden, the sampling rate is minimized at all stages. At the demodulator input the signal $r(t)$ is represented by its samples obtained with a sampling rate of four times the symbol rate. In order to simplify the hardware, the sampling phase is not synchronized with the symbol rate (nonsynchronous sampling). The four synchronization levels employ feedback loop structure and all of them have three elements: parameter error detector, filtering of the estimated error and parameter adjuster [2].

Carrier Frequency Synchronization

A second order phase locked loop (PLL) will be unable to phase lock in the presence of large carrier frequency offset. For that an automatic frequency control (AFC) loop is used to reduce the initial frequency error and then a PLL removes the remaining error. This strategy allows fast and reliable carrier phase acquisition. In both cases the frequency errors are corrected by frequency translation of the incoming signal $r(t)$. The AFC loop is driven by a frequency error detector which is described in ref. [3]. The frequency error is computed as:

$$f_e = |r_+(t)|^2 - |r_-(t)|^2 \quad (9)$$

with,

$$r_\pm(t) = r(t) \otimes g(t)e^{\pm j2\pi F_s t} \quad (10)$$

where $\otimes$ denotes convolution, $g(t)$ is the impulse response of the data filters and F_s is the symbol rate. $r_\pm(t)$ satisfy the conditions found in [3] for pattern jitter free frequency error detector which ensures improved performance. With small frequency error, the required frequency translation is controlled by the phase error detector similar to the case of a second order PLL. In both cases the estimated frequency error is integrated in order to control the Number Controlled Oscillator (NCO). The loop bandwidths of the AFC and PLL are controlled by the coefficients **D** and **C** as shown in Fig.2.

Symbol Timing Synchronization

Since nonsynchronous sampling is employed, correct adjusted samples are computed by interpolation through the available samples. This process is equivalent to a variable delay line in the signal path. Interpolation is implemented by convolution of the incoming signal $r(t)$ with an interpolating function:

$$r(t-\tau_d) = r(t) \otimes d(t+\tau_d) \quad (11)$$

where $\otimes$ denotes convolution, $d(t)$ is the interpolating function and τ_d is the time delay. The interpolator is implemented by a FIR transversal filter and the required delay is selected by changing its coefficients.For efficient implementation, the timing correction resolution is 1.5% of the symbol period. The timing phase error is computed by [3]:

$$\tau_e = \frac{1}{2\pi} arg\left\{ \int_{-LT}^{LT} r_+(t)r_-^*(t)e^{-j2\pi F_s t}dt \right\} \quad (12)$$

where * denotes complex conjugation, LT is a time period of L symbols and, $r_\pm(t)$ and F_s as in (10). The computed timing error τ_e is an unbiased estimate of the timing error, free of pattern noise and carrier amplitude,frequency and phase independent. The error τ_e is integrated and then is used for the appropriate selection of interpolator coefficients. The response of the loop is controlled by the coefficient **A** of Fig.2. Finally, hangup is avoided because the timing correction resolution is discrete and the error τ_e is computed by (12) every $2L$ symbols.

Carrier Amplitude Control

An automatic gain control (AGC) loop is employed to compensate the short term fading attenuation. The long term attenuation is assumed to be handled in the RF and IF sections by a more standard average energy type AGC loop. The carrier amplitude error gain is computed as:

$$A_e = |C_n| / |\hat{C}_n| \qquad (13)$$

where C_n and $\hat{C}_n$ are the n-th complex sample at the input and output of the decision unit. When the signal level and the receiver attenuation are expressed as a logarithm, the AGC loop becomes a linear system. The response of the loop is controlled by the coefficient $\mathbb{E}$ of Fig.2.

Carrier Phase Synchronization

The phase error is computed as:

$$\phi_e = arg\{C_n \hat{C}_n^*\} \qquad (14)$$

where * is complex conjugation, C_n and $\hat{C}_n$ as in (13). (14) can be implemented efficiently by a look-up table [4]. The estimated phase error is integrated and then is used to correct the phase of the matched filter sample. The response of the loop is controlled by the coefficient $\mathbb{B}$ of Fig.2. At this point, a first order loop is sufficient as long as the frequency error is corrected in a separate loop [2].

COMPUTER SIMULATION RESULTS

Acquisition Performance

The acquisition performance of the system is computed in terms of the time period required by the system to obtain synchronization starting from practical maximum timing, frequency and gain errors. Good acquisition performance will also guarantee better steady performance in terms of BER because the reacquisition of the system after a deep fade will be faster causing less errors. The convergence of the system operating over a AWGN channel with E_b/N_o=10dB, starting with a timing error of 50% of the symbol time period, 1.6KHz carrier frequency offset and 6dB signal gain, is shown in Fig.3. The timing recovery loop locks in less than 300 symbol periods. The AFC loop reduces the frequency offset in less than 300 symbols. The PLL and AGC lock in less than 40 symbols giving a total phase lock period of 340 symbols.

Steady State Performance

For a AWGN channel, the BER performance shown in Fig.4 a), demonstrates the small degradation caused by the system.

The BER performance over a Rician channel is shown in Fig.4 b) and c) for various channel parameters. From these figures it is evident that this system is suitable for channels that are not severely faded or vary slowly because of their relaxed requirements in amplitude control and phase synchronization.

CONCLUSIONS

The structure of a low bit rate 16-DEQAM coherent demodulator suitable for all digital implementation is given. Its evaluated acquisition and BER performances over AWGN and Rician channel show the suitability of this bandwidth efficient modulation scheme over mildly faded channels. The main source of degradation was found to be the necessity of good amplitude as well as phase tracking mainly during fast and deep fades.

REFERENCES

1. **Davarian, F.** 1987. Channel Simulation to Facilitate Mobile Satellite Communications Research. IEEE Trans. Commun. VOL. COM-35, NO. 1, pp. 47-56.

2. **Gardner,F.M.** 1988.Demodulator reference recovery techniques suited for digital implementation. Final report ESTEC Contract 6847/86/NL/DG.

3. **Alberty, T., Hespelt, V.**, 1989. A new pattern jitter free frequency error detector. IEEE Trans. Commun., VOL. COM-37, NO. 2, pp. 159-163.

4. **Fines, P., Aghvami, A. H.**, 1990. Implementation of a 16-ary QAM Demodulator for Medium Data Rates. Accepted for publication at the IEE Fifth International Conference on 'Radio Receivers and Associated Systems', Cambridge, July 1990.

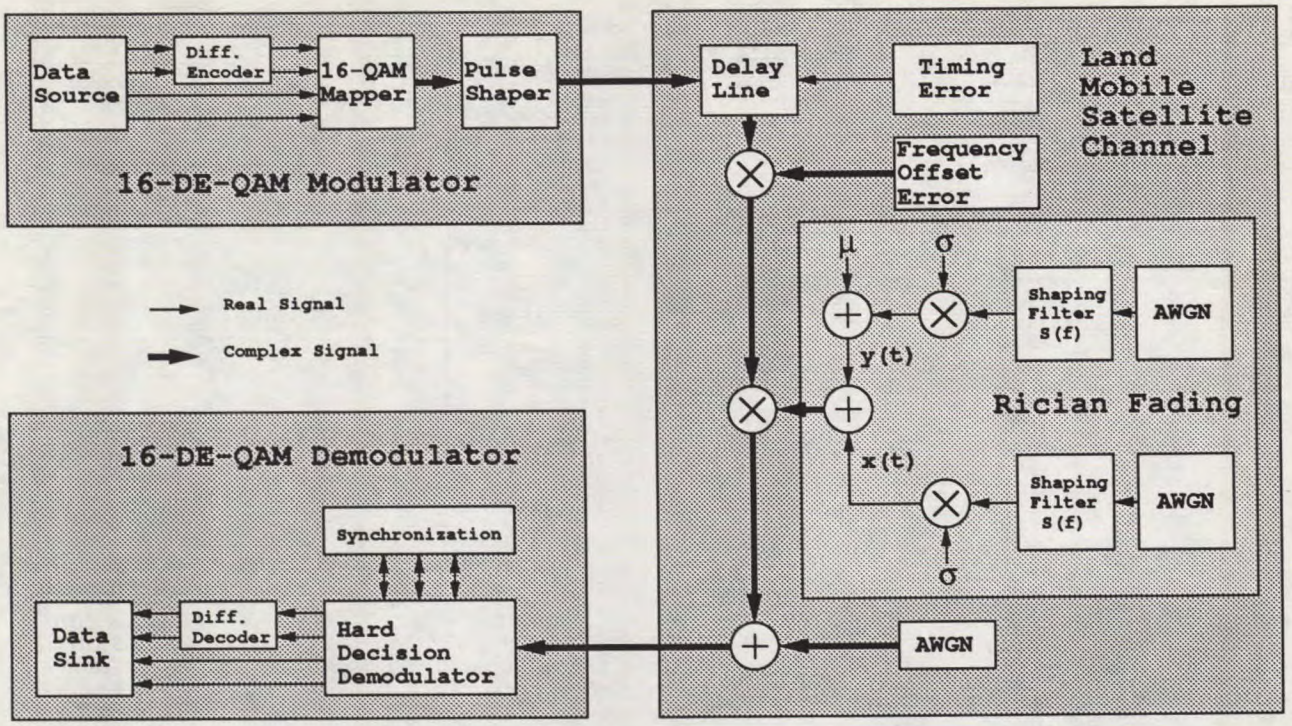


Fig.1 System model under consideration

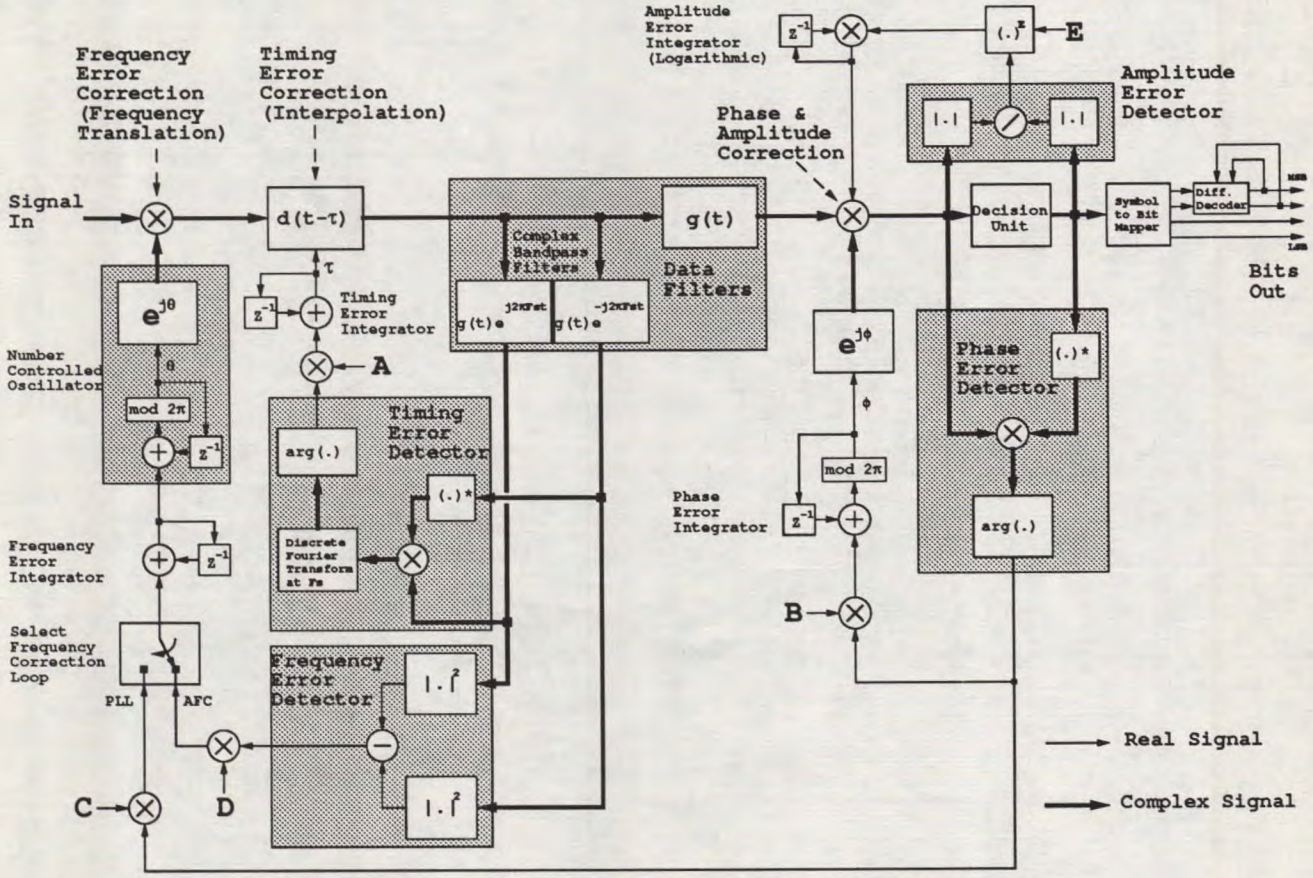


Fig.2 16-DE-QAM All Digital Demodulator

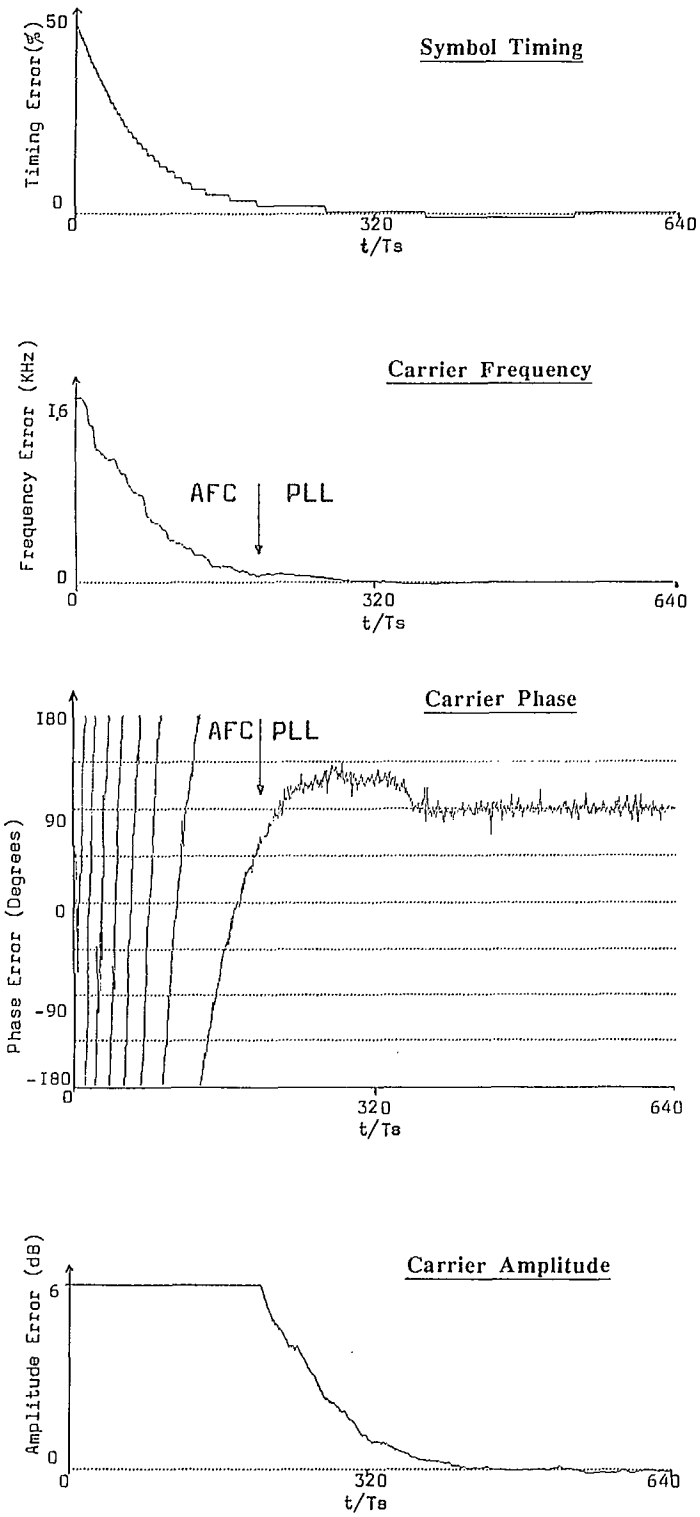


Fig.3 Acquisition performance (E_b/N_o = 10 dB)

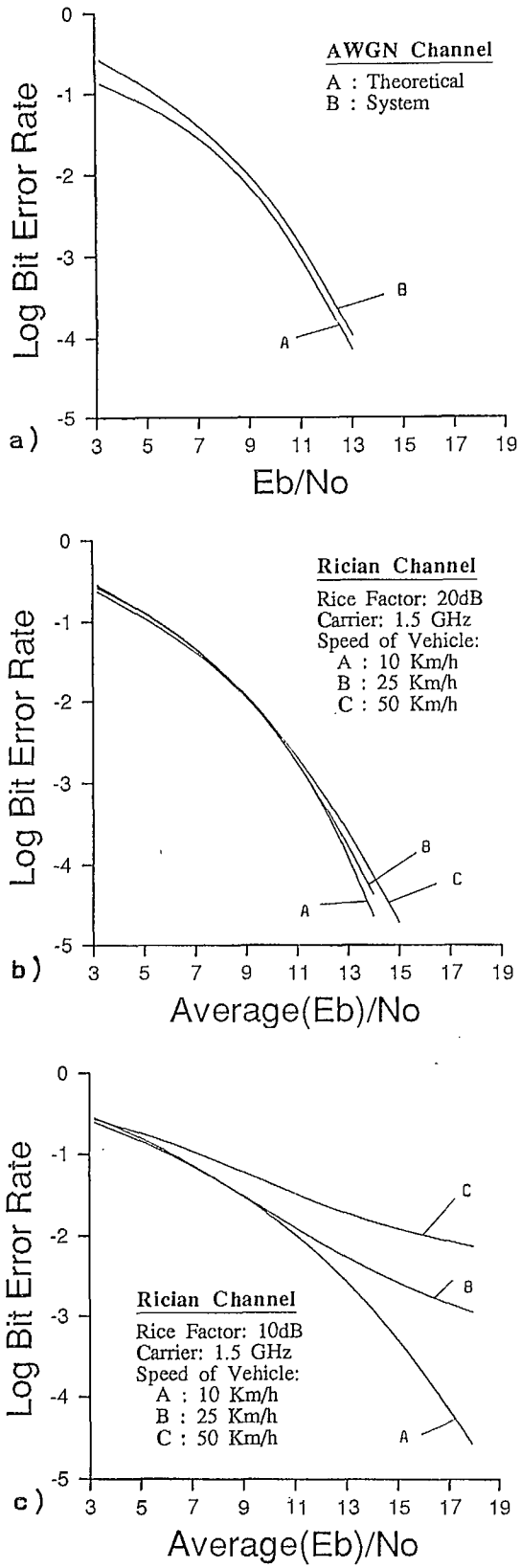


Fig.4. BER over AWGN and Rician Channels

Carrier Recovery Techniques on Satellite Mobile Channels

B. Vucetic
and
J. Du
Department of Electrical Engineering, Sydney University
Sydney 2006, Australia
Phone: 61-2-6923514
Fax: 61-2-6923847

Abstract

An analytical method and a stored channel model were used to evaluate error performance of uncoded QPSK and MPSK trellis coded modulation (TCM) over shadowed satellite mobile channels in the presence of phase jitter for various carrier recovery techniques.

1 Introduction

In coherent digital transmission systems accurate carrier phase and symbol timing information are essential for successful demodulation. The reconstruction of coherent reference can be accomplished in many ways. In general, carrier and clock are restored by pilot tone and self-synchronizing techniques. Self-synchronizing methods could be quite roughly divided into decision-directed algorithms and feedworward nonlinear phase estimators. This paper compares three representatives of the above classes and analyzes their relative merits on satellite mobile channels with shadowing. The impact of phase jitter on coherently demodulated QPSK and M-PSK trellis coded modulation (TCM) with feedforward (FF), decision directed (DD) and TTIB techniques for carrier phase recovery on a is considered.

A stored experimental channel [2] was used to simulate the performance of a satellite mobile communication system. The experimental data for the stored channel were obtained in land mobile satellite propagation measurements at L band in Sydney suburban areas via the Japanese ETS-V and INMARSAT-POR geostationary satellites.

2 System Model

2.1 System Description

In the systems under consideration a voice codec generates data digits at a rate of $1/T_b$. If the QPSK modulation is used the 2-bit input sequences modulate a QPSK waveform at a rate $1/2T_b$ symbols/s. In the 8-PSK TCM system the binary sequence is encoded in a rate 2/3 convolutional encoder. The 3-bit symbols at the output of the encoder are then block interleaved and used to modulate an 8-PSK waveform at a rate $1/2T_b$ symbols/s according to the set partitioning rules [6]. The interleavear size is chosen on the order of the worst case fade depth in the simulations and assumed infinite in the analysis. We consider three different methods for carrier phase recovery, FF, DD and TTIB techniques.

2.2 Model of Shadowed Channel

The signal in a shadowed channel is obtained as the sum of a lognormally distributed random phasor $ye^{j\phi_0}$ and a Rayleigh phasor $we^{j\phi}$:

$$ae^{j\theta} = ye^{j\phi_0} + we^{j\phi} \qquad y, w > 0$$

where the phases ϕ_0 and ϕ are uniformly distributed between 0 and 2π.

The conditional probability of a given y is Rician:

$$p(a \mid y) = \frac{a}{\sigma_r^2}\exp(-\frac{r^2+y^2}{2\sigma_r^2})I_0\left(\frac{ay}{\sigma_r^2}\right)$$

where $2\sigma_r^2$ represents the average scattered power due to multipath propagation.

The total probability is given by

$$p(a) = \int_0^\infty p(a \mid y)p(y)dy$$

It is assumed that p(y) is lognormally distributed:

$$p(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_1 y} exp\left(-\frac{(lny-\mu)^2}{2\sigma_1^2}\right) & 0 \geq y \geq \infty \\ 0 & otherwise \end{cases}$$

with the mean

$$\mu_y = E\{y\} = \exp(\mu + \sigma_1^2/2)$$

and the variance

$$\sigma_y^2 = E\left\{(y-\mu_y)^2\right\} = \exp(2\mu+\sigma_1^2)\left(\exp(\sigma_1^2)-1\right)$$

3 Phase Estimation in FF Technique

Let us assume that v_i represents an MPSK symbol sample transmitted at time i. The phase of the ith transmitted sample can be written as

$$\alpha_i = k_i\frac{2\pi}{M}$$

where k_i is an integer ($k_i \in \{0,1,...,M-1\}$.

The corresponding received sample at the input of the coherent demodulator is:

$$r_i = a_i v_i + n_i$$

where a_i is a real random variable equal to the envelope of the channel attenuation normalized to the transmitted signal, n_i is a sample of a zero mean complex additive white Gaussian noise (AWGN). The recevied sample can be expressed as

$$r_i = r_{Ii} + r_{Qi} = \rho_i\exp(j\phi_i)$$

$$\rho_i = \sqrt{r_{Ii}^2 + r_{Qi}^2}$$

and

$$\phi_i = \tan^{-1}\frac{r_{Qi}}{r_{Ii}}$$

On the other hand the received sample phase can be expressed as

$$\phi_i = \alpha_i + \varphi_i + \epsilon_i$$

where φ_i is the phase jitter and ϵ_i is the phase error caused by Gaussian noise.

In the FF algorithm for each sample we multiply the phase ϕ_i by M and perform a nonlinear transformation F on the amplitude ρ_i [8]. As the result we obtain the signal

$$r'_{Ii} + jr'_{Qi} = F(\rho_i)\exp(jM\phi_i)$$

The phase of the resulting signal is

$$M\phi_i = M\alpha_i + M\theta_i = 2k_i\pi + M\theta_i$$

The phase error θ is then estimated over 2N + 1 symbols as

$$\theta = \frac{1}{M}\tan^{-1}\frac{\sum_{i=-N}^{N}\rho_i sinM(\epsilon_i+\varphi_i)}{\sum_{i=-N}^{N}\rho_i cosM(\epsilon_i+\varphi_i)} \quad (1)$$

θ represents the phase by which the corresponding symbol is rotated.

The above operations cause an M-fold ambiguity in the phase estimate. It can be resolved by differential decoding for M-PSK modulation and $2\pi/M$ phase rotationally invariant M-PSK modulation. If non-rotationally invariant codes are used the phase ambiguity can be removed by monitoring Viterbi decoder metrics.

4 Phase Estimation in DD Technique

The DD is a recursive algorithm which is used for both phase and symbol timing estimation [4]. The algorithm operates under the assumption that the incoming symbols are known to the receiver. This condition is met for small error probabilities or alternately if some known pattern sequence is transmitted in a training period.

If the current phase estimate is θ_i the next symbol phase rotation is determined as

$$\theta_{i+1} = \theta_i + \Gamma im\{r_i\hat{r}_i^*\} \quad (2)$$

where Γ is a small number determined to achieve a trade-off between the phase error and the convergence rate and $\hat{r}_i$ is the ith sample estimate.

The symbol sampling time is determined by using a second order digital filter. The sampling time sequence $\{\tau_i\}$ is determined as

$$\tau_{i+1} = \tau_i + \lambda \mathrm{Re}\{r_i \hat{r}^*_{i-1} - r_{i-1}\hat{r}^*_i\} \quad (3)$$

where λ is a small number.

The phase error can be approximated as

$$\theta_{i+1} \simeq \Gamma(\epsilon_i + \varphi_i) \quad (4)$$

where ϵ_i is the phase noise induced by Gaussian noise and φ_i is the phase jitter.

In block structered data transmission the performance of the algorithm can be improved by sending a preamble at the beginning of each block [5].

5 Phase Estimation in TTIB Technique

The filtering in the TTIB algorithm is achieved by using two bandpass filters with a combined frequency response equivalent to a square root raised cosine transfer function [7]. Each bandpass filter is used to isolate one of two subbands with bandwidths equal to one half of the symbol transmission rate. The subband signals are further frequency translated to enable the insertion of the pilot tone. The bandstop margin is chosen large enough to accommodate the worst case Doppler shift and meet the acquisition time requirements of the receiver. The reference tone is positioned centrally within the signal bandwidth to assure an accurate estimation of the signal degradation.

In the TTIB demodulator the received pilot tone is used to remove phase error. This process is accomplished by computing the pilot tone phase and multiplying each symbol by the negative pilot tone phase. It is essential to choose an appropriate pilot tone level and filter bandwidths . As the pilot tone level increases more power is waisted on the tone transmission. On the other hand if the pilot tone power is reduced, it will degrade its signal-to-noise ratio and phase estimate. Similarly, the pilot tone bandwidth has an optimum. In order to reduce the impact of Gaussian noise it is desirable to have a small pilot tone bandwidth. However, the bandwidth is determined by maximum Doppler shifts in the channel to enable pilot tone extraction.

The insertion of pilot tone, however, will result in a non-constant envelope signal and degrade the performance of this algorithm on a nonlinear channel. The BER degradation on a satellite channel with a TWT nonlinear amplifier is of the order of 0.5 dB [1].

6 Probability of Error for QPSK

The conditional bit error probability for QPSK modulation given the fading attenuation due to shadowing a and the phase jitter θ is given by

$$P_{bc} = Q\left(\frac{a\cos\theta}{\sigma}\right)$$

where σ is the variance of the Gaussian noise in each signal space coordinate.

The total bit error probability is given by

$$P_b = \int_0^\infty \int_{-\pi}^{\pi} Q\left(\frac{a\cos\theta}{\sigma}\right) p(a)P(\theta)da d\theta$$

where p(a) and p(θ) are the probability distributions of fading attenuation and phase jitter, respectively.

7 Error probability for TCM Schemes

Let C denote the set of all possible coded signal sequences $\{z_N\}$, where z_N is a coded symbol sequence of length N:

$$z_N = (z_1, z_2, \ldots, z_i, \ldots, z_N)$$

Let us assume that v_N:

$$v_N = (v_1, v_2, \ldots, v_i, \ldots, v_N)$$

an element of C, is the transmitted sequence. Then v_i represents the MPSK symbol transmitted at time i.

$$v_i = v_{Ii} + jv_{Qi}$$

where v_{Ii} and v_{Qi} are I and Q channel components, respectively.

For codes with large minimum number of symbols on error paths or for any code at small Eb/No the pairwise error probability can be upperbounded by [9]

$$P(v_N \to z_N) = P_{d,D,T}$$

$$\leq \frac{1}{2}\frac{2\sigma d}{\sqrt{4\sigma^2 d^2 + \sigma_x^2}} \exp\left(-\frac{1}{2}\frac{m_x^2}{4\sigma^2 d^2 + \sigma_x^2}\right) \quad (5)$$

where

$$m_x = d^2 m_a m_{\theta 1} + T m_a m_{\theta 2}$$

$$T = \sum_{i=1}^{N} T_i \sum_{i=1}^{N} (z_{Ii} v_{Qi} - z_{Qi} v_{Ii})$$

and

$$\sigma_x^2 = D(m_{\theta 1^2} - m_a^2 m_{\theta 1}^2) + T_2(m_{\theta 2^2} - m_a^2 m_{\theta 2}^2)$$

$$D = \sum_{i=1}^{N} d_i^4$$

$$T_2 = \sum_{i=1}^{N} T_i^2$$

$$m_{\theta 1} = E[\cos\theta]$$

$$m_{\theta 1^2} = E[\cos^2\theta]$$

$$m_{\theta 2} = E[\sin\theta]$$

$$m_{\theta 2^2} = E[\sin^2\theta]$$

The bit error probability can be upperbounded by

$$P_b \leq \frac{1}{k} \sum_{d,D} b_{d,D,T} P_{d,D,T} \qquad (6)$$

where $b_{d,D,T}$ is the total number of erroneous information bits on all paths characterized by the Euclidean distance d, and the parameters D and T.

8 Results

The bit error rate performance with the QPSK modulation and FF algorithm relative to the demodulation without carrier recovery on a Gaussian channel with variable degrees of slowly varying phase jitter is given in Fig. 1 The algorithm removes almost completely the effect of phase jitter up to phase jitter values of 30°. The corresponding results for the 16-state Ungerboeck TCM code are illustrated in Fig. 2. The algorithm operates succesfully up to phase jitter values of 15°. The impact of slowly varying phase jitter on the DD algorithm on a Gaussian channel is illustrated in Fig. 3. The algorithm is able to recover the signal phase up to 40° phase jitter. The BER for larger phase jitter exhibits high variations due to unreliable estimates of the signal phase. For a Gaussian channel with relatively low BER the algorithm works reliably and there is no need for transmiision of preambles.

The performance of the TTIB technique depends on the pilot tone bandwidth. The pilot tone bandwidth should be chosen to match worst case Doppler shifts. For the ratio of the signal to the pilot tone bandwidth $\beta < 50$ the degradation is high. The sensitivity of the TTIB technique to phase jitter for QPSK and TCM is shown in Fig. 4. The TTIB algorithm is can remove slowly varying phase errors, but cannot eliminate the effect of Gaussian noise due to finite pilot tone bandwidth. The loss for the QPSK is about 1dB for $\beta = 50$. For all TTIB tests the power of the pilot tone level was 10 dB below the total output power.

The comparative results of the three algorithms for the QPSK scheme on a satellite mobile shadowed channel are presented in Figure 5. The performance is evaluated by the average block error rate (BLER) which is a suitable parameter for digitized voice transmisssion, (the block length corresponds to 30ms frames). The DD algorithm requires accurate esimates of the received signal and the error performance is degraded by heavy shadowing. The additional information from a known received sequences can improve the reliability of signal estimation. Since the voice codec generally provides preambles at the beginning of each transmitted block of data it is possible to include this information in the DD algorithm. It should be noted that most degradation in the error rate comes from the amplitude variations rather than phase errors, since phase errors in the channel were generally less than 10°. Therefore, carrier recovery methods could not achieve significant performance improvement. Clearly, the TTIB technique shows the superior performance. However, an additional loss of about 2dB should be added due to the pilot tone transmission, nonlinearity and Doppler shift degradations.

The corresponding results for the 16-PSK TCM code is shown in Fig. 6.

9 Conclusions

An investigation of carrier recovery techniques on satellite mobile channels with shadowing has been performed.

The TTIB algorithm achieves the excellent phase estimates for low pilot tone filter bandwidth. accommodate the worst case Doppler shifts.

The DD algorithm with preamble transmission offers good performance, while the implementation is simple and the algorithm with the second order loop provides symbol timing and frequency tracking which is critical on mobile channels.

10 Acknowledgements

This work was carried with support from IR&D Board Grant No. 17001.

References

[1] Nicolas, J.J. and Vucetic B., Performance of MPSK Trellis Codes over Nonlinear Fading Mobile Satellite Channels, International Conference on Communication Systems, Conference Proceedings Singapore ICCS'88, pp.695-699, Singapore, November, 1988.

[2] Data Files with I & Q data, AUSSAT

[3] J.P. McGeehan and A.J. Bateman, Phase-Locked Transparent-Tone-in-Band (TTIB): A New Spectrum Configuration Particularly Suited to the Transmission of Data Over SSB Mobile Radio Networks, IEEE Trans. Comm. Techn., Vol. COM-32, Jan. 1984., pp. 81-87.

[4] H. Kobayashi, Simultaneous Adaptive Estimation and Decision Algorithm for Carrier Modulated Data Transmission Systems, IEEE Trans. Commun. Tech., pp. 268-280, June 1971.

[5] H.C. Guren, Burst Channel Receiver for Aeronautical Communications, ASSPA'89, Adelaide, Australia, 1989.

[6] G. Ungerboeck, Channel Coding with Multilevel/Phase signals, IEEE Trans. Inform. Theory, vol. IT-28, Jan. 1982, pp. 55-67.

[7] J.H. Lodge, M.L. Moher and S.N. Crozier, A Comparison of Data Modulation Techniques for Land Mobile Satellite Channels, IEEE Trans. on Vehicular Techn., vol. VT-36, No. 1., Feb. 1987, pp. 28-35.

[8] A.J. Viterbi and A.M. Viterbi, Nonlinear Estimation of PSK-Modulated Carrier Phase with Application to Burst Digital Transmission, IEEE Trans. Inf. Theory, vol. IT-29, No. 4, pp. 543-551.

[9] B. Vucetic and J. Nicolas, "Construction of M-PSK Trellis Codes and Performance Analysis over Fading Channels", *ICC'90*, Atlanta, USA, April 1990.

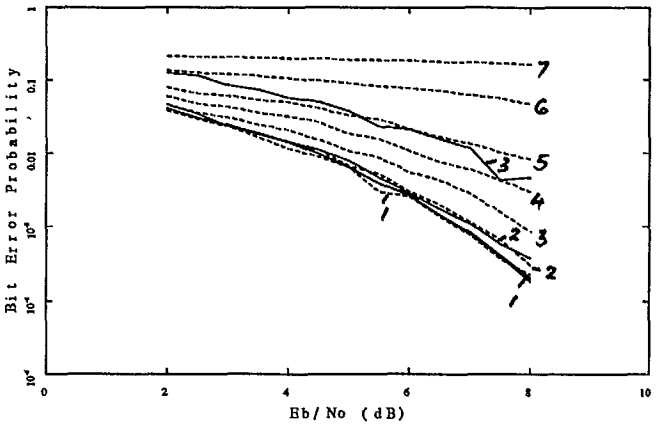


Figure 1: BER performance of QPSK with phase jitter on an AWGN channel:
Dash: without FFD; 1 to 7: 0, 5, 10, 15, 20, 30 and 40 degrees phase jitter
Solid: with FFD; 1: 5 to 20 degrees; 2: 30 degrees; 3: 40 degrees phase jitter

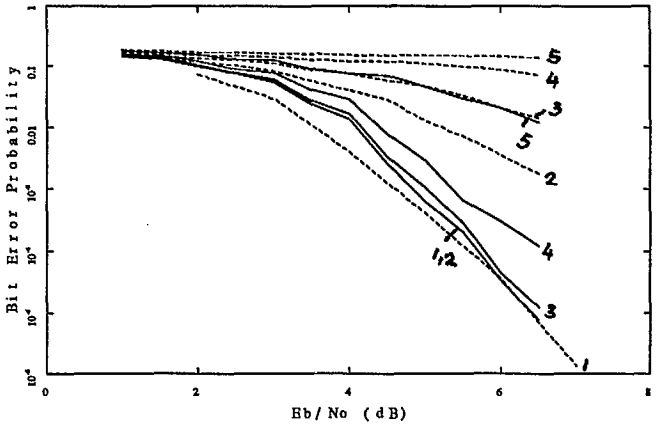


Figure 2: BER performance of TCM with phase jitter on an AWGN channel:
Dash: without FFD, Solid: with FFD ; 1 to 5: 0, 5, 10, 15 and 20 degrees phase jitter

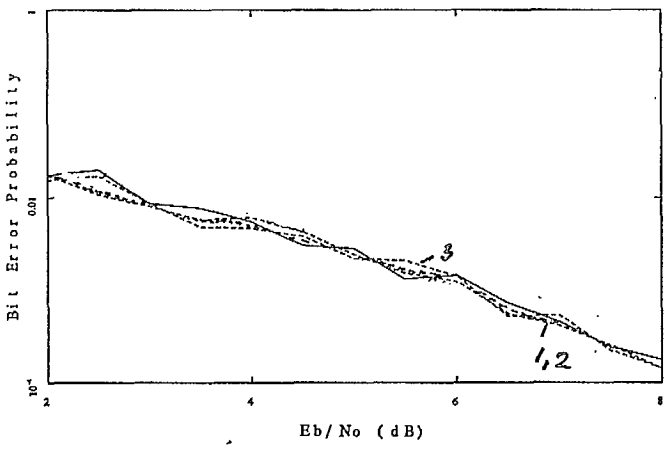


Figure 3: BER performance of the QPSK on an AWGN channel with phase jitter and DD scheme; Solid: Ideal QPSK; Dash: 1: phase jitter=10° 2: phase jitter=20° 3: phase jitter=30°

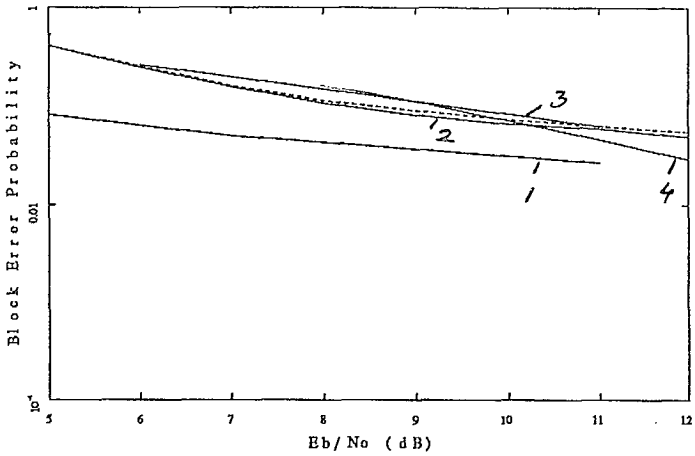


Figure 5: BLER performance of the QPSK modulation with the various carrier recovery schemes for experimental data; Dash: No carrier recovery; Solid: 1: TTIB 2: FF 3: DD 4: DD with preamble

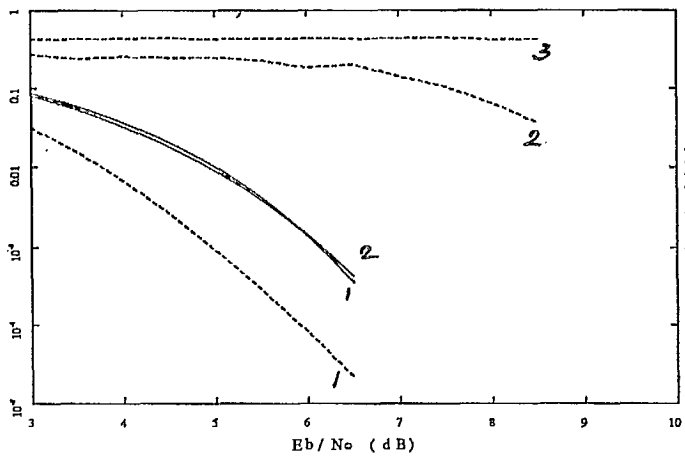


Figure 4: BER performance of the 16-state TCM code on an AWGN channel with phase jitter and TTIB scheme; $\frac{B}{B'} = 50$
Dash: No phase recovery 1: phase jitter=0° Ideal TCM; 2: phase jitter=30° 3: phase jitter=50° Solid: TTIB 1: phase jitter=30° 2: phase jitter=50°

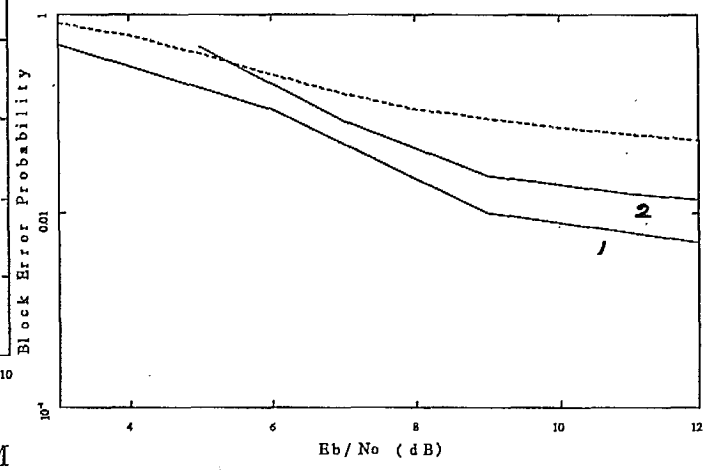


Figure 6: BLER performance of the 16state Ungerboeck TCM code carrier recovery for experimental data; Dash: Uncoded QPSK with ideal phase recovery; Solid: 1: TTIB 2: FF

Performance Evaluation of a Mobile Satellite System Modem using an ALE Method

Tomoki Ohsawa
C&C System Research Labs.
NEC Corporation
1-1, Miyazaki 4-chome
Miyamae-ku, Kawasaki-city 213 Japan
Phone: 81-44-856-2255 Fax: 81-44-856-2061

Motoya Iwasaki
Microwave & Satellite Communications Division
NEC Corporation
4035, Ikebe-cho, Midori-ku
Yokohama-city, 226 Japan
Phone: 81-45-939-2104 Fax: 81-45-939-2117

Abstract

This paper presents experimental performance of a newly designed demodulation concept. This concept applies an Adaptive Line Enhancer (ALE) to a carrier recovery circuit, which makes pull-in time significantly shorter in noisy and large carrier offset conditions. This new demodulation concept has actually been developed as an INMARSAT standard-C modem, and has been evaluated. On a performance evaluation, 50 symbol pull-in time is confirmed under 4dB Eb/No condition.

1 Introduction

Mobile Satellite Communications systems (MSS) have a lot of special characteristics especially in the demodulation field. First, received carrier frequency ambiguity, relative to symbol rate,f_b, becomes very large for low bit rate communications system such as MSAT or INMARSAT. Secondly, time variant channel characteristics, such as multipath fading and shadowing, cause signal amplitude and phase variations and signal dropouts. Thirdly, there are frequency change, caused by mobile acceleration or slowdown and oscilator instability. Therefore, in addition to low Eb/No operation, to maintain its carrier synchronization in the above mentioned conditions, a new carrier recovery circuit is required.

Considering INMARSAT standard-C specifications, the modem has to operate under 4dB or lower Eb/No, large carrier frequency offset ($\pm$2kHz), larger than signal baud rate (1200bps), 65Hz/sec carrier frequency change rate, and 7dB C/M deep multipath fading environment.

Generally, a Phase Locked Loop (PLL) is used as a carrier phase synchronizer. In low bit rate communications, however, pull-in and robust performance for a PLL demodulator becomes poor with low Eb/No. Therefore, a new coherent demodulation method, superior than conventional, is anticipated.

This paper presents a new coherent demodulation concept, referred to as an Adaptive Carrier Estimation DEMOdulator (ACE-DEMO)[1]. The ACE-DEMO makes pull-in time significantly shorter and has a wide pull-in range. This paper reports developed hardware performance, operating to INMARSAT standard-C specifications.

2 Carrier Recovery using ALE

2.1 ACE-DEMO Concept Figure 1 shows a proposed carrier recovery blockdiagram of ACE-DEMO[1]. The received M-PSK signal is converted to base band complex signal, $r(t)$, by a fixed frequency oscillator. Then the signal is changed to a time discrete signal, $r(i)$, by a sampler whose interval is equal to a signal baud interval (T_b). Received signal components are Additive White Gaussian Noise (AWGN) and modulation signal, whose center frequency shifts from nominal value by a carrier frequency offset value, ω. Its spectrum is shown at 1 in Fig.1. In order to extract the carrier signal, received complex signal, $r(i)$, is converted to frequency-multiplied complex signal, $x(i)$ by a multiplier. Its operation removes the modulation signal component from signal $r(i)$. Signal, $x(i)$ spectrum, as shown

at 2 in Fig.1, includes a line spectrum located at the frequency of M times carrier frequency offset value, ω_0.

An ALE is a kind of adaptively tuned, high Q, narrow BPF, which enhances the line spectrum in the AWGN. Therefore, AWGN is suppressed at its output signal, $y(i)$. Its spectrum is shown at 3 in Fig.1. To recover the original carrier signal, $z(i)$, $y(i)$ is frequency-divided to $1/M$ by a frequency divider. The $z(i)$ spectrum is shown at 4 in Fig.1.

In the multiplier, received signal, $r(i)$, and complex conjugate signal of recovered carrier signal are multiplied. Finally, demodulated PSK signal, $s(i)$, is obtained.

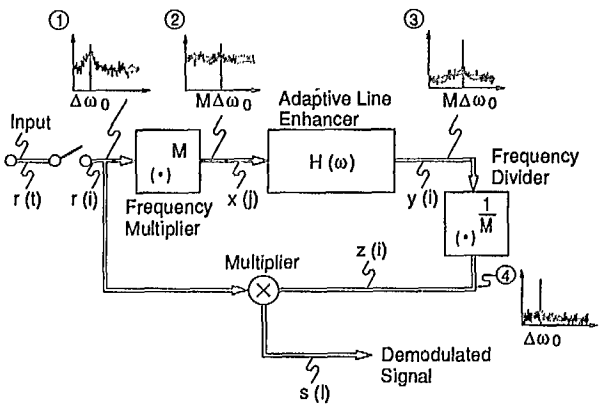


Fig. 1 ACE-DEMO Functional Blockdiagram

2.2 Adaptive Line Enhancer The ALE has developed for an adaptive noise canceler which operates as a self-tuning filter[2]. Figure 2 shows the ALE blockdiagram. the ALE consists of a delay operator, a Finite Impulse Response (FIR) filter, an adder, an adaptation constant multiplier and a coefficients controller. The delay operator, whose delay is set to T_d ($=mT_b$; 'm' is an integer number), accomplishes decorrelation between signal $x(i)$ and $y(i)$, except for the carrier signal component.

The FIR filter, a transversal filter, has L coefficients ($c_1 - c_L$) and the output signal $y(i)$ is expressed as

$$y(i) = C^T X_i, \tag{1}$$

where $C = [c_1, c_2, ...c_L]^T$; coefficient vector,
$X_i = [x(i-T_d), x(i-1-T_d), ...x(i-L-T_d)]^T$,
A^T is a transpose of vector A

Error signal, $e(i)$, derived from the difference between $x(i)$ and $y(i)$, is

$$e(i) = x(i) - y(i). \tag{2}$$

Coefficient vector C is controlled so that the mean value of $e(i)X_i$ may maintain zero. Then, the ALE becomes a narrow BPF which enhances the carrier signal, multiplied M.

There are many algorithms for controlling the coefficient. In this paper, Least Mean Square (LMS) algorithm, which minimizes a mean square error for $e(i)X_i$, is used. In this algorithm, the coefficient vector is controlled in the following manner.

$$C_i = C_{i-1} - \mu e(i) X_i^*, \tag{3}$$

where μ is the adaptation constant and C_i is $i-th$ coefficient vector.

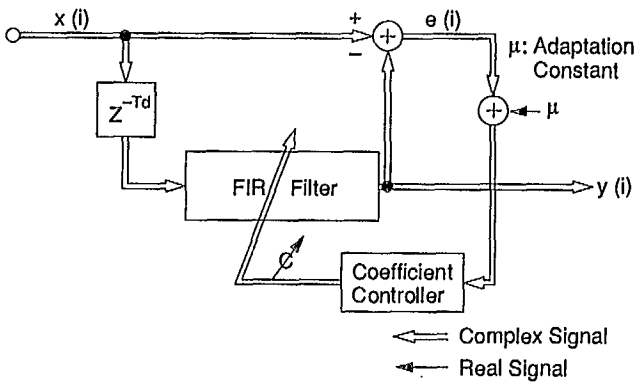


Fig. 2 ALE Blockdiagram

2.3 Pull-in Analysis In the previous analysis[1], ACE-DEMO pull-in equation is expressed as follows,

$$i_{ACE} = -0.3N_x/N_{req} log_e(0.17N_{req}/N_x). \tag{4}$$

$$L = 1.2N_x/N_{req} \tag{5}$$

$$\mu_o = 1 \tag{6}$$

where i_{ACE} is a number of symbols for pull-in,
L is a number of taps for the ALE[2],
μ_o is a normalized adaptation constant[1][2],
N_x is input noise power to the ALE
and N_{req} is required output noise power.

Pull-in symbol, i_{ACE}, is evaluated quantitatively and compared with that of conventional PLL-DEMO. Figure 3 shows a pull-in characteristics comparison between ACE-DEMO and PLL-DEMO[1]. In the comparison, each phase jitter is equally set to 0.16[rad], at 0dB Eb/No condition. In the figure, the PLL rapidly increases the pull-in symbol by 2nd power of carrier frequency offset, ω. Even if there is no carrier frequency offset, non-linear characteristics for the PLL phase detector makes acquisition longer, and it is observed that the PLL pull-in symbol is about three times longer than that of the ACE-DEMO. ACE-DEMO has constant pull-in symbol. Its pull-in frequency range, F_{pir}, is,

$$F_{pir} = \pm f_b/2M \qquad (7)$$

where f_b is signal baud rate.

Therefore, the ACE-DEMO realizes both wide pull-in range and short pull-in time simultaneously.

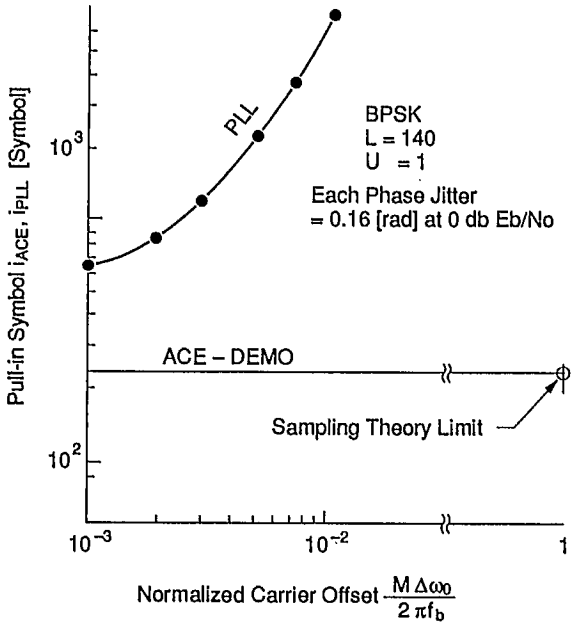


Fig. 3 Pull-in Range Comparison Between PLL and ACE – DEMO

3 Hardware Design

A demodulator using the ALE method has been developed for the INMARSAT standard-C (STD-C) coast earth station system. Table 1 and Figure 4 show STD-C specifications and blockdiagram, respectively. By substituting f_b=1200Hz and M=2 into Eq.(7), ±300Hz pull-in range for ACE-DEMO is derived. Since the carrier frequency offset value ±2kHz as shown in Table1 exceeds this pull-in range limit, a rough carrier frequency offset estimation is necessary. This estimation is carried out by the upper block of Fig 4, which includes a *filterbank* to improve the S/N ratio, a *frequency-multiplier* to remove the modulation component, a *FFT*, and a *max power detector* in which the carrier frequency offset can be detected as a single peak. The clock timming is estimated by the middle block, include a *LPF*(low pass filter) to extract the clock signal component and a *correlator* to calculate the clock phase error[3].

Input signal is first frequency-shifted by the estimated carrier frequency offset. This frequency-shifted signal is filtered by a *I&D* (Integrate and Dump) *filter* according to the estimated clock timming and fed into the *ACE-DEMO*.

Table 1 STD-C specification

moduration ;	unfiltered BPSK
symbol rate ;	1200baud or 600baud
carrier frequency uncertainty ;	±2kHz
carrier frequency change rate ;	65Hz/sec
signal to noise ratio ;	Eb/No = 4.7dB
multipath Rician fading ;	
	C/M =7dB pitch = 0.7Hz
packet structure (signaling channel)	
unique word ;	64symbol
data length ;	252symbol

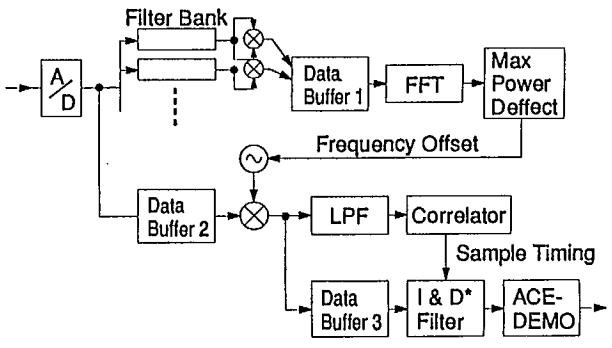


Fig. 4 STD-C Hardware Functional Blockdiagram

Since no preample sequence are used for the signaling channel of STD-C, a switch-back preambleless demoduration technique [4] is adopted. The data is once stored to a data buffer and fed into the ACE-DEMO from the end of the data in reverse order(the backward processing). When the carrier is recovered and the unique word is detected, the data is fed into the ACE-DEMO from the beginning of the data again(the forward processing). The backward proseccing is for carrier frequency acquisition, therefore the acquisition must be achieved within 252 symbols, the packet data length. The data is demodulated in the forward processing.

Figure 5 shows the developed hardware board. It is composed of single chip DSP(TMS320C26), instruction ROM, data RAM, and I/O ports. Number of ALE taps is 128.

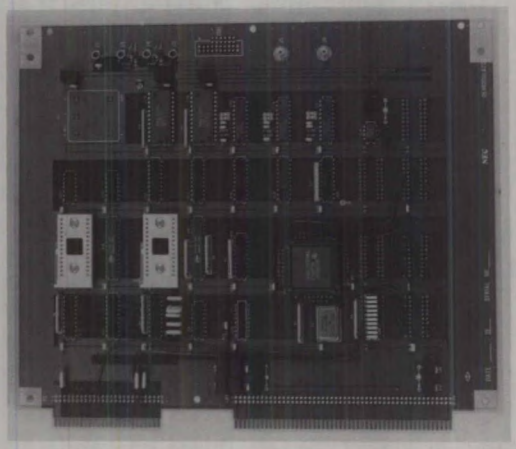


Fig.5 Developed Hardware Board

4 Measured Performance

4.1 BER performance Figure 6 shows the measured BER performance of the developed hardware. In this measurement, Rician fading factor is set to 15,10,7dB C/M (Carrier to Multipath power ratio) and 0.7Hz fading pitch. As shown in this figure, it is obvious that the developed demodulator tracks the carrier changes due to Rician fading with low Eb/No condition.

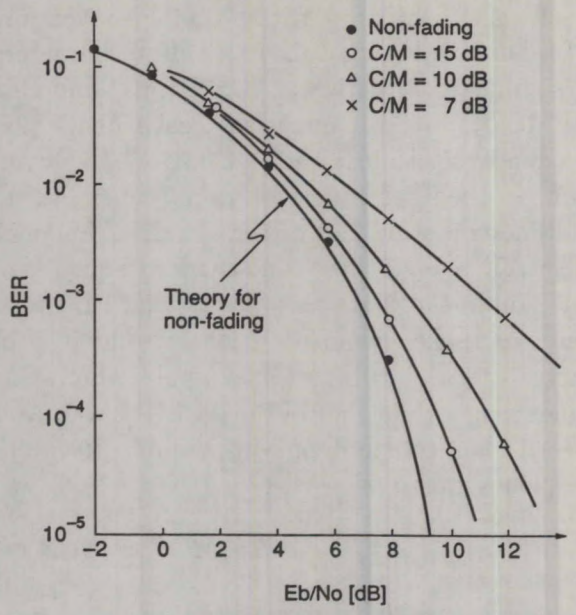


Fig. 6 BER Performance

4.2 ACE-DEMO Pull-in Performance Figures 7 and 8 show the measured pull-in performance for the signaling channel. In these measurements, in order to evaluate the performance of the ACE-DEMO itself, the rough carrier frequency offset estimation is not used and the clock timming is extracted by the transmission clock.

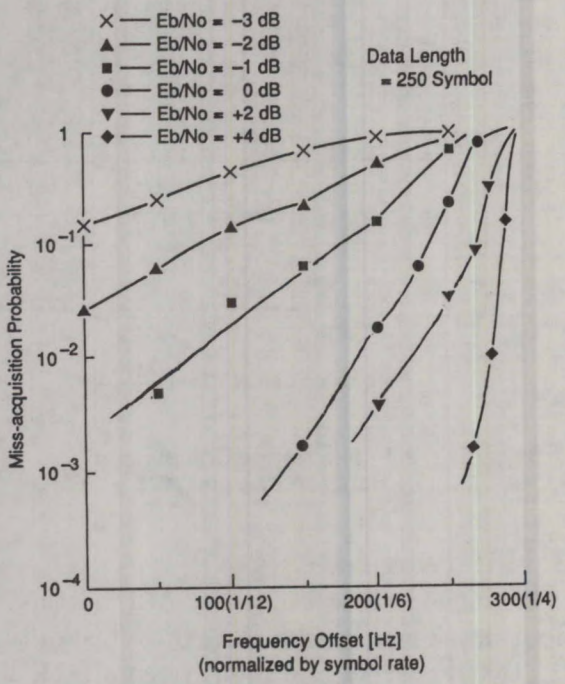


Fig. 7 Pull-in Performance (1)

The packet structure is the same as a STD-C signaling channel.The pull-in is decided when the unique word is detected.

Figure 7 shows the relationship between miss-acquisition probability and carrier frequency offset. In principle, the ACE-DEMO pull-in characteristic has no relation with the carrier offset up to ±300 Hz as mentioned before, but in this mesurement, it degrades as carrier frequency offset becomes large. It is due to fact that the phase jump occurs at the frequency divider of ACE-DEMO by the noise and that there are the power loss at the I&D filter caused by frequency offset. 270Hz ACE-DEMO pull-in range is obtained at 10^{-3} miss-acquisition probability under 4 dB condition, so that it has a much wider pull-in frequency range than the conventional PLL. For STD-C signaling channel application, the accuracy of rough carrier frequency offset estimation is allowed to this level.

Figure 8 shows the relationship between miss-acquisition probability and data length. The acquisition is achieved within 150 symbols for 0dB Eb/No and within 50 symbols for 4dB Eb/No. It is a much shorter pull-in time than the conventional method. It satisfies to STD-C signaling channel specifications

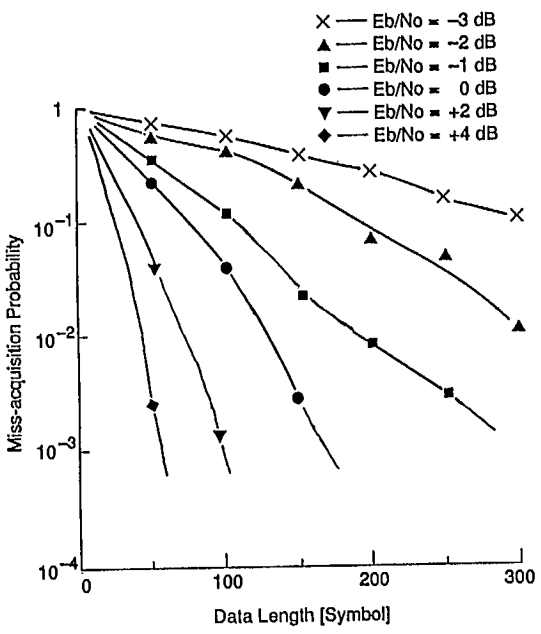


Fig. 8 Pull-In Performance (2)

5 Conclusion

This paper presents a new coherent demodulation concept, ACE-DEMO, based on ALE. The proposed demodulator has many merits. they include;

(1) wide pull-in frequency range,
(2) short pull-in time,
(3) pull-in time independent of carrier frequency,
(4) short pull-in time for reacquisition from outage.

Hardware using this concept has been actually developed for the INMARSAT standard-C coast earth station system, using a 320C26 (T.I.) DSP. In measurement evaluation, excellent BER performance is derived in the severe environment called for by STD-C specifications, which include Rician fading of 7dB C/M, and 4dB Eb/No. The ±270 pull-in frequency range and 50 symbol pull-in time of ALE method is confirmed, so that this demodulator presents attractive advantages especially when apply to a low bit rate communication systems such as STD-C signaling channel.

Acknowledgment

Authors wish to thank Mr. Furuya, Mr. Tomita and Dr. Hirosaki of C&C systems research Labs., NEC Corporation, for their helpful discussions. They also thank Mr. Otani of Microwave & Satellite Communication Division, NEC Corporation, for his assistance.

Reference

[1] **T. Ohsawa** , "Wide Range and Short Time Pull-in Coherent Demodulation Method for PSK signal –ACE DEMO– ", 304.4, Vol.1, PP , SUPERCOMM/ICC'90, April, 1990

[2] **B.Widrow** and **S. D. Stearns** , "Adaptive Signal Processing", Prentice-Hall, Inc., 1985, N.J.

[3] **H. Tomita** and **J. Namiki** , "Preambleless Demodulator for Satellite Communications", ICC'89 Conf. Rec., 16.4.1, pp. 504-508, June, 1989

[4] **T. Ohsawa** and **J. Namiki**, "Preambleless Demodulation Method", Conf. Rec. National Convention, IEICEJ, 2395, pp. 10-95, 1984

Session 12 - Vehicle Antennas

Session Chairman - *L. Shafai,* University of Manitoba, Canada
Session Organizer - *Ken Woo,* JPL

A Phased Array Tracking Antenna for Vehicles

Shingo Ohmori, Kazukiko Mano and Kenji Tanaka
Kashima Space Research Center
Communications Research Laboratory
Ministry of Posts and Telecommunications
Kashima, Ibaraki, 314 JAPAN
Phone:+81-299-82-1211, FAX:+81-299-83-5728

Makoto Matsunaga and Makio Tsuchiya
Mitsubishi Electric Corporation
Amagasaki, Hyogo, 661 JAPAN

ABSTRACT

An antenna system including antenna elements and a satellite tracking method is considered a key technology in implementing land mobile satellite communications. In the early stage of land mobile satellite communications, a mechanical tracking antenna system is considered the best candidate for vehicles, however, a phased array antenna will replace it in the near future, because it has many attractive advantages such as a low and compact profile, high speed tracking and potential low cost. Communications Research Laboratory is now developing a new phased array antenna system for land vehicles based on research experiences of the airborne phased array antenna , which was developed and evaluated in satellite communication experiments using the ETS-V satellite.

This paper describes the basic characteristics of the phased array antenna for land vehicles.

I. INTRODUCTION

An antenna system including antenna elements and a satellite tracking method is considered a key technology in implementing land mobile satellite communications. In the early stage of land mobile satellite communications, a mechanical tracking antenna system is considered the best candidate because of its characteristics such as a simple configuration, a wide beam coverage and an easy installation [1]. However, a phased array antenna will replace it in the near future, because it has many attractive advantages such as a low and compact profile, high speed tracking and potential low cost. On the other hand, it also has such disadvantages as lower G/T caused by complex feed lines, narrow beam coverage and so on.

Communications Research Laboratory is now developing a new phased array antenna system for land vehicles based on research experiences of the airborne phased array antenna [2], which was developed and evaluated in aeronautical satellite communication experiments using the ETS-V satellite [3].

This paper describes the basic characteristics of the phased array antenna for land vehicles.

II. MAIN FEATURES OF THE PHASED ARRAY ANTENNA

The main features of the phased array antenna are as follows:

(1) Antenna elements are excited by electromagnetic coupling with microstrip feed lines. Because of its easy configuration, production cost will be reduced greatly enough for land vehicles including private cars.

(2) A total number of digital phase shifters will be reduced to one-half using a newly developed phase shifter. The phase shifter is printed on a substrate, which will make the present complex feed lines simple. The development of a new phase shifter is in the first stage, where a 3-bit phase shifter with having only 6 PIN-diodes have been developed.

(3) Tracking error between transmitting and receiving frequencies, which is inevitable for a phased array antenna, is potentially eliminated using a frequency dependent phase shifter [4]. The new phase shifter is scheduled to be developed on a second stage of the project.

III. ANTENNA ELEMENT

An electromagnetic coupled antenna is adopted as an antenna element, because it will reduce procedures of assembling the antenna system and it will reduce a cost of the phased array antenna to compete with a conventional mechanical steering antenna. Figure 1 shows a configuration of the antenna element. A feed line is a microstrip line printed on a substrate, which excites a radiating element by electromagnetic coupling through a coupling aperture. A radiating element, which radiates circularly polarized waves, is also printed on a thin film substrate. Figure 2 shows a

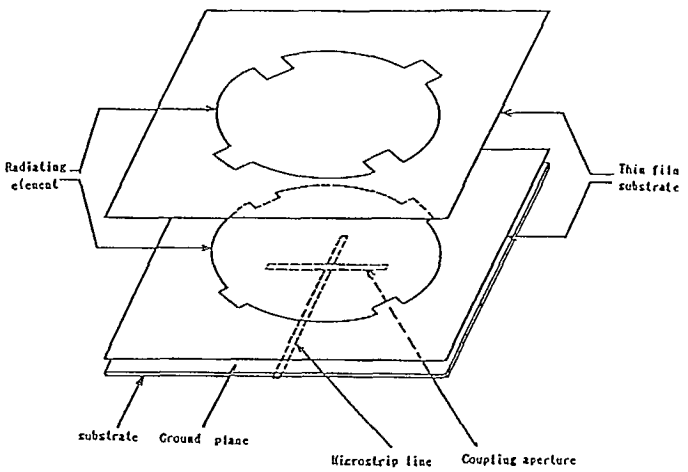


Fig.1 Configuration of Antenna Element.

frequency dependence of return loss of the element. The return loss in the frequency range of 1530-1660.5 MHz, which is required in mobile satellite communications, is found to be below 20 dB.

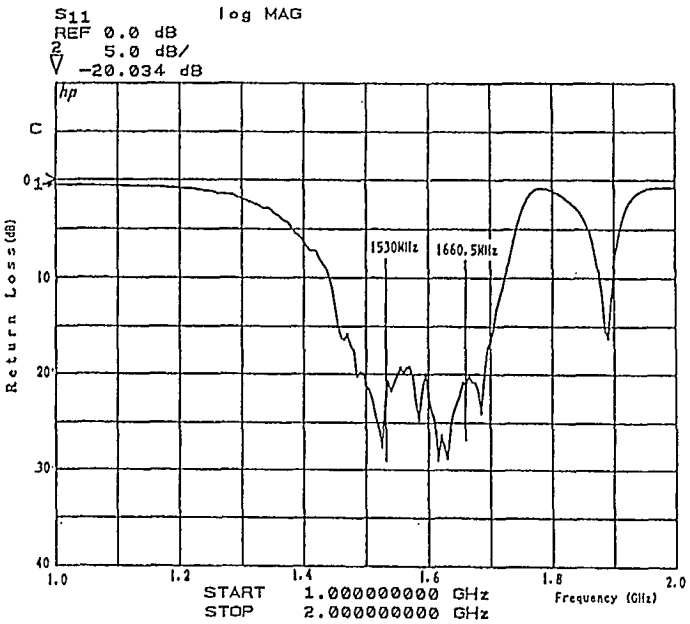


Fig.2 Return Loss of Antenna Element.

IV. PHASED ARRAY

The phased array antenna is under development and a prototype is assembled as shown in Fig. 3. A gain of the array antenna is designed to be 18 dBi in a non-scanned direction (Elevation angle: 90 deg.), and 10 dBi in a scanned angle of 60 degrees (Elevation angle: 30 deg.), which is determined by a system requirement of the ETS-V experiment. Main characteristics of the array antenna are shown in Table 1. The array consists of 19 elements, which are fed by microstrip feed lines as shown in Fig. 4. A theoretical calculation of gains and radiation patterns of the array antenna is shown in Fig. 5, where θ denotes a scanned angle from the boresight direction (Elevation angle: 90 deg). The gain is shown in a directive gain, which does not include losses of feed lines including phase shifters.

Fig.3 Prototype of the Phased Array Antenna.

Tabel.1 Main Characteristic of the Array Antenna

Frequency	f_R 1530.0MHz～1559.0MHz f_T 1626.5MHz～1160.5MHz
Polarization	Left Hand Circular
Scanned Angle	Ele. 30° ～ 90° Az. 0° ～360°
Gain	18dB: (Ele.90°) 10dB: (Ele.30°)
System Temp	200K
Axial Ratio	4dB(Ele.=30°)
Volume	60cmϕ ×4cm(H)
Weight	5kg

V. TRACKING METHOD

A tracking method of a phased array antenna is a very important key technology in land mobile satellite communications, and two different methods and an integrated method will be evaluated in the ETS-V experiment. First method is a closed loop system, which can track a satellite by receiving a signal from a satellite. Second

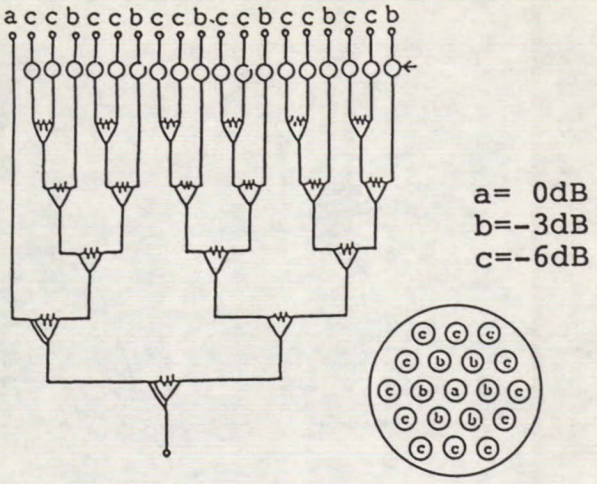


Fig.4 Configuration of Feed Lines.

system is an open loop method, which uses output signals from sensors such as magnetic compass and/or optical gyro. Table 2 shows comparisons of sensors which will be evaluated in the ETS-V experiments. An integrated method using mixture of open and closed methods is also evaluated.

Table.2 Sensors used in the Experiment

Sensor	Accuracy	Resolusion	Cost
Optical Gyro	±0.05° /sec	0.025° /sec	¥1,000,000
Magnetic Sensor	±0.1°	0.1°	¥420,000
Magnetic Compass	±10°	0.25°	¥9,000
Inclination Detector		0.6°	¥340,000
Angular Velocity Detector		0.02° /sec	
Vehicular Speed Detector	400pulses at 60km/h	1pule/6.28cm	

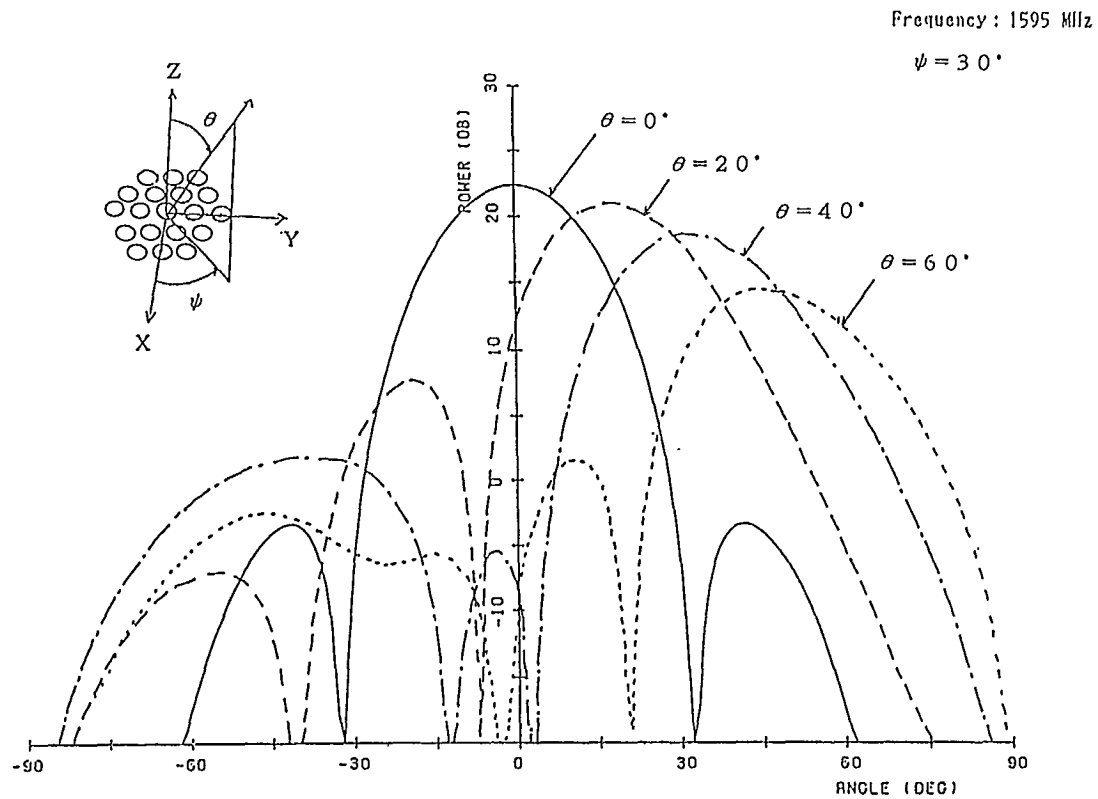


Fig.5 Calucurated Radiation Patterns and Gains.

VI. CONCLUSION

The phased array antenna for land vehicles is under development, and basic characteristics of a prototype are shown. The main aims of developing the phased array antenna are (1) to reduce the tracking error between receiving and transmitting frequencies, (2) to reduce a number of phase shifters in the feed lines and (3) to realize an array fed by electromagnetic coupling.

REFERENCES

[1] J. Huang, "L-Band Phased Array Antennas for Mobile Satellite Communications," IEEE Vehicular Technology Conference, Tampa, Florida, May 1987.

[2] S. Taira, M. Tanaka and S. Ohmori, "High Gain Airborne Antenna for Satellite Communications," To be published in IEEE Transactions on Aerispace and Electronic Systems.

[3] Y. Hase S. Ohmori and K. Kosaka, "Experimental Mobile Satellite System using ETS-V," IEEE Denshi Tokyo, No.25, 1986.

[4] S. Ohmori, S. Taira and M. Austin, "Tracking Error of Phased Array Antenna," To be published in IEEE Transactions on Antenna and Propagation.

MSAT Vehicular Antennas with Self Scanning Array Elements

L. Shafai
Dept. of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2
Tel. (204) 474-9615 FAX (204) 261-4639

ABSTRACT

This paper presents a new approach for designing low profile antennas for MSAT applications. It is based on stacking two microstrip antennas that operate at two adjacent modes. The beam scanning is achieved by introducing a phase shift between the stacked elements and consequently a low cost self scan array element is developed. The concept is used to investigate two different antenna types: 1) a single element unit, with its phase shifter, that provides a moderate gain of about 7 dBic, and 2) a seven element array with a peak gain of about 14 dBic. Computed and measured data for each design are presented and discussed

BEAM SCANNING CONCEPT

Assume N-stacked circular patches, where each operates at one of TM_{n1} modes. If the radiation is circularly polarized, the components of the far field pattern can be written as

$$E^c_{\theta,\phi} = \sum_{n=1}^{N} f_n(\theta)\, e^{in\phi} \qquad (1)$$

where

$$f_n^{\theta} = \frac{vak_0 j^n}{2} \frac{e^{-ikr}}{r} \left[J_{n+1}(k_0 a \sin\theta) - J_{n-1}(k_0 a \sin\theta)\right] \qquad (2)$$

and

$$f_n^{\Phi} = \frac{-vak_0 j^n}{2} \frac{e^{-ikr}}{r} \cos\theta \left[J_{n+1}(k_0 a \sin\theta) - J_{n-1}(k_0 a \sin\theta)\right] \qquad (3)$$

when the excitation of each TM_{n1} mode (i.e. each patch) has a relative phase difference of δ_n, equation (1) modifies to

$$E^c_{\theta,\phi} = \sum_{n=1}^{N} f_n(\theta)\, e^{in\phi - j\delta_n} \qquad (4)$$

and selecting $\delta_n = n\delta_1$ provides

$$E^c_{\theta,\phi} = \sum_{n=1}^{N} f_n(\theta)\, e^{in(\phi - \delta_1)} \qquad (5)$$

For equation (1) the beam peak is located in the $\phi = 0$ plane, and moves to $\phi = \delta_1$ plane for the pattern of equation (5). Now, if the phase differences of $n\delta_1$ are generated by external phase shifters, the array beam can be scanned by varying the relative phase of each stacked element. In particular, increasing the number of stacked elements reduces the beamwidth and consequently increases the array gain.

The above array formation by stacking the TM_{n1} patches also alters the radiation patterns in the θ- direction. Only the TM_{11} mode radiates axially in the $\theta = 0$ direction. All other TM_{n1} modes radiate conical beams, with a null at $\theta = 0$ direction. Consequently, the pattern peak of the stacked arrays occurs, somewhere between $\theta = 0$ and 90°, depending on the number of stacked elements and the substrate dialectric constant.

SINGLE STACKED ELEMENT

In the present study, since only two stacked microstrip patches are utilized, the operating modes are the TM_{11} and TM_{21} modes. Fig. 1 shows the geometry of the structure. Their individual radiation patterns, when each patch is over an infinite ground plane, are shown in Figs. 2 and 3. The substrate dielectric constant is $\epsilon_r = 2.52$. Consequently, the principal plane patterns are not identical, but can be made to be similar by selecting an appropriate ground plane size. For the stacked configuration the computed circularly polarized patterns, in the $\phi = 0$, 180 plane, are shown in Figs. 4a and 4b. The peak gain is 7.9 dBic, which occurs at about $\theta = 32°$, Fig. 4a, and the beamwidth in the θ-direction is about 30 degrees. The generated sidelobe is small, about -20 dB, and occurs at an angle of 70 degrees behind the main beam, $\phi = 180°$ plane. The horizontal copolar pattern, in a plane containing the beam peak, i.e. $\theta = 32°$, is shown in Fig. 4b. The beamwidth is about 180 degrees, and the pattern has no sidelobes.

An experimental unit was fabricated and tested. Fig. 5 shows the measured copolar patterns at two adjacent frequencies. The centre frequency of the patches was selected to be 3.165 GHz, to reduce the element size. The antenna ground plane was small, 4 in x 4 in, but the measured patterns still agree well with computations. The achieved peak gain of 7.56 dBic is also close to computed value of 7.9 dBic.

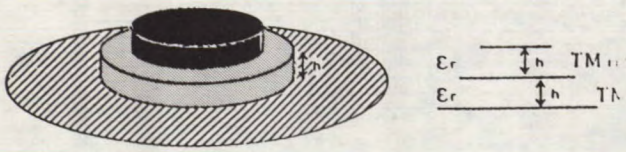
SEVEN ELEMENT ARRAY

Using the above stacked TM_{11}, TM_{21} mode element, a 7-element array was formed, that is shown in Fig. 6, and its performance, with and without mutual coupling effects was investigated. For the symmetric planes, i.e. $\phi = 0$ and 30°, the computed results are shown in Figs. 7 and 8. Fig. 7a shows the copolar pattern with and without mutual coupling in the $\phi = 0$ plane. For the selected dimensions the effect of mutual coupling is small and does not alter the gain or the pattern shape significantly. The beam peak is at $\theta = 47°$ (elevation of 43°), where the array gain is about 13.8 dBic and reduces to 11.0 dBic at the elevation angle of 22°. Since the ground plane is selected to be an infinite one, the co- and cross-polar fields become equal at the horizontal plane. The computed patterns in the $\theta = 47°$, are shown in Fig. 7b, and for the selected beam the sidelobe levels are reasonably high.

Figs 8a and 8b show the corresponding results when the beam is scanned to $\phi_0 = = 30°$, $\theta_0 = 60°$ direction. For this beam the grating lobe of the array falls outside the real space and the array gain and sidelobes improve. The beam peak occurs at $\phi = 30°$, $\theta = 48°$ (elevation of 42°), where the array gain is increased to 14.3 dBic. The copolar patterns in the $\theta = 48°$, are shown in Fib. 8b and again have reduced sidelobe levels.

The 7-element stacked array was fabricated and tested for performance at both beams. Sample of measured results are shown in Figs. 9 and 10, respectively for ($\phi_0 = 0$, $\theta_0 = 60°$) and ($\phi_0 = 30°$, $\theta_0 = 60°$) beams. The measured peak gain of these beams, excluding circuit losses, were at 12.34 dBic and 13.80 dBic, respectively, which are somewhat below the computed values. The discrepancies are partly due to the mismatch and polarization losses, and partly due to the phase errors and finite ground plane size. The effect of phase error is evident from the beam peak, where the pattern is nearly flat between 48° and 60° range.

To scan the array beam one needs two different sets of phase shifters. Array elements require seven corse phase shifters (2 or 3 bits), since their azimuthal pattern is broad and has a 3 dB beamwidth of about 180°. For interelement phasing, six finer (4 bit) phase shifters are necessary to scan the array beam. That is in total one requires 13 phase shifters, which is five less than the number needed for a 15-element TM_{11} mode array. Both arrays have similar gains for the 20° to 60° elevation range.



Geometry of Two Stacked Microstrip Antennas

Fig. 1.

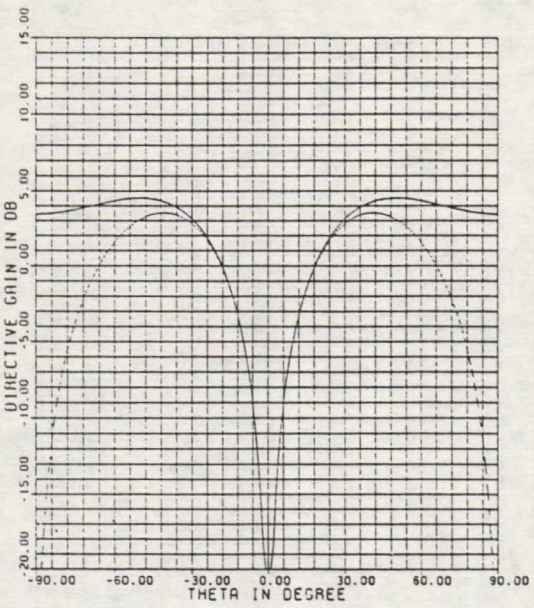


Fig. 3. E and H plane patterns for the TM_{21} mode patch,

——— E-plane, - - - H-plane

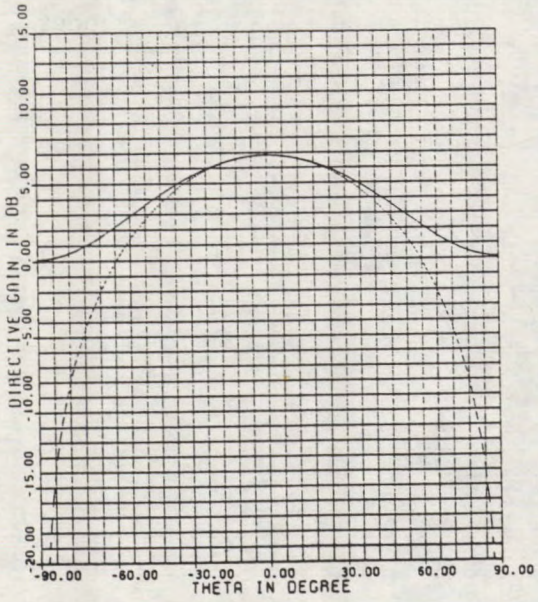


Fig. 2. E and H plane patterns of the TM_{11} mode patch over an infinite ground plane, $\epsilon_r = 2.52$

——— E-plane
- - - - H-plane

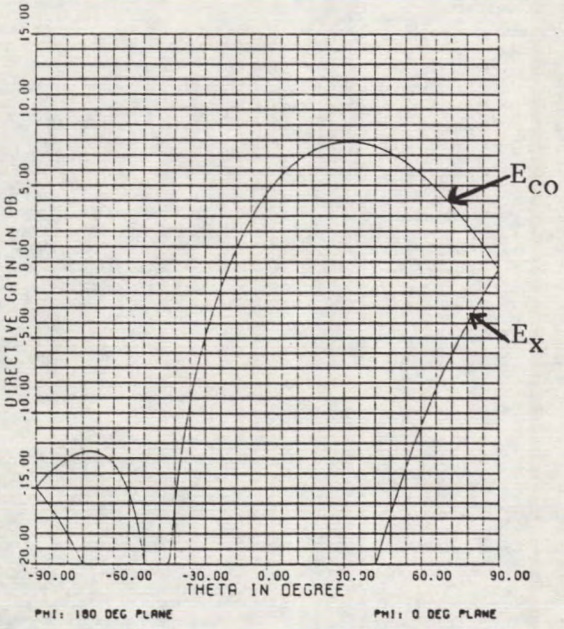


Fig. 4a. Co- and cross-polar (circular) patterns of stacked TM_{11} + TM_{21} modes.

ϕ = 0.0/180, $\epsilon_r = 2.52$

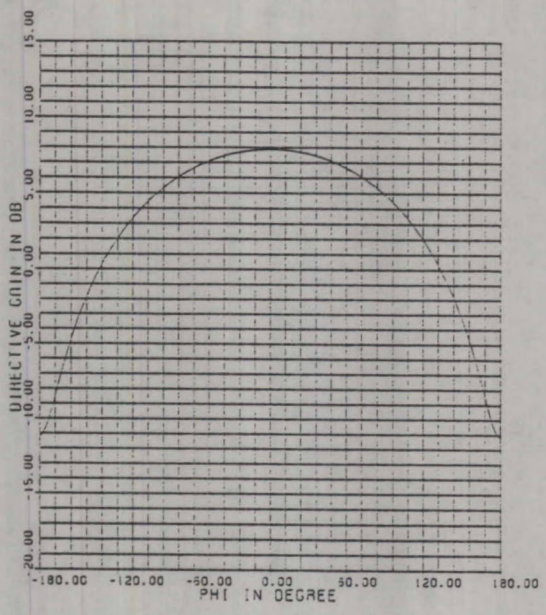


Fig. 4b. Co- and cross-polar (circular) patterns of stacked $TM_{11} + TM_{21}$ modes.

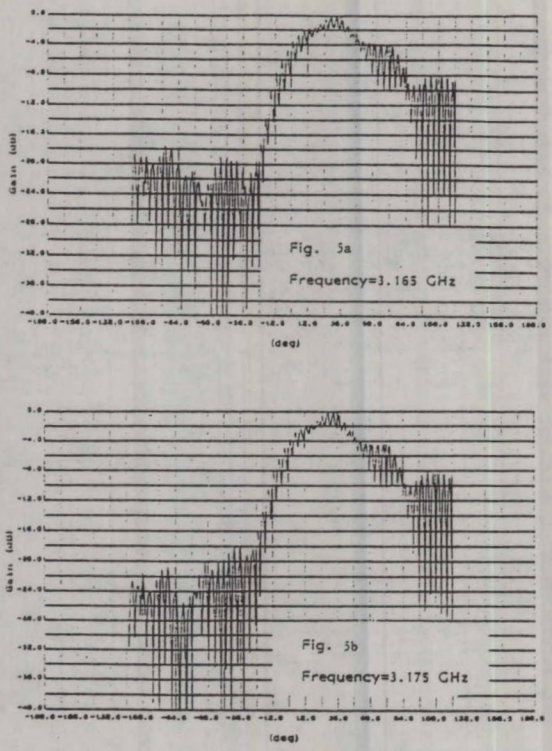


Fig. 5. Experimental patterns of stacked $TM_{11} + TM_{21}$ modes, $\epsilon_r = 2.52$, $h = 1.6$ mm, gain = 7.56 dBic

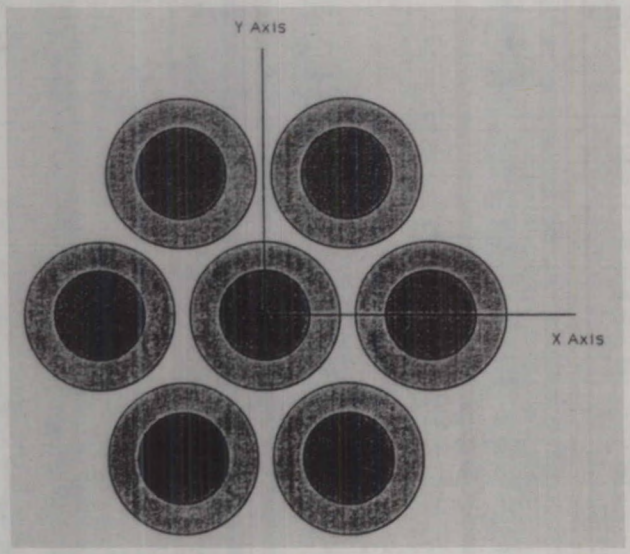


Fig. 6.

Triangular Array of Seven Dual Mode Microstrip Elements

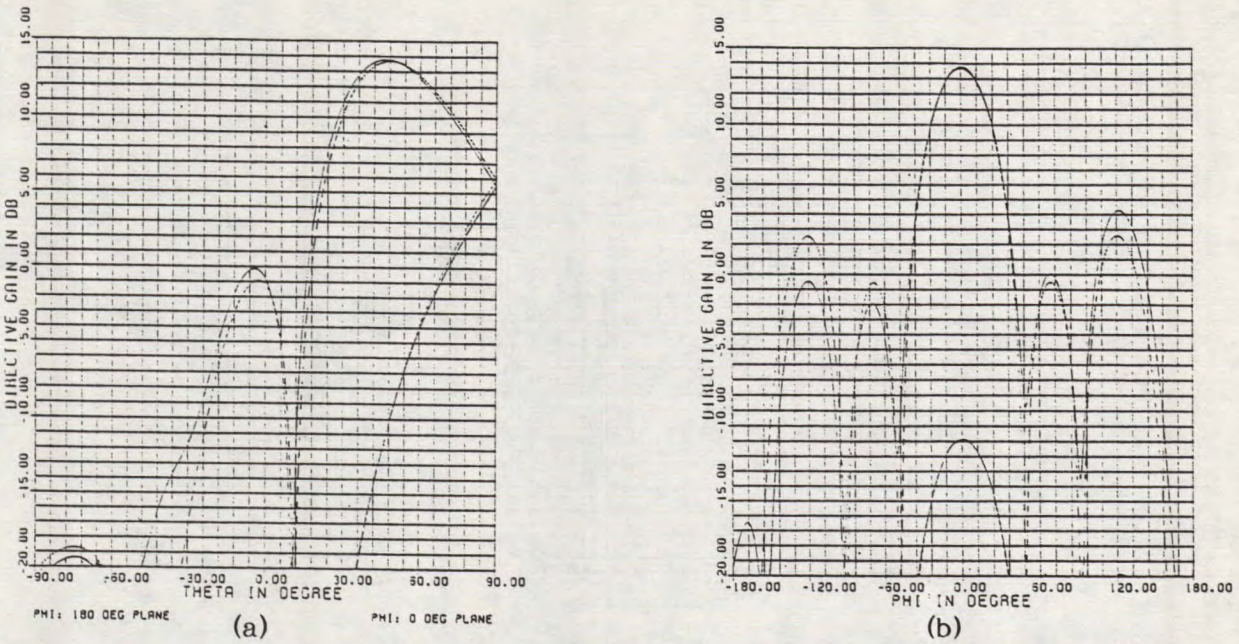


Fig. 7. Computed co-polar and cross-polar patterns of the 7-element array, $\epsilon_r = 2.52$, 0.0 degree beam ($\phi_o = 0.0$, $\theta_o = 60.0$)
——— without mutual coupling, - - - - - with mutual coupling

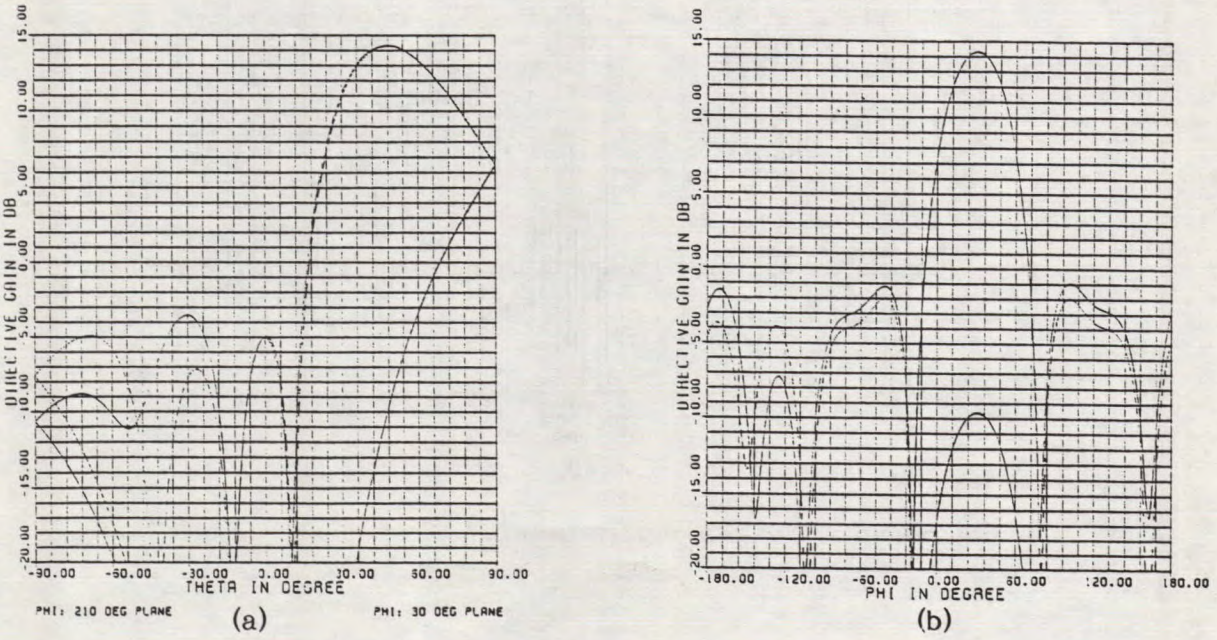


Fig. 8. Computed co-polar and cross-polar patterns of the 7-element array, $\epsilon_r = 2.52$, 30-degree beam ($\phi_o = 30.0$, $\theta_o = 60.$)
——— without mutual coupling, - - - - - with mutual coupling

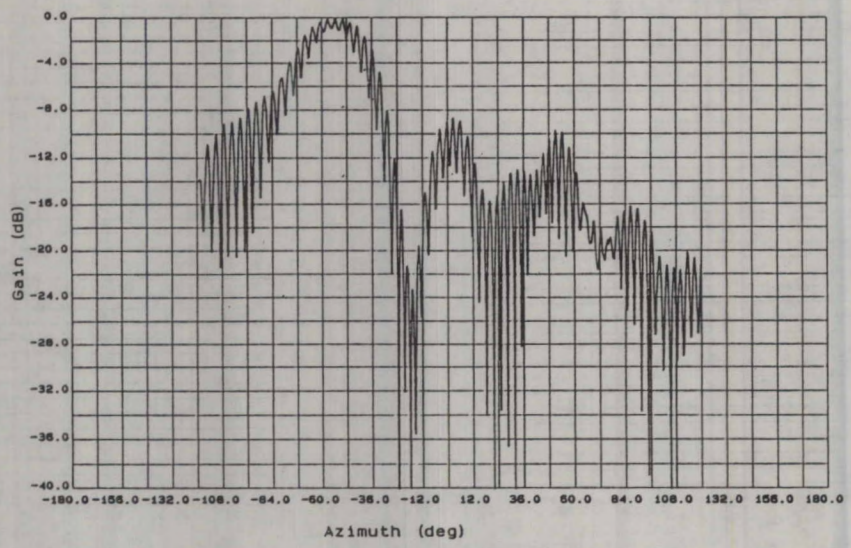


Fig. 9. (ϕ_o = 0.0, θ_o = 60.), gain = 12.34 dBic

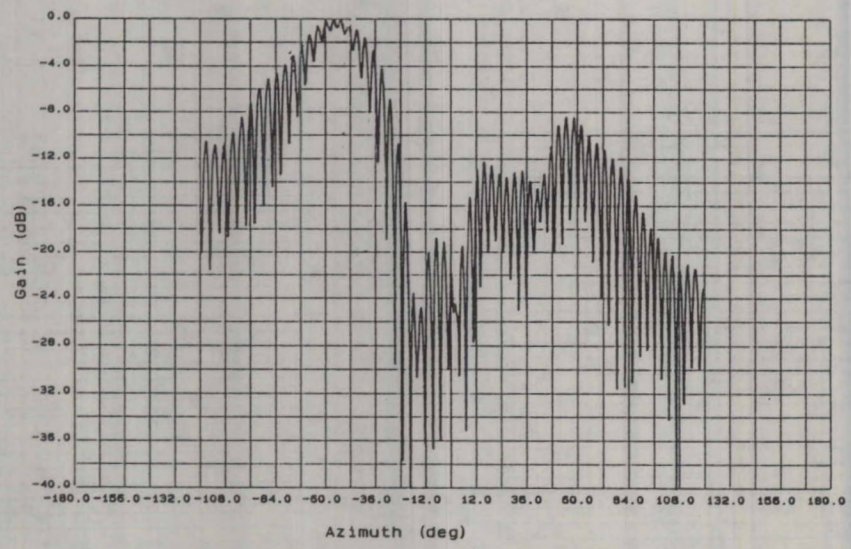


Fig. 10. (ϕ_o = 30.0, θ_o = 60.0), gain = 13.80 dBic

Fig. 9 and 10. Measured patterns of the 7-element array, ϵ_r = 2.52, f = 3.16 GHz

An Adaptive Array Antenna for Mobile Satellite Communications

Robert Milne
Communications Canada
Communications Research Centre
3701 Carling Avenue
Ottawa, Ontario K2H 8S2 CANADA
Phone: 613-998-2434
FAX: 613-990-0316

ABSTRACT

This paper describes the design of an adaptive array antenna for land vehicle operation and its performance in an operational satellite system. Linear and circularly polarized antenna designs are presented. The acquisition and tracking operation of a satellite is described and the effect on the communications signal discussed. A number of system requirements are examined that have a major impact on the antenna design. The results of environmental, power handling and RFI testing are presented and potential problems identified.

INTRODUCTION

The linearly polarized adaptive array antenna consists essentially of a driven $\lambda/4$ monopole surrounded by concentric rings of parasitic elements all mounted on a ground plane of finite size. The parasitic elements are connected to ground via pin diodes. By applying suitable biasing voltages the desired parasitic elements can be activated and made highly reflective. The directivity and pointing of the antenna beam can be controlled both in the elevation and azimuth planes using high speed digital switching techniques.[1] By adding a circular polarizer to the linearly polarized design an increase in gain can be realized at the expense of an increase in antenna height.

The antennas are designed for land mobile applications and provide an angular coverage of 360° in azimuth and between 15° and 60° in elevation over an operating band of 1530 MHz to 1660 MHz. A maximum of 32 beams can be generated in azimuth and 2 to 3 beams in elevation depending on antenna size. The antennas are designed to have low sidelobe levels at low elevation angles to minimize the degrading effects of multipath on the tracking and communications performance. The RF losses are negligible and the measure antenna noise temperature is 60K. The antennas are designed to handle a maximum RF power of 100 watts. The electrical characteristics are simple to measure and are highly repeatable. The antenna manufacturing cost in quantities of 10,000 per year is estimated to be less than $1000. (Can.).

DESCRIPTION

A 5 ring linearly polarized design is shown in Figure 1. The antenna incorporates sufficient electronics to

control the radiation patterns and pointing on command. It is designed to be mounted on the metallic roof of a vehicle where the effective ground plane can significantly enhance antenna gain at low elevation angles. The antenna has a diameter of 20", a height of 2" above the ground plane and 1/2" below. It weighs 11 lbs. and requires less than 10 watts of power. The antenna gain varies between 9 and 11 dBic and has a maximum return loss of -12 dB. over the required angular coverage and operating frequency bands. A circularly polarized version of the antenna, shown in Figure 2, is obtained by adding a polarizer to the linearly polarized array. The antenna has a diameter of 24" and a height of 8" and weights 16 lbs. the antenna gain varies between 10 and 13 dBic and has a maximum return loss of -11 dB over the required angular coverage and operating frequency bands. The antenna pattern and return loss measurements are shown in Figures 3 to 8.

ACQUISITION/TRACKING OPERATION

Commands are executed by transmitting to the antenna a serial bit stream, each bit representing the conducting state of a corresponding pin diode and associated parasitic element. Each command defines the switching configuration to point the antenna beam in a particular direction and takes less than 100 micro-seconds to execute. Diagnostic checks are carried out prior to operation.

The satellite is initially acquired by stepping through 16 azimuth beam positions and selecting the beam with the strongest signal. In the event that the signal falls below a given threshold, the acquisition sequence is again initiated until the signal is re-acquired. The speed of operation is determined by the terminal C/No ratio and the signal to noise requirements in the control loop bandwidth. Currently it takes less than 0.1 seconds to acquire the satellite after initial phase lock. The satellite is subsequently tracked using one of 3 modes of operation.

Signal Sensing

A DC signal proportional to the RF signal is derived at the terminal receiver. The satellite is tracked by periodically switching on either side of the current beam position and selecting the beam with the strongest signal. Currently this is accomplished in less than 10 msec. and in normal vehicle operation occurs twice per second. A number of algorithms have been devised to minimize any perturbation of the communications signal to less than 1% of the time. The maximum phase transients when tracking in azimuth can be kept to less than ±10 degrees over the required angular coverage and operating frequency bands.

Signal and Angular Rate Sensing

The tracking operation is similar to the above but the up-dating frequency of the antenna beam position is governed by the angular rate of change of the vehicle as determined by a flux-gate compass. This minimizes the perturbation of the communications signal and provides the necessary up-dating at high angular velocities.

Flux-gate Compass

The flux-gate compass can be used as a primary sensor or as an adjunct or back-up to the signal sensing system.

FIELD TRIALS

Extensive field trials have been conducted using INMARSAT's MARECS B satellite. The three basic tracking modes have been successfully demonstrated. Calibrated duplex tests have been conducted using the Standard "C" transmission for tracking purposes and transmitting CW signals at 24 dBw from the vehicle and recording the transmission at the base station. The measured C/No ratios were within 1 dB of the predicted values. Field trials using ACSSB modulation are presently being conducted and will be extended in near future to include digital modulation.

SYSTEM CONSIDERATIONS

Current antenna specifications for vehicle antennas are tentative and subject to change. There are a number of uncertainties that have a major impact on the antenna design.

Cost

The cost of the antenna for airborne applications is not likely to be a major consideration. In the case of land vehicles and to a lesser extent maritime applications, the cost of the antenna, when manufactured in quantity, must be low enough to be commercially viable. A manufacturing cost of less than $1,000. (Can.) is the design goal for the Canadian MSAT program.

Polarization

Satellite operation at L-band uses circular polarization to overcome the effects of Faraday rotation and to eliminate the problem of polarization alignment at the mobile terminals. A requirement that the land vehicle antenna be circularly polarized is based in part on the premise that by alternating the sense of polarization of adjacent satellites, a degree of isolation is achieved between satellites (20 dB), that permits the re-use of the available RF spectrum. When the mobile terminal antenna points at the desired satellite, the adjacent satellites are illuminated by the antenna sidelobes. Most electronically steered array designs have difficulty in controlling cross-polarized sidelobe levels over the relatively large angular coverage and operational frequency bands. In some cases the cross-polarized levels can be as high as the co-polarized sidelobes. This limitation together with the depolarizing effects of scattering by objects in close proximity to the antenna, fading and other propagation effects, questions the feasibility of using orthogonal circular polarizations to enhance inter-satellite isolations.

Terminal G/T

The current MSAT and Inmarsat minimum G/T requirements for land mobile terminals are -17 dB/°K and -12 dB/°K respectively with corresponding nominal antenna gains of 8 dBic and 12 dBic respectively. The G/T requirements have been arrived at after allowing for a systems operational margin. A dominant contribution is an allowance for the degrading effects of multipath and fading which occur at low elevation

angles in land mobile operation. A directional antenna pointing at high elevation angles, will experience less multipath and fading and will have a lower antenna temperature than when pointing close to the horizon. A relaxation in G/T requirements at high elevation angles of between 1 and 2 dB vis `a vis low elevation angles may be possible for land mobile operation.

ENVIRONMENTAL TESTING

Vibration

The antenna has been subjected to vibration testing and meets land vehicle MIL specification 810C.

Thermal.

Solar simulation tests have been conducted. The highest temperatures recorded within the radome and electronic enclosure were 50°C and 70°C respectively well within the component design limits. No measurable shift in antenna boresight was observed. No problems were experienced during field trials in the winter months when temperatures were as low as -20°C.

Power

High power testing was conducted by sweeping an 80 watt CW signal across the transmit band via a diplexer and monitoring the receive band. There was no observable increase in noise level or evidence of spurious responses.

RFI

There is the possibility of VHF transmitters (70 - 130 MHz) mixing with the L-band transmit frequencies to produce interference in the receive band. Tests were conducted in the Ottawa area where there are about 10 VHF transmittters within a 10 mile radius with EIRP's of 50 dBw that could cause potential interference. The antenna, radiating 40 watts of CW power in the transmit band, was pointed in direction of the VHF transmitters and the received band monitored. Although no measurable interference was observed, it can be expected that when operating in close proximity to VHF transmitters, interference may occur using even passive antenna designs.

ACKNOWLEDGEMENTS

The author is grateful for the support provided by Martial Dufour, Bert Schreiber and Keith Elms in designing the tracking software and hardware and the magnetic sensor.

REFERENCES

1. Milne, R. 1988. An Adaptive Array Antenna for Mobile Satellite Communications. NASA/JPL Mobile Satellite Conference, Pasadena.

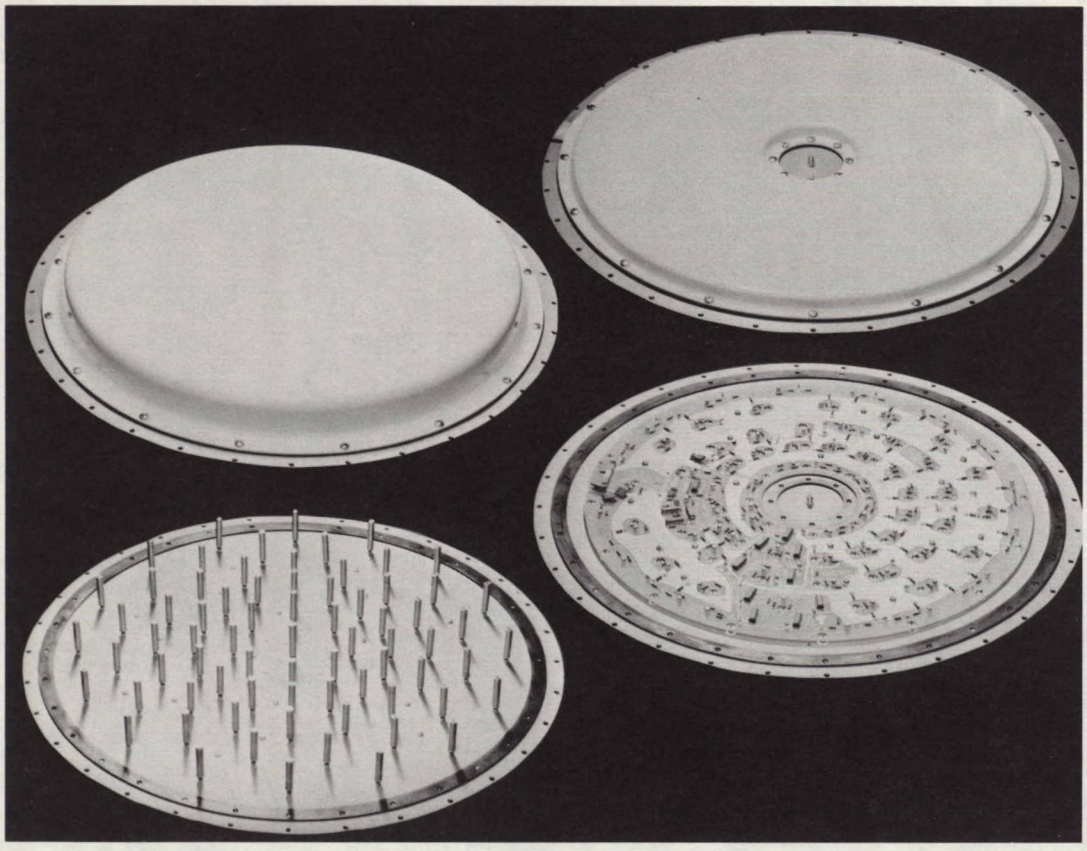


Fig 1 Linearly Polarized Antenna

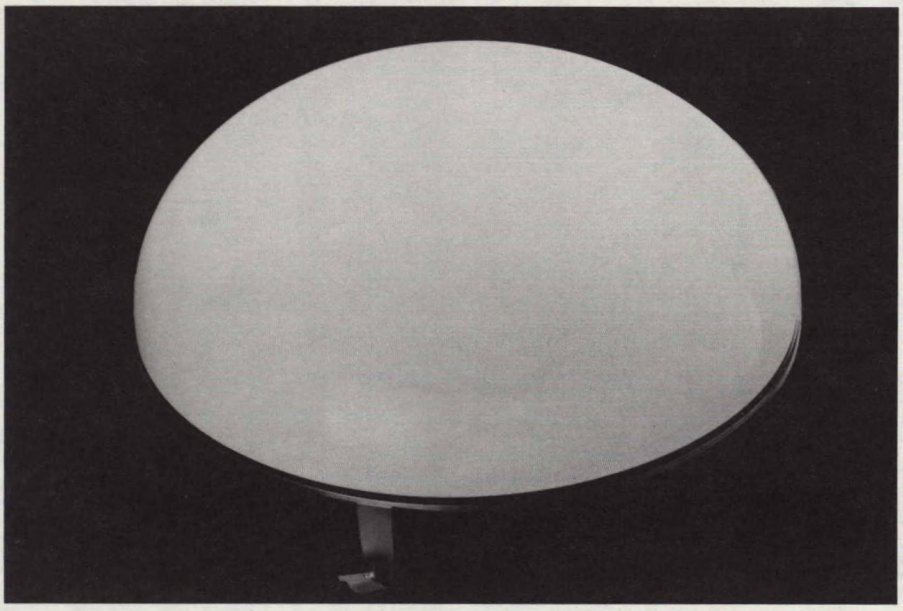


Fig 2 Circularly Polarized Antenna

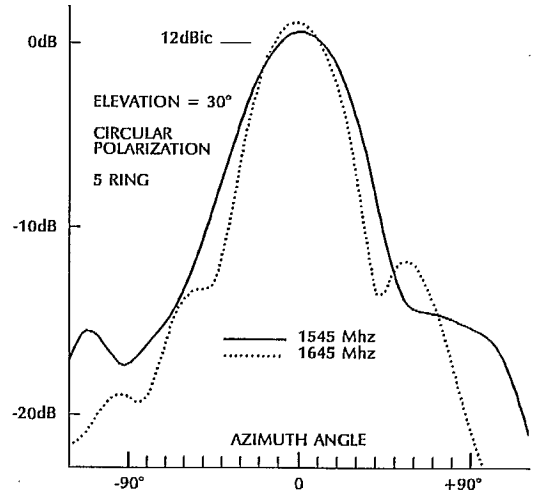


FIG. 3 AZIMUTH PATTERN - LOW BEAM

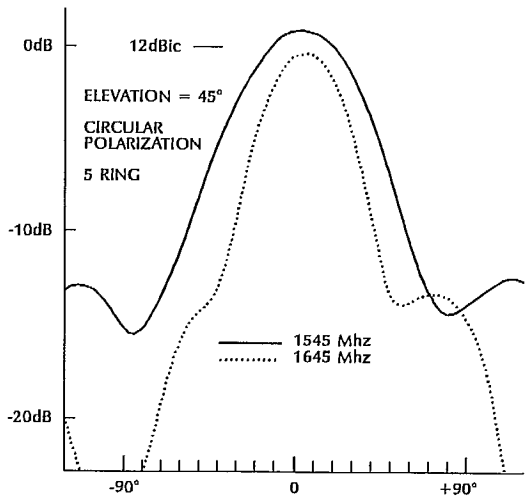


FIG. 4 AZIMUTH PATTERN - INTERMEDIATE BEAM

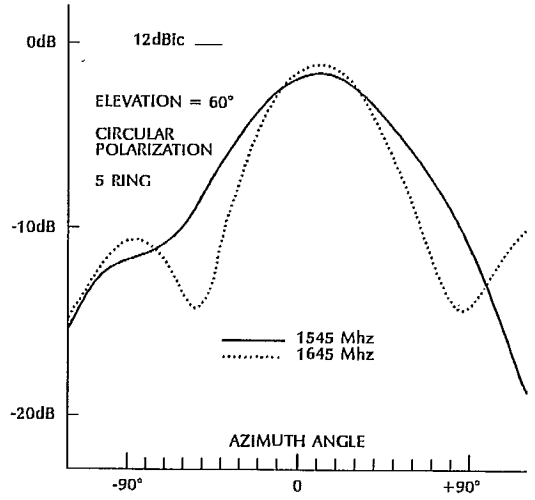


FIG. 5 AZIMUTH PATTERN - HIGH BEAM

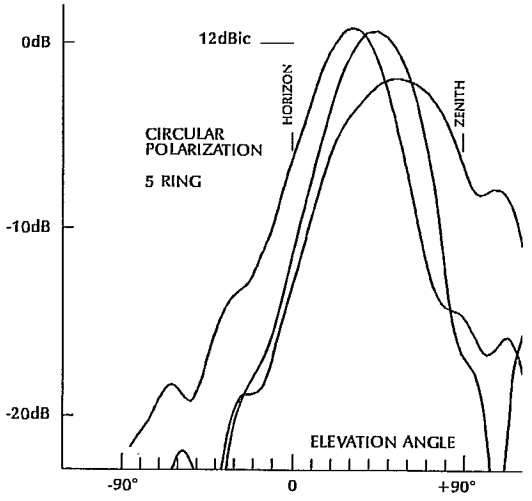


FIG. 6 ELEVATION PATTERNS - 1545 Mhz

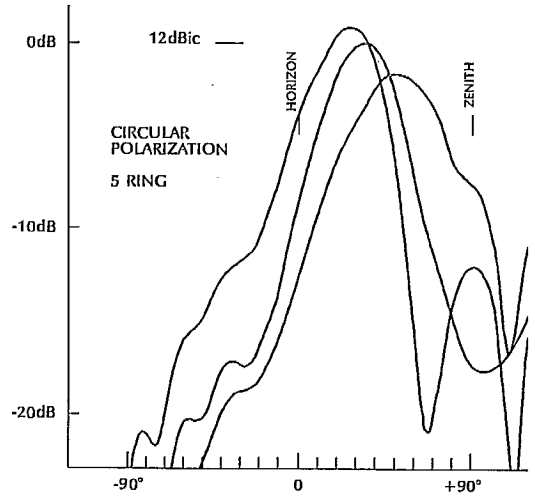


FIG. 7 ELEVATION PATTERNS - 1645 Mhz

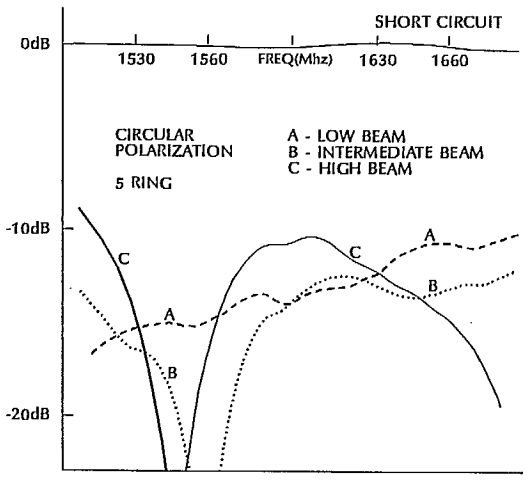


FIG. 8 RETURN LOSS VERSUS FREQUENCY

Standard-M Mobile Satellite Terminal Employing Electronic Beam Squint Tracking

G.J. Hawkins M.A. Beach G.S. Hilton
University of Bristol
Centre for Communications Research
Queens Building, University Walk
Bristol BS8 1TR, United Kingdom
Tel: +44 272 303258, Fax: +44 272 255265

Abstract

In recent years, extensive experience has been built up at the University of Bristol in the use of the Electronic Beam Squint (EBS) tracking technique, applied to large earth station facilities. The current interest in land mobile satellite terminals, using small tracking antennas, has prompted the investigation of the applicability of the EBS technique to this environment. In this paper the development of an L-band mechanically steered vehicle antenna is presented. A description of the antenna is followed by a detailed investigation of the tracking environment and its implications on the error detection capability of the system. Finally, the overall hardware configuration is described along with plans for future work.

Introduction

Currently at the University of Bristol work is progressing on the development of a mechanically steered array for use in L-band MSAT applications. This antenna configuration offers a number of advantages when compared to either an omni-directional antenna or electronically steerable phased array systems. Compared to an omni-directional antenna it offers improved gain characteristics reducing the requirement on satellite transponder power. A high degree of directivity provides better discrimination against multipath effects, and further reduces off-axis radiation with a consequent easing of frequency co-ordination problems. On the receive side, similar considerations offer an improvement of antenna G/T and hence a possible widening of link budget margins. When compared against the electronically steerable phased array, although not offering as elegant a solution to the vehicle mounted antenna, it is more realisable in terms of current technology and hence offers lower production costs.

The antenna gain requirement was a minimum of 10dBic in the direction of the satellite throughout the coverage region, which extends 0° – 360° in azimuth and 20° – 60° in elevation.

Error Detection Systems

Recently the EBS tracking scheme [1,2] has been developed for application to reflector antennas. With this system, a number of mode generators coupled to the primary-feed of the antenna are controlled electronically to sequentially produce, from higher order modes generated in the feed at the beacon frequency, a secondary beam deviation (squint) in a series of predetermined directions. Step scanning of the secondary beam axis is thus achieved without movement of the antenna or the primary beam. Hence, a tracking error signal can be derived by measurement of the beacon signal strength in much the same way as with a step-track system. The essential difference, however, is that the electronic deflection can be achieved at high speed and therefore EBS represents a form of pseudo-simultaneous amplitude sensing scheme. This yields the potential of operational tracking accuracies approaching that of a monopulse based system, at an implementation cost similar to step-track. Further, EBS does not suffer from the problems of dynamic response associated with a step-track facility [3].

If it is assumed that the main beam shape of the antenna approximates to a parabola, then the signal fall off in dB is given by:

$$G = 12\left(\frac{\theta}{\theta_{3dB}}\right)^2 \qquad (1)$$

where θ is the angle away from the antenna boresight

and θ_{3dB} is the half power beamwidth of the antenna. By squinting the beam to two positions equally placed about the antenna boresight, a gain loss of G_1 and G_2 will be observed given by:

$$G_1 = 12\left(\frac{\theta-\alpha}{\theta_{3dB}}\right)^2 \quad (2)$$

$$G_2 = 12\left(\frac{\theta+\alpha}{\theta_{3dB}}\right)^2 \quad (3)$$

where α is the squint angle. Re-ordering and subtracting yields the gain difference equation:

$$\Delta G = \frac{48\theta\alpha}{\theta_{3dB}^2} \quad (4)$$

Simple rearrangement of this relationship therefore gives the present satellite position, θ, in terms of the observed signal variation between squint positions, ΔG, as:

$$\theta = \frac{\Delta G\theta_{3dB}^2}{48\alpha} \quad (5)$$

Employing this relationship in a signal processing controller it can be seen that a simple, yet effective, closed loop control system is formed.

Microstrip Patch Array

A single-layer microstrip array antenna has been developed to provide the beam squint operation described above, in the land mobile environment. The antenna feed network, shown in Figure 1, consists of four circularly polarised array elements (centre-to-centre separation of 0.8λ) connected to a four-way power splitter/combiner with two phase shifters inserted in the outer feed lines. Under normal operation, the phase shifters are not activated; therefore, with no additional phase shifting in the feed network, the array has a broadside main beam. By applying a 50° phase delay in either one or other of the outer feed lines, the position of the azimuth beam can be moved off broadside by a few degrees. This achieves the beam squint action necessary to derive the required azimuth tracking error signals.

As with the azimuth radiation pattern, the elevation pattern has a maximum at the array broadside with the arrangement shown in Figure 1. Thus, it is necessary to mount the array at 40° to the horizontal to achieve the desired elevation coverage (20° – 60°). In order to mount the antenna flat on the vehicle, the use of parasitic elements is being explored as a means of pulling the beam away from elevation broadside.

The unsquinted azimuth radiation pattern for the prototype microstrip array antenna is shown in Figure 2. The gain of the antenna was found to be 11dBic, with a −3dB beamwidth of 16° and sidelobe levels at −10dB relative to peak. The associated elevation pattern is given in Figure 3. A −1dB beamwidth in excess of 40° is observed, and this indicates that a gain of 10dBic could be maintained in the direction of the satellite throughout a coverage region of 360° in azimuth and 40° in elevation. Activation of the individual phase shifters produced a squint magnitude of approximately 1/5th θ_{3dB}, with a gain loss due to the squinting action of less than 0.1dB. Under these circumstances maximum sidelobe levels increased by no more than 2dB.

Computer Simulation

In order to evaluate the likely tracking performance of the described EBS microstrip array antenna system, a computer simulation has been developed; one aspect of which allows the realistic modelling of the land mobile environment. A full description of the simulation is outside the scope of this paper, however, modelling of the beam squint mechanism, channel signal-to-noise ratio (SNR), atmospheric scintillation and the dynamics of a land mobile have all been incorporated.

The tracking performance has been investigated for various link specifications appropriate to the land mobile satellite service. Primarily, this has addressed the problem of acquisition for given terminal dynamics and signal conditions, as well as long term tracking performance, for various SNR's.

The selection of the appropriate sampling frequency is critical for the satisfactory operation of an EBS based antenna tracking system. Derivation of this parameter begins by establishing the relationship between the effective dynamic angular movement of the target and the induced signal level fluctuation. This is possible by considering the differential of the antenna gain loss equation (Equation 1) which yields:

$$dG = 24\left(\frac{\theta}{\theta_{3dB}^2}\right)d\theta \quad (6)$$

This equation can be re-expressed, using the relationship $dy/dx = (dy/dz)(dz/dx)$, as:

$$\frac{dG}{dt} = 24\left(\frac{\theta}{\theta_{3dB}^2}\right)\frac{d\theta}{dt} \quad (7)$$

where dG/dt is the observed amplitude fluctuation per second (dB/sec), θ is the angle of the satellite away from the boresight axis of the antenna and $d\theta/dt$ is the

rate of change of the dynamic angular motion between the target and antenna positions.

An initial indication of the required sampling frequency can be obtained by considering the terminal working to a satellite at a relatively high elevation angle, say 50°, over a line-of-sight path (non-urban environment). In these circumstances, even though the antenna has a wide elevation beamwidth, multipath effects can be shown to be negligible and unwanted signal fluctuations will derive either from atmospheric scintillation effects or vehicle motion, if roadside vegetation shadowing is also ignored. By considering the case of an angular target velocity of 20°/sec (representative of the car turning on a radius of 40m at a speed of 48km/hr), use of Equation 7 shows a maximum amplitude fluctuation of 21.8dB/sec will be observed, with the satellite θ_{3dB} degrees away from the antenna boresight. This amplitude fluctuation is considerably greater than that which will result from worst case atmospheric effects. Substitution of this value into the single axis form of the sampling frequency equation for EBS systems, derived by Hawkins and Edwards [4]:

$$f_s(max) = \frac{\frac{dG}{dt} 10^{(0.05SNR+1)}}{8.69} \qquad (8)$$

yields a maximum required sampling frequency $f_s(max)$ of 250Hz with a SNR = 20dB. This SNR is achievable within the necessary bandwidth of the tracking receiver (1.5 x sampling frequency) if a carrier-to-noise density of 45dBHz is assumed. Such a carrier-to-noise density is typical of that which might be seen on a Standard-M LMS link, as laid down in the INMARSAT 'Strawman' specification [5].

At low elevation angles, the full severity of the LMS tracking environment becomes apparent. Work by Beach *et al.* [6] investigated both line-of-sight (LOS) and multipath propagation for a Standard-M antenna system receiving transmissions from the MARECS B2 satellite. Experimental trials were carried out in southern England which resulted in a working elevation angle of approximately 25°. Results shown in Figure 4 relate to the LOS operation. The signal seen in either channel of the receiving equipment (data for two channels is presented as the work was concerned primarily with diversity combining techniques) consists of a dominant LOS component, a small multipath contribution obtained from signal reflections off grassland surrounding the road being used by the vehicle, as well as some shadowing effects from roadside trees.

Examination of the signal characteristics seen in this figure reveal maximum amplitude fluctuations in the order of 4dB for a spatial movement of the vehicle of 0.2λ. Considering a vehicle travelling with a velocity of 48km/hr this yields the maximum rate of observed signal variation of approximately 1400dB/sec. Such a variation considerably increases the required sampling frequency of the EBS system and because of the wider receiver bandwidth needed to preserve the tracking modulation, means much lower operational SNR's have to be considered. If an SNR of 10dB is initially assumed, use of Equation 8 indicates that a sampling frequency of 5000Hz should be implemented. With a carrier-to-noise density of 45dBHz this means that the actual SNR 'seen' at the input to the tracking controller will be 8dB.

Since it has been already shown that the vehicle motion dynamics impose only a required sampling frequency of 250Hz, and hence an antenna positioning update rate of approximately 125Hz, use of a 'sample mean estimator' as part of the tracking controller algorithm can be shown to reduce the variance of the tracking error estimates. Simulation results obtained using the 5000Hz sampling frequency, SNR = 8dB and the averaging factor in the sample mean estimator = 20, indicate an achievable beam radial error (BRE) in the order of 1.4° with a maximum antenna slew rate of 200°/sec.

Although the LOS propagation characteristics have not been implemented in the simulation, since the maximum LOS amplitude fluctuation that will be seen during the measurement period corresponds to 0.28dB, which using Equation 5 gives a positioning uncertainty of 0.9°, it would not be unreasonable to expect that the overall BRE should not degrade much beyond 1.8°, ie. 1/9th θ_{3dB}. It is apparent that this tracking performance is comparable with that currently required for the land mobile environment [7].

Multipath dominated propagation results (representative of conditions that would be seen in an urban environment) are also reported by Beach *et al.* [6]. It is clear that observed signal fluctuations are much more severe than those for the LOS case and as such would make the use of the EBS tracking technique in this environment impracticable.

Work is currently progressing to further modify the simulation package to include a realistic propagation model to verify the tracking performance indicated above. This model will describe the direct LOS path, with little or no shadowing and only a small multipath contribution by a Rician distribution. For a received signal experiencing only light shadowing, the received signal envelope will be modelled as the sum of log-normal and Rayleigh random processes, while for the case of heavy shadowing the received envelope distribution will tend towards a Rayleigh function.

Overall Hardware Configuration

The overall configuration of the terminal hardware will be very similar to that already proposed by Berner [8] and Bell *et al.* [9], who discuss the JPL mechanically steered array. The rotating antenna platform will be mounted on a fixed supporting structure which will be located in the roof of the test vehicle. Drive to the rotating platform will be via a stepping motor. The received RF and control signals to the phase shifters will be supplied to the antenna via a rotary joint and slip rings respectively. The tracking controller will be based on a TMS 320C25 digital signal processor which will handle the tracking error derivation plus the closed and open loop control functions. This unit is easily capable of dealing with the data acquisition rates discussed and will receive inputs, not only from the tracking receiver, but also from a vehicle mounted angular rate sensor and/or flux gate compass.

Two control strategies are currently being considered for the proposed antenna. The first, will utilise the closed loop EBS element as the primary tracking system. In the case of the received signal dropping below a certain threshold, tracking will be transferred to open loop control based on information available from the rate sensor and/or flux gate compass. The second approach is to use a hybrid tracking system of the type employed in a number of maritime terminals [10,11]. This scheme will optimally combine (using a Kalman filter estimator) the data available from both the closed and open loop tracking elements, to produce the resultant drive signal for the antenna servo system. Results presented [10,11] indicate an improvement in tracking accuracy of 100 – 200% for the cases considered.

Conclusions

The development of an L-band mechanical steered antenna for MSAT applications has been described. Results presented for the prototype microstrip patch array fulfil the required terminal specification and demonstrate the very favourable electronic beam squint operation. Examination of the non-urban land mobile tracking environment has shown that, although this represents probably the most severe test of the tracking ability of an EBS system considered to date, suitable accurate tracking performance can be obtained. Finally, the overall terminal hardware configuration has been outlined and tracking control strategies discussed.

A complete hardware demonstrator is now being constructed in the University. Field trials are expected to take place in the near future and it is hoped these will confirm the optimism currently held in the system described.

References

[1] **Dang, R., Watson, B.K., Davies, I., Edwards, D.J.,** "Electronic tracking systems for satellite ground stations" *15th European Microwave Conference*, Paris, France, September 1985, p.681-687.

[2] **Edwards, D.J., Watson, B.K.,** "Electronic tracking systems for microwave antennas" British Patent Application No.8414963, 12th June 1984.

[3] **Hawkins, G.J., Edwards, D.J., McGeehan, J.P.,** "Tracking systems for satellite communications" *IEE Proceedings*, Vol.135, Pt.F, No.5, p.393-407, October 1988.

[4] **Hawkins, G.J., Edwards, D.J.,** "Operational analysis of electronic tracking scheme" *IEE Proceedings*, Vol.136, Pt.I., No.3, June 1988, p.181-188.

[5] "Land mobile satellite communications" Facts Sheet published by INMARSAT, March 1988.

[6] **Beach, M.A., Swales, S.C., McGeehan, J.P.,** "A diversity antenna array for a land mobile satellite terminal" *Fifth International Conference on Mobile Radio and Personal Communications*, 11-14th December 1989, p.177-181.

[7] **Woo, K.,** "Vehicle development for mobile satellite applications" *Fourth International Conference on Satellite Systems for Mobile Communications and Navigation*, Conference Publication 294, 17-19th October 1988, p.23-27.

[8] **Berner, J.B.,** "The JPL mechanically steered antenna" *Proceedings of Mobile Satellite Conference*, May 3rd-5th 1988, p.241-247.

[9] **Bell, D.,** "Reduced-height mechanically steered antenna development" *MSAT-X Quarterly*, JPL No.18, p.3-10.

[10] **Choi, J., Kasprzak, J.A.,** "Dynamic analysis of shipboard-mounted tracking antennas" *American Control Conference* , Vol.3, 1984, p.1391-1397.

[11] **James, M.R., Maney, J.J.,** "Adaptive alignment of a shipboard satellite terminal" *IEEE Military Communications Conference, MILCOM '85*, Vol.1, 20th-23rd October 1985, p.300-305.

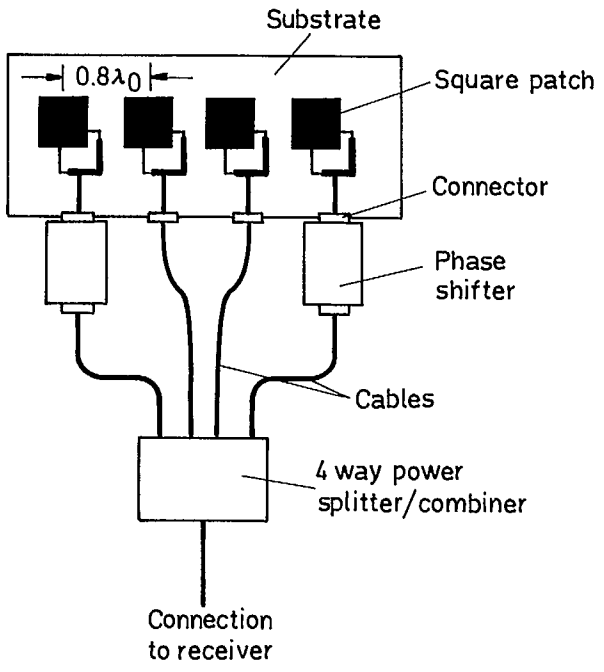


Figure 1: Four element circularly polarised microstrip patch array antenna.

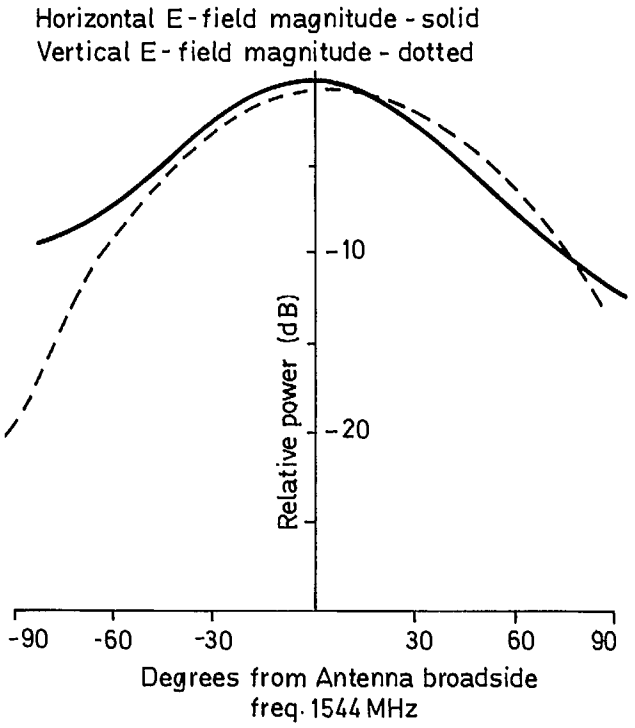


Figure 3: Measured elevation radiation pattern for circularly polarised microstrip patch antenna.

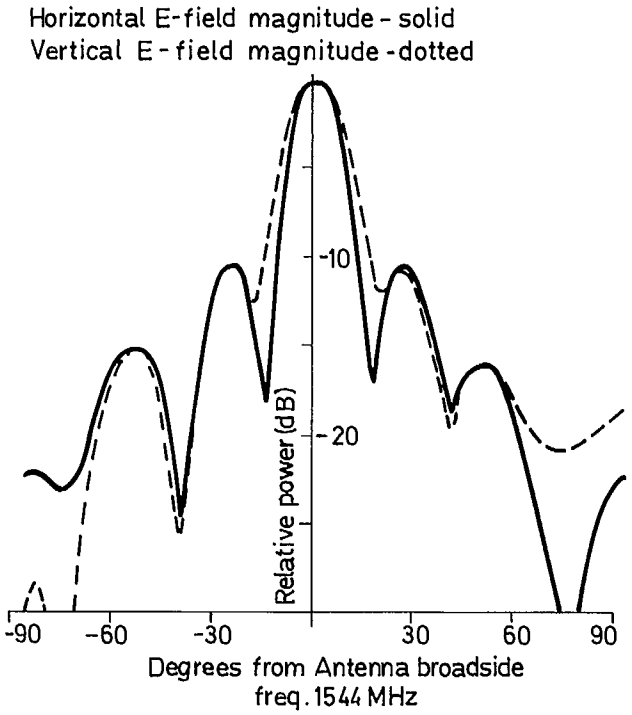


Figure 2: Measured azimuth radiation pattern for circularly polarised microstrip patch antenna.

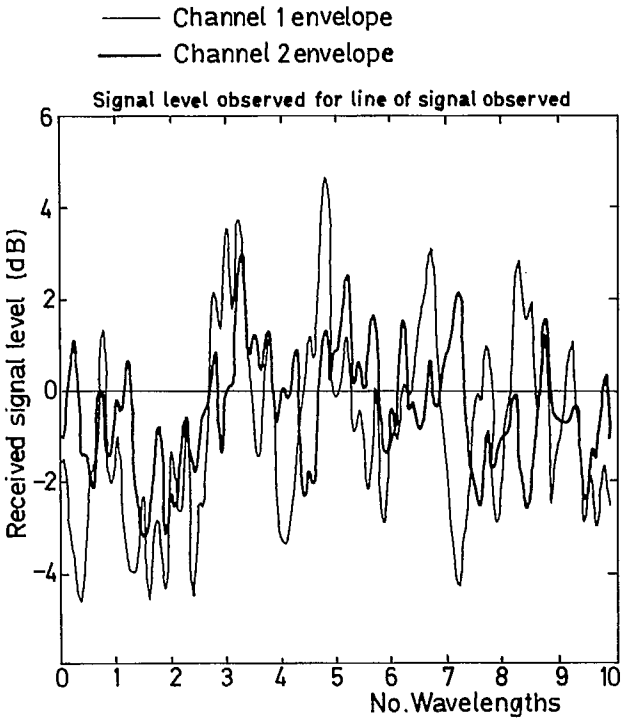


Figure 4: Fading signal envelopes observed for line-of-sight operation.

Performance of a Family of Omni and Steered Antennas for Mobile Satellite Applications

K. Woo, J. Huang, V. Jamnejad, D. Bell, J. Berner,
P. Estabrook, and A. Densmore

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109 USA
Phone: (818) 354-3835
FAX: (818) 393-6875

ABSTRACT

This paper describes the design and performance of a family of vehicle antennas developed at JPL in support of an emerging U.S. Mobile Satellite Service (MSS) system. Test results of the antennas are presented. Trends for future development are addressed. Recommendations on design approaches for vehicle antennas of the first generation MSS are discussed.

INTRODUCTION

The emerging U.S Mobile Satellite Service system will provide telephone and data services for a variety of users across the Continental United States (CONUS). To implement this system, high performance vehicle antennas are needed to provide the communications link between the satellite at the geostationary orbit and the mobile vehicles on the ground. Over a number of years, JPL has been conducting research and development on a family of vehicle antennas for meeting the requirements of diverse potential users. These antennas are required to be circularly polarized and to provide coverage from 20° to 60° above the horizon in elevation, and a full 360° in azimuth. They are to operate at the L-band, 1.5450 to 1.5590 GHz for receive, and 1.6465 to 1.6605 GHz for transmission. Both low-gain omni-directional and medium-gain steerable antennas have been developed. Test results of earlier designs of these antennas have been summarized and reported in Ref. [1]. This papers reviews the designs and addresses recent development of the antennas.

OMNI-DIRECTIONAL ANTENNAS

The purpose of developing omni-directional antennas is to provide users with antennas that are simple, reliable, and low cost. Circularly polarized omni antennas of crossed drooping-dipole, quadrifilar helix, and microstrip designs have been developed[1]. These antennas were found to be able to provide 3.5 to 6.0 dBic gain for CONUS coverage.

The crossed drooping-dipole antenna is most versatile in use. By varying the separation between the dipole elements and the ground plane (top of vehicle) the elevation pattern can be adjusted for optimum coverage for the coverage region of interest. Both right-hand circularly polarized (RCP) and left-hand circularly polarized (LCP) designs were developed. Figure 1 shows an LCP design used in the AUSSAT mobile experiment, described below. The antenna provides a peak gain of 5.3 dBic at the AUSSAT look angle of 55° above the horizon.

STEERABLE ANTENNAS

Medium-gain steerable antennas were developed for mobile vehicles. The purpose of the development is two-fold: (1) to provide higher antenna gain for the mobile terminal, and (2) to provide a directive beam to effect sufficient intersatellite isolation for interference-free communications. The goals are to achieve: (1) a minimum of 10 dBic gain throughout the elevation angles 20° to 60° above the horizon, and (2) 20 dB intersatellite isolation for two satellites with opposite polarization placed approximately 30° apart in the geostationary orbit. Two classes of steerable antennas have been designed and breadboarded. These are the electronically steered phased-array antennas and the mechanically steered tracking antennas. Low cost and low profile were the two principal drivers in designing these antennas.

Electronically Steered Phased Array Antennas

Phased array antennas were developed principally to provide a thin antenna that can be installed conformal to the top of the vehicle for aesthetic or security

reasons. These antennas are well known for their complexity and high cost. As a result, emphasis was placed, in addition to meeting the RF and pointing requirements, on the selection of manufacturing techniques, materials, and component types, so as to keep the cost down.

Two RCP phased-array antennas were separately developed by Ball Aerospace Systems Division[2] and Teledyne Ryan Electronics[3] through contracts and technical guidance by JPL. The antennas developed by these two companies, as shown in Figure 2, exhibit several common features. Both antennas use 19 radiating elements with 18 3-bit diode phase shifters. The RF construction of both antennas employ the "layer" approach. Satellite tracking is achieved in the azimuth plane by the sequential lobing technique and in the elevation plane by a slow amplitude search mechanism. The tracking systems of both antennas are augmented by an angular rate sensor for combating short signal drop-outs or fades.

There are also several distinct differences in technology and design between the two antennas developed by the contractors. Ball uses a dual-resonant stacked half-inch-thick circular microstrip element to cover both the transmit and receive bands, while Teledyne employs a quarter-inch-thick stripline cavity-backed crossed-slot radiator. The microstrip element has a 3-dB beamwidth of about 90°, and the crossed-slot element has a beamwidth of 140°. The crossed-slot element is thinner and allows the array beam to scan lower in elevation due to its wider element beamwidth. The dual stacked microstrip element is lower in cost because of its simplicity and the use of low-cost foam material. Another difference between the two antennas is that Teledyne uses the switched-line 3-bit diode phase shifters, while Ball employs the hybrid reflection/loaded-line type 3-bit phase shifters. The former use 6 diodes each, while the latter require 12 diodes each. The hybrid phase shifter has slightly higher insertion loss and phase error as compared to the switched-line design. Overall, it seems that Ball emphasized a lower cost design, while Teledyne concentrated on performance.

In addition to the RCP phased arrays developed by Ball and Teledyne, an LCP unit was also produced by Teledyne under a follow-on JPL contract. This LCP unit was successfully tested in the AUSSAT mobile experiment in Australia with the Japanese satellite ETS-V which has an LCP antenna. The construction and design of this Teledyne LCP antenna are exactly the same as the RCP unit except the hybrid circuit where the two orthogonal ports are interchanged. As a result, the performance of the LCP unit is very similar to the RCP antenna. Several important characteristics and measured performance of the Ball and Teledyne phased array antennas are listed in Tables 1 and 2. The measured radiation patterns of the LCP phased array antenna are shown in Figure 3.

Mechanically Steered Tracking Antennas

Mechanically steered tracking antennas were developed for the purpose of providing a tracking antenna with considerably lower cost than the phased array antenna. The challenge here is to achieve a low-profile and low-cost design in addition to meeting the RF and pointing requirements.

Three mechanically steered tracking antennas were developed. The original design was a tilted array antenna developed entirely at JPL[4]. An iteration of this antenna produced a lower-profile, lower-weight antenna by redesigning the pointing platform and employing integrated stripline feed circuitry[5]. Another antenna was a hybrid mechanically/ electronically steered antenna built by Teledyne Ryan Electronics under contract with JPL[6].

Figure 4 presents a pictorial view of the two mechanically steered antennas. Both antennas employ the same monopulse technique for tracking the satellite in azimuth. The full antenna system with either antenna installed is effectively described by the block diagram in Figure 5. The array is designed for the simultaneous transmission and reception of communication signals as well as the reception of the "error" signal required for monopulse tracking. The functioning of the antenna RF subsystem in the receive mode is summarized as follows.

The radiating array is divided into two equal subarrays, each composed of two radiating square microstrip patch elements. The signals from the two subarrays pass through a sum/difference hybrid whereby their sum and difference are obtained. The difference or "error" signal is then channeled through a one-bit (0-180 degrees) phase shifter which will be modulated at a rate of between 200 to 4800 Hz. It is then passed through a 30-dB ferrite isolator. The difference signal is then combined with the sum signal via a 10-dB coupler in which the difference signal strength is reduced by approximately 10 dB before combining with the sum signal. The combined output of the coupler is then passed to the single-channel rotary joint. Thus, since the difference signal is modulated before addition to the sum signal, it can be

extracted from the composite signal at the receiver by standard filtering techniques. This normalized "error" signal is then supplied to the pointing control circuitry, which commands a stepper motor to rotate the antenna array in order to maintain pointing toward the source of the incoming signal. This is basically the closed-loop operation of the antenna system. An azimuth rate sensor is used to provide for the open-loop operation of the system, namely to maintain pointing during the tracking phase when the received signal fades below a prescribed threshold.

The design and performance of the original JPL tilted array antenna has been reported extensively. The antenna was successfully tested in the Tower experiments at Erie, Colorado, and Satellite experiments at Santa Barbara, California.

The success of the original breadboard antenna led to additional research aimed at reducing the antenna height as well as integrating some of the discrete RF components into the stripline-fed microstrip patch array. Two stripline layers flush-mounted behind the microstrip patch array provide most of the required RF circuitry. The quadrature hybrids for four microstrip patches are on one layer, while the array divider/combiner, the sum/difference hybrid and the sum/ difference channel coupler are all implemented on a second layer. Discrete components are used only for the modulator and the isolator. In addition to somewhat improved RF performance, this integrated design is lighter and less expensive. The major modification of the pointing platform involved the use of a very low profile pancake stepper motor and associated gear drive mechanism, as well as a low-profile rotary joint. A low-cost rate sensor by ETAK corporation was also incorporated in the design. The modification of lower platform provided a height reduction from 23 to 15 centimeters and a weight reduction by about 4.5 Kg for the overall antenna system.

Both RCP and LCP units were breadboarded for the reduced-height antenna. They have approximately the same RF performance except for the sense of polarization. The LCP units were used in the AUSSAT mobile experiment. Several important characteristics and measured performance of the antennas are listed in Tables 1 and 2. The measured radiation patterns of an LCP unit are shown in Figure 5. The same antenna was also tested for power-handling capability. Over 20 watts of CW RF power was applied to the antenna at the rotary joint input. The antenna successfully passed the test without problems.

For further reducing the height of the antenna, a low-profile planar hybrid mechanically/electronically steered tracking antenna was fabricated by Teledyne Ryan Electronics under contract with JPL. Figure 6 shows a pictorial view of the antenna using 32 crossed-slot stripline elements and pancake stepper motor drive. The array elements are arranged with complete symmetry with respect to the two major axes of the array. Two phase shifters connected to the two identical halves of the array provide for sequential lobing in azimuth. Another two pairs of phase shifters are used for beam switching in elevation. Signal acquisition is based on seeking the maximum received signal strength.

This hybrid-scanned antenna did not achieve the RF performance goals. Factors such as high circuit losses, array grid configuration, and mutual coupling effects contributed to serious performance degradation. The design, however, seems basically sound; and a more careful implementation with some modifications can provide a successful antenna.

Noise Temperature Measurement

The noise temperature of the developed antennas have been measured using the set-up as shown in Figure 7[7]. The antenna under test is mounted to the roof of the Pilot Field Experiment (PiFEx) van and connected via a series of low noise amplifiers and filters to a spectrum analyzer. The noise temperature of the antenna is calculated by comparing the noise power read on the spectrum analyzer when the antenna is connected to the subsequent amplifiers and when a 50-ohm load is connected to the same amplifiers.

Both RCP and LCP units of the antennas were measured. Table 1 lists the measured noise temperature of the RCP units and Table 2 shows those of the LCP units used in the AUSSAT mobile experiment. The measurements were taken at the JPL Mesa Antenna Range. Table 2 also shows the corresponding mobile receiver G/T assuming that the antenna is connected to an amplifier with a 2.0 dB noise figure by a cable with a loss of 0.3 dB. The noise temperature of the receiver and cable (without the antenna) is 203K.

AUSSAT ANTENNA EXPERIMENTS

The LCP units of the reduced-height mechanically steered tilted array antenna, the Teledyne phased array antenna, and the crossed drooping-dipole antenna were field tested during the AUSSAT mobile experiment. Overall, the antennas performed very well, especially in providing the automatic antenna

pointing function. However, both the mechanical antenna and the phased array antenna (after 5 hours of operation) developed excessive noise when transmitting high power. The antennas were therefore used in receive mode only, with the crossed drooping-dipole antenna as the transmitting antenna. In this manner, a great variety of experiments were conducted during which the antenna pointing systems of both antennas performed flawlessly.

The original JPL mechanically steered tilted array antenna was successfully tested in duplex mode for the Tower experiments at Erie, Colorado, that required less transmit power. The noise associated with high power transmission has been traced to the monopulse modulator (phase shifter)[8]. The commercially purchased unit switched too fast for this application, generating extremely high order harmonic of the monopulse tone, which fell into the receive band. This was resolved by redesigning the modulator with slower phase transitions by using slower diodes and slower charge sweeping of the diode junction. Figure 8 shows the new modulator circuit and the noise level before and after the replacement. The noise in the Teledyne phased array antenna has been traced to one of the diode driver integrated circuits that had disconnected from the circuit and possibly microplasma noise of the PIN diodes.

CURRENT DEVELOPMENT ACTIVITIES

To further strive for low cost, low profile vehicle antennas, development work for low cost planar arrays are being carried out at JPL. These are the planar microstrip Yagi array, and the ANSERLIN (Annular Sector Radiating Line) array, as shown in Figure 9. These arrays are designed for direct replacement of the tilted array of the reduced-height antenna shown in Figure 4. The resulting antenna is nearly conformal (~ 4 cm high). A preliminary version of a mechanically steered tracking antenna using the planar microstrip Yagi array is shown in Figure 10. The antenna is currently under test.

The two types of arrays are expected to be low cost in that they both need only a very simple beamformer. In lieu of feeding each element, it is only necessary to feed rows of linear arrays. The planar Yagi array is based on mutual coupling effect, and the ANSERLIN array is based on travelling wave type of operation. The details of the design and the theory of operation can be found in the companion papers [9, 10] of this Conference.

A planar microstrip Yagi array designed to meet the specifications of the MSS has been breadboarded. The array meets the 10 dBic minimum gain requirement for CONUS coverage.

CONCLUSIONS AND RECOMMENDATIONS

Based on the design configurations and performance figures of vehicle antennas developed to date, the following observations are made.

The omni-directional crossed drooping-dipole antenna is small, simple, reliable, and low cost. Although the antenna is relatively low in gain, it does provide a reasonable antenna G/T. The antenna may not be adequate for voice applications, but it is certainly good for low rate data transmission. This antenna will certainly become a very important component for mobile communications.

The electronically steered phased arrays are conformal and therefore are aesthetically pleasing and good for security applications. But at this stage, they are of high cost and have some reliability problems. The cost drivers of the antenna are dielectric circuit boards, phase shifters, diodes, beam-pointing electronics, and assembly labor. Although their costs can be reduced by significant increase in market demand, new technologies are needed to develop low cost manufacturing processes and reliable components for minimizing the cost drivers and improving reliability.

The mechanically steered tracking antennas offer similar gain and G/T as phased array antennas, but they are considerably lower in cost. Although existing designs are relatively high-profile, the success of the low-cost, low-profile arrays such as planar microstrip Yagi, or ANSERLIN, or the like, will lead to a conformal antenna similar in height to the phased array antenna.

Comparing the relative merits, it seems that for the application to the first generation MSS, the mechanical steered tracking antennas using low-cost planar arrays would be a logical choice. However, phased array antennas could also provide distinct services when and if the cost can be brought down. Research should continue to develop new low-cost manufacturing processes and new reliable components/devices so that the phased array antenna can be a choice for future generations of MSS.

ACKNOWLEDGEMENT

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The authors wish to acknowledge the

excellent technical support of R. Thomas and C. Chavez in carrying out the development of the vehicle antennas.

REFERENCES

[1] Woo, K. 1988. Vehicle Antenna development for Mobile Satellite Applications. Proceedings of the IEE Fourth International Conference on Satellite Systems for Mobile Communications and Navigation, pp. 23-27.

[2] Schmidt, F., et al. 1988. MSAT Final Report, Ball Aerospace Systems Division. JPL Contract No. 957467.

[3] Peng, S. Y., et al. 1988. Final Report, Vehicle Antenna for the Mobile Satellite Experiment, Teledyne Ryan Electronics, JPL Contract Nol. 957468.

[4] Jamnejad, V. 1988. A mechanically steered Monopulse Tracking Antenna for PiFEx. MSAT-X Quarterly, 13, Jet Propulsion Laboratory, pp. 18-27.

[5] Bell, D., et al. 1989. Reduced-height Mechanically Steered Antenna Development. MSAT Quarterly, 18, Jet Propulsion Laboratory, pp. 3-10.

[6] Francian, et al. 1989. A Hybrid Scanned Conformal Planar Array for Mobile Satellite Communications. IEEE AP-S Symposium Digest, pp. 686-689.

[7] Estabrook, P. and Rafferty, W. 1989. Mobile Satellite Vehicle Antennas: Noise Temperatures and Receiver G/T. Proceedings of 39th IEEE Vehicular Technology Conference, pp. 757-762.

[8] Densmore, A., et al. 1990. Performance Evaluation of the Antennas used in the AUSSAT Land Mobile Satellite Experiment. MSAT-X Quarterly, 24, Jet Propulsion Laboratory (to be published).

[9] Huang, J., et al. 1990. Microstrip Yagi Array for MSAT Vehicle Antenna Applications. Proceedings of 1990 International Mobile Satellite Conference (these proceedings).

[10] Colomb, F. Y., et al. 1990. An ANSERLIN Array for Mobile Satellite Applications. Proceedings of 1990 International Mobile Satellite Conference (these proceedings).

Table 1. Performance of vehicle antennas designed for CONUS coverage

Antenna[1]	Gain[2] (dBic)	Intersatellite Isolation[3] (dB)	Noise Temperature (K)	Size		Estimated Manufacturing Cost[4] ($)
				Height (cm)	Diameter (cm)	
Mechanically steered tilted array antenna						
Original design	≥10.0	≥24	172	23	51	600
Reduced-height design	≥10.0	≥24	164	15	51	600
Electronically steered phased array antenna						
Ball design	≥ 8.0	≥19	161	3.3	61	1600
Teledyne design	≥ 8.0	≥25	186	1.8	54	1800
Crossed drooping-dipole antenna	≥ 4.0	≥10	45	≤12	8.0	50

1 All antennas are right-hand circularly polarized.
2 Averaged in azimuth at various elevation angles of interest for both transmit and receive frequencies.
3 For two satellites with opposite circular polarization separated by 30° in geostationary orbit.
4 Based on producing 10,000 units per year over a five-year period.

Table 2. Performance of vehicle antennas designed for AUSSAT mobile experiment

Antenna[1]	Gain (dBic)	Noise Temperature (K)	G/T[2] (dBic/K)
Reduced-height mechanically steered titlted array antenna (stripline-fed)	11.3	134	-14.0
Teledyne phased array antenna	10.5	186	-15.4
Crossed drooping-dipole antenna	5.3	37	-18.5

[1] All antennas are left-hand circularly polarized.

[2] Based on receiver and cable noise temperature of 203K.

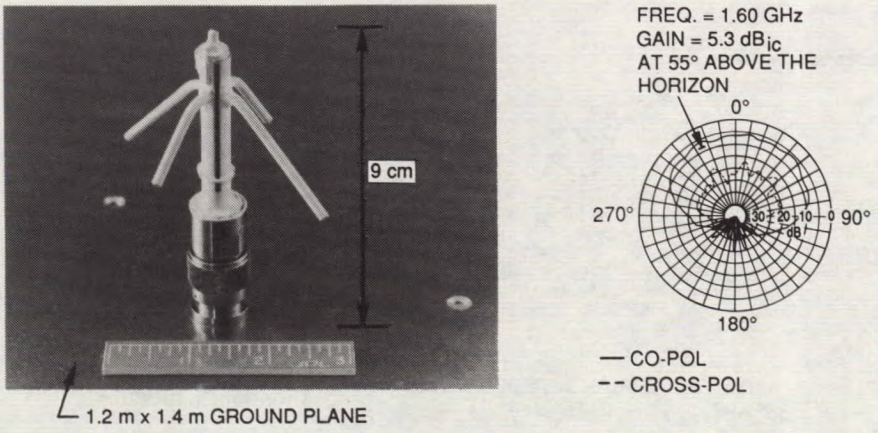


Figure 1. Crossed Drooping-Dipole Antenna, LCP

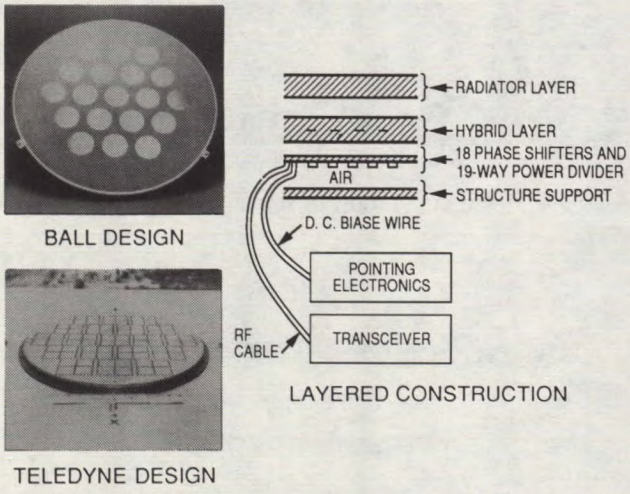


Figure 2. Electronically steered phased array antennas

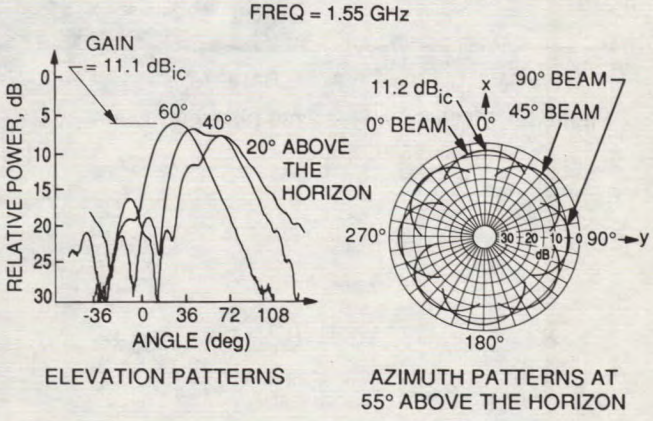
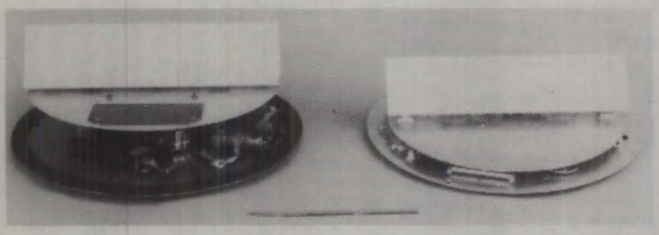
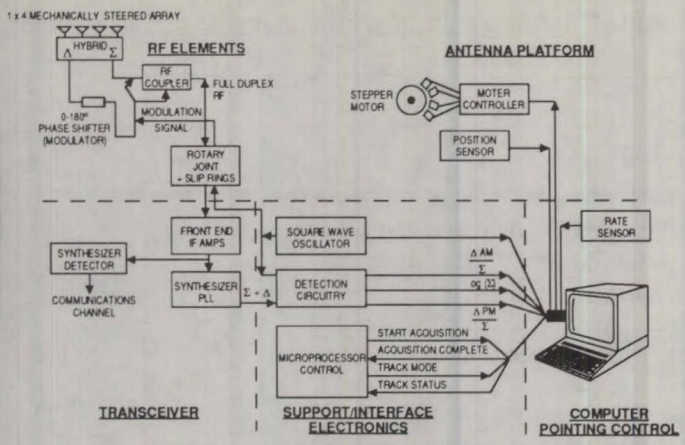


Figure 3. Co-pol patterns of Teledyne phased array antenna, LCP

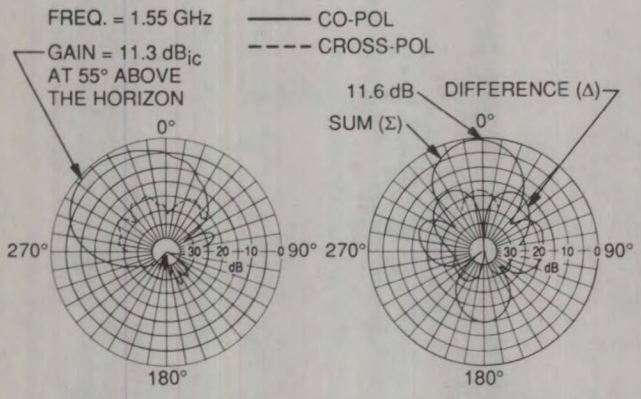


ORIGINAL DESIGN REDUCED-HEIGHT DESIGN



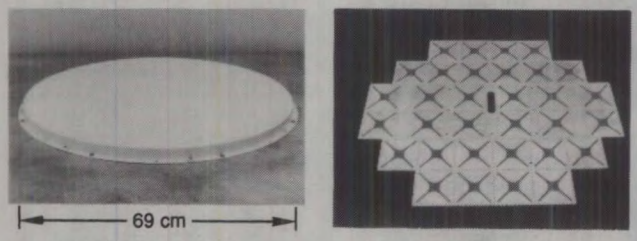
ANTENNA SYSTEM BLOCK DIAGRAM

Figure 4. Mechanically steered tilted array antennas



ELEVATION PATTERNS AZIMUTH PATTERNS AT 55° ABOVE THE HORIZON

Figure 5. Radiation patterns of reduced-height mechanically steered tilted array antenna, LCP



OVERALL VIEW RADIATING ELEMENTS

Figure 6. Mechanically steered planar array antenna (Teledyne design)

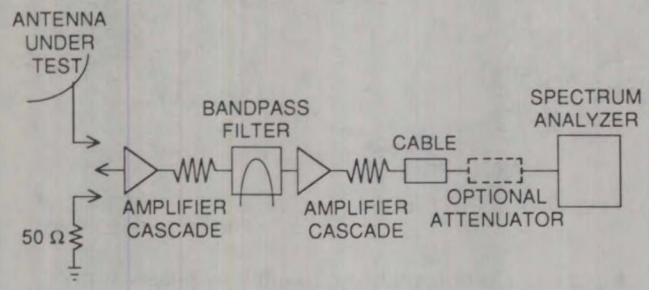


Figure 7. Noise temperature measurement set-up

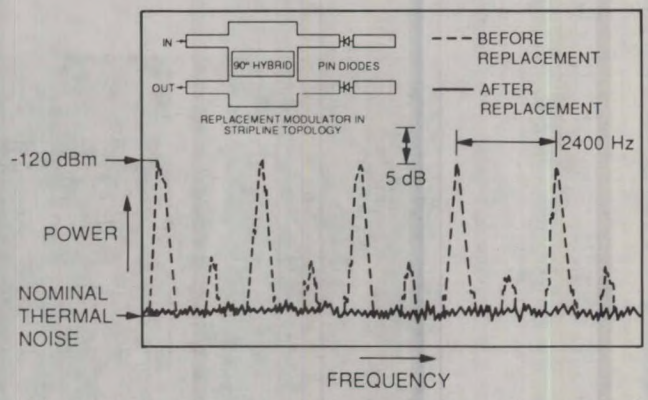


Figure 8. Antenna noise before and after replacement of monopulse modulator

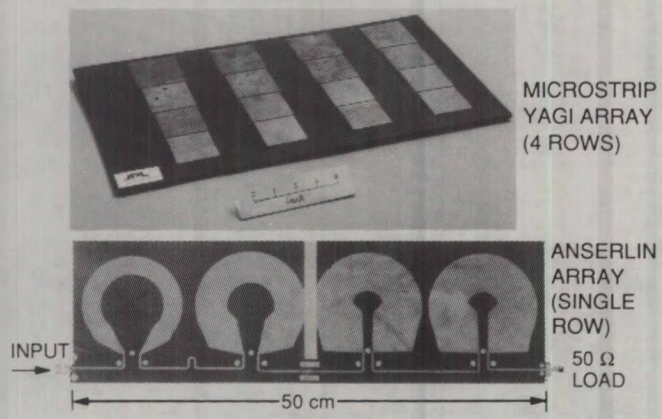


Figure 9. Low cost planar arrays

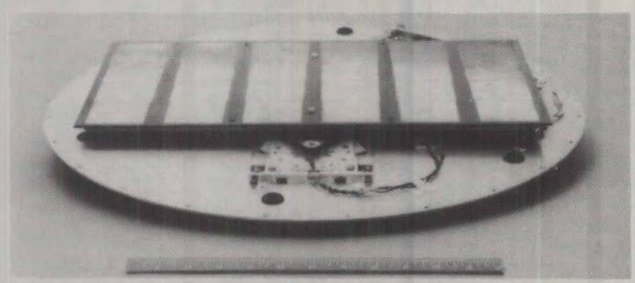


Figure 10. Conformal mechanically steered tracking antenna

An ANSERLIN Array for Mobile Satellite Applications

F. Y. Colomb, D. B. Kunkee, P. E. Mayes, and D. W. Smith
Electromagnetics Laboratory
University of Illinois
Urbana, Illinois 61801
Phone: 217-244-0543

V. Jamnejad
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109
Phone: 818-354-2674

ABSTRACT

Design, analysis, construction and test of linear arrays of ANSERLIN (ANnular SEctor, Radiating LINe) elements are reported and discussed. Due to feeding simplicity and easy construction as well as good CP performance, a planar array composed of a number of such linear arrays each producing a shaped beam tilted in elevation, is a good candidate as a vehicle-mounted mechanically steered antenna for Mobile Satellite Applications.

A single level construction technique was developed that makes this type of array very cost competitive with other low-profile arrays. An asymmetric 19.5" long four-element array was fabricated and tested with reasonable performance. A smaller five-element symmetric array (16" long) was also designed and tested capable of operating in either sense of circular polarization. Mutual coupling, however, seems to be a problem. Efforts have been made to successfully reduce this effect.

INTRODUCTION

(a) System Requirements

Ground-based antennas for mobile satellite applications must meet rather severe requirements, not the least of which is low cost. They must also provide adequate gain for circular polarization and, to operate at an arbitrary location in CONUS, the gain must be maintained at all elevation angles between 20 and 60 degrees above the horizon. Furthermore, the beam must track the satellite throughout maneuvers of the vehicle. Azimuth tracking has been previously demonstrated using mechanically steered antennas[1]. However, the antenna used for that system was a tilted broadside array with a relatively high profile. The goal of the study reported here is to provide a low-profile planar array with equivalent or better performance.

(b) ANSERLIN antennas

The annular sector, radiating-line (ANSERLIN) antennas are well suited to this application[2]. They have low-profile, low axial ratio, are well matched over a wide band, and can be designed to provide wide range of aperture tapers in a single series-fed configuration. The geometry of a single ANSERLIN element is shown in Figure 1. An annular sector of strip conductor over a ground plane is fed by a triangular fin transmission line of the same characteristic impedance. The fin transmission line can be excited at its small end by the center conductor of a coax cable, or by a narrow microstrip line. A second port, identical in geometry to the first, is provided at the other end of the annular conductor. In this manner, the entire antenna is constructed to be of uniform characteristic impedance. When one port is terminated in the matching impedance, the input impedance at the other port is invariant with frequency and equal to the characteristic impedance of the lines comprising the complete structure. Return losses of over 20 dB across bands as wide as 8:1 are typical.

Since no reflections occur at any point along the annular sector, a wave produced by excitation of one port travels around the sector with progressive phase that is proportional to the azimuth angle (ϕ) traversed. Traveling-wave sources with a phase shift of one degree per degree azimuth angle (m' = 1 in exp[-jm'ϕ]) are known to produce a low axial ratio

(AR) over an appreciable portion of the single-lobed radiation pattern. This condition is achieved by the ANSERLIN antenna at a particular frequency. However, a wide pattern bandwidth is achieved when frequency is varied about that frequency. Neither power gain nor axial ratio suffer significantly over bandwidths as great as 30%. The first consequence of changing frequency away from the m' = 1 operating point is a squinting of the beam in the plane orthogonal to the plane of symmetry of the element. For the MSAT application this squint of the element pattern can be used to advantage in achieving the desired low-angle coverage.

ARRAY DESIGN PROCEDURES

(a) Sector Patterns by Fourier Series

A Fourier synthesis procedure is used to calculate the required excitations coefficients of the array for obtaining a sector pattern. Both symmetric and asymmetric coefficients about the center of the array can be obtained using this synthesis procedure.

The asymmetric realizations results in excitation coefficients that are monotonically decreasing along the array, more easily realized with ANSERLIN elements.

(b) Tilted Element Patterns

Using the above synthesis procedure it is possible to find excitation coefficients that would yield a space factor with approximately uniform coverage of the angles from 20 to 60 degrees above the horizon. However, if the element pattern has a maximum at zenith (90 degrees above the horizon), then the space factor must be distorted to emphasize the low angles to compensate for the reduction in directive gain of the element at those angles. This makes the space factor much more difficult to realize with only a few elements. It was thus considered important to make use of the tilted element patterns that are readily achieved by operating ANSERLIN elements at frequencies where m' is either greater or less than unity.

Since the tilt is more pronounced in the case of m' > 1, the initial designs for the MSAT array were carried out for this case. Figure 2 shows typical tilted element patterns for two orthogonal linear polarizations. These patterns were computed using patch basis functions to represent the current on the ANSERLIN element. The expansion coefficients for the current were found by implementing a moment method code on a Cray supercomputer. The computed results are in good agreement with measured patterns. The maximum tilt, at m' = 1.25, is $\theta_0 \approx 15°$. A pattern that is tilted in the opposite direction (m' = 0.8) is shown in Figure 3. In this case the achievable tilt is less and the difference between θ and ϕ polarizations is somewhat greater, but the AR in the coverage region remains very good. In both cases, the field at θ = -70° (20° above the horizon) is several dB higher than it would be for m' = 1.

(c) Computer-Aided Synthesis

The synthesis procedure produces excitation coefficients that give a radiation pattern that approximates the desired one. However, due to the limited number of elements, this procedure provides only a rather rough approximation of the pattern and there is still room for significant improvement by minor adjustments in the excitation coefficients. Rather than using a time-consuming experimental procedure to make these adjustments, a numerical method was used. The procedure is built around a computer program that minimizes a multivariable non-linear function. The routine performs a minimization of a performance index, PI, which is the single-valued sum of the differences of squares

$$PI = \Sigma (a_n^2 - b_n^2), n = 0,1,...,N$$

where a_n represents one of N values of the calculated magnitude of the array pattern at an elevation angle, θ_n, and b_n represents the magnitude of the desired pattern at the same location. To simplify matters, uniform spacing is assumed during the computer optimization of the array design. The element spacing is fixed and an initial set for the excitation coefficients is determined. The starting point can be obtained by using an approximate synthesis technique, such as referred to in (a).

In the program written to implement this process, the desired pattern can be entered in one of two ways. The simplest is to enter a sector pattern. To do this, the upper and lower limits of $\cos(\theta)$ are entered with the number of sample points. The program then constructs a desired pattern with a magnitude of unity over the desired limits of θ and zero elsewhere. The second method is useful in fine

tuning the pattern. it allows the normalized value of the desired pattern to be entered at incremental steps of θ. This is particularly appropriate MSAT array design which requires coverage at lower elevation angles. The beam can be "pushed" toward the horizon to improve coverage at these lower angles. This is done by increasing the weight of the error function at lower elevation angles. This technique was employed to obtain the excitation coefficients used in producing the pattern shown in Figure 4.

(d) Realization of Excitations Through Scattering Parameters

Since each ANSERLIN element is a two-port device, it is convenient to characterize its performance in the feeder network of the array in terms of scattering parameters. Realization of the excitation coefficients is then attempted by means of variations in the ratios of outer to inner radii of the annular sectors[3]. At m' = 1.25, the magnitude of S_{21} can be changed between approximately 0.5 and 5.0 dB by varying the diameter ratio between 1.5 and 7.0. The particular ratio that is required to obtain a given magnitude for S_{21} is determined from data acquired by measuring the scattering parameters of several ANSERLIN elements[3].

Neglecting mutual coupling, the S_{12} parameters across the array elements can be easily related to their excitation coefficients using simple power relations. Once the excitation coefficients are determined by the synthesis procedure, the required scattering parameters are found which are then used in deciding the geometry of the elements.

(e) Progressive and Quasi-Progressive Phasing

One drawback in using a tilted beam with m' > 1 is the size of the ANSERLIN elements. The overall size of the array can be reduced significantly by using an element with m' < 1. Another factor not considered in the above procedures, is the possibility of reducing production costs by being able to use the same array for either sense of circular polarization. A series-fed array of ANSERLIN elements is capable of giving identical shape of radiation patterns for both right-hand and left-hand senses if the array is made so that it is symmetrical about its center. However, for an array limited to only four or five elements, the number of parameters that can be varied to optimize the design becomes very limited. A preliminary design has been done for a symmetric array of five elements with m' < 1. The elements were assumed identical so the only variable parameters are the magnitude of S_{21} and line lengths between elements. For simplicity, the feedline lengths were first taken to be equal (progressive phasing), and subsequently adjusted to give very near endfire condition at the lowest frequency (1.545 GHz). The array factor so obtained produces a high backlobe. It was determined, however, that by introducing a symmetric deviation from progressive phase this lobe can be greatly reduced.

(f) Non-Uniform Spacing

One way to add degrees of freedom to the design of symmetric arrays is to relax the requirement for constant spacing. This was considered worthy of consideration since the uniform spacing of ANSERLIN elements operating with m' < 1 is necessarily very close to one-half wavelength. As a result, an array with near-endfire phasing will have a space factor with a grating lobe intruding into the visible region. Some numerical experiments were performed with non-uniform spacing, but the most effective way to reduce the grating lobe appeared to be the reduction of inter-element spacing.

ARRAY CONSTRUCTION

(a) Microstrip Feed Lines

Although the original ANSERLIN elements were fed and terminated with small-diameter coaxial cable, this method is not conducive to a low-cost mass-produced fabrication process. Hence, all the array designs incorporate microstrip lines at the input and output, and in between the elements. The design of the transitions from microstrip to the annular sector is best accomplished with the aid of a time-domain reflectometer (TDR). After designing the components (microstrip, triangular fin, and annular strip) separately, to have the same characteristic impedance, they are joined to one another. The TDR display then indicates regions of the composite structure where significant deviations from the characteristic impedance occurs which can be fixed by slight alterations in the geometry in those regions.

(b) Double-Level Arrays

The microstrip feedlines must be placed very close to the ground plane to provide the proper

impedance in a reasonable width and to prevent radiation. However, the annular sector of the ANSERLIN element has a higher separation from the ground plane to enhance the radiation. Originally, in ANSERLIN array designs, the microstrip lines and the annular sectors were placed on different levels. This technique is still preferred during initial stages of testing a newly designed array since it facilitates the alteration of spacing and feedline lengths. However, mass production of such arrays is not desirable.

(c) Single-Level Arrays

A new construction technique has been developed for this project which reduces the complexity and makes ANSERLIN arrays cost competitive with other approaches. The entire array, including the interconnecting microstrip phasing lines, is etched on one side of a thin dielectric substrate. Ground, plane for the phasing lines is provided on the back side of the substrate. However, the conductor is etched away underneath the radiating elements to provide larger spacing between them and a second ground plane. This variation in ground plane height inhibits radiation from the phasing lines and enhances radiation from the ANSERLIN elements. Metallic ramps are used under the input and output transitions between the phasing and radiating lines. The impedance at any point in the array, from feed end to termination, is maintained very close to the desired constant value. Hence, the input impedance is very nearly constant over a very wide band.

Figure 5 shows the artwork that was used to construct a four-element array. Note that m' > 1 type elements used in this array require large inter-element spacing approximately equal to 0.60 wavelength. Experience with previous arrays indicated that coupling between elements should not be a problem with this spacing. The difference in the geometry of the elements should also be noted. In this way, not only the main beam of the pattern is shaped but the grating lobe is also effectively controlled.

(d) Symmetric Arrays with Dual Polarizations Capability

Some drawbacks of the array of Figure 5 were addressed by studying symmetric arrays using ANSERLIN elements operating at m' < 1 mode. These elements are significantly smaller than those of Figure 5. However, the range of variation in the magnitude of S_{21} is also much smaller. This, together with the requirement for symmetry, leads to the conclusion that the practical elements of this variety will be almost identical. Hence, the starting point for designing arrays of this type is to take both the element geometry and the inter-element spacing as constant.

Figure 6 shows computed patterns for θ and ϕ polarization at mid-band for a symmetric five-element array using quasi-progressive, near-endfire phasing. The major problem with this design is the θ-polarized lobe near 20 degrees above the horizon ($\theta = 70°$) in the backlobe region. The problem is most severe at the lower band limit and essentially disappears at the upper band limit. A controlled deviation from progressive phasing (note the differences in the feedline lengths) has been used to reduce the response over most of the reverse coverage region, but the intrusion of the grating lobe into the visible region does not readily respond to this particular manipulation. However, it was considered unwise to be overly concerned about the theoretical behavior of the array at very low angles since the influence of the finite ground plane will alter significantly the array performance in that region.

TEST RESULTS

(a) Single-Level Array with Synthesized Pattern

Plots of realized gain versus elevation angle were obtained for the array of Figure 5 at JPL test facility. It was found that the antenna performed best over a band that was about three percent lower in frequency than desired. This is likely due to effects of the substrate material on the phase shift through the elements, an effect that was only estimated due to the lack of experimental data. Figure 7 shows the pattern measured at 1.55 GHz, the frequency corresponding to mid-band. The maximum gain is approximately 12.0 dB relative to a circularly polarized isotropic source, and the gain at 60 degrees above the horizon is about 9.0 dB, both well above the desired value of 6 db. However, the gain at 20 degrees above the horizon is close to 0 dB, and decreases further as frequency decreases. While this is lower than desired, the influence of the particular ground plane used in these measurements is certainly the greatest in this part of the pattern.

At JPL, the array was mounted on a 48 x 56 inch rectangular ground plane. Pattern measurements

made at the University of Illinois with the same array mounted on a 36-in diameter circular ground plane agreed very well with the JPL data, except near the horizon. On the smaller ground plane the field at 20 degrees above the horizon was only 8 dB below the maximum, rather than almost 12 dB as shown in Figure 6. Other aspects of this pattern look particularly good. The level of cross polarization is quite low across the entire coverage region and the secondary lobes are down by 10 dB or more. As frequency increases, the coverage at 20° elevation improves, but the level of the lobes increases. The minor lobes are down by more than 15 dB at 1.475 GHz, but the gain at 20° above the horizon is -5 dB. A more realistic evaluation of performance at low angles can probably best be made by utilizing a ground surface that more closely approximates the size and shape of the vehicle in the vicinity of the mounting location. The measured return loss was approximately 10 dB over the entire band. The fact that this is significantly lower than the values observed for individual ANSERLIN elements is likely due to mutual coupling.

(b) Smaller Symmetric Array - Coupling Effects

Some preliminary measurements have been done on the array outlined in Figure 8. Note that the overall length of this five-element array is 16 inches as opposed to 19.5 inches for the four-element array of Figure 5. This reduction in size is made possible by using elements operating near m' = 0.8 so that they are much smaller in terms of the operating wavelength. To increase the radiation from these smaller elements, the height above the ground plane was increased beyond that used in the four-element array. The measurements on the array indicate detrimental effects possibly due to mutual coupling. The return loss is reduced and the pattern shape, particularly in the backlobe region, does not agree well with the computed array patterns which use the actual patterns of isolated elements but ignore coupling. The return loss was increased and better agreement with theoretical patterns were obtained after surrounding each element by a conducting fence. For a more practical solution, attention is being given to reducing the height, and perhaps adding very short fences (ridges in the ground plane).

CONCLUSIONS

Design, construction and test of a four element array of ANSERLIN elements demonstrated the successful application of synthesis methods that yield excitation coefficients for a difficult-to-realize radiation pattern. A single-level construction technique was developed that makes the production of array of ANSERLIN elements competitive in cost with arrays of other types of low-profile elements.

Work with a smaller five-element array has shown that reasonable pattern shaping can be achieved with a symmetric array that is capable of operating in either sense of circular polarization. Using smaller ANSERLIN elements produces potential problems with mutual coupling. Achieving good realization with present design procedures, that ignore mutual coupling, will depend upon developing practical ways to reduce the effects of the coupling. Some encouraging results in this direction have already been obtained.

Most of the requirements for MSAT applications, particularly those for low axial ratio and high return loss, have been met or exceeded by these arrays. The gain and coverage requirements have also been satisfied over a large portion of the frequency band. Several promising ideas, not yet been fully explored, indicate that ANSERLIN arrays can be successfully employed on mobile vehicles in the MSAT system.

ACKNOWLEDGMENT

This work was carried out jointly by the University of Illinois, Urbana, and the Jet Propulsion Laboratory, California Institute of technology, under contract with National Aeronautics and Space Administration.

REFERENCES

1. **Jamnejad, V.** 1988. A mechanically steered than tracking antenna for PiFEx. MSAT-X Quarterly No. 13, Special Issue on PiFEx Tower-l Experiments. (Pasadena, California; The Jet Propulsion Laboratory), pp. 18-27.

2. **Drewniak, J. L., Mayes, P. E.** 1989. ANSERLIN: a broadband, low-profile, circularly polarized antenna. IEEE Trans. Antennas Propagat., vol. AP-37, pp. 281-288.

3. **Drewniak, J. L., Mayes P. E.** 1990. The synthesis of patterns using a series-fed array of ANnular Sector Radiating Line (ANSERLIN) elements, low-profile, circularly polarized, radiators. To be published in IEEE Trans. Antenna Propagat.

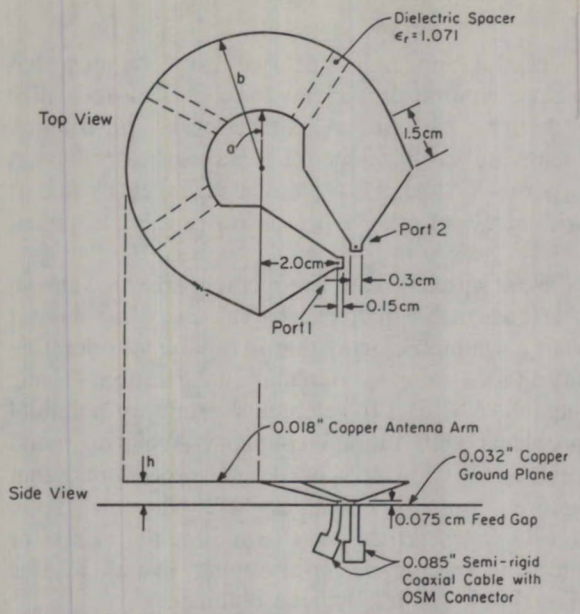


Figure 1. Geometry of an ANSERLIN element

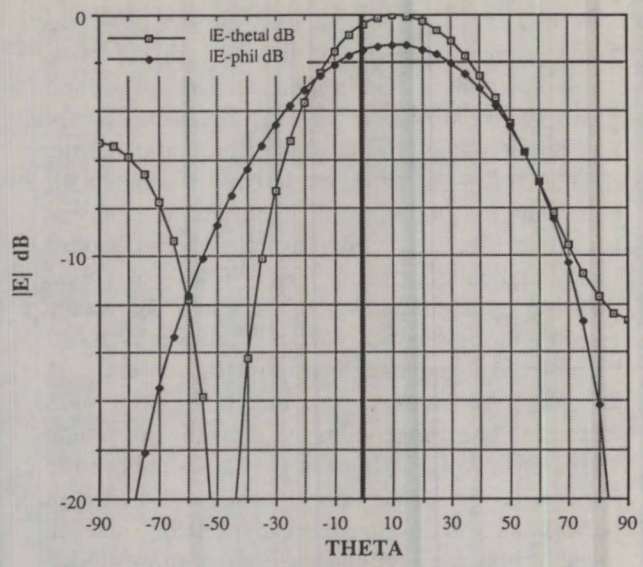


Figure 3. A typical element pattern in the plane of the array at 1.6 GHz (m' = 0.8).

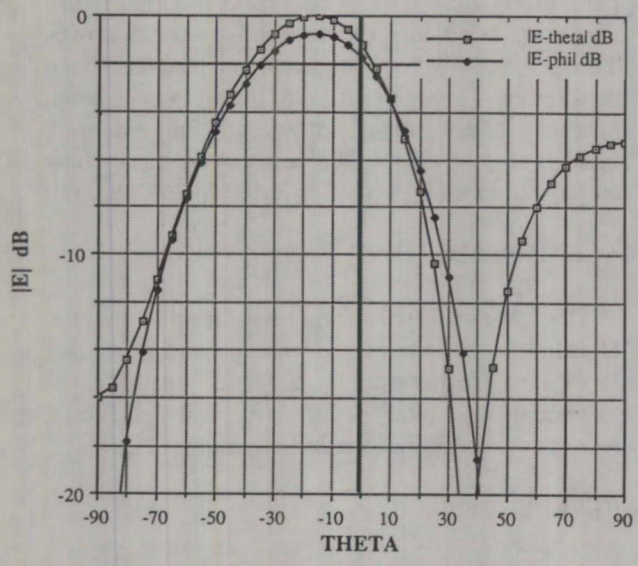


Figure 2. A typical element pattern in the plane of the array at 1.6 GHz (m' =1.25).

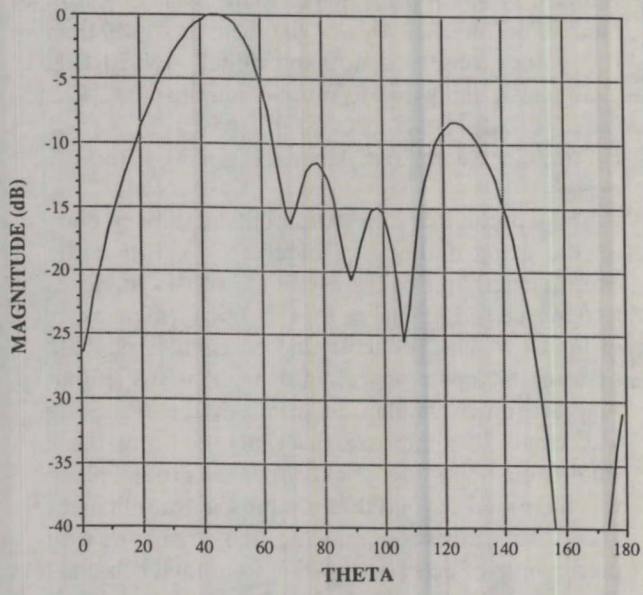


Figure 4. Predicted elevation patterns of 4-element array at 1.55 GHz.

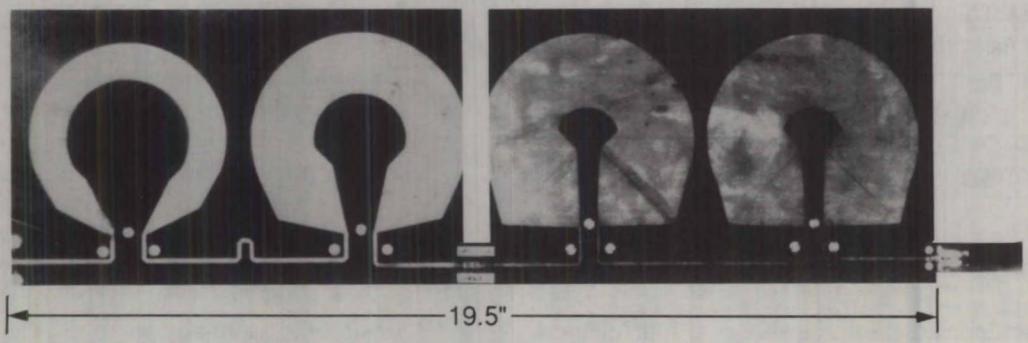


Figure 5. A pictorial view of the fabricated 4-element array.

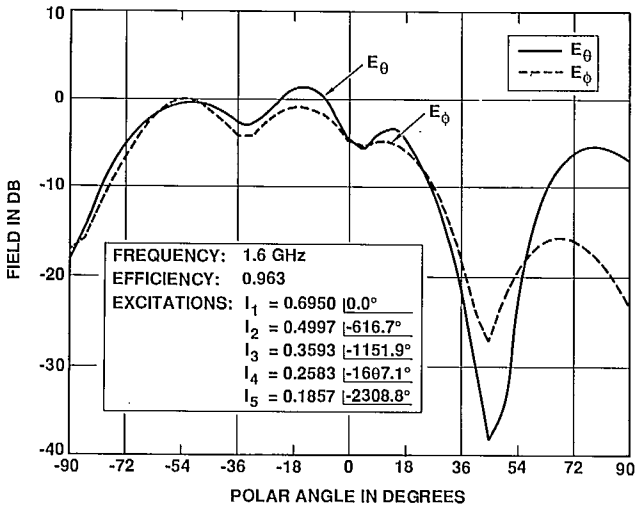


Figure 6. Computed elevation patterns of the 5-element ANSERLIN array at 1.6 GHz.

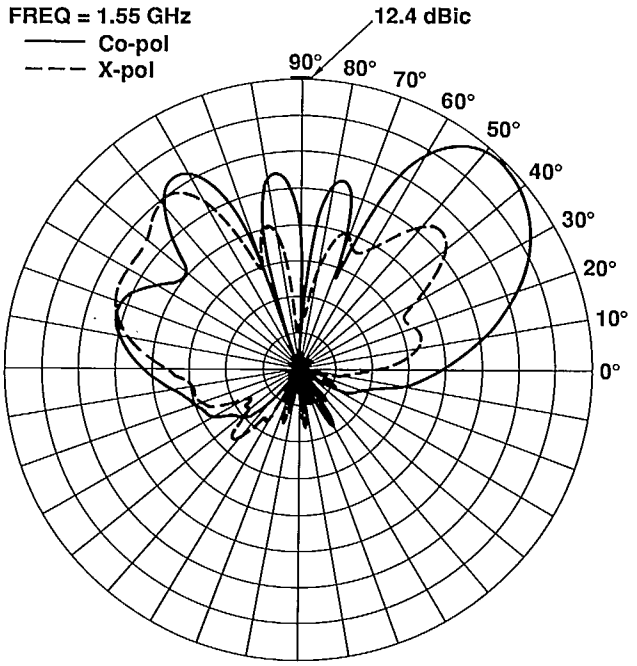


Figure 7. Measured elevation patterns of the 4-element ANSERLIN array at 1.55 GHz, on a 48" x 56" ground plane.

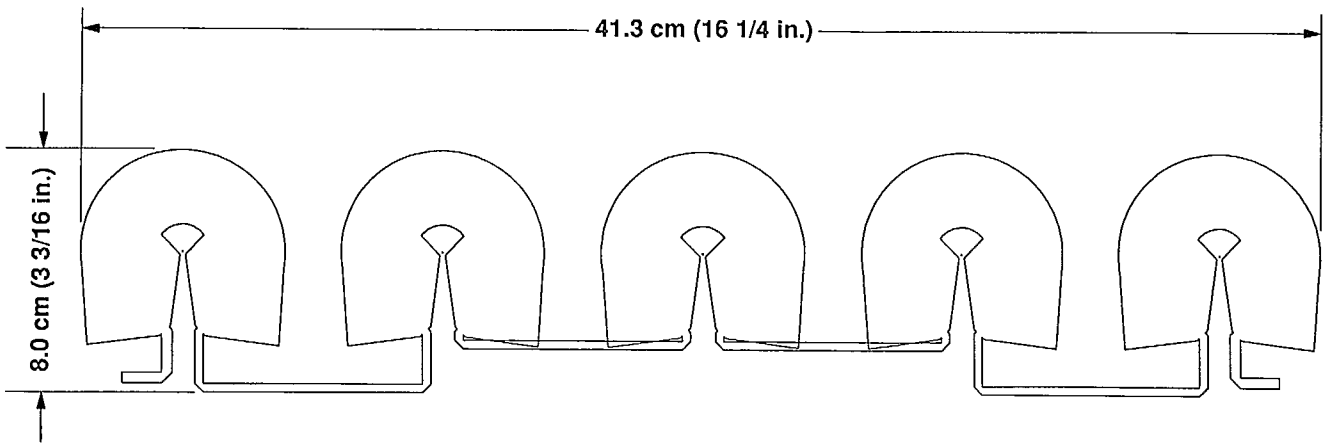


Figure 8. Outline of the symmetric 5-element array.

Microstrip Yagi Array for MSAT Vehicle Antenna Application

John Huang and Arthur Densmore
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

David Pozar
University of Massachusetts
Amherst, Massachusetts 01003

ABSTRACT

A microstrip Yagi array has been developed for the MSAT system as a low-cost mechanically steered medium-gain vehicle antenna. Because its parasitic reflector and director patches are not connected to any of the RF power distributing circuit, while still contributing to achieve the MSAT required directional beam, the antenna becomes a very efficient radiating system. With the complete monopulse beamforming circuit etched on a thin stripline board, the planar microstrip Yagi array is capable of achieving a very low profile. A theoretical model using the Method of Moments has been developed to facilitate the ease of design and understanding of this antenna.

INTRODUCTION

A major element of the Mobile Satellite (MSAT) program has been the development of several types of medium-gain L-band vehicle antennas. Currently, two medium-gain antennas have been developed and field tested with successful results. One is the electronically steered planar phased array[1], and the other is the mechanically steered 1x4 tilted microstrip patch array[2]. The phased array offers the advantage of low profile (one-inch tall) and beam agility at the expense of higher production cost. On the other hand, the mechanically steered 1x4 array antenna offers the lowest production cost with the highest profile (six-inch tall).

To combat the disadvantages of high cost or high profile of the above antennas, a new antenna concept is being introduced, which not only offers both the advantages of low profile and low cost, but shows excellent efficiency in its beamforming circuitry as well. This antenna is a mechanically steered planar Microstrip Yagi array[3]. It is composed of twelve parasitic director and reflector patch elements and four driven elements. The reflector and director patches, based on Yagi's principle[4], tilt the array's broadside beam towards endfire for satellite pointing. Because only the few microstrip driven element patches are directly connected to and driven by the RF power distributing circuitry, the complexity and RF loss of the power distributing circuit are dramatically reduced, resulting in improved antenna noise temperature and antenna gain. The overall height of the antenna is only 1.5 inches, which includes both the RF portion and the mechanical rotating platform. The rotation in azimuth is accomplished by a thin pancake motor. The control of this motor, or the beam pointing, is done by a monopulse system that is uniquely designed for the

communication antenna to provide simultaneous transmit and receive signals.

To facilitate the ease of design and have a complete understanding of this microstrip Yagi array, a theoretical model based on the Method of Moments has been developed[5]. The purpose of this article is to present both the theoretical and experimental results of the MSAT microstrip Yagi planar array antenna, as well as its monopulse satellite tracking system.

Principle of Microstrip Yagi Array

The microstrip Yagi array antenna, as illustrated in Figure 1, consists of a driven patch element and a few parasitically coupled reflector and director patch elements. It utilizes the same principle as a conventional dipole Yagi array where the electromagnetic energy is coupled from the driven element through space into the parasitic elements and then re-radiated to form a directional beam. The sizes of the parasitic dipole elements and their relative spacings are designed so that the phases of all the elements are coherent toward the end-fire direction (in the direction of the director elements). For the microstrip patches to have the same Yagi effect as the dipole elements, the adjacent patches need to be placed very close to each other so that significant amount of coupling can be formed through surface wave and radiation mechanisms. It was found experimentally[3] that the gap distance between two patch edges should be equal to or less than the dielectric substrate thickness. Since the amount of surface wave is a strong function of the dielectric constant and substrate thickness, the pattern shape of the microstrip Yagi array is also a function of these two parameters. In addition, the sizes of the parasitic reflector and directors are factors in determining the pattern shape and beam direction. When the size of the parasitic patch is different from the resonant dimension (driven element size), a reactance component is added into the radiation impedance of this parasitic patch, which introduces an additional phase to its radiation field. It is this phase, together with the coupling phase, that contribute to the Yagi array's beam tilt from its broadside direction. Because the microstrip antenna requires a ground plane, the beam direction of the Yagi never becomes true end-fire. This beam tilt angle from the broadside, as explained above, is a function of dielectric constant, substrate thickness, patch separation, and reflector and director sizes. In general, the patch separation and parasitic patch size govern the relative phase, while the gap distance determines the current amplitude of each Yagi element.

Several important and interesting characteristics of the microstrip Yagi antenna have been found through experiments, first is that, if good circular polarization (c.p.) is desired in the Yagi's main beam region, the phase difference between the two orthogonal feeds needs to be deviated from the conventionally required 90-degree. This is due to the fact that the main beam is no longer in the broadside direction, and the E-plane and H-plane coupled fields have different coupled phase terms. Figure 2 shows the circularly polarized microstrip Yagi array pattern where the two orthogonal probes are fed 90° apart in phase. Figure 3 illustrates the pattern of the same array with improved c.p. when the two probes are fed 115° apart in phase. The second interesting finding is that, because the E-plane (along the array axis) has a much stronger surface wave coupling than the H-plane, the input impedance of the E-plane probe is strongly affected by the parasitic patches. As a result, to achieve the nominal 50 ohm input impedance match, the position of the E-plane probe in the driven element of the Yagi array is different than that of the

classical single microstrip patch. This re-adjusted feed position has been found to be toward the edge of the patch. Finally, the microstrip Yagi array has a wider bandwidth than the driven element by itself. This is because the parasitic director and reflector, having different sizes from the driven element, are causing a "log-periodic frequency independent" phenomenon to occur. Figure 4 is a Smith chart that gives the input impedance as a function of frequency for a single patch (2.2-inch square) with dielectric constant of 2.5 and substrate thickness of 0.25 inch. It shows a single-loop response with a relative narrow bandwidth. Figure 5 gives the input impedance plot of the same element with H-plane coupled parasitic director and reflector patches. The plot shows a double-loop response which has a double-resonance and covers both the MSAT's transmit and receive frequencies.

Theoretical Analysis

The microstrip Yagi array is analyzed using the full-wave moment method solution, similar to solutions that have been previously used for a variety of microstrip antenna problems[5]. Since the parasitic elements in the Yagi are excited solely through mutual coupling from the driven element, it is necessary that the analysis be able to predict mutual coupling in an accurate manner. Surface wave prediction is especially important here because the gap distance between adjacent patches is relatively small and hence, cause a strong surface wave coupling. The solution also models the probe feed, as well as the circularly polarized dual feeds.

A computer code has been developed for the analysis of the microstrip Yagi geometry having one driven element and up to four parasitic elements with arbitrary sizes and spacings. The software uses the moment method solution to compute the currents on the patch elements; from these quantities the input impedance, far-field patterns, and directivity can be determined. The analysis has shown that the director patch has a much stronger coupling than the reflector's coupling from the driven element. This is why, as illustrated in Figure 1, two directors are used to effectively tilt the beam, while only one reflector is needed. Adding more reflectors to the array does not achieve any significant effect. The comparison between the calculated and measured patterns of a single column of microstrip Yagi array at the frequencies of 1552 MHz are shown in Figure 6. The dimensions of this antenna are given in Figure 7. The agreement of the co-pol is quite good. The disagreement, especially in the cross-pol, is mostly attributed to the fact that the measurement was performed on a finite ground plane (5-wavelength diameter), while the calculation was carried out with an infinite ground plane.

Description of Antenna and Monopulse Beamformer

In order to meet the MSAT required minimum gain of 10 dBic between 20° and 60° elevation angles, four microstrip Yagi array columns are needed. The dimensions of each column are given in Figure 7, and the spacing between adjacent columns is 4.5 inches (0.61 wavelength at 1.6 GHz). The overall antenna aperture is a rectangle with dimensions of 10 by 17 inches. A 1/8-inch thick stripline monopulse beamformer located flush below the array, is designed to provide uniform excitations through feed probes to the 4 driven elements of the Yagi array. The RF block diagram of the overall antenna is shown in Figure 8, and its corresponding stripline beamformer layout is presented in Figure 9. A pictorial view of the fabricated antenna and beamformer is illustrated in Figure 10.

In Figure 9, the antenna feed network distributes RF power through Wilkinson dividers to the four driven elements of the Yagi array and establishes the required differential phase (115° in this case) between the two orthogonal feeds of each driven element. With the 180-degree hybrid coupler combining the incoming signals from the dividers, the monopulse sum and difference signals are generated. The bi-phase modulator implements a half-wavelength phase delay that is switched on and off by the pin diodes at a frequency around 1 KHz. With the simple single-line bias design, no slip rings are needed. Both the RF and diode control current go through a single-channel rotary joint. The modulated difference signal is coupled through a 10-dB coupler onto the sum channel, which are together sent down to the rotary joint. The purpose of the filter is to stop any portion of the transmit signal, coming up from the rotary joint, to the difference channel.

The antenna and the beamformer are together mounted onto a half-inch thick pancake motor that is controlled by the monopulse feedback pointing error signal to track the satellite. The complete antenna system is covered by a radome having a height of 1.5 inch and a diameter of 21 inches. Presented in Figures 11 and 12 are the measured antenna patterns at the receive and transmit frequencies, respectively. One noticeable shortcoming of the Yagi array is its high backlobe over a relative wide frequency bandwidth. Fortunately, investigation has led to the conclusion that these high backlobes will not present any adverse effect to the MSAT system in terms of tracking, multipath, or user interference. Certainly, power is being wasted in these backlobes. However, due to the excellent efficiency of the Yagi array with only 1.5 dB of loss in the beamformer, the required 10 dBic of gain has been achieved as shown in Figures 11 and 12.

Conclusion

A mechanically steered antenna with low cost, low profile, and low insertion loss has been developed for the MSAT system. This antenna combines the excellent efficiency of the microstrip Yagi concept and an elegant design of the monopulse beamformer to achieve the required antenna gain and high G/T ratio. Experiment has indicated that, with a slightly different array design, the microstrip Yagi antenna can also provide the required performance for the Canadian region where a lower elevation coverage is encountered.

ACKNOWLEDGMENT

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. Huang, J. 1987. L-band phased array antennas for mobile satellite communications. IEEE Vehicular Technology Conference, PP. 113-117.

2. Jamnejad, V. 1988. A mechanically steered monopulse tracking antenna for PiFEx: RF system design. MSAT-X Quarterly, No. 13, JPL 410-13-13.

3. Huang, J. 1989. Planar microstrip Yagi array antenna. IEEE AP-S/URSI Symposium digest, PP. 894-897.

4. Yagi, H. 1928. Beam transmission of ultra short waves. Proc. IRE, Vol. 16, PP. 715-741.

5. Pozar, D. M. 1982. Input impedance and mutual coupling of rectangular microstrip antennas. IEEE Transactions Antennas & Propag., Vol. AP-30, No. 6, PP. 1191-1196.

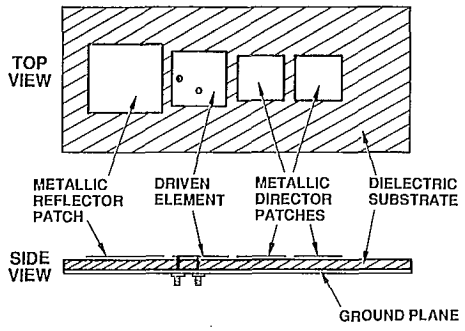


Figure 1. Microstrip Yagi array configuration.

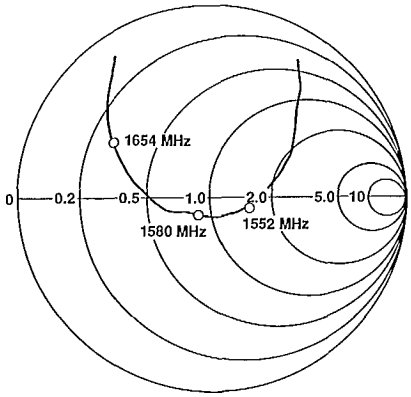


Figure 4. Input VSWR of a single microstrip patch.

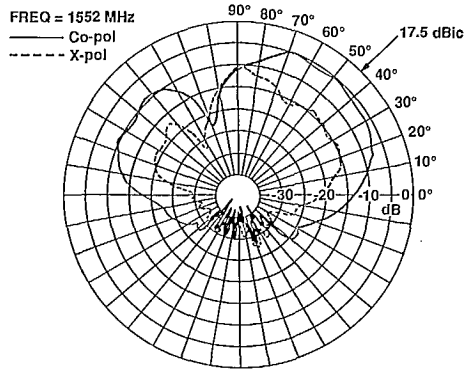


Figure 2. Measured C.P. microstrip Yagi array elevation patterns with 2 feeds separated 90° in phase.

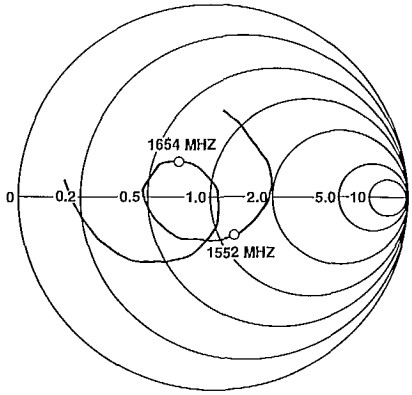


Figure 5. Input VSWR of a microstrip patch with parasitic Yagi patches.

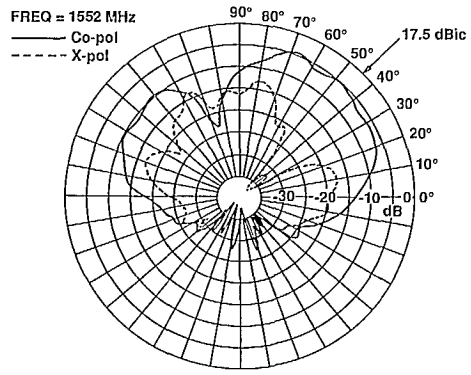


Figure 3. Measured C.P. microstrip Yagi array elevation patterns with 2 feeds separated 115° in phase.

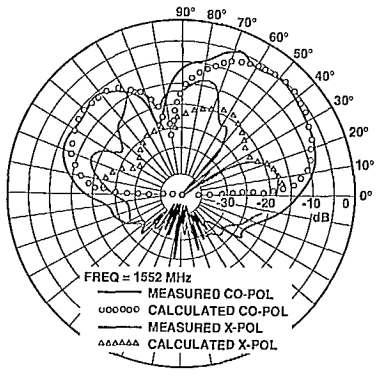


Figure 6. Comparison of calculated and measured patterns of the microstrip Yagi array shown in Figure 7 at 1552 MHz.

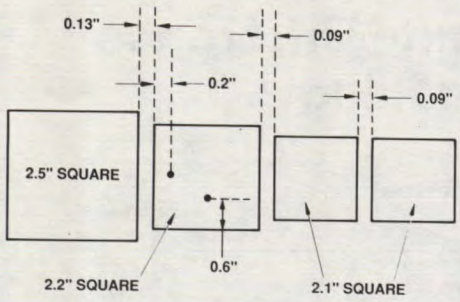


Figure 7. Dimensions of a single column L-band microstrip Yagi array.

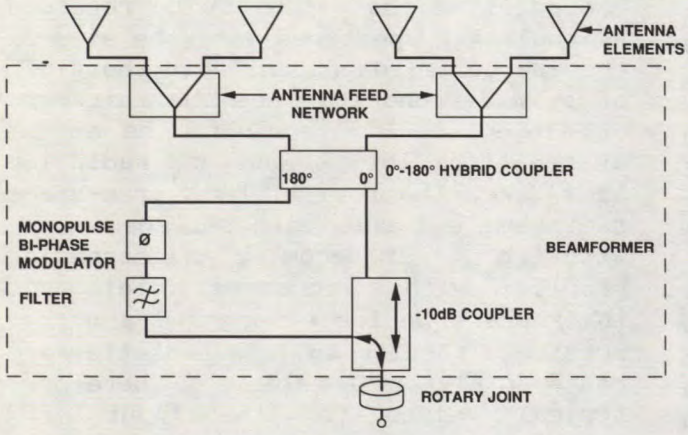


Figure 8. Block diagram of antenna and beamformer.

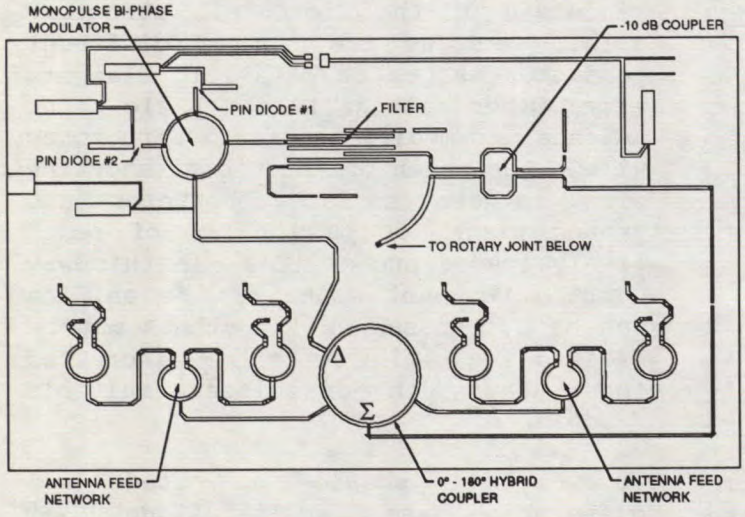


Figure 9. Stripline monopulse beamformer circuitry.

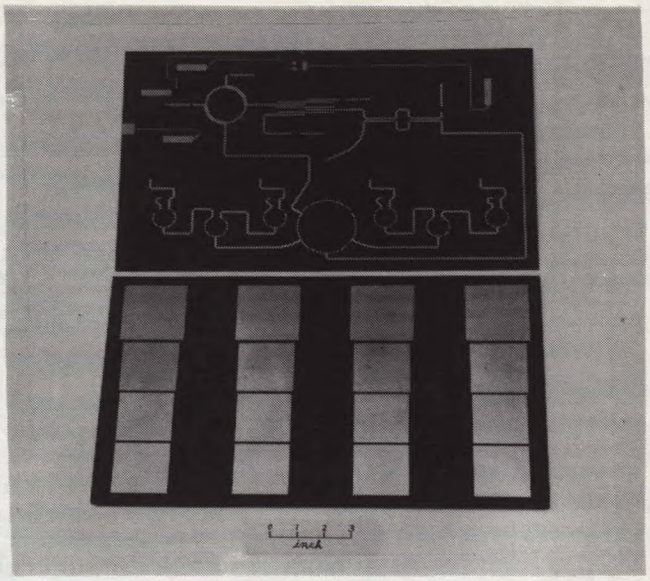


Figure 10. MSAT microstrip Yagi antenna and beamformer.

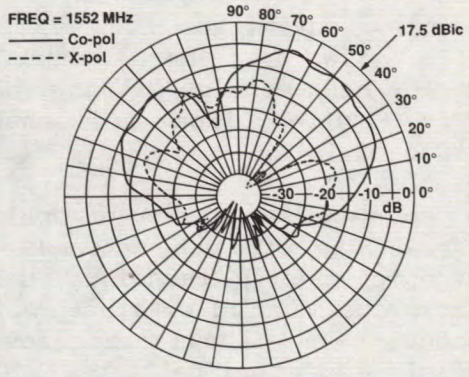


Figure 11. Measured elevation pattern of the MSAT microstrip Yagi antenna at 1552 MHz.

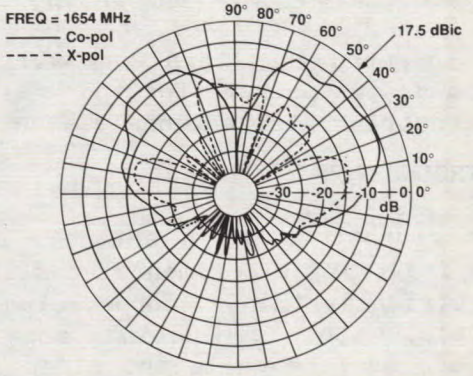


Figure 12. Measured elevation pattern of the MSAT microstrip Yagi antenna at 1654 MHz.

The Analysis of Reactively Loaded Microstrip Antennas by Finite Difference Time Domain Modelling

Hilton, G.S., Beach, M.A., & Railton, C.J.
Centre for Communications Research,
Faculty of Engineering,
University of Bristol,
Bristol, U.K.
Tel. Int.: +44 272 303258
Fax. Int.: +44 272 251154

ABSTRACT

In recent years much interest has been shown in the use of printed circuit antennas in mobile satellite and communications terminals at microwave frequencies. Although such antennas have many advantages in weight and profile size over more conventional reflector/horn configurations, they do, however, suffer from an inherently narrow bandwidth.

This paper examines a way of optimising the bandwidth of such antennas by an electronic tuning technique using a loaded probe mounted within the antenna structure, and shows the resulting far-field radiation patterns. Simulation Results from a two-dimensional Finite Difference Time Domain (FDTD) model for a rectangular microstrip antenna loaded with shorting pins will then be given, and compared to results obtained with an actual antenna.

It is hoped that this work will result in a design package for the analysis of microstrip patch antenna elements.

INTRODUCTION

The microstrip patch antennas used in the following experiments consist of an electrically-thin, copper-clad substrate, with a conducting patch etched on one surface while the other is left as a ground plane. Thus, the structure resembles a thin, open-circuited resonant cavity whose losses are due mainly to radiation from the patch, but also to a lesser extent from both the conductor and dielectric.

The edges of the patch can be regarded as radiating apertures with the size of the aperture (dependant upon the width of surface conductor and the thickness of dielectric) determining the amount of radiation loss. However, radiation loss takes the form of both free-space radiation and substrate surface waves, with the latter becoming an increasing problem with increased dielectric thickness. On the other hand a narrow substrate results in less radiation and hence a higher Q factor. Therefore, typical values for the input VSWR bandwidth of a rectangular patch microstrip antenna are in the region of 1 - 4 % .

Various schemes to increase the bandwidth of the microstrip structure exist; such as the use of different conductor shapes or parasitic elements [1]. Other methods take a basic patch antenna and modify the field structures within the patch by the use of shorting pins between surface conductor and ground plane, or by the use of reactively loaded probes [2]. In this way certain resonant modes of the antenna can be suppressed whilst others may be excited, resulting in either increased input bandwidth or closer multiple resonances.

However, it should be noted that although these modifications may improve the input feed/antenna match over a wider frequency, this may result

in the formation of unsuitable far-field radiation pattern and/or no real increase in main beam gain. Another drawback might be the increase of cross-polar products, hence a compromise between input match and radiation fields may have to be sought.

PROJECT OVERVIEW

The overall aim of the project is to develop a method of analysing microstrip antenna elements and arrays for use in mobile satellite and communication terminals. These antennas may have to be frequency adjustable and/or multi-resonant.

For instance, with the INMARSAT Standard-C satellite system, it was envisaged that a single, high Q patch could be used to cover the bands 1530-1545 MHz and 1626.5-1646.5 MHz - a basic circular polarised square patch design has an input bandwidth of greater than 20 MHz but does not produce a suitable match at both the required bands. Therefore basic dual feed, 'single' frequency, microstrip patches of the type previously used at Bristol University for satellite signal acquisition [3,4], would not be suitable, so a more complicated structure would be required.

LOADED PROBE MICROSTRIP ANTENNA

The initial task was to prove that the use of a single, reactively loaded probe within a microstrip patch element could produce a suitable tuning capacity to give an improved broadside gain response. Methods of generating circular polarisation are to be dealt with at a later stage in the work, and are not included in this paper.

A series of tests were carried out on a rectangular microstrip patch antenna etched on a RT/Duroid substrate with a dielectric permittivity of 2.2 . The dimensions of the patch were chosen purely to fit the grid of the mathematical model to be described later in this paper, and not for any specific frequency requirement other than general operation in the L band.

The antenna consisted of a rectangular patch, 86.4 x 57.6 mm, etched onto a 1.6 mm thick board. Two probes were placed within the substrate and attached to SMA connectors on the ground plane; the positions of the feed and load probes are shown in Fig. 1. The feed point was chosen to give a relatively good input match and low excitation of modes giving rise to cross-polar radiation, while the load position was experimentally determined such as to not encourage cross-polar radiation and yet still maintaining a useful tuning operation.

In order to vary the reactive loading on the probe, a phase-shifter was attached to the loading port. This consisted of a hybrid coupler incorporating two varactor diodes connected to ground in the output arms, while the 'isolated' port remained open-circuited.

By variation of the bias voltage applied to the phase-shifter, the effective length of the open-circuit is either increased or decreased resulting in a change in the reactive load as seen by the antenna. On full bias the total phase shift, at the original antenna resonance frequency, was increased by approximately 130°.

Results for loaded probe antenna

Measurement of the input feed/antenna match of the microstrip antenna revealed a VSWR of less than 1.1 for the unactivated phase shifter at the resonant frequency of 1674 MHz; the bandwidth (for an input VSWR<2) being 23.3 MHz. Activation of the phase shifting network, giving a maximum loading of the antenna, moved the resonant frequency to 1708 MHz, with a corresponding VSWR of 1.33 and a bandwidth of 18.6 MHz.

Before activation of the phase-shifter, the VSWR at 1708 MHz had been 5.5; therefore, this improved match should have resulted in 2.8 dB more radiated power. Subsequent measurement of far-field radiation patterns gave an

improvement in broadside gain of 2.6 dB. However, the received power at this frequency was still 2 dB less than that achieved at the original lower frequency. By reduction of the loading at the probe an improvement in gain of 0.8 dB was obtained, although the resonant frequency of the antenna was now reduced to 1695 MHz. Thus, it can be seen that best input VSWR match of the antenna does not necessarily correspond to the best achievable signal strength.

The co-polar and cross-polar far-field radiation patterns for one of the orthogonal antenna planes at 1674 MHz (without loading) is shown in Fig. 2, while the equivalent plots at 1708 MHz (phase-shifter operating with a loading below maximum) are shown in Fig. 3.

It can be seen that the shapes of the two co-polar plots are very similar, indicating that there has been no major excitation of additional modes within the patch. The gain at the higher, tuned, frequency is 1 dB down on the level achieved at the lower frequency (without loading) rather than 4.4 dB as would be the case without the loading. Therefore, there has been a 3.4 dB improvement with the loaded antenna in the gain achieved 34 MHz above the original resonant frequency. In all cases the cross-polar radiation levels were more than 20 dB below their corresponding co-polar responses.

It can be seen that this simple method can be used to optimise signal strengths without causing any major change in radiation pattern shape. Further investigation is now being undertaken to determine the capability of such a method to tune circularly polarised microstrip patch elements.

FINITE-DIFFERENCE TIME-DOMAIN MODELLING

The regularly shaped microstrip patch antennas (e.g. rectangular, circular) can be reasonably matched to a feed transmission line using modal analysis of the structure [5,6]. However, the addition of shorting pins and reactively loaded probes, leading to excitation of non-standard resonances, requires a more general analysis technique.

Both finite-element and finite-difference techniques were considered, though a finite-difference model was initially chosen for evaluation.

The basic algorithms for the finite-difference solution of Maxwell's equations are given below :

$$\nabla \times \underline{H} = \sigma \underline{E} + \varepsilon(\Delta \underline{E}/\Delta t) \qquad (1)$$

$$\nabla \times \underline{E} = -\mu(\Delta \underline{H}/\Delta t) \qquad (2)$$

and described in [7] and [8].

The space surrounding the conducting elements is divided into cubic grids, with the grid meshes usually being concentrated at areas of major field changes, i.e. discontinuities within the structure. A signal (or signals) are applied to the model while the electric and magnetic fields are monitored as a function of time. After many iterations of the basic algorithm the simulation model will settle down to a steady-state condition. At this time the voltage and current at the feed point can be used to determine the impedance at that point.

LOADED MICROSTRIP PATCH SIMULATION

Extensive work has been carried out at Bristol University on the modelling of enclosed microstrip discontinuities [9], and this is being extended to cover perfect absorbing boundaries for use in antenna work. However, for the analysis of shorted microstrip structures a slightly different approach was used that required a coarser and simpler grid, and hence less processing time.

The method used was only required to give a useful trend to the field resonances within a loaded patch structure over a limited bandwidth. It was therefore considered that only a two-dimensional finite-difference time-domain (FDTD) model would be used and hence only three fields (Ez, Hx & Hy) required for processing. As the thickness of the dielectric is small compared with the dimensions of the microstrip conductor, the Ez field can be assumed to be uniform beneath the strip while the microstrip structure can be assumed to have magnetic walls.

Unfortunately, the model does not directly allow for fringing effects at edge discontinuities, but should give a reasonable response for loaded regions within the patch structure. However, by use of the ε and σ terms at the structure boundary, susceptive loading due to evanescent fields, and radiation can in some way be modelled - these values being derived from experimental observation. Fig. 4 shows the grid model used in the initial simulation runs.

For comparison with the simulation model, a rectangular microstrip patch antenna, 86.4 x 57.6 mm, was again used; though this time fed, via a 50 Ω microstrip transmission line, to a rough matching point for the second lowest order natural resonance. The input response was then recorded for a 200MHz region about this frequency.

The results from the practical work were compared with the simulation, and the model modified to give suitable results. Now it could be reasonably assured that edge fringing and coupling had been taken into account.

Simulation and Experimental Results

The next stage involved shorting out areas of both the simulation model and microstrip antenna, and comparing results obtained. Fig. 5 shows the input impedance simulation results for the model for two cases: (1) shorting at region m=8, n=12, and (2) shorting at regions m=8, n=12 and m=11, n=14. Results obtained by locating shorting pins in equivalent positions in the microstrip antenna are given in Fig. 6. Both sets of results have been adjusted to the impedance monitor point on the feed, Fig. 4.

Location of the first pin produced little effect, but did improve the antenna/feed match slightly and increased the resonant frequency by 2 MHz. Therefore, a second pin was added to cause a larger change in the input response. This time the match was made worse and the resonant frequency increased by a further 7 MHz.

Although the simulation model was originally tuned to give a good response for the unloaded antenna over approximately 200 MHz region around resonance, there was still a good correlation between the theoretical and experimental results for the loaded microstrip cases. The resonant frequencies obtained were only a few MHz adrift, while the Q of the experimental antenna was slightly higher than that obtained in the simulation. However, the results obtained certainly form a basis for further work in this area in order to achieve a more realistic model.

CONCLUSION

A description has been given of a model for the analysis of loaded microstrip structures by a method of tuning a two-dimensional FDTD model. Initial results have shown that this method has proved suitable for analysis of the fields within microstrip antennas incorporating shorting pins, and is currently being evaluated for the electronically controlled reactively loaded probe. At the moment, good results have only been obtained for a limited frequency band around the desired resonant point. However, the model is currently being further developed to allow for a more thorough analysis.

REFERENCES

[1] Carver, K.R., & Mink, J.W., January 1981, "Microstrip Antenna Technology", *IEEE Trans. Antennas & Propagation.*, Vol. AP-29, No. 1, pp. 2-24.

[2] Richards, W.F., Davidson, S.E., & Long, S.A., May 1985, "Dual-Band Reactively Loaded Microstrip Antenna", *IEEE Trans. Antennas & Propagation.*, Vol. AP-33, No. 5, pp. 556-561.

[3] Beach, M.A., Swales, S.C., Bateman, A., Edwards, D.J., & McGeehan, J.P., May 1989, "A Diversity Combining Antenna for Land Mobile Satellite Communications", *IEEE 39th Int. Conf. on Vehicle Technology,* San Francisco, California, USA, pp. 749-756.

[4] Hilton, G.S., Hawkins, G.J., & Edwards, D.J., December 1989, "Novel Antenna Tracking Mechanism for Land Mobile Satellite Terminals", *IEE 5th Int. Conf. on Mobile Radio and Personal Communications.*, University of Warwick, Coventry, UK, 315, pp. 182-186.

[5] Lo, Y.T., Solomon, D., & Richards. W.F., March 1979, "Theory and Experiment on Microstrip Antennas", *IEEE Trans. Antennas & Propagation.*, Vol. AP-27, No. 2, pp. 137-145.

[6] Richards, W.F., Lo, Y.T., & Harrison, D.D., January 1981."An Improved Theory for Microstrip Antennas and Applications", *IEEE Trans. Antennas & Propagation.*, Vol. AP-29, No. 1, pp. 38-46.

[7] Yee, K.S., May 1966, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media", *IEEE Trans. Antennas & Propagation.*, Vol. AP-14, pp. 302-307.

[8] Taflove, A., & Brodwin, M.E., August 1975, "Numerical Solution of Steady-State Electromagnetic Scattering Problems Using the Time-Dependent Maxwell's Equations", *IEEE Trans. Microwave Theory & Tech.*, Vol. MTT-23, No. 8, pp. 623-630.

[9] Railton, C.J., & McGeehan, J.P., June 1989, "Analysis of Microstrip Discontinuities Using the Finite Difference Time Domain Technique", *Dig. IEEE MMT-S,* Long Beach, USA, pp. 1009-1012.

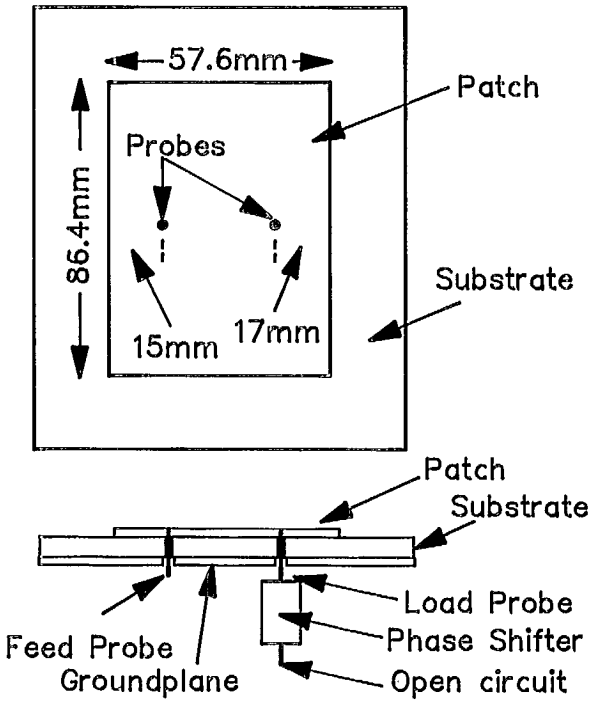


Fig. 1. Reactively Loaded Microstrip Patch Antenna.

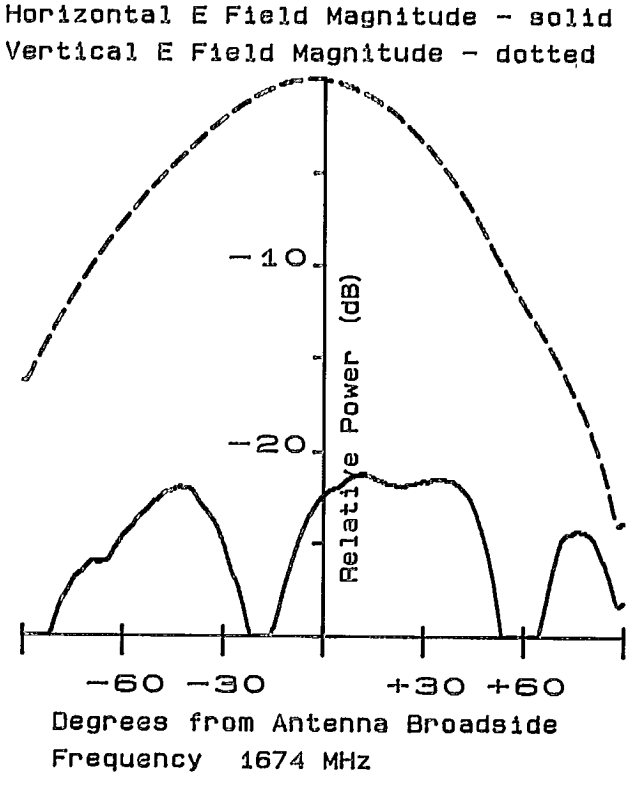


Fig. 2. Far-Field Radiation Plots (phase shifter off)

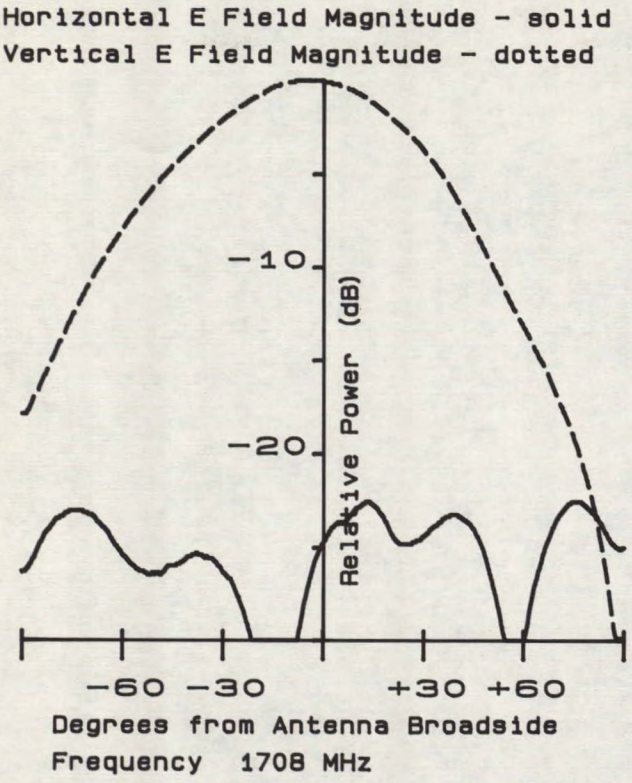
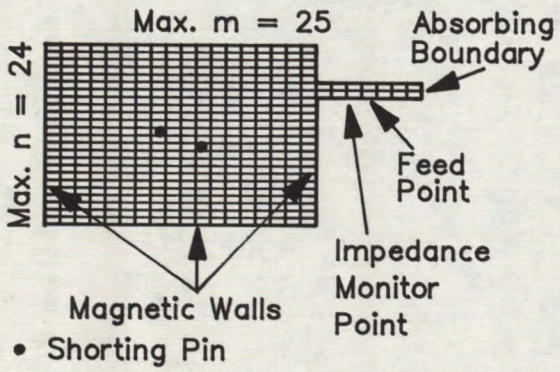


Fig. 3. Far–Field Radiation Plots (phase shifter on)



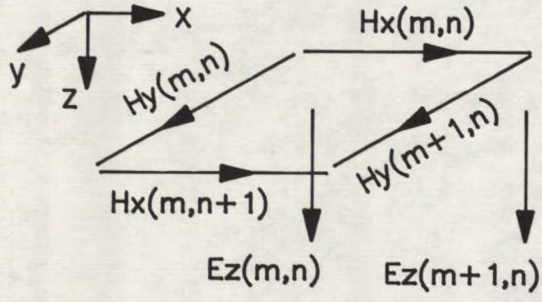


Fig. 4. Patch Simulation Grid and Field Co–ordinates.

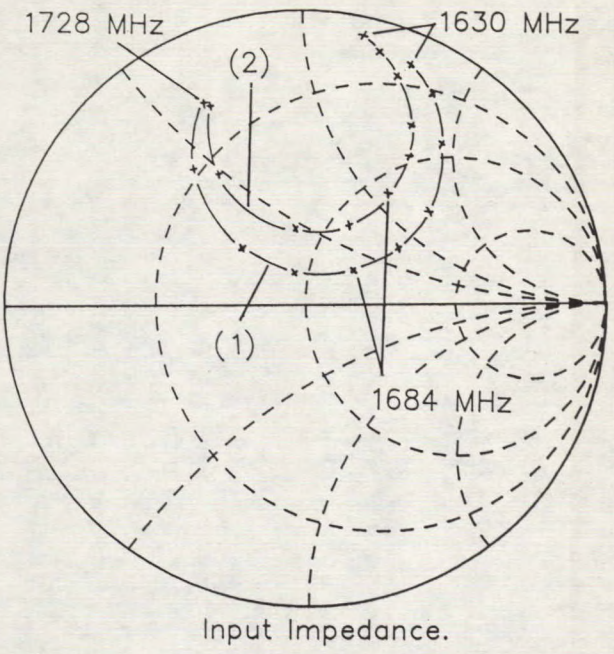


Fig. 5. Simulation Results for Coarse Grid Model with shorted regions.

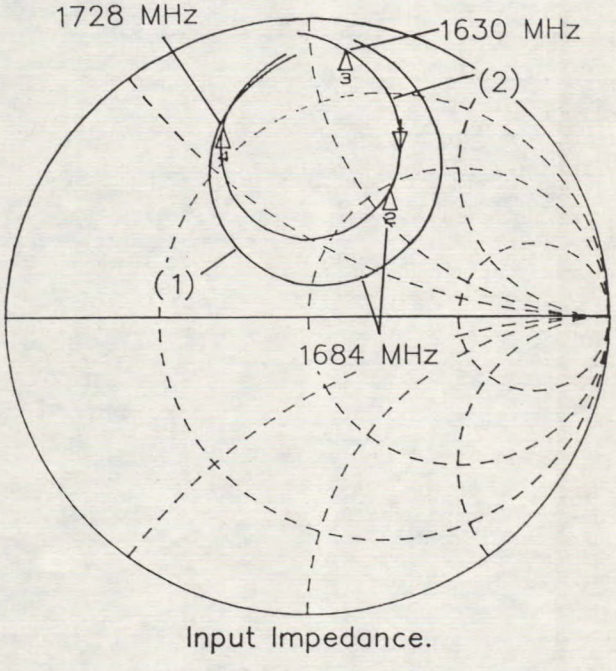


Fig. 6. Experimental Results for Microstrip Antenna with shorting pins.

Session 13
Aeronautical Applications - II

Session Chairman - *Clyde A. Miller,* Federal Aviation Administration, USA
Session Organizer - *Jack Rigley,* DOC

System Considerations, Projected Requirements and Applications for Aeronautical Mobile Satellite Communications for Air Traffic Services

K.D. McDonald, C.M. Miller
W.C. Scales, D.K. Dement
Federal Aviation Administration
800 Independence Ave., SW
Washington, DC 20591
United States
Tel: (202) 267-9840 Fax: (202) 267-5418

SYNOPSIS

This paper addresses the projected application and requirements in the near term (to 1995) and far term (to 2010) for aeronautical mobile services supporting air traffic control operations. The implications of these requirements on spectrum needs, and the resulting effects on the satellite system design and operation are discussed. The United States is working with international standards and regulatory organizations to develop the necessary aviation standards, signalling protocols, and implementation methods. In the provision of aeronautical safety services, a number of critical issues have been identified, including system reliability and availability, access time, channel restoration time, interoperability, preemption techniques and the system network interfaces. Means for accomplishing these critical services in the aeronautical mobile satellite service (AMSS), and the various activities relating to the future provision of aeronautical safety services are addressed.

The full text of this paper was not available at press time

Aeronautical Mobile Satellite Service: Air Traffic Control Applications

Dave Sim
Transport Canada AANFBL
Ottawa, Ontario
K1A 0N8 CANADA
Phone: 613-998-4870

ABSTRACT

Canada's history both in aviation and in satellite communications development spans several decades. The introduction of aeronautical mobile satellite communications will serve our requirements for airspace management in areas not served by line-of-sight radio and radar facilities. The ensuing improvements in air safety and operating efficiency are eagerly awaited by the aviation community.

AVIATION IN CANADA

Canadians reflect with pride upon our aviation heritage. We recall our triumphs in developing aircraft to meet our varied needs, including the de Havilland Beaver, Otters, Dash-7 and Dash-8, the AVRO Jetliner and the Arrow, and the Canadair Challenger corporate jet. Canadian aviators have traversed the vastness of our nation, bringing travellers, mail and cargo to the most isolated points of human settlement.

There are over twenty eight thousand civil aircraft registered in Canada today. Twenty-five percent of these are registered for commercial operation, and of these some five hundred and fifty are classed as heavy commercial, over 13,500 kilograms. Although ninety percent of our population occupies the ten percent of our land along the Canada - United States border, and is well served by major airlines, the aviation community acknowledges the immense, sparsely populated northern region, with many isolated settlements. For many Canadians, the only regular contact with the world outside their community is the 'sched flight' which brings them mail, groceries, and medical attention.

For most of us, though, domestic and international air carriers provide the travel services we need, to fly to another city for a business meeting, or to warmer climates for a relaxing holiday. Transport Canada's Aviation Industry Review[1] reports that over sixty-five million passengers passed through Canadian airports in 1988, evidence of the popularity of air travel.

For airline, commercial and private pilots alike, there are two fundamental requirements for safe and efficient flying - the ability to navigate from place to place, and the ability to communicate with ground support facilities.

WHY AIR TRAFFIC CONTROL?

Airspace is a resource which must be managed to maintain the highest level of safety to the public. It is necessary to control the movement of air traffic where conflicts may occur, and where airspace and airport facilities must be shared. Transport Canada is the federal government department responsible for regulating and managing air traffic in Canada's domestic airspace. Transport Canada provides communications, navigation aids, and surveillance systems for air traffic control. As well, the International Civil Aviation Organization has delegated to Canada and Britain, the responsibility to provide air traffic services for aircraft traversing the northern portion of the North Atlantic Ocean.

COMMUNICATIONS, NAVIGATION AND SURVEILLANCE

Pilots must be able to communicate - to express and acknowledge instructions, to indicate intentions and threatening situations, and to report their locations. They must be able to navigate - to know where they are and to establish routes between locations. They must come under surveillance, so that air traffic control agencies are aware of their positions and can safely manage the traffic within their domains.

In southern Canada, aircraft utilize double-sideband, amplitude modulated, very high frequency (VHF) radios for communications, where ground facilities are within the line-of-sight range of about three hundred kilometres. Ground-based radio aids to navigation and surveillance radar coverage coincide with VHF radio coverage. These three elements of communication, navigation and surveillance systems allow for the efficient aircraft operations in more densely populated areas.

During flights over isolated areas such as the Canadian North or the oceanic regions, communication over several thousand kilometres is possible with single-sideband, high frequency (HF) radios. Since HF radio propagation is subject to fading, the quality of these links may vary from excellent to barely intelligible. Yet, HF radio remains our standard for long distance aeronautical communications. Ground based navigation aids are not available, and radar surveillance is replaced by voice radio position reports from aircrews. It is in such areas that satellite services will offer major improvements in air traffic control.

CANADA IN SPACE

In 1959, as Canadians watched SPUTNIK cross our night sky, we committed ourselves to join the race for space. A milestone was set in 1962 with the launch of the first Canadian satellite, Alouette 1. From this early work in space technology, Canada has maintained a leading role in the development and implementation of space-based communications. The family of Anik satellites has established C and Ku-band networks for fixed land services. The launching of MSAT in 1993 will introduce domestic L-band mobile satellite services.

INITIAL AERONAUTICAL EXPERIMENTS

Clearly, our capabilities in

satellite communications should combine with our requirements for long range aeronautical communications to form an air traffic control communications network. Indeed, there is a global requirement for such a system, for air traffic is world wide. Let us summarize the efforts have been undertaken to date.

Initial satellite communications flight trials were undertaken early in the 1960s. The International Civil Aviation Organization (ICAO) established a panel of experts knows as ASTRA - the Application of Space Techniques Related to Aviation. From the work of this panel came the Aerosat program, undertaken by a consortium of Canada, the United States, and the forerunner of the European Space Agency. The Aerosat program was comprised of a space segment and a coordinated ground segment. The launch of the satellite was scheduled for 1979, but the program was terminated in 1976. The international fuel crisis of the early 1970s caused a withdrawal of funding for the project. It was agreed, however, that the investigations into the potential uses of satellites by civil aviation should continue, so an ICAO committee known as ARC - the Aviation Review Committee, was formed.

From the Aviation Review Committee studies between 1978 and 1982 came the recommendation for the use of a shared space segment for air-ground data link communications. Shared access avoids the high costs associated with a dedicated satellite, making satellite services more affordable for the aviation community. In 1983, ICAO accepted a second ARC recommendation, and formed the Special Committee on the Future Air Navigation Systems, knows as the FANS Committee. The FANS Committee has identified, studied and advised on the development of new aeronautical communications, navigation and surveillance systems for use over the next twenty five years.[2] As suggested by the FANS Committee, ICAO has approved the Aeronautical Mobile Satellite Services Panel which is developing Standards and Recommended Practices (SARPS) for aeronautical satellite communications.

SYSTEM ARCHITECTURE

A system architecture has been defined which will provide both data link communications for air traffic services and aeronautical operations control, and the capacity to accommodate high data rates, digital voice and passenger correspondence. The RF links from aircraft to satellite at 1646.5 - 1656.5 MHz and from satellite to aircraft at 1545.0 - 1555.0 MHz will provide for several categories of communications which will be assigned different priorities. These aeronautical frequencies are extensions to existing L-band mobile services which are available through INMARSAT satellites from Teleglobe Canada. INMARSAT has defined the system in their Aeronautical System Definition Manual.[3] The airline industry is applying the system definition in ARINC Characteristic 741.[4] Regional systems, such as MSAT, will service areas not covered by INMARSAT's global beams.

SATCOM APPLICATIONS

Civil aviation authorities, airlines and satellite service providers are presently

undertaking trials to establish operational requirements and procedures for satellite communications. Many applications are being investigated, including direct controller to pilot data communications, automatic aircraft position reporting, and air traffic services general message handling.

When aircraft journey beyond the range of VHF radio coverage, HF radio is used to communicate between air traffic controllers and aircrews. The controller no longer speaks directly to the crews, as messages are relayed through HF radio operators. The variable quality of HF reception and the extra link in the communications chain causes delays in aircraft position reporting and in requesting and granting clearances for manoeuvres. The uncertainties which result lead to the allocation of large volumes of airspace for each aircraft to insure that no conflict can exist. This generous use of airspace limits the ability to fly fuel efficient direct routes. Over the North Atlantic, several hundred daily flights must compete for slots in the airway track system which dictates a sixty mile lateral separation between aircraft.

Some of the information which is periodically reported by aircrews is the location of their craft in space. Since radar coverage in unavailable, this position reporting is the controller's means of managing traffic within his airspace. Because of delays, this data may not be up-to-date. The planned implementation of Automatic Dependent Surveillance will provide automatic, regular position reporting by satellite data link from an aircraft's flight computers to the controller. As aircraft navigation accuracy improves with the expansion of the Global Positioning Satellite system (GPS), the controller will have at hand information as accurate as surveillance radar position data.

CANADIAN ACTIVITIES

Transport Canada has established a program to investigate the application of aeronautical satellite services for direct communications, for position reporting and for other data link messages. The development of ICAO Standards and Recommended Practices is being supported. Canadian manufacturers are under contract to develop satcom avionics with advanced modulation schemes. Clearance information is delivered by VHF data link to eastbound trans-Atlantic aircraft from the Gander, Newfoundland, Oceanic Control Centre.

In an experiment with Air Canada, several Boeing 767 aircraft are tracked across North American airspace using automatic position reporting through existing VHF data link facilities. With the introduction of Boeing 747-400 aircraft with satcom avionics in 1991, it is planned to extend this experiment to include North Atlantic flights which will access INMARSAT and Teleglobe Canada facilities. As part of this preliminary Automatic Dependent Surveillance work, a message processor and graphical situation display are being developed for trials at the Gander Oceanic Control Centre. This system, to be completed by the Fall of 1990, will process and display position reports transmitted from aircraft

navigation computers and will enable direct message communications between air traffic controllers and aircraft flight deck displays. Position reports will automatically be compared to flight plans to check for conformity to assigned routes.

An agreement is in place between Transport Canada and the United States Department of Transportation - Federal Aviation Administration, to share information on the development of aeronautical mobile satellite services, and to cooperate in pre-operation Automatic Dependent Surveillance trials over both the Pacific and Atlantic oceans.

Development work now ongoing in several ICAO member states will lead to operational Automatic Dependent Surveillance and direct controller to pilot communications by the second half of this decade. Canadian agencies will interface with Teleglobe and the European consortium SITA for international satellite traffic over INMARSAT spacecraft, and with MSAT service providers for domestic communications.

WHAT DO WE GAIN?

The benefits of satellite communications to the aviation community will be substantial. Aviation safety, always the prime concern, will improve as air traffic controllers and pilots exchange reliable and timely information necessary for airspace management. The efficiency of airline operations will improve as delays and ambiguity are reduced, and routing and scheduling are optimized.

Aeronautical communications is about to enter the space age. The introduction of operational satellite services for air traffic control will be one of the most important advances of the past forty years of aviation.

REFERENCES

1. **Transport Canada.** 1988. Aviation Industry Review 1988.
2. **International Civil Aviation Organization.** 1988. Future Air Navigation Systems - FANS/4 Report.
3. **INMARSAT.** 1990. Aeronautical System Definition Manual, Module 1: System Definition.
4. **Aeronautical Radio Inc.** 1990. ARINC Characteristic 741: Aviation Satellite Communication System.

Spectrum Sharing between AMSS(R) and MSS

Roy E. Anderson
Anderson Associates
P.O. Box 2531
Glenville, NY 12325 U.S.A.
Phone: (518) 384-1212
FAX: (518) 384-1211

INTRODUCTION

Generic satellite systems will serve aeronautical, land and maritime users in the United States and Canada. One important service, Aeronautical Mobile Satellite (Route) Service (AMS(R)S), pertains to the safety and regularity of flight. The North American systems are designed to assure that this vital safety service is not impaired in any way as it shares the spectrum and satellites with a large number and great variety of other users in other services.

AMS(R)S has priority and preemption rights over all other communications, including aviation non-safety, in the shared band. Internationally, those rights are in the band 1545-1555/1646.5-1656.5 MHz and in the United states 1545-1559/1646.5-1660.5 MHz. The lower portions are the satellite-to-aircraft links, the upper portions the aircraft-to-satellite links.

Safety related aeronautical communications impose stringent requirements on the design and operation of aeronautical mobile satellite systems. A pilot must have reliable, near instantaneous communication with an air traffic controller. There must be smooth transitions from one sector or flight information region to another. An aircraft must be able to fly anywhere in the world and communicate with its one set of equipment through any satellite that provides aeronautical communications.

The satellite service providers and the aviation community are working together to assure that the stringent requirements will be met. Minimum Operational Performance Standards are being developed by the Radio Technical Commission for Aeronautics. Standards and Recommended Practices are being developed by the Aeronautical Mobile Satellite Service Panel of the International Civil Aviation Organization. Avionics standards, "form, fit, and function", are specified by the Airlines Electronic Engineering Committee. The engineering bases for radio frequency regulations are developed by the International Radio Consultative Committee (CCIR).

All of the organizations are concentrating their efforts on AMS(R)S which consists of Air Traffic Services (ATS) and Aeronautical Operational Control (AOC). ATS, which includes air traffic control, is the responsibility of civil aviation authorities. AOC is the responsibility of the airlines and other users of the service. In addition to the AMS(R)S

safety related communications, there are non-safety Aeronautical Administrative Communications (AAC), primarily airline business operations, and Aeronautical Public Correspondence.

Aeronautical mobile satellite communications will also be provided through the satellites of global systems, such as INMARSAT, which serve oceanic and continental areas that do not have regional systems. Regional systems, such as those of the American Mobile Satellite Corporation (AMSC) in the United States and of Telesat Mobile Inc. (TMI) in Canada, will provide the satellites for high traffic density areas.

In the United States AMS(R)S will be served by a network separate from all other uses of the AMSC satellites. The architecture of the system will interconnect the AMSC Network Control System and the AMS(R)S network to assure real time access to all the spacecraft power and bandwidth that will be needed. Coordination with other satellite systems will facilitate the air traffic handover between flight information regions and enhance efficiency in the use of the spectrum. The bandwidth and linearity of the satellite transponders will assure that every kind of communication signal and protocol used by aviation will be accommodated.

AMS(R)S REQUIREMENTS

The first use of satellites for AMS(R)S will be over the oceans. Automatic Dependent Surveillance (ADS) will provide automatic, frequent position reports to air traffic controllers. They can then track aircraft, thus supporting separation reduction and allowing more aircraft to follow the most fuel efficient routes. The position fixes are derived from on-board navigation systems, such as the aircraft inertial navigation system, Loran-C, or the Global Positioning System. The transmissions of the reports to the ground will be short data packets sent at timed intervals.

ICAO has specified the use of digital techniques for all voice and data communications. The minimum "core" capability requirement is duplex data communications between aircraft and ATS earth stations at 600 bits per second with aircraft equipped with 0 dBi omni-directional antennas. A signal power density at the aircraft of -165 dBW/Hz/m^2 is specified. Assuming global beam satellites, the international airlines have considered that digital communications at bit rates high enough to provide toll quality voice would require 12 dBi aircraft antennas with steerable beams.

The Airlines Electronic Engineering Committee has prepared ARINC Characteristic 741 describing the avionics equipment to be used on transport aircraft. Several manufacturers are building equipment according to this specification, and some aircraft that fly oceanic routes are being equipped. Aeronautical administrative and air passenger telephone communications are being introduced using the INMARSAT satellites. Oceanic air traffic services will be tested, but will not be operational until they have been qualified by the traffic control authorities.

There is no current plan to introduce AMS(R)S service in U.S. domestic airspace. However, studies by the Federal Aviation Administration are postulating the use of Automatic Dependent Surveillance as an alternative to the present en route radar surveillance for the time period circa 2010. A peak instantaneous aircraft count in North America may then be between 17,000 and 34,000. If those numbers of aircraft participate over North America and each aircraft transmits its location every ten seconds, the spectrum needed to handle the ADS and associated ATS and AOC communications

would be between 8 and 13 MHz assuming that there is a frequency reuse factor of 3. The traffic load would have a large variation during the course of the day as well as seasonally. It would also vary with weather conditions.

Service areas of the INMARSAT, USSR and possibly other satellites overlap the service area of the AMSC satellites. International coordination according to International Telecommunication Union rules will be necessary to avoid interference between the systems. That coordination will be the responsibility of AMSC.

As aircraft move from the service area of one satellite to another, or between beams of a single satellite, and as they move from the region of one traffic control authority to another, a dynamic handover coordination procedure must be followed. That coordination is primarily the responsibility of the AMS(R)S network operator.

AMS(R)S OPERATION THROUGH AMSC SATELLITES

Each earth station in the AMS(R)S network will be interconnected with the Aeronautical Telecommunications Network that will use Open Systems Interconnection protocols. AMS(R)S authorities, through their network operator, will specify the portion of AMSC satellite power and bandwidth needed for AMS(R)S. The amount may vary diurnally and seasonally due to characteristic patterns of aviation traffic. A factor influencing the amount is the low blocking rate required for the service. AMSC will insure that no other traffic is allowed in the spectrum assigned to AMS(R)S.

The AMS(R)S authorities will operate an Aviation Priority Demand Assigned Multiple Access (AV PDAMA) facility that controls the communication traffic flow between the aircraft and the ATS and AOC authorities. The Network Control Center of AMSC will contain a Priority Demand Assigned Multiple Access (PDAMA) facility that has a robust interconnection with the AV PDAMA.

The ATS and AOC authorities will monitor the blocking rates of AMS(R)S traffic through the AMSC satellites. When they exceed acceptable limits AMSC will be notified and required to immediately preempt additional bandwidth and power for reassignment to AMS(R)S.

The AMSC system design complies with the preemptive access requirement. Every Mobile Earth Terminal that uses the AMSC satellites will be type approved and individually commissioned. All of them will be under positive control of the Network Control System (NCS). Positive control means that the NCS can issue commands at any time to one or groups of mobile terminals that must be acted upon immediately.

The aviation community, through the organizations listed in the Introduction, have defined in great detail the performance requirements, the operating procedures, and the equipment specifications for AMS(R)S. It is an ongoing process, and an important one because it is the introduction of new technology and capability into one of the world's major industries. The plans must anticipate developments over the next several decades in aviation and in mobile satellite technology.

Whenever it can play an appropriate role, AMSC participates actively in all of the aviation fora involved in the process. Members of the AMSC ownership consortium have participated in experiments and in the planning process since aeronautical communications by satellite were first suggested. Through that participation AMSC has contri-

buted to the process, developed its system to be fully compliant with AMS(R)S requirements, and laid the groundwork for future system growth in capacity and performance that will keep pace with aviation's advances.

Recent Technical Advances in General Purpose Mobile Satcom Aviation Terminals

John T. Sydor, Project Leader
Mobile Terminal Development
Department of Communications
Communications Research Centre
3701 Carling Avenue
P.O. Box 11490, Station H
Ottawa, Ontario
K2H 8S2
Phone: (613) 998-2388
Fax: (613) 990-6339

ABSTRACT

A second general aviation ACSSB aeronautical terminal has been developed for use with the Ontario Air Ambulance Service (OAAS). This terminal is designed to have automatic call set up and take down and to interface with the PSTN through a Ground Earth Station hub controller. The terminal has integrated RF and microprocessor hardware which allows such functions as beam steering and automatic frequency control to be software controlled. The terminal uses a conformal patch array system to provide almost full azimuthal coverage. Antenna beam steering is executed without relying on aircraft supplied orientation information.

GENERAL DESCRIPTION OF THE TERMINAL

The first general purpose aeronautical satellite terminal developed by CRC for the Government of Ontario Air Ambulance Service (OAAS#1) proved the feasibity of aeronautical satcom for small aircraft by using L-Band mobile satellite technology (Ref.1). However, the technology developed for this demonstration has limited operational characteristics. The setting up and taking down of calls between the aircraft and the Medical Communications Centre in Toronto requires considerable human intervention at the Ground Earth Station. Similarly, the requirement for ultra-stable frequency control necessitated by the narrow band modulation technology and the demands of the satellite provider (INMARSAT), forced the use of an ovenized oscillator with a long stabilization time which is undesirable for a terminal designed for emergency purposes. Lastly, the use of fixed window antennas on each side of the aircraft limited the azimuth angle of satellite coverage, EIRP, and the effective G/T of the terminal. Unless the pilot was willing to steer the aircraft in a usually less than favoured direction, achieving reliable satellite communications proved to be difficult.

The purpose in developing a second air ambulance satcom terminal was to address these operational limitations. Our solution was to design a call set up and take down system which

would be automatic and allow channel access to be arranged by a Carrier Sense Multiple Access protocol. The protocol would work from a hub controller located at the Teleglobe Canada Ground Earth Station in Weir, Quebec. In order to minimize the long frequency stabilization times experienced with the first terminal, the second terminal would use a satellite pilot derived frequency reference that would reduce waiting time from 15 minutes to less than one. Finally, the second terminal was to be designed to provide the aircraft with almost full azimuth coverage by using a configuration of low cost, low drag conformal patch antennas that would surround the aircraft with a series of steerable high gain beams.

The terminal (OASS#2) was built in the fall of 1989 with software development and testing starting in the spring of 1990.

ANTENNA AND RF SYSTEM

A block configuration of the terminal is shown in Figure 1. The antenna system is comprised of three microstrip patch array antennas each equipped with a diplexor, low noise amplifier, and for those antennas having multiple beams, a coaxial switch. This configuration of front end RF was chosen because it maximized the G/T and EIRP. A small phased array was considered but the losses of such a system would have resulted in poorer performance.The antenna system implemented generates 9 beams that provide the aircraft with about 320 degrees of useable coverage (Figure 2).Over this azimuth angle the G/T of the terminal varies from -8 to -18 dBK. Similarly, the EIRP goes from 48 to 58 dBm, resulting in a received C/No of 40 to 50 dBHz at the ground earth station.

The multiple beam patch antennas are mounted on the port and starboard aft fuselage of a Beechcraft King Air 200 airplane. A third antenna is mounted in the nose of the aircraft above the radar. The forward antenna almost completely fills the forward keyhole left vacant by the patch arrays. All the antennas are made of air dielectric microstrip patch elements and are 1.25 inches thick. The port and starboard antennas are conformally mounted on the skin and protected by a radome that measures approximately 18 by 24 inches having a height of about 1.5 inches, resulting in negligible air drag. The port and starboard antennas each produce 4 beams having gains that range from 13 to 16 dBic. The forward antenna located in the nose of the aircraft has a gain of 12 dBic. The gains are measured from 1550 to 1650 MHz. The beams are squinted at elevation angles of 20 degrees, corresponding to the satellite look angle in the Northern and Western Ontario service area of the ambulance.

The diplexor/low noise amplifier switch boxes are mounted within 10 inches of the antennas. This assembly is light weight (approx. 2 lbs.) and has a receive path insertion loss of 1.0 dB and a transmit path loss of 1.3 dB. Each receive RF path is a concatenation of antenna, 10 inches of 0.141 semirigid coaxial cable, 4 to 1 low loss coaxial switch (the nose antenna does not need a switch), circulator, 6 pole cavity filter, an 0.8 dB noise figure low noise amplifier, followed by the RF cable that can be anywhere from 5-30 feet long directing the signal back to the

main terminal where a final 3 to 1 coaxial switch is encountered. The transmit path begins with a 6 pole transmit filter located next to the power amplifier, a 1 to 3 coaxial switch, RF cable, circulator, 4 to 1 switch and finally the feeder coaxial cable and antenna.

The desired beam was chosen by first switching the 3 to 1 coaxial switch to select either the port,starboard, on forward antennas. If the port or starboard antenna is selected the 4 to 1 switch is used to select the individual beam. The total switching time is 25 milliseconds. The loss in a coaxial switch is less than 0.1 dB over the 1550 to 1650 MHz band.

The up and down conversion system and L-Band voltage controlled local oscillator are constructed of low complexity microstrip circuitry. The design considerations here were for low cost and robustness, to satisfy the eventual land mobile applications for these circuits. The L-Band local oscillator has an output of 13 dBm at 1472.750 Mhz, a tuning range of approximately 10 KHz, and a phase noise specification of -95 dBC/Hz at 1 KHz. The RF power amplifier is a commercially available (Canadian Astronautics Limited) class AB linear L-Band device capable of delivering a peak output power of 49 dBm. To accomodate the linearity requirements of the Amplitude Companded SSB modulation, the amplifier is set at an average output of 43 dBm.

SIGNALLING AND ACCESS PROTOCOLS

The modems and protocol controller of the terminal are off-the-shelf ACSSB channel units procured from Skywave Electronics Limited. The units were modified to accept DTMF signalling and redesigned with a robust access protocol that would ensure operation at low C/No. The DTMF signalling was chosen because of its implementation simplicity and robustness. The probability of successful signalling with the DTMF tones is approximately 78.7 % at 33 dB Hz and 99.8 % at 38 dB Hz. This coupled with the observation that the limit of intelligible ACSSB is about 38 dB Hz, has caused us to set this as the lower limit of terminal operation.

The terminal is designed to work in a carrier sensed multiple access system controlled by a 386 microprocessor family computer. The computer is linked with channel units that service air ambulance calls. The system mediates PSTN signalling and controls call set up and take down between the Medical Communications Center in Toronto and the aeronautical terminal. The computer also collects usage information for billing purposes and controls forward link EIRP from the satellite. It is located at the Teleglobe Canada Weir Ground Earth Station and has a single INMARSAT 25 KHz - 21 dBW EIRP L-Band channel allocated to it. The channel is subdivided into three 8.33 KHz channels that are shared on a demand assigned basis by 2 Air Ambulance terminals (OAAS #1 and #2) and 18 briefcase portable ACSSB/DMSK terminals used for the Canadian MSAT trials program.

SATELLITE TRACKING FACTORS

There were several factors motivating the design of the beam steering hardware and software for the terminal. One factor was the need to maintain the the best C/No for the received signal. The operating margins for the forward and return links are minimal and even these change on a daily basis as the INMARSAT satellite becomes loaded with maritime traffic. The second factor driving the design was the limited lifetime of the coaxial switches for the antennas. These switches are rated for one million switching operations which is a number that can be quickly reached, especially if the terminal dithers excesively from beam to beam.

PILOT TRACKING HARDWARE

The terminal tracks the INMARSAT Standard A System channel 110 BPSK pilot. This pilot is common to the worldwide maritime satcom service and will be present as long as the Standard A service remains viable. The pilot has a low forward EIRP (only 13 dBW) and, as a consequence, is difficult to monitor. The BPSK signal is detected by a quadrature demodulator which produces signal level and frequency information. The demodulator contains a phase lock loop that has a tracking range of 300 Hz and a loop bandwidth of 50 Hz. The output of these circuits is fed into a 8 Bit A/D and a microcontroller where signal analysis is performed. The BPSK pilot signal can be detected and measured to a C/No of 24 dB Hz, though the limits for terminal operation are set at 27 dB Hz. It should be noted that the BPSK pilot has a mean power 8 dB below the average power of the ACSSB communications signal. As such, since the lower limit of ACSSB operation is set at 38 dB Hz, the BPSK pilot must be useable down to at least 30 dB Hz.

The sampling interval is 200 milliseconds long, over which time measurements were continually taken and averaged by the microcontroller. This interval was chosen to accomodate a 5 Hz ripple in the signal power of the BPSK. The ripple is due to the repitition of an unchanging identity work on the INMARSAT Standard A.

BEAM STEERING

Unlike the commercial aviation L-Band satellite terminals that are being currently developed, the OAAS#2 terminal does not switch antennas using orientation information provided by the aircraft. Steering information is derived solely from the quality of the received pilot signal and switching decisions are made on the basis of the quality of the received signal dropping below thresholds which are adaptively determined by the terminals' microprocessor.

Adaptability of switching thresholds is a key requirement with this terminal because of antenna beam scalloping and gain variation. The 9 antenna beams have gains which vary from 12 to 16 dBic. These gains and the symmetry of the patterns are also altered by the metallic structure of the aircraft. Additional complications arise from the fact that each beam represents a unique RF signal path, which can vary independently in both noise content and gain over time and temperature. To counter these

variations the terminal bases its decision processes solely on the basis of received C/No. Each beam, from the onset of terminal activation, is calibrated for noise density and this calibration is continually updated and averaged every 10-16 seconds for the beam that is currently used. Similarly, the received pilot signal is sampled and averaged over six 200 millisecond intervals taken within a 2 second period. These averages are used to create a mean C/No value for the current beam, representing the quality of the signal over the previous 2 seconds. Experiments with the stationary OAAS#2 terminal have shown this sampling system capable of measuring the mean C/No to an accuracy of +/- 0.5 dB over a 20 dB dynamic range from 28 to 48 dB Hz. The nominal pilot signal level with a 14 dBic beam would be 39 dB/Hz.

TRACKING ALGORITHM

Ideally, the tracking algorithm should switch whenever the pilot signal is higher in an adjacent rather than current beam. Each beam (labelled n, where n=1....9) has two thresholds T(n+1) and T(n-1) with adjacent beams or beam pattern keyholes. These thresholds are determined by the terminal as it operates, and are updated with each successful switching that the terminal makes. When a threshold is reached the terminal determines the C/No in the current beam, then switches to the adjacent beam and samples the signal in it for an interval of 400 milliseconds. If this latter sample promises to exceed in C/No that of the original beam, sampling for a full 2 second interval is maintained. Providing the new sample is greater than one in the original beam, a threshold is created by calculating the difference between the two samples and adding half of this to the lower sample. The thresholds are further reduced by about 1-1.5 dB to increase the probability of switching success by increasing the probable difference between the current and adjacent beam samples (see Figure 3).

One difficulty that exists with this algorithm is that the terminal has only a 50% chance of moving to the correct beam whenever the signal level is between thresholds. That is, given that the level of the signal is S in beam (n), we have two possibilities:
$T(n+1) < S < T(n-1)$ or
$T(n-1) < S < T(n+1)$ (Figure 3).

As an example of a wrong more, the signal in the current beam may be entering the beam on its side adjacent to the n+1 beam, however this signal may be below T(n-1). Unless some preventative action is taken, the terminal will switch after each sampling cycle to the n-1 beam, where the signal is virtually non-existent. Switching will then oscillate with the periodicity of the sampling /switching cycle until the signal either goes below the T(n+1) threshold or above T(n-1). This situation will occur half the time. The other half of the time the signal will occur on the other side of the beam where the move will result in a greater signal.

A closer examination of this problem reveals that nothing can be done to prevent such oscillations, other than extracting additional information from the received signal which would indicate the direction of

aircraft turn. This level of processing was not attempted with the current terminal. Instead we opted for a solution which would inhibit the terminal from making incorrect moves by dynamically varying the interval of time between wrong switching decisions. Thus after one wrong move the terminal will wait for two sampling cycles to elapse before undertaking a second wrong move. After the second wrong move the terminal will wait 10 sampling cycles after which wrong moves will be allowed only every 14 cycles (28 seconds). If the terminal moves above the threshold forcing the incorrect moves then it will quickly remove the inhibition. If the terminal successfully changes beams, then all the inhibitions for the previous beam will be set to zero.

The lost signal threshold for the terminal is set at 38 dB Hz (30 dB Hz for the BPSK pilot). No matter what level of switching inhibition the terminal is in, if the lost signal threshold is reached the terminal goes into a signal search mode beginning its scan with the adjacent beams. A complete scan of the 9 beams can be achieved in about 1.8 seconds. At times, these may be unsuccessful in acquiring the lost signal, especially if the aircraft is flying with the signal entering a keyhole. During such instances, the scan times are slowed to one every thirty seconds.

FREQUENCY CONTROL

In the design of the frequency control system for the OAAS#2 terminal it was decided that ovenized oscillators were not to be used. In consideration of future low cost terrestrial mobile terminal applications, a frequency control system using a satellite based pilot reference was preferred.Pilot reference systems using either frequency or phase lock loop techniques combine the advantages of a single, ultra-stable long term system reference with the low phase noise characteristics achievable using low cost crystal or SAW oscillators resident within the terminal.

To lower the complexity of a pilot based frequency acquistion and control system, we combined simple microstrip multiplier circuits with a microcontroller, D/A, and voltage controlled crystal oscillator to produce a frequency locked reference that is able to lock and track the Standard A BPSK pilot to an accuracy of +/- 50 Hz at a C/No of 27 dB Hz.

FREQUENCY ACQUISTION

The INMARSAT Standard A BPSK pilot at 1537.750 MHz has an absolute uncertainty of +/- 200 Hz. The pilot is detected using a 2 stage downconverter feeding the BPSK demodulator and microprocessor. The microprocessor monitors the demodulated pilot centered at a baseband frequency of 9600 Hz and provides frequency correction using a voltage controlled crystal oscillator. The 9600 Hz signal is actually a square wave whose periodicity is easily determined by the microcontroller by a simple counting routine. Offset of the 9600 Hz represents the difference between the satellite pilot and the terminals main local oscillator.

An L-Band multiplier takes the output of the oscillator and produces the 1467.750 MHz local

oscillator of the terminal. Using a voltage controlled crystal oscillator was advantageous because it eliminates the need for any phase-lock loop circuits and in conjuction with the microstrip circuit multiplier, provides a highly stable, low cost L-Band reference. The circuits provide very quick frequency acquisition. At a pilot level of 27 dB Hz the INMARSAT pilot could be detected and used as a terminal reference within 30-45 seconds 90 % of the time.

The total absolute error of the terminals local oscillator after lock-up is in the order of 2.4 parts per 10(7) and is attributable to three factors, the absolute error of the INMARSAT pilot (+/- 200 Hz), the resolution of the 8 bit microprocessor system (+/- 50 Hz), and the uncertainty of the second LO in the downconversion chain (+/- 70 Hz). Secondary errors due to the drift in the temperature compensated voltage controlled oscillator were on the order of 30 Hz prior to stabiliztion of the terminal. This latter error is corrected out by periodic frequency recalibrations of the terminal flight.

The terminal acquires frequency lock to the INMARSAT pilot during those moments when the aircraft powered up. The aircraft at such moments is relatively stationary allowing the terminal to acquire a pilot free of Doppler offset. As the aircraft flies there are occasions when its orientation toward the satellite is such that Doppler offset on the received pilot is minimal. During these occasions the port or starboard antenna beams 3 or 7 are in use (see Figure 2) and the microprocessor takes advantage of this situation to perform minor frequency corrections. This procedure in reality worsens the absolute frequency error of the terminal by the acquired Doppler error (approx 90 Hz) , but it reduces the relative error between the terminal and the satellite pilot and as well, corrects any minor drift problems associated with the temperature compensated voltage controlled L-Band local oscillator.

Experiments with frequency resolution were tried with the frequency monitoring system of the terminal. The counter used for frequency measurements in the microcontroller has a 16 bit resolution allowing the pilot to be measured to a 0.5 Hz accuracy. Such accuracy could allow the terminal to quickly determine whether the aircraft is turning either toward or away from the satellite simply by monitoring the change inthe Doppler frequency of the pilot. It is proposed that incorporating this information with the steering algorithms could provide significant improvement in antenna steering performance especially in preventing moves into incorrect beams when as discussed earlier. Developing terminals and algorithms to do this processing is a future endeavour for us.

CONCLUSIONS

The above brief description has not covered many of the background functions of the terminal such as its call set up and take down processing, RF power control, and a temperature control system which ensures terminal operation over a wide range of operational and environmental conditions. Many of

these functions are controlled by the same microcontroller that controls sampling, beam steering, and frequency correction.

The OAAS#2 terminal demonstrates that elements of operation such as frequency control and beam steering can be carried out using relatively simple hardware. The key to the success being control and monitoring by a microprocessor and its associated software.

Future L-Band terminals will have an even greater integration of beam steering and frequency control hardware and microprocessor software. Ultimately the goal will be to provide a simple,low cost RF system that use baseband signals at a frequency where they can be digitally signal processed without any intermediary hardware such as hardware demodulators or voltage controlled oscillators.

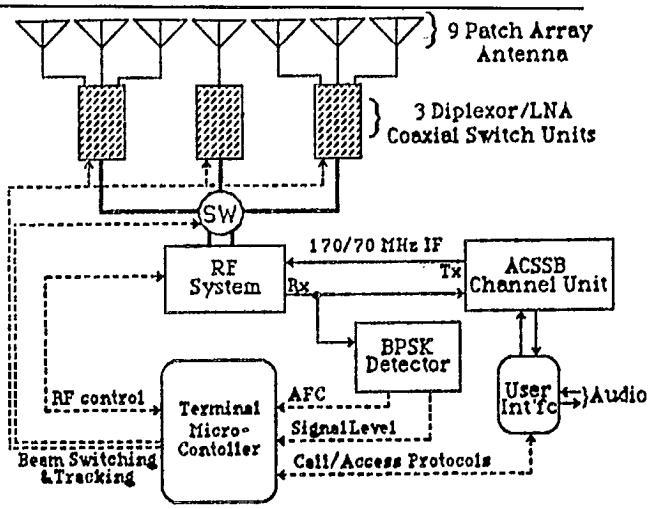


Figure 1

OAAS #2 Block Level Organization

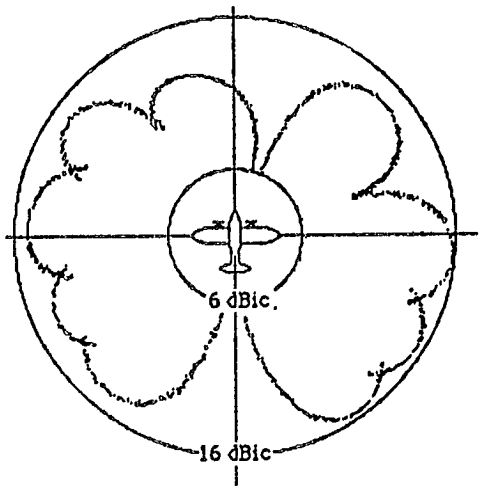


Figure 2
9 Beam Antenna Pattern of OAAS#2

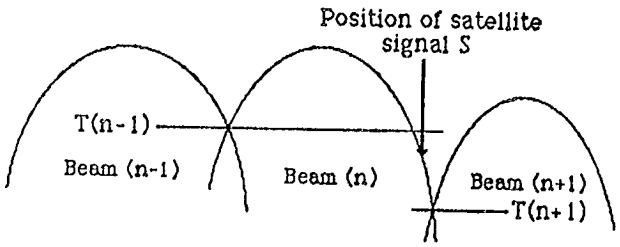


Figure 3
T(n-1)< S <T(n+1) Condition for Switching Oscillation into Beam (n-1)

REFERENCES

1. Butterworth, J. S. 1988. Satellite Communications Experiment for the Ontario Air Ambulance Service. Proceedings of the Mobile Satellite Conference May 3-5, 1988. JPL Publication 88-9

Future Developments in Aeronautical Satellite Communications

Peter Wood
Inmarsat
40 Melton Street
London, NW1 2EQ
United Kingdom
Phone: +44 71 387 9089
FAX: +44 71 383 0849

ABSTRACT

Very shortly aeronautical satellite communications will be introduced on a world-wide basis. By the end of the year voice communications (both to the cabin and the cockpit) and packet data communications will be available to both airlines and executive aircraft. During the decade following the introduction of the system there will be many enhancements and developments which will increase the range of applications, expand the potential number of users, and reduce costs. This paper presents a number of ways in which the system is expected to evolve over this period. Among the issues which are covered are the impact of spot beam satellites, spectrum and power conservation techniques, and the expanding range of user services.

INTRODUCTION

Aeronautical satellite communications have been the subject of tests and studies for more than 25 years. The earliest test was in 1964, when communications were successfully established with an PanAm aircraft in flight, using the SYNCOM III Satellite. The most recent occurred earlier this year when an aircraft operated by the U.S. National Science Foundation was able to maintain communications to its home base to within 50 miles of the South Pole. Studies have included the Oceanic Area System Improvement Study[1], the RTCA SC-155 User Requirements for Future Communications, Navigation, and Surveillance Systems[2], and the report of the ICAO Special Committee on Future Air Navigation Systems[3].

Approximately five years ago Inmarsat began to develop the concepts for a global aeronautical satellite service, and contracted for the transponders in its Inmarsat-II satellites to include 6 MhZ allocated to AMSS(R) use (3 MHz in each direction). Since that time the system has evolved to the point where it will enter into service in the very near future, and will provide both voice and data communications to aircraft operating world-wide, with the exception of a limited area near the poles.

The Inmarsat approach has been to use a single-channel per carrier for voice communications to and from the aircraft, a combination of slotted Aloha and TDMA for data communications from the aircraft, and TDM for data communications to the aircraft. During the course of development an alternative to the Inmarsat approach based entirely on TDMA techniques was proposed by the AvSat Corporation, but although the concepts are still retained in some

specifications[4], development has been terminated, and it is unlikely that a system based entirely on TDMA will be introduced in the forseeable future.

OUTLINE OF THE SYSTEM

The INMARSAT Aeronautical Satellite System is based upon two types of antenna, a nominally omnidirectional low gain antenna, and a steerable high gain antenna (nominally 12 dB), together with a range of channel transmission characteristics (Table 1) designed to meet a wide range of applications.

Table 1. Channel Transmission Characteristics

Channel Rate (bit/s)	Channel Spacing (kHz)	Modulation
21000	17.5	A-QPSK[a]
10500	10.0/7.5	A-QPSK
6000	5.0	A-QPSK
5250	5.0	A-QPSK
4800	5.0	A-QPSK
2400	5.0	A-BPSK[b]
1200	5.0/2.5[c]	A-BPSK
600	5.0/2.5	A-BPSK

a A-QPSK is Aviation QPSK, a particular form of O-QPSK[5, 6]

b A-BPSK (Aviation Binary Phase Shift Keying) is a from of differentially encoded BPSK in which alternate modulation signals are transmitted in notional In-Phase and Quadrature channels[7]

c 5.0 applies to P-channel, 2.5 to R- and T-channels

Four types of channels are supported:

- P-Channel: Packet mode time division multiplex (TDM) channel, used in the forward direction (ground-to-air) to carry signalling and user data
- R-Channel: Random access (slotted Aloha) channel, used in the return direction (aircraft-to-ground) to carry some signalling and user data (short messages only)
- T-Channel: Reservation Time Division Multiple Access (TDMA) channel, used in the return direction only.
- C-Channel: Circuit Mode single channel per carrier (SCPC) used in both the forward and return directions. The use of the channel is controlled by assignment and release signalling at the start and end of each call The C-channel includes sub-band data fields for signalling purposes.

In the initial system, installations using the omnidirectional low gain antenna will only support low data rate communications over the P-, R- and T-channels. High gain antennas will support both low and high rate data, and voice communications over the C-channel.

The majority of channel types use Forward Error Correction (FEC) coding, consisting of a convolutional encoder of constraint length k=7 and an 8-level soft decision Viterbi decoder. The FEC coding rate is 1/2. Because of the multi-path fading characteristics of the aeronautical transmission path interleaving is applied to all channels using FEC coding in order to preserve the FEC coding gain.

Signalling and user data messages on the P, T, and sub-band C-channels are formatted into standard length signal units of 96 bits. Signalling and user data messages on the R-channel are formatted into extended length

signal units of 152 bits. These allow for the most common transactions to be carried out within only one signal unit with a minimum of spare unused capacity. More complex messages (including use data) can be carried by sequence of several signal units, up to a maximum of 64.

Voice signals are digitally encoded using an APC algorithm at a bit rate of 9.6 kbits/s. The encoded voice data is combined with the sub-band signalling channel for transmission over a C-channel. The system is designed to allow operation through the Public Switched Telephone Network, although for operational reasons initially it will only by possible to initiate calls from the aircraft. This restriction does not apply to communications to and from the cockpit or with executive aircraft.

The system design is based upon the use of a linear (class-A) High Power Amplifier in the aircraft. This will support operation with several channels simultaneously. The number of channels that the amplifier can support is dependent upon the satellite through which an aircraft is operating, and the location of the aircraft with respect to the satellite. It is also possible to use a class-C amplifier for single channel installations.

SATELLITES

During the first few months of service aeronautical communications will use the first generation of Inmarsat satellites (Marecs and Intelsat-MCS), and communications will be in the maritime band. The main operational limitation with these satellites is that they will only support two channels simultaneously (reduced to one channel at satellite elevation angles less than 20°).

The first major operational change will result from the relocation of one of the Marecs satellites to 55° West, resulting in the creation of four ocean regions (Atlantic-E, Atlantic-W, Pacific and Indian). The effect of this relocation will be to close the gap that currently exists over part of North America and the eastern Pacific Ocean.

This will be shortly followed by the launch of the first of four Inmarsat-II satellites. Each of these satellites will provide 3 MHz capacity in the aeronautical band and, more importantly, also have greater sensitivity, which will result in doubling the number of channels that can be supported by single HPA.

The third stage in the evolution of satellite support for aeronautical communications will be the launch of spot-beam satellites. Currently AMSC in the United States, M-Sat in Canada, and Inmarsat are planning to launch spot beam satellites. The first two will provide regional coverage; Inmarsat plans to provide global coverage through its Inmarsat-III spacecraft. The first Inmarsat-III spacecraft is expected to enter service late 1994 or early 1995.

Two important operational enhancements will result from the introduction of spot beam satellites. The first is an increase in the number of channels that can be supported by a single aircraft installation that complies with industry standards. The second is the possibility of supporting voice services through a low-gain omnidirectional antenna. Although Inmarsat does not see this as being the primary method of operation (if for no other reason than the increased interference to global beam satellites that would result), nevertheless this would be a method of maintaining a minimum level of voice service for emergency use in the event that the high gain steerable antenna failed for some reason.

SYSTEM DEVELOPMENTS

System Management

Initially each Ground Earth Station (GES) that supports the Inmarsat system will operate on a stand-alone basis, using a dedicated set of frequencies. Each GES will be required to support a high-power P-channel to communicate with aircraft that are only fitted with low-gain antennas.

As usage of the system develops, Inmarsat will be introducing Network Coordination Stations (NCSs) which will allocate frequencies on a dynamic basis according to demand. It would also be possible to use space segment resources more efficiently by providing a common high-power P-channel at NCSs only.

As other satellite service providers begin to offer service, the NCSs will also be responsible for inter-system coordination.

Voice Coding

At the time the concepts for Inmarsat's aeronautical system were maturing, the lowest bit rate at which it appeared an acceptable quality for voice communications could be achieved was in the order of 9.6 kbits/s. Extensive tests were carried out on behalf of the Airlines Electronic Engineering Committee to determine which of the algorithms available had the highest performance in an aeronautical environment, and an APC based algorithm developed by British Telecom Research Laboratories (BTRL) was selected as the industry standard.[8]

Subsequently developments in voice codec technology have led to the possibility that acceptable quality could be achieved at bit rates as low as 4.8 kbit/s. Formal comparative tests between 9.6 kbit/sec and 4.8 kbit/s have still to be performed. If the performance of these lower data rate codecs lives up to their promise then it is probable that they will be eventually selected as a future standard. No major modifications to the basic system design would be required, and the 9.6 kbit/sec codec would continue to be supported while it was in general use. The lower bit rates would, of course, lead to a more efficient use of the available spectrum; a key issue as the uses of aeronautical satellite communications expands to fill the frequencies allotted by the ITU for that use.

Circuit Mode Operation

Circuit-mode data service is an optional enhancement to Inmarsat's aeronautical system. This service provides a basis circuit mode data channel with a call set-up procedure similar to that used for voice calls. The nature of the end-to-end service between the aircraft and the end user depends upon the type of interconnection with the terrestrial network. The use of voice-band analog modems provides an analog-interconnect data service. Digital interconnection is possible where a digital path is available through the terrestrial network (circuit switched digital data network, dedicated digital network, or ISDN).

Each circuit-mode data call utilizes a pair of C-channels for the duration of the call. The digital bit stream of the C-channels, without the use of voice band modems over the satellite channel, is directly employed to transport the data traffic. The voice codecs normally associated with the C-channels are not used during the data transfer state.

Rate adaptation must be performed to match the user data bit rate, taking into account clock rate variations, to the available satellite channel capacity. The bit rate stability of the received data from a PSTN modem is, in general, worse than the required clock rate stability of the satellite channel. It is therefore necessary to

implement a plesiochronous interconnection arrangement in the forward direction. A plesiochronous buffer is not required in the return direction provided the PSTN modem is synchronized to the incoming data stream from the C-channel.

NEW SERVICES

When the concepts for the system were originally developed it was considered that there were two basic types of service required; packet mode data for air traffic service and airline use, and voice for public correspondence and, to a more limited extent, for non-routine air traffic service and airline messages. As more users became aware of the potential of aeronautical satellite communications a wider range of services appeared to justify support.

Examples of the services which will be offered in the future are:

Facsimile

One of the earliest to be introduced will be facsimile. An interface unit to permit operation of standard Group-III facsimile equipment is now under development. The ability to successfully transmit facsimile over a satellite link has already been demonstrated.

Broadcast Messages

There is also a demand to broadcast voice and data messages to a group of aircraft. Voice messages can be accommodated quite easily since only reception is required; this allows the service to be introduced without absorbing any of the limited power available from the aircraft HPA. Data messages (similar to the Oracle and Seefax messages currently being transmitted over TV channels in the U.K.) could also be transmitted for subsequent display over the aircraft's video equipment.

Computer-to-Computer Communications

With the increased use of portable and lap-top computers a demand has arisen to support their use by passengers while they are on the aircraft. A simple interface unit would permit this; the main limitation is the feasibility of certifying computer equipment for use by the public on aircraft.

Mobile Telephone Service

There would be considerable benefits to allowing the use of portable telephones on an aircraft. Again, the main problem is related to licensing of the equipment rather than technical problems in its use. If this could be overcome such a service would most likely be limited to the air-to-ground direction.

CONCLUSIONS.

Over the next few years, enhancements to the Inmarsat aeronautical satellite system will be incremental in nature. More exotic developments, such as direct television broadcasts from satellite to aircraft, will eventually be introduced, but will required a switch to a new portion of the spectrum. All projections show that the L-band spectrum currently allocated to aeronautical mobile services will only be adequate for essential demands by the end of the decade. As the amount of traffic required to support safety and regularity of flight increases there will be pressure to move public correspondence to these other frequencies. This could well be the catalyst for introducing these new services.

REFERENCES

1. **Couluris, G.J. and B. Conrad.** 1981 *Oceanic Area System Improvement Study (OASIS)--Volume 1 Final Report* FAA-EM-81-17 (Menlo Park, California, SRI International)

2. ----1986. *Report of RTCA Special Committee 165--User Requirements for Future CNS Systems, Including Space Technology Applications.* DO-193 (Washington D.C., Radio Technical Commission for Aeronautics)

3. ----1988. *Special Committee on Future Air Navigation Systems--Fourth Meeting Report* Doc 9524, FANS/4 (Montreal, Quebec, International Civil Aviation Organization)

4. ---1989. *Aviation Satellite Communication System: Part 2, Equipment Design and Equipment Functional Description.* ARINC Characteristic 741 (Annapolis, Maryland, Aeronautical Radio Inc.)

5. **Gronemeyer, S.A.** 1976. MSK and O-QPSK Communications *IEEE Transactions on Communications* Vol. COM-24, pages 809-830

6. **Fang, R.J.F.** 1981. Quarternary Transmission Channels over Satellite Channels with Cascaded Nonlinear Elements and Adjacent Channel Interference *IEEE Transactions on Communications* Vol. COM-29, pages 567-581

7. **Winters, J.H.** 1984. Differential Detection with Intersystem Interference and Frequency Uncertainty. *IEEE Transactions on Communications* Vol. COM-32, pages 25-33

8. **Crowe, D.P.** 1988. *Selection of Voice Codec for the Aeronautical Satellite Service. (Ipswich, British Telecom Research Labs.)*

Session 14
Modulation and Coding - III

Session Chairman - *Suzo Kato,*
Nippon Telegraph & Telephone Corp., Japan
Session Organizer - *John Lodge,* DOC

14

Considerations of Digital Phase Modulation for Narrowband Satellite Mobile Communication

Knut Grythe
ELAB-RUNIT
N-7034 TRONDHEIM, NORWAY
Phone: +47-7-592683
Fax : +47-7-594302

ABSTRACT

The Inmarsat-M system[1] for mobile satellite communication is specified as a FDMA system, applying Offset QPSK for transmitting 8 kbit/sec in 10 kHz user channel bandwidth.

In this paper we consider Digital Phase Modulation, DPM[2], as an alternative modulation format for Inmarsat-M. DPM is similar to Continuous Phase Modulation, CPM[3], except that DPM has a finite memory in the premodulator filter with a continuous varying modulation index.

It is shown that DPM with 64 states in the VA obtains a lower Bit Error Rate, BER.

Results for a 5 kHz system, with the same 8 kbit/sec transmitted bitstream, is also presented.

Introduction

The specifications for the Inmarsat-M system is based upon the use of OQPSK as modulation format, combined with convolutional codes to obtain the wanted BER of the speech and data services. OQPSK is selected due to its spectral properties when exposed to a nonlinear transmitting chain.

As an alternative to OQPSK, ELAB-RUNIT has been considering DPM for the same transmission specifications, as those defined for OQPSK. This is a required adjacent interference level, under specific relative power levels between the centre channel and the two neighbours.

In this paper, we first review DPM. The performance of DPM compared to OQPSK for the 10 kHz channel is shown along with results for the imaginary case of lowering the bandwidth to 5 kHz and maintaining the transmitted bitrate.

Digital Phase Modulation

Here we present DPM and some of its properties. Reference 1 has further informations on DPM.

DPM Transmitter

The DPM signal is equal to

$$\mathrm{Re}(s(t)) = \mathrm{Re}\left(\exp\left(j2\pi h \sum_{l=1}^{\infty} a_l h_p(t-lT_s)\right)\right) \qquad (1)$$

where

a_l - transmitted symbols, $a_l \in \{(M-1),..,1,-1,..,-(M-1)\}$

$h_p(t)$ - phase shaping filter

$$h_p(t) \begin{cases} 0, & 0 \le t \le T_L \\ 0, & \text{otherwise} \end{cases}$$

T_s - symbol duration

h - modulation index

$\mathrm{Re}(x)$ - real part of x

Observe that (1) has constant envelope and is of the same form as CPM, but with one important difference. The duration of $h_p(t)$, which for CPM is infinite, is for DMP finite of length T_L.

In a practical realization, s(t) is digital, that is s(nD). The oversampling is defined by

$$\eta = \frac{T_s}{D} \qquad (2)$$

where a typical value of η is 4 to 12, depending on the filter $h_p(t)$ and h. D gives the sampling clock relative to the symbol rate.

The value of the modulation index, h, influences the phase excursions. We have

$$(M-1)\pi h = \max\{\phi(t+T_s) - \phi(t)\} \qquad (3)$$

When discussing h, we should emphasize one important practical difference between DPM and CPM. In CPM, h must have a value[2],

$$h_{CPM} = \frac{2k}{q} \qquad (4)$$

due to the receiver implementation, while in DPM

$$h_{DPM} \in R^+ \qquad (5)$$

where R^+ denotes the set of non-negative real numbers.

This valuable freedom which DPM have, proves to be very useful in designing systems under various constraints like tuning to spectral requirements and/or BER performance curves.

The spectrum of s(nD) has a continuous part and a discrete part[1]. In Fig. 1 we show an example of the continuous part. The modulation parameters are

$$M = 8, \quad T_L = 3, \quad h = 0.56, \quad \eta = 8$$

and the filter is a raised cosine (RC) with

$$h_p(t) = \frac{1}{4T_L}(1-\cos(\frac{2\pi t}{T_L})), \quad 0 \le t \le T_L \qquad (6)$$

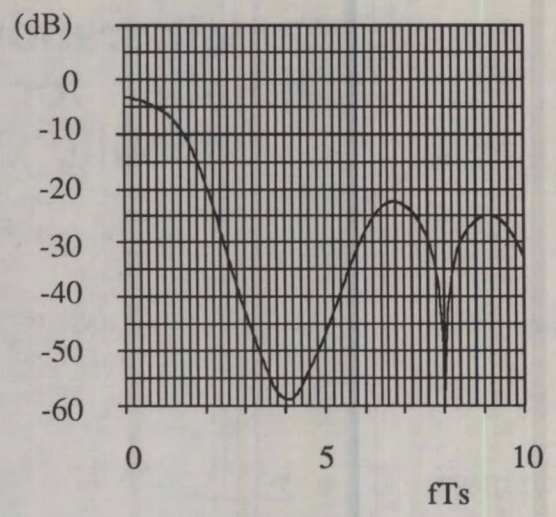


Fig. 1. Spectrum of DPM, continuous part. M=8, $T_L=3\cdot T_s$, $\eta=8$, h=0.56.

In a real system, s(nD) is lowpass filtered so that only the main lobe are transmitted. This results in minor amplitude variations.

General considerations concerning the spectral properties are:

larger h : increased bandwidth
larger T_L: more compact spectrum

DPM Demodulation

Demodulation of DPM is usually based on the application of the Viterbi algorithm (VA). This requires a coherent demodulator for being optimum in the maximum likelihood (ML) sense, since the VA does not give ML performance using a noncoherent metric.

The coherent metrics are in white gaussian noise defined as

$$\lambda_n = \mathrm{Re}(r_n \cdot s^*_{n,j}) \qquad (7)$$

where

r_n - received signal sample (complex) at time n.

$s_{n,j}$ - one of the possible phases, signals, at time n.

$$j = 1:M^r, \quad r = \left\lfloor \frac{T_L}{T_s} \right\rfloor$$

where we defined $\lfloor\ \rfloor$ as the next closest integer, e.g. $\lfloor 6.7 \rfloor = 7$.

The number of states and transition in the VA for DPM are equal to

$$N_S = M^{r-1}$$
$$N_{tr} = M^r \qquad (8)$$

The states are only correlative due to the finite memory of $h_p(t)$ making N_s and N_{tr} independent of $^p h$. CPM has in addition phase states, caused by the infinite memory.

The shape of the BER curve of Fig. 2 is typical for DPM. However, it is possible to shift the crossing point with the QPSK curve and to change the tilt

$$\frac{\partial BER}{\partial (E_b/N_0)} \qquad (9)$$

by varying the parameters involved. These are, as we know from the definition of the DPM signal, h, M and T_L.

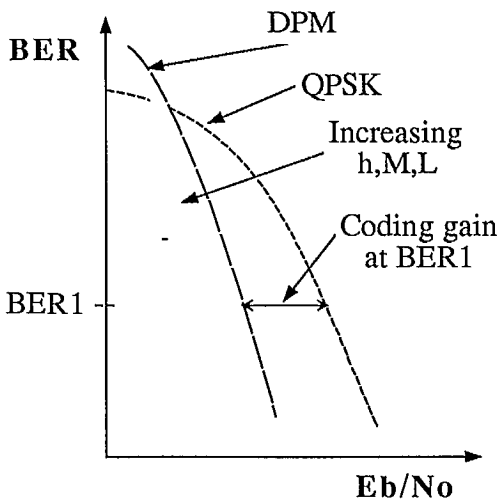


Fig. 2. The general trend of the behaviour of the DPM BER curve for variations of the parameters h, M and T_L.

One important observation is that the parameters are not independent. Therefore, Fig. 2 must be interpreted as an illustration of the fact that the BER curve is controlled by all the three parameters.

Practical systems exhibit a varying phase error between the received carrier phase and the reference in the receiver.

The VA requires a tracking error, $\delta\theta$, below a given value, β, to keep the performance degradation as low as possible. That is

$$|\delta\theta| \leq \beta \qquad (10)$$

Examples of simulations with a <u>non-coherent</u> metric applied in the VA, show degradations in E_b/N_0 versus $\delta\theta$ as indicated in Table 1.

Table 1. Degradation in E_b/N_0 at BER = 10^{-2} for two values of $\delta\theta$, using the noncoherent metric in the VA algorithm. M=8, T_L=3, h=0.26.

	[dB]	
$\Delta E_b/N_0$	0.1	0.5
$\delta\theta[^0]$	5	10

The incoherent metric is defined as

$$\lambda_j = \left| \sum_{i=1}^{N} r_i \cdot s^*_{i,j} \right|, \ V_j \qquad (11)$$

where the absolute value is applied instead of the real part from (8).

Using VA and the incoherent metric is an approximate solution since Eq. (11) is not directly summable, but several simulations show that under phase errors it is more tolerable than the coherent metric. Under zero phase error, a minor difference between the performance of the coherent and non-coherent matric is observed.

To compensate for a phase shift, $\delta\theta(t)$, due to doppler, a wide loop bandwidth must be applied. This very often leads to a contradiation with

the oscillator phase noise requirements where a narrower bandwidth may be more optimum.

One possible way of overcoming or reducing this problem is to extend the DPM trellis with phase states. The resulting trellis structure then becomes a parallel structure consisting of a number of similar trellis units. Each trellis unit is optimum for a receiver coordinate system offset with Θ_i from the transmitter axis where

$i = 1:N_{PT}$

N_{PT} - number of parallel trellis structures

If we now tolerate $\pm 10^0$ for each unit, $N_{PT}=5$ covers

$\Theta_{max} - \Theta_{min} = 100^0$

Consequently, a much wider phase variation range is available as input to the VA, lifting some of the burden off the phase tracking system. At the same time we see that if $\delta\Theta$ reaches a value of $20\text{-}30^0$, a PLL solution reaches its linear range limit, making a forward carrier phase estimator solution more attractive[4]. Due to the delay requirements on a satellite link, we can not adapt the same tracking methods as available in a landmobile system.

Performance Comparisons for Inmarsat-M, 8 kb/s in 10 kHz.

The following[1] transmission parameters are valid

- Bitrate 8 kb/s
- Bandwidth 10 kHz
- Adjacent channel interference degradation $\Delta E_b/N_0 \leq 0.1$ dB.
- 3/4 rate convolutional code, constraint length 7.

Fig. 3 gives the result.

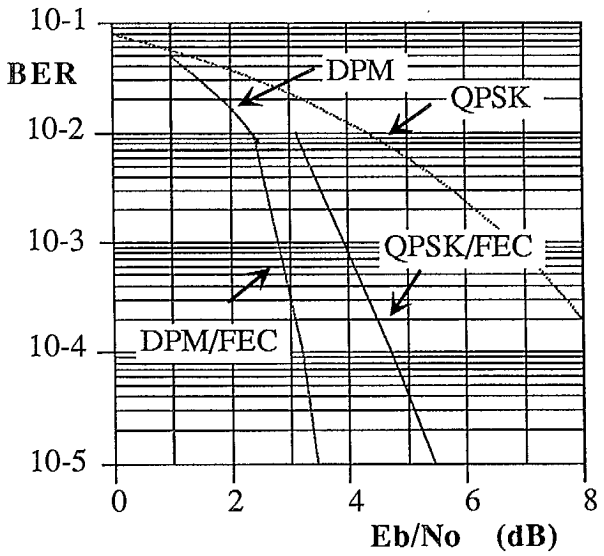


Fig. 3. 8 kbit/sec in 10 kHz user channels. Bit error rate for QPSK, QPSK/FEC and DPM/FEC. FEC: 3/4 rate, K=7, number of states = 64, DPM: M=8, Filter 3RC, h=0.56.

The results exhibit two important properties.

- We can reduce satellite power for a given BER.
- The BER curve for DPM is steeper than for OQPSK.

A consequence of this, is that for a given probability density of the received signal power, p(P), we obtain

$T_{DPM}(BER \leq BER_1) > T_{OQPSK}(BER \leq BER_1)$

where $T(BER \leq BER_1)$ is the percentage of time the BER is lower or equal to BER_1.

For the system this implies that the user will have access to a given quality a larger portion of time with DPM than with OQPSK or put another way the average quality obtained by DPM is expected to be higher than in a OQPSK system.

Performance of 8 kb/s in 5 kHz Bandwidth

Result for this imaginary case is shown in Fig. 4. The DPM parameters are M=8, filter 3RC and h=0.26.

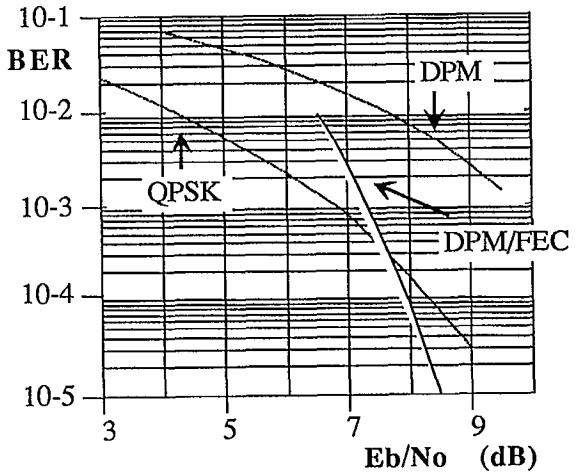


Fig. 4. BER curve for QPSK, DPM and DPM/FEC. DPM: M=8, filter= 3RC, FEC: 3/4 rate. Coherent demodulation.

If we compare Fig. 3 with Fig. 4, we see that the reduction of bandwidth from 10 kHz to 5 kHz and maintaining the 8 kb/s rate transmitted requires an increase in available power per channel of approximately 4 dB and 10^{-2} and 5 dB at 10^{-5}.

With respect to unfiltered 8PSK, DPM with h=0.26 has a negative gain of 0.3 dB at 10^{-2} and a crossing to positive gain around E_b/N_0=8.7 dB.

Conclusion

DPM is an alternative modulation method for narrowband satellite communication. We have shown that for the Inmarsat-M spesifications, it performs better than OQPSK. With respect to carrier phase estimation, the type of estimator has to be considered by taking into account the demodulation metric and the complexity of the VA unit.

Acknowledgement

The project has been financed by the Norwegian Telecommunications Administration (NTA). The author would like to thank S.R. Skaland, NTA, for many valuable discussions. Also B. Risløw at ELAB-RUNIT who has programmed the simulation software.

References

1 INMARSAT: "Standard-M system definition manual, Modul 1, Draft 1.0". Jan. 1989.

2 T. Maseng: "Digitally Phase Modulated (DPM) Signals". IEEE Trans. on Comm., Sept. 1985, pp. 911-918.

3 Anderson, Aulin, Sundberg: "Digital Phase Modulation". Plenum Press. 1986.

4 G. Ascheid et al.: "An all digital receiver architecture for bandwidth efficient transmission at high data rates". IEEE Trans. on Comm., August 1989.

Performance of Concatenated Reed-Solomon Trellis-Coded Modulation over Rician Fading Channels

Michael L. Moher, John H. Lodge
Communications Research Centre
P.O. Box 11490, Station H, Ottawa, Canada K2H 8S2
Ph.: (613) 998-8669, FAX: (613) 990-6339

ABSTRACT. In this paper a concatenated coding scheme for providing very reliable data over mobile-satellite channels at power levels similar to those used for vocoded speech is described. The outer code is a shortened Reed-Solomon code which provides error detection as well as error correction capabilities. The inner code is a one-dimensional 8-state trellis code applied independently to both the inphase and quadrature channels. To achieve the full error correction potential of this inner code, the code symbols are multiplexed with a pilot sequence which is used to provide dynamic channel estimation and coherent detection. The implementation structure of this scheme is discussed and its performance is estimated.

1.0 INTRODUCTION

There are several considerations to be made when designing modulation and coding strategies for reliable communications in the mobile-satellite environment. One of the major technical considerations is the fading of the received signal due to vehicle motion. Conventional modulation and detection schemes, such as differentially detected MSK and BPSK, provide poor performance over moderately and severely faded channels without the use of coding. Furthermore, current mobile-satellite systems are typically power-limited but it is expected that future systems will be bandwidth-limited which implies that bandwidth conservation is an important consideration when choosing a coding scheme. In addition, mobile-satellite systems are concerned about minimizing the mobile terminal complexity and cost while maintaining its flexibility.

One way of satisfying these constraints and to meet the requirement of providing highly reliable data is to use a concatenated coding scheme where the inner codec is chosen to be bandwidth efficient and to provide good performance by itself, performance acceptable for vocoded data, while the outer codec would be a completely independent codec to be used only when very reliable data is required.

2.0 THE CODES

The Inner Code

Trellis coded modulation techniques are strong candidates for the inner code because they include low complexity codes which provide significant coding gain and diversity without increasing the transmission bandwidth. The drawback of applying such techniques to fading channels is that they require coherent detection to achieve their full potential. However, this drawback can be overcome by interleaving a known pilot sequence with the trellis coded (TCM) sequence in an scheme known as TCMP [1],[2]. In TCMP, the inclusion of a pilot sequence, at the expense of a fractional increase in power and bandwidth, allows the receiver to dynamically estimate the channel and perform a form of coherent detection. The bit error rate performance of TCMP with the optimum two-term 8-state one-dimensional code [3] applied to both the inphase and quadrature channels is shown in Figure 1. The scheme clearly provides very robust performance under Rician fading conditions.

The Outer Code

The factors to consider when choosing candidates for the outer code include the tendency of the inner code to produce short error bursts because of its

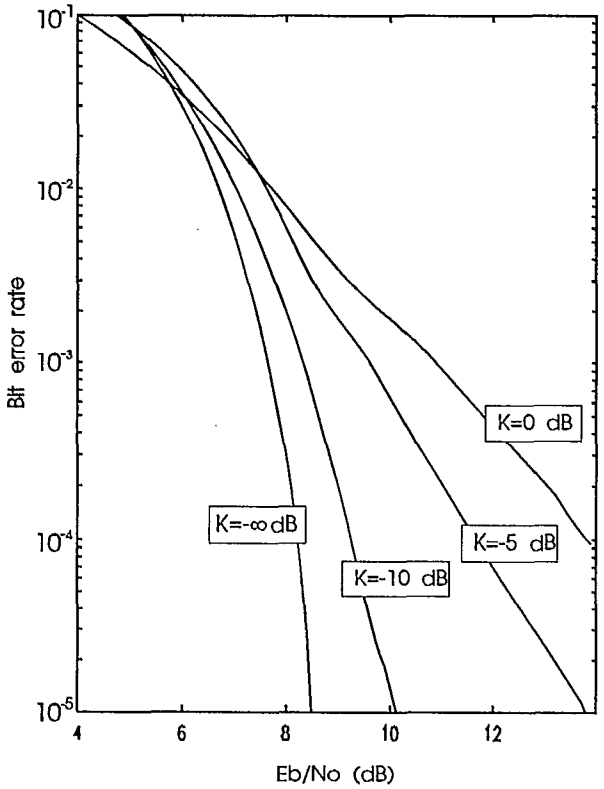


Figure 1. Performance of TCMP modulation and coding strategy over Rician fading channels.

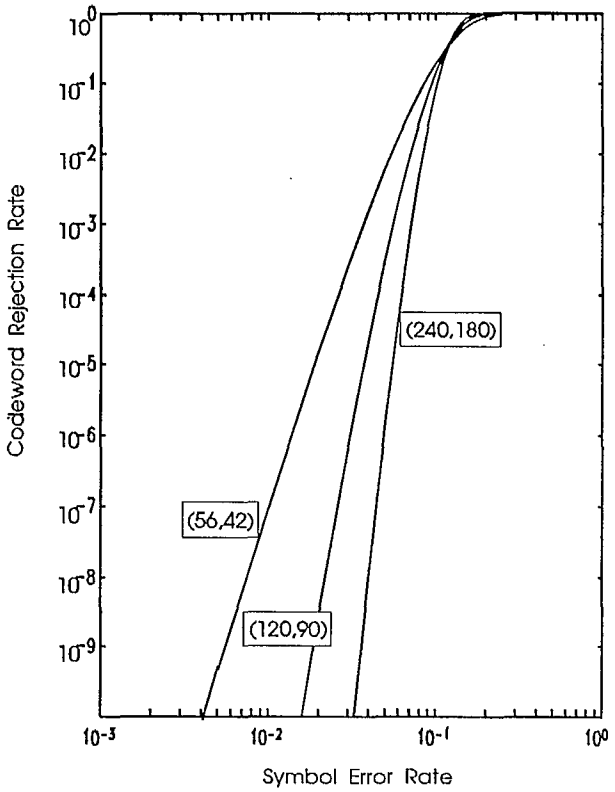


Figure 2. Performance of three shortened Reed-Solomon codes as a function of symbol error rate.

trellis-like structure, and the requirement for both error correction and error detection capabilities to provide highly reliable data. These factors suggest using Reed-Solomon block codes, which are among the best linear block codes for correcting bursts of errors up to a given length. The data to be transmitted will generally be in a byte format thus a symbol size of 8 bits would be the most convenient. This suggests a primitive Reed-Solomon (RS) code (255,k) where k is the number of information bytes in a codeword but it is also of interest to have the code rate of the outer code as a low order rational.

A code satisfying these requirements is the shortened RS code (240,180), shortened from (255,195). This is a rate 3/4 code but it also has the advantage that the same codeword length can be used for rate 7/8 and 15/16 codes. The RS(240,180) code can correct a codeword having up to 30 bytes in error, and its codeword error rate performance as a function of the input symbol error rate, assuming independent symbol errors, is shown in Figure 2.

The other performance criterion is the probability of decoding error. Since the non-shortened RS codes are maximum distance codes, their probability of a decoding error can be determined analytically [4]. These probabilities of decoding error for the full length code provide upper bounds on the error detection capabilities of the shortened RS codes, and are shown in Table 1. These results indicate that the probability of decoding error for the (240,180) code is less than 10^{-33} which, for practical purposes, is zero.

Prob. of Symbol Error	*Probability of Decoding Error*			
	(31,23)	(63,47)	(127,97)	(255,195)
.6	$2E-2$	$7E-6$	$2E-13$	$3E-34$
.5	$1E-2$	$6E-6$	$1E-13$	$3E-34$
.4	$1E-2$	$6E-6$	$1E-13$	$2E-34$
.2	$6E-3$	$3E-6$	$7E-14$	$1E-34$
.1	$1E-3$	$4E-7$	$7E-15$	$6E-36$
.01	$7E-8$	$2E-14$	$3E-27$	$3E-59$

Table 1. Probability of decoding error versus probability of input symbol error for various Reed-Solomon codes.

For comparison purposes, the performance of the two shortened RS codes (120,90) and (56,42), which are shortened from (127,97) and (63,49) RS codes, respectively, are also shown in Figure 2 and Table 1. These also are rate 3/4 codes, and even with these shorter codes, the performance and probability of decoding error are still very good.

3.0 MODEM-CODEC STRUCTURE

To maximize the benefits obtained from a concatenated coding strategy, the two codes clearly must be combined in a manner which emphasizes the strengths of each. Trellis decoders are designed with the basic premise that errors in the input data sequence are independent. Consequently, to effectively use these codes on a slow fading channel interleaving must be performed to break up the error bursts caused by the channel and to create a sequence of errors which are approximately independent. In Figure 3 the interleaving of the trellis code is performed using a multiplexed codec approach which allows simple insertion of the pilot sequence.

When errors occur in a trellis codec, they tend to occur in bursts. To take advantage of the burst error-correcting potential of the RS code, the input to the multiplexed codecs were interleaved on a subsymbol (RS) basis as shown in Figure 3. Thus the bursts of bit errors produced by a trellis decoder tend to be combined into a smaller number of RS symbols.

In the simulation of the concatenated coding scheme described in Figure 3, an information rate of 3600 bits per second was assumed. This incoming data was buffered and coded using the rate $\frac{3}{4}$ RS code to a rate of 4800 bits per second. Each of the trellis codecs consisted of two one-dimensional

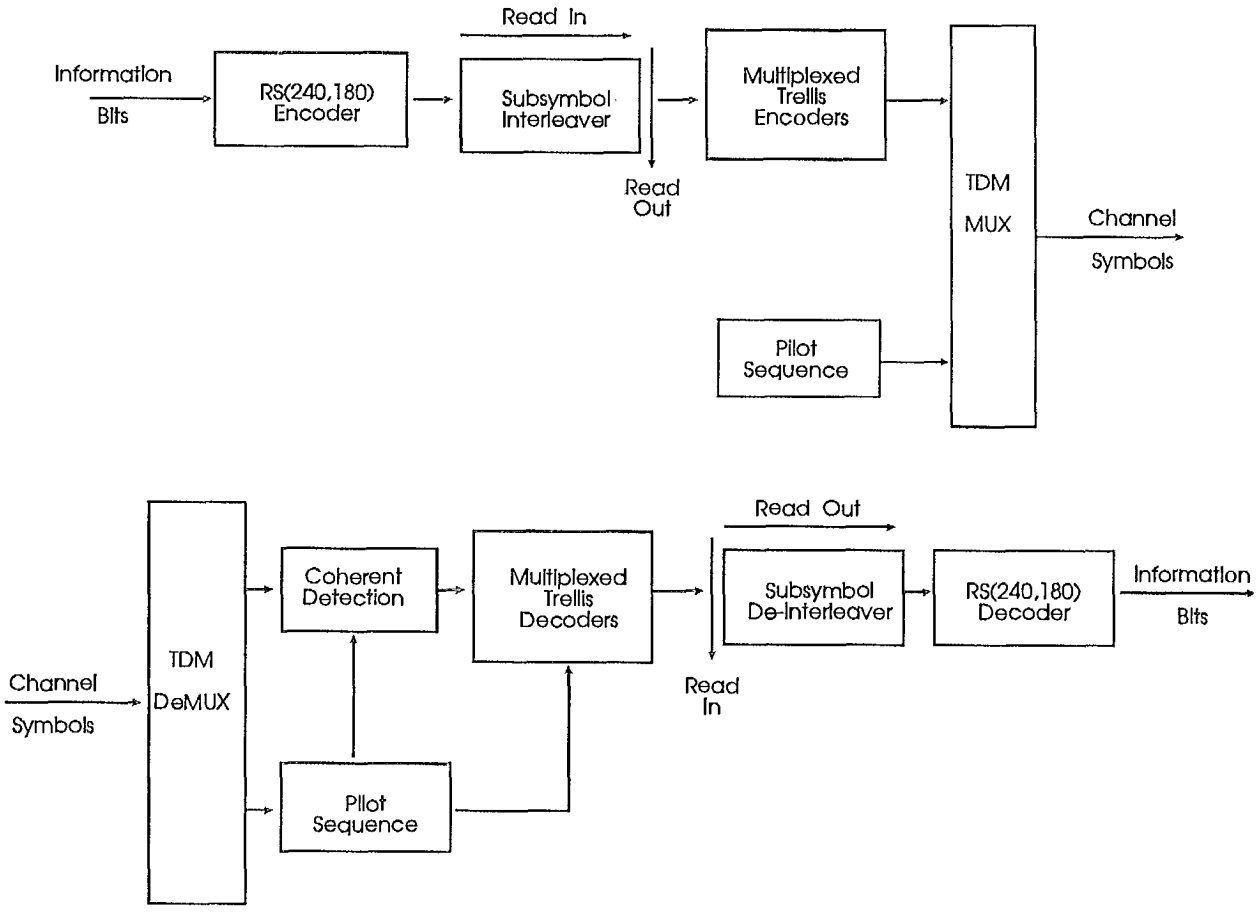


Figure 3. Illustration of concatenated coding scheme with subsymbol interleaving matched to the interleaving depth of the trellis code.

encoders whose outputs are combined in quadrature to produce a 16-QAM constellation. The output symbols of the RS code were allocated sequentially to the parallel trellis encoders. The output symbols of the multiplexed codecs — at a combined rate of 2400 symbols per second — were multiplexed with a pilot sequence of one quarter that rate, filtered to obtain the desired square root raised cosine spectral shape having a 50% rolloff factor, and transmitted at a rate of 3000 symbols per second.

The detection strategy for TCMP described in [1] used the pilot symbols to estimate the phase and gain of the channel and bypassed the need for explicit carrier recovery and short-term gain control. The output of each of the multiplexed trellis codecs was stored in an 8-bit buffer; which served as the symbol inputs to the RS decoding algorithm.

4.0 CONCATENATED PERFORMANCE

If the symbol errors are independent then the performance of the outer RS code can be characterized by the binomial distribution with parameter, θ, the symbol error rate. However, unless the interleaving strategy is perfect, the errors in the input symbol stream to the RS decoder will be correlated due to the combination of the trellis inner code and the slow fading channel. Also note that the average symbol error rate is not a simple function of the output bit error rate of the inner trellis decoder, since the correlation properties of these latter errors are a function of the channel conditions. However, a lower bound on performance can be obtained by noting that in the ideal case, the symbol error rate and the bit error rate will be the same (eight bit errors are mapped into a single symbol error), and the symbol errors are uncorrelated. An upper bound is provided by the case when the bit errors are totally uncorrelated and the symbol error rate is eight times the bit error rate. These bounds are shown in Figure 4 as a function of the input bit error rate, together with some simulation results. These simulated results are also reported in Table 2 as a function of the channel fading and noise parameters. The channel conditions simulated were those of a Rician fading channel with a fading bandwidth of 120 Hz and with additive white gaussian noise (AWGN). The fading conditions were varied from a static channel ($K=-\infty$ dB) to a severely faded channel (K=0 dB). Over the range of E_b/N_o simulated the percentage of blocks rejected ranges from 0 to 100% and the knee

E_b/N_o (dB)	K-factor (dB) $-\infty$	-10	-5	0	cdwd. sim.
6	.94	.96	.96	.47	168
7	.012	.06	.13	.04	336
8	.0	.0	.002	.002	504
9	.0	.0	.0	.0	1472

Table 2. Fraction of RS(240,180) codewords rejected in a 120 Hz Rician fading channel.

of the performance curve lies between 7 and 8 dB in E_b/N_o. Correlating this with the output error rates of the trellis codec, given in Table 3, it is noted that the performance of the Reed-Solomon codec starts to degrade at an average input bit error rate of 1×10^{-2}. To achieve a codeword rejection ratio of 1 in 10^7, the lower bound in Figure 4 together with the worst case performance of the TCMP strategy (K = 0 dB in Figure 1) indicates that an E_b/N_o of 9.6 dB is required. (An additional 1.2 dB is added to the performance shown in Figure 1 to account for power loss due to the rate $\frac{3}{4}$ RS code.)

At first glance, this level of performance may not appear to be outstanding until one realizes that it is obtained in a severely faded channel. Under AWGN

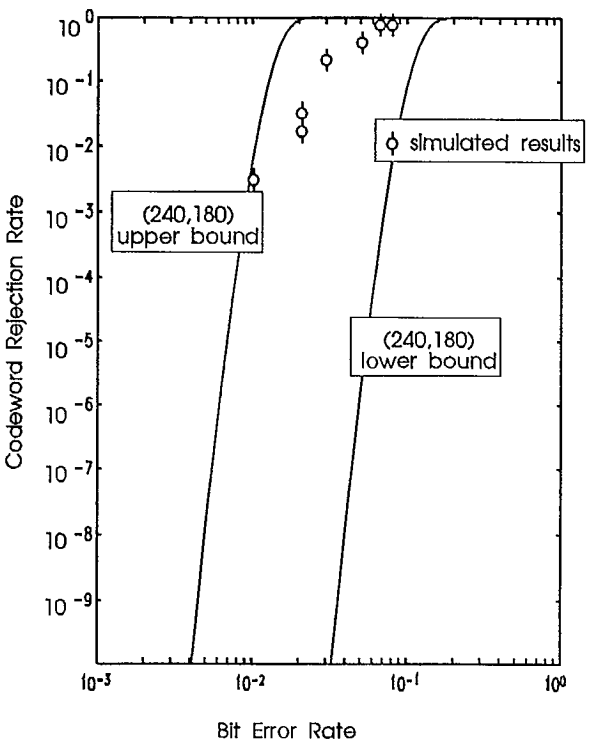


Figure 4. Bounds of performance of the (240,180) code as a function of the bit error rate.

E_b/N_o (dB)	K-factor (dB) $-\infty$	-10	-5	0
6	$8E{-}2$	$8E{-}2$	$7E{-}2$	$5E{-}2$
7	$2E{-}2$	$3E{-}2$	$3E{-}2$	$2E{-}2$
8	$5E{-}2$	$8E{-}3$	$1E{-}2$	$1E{-}2$
9	$6E{-}4$	$2E{-}3$	$3E{-}3$	$5E{-}3$

Table 3. Average output error rates of the inner trellis code which produced the results in Table 1.

E_b/N_o (dB)	K-factor (dB) $-\infty$	-10	-5	0	cdwd. sim.
6	.67	.69	.61	.32	360
7	.067	.11	.18	.10	720
8	.0	.007	.010	.016	1432
9	.0	.0	.001	.003	2348

Table 5. Fraction of RS(56,42) codewords rejected in a 120 Hz Rician fading channel.

conditions it can be shown using Figures 1 and 4 that a codeword rejection rate of 10^{-7}can be obtained with an E_b/N_o of between 6 and 7 dB. Furthermore, in AWGN the pilot sequence, which is crucial to the performance under fading conditions, could be removed for an approximate 2 dB [1] improvement in performance.

In Tables 4 and 5, the codeword rejection rates for the other rate 3/4 codes, RS (120,90) and RS (56,42), are shown. It may be surprising that the simulated performance of these codes is almost as good as, and in some cases better than, the longer code. However, if one considers that the number of bit errors in a received codeword will have greater variation with shorter block lengths then the results can be explained. At high input error rates, shorter codewords have a greater probability of having a correctable number of errors. Similarly, at low input error rates, shorter codewords have a greater probability of having an uncorrectable number of errors which would result in either a decoding failure or a decoding error. The E_b/N_o range for the simulation results happens to be the range in which the performance curves of the codes of different blocklengths crossover, as illustrated by Figure 2.

E_b/N_o (dB)	K-factor (dB) $-\infty$	-10	-5	0	cdwd. sim.
6	.78	.78	.73	.33	96
7	.05	.11	.12	.12	144
8	.0	.0	.005	.008	384
9	.0	.0	.0	.0	576

Table 4. Fraction of RS(120,90) codewords rejected in a 120 Hz Rician fading channel.

5.0 CONCLUSIONS

In this paper a concatenated coding scheme for providing very reliable data over Rician fading channels has been described. The inner code is a simple 8-state trellis coded modulation scheme with an interleaved pilot sequence. This code takes full advantage of soft decision and channel state information to provide robust performance with significant coding gain and little bandwidth expansion. The outer code is a shortened Reed-Solomon code which provides significant error correction capabilities as well as virtually perfect error detection performance. This concatenated scheme can provide very reliable data at power levels similar to those required for vocoded speech.

References

[1] M.L.Moher and J.H.Lodge, "TCMP — A Modulation and Coding Strategy for Rician Fading Channels", *IEEE J. Sel. Area. Comm.*, Vol.7, pp.1347-1355, December 1989.

[2] J.H.Lodge and M.L.Moher, "Time Diversity for Mobile Satellite Channels using Trellis Coded Modulations", *IEEE Global Telecom. Conf.*, pp. 8.7.1-8.7.5, November 1987.

[3] R.Calderbank and J.Mazo, "A New Description of Trellis Codes", *IEEE Trans. Inf. Th.*, Vol.28, pp.784-791, November 1984.

[4] R.E.Blahut, *Theory and Practice of Error Control Codes*, Reading MA: Addison-Wesley, 1983.

A Novel Scheme to Aid Coherent Detection of GMSK Signals in Fast Rayleigh Fading Channels

Patrick S.K. Leung*, Kamilo Feher
Digital Communications Research Laboratory
Department of Electrical Engineering and Computer Science
University of California, Davis
Davis, CA 95616, U.S.A.

Abstract.
A novel scheme to insert carrier pilot to GMSK signal using Binary Block Code (BBC) and a highpass filter in baseband is proposed. This allows the signal to be coherently demodulated even in a fast Rayleigh fading environment. As an illustrative example, the scheme is applied to a 16 kb/s GMSK signal, and its performance over a fast Rayleigh fading channel is investigated using computer simulation. This modem's "irreducible error rate" is found to be $Pe=5.5*10^{-5}$ which is more than one order of magnitude lower than that of differential detection. The modem's performance in Rician fading channel is currently under investigation.

1. Introduction.

GMSK [1] is a digital modulation with appealing characteristics for mobile communications. In slow fading channels, GMSK can be detected either coherently or non-coherently. In a fast fading channel environment however, accurate carrier recovery is difficult and non-coherent detection is preferred [15].

---

* On leave from Dept. of Electrical Engineering, Footscray Institute of Technology, Melbourne, Victoria, Australia.

A limiting factor in relatively low bit rate modem operation over fast Rayleigh fading channel is the "irreducible error rate". In coherent detection, this is caused by the deep fades that force the carrier recovery loop to hang up and lose synchronization even in the absence of channel noise [14]. In differential detection, it is caused by random FM noise that partially destroys the carrier phase coherency in adjacent symbol time. Theoretical and simulation studies [4] show that the irreducible error rate for a 16 kb/s GMSK system with BT=0.25 and 2T differential detection is approximately 10^{-3} when the Doppler frequency is f_D=100Hz [4].

A useful technique to lower the "irreducible error rate" is to transmit an in-band pilot tone at the carrier frequency.

Provided that mutual interference with the data signal is avoided, the received pilot can be extracted using a simple bandpass filter to aid accurate coherent demodulation. The main drawbacks of pilot schemes are (1) the pilot takes up additional power and/or bandwidth, and (2) it destroys the otherwise constant envelope property of certain modulations. However, under some

circumstances, the benefit of coherent detection can outweigh these shortcomings. Some well-known examples are the Tone Calibrated Technique (TCT) [5] and the Transparent-Tone-In-Band (TTIB) technique [6].

In the literature, there are reports on in-band carrier pilot insertion for linear modulations such as BPSK [7] ,MPSK[8], QPRS[9], and QAM[10]. But there appears to be little study on it's application to non-linear modulation. In this paper, we address the issue by extending the technique to GMSK. We propose a modified modem structure which allows a carrier pilot to be transmitted with GMSK signal in band. Our study demonstrates that for a given application over fast Rayleigh fading channel, the modem's irreducible error rate is reduced to $Pe=5.5*10^{-5}$ which is more than one order of magitude lower than that of differential detection.

2. BBC coded GMSK modem with in-band pilot.

GMSK is traditionally viewed as a nonlinear digital FM signal with Gaussian pre-modulation lowpass filtering (see Fig. 1). TTIB is a well-known carrier pilot insertion scheme with potential application to GMSK, however, its sophistication may not be warranted in some applications. The TCT technique is more straightforward, but it is not immediately obvious how the technique can be applied to non-linear modulations.

Fortunately, Laurent [12] described a method that allows some digital phase modulations to be constructed by superposition of amplitude modulated pulses p(t) in quadrature. In this paper, we apply Laurent's method to GMSK, and construct the signal as an offset quadrature amplitude modulation. This alternative view of GMSK suggests a convenient way to insert an in-band carrier pilot. We accomplish this by applying BBC line code and highpass filtering to the I/Q channel data for the creation of a spectral void zone at the carrier frequency.

The block diagram of the proposed modem is shown in Fig.3. The binary input data $\{b_k\}$ is split into I and Q channel streams with the latter delayed by Ts/2. Ts=2T(M-1)/M where T is the input data bit duration. The I/Q channel data are convolved with the signaling pulse p(t) before mixing with the quadrature carriers to form the GMSK signal. The waveforms p(t) corresponding to several BT values are shown in Figure 2. They are smooth functions with limited time duration, and their pulse width decreases with increasing BT value. In the limiting case where BT approach infinity, p(t) becomes a sine pulse that corresponds to MSK.

To facilitate carrier pilot insertion, both I/Q channel data are encoded by the Binary Block Code (BBC) [13] coder. BBC suppresses the signal's low frequency spectral component by restraining its running digital sum to be within bound. This code has simple codec structure, and low redundancy requirement. Its power spectrum $S_u(f;t)$ is given by [13]:

$$TS_u(f\,;T)=1-\left(\frac{\sin(M\omega T/2)}{M\sin(\omega T/2)}\right)^2 F(\omega T\,;M),$$

$$F(\omega T\ ;\ M) = 0.645\cos(M\omega T) - \frac{0.88}{M}\cos(2M\omega T)$$

$$- 0.85\frac{\cos(2M\omega T) - 0.2}{\cos(2M\omega T) - 2.6} \cdot \frac{1}{13 + 12\cos(2M\omega T)}$$

$$\times\left[\frac{2}{M}\left[-7.2+6.4\cos(M\omega T) - 10.8\cos(2M\omega T)+0.6\cos(3M\omega T)\right]\right.$$

$$\left. + 0.05\left[12\cos(4M\omega T)+2\cos(5M\omega T)+18\cos(6M\omega T)+3\cos(7M\omega T)\right]\right]$$

M odd and ω is frequency in rad/s.

Figure 4 shows $S_u(f;t)$ for M=3,7 and 11 where 1/(M-1) is the code's redundancy. In this study, the BBC code with M=11 is used. While it has a reasonable spectral gap (see Fig 4), this code only requires a 10 % redundancy. The highpass filter (with cutoff frequency f_H Hz) further limits the interaction between data and carrier pilot signal.

The received signal S(t) is filtered by the receive BPF with bandwidth equal to the bit-rate. The carrier pilot is recovered by the pilot BPF whose bandwidth f_B is set to be $f_B=f_H$. After appropriate conditioning, the recovered pilot coherently demodulates the received signal.

3. Performance in fast Rayleigh fading channel.

As an illustrative example, the performance of a 16kb/s GMSK modem with BT=0.25 and M=11 BBC coding in a fast Rayleigh fading environment is studied using computer simulation. The Rayleigh fading process is generated using the model described in [2].

(a) Power spectrum considerations.

Figure 5a. shows the normalized spectrum of the transmit signal where the carrier pilot power level is 10 dB below the data signal. Here f_c is the carrier frequency and $f_s=1/Ts$. This spectrum features a spectral void zone of 800 Hz that has been craved out by the f_H=400Hz ideal HPF and M=11 BBC coder. Figure 5b shows the spectrum of the same signal when fast Rayleigh fading with Doppler frequency f_D = 100 Hz is introduced. Both the pilot and data signal are seen to suffer spectral spreading as a result of fast fading. In this case though, the carrier pilot is still distinguishable from data signal, and it can be recovered with a suitable bandpass filter. Figure 5c shows the spectrum when the spectral void zone is only 400 Hz. Here, the carrier pilot merges with the data signal and carrier recovery would be difficult.

(b) BER Performance.

BER performance of the modem in a fast Rayleigh fading environment with f_D = 100 Hz is shown in Figure 6. Curve (a) plots the performance when the pilot BPF bandwidth is f_B=300 Hz. Curve (b) shows the performance with f_B=400 Hz. The two curves crossover at approximately C/N = 28dB. Below this crossover point, the modem with f_B=400 Hz has a slightly inferior performance. But above the crossover, the system with f_B=400Hz clearly performs better. This behavior can be explained as followed. When C/N is small, channel noise dominates the contributions to the phase jitter in the recovered pilot. The modem with f_B=400Hz admits more noise and hence causes more error. However, when C/N is large, the modem performance is largely determined by how accurately the recovered pilot tracks the channel fading characteristics. Here, the f_B=400Hz BPF recovers more completely the fading pilot's spectral contents, and

it brings about a lower bit error rate and a lower error floor ($Pe=5.5*10^{-5}$ Vs. $Pe=2.9*10^{-4}$). Of course, one needs to exercise caution in employing too large a pilot recovery BPF bandwidth. Firstly, it admits more noise, but more importantly, it demands a higher cutoff frequency in the HPF of the transmitter. This introduces more ISI and performance degradation.

For comparison, we also plot the performance of a BT=0.25 GMSK modem with 2T differential detection as curve (c). The Doppler frequency is the same at f_D = 100 Hz. Comparison between curves (b) and (c) demonstrates the advantage of pilot-aided coherent demodula-tion. For small C/N (<20dB), the coherent modem is about 2 dB better than differential detection. But more importantly, the coherent modem has an error floor of $Pe=5.5*10^{-5}$ which is more than one order of magnitude below that of the differential detection at $Pe=1.5*10^{-3}$.

Figure 7 plots the modem's error floor against the pilot recovery BPF's bandwidth f_B, with f_D=100 Hz. The error floor is found to be a concave function, and it reaches its minimum at f_B=400Hz.

(c) Effect of channel nonlinearity.

Figure 8 shows the effect of channel nonlinearity on the signal spectrum in a fast Rayleigh fading environment. The nonlinear characteristics is:

$$so(t)=si(t)/(1+0.414*|si(t)|)$$

where si(t) and so(t) are the input and output signal respectively [11]. The Doppler frequency is f_D=100Hz and f_B=400Hz. Comparison with Fig.5b indicates the presence of further spectral spreading. This is caused by the non-constant signal envelope resulting from the inserted pilot. However, the degree of spreading in this case is only minor.

4. Conclusion.

A scheme to insert carrier pilot to GMSK using BBC coding and highpass filter for coherent demodulation is proposed. As an illustrative example, the scheme is applied to a 16Kbps BT=0.25 GMSK signal over a fast Rayleigh fading channel with f_D=100Hz, and its performance is investigated. The modem is found to have an error floor of $Pe=5.5*10^{-5}$ which is more than one order of magnitude lower than that of differential detection. The effect of the pilot recovery BPF on modem performance is also investigated. The optimum bandwidth is found to be f_B=400Hz when the Doppler frequency is 100Hz. Finally, the effect of channel nonlinearity on the signal spectrum is also investigated. With the pilot power 10dB below signal, the amount of spectral spreading is found to be minor for a certain channel nonlinearity.

References:

[1] K. Murota, K. Hirade, "GMSK modulation for digital mobile radio telephony," IEEE Trans on Comm. Vol. COM-29, No.7. pp1044-1050, July 1981.

[2] J.R. Ball, "A real-time fading simulator for mobile radios," The Radio and Electronic Eng., Vol.52, pp.475-478, October 1982.

[3] A. Bateman, "A general analysis of bit error

probability for reference based BPSK mobile data transmission," IEEE Trans. on Comm. Vol. COM-37, pp398-402, April 1989.

[4] R. Vaughan, "Signals in mobile communications: A review," IEEE Trans. on Veh. Tech. Vol. VT-35, No.4. pp.133-145, Nov.1986.

[5] F. Davarian, "Mobile digital communications via Tone Calibration," IEEE Trans. on Veh. Tech. Vol. VT-36, pp.55-62, May 1987.

[6] J. McGeehan, A. Batman, "Phaselocked transparent tone-in-band (TTIB): a new spectrum of data over SSB mobile radio networks," IEEE Trans. on Comm. Vol. COM-32, pp.81-87, Jan. 1984.

[7] W. Rafferty, J.B. Anderson, G.J. Saulnier, J.R. Holm, "Laboratory measurements and a theoretical analysis of the TCT fading channel radio system," IEEE Trans. on Comm. Vol. COM-35, pp172-180 Feb. 1987.

[8] I. Korn, "Coherent detection of M-ary phase-shift keying in the satellite mobile channel with tone calibration," IEEE Trans. on Comm. Vol. COM-37, pp.997-1002 Oct. 1989.

[9] S. Gurunathan, K. Feher, "Rayleigh fade compensated QPRS coherent mobile radio modems operated at a practical 2b/s/Hz efficiency," IEEE Vehicular Technology Conference, Orlando, May 1990.

[10] P.S.K. Leung, K. Feher "Block-Inversion-Coded QAM systems," IEEE Trans. on Comm., Vol.36, No.7, pp.797-805, July 1988.

[11] J. Lodge, M. Moher, S. Crozier, "A comparison of data modulation techniques for land mobile satellite channels," IEEE Trans. on Veh. Tech. Vol. VT-36, No.1, pp.28-35 Feb. 1987.

[12] P. Laurent, "Exact and approximate construction of digital phase modulations by superposition of amplitude modulated pulses (AMP)," IEEE Trans. on Comm. Vol. COM-34, No.2, pp.150-160, Feb. 1986.

[13] L.J. Greenstein, "Spectrum of a binary signal block code for DC suppression," BSTJ Vol.53, No.6, pp.1103-1126, July-Aug. 1974.

[14] H. Suzuki, Y. Yamao, H. Kikuchi, "A single-chip MSK coherent demodulation for mobile radio transmission," IEEE Trans. on Veh. Tech. Vol. VT-34 No.4. pp.157-168, Nov. 1985

[15] K Hirade, "Mobile radio communications," in K. Feher "Advanced digital communications, systems and signal processing techniques," Prentice-hall, Englewood, 1987..

[16] M.K. Simon, C. Wang, "Differential detection of Gaussian MSK in a mobile radio environment," IEEE Trans. on Veh. Tech. Vol. VT-33, no4, pp307-320, Nov 1984..

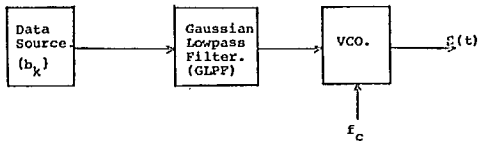


Figure 1. Conceptual generation of GMSK signal.

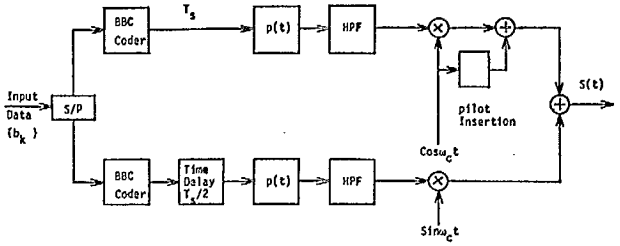


Figure 3a. Transmitter of the GMSK modem with carrier pilot insertion.

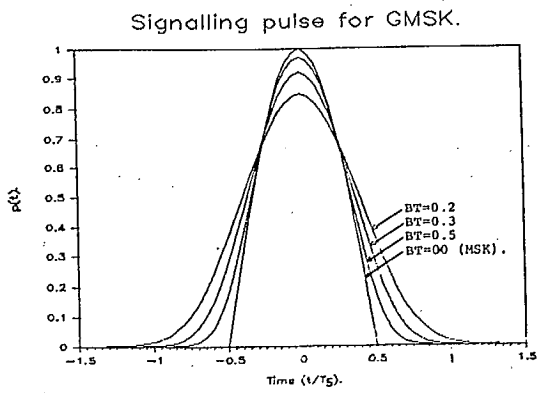


Figure 2. Signalling pulse p(t) for GMSK.

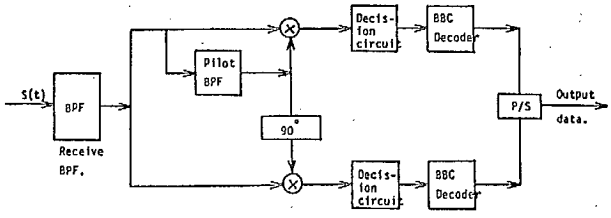


Figure 3b. Coherent detector of the coded GMSK modem.

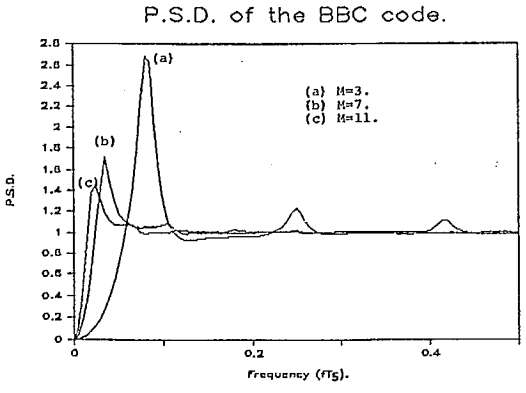


Figure 4. Power spectral density of BBC codes.

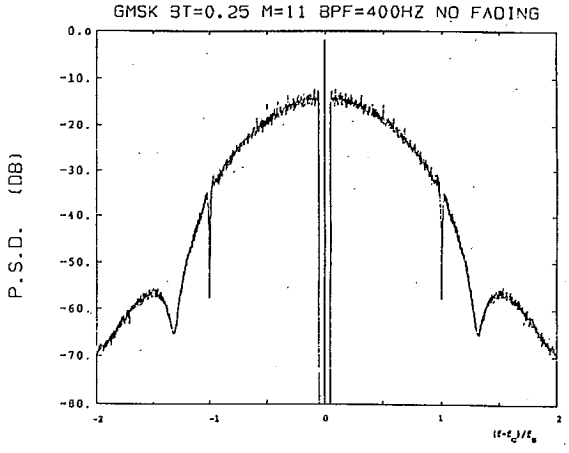


Figure 5a. Spectrum of M = 11 BBC coded GMSK signal with pilot, f_B = 400 Hz, No fading.

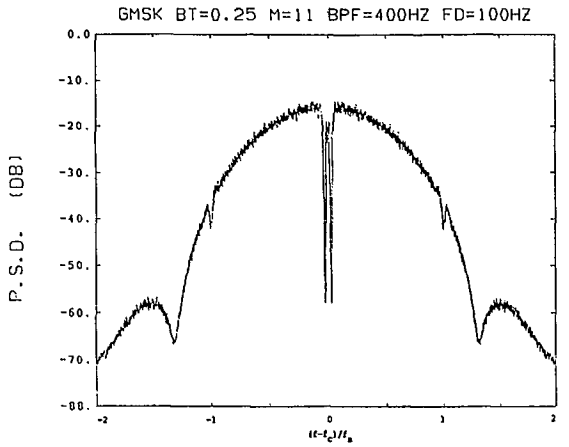


Figure 5b. Spectrum of M = 11 BBC coded GMSK signal with pilot. f_B = 400 Hz. Fading channel with f_D = 100 Hz.

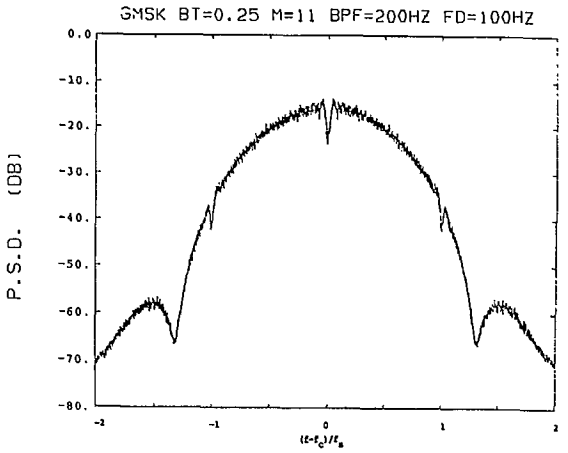


Figure 5c. Spectrum of M = 11 BBC coded GMSK signal with pilot. f_B = 200 Hz. Fading channel with f_D = 100 Hz.

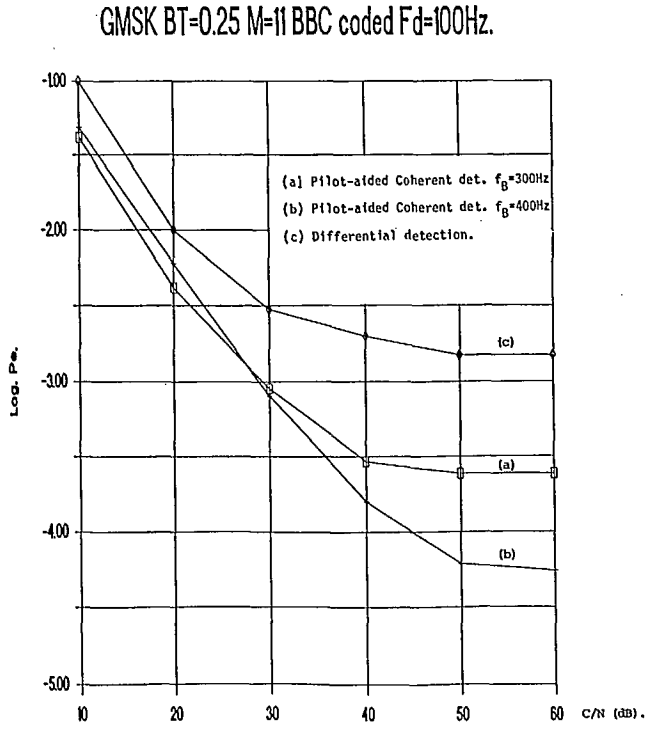


Figure 6. Performance of M = 11 BBC coded GMSK modem in fast Rayleigh fading channel. 16 kb/s. BT = 0.25. Doppler frequency f_D = 100 Hz. Receive BPF bandwidth = bit rate.

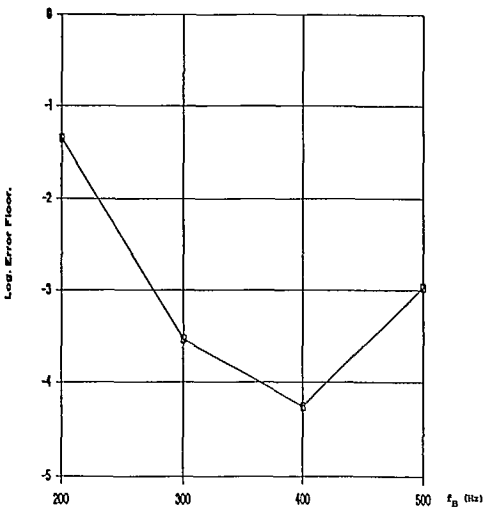


Figure 7. Error floor of M = 11 BBC coded GMSK signal vs pilot recovery. BPF bandwidth.

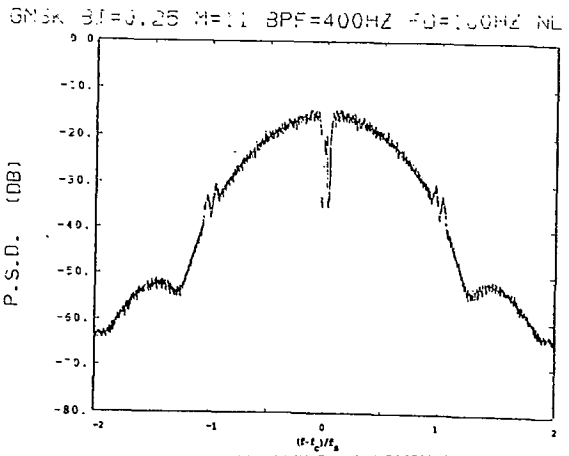


Figure 8. Spectrum of M = 11 BBC coded GMSK signal with pilot. f_B = 400 Hz. Nonlinear channel with f_D = 100 Hz.

Analogue and Digital Linear Modulation Techniques for Mobile Satellite

W.J. Whitmarsh, A. Bateman and J.P. McGeehan
University of Bristol
Centre for Communications Research
Queens Building, University Walk
Bristol BS8 1TR, United Kingdom
Tel: +44 272 303726, Fax: +44 272 255265

Abstract

The choice of modulation format for a mobile satellite service is complex. This paper summarises the subjective performance of a variety of candidate schemes and voice coder technologies. It is shown that good performance can be achieved with both analogue and digital voice systems, although the analogue system gives superior performance in fading. The results highlight the need for flexibility in the choice of signaling format. Linear transceiver technology capable of utilising many forms of narrowband modulation are described.

Introduction

The mobile satellite environment places severe constraints on the selection of modulation systems. It is important to balance the power efficiency of any scheme that is chosen with its bandwidth utilisation. The selection should seek to minimise the bandwidth required for a single voice channel (or corresponding data channel) while maximising the quality of performance at any given carrier to noise level. This paper discusses the subjective performance of various modulation schemes and speech processing techniques with application to the mobile satellite service. The discussion is based on the mean opinion score (MOS) assessment of voice quality with measurements drawn from a number of recent publications and laboratory measurements. A comparison is made between the schemes on the basis of both bandwidth and speech quality at given carrier to noise levels.

It is interesting to note how the performance of an analogue form of linear modulation (ACSSB) compares with that of a digitised voice source used in conjunction with an appropriate data modulation scheme. For a given power/bandwidth budget, the linear analogue system performs at least as well as digital equivalents using the present generation of voice coder technology. This is perhaps surprising given the current trend towards digital implementations. It should be noted that the analogue scheme is most favourable in a fading environment (Ricean) due to the absence of any sharp performance threshold as received signal strength is lowered.

Future advances in vocoder technology will undoubtedly make digital voice more attractive and it is prudent to select a modulation system, and consequently a transceiver design, that can efficiently convey both digital and analogue traffic with the same hardware. A pilot aided form of linear modulation based on the Transparent Tone in Band (TTIB) technique [1] has been shown to be directly compatible with both analogue and digital traffic and would allow a smooth and cost effective transition between analogue and digital voice operation. The design of mobile terminal equipment intended for pilot aided linear modulation is briefly reviewed here before discussing the comparison of modulation systems.

The Linear Modulation Mobile Terminal

Recent developments in digital signal processing technology has made the exciting prospect of a reconfigurable transceiver, capable of handling a variety of analogue and digital signaling formats a realistic goal. Prototype equipment based on direct conversion techniques has been developed at the Centre for Communications Research at the University of Bristol for use in a variety of frequency bands [2]. Such a system is normally described as *linear* so long as its operation is

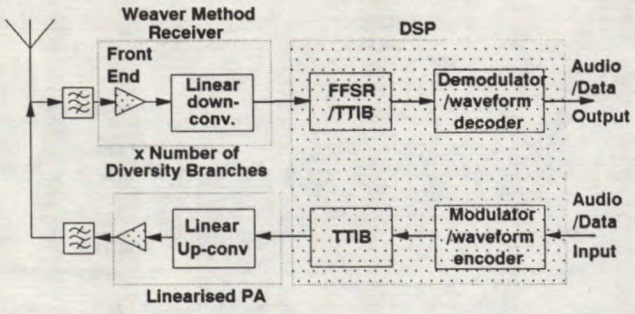


Figure 1: A Pilot Aided Linear Transceiver

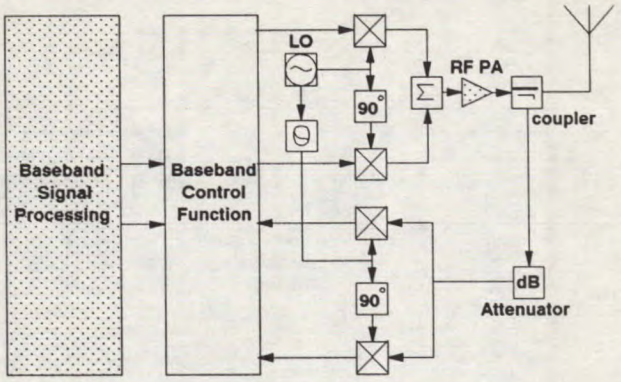


Figure 2: Feedback Linearisation of an RF Amplifier

independent of the amplitude and phase characteristics of the signaling format.

Both the transmitter and receiver systems are based on the Weaver frequency translation technique. This means that each frequency translation requires a quadrature local oscillator to drive the mixers. An advantage of using the Weaver method is that the image products fall within the users own channel, thus minimising the accuracy needed in the gain and phase match on the two quadrature signal paths.

Figure 1 gives a block diagram of a proposed transceiver architecture for a pilot aided linear transceiver. In such a system, distortions due to fading are removed in the receiver baseband processing by ascertaining the fading characteristics from the nature of the received pilot tone. The transceiver is divided into the fixed hardware sections of the linearised transmitter and the zero IF/single IF receiver, and the software reconfigurable baseband processing sections. In a pilot aided system, the baseband processing includes the generation of TTIB and fading correction using feedforward signal regeneration [3]. Processing is also required for whatever signaling format is to be transmitted through the system. It is intended that the particular signaling format chosen is not limited by the transceiver design.

Linear Transmitters

The transmit section of the transceiver must meet stringent requirements for linearity and power efficiency. Considerable success has been achieved at Bristol using feedback amplifier linearisation techniques, although feedforward techniques also have potential. A block diagram of feedback amplifier linearisation is shown in figure 2. The baseband control function may take a number of forms, the simplest and most well understood being the cartesian feedback system [4] in which the control function consists of a pair of low-pass differential amplifiers. A well known alternative is adaptive predistortion [5] in which the feedback path is sampled (not necessarily continuously) and the gain and phase response of the amplifier assessed across its operating range. The transmitted signal is predistorted to correct for any non linearities in the amplifiers characteristic.

With careful design, feedback linearisation can give excellent performance. Linearised class C amplifiers yield third order intermods better that 60dB down on PEP for a two tone test. Performance is limited by the loop gain that can be achieved within the limits of stability.

The gain and phase match of the I/Q downconverter limits the image rejection at the transmitter output. 40 dB image rejection requires a phase match of 1 degree and an amplitude match of 0.1 dB. This can be achieved over a reasonable bandwidth using RF techniques alone. To extend the bandwidth further, an automatic calibration procedure can be carried out in production to calculate correction factors for any gain and phase mismatch. In service, the baseband processing can use these to predistort the drive to the amplifier. In a transmitter fed with I and Q baseband signals, DC offsets will coincide with the pilot tone and can easily be made small enough not to dominate (offsets 45 dB down on PEP can be achieved, while the pilot is often around 10 dB down).

Linear Receivers

Weaver method zero IF receiver systems (direct conversion) have been suggested for linear transceivers [6]. They offer the advantage of minimising the component count, but require two linear mixers operating at the incoming frequency and an accurate 90 degree phase

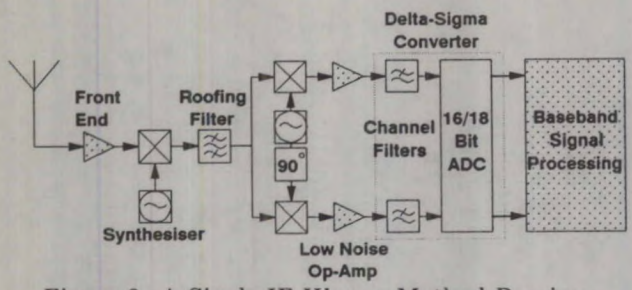


Figure 3: A Single IF Weaver Method Receiver.

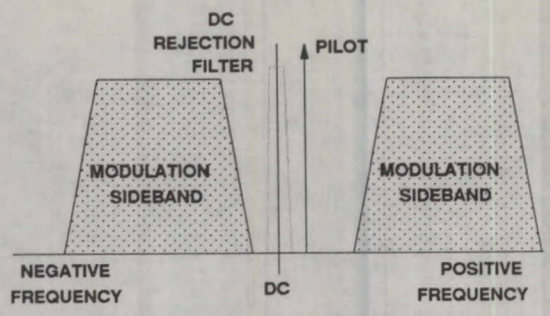


Figure 4: Placing DC in the demodulated I/Q signals.

shift network which operates across the entire range of the systems frequency synthesiser. Carrier leakage can also be problematical at high frequencies, causing large DC offsets in the demodulated I/Q signals. These problems mean that direct conversion will probably be unsuitable for use at mobile satellite frequencies for the immediate future.

Figure 3 shows a block diagram for a single IF linear receiver. This alleviates many of the problems by performing the I/Q generation at a low, fixed IF. If a suitable frequency is chosen, then mixers with good linearity, low carrier leakage and low cost can be obtained. The quadrature phase shift network need only operate at a single frequency, making accuracy and stability easy to achieve. Unfortunately, a roofing filter will now be required to remove the image frequencies which, after the first mixing stage, do not lie within the wanted channel. Such filters can be obtained cheaply if a prudent IF is chosen.

Carrier leakage around the final mixers results in a DC offset in the resulting I/Q baseband signals. Such offsets can be removed by filtering, but they can alter slightly with incoming signal strength. When operating in weak signal areas, this variation means that DC rejection must be achieved up to 5 to 10 Hz. Such filtering removes some of the information content of the received pilot. It may be prudent to frequency lock the receiver so that DC falls in the gap between the signaling modulation and the pilot. DC rejection can then be achieved with no information loss and removal of the frequency offset is straightforward (figure 4).

The availability of modern delta-sigma analogue to digital converters with 18 bit resolution means that conversion into the digital domain can be achieved with dynamic range well in excess of that required for mobile satellite receivers. The signals can then be filtered, amplified and mixed as required with no fear of instability. (High resolution ADCs are becoming cost effective due to the economies of scale in the digital audio market). The limitation in receiver performance; adjacent channel performance, sensitivity, selectivity and overload point. are now largely dependant on the linearity of the RF pre-amplifier and subsequent mixer stages. Field testing of a 900MHz single IF linear receiver is currently in progress at Bristol and prototype hardware for 1.7 GHz operation is under development.

Subjective Comparison of Modulation Systems

Before comparing some modulation formats which may be used in a mobile satellite system (whether linear of not) it must be stated that voice assessment based on MOS is only truly valid if all modulation candidates are tested by the same panel of listeners under the same test conditions. (Clearly, the source tape must be common for all systems under trial and the playback to assessment panels, customers etc, must be randomised to prevent learning of the information content). As the results for this assessment are drawn from a number of sources [7,8,9,10], this condition is not fully satisfied, however, the subjective performance of analogue linear modulation with respect to a high quality 9.6kbps codec has been validated at the laboratories at Bristol University and thus provide a good reference point for scaling the other results. It is on this basis that the results are presented and discussed, giving, it is felt, and accurate picture of modulation and coder performance with current technology. Further work is in progress at Bristol to complete the analysis for a full range of coder types and modulation techniques.

Modulation Systems

All modulation systems will be compared on the basis of average signal power with respect to noise power density (dB/Hz) required to maintain a given voice quality (MOS). Three digital modulation schemes have been studied, namely:

- 25 % raised cosine filtered 16-QAM
  (coherent detection).
  Bandwidth efficiency = 3.2 bps/Hz
  BER of 10^{-3} at an $\frac{E_b}{N_0}$ of 13.5 dB
  BER of 10^{-2} at an $\frac{E_b}{N_0}$ of 11.5 dB

- 25 % raised cosine filtered QPSK
  (coherent detection).
  Bandwidth efficiency = 1.6 bps/Hz
  BER of 10^{-3} at an $\frac{E_b}{N_0}$ of 9 dB
  BER of 10^{-2} at an $\frac{E_b}{N_0}$ of 7 dB

- 100 % raised cosine filtered trellis coded 8-PSK
  Bandwidth efficiency = 1 bps/Hz
  BER of 10^{-3} at an $\frac{E_b}{N_0}$ of 5.8 dB
  BER of 10^{-2} at an $\frac{E_b}{N_0}$ of 3.8 dB

All three systems assume a 2 dB implementation loss.

Three vocoder rates are investigated of 9.6kbps, 4.8kbps, and 2.4kbps. Channel bandwidth for the digital systems is calculated as the vocoder rate divided by bandwidth efficiency plus 300 Hz guard band to allow for Doppler spread.

The carrier to noise ratio, $\frac{C}{N_0}$ is calculated as $\frac{E_b}{N_0} \times 10\log(\text{vocoder rate})$

The fourth system is an analogue form of linear modulation based on TTIB. The $\frac{E_b}{N_0}$ values for this system are measured as the average received signal power when the speaker is active (including pilot tone) divided by the received noise power density. It should be noted that this figure is somewhat pessimistic as it does not take account of periods of voice inactivity. Clearly, the average received power would be lower if voice inactivity were accounted for resulting in lower power consumption in both the mobile and the satellite. For the equivalent power saving in a digital voice system, a robust voice activity detection is required.

The channel bandwidth for the analogue system is 3.7 kHz, consisting of a 300 to 3000 Hz voice bandwidth, a 600 Hz TTIB notch and a 400 Hz guard band.

Figure 5 shows the results for a static AWGN. As would be expected, the analogue system out-performs the digital systems at high signal strengths, the latter exhibiting an upper threshold due to vocoder algorithm limitations which degrade as the coder rate is reduced. As the signal strength falls, certain digital systems out-perform the analogue scheme. For example, TC-8PSK with a 9.6 kbps coder is much better than ACSSB at 45 dB/Hz. The improvement in performance, however, is achieved at the price of bandwidth expansion. The TC-8PSK system with a 9.6 kbps coder requires a channel bandwidth of approximately 9.9 kHz, compared with the 3.7 kHz channel bandwidth for ACSSB.

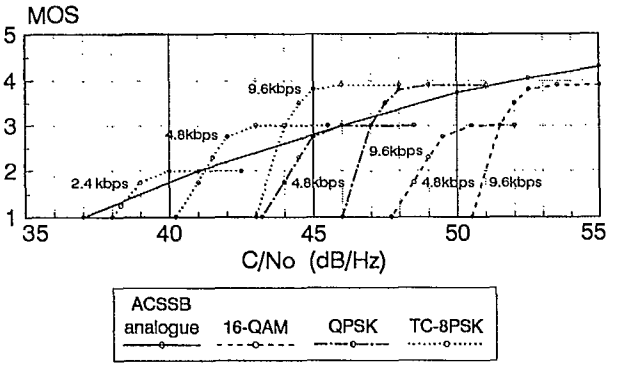


Figure 5: Subjective Performance of all Systems.

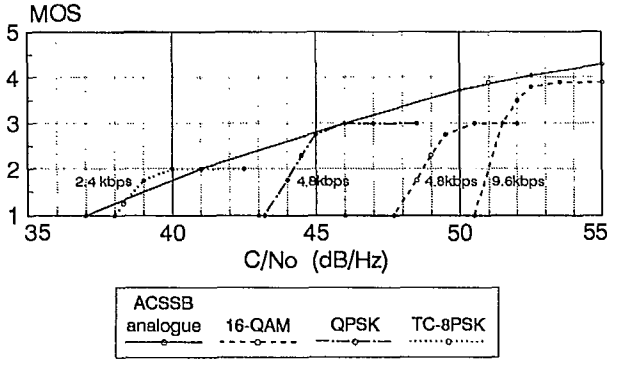


Figure 6: Subjective Performance of Systems with Channel Bandwidth below 4 kHz.

Taking into account the narrow band requirement for mobile satellite operation, figure 6 highlights the performance of those systems which require a channel bandwidth of less than 4 kHz. It is now clear that ACSSB performs as well as any of the digital competitors, except for a 2.4 kbps coder with TC-8PSK at very low signal strengths. It is well recognised that 2.4 kbps coders do not, at present, provide sufficient quality of service for a public telephony mobile satellite service. If services with a MOS below 2.5 are neglected, then the choice of modulation system becomes that of figure 7. It is now clear that ACSSB performs at least as well as present generation modulation and voice coding technologies. At low signal strengths the amplitude companding maintains a usable voice channel without recourse to bandwidth expansion.

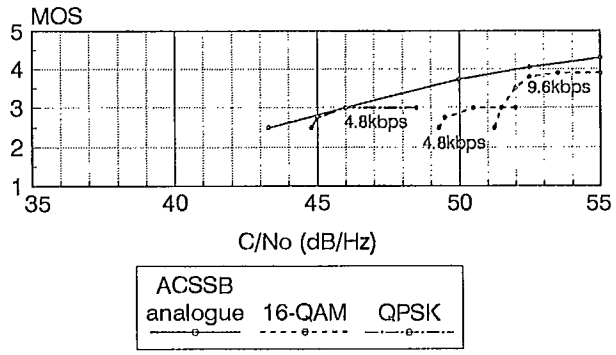


Figure 7: Subjective Performance of Systems with Channel Bandwidth below 4 kHz and MOS above 2.5.

Performance in Fading

Taking an operating point of 51 dB/Hz as a typical link budget for the M-SAT system, the 9.6 kbps 16 - QAM system will not operate satisfactorily, 4.8 kbps 16 - QAM will allow a fade margin of 2 dB, 4.8 kbps QPSK will give a fade margin of 6 dB, and ACSSB a fade margin of 7 - 8 dB.

Conclusions

With the application of feedback amplifier linearisation techniques, along with recent advances in device technology (particularly ADC, DAC and DSP technology) a universal linear transceiver can be realised. This could convey voice traffic via an analogue or a digital modulation format. Thus a smooth and cost effective transition between signaling formats could be achieved as and when superior speech vocoders become available for the mobile satellite service.

Based on some recent subjective performance measurements of candidate modulation schemes and speech coders for mobile satellite, it is concluded that digital vocoder technology, with a suitable data modulation format can now approach the performance af an analogue linear modulation system in a given power/bandwidth budget. Speech quality, channel occupancy, fade margin and power consumption, however, are all superior in the analogue implementation.

References

[1] A. Bateman, "Feedforward Transparent Tone In Band: Its Implementation and Applications," *to be published, IEEE Transactions on Vehicular Technology,*

[2] A. Bateman, D. Haines and R.J. Wilkinson, "Direct Conversion Linear Transceiver," *5th International Conference on Mobile Radio and Personal Communications*, December 11-14 1989, United Kingdom.

[3] J.P. McGeehan and A. Bateman "Theoretical and Experimental Investigation of Feedforward Signal Regeneration as a means of Combating Multipath Propagation Effects in Pilot-Based Mobile Radio Systems," *IEEE Transactions on Vehicular Technology*, VT-32, 1983, pp.106-120

[4] A. Bateman, R.J. Wilkinson and J.D. Marvill "The Application of Digital Signal Processing to Transmitter Linearisation," *8th European Conference on Electrotechniques*, June 13-17, 1988, Sweden.

[5] Y. Nagata "Linear Amplification Technique for Digital Mobile Communications," *39th IEEE Vehicular Technology Conference*, May 1-3, 1989, USA.

[6] A. Bateman and D.M. Haines "Direct Conversion Transceiver Design for Compact Low-Cost Portable Mobile Radio Terminals." *39th IEEE Vehicular Technology Conference*, May 1-3, 1989, USA.

[7] "AEEC Subjective Test Report - Selection of Voice Codec for the Aeronautical Satellite Service," *British Telecom R&T Memorandum No. RT53/001/88.*

[8] "On the Subjective Voice Quality Test of ACSSB Modulators and Digital Vocoders," *INMARSAT Internal Document*, 1989.

[9] "Relative Performance of Linear Modulation (analogue) vs a 9.6kbps digital vocoder," *Centre for Communications Research, Bristol University, Internal Report*, May 1989.

[10] "Trials Report on the Subjective Performance of Modulation Techniques for Mobile Satellite," *Australia Telecom Internal Report*, 1989.

MSAT Voice Modulation Considerations

Dan Bossler
Telesat Mobile, Inc.
P.O. Box 7800
Ottawa, Ontario
K1L 8E4
Phone: 613-756-5601
FAX: 613-746-2277

ABSTRACT

The challenge for MSAT voice services is to provide near-toll quality voice to the user, while minimizing the power and bandwidth resources of the satellite.

The options for MSAT voice can be put into one of two groups- Analog and Digital. Analog, nominally narrowband single sideband techniques, have a demonstrated robustness to the fading and shadowing environment. Digital techniques, a combination of low-rate vocoders and bandwidth efficient modems, show the promise of enhanced fidelity, as well as easier networking to the emerging digital world.

The problems and trade-offs to designers are many, especially in the digital case. Processor speed vs. cost and MET power requirements, bits to be allocated to the voice encoding vs. channel coding, bandwidth efficiency vs. power efficiency etc. While the list looks daunting, in fact an acceptable solution is well within our technology.

This paper sets the objectives that the MSAT voice service must meet, and reviews the options that we see now and in the future.

OUTLINE

This paper first addresses voice service on MSAT from the point of view of users to gain an appreciation of the requirements, then from a system designer's viewpoint as he/she tackles the requirements.

USER'S VIEW

The most eager adoptors of MSAT will be people who, because of where they live or work, receive very poor service from existing terrestrial systems, if they get service at all. "Poor service" typically means low availability (does not work everywhere), unreliable (doesn't work when he wants it), low quality (cannot understand or recognise the talker).

MSAT is a natural service to satisfy the availability requirement, since it will provide coverage of all of North America. MSAT will also provide a very reliable link, compared to HF communications which are the only available in many areas. Although the link will suffer when shadowed with trees, the user has some control since he can move to a clear location. Unfortunately, he cannot change the atmosphere to create a better HF channel.

Voice quality is the most nebulous, hardest-to-get-a-grip-on issue facing the design of the voice service, because there are so many variables that affect quality, and the degree to which it affects quality is different for each person. Furthermore, some customers will accept any quality as long as the voice is intelligible, while others are more discriminating. Most users will ask for "near-toll quality".

SYSTEM'S VIEW

The system designer looks first at the user needs, then factors in space and ground segment costs, the propagation channel, modulation, available technologies, coding and voice

encoding; then attempts to come up with the most cost effective solution. All the above factors interact with each other, making the solution difficult. Given that we cannot objectively quantify the interaction makes finding an optimal solution nearly impossible.

Defining the problem

The goal of any non, non-profit company is to make an acceptable return on investment. TMI's business plan calls for providing a high quality service at a modest user cost.

The cost to TMI is dependent on the cost of network (satellite and earth stations) and the number of channels, which in turn is dependant on the total resources of the satellite (bandwidth and power) and the resources required per channel. But a factor in the per channel power and bandwidth is the antenna gain and the quality the channel is to provide. Generally speaking, a more complex ground segment will reduce the resource requirements, while increasing the quality will use more resources.

In addition to these technical trade-offs, we must also take into account the behaviour of the users. Customers will naturally pay more for higher quality. But how much more? The "trick" is to find the area where the value to customer is greater than the service provider's cost by a sufficient margin to ensure a reasonable profit. TMI believe this margin does not lie in providing a "low budget" service.

Developing solutions

The first trade-off is essentially one of mobile antenna gain vs satellite EIRP. Potential users are sensitive to the upfront cost of a MET and, therefore, one is not likely to specify a high gain (i.e. cost) antenna. The exact value of gain is still under discussion, however, a 10 to 12 dBic gain is thought to be reasonable.

In Figure 1, satellite capacity in assignable channels is related to MET antenna gain. At 10 dBic gain, the capacity of the system is about 1480 assignable channels.

More complicated trade-offs are dependent on the modulation. Therefore, we will examine potential analog and digital solutions to voice transmissions.

Analog solution. The capacity example of Figure 1 is derived assuming Amplitude Companded Single Sideband (ACSSB). ACSSB, described in [1], is an analog modulation technique that requires only 5 kHz of RF bandwidth. The demodulator makes use of a pilot tone to deal with the propagation effects. The result is a highly robust signal that can tolerate fades of 12 to 14 dB while maintaining intelligible voice.

Figure 2, adapted from the results in [2], depicts the variation of Mean Opinion Score, MOS, with carrier-to -noise density ratio, C/No. In the MOS method, speech samples are played for a large population of evaluators who are asked to rate the quality from 1-BAD to 5-EXCELLENT. The average of each person's rating is the mean opinion score.

The figure shows that quality peaks at about 51-52 dB-Hz with a rating of about 4.1 (GOOD). It also shows a graceful degradation down to a MOS of 2 (POOR) at about 42 dB-Hz. Recent improvements to ACSSB promise to shave 3 to 5 dB off the necessary power.

Because ACSSB is a linear modulation technique, a linear amplifier is required. The peak power of the signal is approximately 6 dB above the average power.

ACSSB also provides a transparent channel that allows direct operation of facsimile, telephone modems, and DTMF over-dialling. From this point of view, ACSSB is more natural for interfacing to the PSTN.

In order to save satellite power, the North American MSAT system will employ voice activation. Figure 3 shows the effect of various levels of voice activation on satellite capacity. The residual pilot level is level of the pilot, in dB, during speech pauses, relative to the power in the signal during speech (including pilot). During speech, approximately half the power of the signal is in the pilot. Thus a pilot level of -3 dBc, indicates no voice activation and the capacity drops by over 40% compared to total voice activation. Even at pilot levels 12 dB down, the capacity drops by about 8%.

Although these calculations were performed assuming ACSSB, digital voice would

experience the similar problems if the carrier remained on during speech pauses. The difference is that higher levels of voice activation should be easier to implement with ACSSB than with digital transmissions. This is another point in favour of ACSSB.

Digital solution. In the discussion of ACSSB above, we saw that while ACSSB is very robust in a faded environment, it requires 51 - 52 dB-Hz to achieve good quality in a static environment. Thus ACSSB is not the best choice in uses where fading and shadowing are relatively light.

The trade-off between digital voice quality and satellite resources is not so easy. In a digital system, resources of the satellite are used to pass bits between users. (Generally speaking, the more bits/sec that the satellite passes, the higher the quality. However, quality does not improve indefinitely, but there is a point beyond which increasing the rate does not buy better quality. What increasing the bit rate can buy is reduced complexity.)

Conventional transmission of digital voice exhibits a threshold effect. When the C/N ratio drops below a certain point, the quality quickly deteriorates. Thus in fading and shadowing, ordinary digital systems are weak. However, digital voice has some properties that can be exploited to over come this threshold to some extent. Error control coding can be applied to specific bits that are most sensitive to errors. This can raise threshold BER performance from about 10^{-3} to 10^{-2} with a minimum of coding overhead. Furthermore, techniques such as frame repeat, or frame substitution can be used to fill in for badly corrupted frames. These techniques can be used to overcome bursts errors of 60 - 90 ms in length.

The issue of modulation schemes for the mobile environment is one which has had considerable study. INMARSAT and AUSSAT have both specified QPSK. The Jet Propulsion Laboratory have developed an 8-PSK Trellis Coded Modulation (TCM) modem as part of their MSAT-X work for NASA [6], and the Communications Research Centre are presently developing a 16-QAM TCM modem [7].

QPSK is relatively easy to implement, and Offset-QPSK, INMARSAT's choice, has a near-constant envelope which reduces the linearity requirement on the amplifier. However, compared to TCM schemes, this modulation is not bandwidth efficient. Since the first generation AUSSAT and INMARSAT satellites will be power limited, QPSK is a logical choice.

The MSAT satellites, on the other hand, will be dedicated high power satellites and may be both power and bandwidth limited. Therefore a power and bandwidth efficient modulation scheme is required.

Voice quality testing. For the testing of digital voice, ideally we would take all likely combinations of vocoders, background noise conditions, modulation, coding, propagation conditions and interleaving, measure the quality of each combination, weight each by the cost (satellite resources) and the probability of occurrence of the impairments (background acoustic noise and propagation conditions). At the end, we would have the best design for the voice service.

Testing the quality of analog voice is considerably simpler, since the voice "coding" and modulation is one process. This eliminates the all the vocoder/modulator combinations.

The real problem is measuring the quality of each combination. Objective measurements such as signal-to-noise ratio can be used. But this can't account very well for the physiological aspects of human hearing. A more accepted solution is the subjective Mean Opinion Score.

MOS is not without its problems either. It is a one-way test; this is especially significant because of the variable (from one design to another) delay of the encoding and interleaving process, on top of the fixed satellite delay. The test cannot judge how people would rate the quality in an actual conversation.

A bigger drawback is that full factorial testing of all the factors is clearly not practical. To reduce the number of tests for the digital solution, it is common to separate the modulation and channel coding from the voice encoding. The effects of the channel are therefore manifested in terms of bit errors. The task is often further subdivided by separately testing the effects of uniform errors, burst errors and background noise environments.

This is the procedure of the vocoder testing that Telecom Australia performed on behalf of INMARSAT and AUSSAT [3,4], and in in-house testing at Telesat Mobile [5].

To compare analog and digital solutions, the vocoder must be matched to an RF modem, and tests made over a propagation simulator. TMI and the Communications Research Centre intend to perform such tests when CRC develops the necessary hardware this summer.

SUMMARY

TMI and AMSC have specified flexibility in the design of the ground segment for the MSAT system. Vocoders and RF modems will continue to improve with time, and it is the objective for the system to take advantage of these improvements.

The MSAT signalling system will allow up to 16 different voice coding/modulation channel types. By taking advantage of Digital Signal Processors, we hope to minimize the amount of hardware that is necessary to support the different types. Wherever economic, DSP software will be used to replace hardware.

TMI will be monitoring closely the vocoder testing in Australia, and the TCM work at CRC. In the fall of 1990, we hope to be able to test the integration of vocoder(s) and the TCM modem over a propagation simulator. We will also test improved versions of ACSSB, which will also be available.

It is expected that comparison of the two will show that ACSSB is superior in moderate-to-heavy shadowing that many customers may experience in Canada due to the low elevation angles. Other customers, who are not using the system in harsh environments may be able to take advantage of a higher fidelity digital system.

The table below summarizes TMI's current view of the analog and digital solutions. Each system has its advantages over the other in different areas, as indicated by check marks. For this reason, we plan to support both ACSSB and a digital scheme when MSAT is operational in 1994.

REFERENCES

1. Lodge, J and Boudreau, D. The implementation and performance of narrowband modulation techniques for mobile satellite applications. In *Proceedings of ICC'86* P44.6.1

2. Sydor, J and Chalmers, B. Co- and adjacent channel performance of ACSSB. *DOC/CRC Technical Memorandum SS #4/87.* June 1987.

3. Wong, S. Inmarsat codec evaluation process, ICEP working document ICEP/WK/1. 18-8-89

4. Wilkinson, M. Voice codec test and evaluation procedure, VCTEP/2. Telecom Research Lab. Australia Telecommunication Corporation. January, 1990.

5. Bossler, D. Vocoder testing report. Telesat Mobile Inc. March, 1990.

6. Jedrey, T., Lay, N., and Rafferty, W. An 8-PSK TCM modem for MSAT-X. In *Proceeding of the Mobile Satellite Conference.* May 3-5, 1988. Page 311-316

7. Moher, M and Lodge, J. TCMP- a modulation and coding strategy for rician fading channels. In *JSAC Vol. 7, No. 9* December, 1989

Comparison of Analog and Digital Solutions

	Bandwidth	Interface to POTS	Interface to Digital Networks	Unfaded Quality	Heavily Shadowed Quality	Potential for Improvement
Analog	✓	✓			✓	
Digital			✓	✓		✓✓

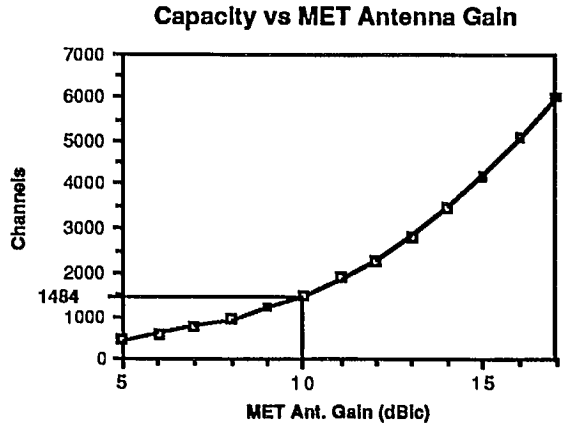


Figure 1

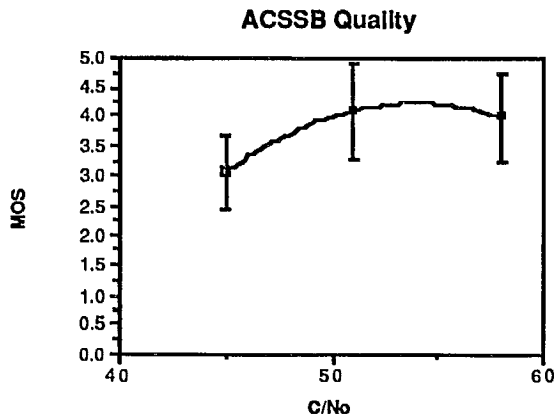


Figure 2. ACSSB Quality (C/No is in dB-Hz)

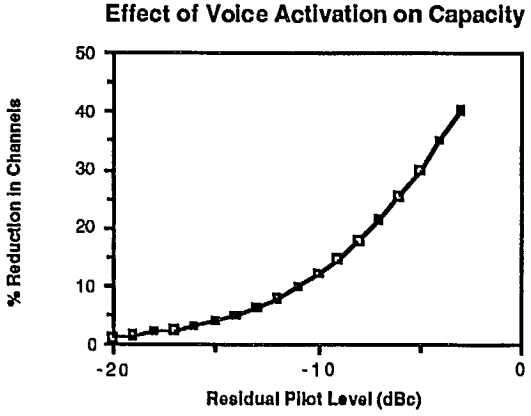


Figure 3

A New Coherent Demodulation Technique for Land-Mobile Satellite Communications

Shousei Yoshida, Hideho Tomita

C&C Systems Research Labs., NEC Corp.
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki-city, 213, JAPAN
PHONE +81-44-856-2255 FAX +81-44-856-2061

ABSTRACT

This paper describes an advanced coherent demodulation technique for land-mobile satellite (LMS) communications. The proposed technique features a combined narrow/wide band dual open loop carrier phase estimator, which is effectively able to compensate the fast carrier phase fluctuation by fading without sacrificing a phase slip rate. Also included is the realization of quick carrier and clock re-acquisition after shadowing by taking open loop structure. Its bit error rate (BER) performance is superior to that of existing detection schemes, showing a BER of 1×10^{-2} at 6.3 dB E_b/N_0 over the Rician channel with 10 dB C/M and 200 Hz (1/16 modulation rate) fading pitch f_d for QPSK. The proposed scheme consists of a fast response carrier recovery and a quick bit timing recovery with an interporation. An experimental terminal modem has been developed to evaluate its performance at fading conditions. The results are quite satisfactory, giving prospects for future LMS applications.

1. INTRODUCTION

Satellite based land-mobile communications have received increasing attention over recent few years because of its extended service to the rural area, where terrestrial mobile systems cannot be incorporated from an economical viewpoint. In LMS systems, the use of power and spectrally efficient transmission scheme is intensely requested to overcome the large space loss, and to conform the limited available frequency band. To meet such requirements, coherent detection combined with forward error correction (FEC) is in common practice. However, in the LMS environment, there exist two serious degrading factors for coherent detection. One is multipath fading, which ranges to 100- 200 Hz at maximum. Since a 4800 bps data rate is sufficient for a current digital voice transmission technique, good phase coherence cannot be counted on over long bit period, thus it results in serious BER degradation. The other is frequent received signal interruption by obstacles (i.e. shadowing). A phase lock loop (PLL) used conventionally in carrier and clock synchronization is very vulnerable to such environments, and often loses its track, taking a lot of time for the recovery. Although the phase tracking speed and/or recovery time can be improved by making loop bandwidth wide, it induces frequent phase slips as a penalty. It is quite difficult to find a good compromise between these two conflicting factors. Authors propose a new coherent demodulation technique, which is almost able to remove PLL's deficiency.

2. NEWLY PROPOSED COHERENT DEMODULATION SCHEME

2.1 Carrier Recovery by Dual Open Loop Phase Estimator

In order to resolve aforementioned problems, authors propose a new carrier phase estimator, which consists of two different bandwidth estimators to compensate the fast phase fluctuation by fading without increasing phase slips. An LMS channel is modeled by the mixture of a strong direct-path signal component and a scattered weak multipath signal component, that is Rician. Moreover the received signal is contami-

nated with thermal noise, as depicted in Fig. 1.

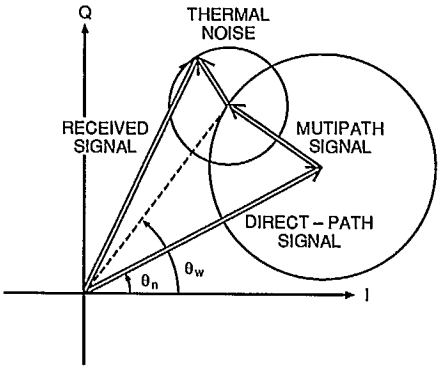


Fig. 1 Received signal vector over LMS channel

Figure 2 shows a block diagram of the proposed carrier recovery by a dual open loop phase estimator. This block diagram features open loop architecture used throughout to make free from acquisition problems. Basic open loop phase estimator's performance is previously analyzed in [1],[2]. A long term frequency offset is assumed to be corrected by an automatic frequency control (AFC) circuit. As a first operation, phase modulation is removed by an Mth-law operator, after that, the signal is routed both to narrow and wide band phase estimators. The narrow band estimator's bandwidth is designed around some tens Hz so as not to induce phase slips by thermal noise, and the estimator's filter is constructed by a finite impulse response (FIR) filter because of its well-defined delay time. As a result, almost all the multipath components as well as thermal noise are smoothed out, because the filter's bandwidth is much narrower than assumed maximum fading pitch 200Hz, and only a direct-path component is derived. A narrow band estimator's output $\theta_n(n)$, n as an integer, is expressed by

$$\theta_n(n) = \arg[\sum_{k=n-(N1-1)/2}^{n+(N1-1)/2} C_n(n-k)\mathbf{S}_M(k)]$$
$$= \arg[\frac{1}{N_1}\sum_{k=n-(N1-1)/2}^{n+(N1-1)/2} \mathbf{S}_M(k)]$$
$$(C_n(i) = 1/N_1 \quad \text{for} \quad {}^{\forall}i), \qquad (1)$$

where $\mathbf{S}_M(n)$ is a modulation removed signal, and $C_n(i)$ and N_1 are tap coefficients and tap number of the narrow band estimator respectively.

The fast phase fluctuation due to the multipath is estimated by the wide band estimator. The wide band estimator takes the same configuration as the narrow band estimator except its bandwidth. A wide band estimator's output $\theta_w(n)$ is also expressed by

$$\theta_w(n) = \arg[\sum_{k=n-(N2-1)/2}^{n+(N2-1)/2} C_w(n-k)\mathbf{S}_M(k)]$$
$$= \arg[\frac{1}{N_2}\sum_{k=n-(N2-1)/2}^{n+(N2-1)/2} \mathbf{S}_M(k)]$$
$$(C_w(i) = 1/N_2 \quad \text{for} \quad {}^{\forall}i), \qquad (2)$$

where $C_w(i)$ and N_2 are tap coefficients and tap number of the wide band estimator respectively. An impulse response of the respective estimators is designed rectangular to reduce a computational load.

A carrier phase is estimated by combining the narrow and wide band estimator's outputs so that the wide band estimator's output cannot induce carrier phase slips. First, the narrow band estimator's output is subtracted from the wide band estimator's output, and the result is divided by modulation phase M, thus a fast phase component scattered by the multipath is obtained within $\pm\pi/M$ radian, for example $\pm\pi/4$ radian for QPSK. The narrow band estimator's output is also divided by M, and then its phase range is extended to $\pm\pi$ radian by quadrant correction. A recovered carrier phase is obtained by adding the estimated fast phase component to the quadrant corrected slow changing phase. Thus, a recovered carrier phase $\theta_r(n)$ is calculated by using the following equation,

$$\theta_r(n) = \mathrm{mod}\{\frac{1}{M}\theta_n(n) + \frac{2\pi}{M}i(n)$$
$$+ \frac{1}{M}\mathrm{mod}[\theta_w(n) - \theta_n(n), 2\pi], 2\pi\}, \quad (3)$$

where $i(n)$ denotes $\mathrm{mod}\{i(n-1) - \mathrm{sign}[\theta_n(n) - \theta_n(n-1), 1], M\}$ for $|\,\theta_n(n) - \theta_n(n-1)\,| > \pi/M$ and $i(n-1)$ otherwise. In this carrier recovery scheme, a carrier phase slip rate depends only on the narrow band estimator's performance. Finally, the recovered phase is converted to a complex form, and a received signal is coherently detected by the complex carrier. Delay time T_1 and

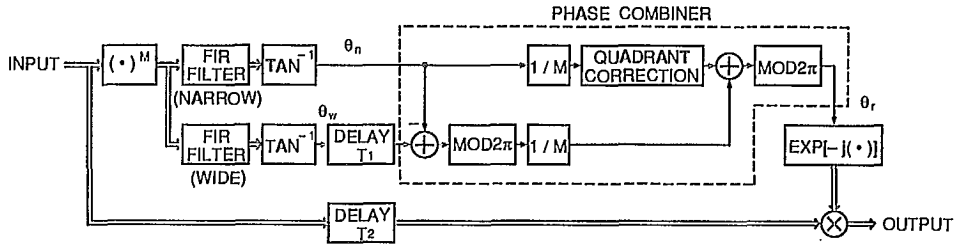


Fig. 2 Proposed carrier recovery block diagram

T_2 in Fig. 2 are determined using baud interval T_b as follows,

$$T_1 = \frac{N_1 - N_2}{2} T_b, \quad T_2 = \frac{N_1 - 1}{2} T_b. \qquad (4)$$

The BER performance for several detection schemes including the proposed scheme is evaluated by computer simulation over the Rician fading channel with carrier-to-multipath ratio (Rician factor) C/M of 10 dB and fading pitch f_d of 1/16 modulation rate for QPSK. These results are shown in Fig. 3.

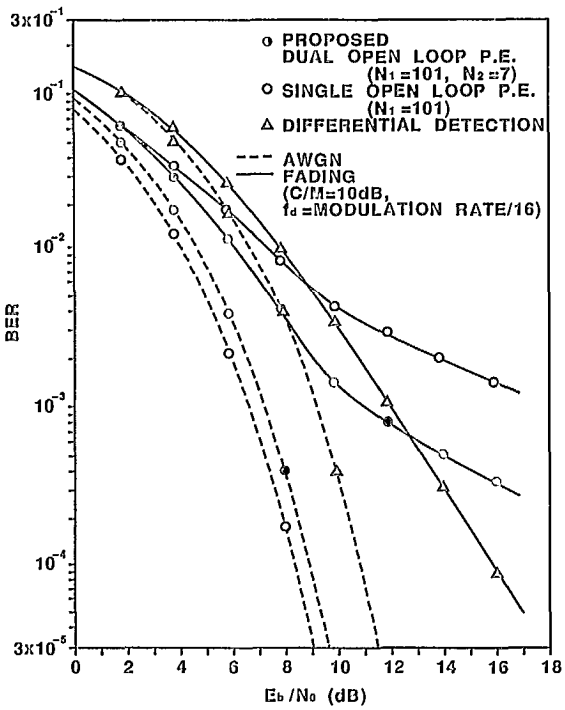


Fig. 3 BER characteristic by simulation

The proposed scheme shows the E_b/N_0 improvement of around 1-2 dB compared with conventional coherent detection using a single open loop phase estimator with narrow bandwidth, and the E_b/N_0 improvement of around 2 dB over non-coherent differential detection, at the operating 1×10^{-2}-5×10^{-3} BER point derived from the high coding rate FEC.

Although static BER performance of the proposed scheme is degraded by the wide band estimator, its amount is relatively small from the conventional scheme, and in an LMS environment its degradation becomes negligible compared with that by fading.

2.2 Quasi-Open Loop Bit Timing Recovery

In bit synchronization, quick recovery from frequent shadowing is another requested characteristic in an LMS environment. To meet such a requirement, authors propose a new bit timing recovery scheme, which is able to make free from acquisition problems by taking quasi-open loop structure. The block diagram is shown in Fig. 4. After root Nyquist filtering, a signal envelope is detected, and then phase correlation between the signal envelope and a complex sinusoidal local reference is calculated. A ROM table is referred to generate the sinusoidal reference at the input sample timing. T_s as sample interval, phase correlation $\mathrm{S}_c(nT_s)$ is expressed by

$$\mathrm{S}_c(nT_s) = V_e(nT_s)\exp(-j2\pi n/N), \qquad (5)$$

where $V_e(nT_s)$ is a signal envelope and N is a sample number per baud interval. The correlated output is fed to a low-pass filter to smooth out the phase jitter, and $\tan^{-1}$ operation of the filter's output is performed. The result represents a bit timing error against the local reference. In a case when a slight frequency difference exists between

a modulation clock and the local reference, the slow bit timing change is observed. To track the bit timing continuously, a wide band first-order PLL is introduced. A $2\pi/N$ radian phase step size is sufficient for controlling a digital VCO, because its request is to specify the nearest sample to real bit timing from the input N samples. Hence, the PLL's loop gain K is represented as

$$K = 1/N. \tag{6}$$

The loop gain K is designed large enough that the PLL's bandwidth is sufficiently wide compared with the open loop phase estimator's bandwidth (low-pass filter's bandwidth). Thus, the whole response is solely determined by the low-pass filter's characteristics, showing this loop virtually open.

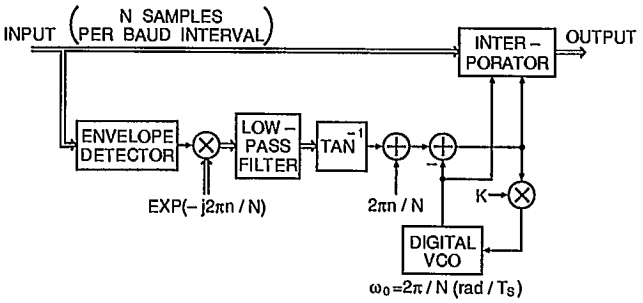


Fig. 4 Bit timing recovery block diagram

A real sampled value at bit timing is obtained by an interporation. The second-order's Lagrange's formula is selected as the interpolation scheme to facilitate the calculation. An interpolated signal $\mathbf{S}_i(r)$ is calculated by using the following equation,

$$\begin{aligned}\mathbf{S}_i(r) &= \frac{r^2-r}{2}\mathbf{S}(-T_s) - (r^2-1)\mathbf{S}(0) + \frac{r^2+r}{2}\mathbf{S}(T_s) \\ &= C_1(r)\mathbf{S}(-T_s) + C_0(r)\mathbf{S}(0) + C_{-1}(r)\mathbf{S}(T_s) \\ &= \sum_{n=-1}^{1} C_{-n}(r)\mathbf{S}(nT_s) \quad (|\, r \,| \leq \frac{1}{2}),\end{aligned} \tag{7}$$

where $\mathbf{S}(0)$ is the nearest sample to real bit timing, $\mathbf{S}(-T_s)$ and $\mathbf{S}(T_s)$ are its adjacent samples, and $r = T_i/T_s$ is normalized interporation timing. This calculation can be implemented by an FIR filter with time variant three tap coefficients $C_i(r)$. These coefficients are written in a ROM table in advance.

3. MODEM IMPLEMENTATION

An experimental terminal modem is developed to evaluate our proposed scheme. From a viewpoint of the spectral efficiency, 5 kHz channel spacing is strongly requested in LMS systems. Therefore, the modulation rate is determined to be 3.2 kbaud for QPSK modulation using band-limit filters with 0.5 roll-off factor. Thus, it enables redundancy for FEC with 3/4 coding rate. Table 1 shows the modem specifications summary.

Items	Specifications
Modulation scheme	3.2 kbaud QPSK
Band-limit filter	Raised cosine roll-off (roll-off factor=0.5)
Operation mode	SCPC continuous mode
Channel spacing	5 kHz
Carrier deviation	Less than ± 3 kHz
A/D sampling rate	606.667 kHz (455 kHzIF × 4/3)
A/D quantization	8 bits
Sample number N	6 samples per baud interval
FEC	K=7, R=3/4, 6/7 Punctured convolutional coding with Viterbi decoding
Voice codec	4.8 kbps multi-pulse coding

Table 1 Modem specifications summary

Figure 5 shows the demodulator block diagram. A 455 kHz intermediate frequency (IF) signal is fed to a single A/D converter, and directly sampled by 4/3×IF clock. Quadrature detection is performed all by digital circuit using these samples, and the digitized orthogonal signal is appropriately decimated. After that, demodulation is performed by three NEC's μPD77230 digital signal processors (DSP) except Viterbi decoding using a dedicated LSI chip, and the following tasks are assigned for each processor.

1) Root Nyquist filtering (FIR filter)
2) Bit timing recovery, AFC and carrier recovery
3) Frame synchronization, de-interleaving, Viterbi decoder interface and Voice decoder interface

Since an open loop phase estimator is very sensitive to a carrier frequency offset, an AFC circuit [3] is necessitated to eliminate this offset before the carrier phase estimation. The AFC circuit is composed of slow and fast tracking parts. The slow tracking AFC part is requested so as to calibrate a local oscillator at all times stably, and the fast tracking AFC part is for the purpose

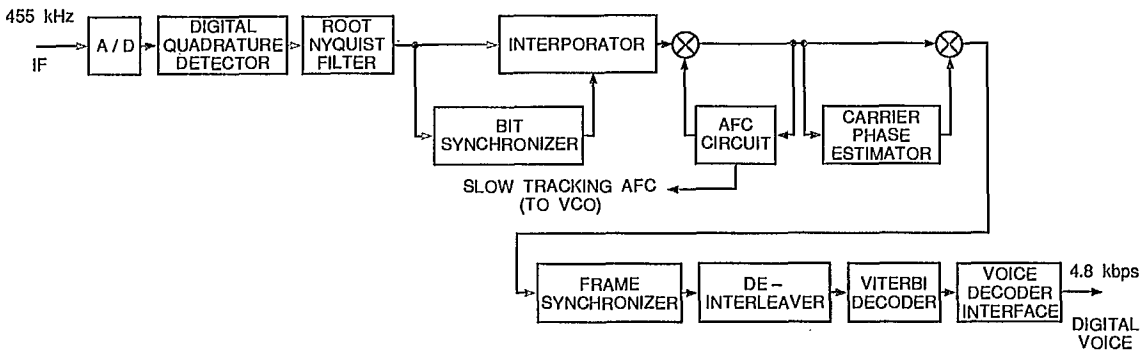


Fig. 5 Demodulator block diagram

of compensating the fast frequency fluctuation by Doppler.

A demodulated signal after Viterbi decoding is sent to a 4.8 kbps voice decoder and converted to a voice signal in it.

4. PERFORMANCE OBTAINED

The modem has been tested using a radio channel simulator. The bit error rate and carrier phase slip rate are measured in the presence of AWGN and Rician fading. The BER performance is shown in Fig. 6 with Rician factor C/M of 10 dB and fading pitch f_d of 200 Hz (1/16 modulation rate). The deviation from computer simulation is well within 1 dB at the operating BER point. To prove the effectiveness of our proposed scheme over the conventional scheme using a single open loop estimator with narrow bandwidth, the performance is measured removing a wide band estimator. As was expected from simulation, the proposed scheme showed the E_b/N_0 improvement of around 1-2 dB over the conventional scheme at the operating 1×10^{-2}-5×10^{-3} BER point. And, a low carrier phase slip rate of less than 5×10^{-2} times/sec is realized at the operating E_b/N_0 of 6 dB with severe Rician fading, as seen in Fig. 7. The demodulated scattering diagrams for the proposed and conventional schemes are depicted in Fig. 8, which shows an appearance that the fast phase fluctuation by fading is effectively compensated by a dual open loop phase estimator.

5. CONCLUSION

A new carrier and bit timing recovery scheme is presented, which remarkably augments the demodulator's performance in an LMS environment. Furthermore, an actual experimental terminal modem has been manufactured to verify the effectiveness of the proposed idea. The results are quite satisfactory, achieving good bit error rate performance and stable synchronization. This scheme is applicable for different modulation schemes.

ACKNOWLEDGMENTS

Authors wish to thank Mr. Furuya, Dr. Hirosaki and other members of C&C Systems Research Labs., NEC Corporation, for their helpful discussions.

REFERENCES

[1] A.J.Viterbi,A.M.Viterbi, "Nonlinear Estimation of PSK-Modulated Carrier Phase with Application to Burst Digital Transmission," IEEE Trans.Inform.Theory, vol.IT-29, pp.543-551, July 1983

[2] W.Hagmann,J.Habermann, "On the phase Error Distribution of an Open Loop Phase Estimator," in Proc.ICC'88, June 1988, pp.1031-1037

[3] F.D.Natali, "AFC Tracking Algorithms," IEEE Trans.Commun., vol.COM-32, pp.935-947, Aug. 1984

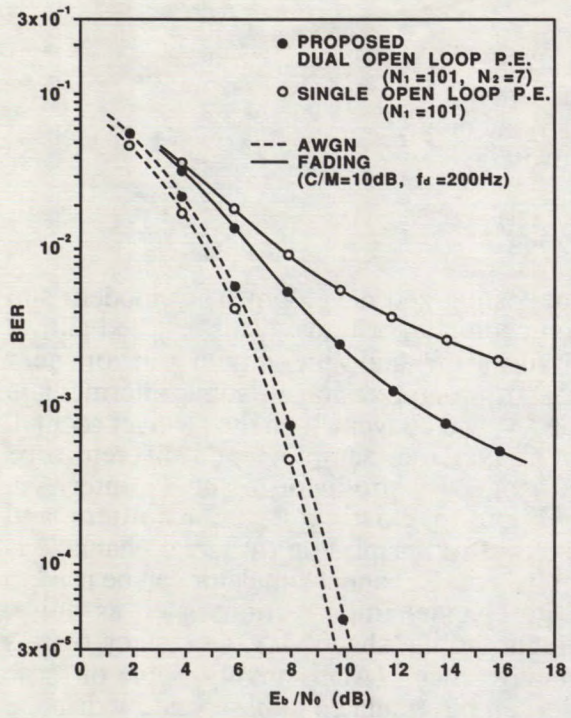


Fig. 6 BER characteristic

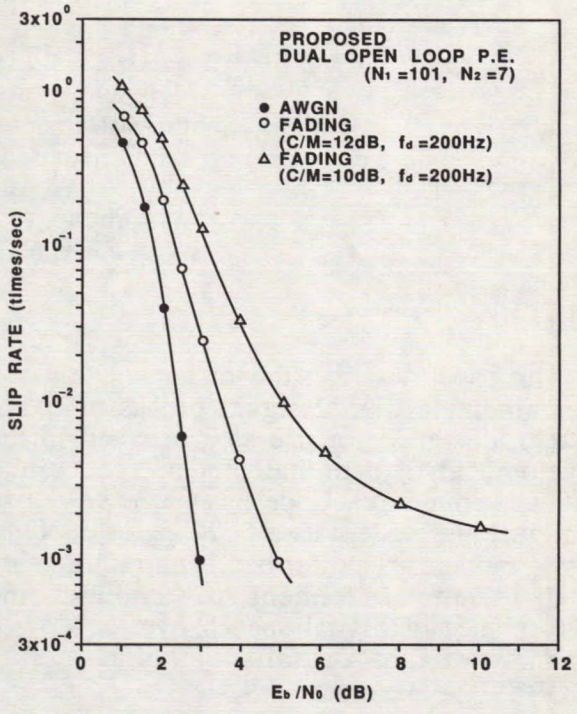


Fig. 7 Carrier slip rate characteristic

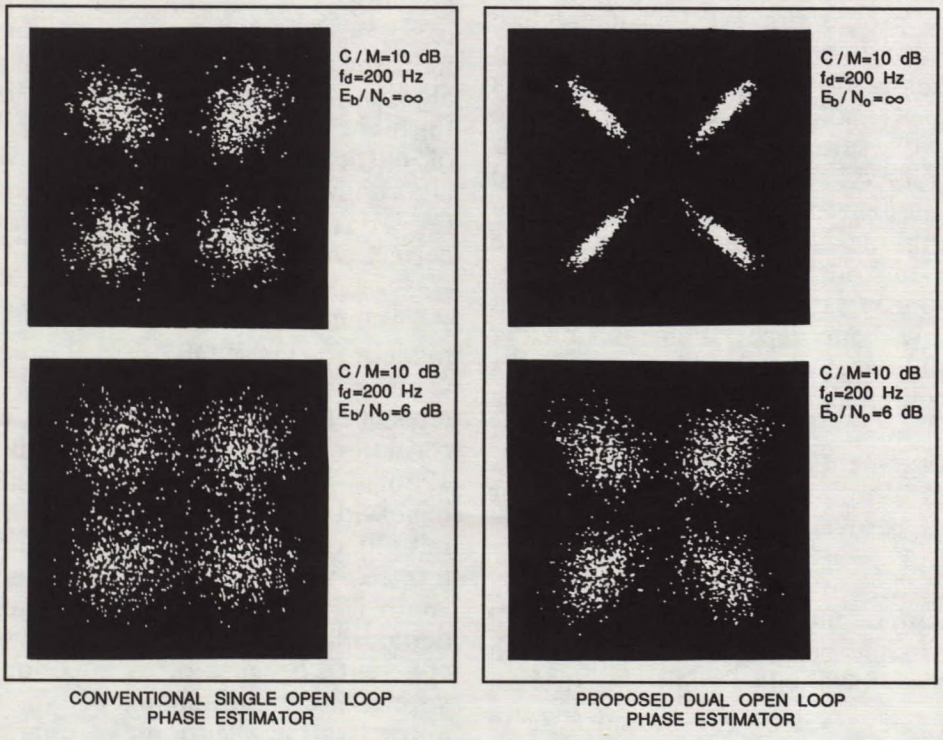


Fig. 8 Demodulated scattering diagrams

Modem Design for a MOBILESAT Terminal

M. Rice, M. J. Miller, W. G. Cowley, D. Rowe
Digital Communications Group,
South Australian Institute of Technology,
P.O. Box 1, Ingle Farm 5098,
Australia.
Phone: (+618) 343 3310
FAX: (+618) 260 4724

ABSTRACT

This paper describes the implementation of a programmable digital signal processor based system, designed for use as a test bed in the development of a digital modem, codec and channel simulator. Code has been written to configure the system as a 5600bps or 6600bps QPSK modem. The test bed is currently being used in an experiment to evaluate the performance of digital speech over shadowed channels in the Australian mobile satellite (MOBILESAT) project.

1 INTRODUCTION

The availability of increasingly powerful digital signal processors (DSPs) has made possible the implementation of what were considered computationally complex tasks, at a reasonable cost and with relatively simple circuitry. Two examples of applications for DSPs are digital modems and codecs for digital speech. In the case of modems, traditionally analogue techniques have been used for the signal processing operations such as mixing, filtering, synchronisation etc. Digital signal processing allows better repeatability and a wider range of possible operations. In addition, DSPs provide a programmable development tool. This greatly facilitates the iterative process of algorithm implementation and development. (Alternatives include lengthy computer simulations and hardware modifications).

This paper describes the development of a programmable modem test bed based on DSP techniques. It has been designed for use in development of terminals for the Australian MOBILESAT system[1]. The intention is to provide a system which can be configured in various modes to provide a powerful DSP facility for the testing and development of modem and speech coding algorithms. This test bed allows generation of analogue signals at or near baseband, up-conversion to some intermediate frequency, and conversely in the receiver section, down conversion, sampling and discrete time processing. By provision of an IF interface, connection to existing RF transmitters and receivers for transmission over real channels is possible. An IF channel simulator can be used to recreate channel imperfections such as noise, multipath fading, shadowing, frequency offsets and interference. (Alternatively, some of these effects can be simulated at baseband by discrete time processes within the DSP. This provides a more controllable and programmable approach to testing, although less of a test for the entire system.) The main intended use for the modem test bed system is to provide a real time test facility for fairly low rate digital modem algorithms. It also allows a fast simulation tool for higher speed modems, and by the pipelining of extra DSPs and suitable break down of algorithms, testing of extra system components (eg digital speech codecs, forward error correction).

The immediate goal has been to provide a test modem for the MOBILESAT speech channel[2]. Current specifications allow for two channel data rates of 5600bps and 6600bps, each with a frame format corresponding to a frame length of 120ms. With the aim of using 5kHz or 7.5kHz bandwidth per speech channel, QPSK was initially chosen as the modulation scheme, with transmit root-Nyquist filtering and 40% excess bandwidth. Coherent detection was the chosen demodulation type (rather than differentially coherent), because of its superior performance with a direct path signal. Both frame formats allows for the insertion of a unique word at the start of each 120ms frame, and this is used to resolve the 4 state phase ambiguity.

The proposed MOBILESAT speech channel has an unfaded C/N_0 of 48 dB-Hz providing E_b/N_0 figures of 10.5 dB and 9.8 dB at 5600bps and 6600bps respectively. The high elevation angle for Australian geo-stationary satellites implies that the main problem associated with this channel at L-band will be shadowing caused by trees and buildings rather than multipath effects. This, together with channel measurements indicate that the propagation can be approximately modelled in terms of an 'on-off channel'. It is therefore desirable for the demodulator to have the ability to acquire signal synchronisation rapidly and 'free wheel' during periods of severe signal attenuation. A channel simulator has been developed that can be driven by amplitude and phase data to simulate shadowing effects at IF. This has successfully been used to replay recorded channel measurements during modem tests.

Section 2 gives a brief description of the hardware contained in the test bed at the time of the AUSSAT speech codec tests. Section 3 discusses the signal processing algorithms performed by the DSP. These include filtering, timing recovery, phase recovery and frame processing. Some results for the modem performance are given in section 4, and finally, in section 5, current work and future proposals for the modem test bed are outlined.

2 MODEM IMPLEMENTATION

The basic structure of the modem test bed is shown in figure 1. On the modulator (transmit) side of the system, the source supplies data via a serial bit stream, at 5600bps or 6600bps. Data is read by the modulator processor as pairs of bits (di-bits), which are formatted into frames and modulated to form a set of discrete baseband complex samples. The samples are converted to inphase and quadrature continuous analogue signal components using 14 bit D/As and analogue reconstruction filters. The signal is then mixed up to an IF of 71.15MHz using a quadrature mixer. The signal at IF can then be passed on to an RF transmitter, or passed through a channel simulator.

On the demodulator side, the converse functions take place. The received IF signal is mixed down to nominal baseband using a quadrature mixer, and then the inphase and quadrature components are low-pass filtered to eliminate aliasing effects, sampled, and the discrete quantised samples are passed on to the DSP. The DSP performs the receive filtering, all the synchronisation tasks (described in the next section) and reconstructs the data stream by unscrambling and differentially decoding, if necessary. The decoded data is sent to interface circuitry which converts it to a serial bit stream compatible with the AUSSAT specification.

Another important module in the system is the microcontroller. This provides a boot facility for the DSPs, control of DSP processes, and allows communication between the modem and a monitoring/controlling PC. This latter function allows downloading of programs to the DSPs, plus real-time monitoring and modification of DSP memory. With the aid of some graphical display software written for the PC, modem parameters can be observed in a convenient form, and recorded if required. For example, scatter diagrams, eye diagrams, frequency and timing tracking, etc can be shown in near real-time. This provides an excellent tool for the assessment of algorithm performance. Test programs have been written for the modem to measure bit error and frame error rates which can be monitored by a PC. Algorithm parameters can be modified during processing and subsequent improvements in performance (or otherwise) noted.

3 DIGITAL SIGNAL PROCESSING ALGORITHMS

The signal processing taking place in the modem test bed is illustrated in figure 2. Brief descriptions of the most important discrete time processes handled by the DSP are given below. The DSP32C was chosen as the most suitable DSP device because of its computational power (12.5MFlops), floating point processing capability, low cost, and the availability of a C compiler. These last two factors help considerably with software development, although careful optimisation of code is necessary to achieve the performance limits of the processor.

3.1 Filtering

Ideally, both the transmit and receive filters have identical root-Nyquist responses with an excess bandwidth of 40% for the purposes of limiting bandwidth, rejecting out of band noise and interference, and minimising inter-symbol interference. A finite impulse response filter is used to approximate the required filter response. This is achieved by a direct software

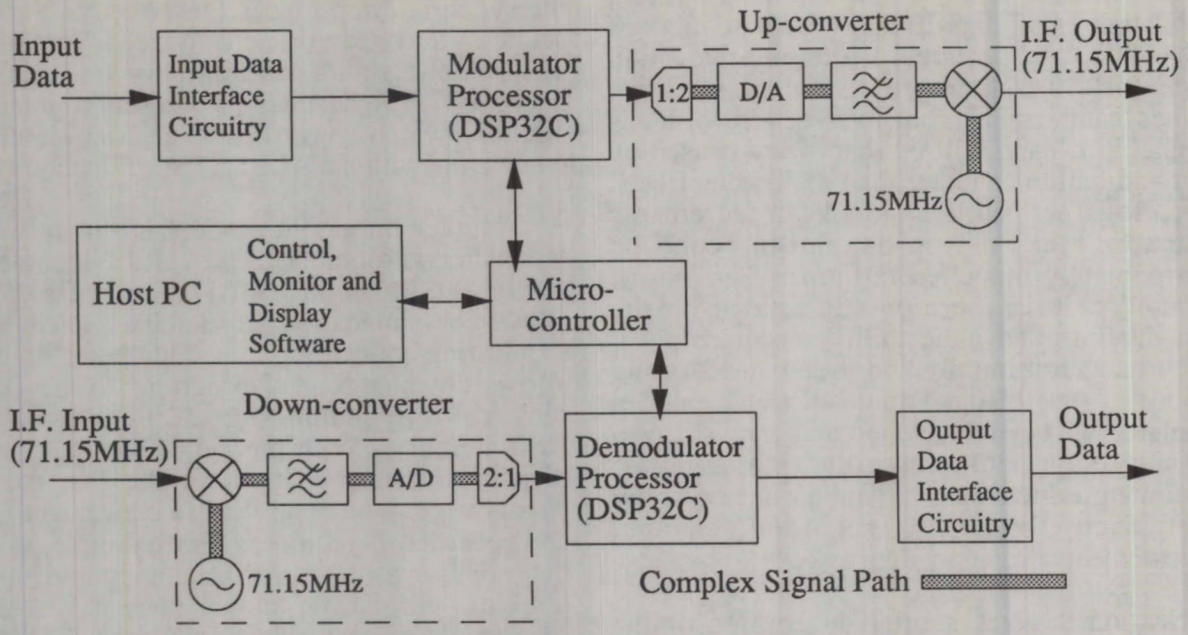


Figure 1: MOBILESAT Modem Test Bed

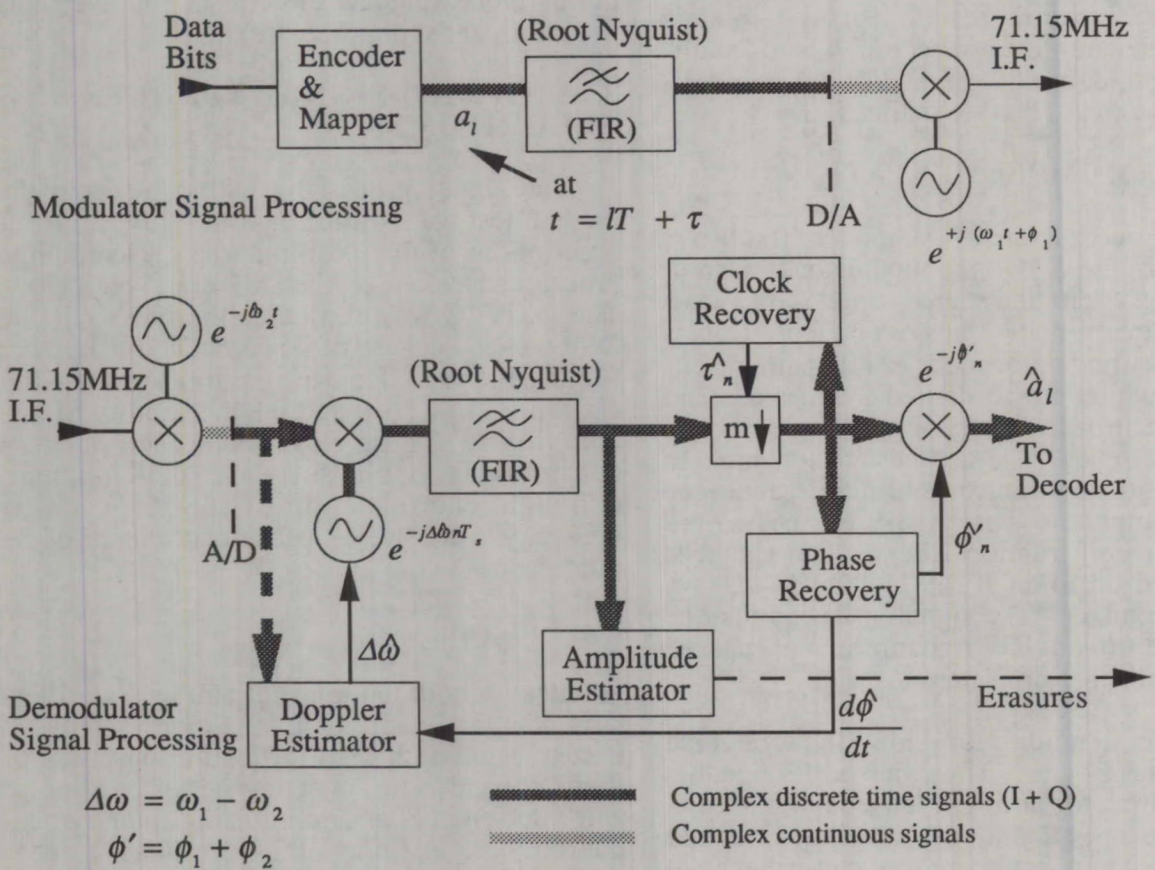


Figure 2: Modem Signal Processing

implementation of the filter equation

$$\overline{r_k} = \sum_{i=0}^{N-1} r_{k-i}\, h_i \qquad (1)$$

where h_i is a sample from the desired impulse response, r_k a received signal sample, and $\overline{r_k}$ the filtered signal. In the present design 4 complex-valued signal samples are taken per symbol period. After digital receive filtering, the symbol midpoint sample (y_n) can be interpolated and then used to update symbol timing and phase offset estimation. The number of taps was chosen to be 36 with a sampling rate of 4 per symbol.

3.2 Symbol Timing Recovery

The current symbol timing recovery algorithm is a one-sample-per-symbol decision-directed method that has been previously described by Cowley[3]. The technique represents a gradient solution to the maximum likelihood estimate of the timing offset between the sampling instant and the true symbol mid-point. The gradient solution leads to an iterative equation for estimating the timing offset, τ_n, associated with the nth received symbol:

$$\tau_n = \tau_{n-1} + \beta\, \mathrm{Re}\{ y'_n \overline{a}_{n-1} - y'_{n-1} \overline{a}_n \} \qquad (2)$$

where overbar indicates complex conjugation, a_i are the estimated symbol values, and β is a small positive adaption constant which controls the size of the symbol timing updates. The values y'_n are the complex signal samples after phase correction. To make the symbol timing independent of phase recovery, an estimate of the phase offset can be made from just the two samples y_n and y_{n-1}. A simple way of doing this is described in reference 4, in which the phase estimate is quantised by $\pi/4$. Although the technique is suboptimum, the degradation is small at the unfaded E_b/N_0 ratio in use.

3.3 Phase Recovery

The phase recovery algorithm used in the current modem was described by Viterbi[4] and provides rapid and robust non-decision-directed feed forward phase recovery. If y_n is represented in polar form by r_n and θ_n then $z_n = r_n e^{j 4\theta_n}$, an unbiased estimate of the phase offset, ϕ, associated with that symbol can be made by estimating the mean values of $\mathrm{Re}\{z_n\}$ and $\mathrm{Im}\{z_n\}$ over $2N+1$ symbols and calculating the corresponding phase angle. The n^{th} estimate of ϕ is given by:

$$\phi_n = \frac{1}{4} \arctan \left(\frac{\sum_{i=-N}^{N} \mathrm{Im}\{ z_{n+i} \}}{\sum_{i=-N}^{N} \mathrm{Re}\{ z_{n+i} \}} \right) \qquad (3)$$

The mean phase offset is used to rotate the nth symbol sample y_n, thus achieving coherent demodulation.

Due to the phase quadrupling, a 4 phase ambiguity is left which can be resolved by observing the phase of the unique word (transmitted at the beginning of every frame).

3.4 Frequency Tracking

The frequency tracking is based on forming an estimate of the frequency offset of the signal mixed down to nominal baseband, and using this to frequency shift the signal closer to true baseband. The frequency offset estimate is updated periodically by calculating the mean phase estimate change per symbol. The maximum rate of change of frequency offset would be caused by the Doppler frequency shift due to mobile movement relative to the satellite. For an automobile this would usually be considerably less than 15Hz/s. Frequency offset updates rates of one per 120ms frame have proved to be more than adequate with the current phase recovery scheme.

3.5 Initial Frequency Estimation

The initial frequency estimation is based on a method involving the differerential power measurement from two passband filters located at either end of the required signal spectrum. Should the received signal have a spectrum that is offset in frequency, the power from the output of each of the filters is unbalanced and an error term can be derived to modify the frequency offset estimate. Simulation has indicated that a frequency offset of ±200Hz can be tracked within 50 symbols at an E_b/N_0 of 0dB using this method.

3.6 Demodulator Frame Processing

The frame processing has two states of operation depending on whether frame synchronisation has occurred. If the modem does not have frame synchronisation, the initial frequency estimation is used, and on each

received symbol a check is made to see if the unique word has been received. This continues until a UW is detected, at which time the modem switches to 'in frame sync' mode. The initial frequency estimation is deactivated, and the frequency tracking algorithm is used exclusively. The phase of the received UW is used to set the phase reference to demodulate future symbols. The proceeding data bits in the frame are then unscrambled and differentially decoded, if that option is set. After receiving an entire frame, the next UW is expected. If this is received successfully then the process continues as before. Otherwise, a counter is incremented to record the number of consecutive UW failures. If this counter reaches a given threshold, the demodulator reverts to its 'out of frame sync' mode. In fact there are two counters and thresholds which are used. One corresponds to a normal signal power level and the other to a shadowed power level. Should the average signal power drop below a predetermined level, then the signal is considered to be shadowed, and a larger number of UW failures are tolerated. The counters are reset as soon as a UW is detected again.

4 MODEM PERFORMANCE

Modem performance is measured in several ways including implementation loss compared to an ideal modem and acquisition time. The bit error rate performance of the 5600bps and 6600bps modems were found to be virtually identical. At E_b/N_o of 5dB and above the implementation loss is nearly 0.5dB, and the performance drops away considerably from the ideal case below 3dB.

Parameters of the modem synchronisation may be adjusted to give the best performance for the specific channel concerned. For example, if N is reduced in (3) the implementation loss increases but the recovery time after a signal outage is reduced. Symbol timing is not affected significantly by fades as the clock accuracy is sufficient to allow timing updates to be inhibited during fades so that symbol timing can 'free wheel'.

Some tests have been performed using a channel simulator allowing the replay of measured signal levels from a real mobile satellite channel. This demonstrated the modem's ability to retain synchronisation despite severe shadowing.

5 FUTURE ACTIVITIES

Several future developments are planned for the modem both in terms of hardware and software. It appears that part of the implementation loss evident in the modem tests could be due to some imperfections in the IF stages. To eliminate some of these, conversion to a low IF signal within the DSP is planned with a double stage conversion up to 71.15MHz. This will make the quadrature mixer redundant, thus getting rid of problems associated with errors in the quadrature arm phase shift and reducing carrier feed through. Similarly, IF sampling is planned for the receiver. Software configurable modem parameters such as sampling clock rates, filter cut-offs, and local oscillator frequency control are also under consideration.

The following modifications are also being made at the moment:

- Development of $\pi/4$ QPSK modem for the MOBILESAT speech channel
- More optimum synchronisation algorithms
- Conversion to low IF signal construction and sampling
- Alternative initial frequency estimation schemes.

The programmable nature of the DSP allows many modulation, coding and synchronisation algorithms to be developed and tested. It is hoped to produce a suite of algorithms that can be called upon to provide implementation of a range of modems and processing capabilities.

6 REFERENCES

1. Dinh, K. et al, 1990. MOBILESAT - The Kangaroo Way. International Mobile Satellite Conference, Ottawa, Canada, June 1990.

2. AUSSAT Pty. Ltd., 1989. MOBILESAT System Specification. Appendix 4: Digital Speech Coding Description, AUSSAT Pty. Ltd., Sydney, Australia.

3. Cowley, W. G., 1989. Synchronisation Algorithms for Digital Modems. Proc. ASSPA 89, Adelaide, Australia, 1989.

4. Viterbi, A. J. and A. M., 1983. Nonlinear Estimation of PSK-Modulated Carrier Phase with Application to Burst Digital Transmission. IEEE Trans. Inf. Th. Vol IT-29, No. 4, pp. 543-551.

Frequency-Offset Insensitive Digital Modem Techniques

S. Dutta and **D. C. Nicholas**
Mobile Communication Satellite System
Rockwell International Corporation
400 Collins Road, NE
Cedar Rapids, Iowa 52498

ABSTRACT

Conventional DPSK systems are adversely affected by transmitter/receiver frequency offsets due to frequency reference errors and Doppler shifts. We present two DPSK modem concepts which avoid the long frequency acquisition process of conventional DPSK. One technique involves a modified demodulator for conventional DPSK signals, while the other involves making minor modifications to both the modulator and demodulator. Simulation results are provided showing performances relative to conventional DPSK.

INTRODUCTION

The desire for small low-gain antennas in mobile satellite systems results in power budgets permitting only relatively low data rates (e.g. 600 bps). Low data rates at the high carrier frequencies of satellite systems (e.g. 1.5 GHz) make DPSK demodulators very susceptible to frequency errors, which may be a significant fraction of the data rate. While in many applications the carrier can be frequency located and tracked with a phase-locked loop, for some applications, such as burst demodulation, the lock-on time is excessive.

We review first the "gold standard" – fully coherent demodulation. We then discuss conventional differential demodulation, which is often preferred because of the difficulty of establishing a sufficiently "clean" local phase reference with noisy input signals. For differential detection, we point out the important distinction between predetection and postdetection matched filtering (MF). It is shown that while predetection MF is required for optimum performance, postdetection MF can ease the frequency acquisition problem.

Two new modem techniques, both derived from DPSK, are suggested. The first utilizes conventional differential modulation and a new split delay line (SDL) technique to recover the signal, independent of carrier offset. The second (DDPSK) utilizes double differential modulation and then recovers data with a double differential demodulator using either pre- or postdetection MF.

These techniques are examined for the case of binary phase shift keying (BPSK) using a square symbol shape. This case is important in satellite applications, which tend to be power limited rather than bandwidth limited.

Conventional Binary Modem Techniques

Phase shift keyed (PSK) modulation and coherent demodulation constitute the most power efficient data transmission technique. In BPSK modulation, the transmitted carrier is shifted in phase by 0 or π radians, depending on the transmitted data. Demodulation is accomplished at the receiver by multiplication with a regenerated, phase-

locked carrier followed by baseband MF. For square modulation, the MF is an integrate and dump device. Figure 1 shows a block diagram of a BPSK demodulator.

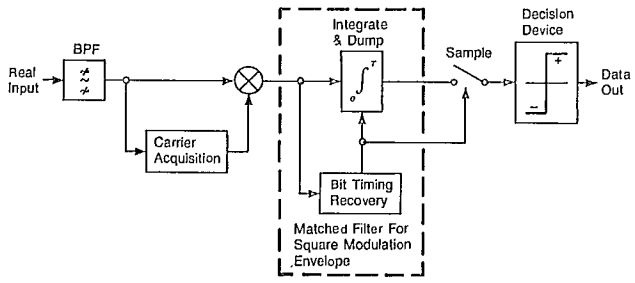


Fig. 1. Coherent BPSK demodulator

The bit error rate (BER) for BPSK in additive white gaussian noise (AWGN), using ideal coherent detection, is given by

$$BER_{BPSK} = 0.5\text{erfc}(E_b/N_0) \qquad (1)$$

Coherent demodulation requires *three* levels of parameter acquisition — carrier frequency, carrier phase and data clock. DPSK modems require only carrier frequency and data clock acquisition.

In DPSK modulation, the reference phase for any transmitted symbol is the previous symbol. Demodulation may be accomplished by differential detection, which is correlation detection using the previous symbol as the reference. A choice exists in DPSK of performing MF before or after differential detection. Predetection MF is superior because detection is a nonlinear process. As a result filtering after detection always yields greater noise power than filtering before detection. Figure 2 shows the block diagram of optimum DPSK, using predetection MF. Postdetection MF would involve moving the integrate and dump device to immediately after the detector. In the latter case, the noise bandwidth at the detector input would be established by the IF filter.

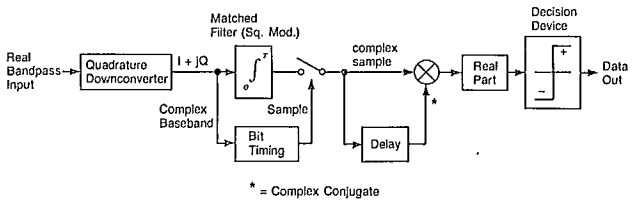


Fig. 2. Optimum DPSK demodulator (predetection MF)

The bit error rate (BER) for optimum DPSK is given by:

$$BER_{DPSK} = 0.5\exp(-E_b/N_0) \qquad (2)$$

The analytical expression for the BER of DPSK using postdetection MF has not been derived and does not appear to be available in the open literature. However, simulation results showing the degradation incurred by postdetection MF are presented in Figures 5, 7 for a IF-bandwidth/data-rate ratio of 8.

As mentioned above, DPSK does require carrier frequency acquisition. In land mobile applications, the Doppler shift due to vehicle motion can be of the order of 100 Hz, while typical data rates are 600-1200 bps. Thus, even without transmitter/receiver frequency drift, Doppler shift can be a significant source of frequency error.

Both pre- and postdetection MF DPSK suffer performance degradations due to frequency offset. However, the susceptibilities are different. In the case of postdetection MF, with a relatively wide IF bandwidth, the error will only be due to the carrier phase error, $\Delta\omega.\tau$, where $\Delta\omega$: frequency offset in radians and τ: symbol period. On the contrary, predetection MF suffers both due to the carrier phase error as well as a reduction in the MF's peak output amplitude. The latter is referred to as correlation loss and has a $\sin(\Delta\omega.\tau/2)/(\Delta\omega.\tau/2)$ dependence on $\Delta\omega$.[1] It is clear that, for frequency offsets less than 25% of the data rate, the correlation loss is a much weaker function of $\Delta\omega$ than the carrier phase error. Therefore, the

degradations of pre- and postdetection MF DPSK are similar for small frequency offsets, but predetection MF suffers greater degradation for large offsets.

NEW MODEM PROPOSALS

Basic Concept

We now present two modifications of conventional DPSK that considerably reduce the sensitivity to frequency offsets. Both techniques use a common overall demodulation philosophy – *double differential detection* – although they were developed independently. The basic concept of double differential detection is that a time-invariant carrier phase error, as defined above, may be cancelled by two cascaded differential detections involving three equidistant points in time. One application of the above concept is termed the Split Delay Line (SDL) demodulator, and applies double differential detection to a conventional DPSK waveform; hence it can be retrofitted to existing DPSK transmission systems. The other technique, Double Differential Phase Shift Keying (DDPSK), involves modifications to both the modulator and the demodulator of conventional DPSK. The advantage of DDPSK over SDL is greater power efficiency, although both techniques are less efficient than DPSK in the absence of frequency offsets. Both SDL and DDPSK can be used with pre- and postdetection MF. The details of the two techniques are given below.

Split Delay Line Demodulation

The following discussion of postdetection MF DPSK and SDL DPSK utilizes the arc tangent and manipulates only the angle. Both approaches are shown in Figure 3.

The input signal is:

$z(t) = x(t) + jy(t)$, where
$x(t) = \cos(\omega t + \phi(t))$;
$y(t) = \sin(\omega t + \phi(t))$;
$\phi(t) = 0, \pi$; and
τ = the bit period.

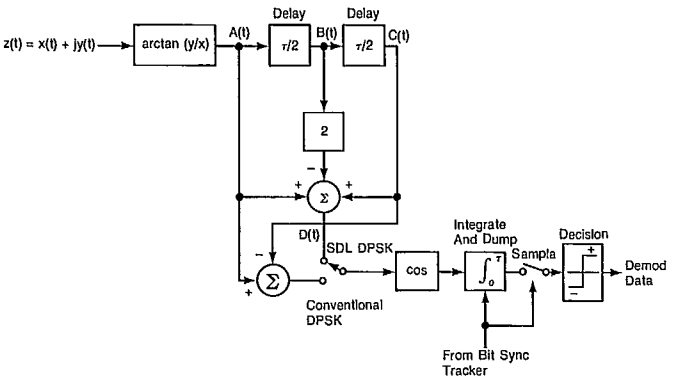


Fig. 3. DPSK and split delay line demodulator (postdetection MF)

In DPSK, we extract the angle using the arc tangent function and, using a single delay line of length τ, form the difference:

$$[\omega t + \phi(t)] - [\omega(t - \tau) + \phi(t - \tau)]$$
$$= \phi(t) - \phi(t - \tau) + \omega\tau$$

We then require that the original modulator choose $\omega\tau = 2\pi n$, where n is an integer. Then

$$\cos [\phi(t) - \phi(t - \tau) + \omega\tau] = \pm 1,$$

which is the desired signal. However, as $\omega\tau$ approaches 90 degrees, the useful output is reduced to zero.

In satellite systems we do not have sufficient control of ω, as received with Doppler shift and tuning errors, to guarantee that $\omega\tau$ will be close to $2\pi n$. This results in a carrier phase error, which is overcome in the SDL approach as shown in Figure 3, which instead forms:

$$\begin{aligned} D(t) &= \cos[\omega t + \phi(t) \\ &\quad - 2[\omega(t - \tau/2) + \phi(t - \tau/2)] \\ &\quad + \omega(t - \tau) + \phi(t - \tau)] \\ &= \cos [\phi(t) - 2\phi(t - \tau/2) + \phi(t - \tau)] \quad (3) \\ &= \pm 1, \text{ the desired signal.} \end{aligned}$$

To show the last step, we observe that the argument of the cosine is $[\phi(t) - \phi(t-\tau)]$ during the first half of the symbol and $-[\phi(t) - \phi(t-\tau)]$ over latter half. Since the cosine is the same in both cases, we obtain the desired signal.

For both DPSK and SDL, the cosine output is passed through an integrate and dump MF (for square envelope) that performs integration over the full symbol. Adequate symbol synchronization is assumed. The output of the MF is sampled for bit decisions at the end of the integration period.

Predetection MF for SDL is similar to that for DPSK (c.f. Figure 2) with the exception that four half-symbol predetection integrations (#1, 2, 3 and 4) are performed over two adjacent symbols. Then, SDL processing (equation (3)) is performed using MF outputs (#1, 2, 3) or (#2, 3, 4). Approximately 1-dB of additional performance improvment may be realized by equal-gain combining the two SDL outputs; this was implemented in the simulation model for predetection MF SDL.

The SDL receive processing expressed by equation (3) is equivalent to two cascaded levels of differential detection involving signals A(t), B(t) and C(t) in Figure 3. The first involves A(t),B(t) and B(t),C(t) , while the second differential detection involves the outputs of the first.

Double Differential Phase Shift Keying

In DDPSK, the data is encoded not on the *first* difference of the phase between *two* adjacent symbols (as in DPSK), but on the *second* difference of the phase involving *three* adjacent symbols. This concept is illustrated in Table 1. It is noteworthy that the rows for BPSK, DPSK 1st Difference, and DDPSK 2nd Difference contain identical entries. This implies identical data outputs from the respective demodulators.

Table 1. BPSK, DPSK and DDPSK modulation/demodulation

DATA IN		1	0	1	1	0	0	0	1
BPSK		π	0	π	π	0	0	0	π
DPSK	ϕ	$\phi+\pi$	$\phi+\pi$	ϕ	$\phi+\pi$	$\phi+\pi$	$\phi+\pi$	$\phi+\pi$	ϕ
1st Dif		π	0	π	π	0	0	0	π
DDPSK	ϕ	$\phi+\pi$	ϕ	ϕ	$\phi+\pi$	ϕ	$\phi+\pi$	ϕ	ϕ
1st Dif	0	π	π	0	π	π	π	π	0
2nd Dif		π	0	π	π	0	0	0	π

Figure 4 shows a block diagram of the DDPSK demodulator using postdetection MF. Note that the MF follows the *first* differential detector. The input to the second differential detector is a complex sample at the symbol rate. Hence, the additional computation load over DPSK demodulation is minimal.

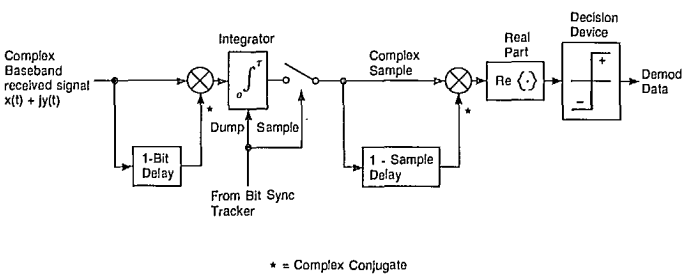


Fig. 4. Double differential PSK (DDPSK) demodulator (postdetection MF)

To implement the predetection MF version of DDPSK demodulation, one simply moves the integrate and dump device to a position prior to the first differential detector. The input to the first differential detector is then a complex sample at the symbol rate. Although accurate frequency acquisition is not required in DDPSK, symbol synchronization *is* required, as in any other data modem. DDPSK symbol synchronization is no more difficult than that of DPSK. Hence, similar techniques may be used.

MONTE CARLO SIMULATIONS

Monte Carlo, discrete time simulations were performed to examine the BER performances of the SDL and DDPSK modems relative to conventional DPSK. Results were obtained with both pre- and postdetection MF, and both with and without frequency offset.

Results

The results are shown in Figures 5-8. The following abbreviations are used for the different modem curves:

A: Theoretical coherent BPSK (no frequency offset).
B: Theoretical DPSK (no frequency offset).
C: Simulated DPSK (frequency offset as specified)
D: DDPSK (frequency offset as specified)
E: SDL (frequency offset as specified)

The curves A and B are shown in each figure for reference.

In Figure 5, the simulation results for DPSK are very close to the theoretical values, which provides a measure of assurance for the validity of the simulation model. Using BER=10^{-3} as the reference, we note that, in the absence of frequency offset, DDPSK is approximately 3.5 dB and SDL 6 dB worse than DPSK. Figure 6 shows that, in the presence of 100-Hz frequency offset (which is 20% of the data rate) the performance of DPSK suffers extreme degradation, whereas the degradations are 0.5 dB for DDPSK and immeasureably small (at these E_b/N_0 values) for SDL. At BER=10^{-3}, the power efficiency of DPSK equals that of DDPSK at a frequency offset of approximately 57 Hz, while it is equal to that of SDL at 72-Hz offset.

The predetection MF results may be explained as follows. For DPSK, the major source of degradation at 20% of the data rate is the carrier phase error, while the contribution of correlation loss is relatively quite small. The SDL and DDPSK modems suffer a small degradation because they are insensitive to the carrier phase error but, for the given offset value, *are* mildly sensitive to the correlation loss.

The major observations from the results for postdetection MF (Figures 7, 8) are as follows. In the absence of frequency offset, the degradation due to postdetection MF (relative to predetection MF) is 2.2 dB for DPSK, 0.7 dB for DDPSK and 2 dB for SDL. The 100-Hz frequency offset produces severe degradation in the case of DPSK. However, there is *no degradation* for DDPSK and SDL. The absolute frequency insensitivity of SDL and DDPSK in the postdetection MF case is due to the absence of prefiltering of the input signals (in the simulation), resulting in zero correlation loss. In practice, some prefiltering will always exist due to the bandpass IF stage, which is necessary to limit the noise power.

It should be noted that the quantitative results presented here for postdetection integration are valid only for the specific ratio of the input noise bandwidth (4 kHz) to the noise bandwidth of the postdetection integrator (500 Hz). A greater ratio between the two will yield greater degradation of postdetection MF relative to predetection MF.

CONCLUSIONS

Two new modem proposals are presented in this paper — DDPSK and SDL. Both are derivatives of binary DPSK but do not share its performance sensitivity to frequency offsets. Both techniques are less power efficient than DPSK in the absence of frequency offsets. However, in the presence of frequency offsets greater than 12% of the data rate, DPSK is worse than DDPSK, while at 15% offset, the efficiency falls below that of SDL.

Both DDPSK and SDL can be used with either pre- or postdetection MF. DDPSK is fairly insensitive to the type of MF. This means that DDPSK can be operated with a wide IF bandwidth (e.g. 4-kHz for a 500-baud data signal), with no post-IF frequency acquisition. SDL offers similar frequency insensitivity to DDPSK but is less power efficient.

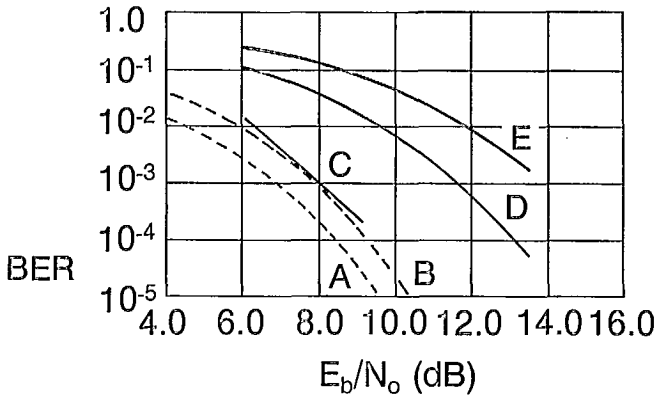


Fig. 5. Predetection matched filtering (no frequency offset)

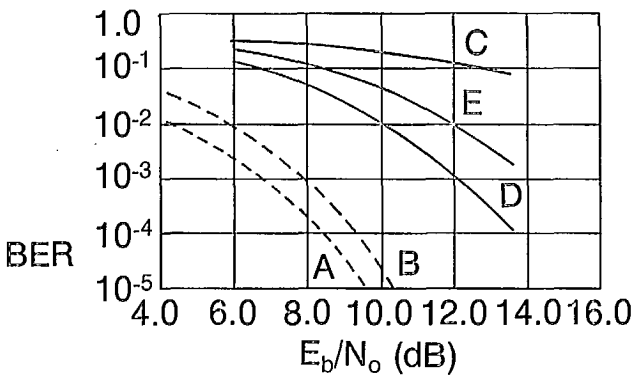


Fig. 6. Predetection matched filtering (100-Hz frequency offset)

REFERENCES

1. **Henry, J. C. III.** Dec. 1970. DPSK Versus FSK with Frequency Uncertainty. *IEEE Trans. Comm. Systems*, pp. 814-816.

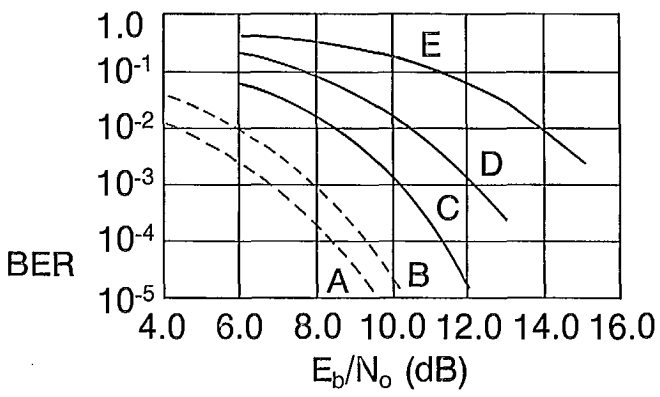


Fig. 7. Postdetection matched filtering (no frequency offset)

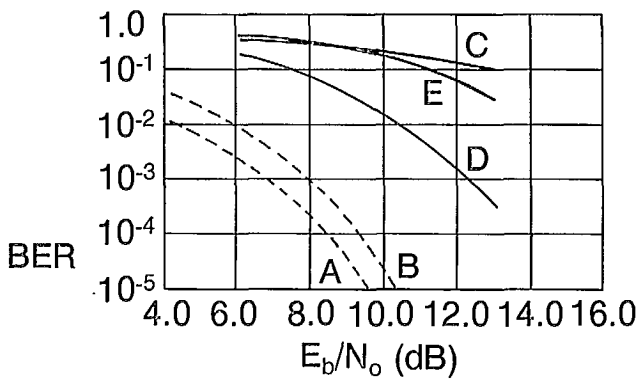


Fig. 8. Postdetection matched filtering (100-Hz frequency offset)

Comparison of FDMA and CDMA for Second Generation Land-Mobile Satellite Communications

A. Yongaçoğlu
Univ. of Ottawa
Dept. Electrical Eng.
Ottawa, Ontario

R.G. Lyons
SkyWave Electronics
300 March Road
Kanata, Ontario

B.A. Mazur
Calian Communications
300 Legget Drive
Kanata, Ontario

Abstract

CDMA and FDMA (both analog and digital) systems capacities are compared on the basis of identical link availabilities and physical propagation models. Parameters are optimized for a bandwidth-limited, multi-beam environment. For CDMA, the benefits of voice activated carriers, antenna discrimination, polarization reuse, return link power control and multipath suppression are included in the analysis. For FDMA, the advantages of bandwidth efficient modulation/coding combinations, voice activated carriers, polarization reuse, beam placement, and frequency staggering have been taken into account.

1. Introduction

In the long term, the land-mobile satellite channel will become bandwidth limited; hence it is important from the outset to identify and plan the introduction of techniques and technology offering maximum overall spectral efficiency. For this purpose, we have studied the CDMA and FDMA system capacities for a bandwidth-limited second generation land-mobile satellite communications system.

This paper presents a capacity analysis and comparison of CDMA and FDMA approaches as applicable to the voice service. Both digital and amplitude companded single sideband (ACSSB) analog modulations have been considered with FDMA. Analysis is conservative to ensure stated performance can be achieved, notwithstanding implemention uncertainties, using today's technology.

The organization of the paper is as follows. Section 2 states an objective basis for comparison and discusses the relevant system parameters. Section 3 presents the capacity analysis for each candidate system. Section 4 compares the performance of the candidate techniques. Potential capacity enhancement techniques are discussed in Section 5, and conclusions are drawn in Section 6.

2. System Parameters

The systems are compared on the basis of identical [1]

- C/N_oW_s= total received signal to noise power ratio (L-band forward or return link),
- $\theta_{3dB} = 3$ dB satellite beamwidth,
- spatial coverage area (relative to θ_{3dB}).

All candidates are assumed to operate in a circuit switched demand assignment mode, with primarily voice traffic. Access control and signalling overhead has been neglected as a differentiating factor in comparing b/s/Hz of usable capacity versus C/N_oW_s. All beams are assumed to generate and terminate the same number of erlangs of traffic. The implications of non-uniform traffic is discussed in [3].

The issue of voice coding rate is avoided by expressing capacity in b/s/Hz rather than as a number of channels/kHz. ACSSB at a specific C/N_o value is treated as a 4.8 kb/s circuit to allow comparison on the basis of b/s/Hz.

The other system parameters and the assumptions made are as follows.

Antenna Pattern

The assumed spacecraft antenna illumination pattern was that of a parabolic reflector with constant illumination, and is given by

$$G(u) = 2(J_1(u)/u)^2 \qquad (1)$$

where

$$u = \pi D \frac{\sin\theta}{\lambda}. \qquad (2)$$

In (1) and (2), $J_1(\ )$ is the Bessel function of the first kind and first order, D is the antenna diameter, λ is the carrier wavelength, and θ is the off-axis angle.

Sensitivity to illumination pattern is not significant to the conclusions stated herein [3].

Multi-Beam Geometry

A hexagonal coverage contour with two apposite sides elongated was considered, as shown in Fig. 1. This contour is illuminated by 10 beams touching at -3 dB or 13 beams touching at -2 dB contours. The minimum antenna discrimination values indicate that FDMA capacity is optimized using 10 beams with 3 frequency bands. In Fig. 1, the notation $1^{s,r}$ refers to frequency band 1, right hand polarization with frequency staggering. Similarly 2^{l} denotes frequency band 2, left hand polarization, no staggering.

For CDMA, channelization by utilizing different frequency bands in different beams, in general, was not found to provide any significant capacity increase. However, using 13 beams touching at -2 dB contours provides higher capacity than 10 beams touching at -3 dB contours. Therefore, 13 beams with a single frequency band was used in the capacity analysis.

Frequency Staggering

Frequency staggering refers to offsetting the FDMA system frequency plans in different beams by one-half channel. For both digital and ACSSB FDMA, the interference reduction achieved by frequency staggering has been included in the analysis.

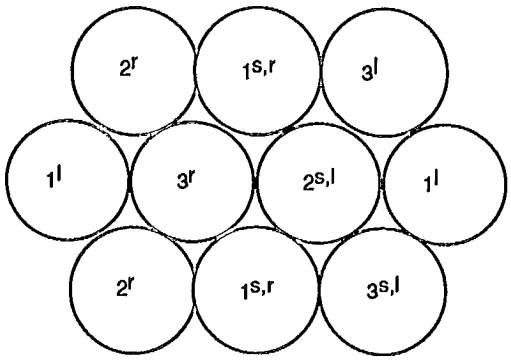


Figure 1: **10-Beam dual polarization configuration with a frequency reuse factor of 10/3.**

Power Control

Applying power control to compensate for the static losses incurred by signals travelling to and from the mobiles is a technique which can reduce interference in some cases. For the FDMA schemes, the effectivenes of power control on the worst case interference condition was found to be minimal, hence was not included in the analysis. Considering also implementation issues, power control was assumed to be applied to the return link of CDMA, but not the forward link.

Dual Polarization

Dual polarization was assumed for CDMA and digital FDMA, but not for ACSSB FDMA.

Grade of Service

The link analysis of a mobile satellite system involves a number of parameters which are subject to statistical variations. We have accounted for long-term and gross spatial variations (e.g., received power differences between the centre and edge of a beam) by assuming nominal worst-case values for such parameters. Small spatial and short-term variations were accounted for by applying a grade of service (GOS) approach.

A primary contributor to short-term variations is shadowing attenuation. We modelled shadowing as a two-state variable. Based on empirical data [2], the channel was assumed to be in the ON-state (i.e., errors are due to

AWGN and multipath but not shadowing) for 90% of the time. Thus, for system grades of service of 90% or less, defined at the voice decoder input, no particular effort to counteract shadowing need be applied. However, GOSs above 90% can be achieved by a power margin and/or interleaving. In this paper we consider a GOS of 90%.

3. Capacity Analysis[1]

3.1 CDMA

The capacity of a CDMA system, η, in units of b/s/Hz is given by

$$\eta = \frac{MR_b}{W_s} = \frac{C}{N_o W_s} \frac{(\frac{N_o'}{E_b})_{\text{req}} \frac{1}{W_n}}{V(L_s + \alpha\rho(\frac{C}{N_o W_S}))} \quad (3)$$

where
M: total number of users in the system,
R_b: information bit rate per user,
W_s: total available bandwidth,
V: voice activity factor,
N_o: thermal noise spectral density,
$(E_b/N_o')_{\text{req}}$: energy per bit to total noise density ratio required for a specific BER,
W_n: ratio of total bandwidth to the chip rate,
L_s: shadowing loss factor,
α: antenna discrimination factor,
ρ: polarization discrimination factor.

The parameters in (3) which are common to forward and return links are calculated as follows.

Coherent BPSK theoretically requires an E_b/N_o' of 6.8 dB to achieve a BER of 10^{-3}. For this we assumed the use of a rate 1/3, constraint length 6 convolutional code for which at a BER of 10^{-3} the coding gain is approximately 4 dB. Assuming an implementation margin of 2 dB, $(E_b/N_o')_{\text{req}}$ was found as 4.8 dB.

The voice activity factor, V, was assumed to be 0.4.

The bandwidth to chip rate ratio, W_n, which depends on the filtering used, was assumed to be 1.125.

[1]Due to space limitations, the interested reader is referred to [3-5] for the derivation of the following equations.

Assuming that half of the users will transmit and receive left-hand circular polarization and the other half will use right-hand circular polarization, the polarization discrimination factor is

$$\rho = \frac{1}{2}(1 + x_{\text{pol}}) \quad (4)$$

where x_{pol} is the cross polarization rejection factor. For a x_{pol} of 0.1, the polarization discrimination factor is $\rho = 0.55$.

The differences between the forward and return link capacity equations mainly lie in the antenna discrimination factors and shadowing loss factors.

In the forward direction, a mobile terminal will receive all the co-channel interference from the other users in its particular beam, plus some additional co-channel interference from all the other beams in the system at a reduced power level. For the beam geometry and the antenna radiation pattern assumed above, the forward link antenna discrimination factor was calculated as $\alpha_f = 0.255$ [3].

In the return direction, a particular beam of the satellite collects co-channel interference from all the users, weighted by the antenna gain in the direction of the co-channel users. The return link antenna discrimination factor corresponding to this situation was calculated as $\alpha_r = 0.169$. Note that these α_f and α_r values are for worst case user locations.

In CDMA, because it has a better antenna discrimination factor and shadowing loss factor, the capacity of the return link was found to be higher than the forward link capacity.

3.2 Digital FDMA

The capacity of a digital FDMA system is given by:

$$\eta = \frac{MR_b}{W_s} = \frac{C}{N_o W_s} \frac{1}{E_b/N_o'} \frac{1}{V} (1 - \frac{E_b}{N_o'} Y_o) \quad (5)$$

where N_o' is the total noise power density falling within the bandwidth of the receiver and is given by

$$N_o' = N_o + Y_o E_b. \quad (6)$$

In (6), Y_o lumps together the co-polar interference, cross-polar interference, and adjacent channel interference (ACI), i.e.,

$$Y_o = (I_{o-co} + I_{o-x} + I_{o-ACI})/E_b \qquad (7)$$

where

I_{o-co}: co-polar interference spectral density,
I_{o-x}: cross-polar interference spectral density,
I_{o-ACI}: adjacent channel interference spectral density.

In FDMA, where the number of interferers is small, treating interference as Gaussian noise and adding it to the thermal noise on a power basis is a conservative approach. However, since the target bit error rate, 10^{-3}, is not very low, this approach is acceptable.

For the assumed beam geometry and antenna pattern, and using dual polarization, the forward and return link (E_b/I_{o-co}) values were calculated as 18.1 dB and 17.1 dB respectively. Similarly the forward and return link (E_b/I_{o-x}) values were calculated as 21.5 dB and 20.5 dB respectively. Intra-system interference densities are related to E_b, because increasing the total transmitted energy proportionally increases the interference. There is little difference between forward and return link capacities at the same C/N_oW_s value.

The upper limit of η is determined by the modulation technique, FEC code rate and frequency reuse factor. For example, a rate r coded PSK signal with q phases has a maximum bandwidth efficiency of

$$\eta = \beta r \frac{\log_2 q}{1+b} \qquad (8)$$

where β is the frequency reuse factor, and b accounts for the excess bandwidth required for practical filters and the guard-bands between FDMA channels. We used $b = 0.2$ as a typical implementable value.

The ACI term is a function of the frequency spacing between carriers, transmit symbol rate (i.e., modulation and coding dependent), transmit and receive filter characteristics, and frequency uncertainty. (E_b/I_{o-ACI}) values were calculated by assuming that the transmit and receive filters have a raised-cosine characteristic with a roll-off of 0.2. Note that since ACI is mainly due to differential frequency offset, it was considered only for the return link. A worst case differential frequency uncertainty of ± 400 Hz was assumed for the return link.

The candidate modulation techniques considered for the digital FDMA were:

- coherent 16-state Trellis-Coded-8PSK (TC-8PSK),
- coherent 4PSK with rate 3/4 convolutional coding (constraint length of 6).

Both rate 3/4 coded 4PSK and TC-8PSK modulations were assumed to employ a short interleaver (with about 60 ms of total delay) to randomize most of the errors due to multipath.

The modem implementation loss was assumed to be 2 dB for 4PSK and 2.5 dB for TC-8PSK.

After AWGN, multipath and implementation loss are taken into account, the E_b/N_o' values required to obtain a BER of 10^{-3} are 7.3 dB for rate 3/4 coded 4PSK and 9.1 dB for TC-8PSK.

The maximum spectral efficiency of rate 3/4 coded 4PSK (including frequency reuse factor of 10/3) is 4.17 b/s/Hz, while TC-8PSK can achieve 5.56 b/s/Hz. Based on equation (5) and the above given E_b/N_o'values, it can be concluded that in the power limited region rate 3/4 coded 4PSK is more efficient. However, if sufficient power is available, TC-8PSK can achieve a better spectral efficiency.

3.3 ACSSB

In ACSSB, the mean opinion score (MOS) is used as a mechanism for quantifying subjective acceptability. C/N_o required per channel, $(C/N_o)_{ch}$, was obtained from experimental data [6,7] by determining the operating point from the channel parameters and the interference levels.

Considering a voice activity factor of 0.4 and a channel spacing of 5 kHz, the total C/N_oW_s can be expressed as

$$\frac{C}{N_oW_s} = 0.8 \times 10^{-4}\beta\left(\frac{C}{N_o}\right)_{ch} \qquad (9)$$

where $\beta = 10/3$ is the frequency reuse factor.

For a MOS of 2.5 and the multi-beam geometry of Fig. 1, the required $(C/N_o)_{ch}$ for the forward and return link are found as 47 dB-Hz and 48 dB-Hz respectively. These values include a margin of 1 dB for phase noise, amplifier nonlinearity and filtering.

Note that for the beam geometry assumed, ACSSB can readily achieve its bandwidth-limited capacity without utilizing dual polarization. Hence only single polarization has been considered.

4. Capacity Comparisons

In this section the forward and return link capacities of CDMA and FDMA schemes are compared. The C/N_oW_s values used in the comparisons are 10 dB and 20 dB. Table 1 presents the capacities for the 10 dB C/N_oW_s value which corresponds to a power limited condition for the candidate systems. Table 2 presents the capacities for the 20 dB C/N_oW_s value, at which the FDMA candidates have reached their bandwidth-limited capacities, and CDMA curves are approaching their asymptotes.

Note that an excess power margin (beyond the bandwidth-limited C/N_oW_s) does not increase capacity but increases the GOS to above 90%.

5. Capacity Enhancement

Some further potential capacity enhancement techniques which were not included in the analysis are briefly discussed below.

5.1 CDMA

For CDMA, using 4PSK rather than BPSK would distribute the self noise among orthogonal dimensions of the signal space. Thus 4PSK could increase the capacity by 60% and 90% at C/N_oW_s values of 10 dB and 20 dB respectively.

Bit or chip synchronization permits one to choose spreading codes which are approximately orthogonal, thus reducing the self noise.

As a consequence of above mentioned enhacements, the capacity of CDMA would significantly improve. However, this may be practical only in the forward link. Since the CDMA sys-

Table 1: **Capacity at C/N_oW_s= 10 dB.**

System	Capacity (b/s/Hz)
CDMA Forward Link	2.3
CDMA Return Link	4.1
Digital FDMA Forward Link	4.17[1]
Digital FDMA Return Link	3.7[1]
ACSSB Forward Link	2.4[2]
ACSSB Return Link	2.4[3]

1. 4PSK with rate 3/4 coding
2. 2 dB excess margin available
3. 1 dB excess margin available

Table 2: **Capacity at C/N_oW_s= 20 dB.**

System	Capacity (b/s/Hz)
CDMA Forward Link	3.1
CDMA Return Link	8.3
Digital FDMA Forward Link	5.56[1,2]
Digital FDMA Return Link	5.56[1,3]
ACSSB Forward Link	3.2[4]
ACSSB Return Link	3.2[5]

1. TC-C8PSK
2. 6.5 dB excess margin available
3. 5.2 dB excess margin available
4. 8.8 dB excess margin available
5. 7.8 dB excess margin available

tem has a lower relative capacity in the forward link, the associated complexity might be justifiable.

5.2 FDMA

Subdividing each beam into many sectors described by constant interference level boundaries and assigning location dependent frequency or polarization is expected to improve the capacity.

Capacity enhancement with ACSSB is limited to reducing spectral occupancy and channel spacing. A 20% increase (corresponding to a reduction in channel spacing from 5 kHz to 4 kHz) is achievable in the bandlimited region; resulting increased sensitivity to interference may be offset by resorting to dual polarization.

6. Conclusions

Based on the results presented above, the following conclusions can be drawn.

- For both FDMA systems, the disparity between the forward and return link capacities is negligible.
- The CDMA capacity for the forward and return links display a wide variance. In all cases the return link provides higher capacity.
- At low C/N_oW_s values (i.e., up to 10 dB) the capacity of digital and ACSSB FDMA are comparable. At higher C/N_oW_s values digital FDMA provides more capacity then its analog counterpart, while ACSSB FDMA has larger excess margins available (which yields better voice quality/higher GOS).
- In the forward link, digital FDMA provides more capacity than CDMA. The differential between digital FDMA and CDMA decreases with increasing GOS.
- In the return link, CDMA provides more capacity than digital FDMA. The differential between digital FDMA and CDMA remains fairly constant as GOS is increased.
- Both FDMA system capacities are sensitive to channelization, while CDMA is not.
- In [3] other beam contours have been considered. Capacity increases with increased coverage area, while the relative performance of the various schemes was the same as for the smaller coverage area.

It can be concluded that all systems considered possess various possibilities for enhancement, but the capacity improvements available to CDMA are probably larger and more easily attainable.

References

[1] A.J. Viterbi, "When Not to Spread Spectrum – A Sequel", IEEE Com. Magazine, Vol. 23, pp. 12 - 17, April 1985.

[2] Standard-M Technical Note TN/1/3; "Simplified Land Mobile Burst Channel Model for Standard-M", August 1989.

[3] "Study of Modulation Techniques in View of Frequency Reuse for Land-Mobile Satellite", Final Report to European Space Agency by SkyWave Electronics, April 1990.

[4] T. Murphy, "Calculation of FDMA and CDMA System Capacity for a Land Mobile Satellite System", MobileStar^{sm} Technical Report TM-C1412, Hughes Communications Inc., April 1988.

[5] K.G. Johannsen, "CDMA versus FDMA Channel Capacity in Mobile Satellite Communication", Int'l. J. of Satellite Commun., Vol. 6, pp. 29 -39, 1988.

[6] S.C. Morris, "A Subjective Evaluation of the Voice Quality of MSAT Modulation Schemes", BNR Report OCTR 85-0006, Dept. of Comm. Contract No, 01SM.36001-4-4005, Dec. 1985

[7] J.T. Sydor and B.AS. Chalmers, "Co and Adjacent Channel Performance of Amplitude Companded Single Sideband Modulation", Comm. Research Centre, Mobile Satellite Communications SS # 4/87, June 1987.

Acknowledgement

This work is based on a study performed for the European Space Agency by SkyWave Electronics Ltd. and Calian Communications Systems. The views expressed are those of the authors.

Session 15
Speech Compression

Session Chairman - *Peter Kabal,* INRS, Canada
Session Organizer - *Hisham Hassanein,* DOC

15

Performance of a Low Data Rate Speech Codec for Land-Mobile Satellite Communications

Allen Gersho
Voicecraft, Inc.
815 Volante Place
Goleta, California 93117 USA
Phone: (805) 683-2800
FAX: (805) 964-5893

Thomas C. Jedrey
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109 USA
Phone: (818) 354-5187
FAX: (818) 354-6825

ABSTRACT

In an effort to foster the development of new technologies for the emerging land-mobile satellite communications services, JPL funded tow development contracts in 1984: one to the University of California, Santa Barbara (UCSB), and the second to the Georgia Institute of Technology, to develop algorithms and real-time hardware for near-toll quality speech compression at 4800 bits per second. Both universities have developed and delivered speech codecs to JPL, and the UCSB codec has been extensively tested by JPL in a variety of experimental setups. The basic UCSB speech codec algorithms and the test results of the various experiments performed with this codec are presented in this paper.

INTRODUCTION

Over the past several years, a significant amount of research and development in the area of low bit rate (4800 bits per second) speech coding has taken place. As a result of this research, the emerging land-mobile satellite communications services will in all likelihood use these codecs to provide voice communications. In an effort to accelerate the development of these codecs, JPL funded two development contracts in 1983 with the University of California, Santa Barbara (UCSB), and the Georgia Institute of Technology to develop the necessary algorithms and real-time hardware for near toll quality speech codecs at 4800 bits per second.

As a result of these contracts, several speech codecs were developed and delivered to JPL for use in the NASA Mobile Satellite Experiment (MSAT-X) Program. These codecs have been integrated into the MSAT-X land-mobile satellite communication terminal, and the UCSB codec has been extensively tested in environments ranging from a simulated satellite (a 1000 foot tower), to a full scale land-mobile satellite channel. In addition to these tests, the UCSB codec has been independently tested by the US Department of Defense [1].

The UCSB speech codec algorithms and test results from the various experiments performed are presented in this paper. Techniques employed in the codec to mitigate the effects of channel errors will be stressed, including frame synchronization and frame repeat strategies. Results from both the aeronautical and land-mobile experiments will be presented.

USCB SPEECH CODEC

Three candidate algorithms were identified at UCSB for the MSAT-X application. Of these three, two algorithms Vector Adaptive Predictive Coding (VAPC) and Pulse Vector Excitation Coding (PVXC) were chosen for hardware implementation. The final algorithm selected for use in the MSAT-X testing was the VAPC algorithm, and all test results and further discussions in this paper are restricted to this algorithm [2].

The VAPC algorithm encodes and decodes telephony bandwidth speech sampled at 8 kHz. The resulting speech at a cumulative data rate of 64 kHz is analyzed without frame overlap at 22.5 ms intervals. As discussed below, the VAPC algorithm is based extensively on the use of vector quantization, a powerful generic technique for efficient coding of sets of parameters that characterize attributes of speech. With vector quantization, a relatively short binary word is often sufficient for accurately specifying the amplitude of a large number of parameter values, or waveform samples needed for reproducing speech sounds at the receiver.

In speech coding below 16 kb/s, one of the most successful scalar coding schemes is Adaptive Predictive Coding (APC) developed by Atal and Schroeder [3]. It is the combined power of vector quantization and APC that led to the development of VAPC.

The basic idea of APC is to first remove the redundancy in speech waveforms using adaptive linear predictors, and then to quantize the prediction residual using a scalar quantizer. In VAPC, the scalar quantizer is replaced with a vector quantizer. The motivation for using the vector quantizer was two-fold. First, although linear dependency between adjacent speech samples is essentially removed by linear prediction, adjacent samples may still have a nonlinear dependency which can be exploited by vector quantization. Secondly, vector quantization can operate at rates below one bit per sample. This is not achievable with scalar quantization, but is essential for speech coding at low bit rates.

VAPC Structure

The basic structure of an early version of VAPC, shown in Figure 1, is quite similar to that of conventional APC. In the transmitter, the redundancy due to pitch quasi-periodicity is first removed by a long delay predictor, or "pitch predictor". A short delay predictor is then used to remove the short term redundancy remaining in the pitch-prediction residual, and the final residual is quantized by a gain-adaptive vector quantizer. In the receiver, the speech waveform is reconstructed by exciting two cascaded synthesis filters with the quantized prediction residual.

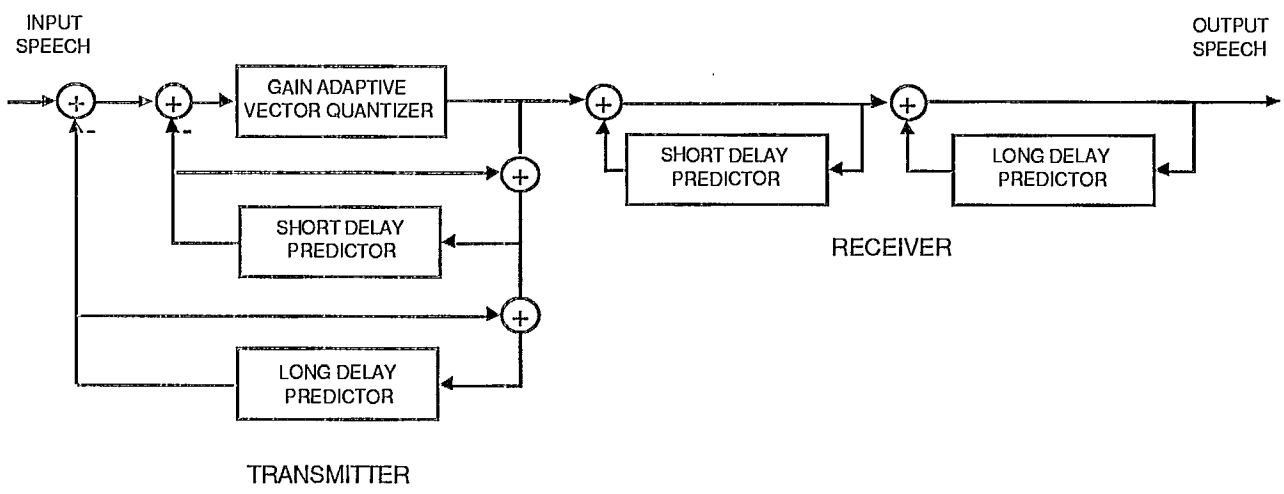


Figure 1 Basic Structure of VAPC

The structure shown in Figure 1, was modified to produce the efficient structure shown in Figure 2. To encode each vector of speech samples, the pitch prediction residual vector is generated, passed through a perceptual weighting filter, and the zero input response vectors are subtracted from it. The resulting vector is then compared with the N stored zero-state response vectors. The index of the nearest neighbor is then used to extract the corresponding vectors in the vector quantization codebook. This codevector is then used to excite the LPC synthesis filter to generate data for use in pitch prediction of the subsequent speech vectors.

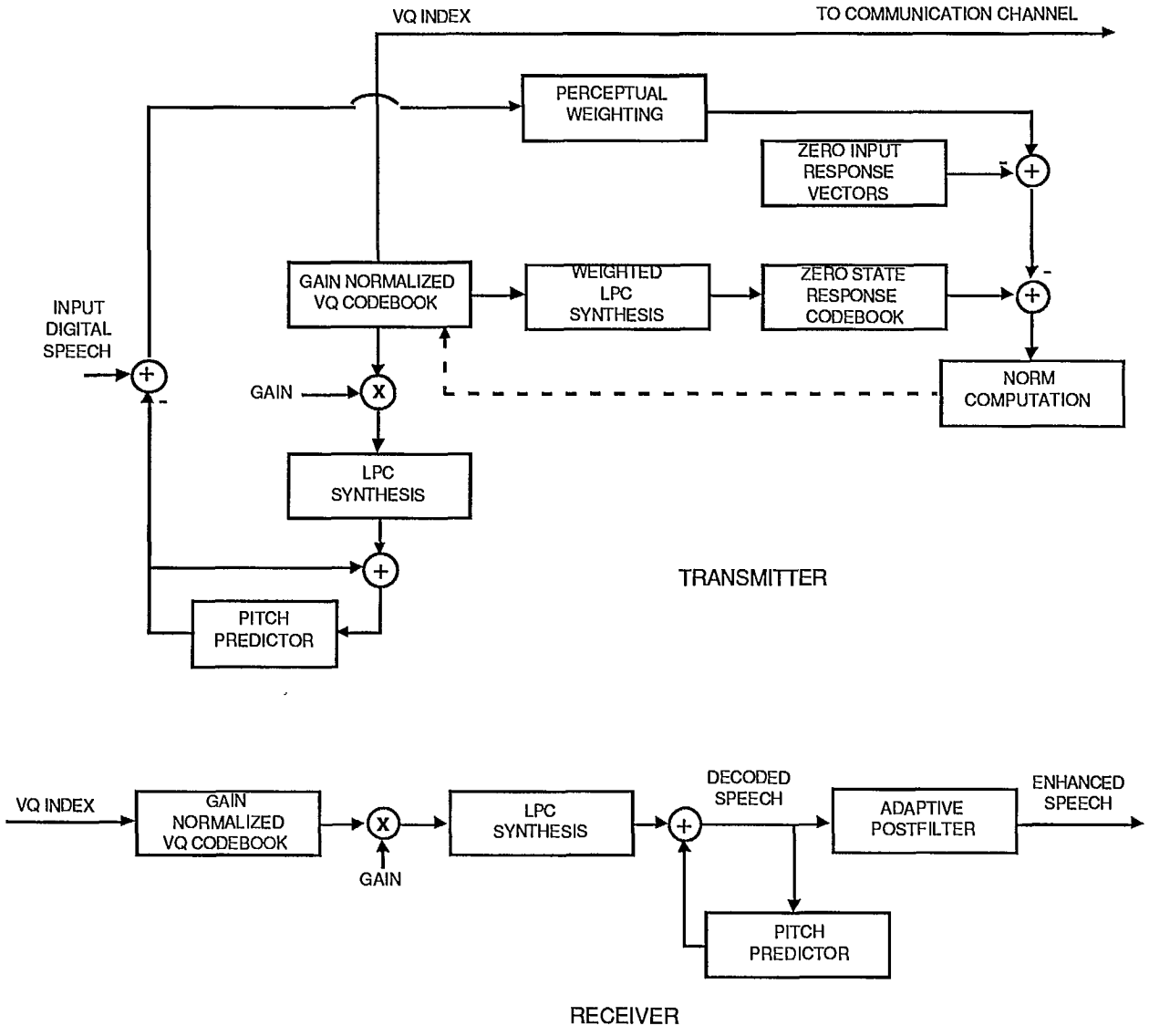


Figure 2 VAPC Transmitter and Receiver

The vector quantization codebook is designed as the gain-normalized codebook of a forward gain-adaptive vector quantizer. The normalized vector quantization codebook is fixed, while the zero-response codebook changes from speech frame to speech frame.

To further improve the perceptual quality of the coded speech, a novel adaptive post-filtering technique was developed that greatly reduces the perceived level of coding noise without introducing significant distortion in the filtered speech [2].

VAPC Channel Optimization

Once the basic algorithm was fixed, optimization of the VAPC codec for the communication channel was considered. The specified channel was a bursty channel with an average bit error rate of 10^{-3}. Several techniques for combatting the channel effects were considered and implemented, including frame synchronization, pseudo-Gray coding, error detection, and frame repeat strategies. However, prior to implementing any error detection/mitigation strategies, the VAPC algorithm was tested in the presence of bit errors from a simulated satellite channel. The results indicated that the basic algorithms were relatively insensitive to isolated errors and even to moderated bursts of errors, depending on the locations of the errors.

As mentioned above the basic VAPC algorithm frame length is 22.5 ms. This corresponds to a 108 bit frame. Of these 108 bits four bits were allocated for frame synchronization and error detection (more bits could have been allocated, however this would have reduced the quality of the coded speech). This translates into an overhead rate of 200 bits per second for link maintenance. Based on the low number of bits allocated per frame for this purpose, it was decided to minimize the number of bits used for frame synchronization (based on the constraint that the received data is initially synchronized) and to restrict the remaining bits (three) to error detection.

In the case of frame synchronization, there were several issues to be considered, including proper detection of an out-of-synch frame, and proper re-synchronization of a frame once the out-of-synch condition has been detected. In addition to these issues, there is the requirement that the reframing time be kept to a minimum. Based on a tradeoff between acceptable reframing time (one second), detection of the out-of-synch condition, proper resynchronization once the out of synch condition has been detected (versus false detection), the desire to keep the link maintenance overhead at a minimum, and computational complexity, an alternating pattern of ones and zeroes was chosen for the synchronization pattern.

The out-of-synch condition is declared by the codec when the received synchronization pattern over an eight frame history differs from an alternating pattern by more the a single error or two consecutive errors. When the out-of-synch condition is declared, the speech decoder produces silence until the in-synch condition is declared.

Once the codec enters the resynchronization state, a pattern matching algorithm is implemented to detect the alternating synch pattern, and this algorithm operates until a sufficient number of synchronization bits (7 out of 8 bits) are correctly received.

Based on the above algorithms and the channel statistics, it has been computed that the minimum time to detect the out-of-synch condition is three frames, and the probability of non-detection of the out-of synch condition after eight frames is approximately 6%. The average resynch time is estimated to be approximately 8 frames.

In the case of error control, as mentioned above, three bits per frame are allocated for error detection. Given the limited number of bits per frame allocated for this purpose (driven by the speech quality constraint), only the most critical channel errors are addressed: that of burst errors (a mitigation strategy for isolated errors - pseudo-Gray coding is discussed below). To that end,

the three bits in each frame are designated as parity bits jointly covering the speech frame. A 108 bit frame is divided into three words as follows. The first word is formed by concatenating the first frame bit with every third bit after it. The second and third words are formed in a similar fashion. The parity bits are then chosen to force the three words formed to have even parity. The probability that an error burst goes undetected is then approximately the probability that an even number of errors occurs in each parity word (i.e., approximately 13%). Although this probability is relatively high, the experimental results on the robustness of the VAPC algorithm in the presence of errors in the absence of any error protection indicate that it is quite robust to isolated errors. When the parity bits do not check, the previous speech frame is repeated if the number of consecutive repetitions is below two, otherwise silence is played until the error burst ends. Experimental results have shown that this frame repeat strategy significantly reduces the perceptual impact of error burst that last two frames or less.

Finally, a technique to mitigate the presence of isolated errors that involves no coding overhead, called pseudo-Gray coding has been studied and implemented. This technique involves assigning the binary indices to code vectors and codebook design so that isolated channel errors produce minimal perceptual errors (very similar to Gray coded QPSK). Simulation results with PCM on the binary symmetric channel with bit error rates between .01 and 10% have indicated a substantial gain of 2-4 dB in SNR, roughly uniform over the error probability range.

Combining the speech coding/decoding and the channel overhead, the overall complexity of the VAPC algorithm is approximately 4 million multiply/adds per second, and the algorithm requires approximately 8 kwords of RAM for fixed and variable data, and program storage. This algorithm is implemented using two DSP32's for the MSAT-X program. It has also been implemented on a single Motorola 56000 DSP chip at Voicecraft, and at Microtel Pacific Research (with appropriate support chips in both cases).

CODEC TESTING

The speech codec testing consisted of laboratory tests at JPL and UCSB, quantitative tests by the US Department of Defense, field tests by JPL in various environments, and quantitative tests by the Australian TELECOM. The qualitative test results from the JPL field tests and the quantitative results from the US Department of Defense and Australian TELECOM tests are discussed below.

US Department of Defense Testing

The final version of the VAPC algorithm was evaluated by the US Department of Defense [1] in 1988 as part of a very extensive and thorough study of 4800 bit per second speech codecs. The testing program consisted of subjective evaluations of quality under a variety of operating conditions. Subjective ratings were made using the Diagnostic Rhyme Test (DRT) and the Diagnostic Acceptability Measure (DAM). The DRT test measures the ability to distinguish between pairs of rhyming words, and is a measure of the intelligibility of the speech. The DAM test uses complete sentences and listeners judge various quality attributes that lead to an overall measure of speech quality. Clearly, in terms of user acceptability, the DAM scores are the most important, while in cases where intelligibility is of prime concern (e.g., air traffic control) the DRT scores are the most important.

As a result of these tests, the VAPC algorithm was found to have the highest DAM scores (of the seven different codecs that underwent detailed testing) for quiet speech (no background noise), office speech (typical office background noise), speech through a carbon microphone, and a noisy aircraft environment. Under the quiet background noise environment, the VAPC algorithm received a DAM score of 65.5. In comparison, the LPC-10 2400 bit per second standard has a DAM score of 48 under the same

conditions. VAPC ranked poorly in the helicopter noise, tandeming, and 1% bit error rate environments. A point to note is that the VAPC algorithm was not designed to operate in the latter environment. For the DRT tests, the VAPC algorithm tended to score somewhat lower, averaging around fifth out the seven codecs tested.

JPL Field Tests

The VAPC codec has been tested in three separate field tests by JPL. These tests range from a simulated satellite environment using a 1000 foot tower as the satellite simulator, to an aeronautical mobile experiment using the INMARSAT Marecs B2 satellite, to land-mobile satellite experiments in Australia in conjunction with AUSSAT, using the Japanese ETS-V satellite.

In all three field trials, the VAPC codec performed well, providing an intelligible, good quality voice link, through which the experimenters communicated between the mobile unit and the fixed ground station. All users of the speech link were impressed with the quality of the speech and the ability to identify to far end speaker (as compared to LPC-10).

During the aeronautical mobile experiment [4], the full-duplex voice link was established often and used as the main (in fact the only available) method for direct communication between the experimenters on the aircraft (an FAA Boeing 727 flying along the East Coast of the United States) and in the fixed ground station. These links were run routinely at the same signal to noise ratio that resulted in 10^{-3} BER. There was no perceptible difference in speech quality or intelligibility between in-flight and operation on the ground. Jet noise had no significant effect on the communications. A formal part of the experiment was the demonstration of the voice link for air traffic control applications. During one of the flights, an FAA engineer on-board the aircraft read a variety of air traffic control-type messages into the VAPC codec. The voice received at the ground station was assessed by FAA personnel and recorded. Live conversations were also recorded. The intelligibility and quality of the speech, and the robustness of the link, were deemed acceptable by the FAA staff. Remarkably, the audio output of the codec at the ground station, which was available on a headphone speaker, was acoustically (not electrically) patched to a telephone headset and through a long-distance line to an FAA listener attending a meeting in Montreal, Canada. The listener found the voice to be intelligible and its quality to be acceptable.

The last field test that the VAPC codec was tested in was the full scale land-mobile experiment conducted in Australia [5]. During this experiment, the satellite based speech link was used as the primary means of communicating between the mobile terminal and the fixed terminal (an available HF link provided at best, poor quality communications). The experimental performance of the codec was similar to that obtained in the previous two field tests and was dictated by the overall bit error rate performance of the mobile and fixed terminals. During the tests, several speech links were established and maintained over periods of two hours while the mobile terminal travelled the Australian countryside. This link was maintained even in the presence of heavy blockage. During these tests as well as the previous two tests a considerable number of voice tapes ranging from DRT and DAM tapes through live conversations were recorded. All users of the system were impressed by the quality of the speech. Indeed, interested parties at AUSSAT (the Australian national satellite systems providers) were very impressed by the performance of the codec, and rated the overall performance of the codec and terminal superior to the other analog and digital systems they were currently reviewing. As a result of these tests, the Australian land-mobile satellite system specification has been changed from an approach based to analog speech (ACSSB) to digital speech at approximately 5000 bits per second, such as that provided by the VAPC codec.

Australian TELECOM Testing

As part of the Australian experiment mentioned above, the MSAT-X modem and the VAPC codec were installed in the Australian TELECOM laboratories where a variety of channel tests were performed. These tests ranged from the codec/modem performance in the heavily shadowed Rician fading environment to a Rician fading environment (K=20). Of significant interest were the results of the codec/modem pair when compared to the performance of one of the best ACSSB modems available over the Rician fading environment. The overall performance was rated based on the Mean Opinion Score (MOS), a subjective measure of the overall quality of the received speech. The basic results were that the codec/modem pair had an average MOS of slightly better than 3.0 (on a 5.0 scale, with toll-quality speech rated at 4.2) for C/N0 values ranging from 45 dB-Hz to 56 dB-Hz. This MOS value fluctuated slightly over this range due to the sample sizes used in the experiment, but was approximately 3.1 at 45 dB-Hz, and at 56 dB-Hz. In comparison, the ACSSB modem achieved a MOS score of approximately 1.8 at 45 dB-Hz and 3.5 at 56 dB-Hz.

CONCLUSIONS

The development program for 4800 bit per second speech codecs under the MSAT-X program in that several different codecs have been developed that provide good quality specch at this data rate. Of particular note is the performance of the VAPC codec as described in this paper. This codec provides good quality speech at 4800 bits per second, and ranks well when compared to other codecs at the same data rate. A very important distinction between this codec and many of the other 4800 bit per second codecs is the required number of computations per unit time [1]. When compared with other codecs with the same level of computational complexity, the VAPC codec appears to be distinctly superior. In particular, the VAPC algorithm has less than half the complexity of the CELP algorithms tested by the US Department of Defense and appears to be the only one implemented with a single fixed point DSP chip.

Modifications of the VAPC algorithm have led to very high quality codecs at 8 and 16 kilobits/second and commercial liscenses of the algorithm have already been issued. In particular, Compression Labs, Inc. uses VAPC at 8 kbits/s for the audio signal in its low bit rate video codecs.

ACKNOWLEDGEMENTS

This work was performed at the University of California, Santa Barbara, and the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and Space Administration.

REFERENCES

[1] D. P. Kemp, R. A. Sueda, T. E. Tremain, "An Evaluation of 4800 BPS Voice Coders", Proceedings Intl. Conf. on Acoustics, Speech, and Signal Processing, pp. 200-203, 1989.

[2] J. H. Chen, A. Gersho, "Real-Time Vector APC Speech Coding at 4800 BPS Using Adaptive Postfiltering", Proceedings Intl. Conf. on Acoustics, Speech, and Signal Processing, pp. 2185-2188, 1987.

[3] B. S. Atal, M. R. Schroeder, "Predictive Coding of Speech Signals and Subjective Error Criteria", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-27, June, 1979.

[4] T. C. Jedrey, K. I. Dessouky, N. E. Lay, "An Aeronautical Mobile Satellite Experiment", JPL Report in preparation.

[5] T. C. Jedrey, W. Rafferty, "The MSAT-X/AUSSAT Land-Mobile Satellite Experiment: An Overview", MSAT-X Quarterly No. 22, JPL 410-13-22, January 1990.

Masking of Errors in Transmission of VAPC-Coded Speech

Neil B. Cox and Edwin L. Froese
M.P.R. Teltech Ltd.
8999 Nelson Way
Burnaby, B.C., Canada, V5A 4B5
Phone: (604) 294-1471
FAX: (604) 293-5787

1. ABSTRACT

This study provides a subjective evaluation of the bit-error sensitivity of the message elements of a Vector Adaptive Predictive (VAPC) speech coder, along with an indication of the amenability of these elements to a popular error-masking strategy (cross-frame hold-over). As expected, a wide range of bit-error sensitivity was observed. The most sensitive message components were the short-term spectral information and the most significant bits of the pitch and gain indices. The cross-frame hold-over strategy was found to be useful for pitch and gain information, but it was *not* beneficial for the spectral information unless severe corruption had occurred.

2. INTRODUCTION

Application-specific information can often be exploited in the design of error-control methodologies for dedicated communication channels. While a concession is made to the generality of the system when such information is used, there are practical applications for which this concession is acceptable. One such application is speech transmission over mobile satellite channels. Here there are four sources of application-specific information: the channel characteristics, the speech coding format, predictable characteristics of the speech signal and the relative importance of signal components in speech perception. With the possible exception of channel characteristics, these options are not exploited if error control is based solely on general-purpose error correction codes.

The following factors should be considered for efficient control of transmission errors in VAPC-encoded speech:

1. The error-free delivery of all message bits is not required for meaningful speech communication, as human listeners are remarkably adept at inferring meaning from context. This implies that the goal of error control should be to reduce the *perceptual* effect of errors.
2. The short-term predictability of speech provides a variety of intuitive approaches to error compensation, such as adaptive smoothing or cross-frame hold-over of parameters.[1,9] While much of the effort in speech coding is devoted to the removal of this predictability, the coding algorithms generally update their parameters at a high enough rate to adequately represent the signal during its most transient conditions. Thus, residual predictability can be expected for a considerable proportion of the speech sequence.
3. The bits of a coded speech message have a widely-varying influence on the perceived speech quality. Ordered parameters are naturally comprised of bits with varying significance. Some parameters are interrelated or dependent on past samples, leading to a propagation of the errors within a frame and across frames. Certain parameters represent fundamental aspects of speech, whereas others only refine the quality.

Methods of accounting for the varying importance of message bits have been proposed in the literature. Numerous examples can be found where error detection and/or correction is applied to a subset of the message bits.[1,9] The parallel application of codes has been used to further concentrate the protection on the most important bits.[10] Rate-compatible punctured convolutional codes provide for selective allocation of code power without the need to switch between coders.[5] All of these approaches require a rank-ordering of message bits. Based in part on informal listening tests, it is common to leave residual information unprotected for linear prediction coders (LPC) and sub-band coders.[9,10] It has been reported for the basic LPC-10 approach that the critical bits are the most significant bits of the first three or four prediction coefficients along with the most significant bits of the gain, pitch and voicing parameters.[1] The more complex LPC approaches are not directly comparable, as the encoding introduces dependence between parameters and between frames. Nonetheless, it is generally

observed that residual information is less sensitive to bit errors than gain, pitch or spectral information.

The purpose of this study is to provide data on the bit-error sensitivity of the message elements in a Vector Adaptive Predictive (VAPC) speech coder. Existing information on this topic is sparse and has generally been acquired in an informal fashion. The sensitivity of each message element to random errors is addressed, along with relative merit of holding-over preceding message elements when errors are present. The evaluation was performed for a random error model and a 2-state Markov simulation of burst errors. The results provide useful guidance in the design of efficient error control techniques for VAPC-encoded speech.

3. VECTOR-ADAPTIVE PREDICTIVE CODING

The VAPC encoding algorithm is illustrated in Figure 1. Briefly, the speech waveform is passed to the encoder in 20 msec frames. The pitch-period is determined using a bounded search for the autocorrelation peak. A 3-tap linear pitch predictor is used to remove signal components that are related to pitch. The prediction points are separated from the predicted point by one pitch-period. This is followed by a 10-th order linear predictive inverse filter that models the spectral envelope. Gain information is derived from the output of the two filters. Finally, residual vectors are selected to minimize the difference between the input signal and a locally-synthesized output. This analysis-by-synthesis approach partially compensates for errors that result from quantization of the pitch, spectral and gain information. Further detail can be found elsewhere.[4,6,12]

The codec evaluated in this study has the following bit allocation. The pitch-period index (*idxp*) is a 7-bit linear quantization of the pitch-period. The pitch prediction vector index (*idxpp*) uses 6 bits to select a pitch prediction filter from a codebook of 64 candidates. The selected predictor provides the largest reduction of signal energy. The LSP error indices (*idxsp*) and the classification index (*idxcl*) are a complex relative representation of the short-term spectral information, where the *idxsp* are scalar quantizations of the difference between computed Line Spectrum Pair (LSP) coefficients and values predicted from the previous frame through first-order linear vector prediction, and *idxcl* chooses one of four sets of coefficients for the vector predictor. This consumes 29 bits. The residual gain index (*idxg*) is a 6-bit logarithmic quantization of the residual energy. Finally, the sixteen 7-bit residual vector indices (*idxr*) are a multi-stage vector quantization of the excitation signal that minimizes the analysis-by-synthesis error.

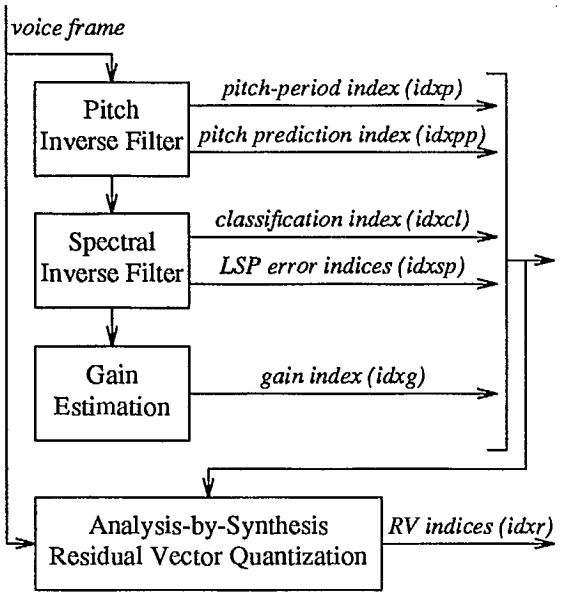


Figure 1: Structure of a Generic VAPC Encoder

4. EXPERIMENTAL SETUP

The codec described above has been implemented for real-time operation on a single 20 mHz Motorola DSP56001 signal processor. Sufficient time is available to also perform adaptive echo cancellation if the 27 mHz version of the chip is used. In fact, a single-chip real-time 2-channel implementation has been developed by excluding the echo cancellation and post filtering, and shortening the residual vector codebooks.

A locally-developed test bed for codec evaluation was used for this study. This test bed provides flexible synchronization, communication and data input/output among general-purpose DSP56001-based processing cards. A separate program can be downloaded to each card and interactively modified when necessary. Three cards were used for the present study. Two of the cards received the VAPC encoder/decoder program and the third card received an error imposition program.

The error imposition program is capable of imposing random errors on specified groups of bits within a frame, or optionally holding over preceding indices instead of imposing the errors. The bit-error probability is adjusted by a thumbwheel switch while the program is running. A special "decoder hold-over mode" was implemented for the short-term spectral information. Here the LPC parameters derived within the decoder are held over whenever an error is imposed on the classification index or the LSP error indices. By using a derived parameter set, the effects of a bit

error are spread to a number of the parameters, thus making it advisable to hold-over the entire LPC parameter set. Due to their relative encoding format, the transmitted spectral indices are not themselves suitable for cross-frame hold-over.

Two models for the bit errors were implemented; a random error model and a 2-state Markov model. The first model simply applies random errors to the data at the rate specified by the thumbwheel switch. The two states for the Markov model are random error models. At the start of each frame, one of the states is selected based on pre-specified state transition probabilities. The test configuration was set up so that an error-free state was chosen 90% of the time. That is, both rows of the state transition matrix were (0.9 0.1). The error rate for the "bad" state is determined by the thumbwheel switch. This simulates a channel with bursty errors, and loosely conforms to data described in an AUSSAT codec test procedure.[11]

5. ASSESSMENT METHODOLOGY

The subjective assessment methodology is similar to the Degradation Category Rating (DCR) procedure.[2] This is a pair-wise comparison procedure where the listeners are asked to judge the degradation of the second sample of each pair relative to the first. The following 5-point degradation scale was used:

5 = no degradation
4 = slightly annoying
3 = annoying
2 = very annoying
1 = extremely annoying

Eight listeners were seated in a quiet room and provided with written instructions about the format of the test. The listeners were not allowed to discuss or compare judgements. The test samples were presented over a high quality audio system. Twelve "practice" samples were presented at the start of the test in order to familiarize the listeners with the task and to expose them to the range of degradations that they will encounter. These judgements were excluded from subsequent analyses.

The test stimuli were recorded in random order on a test tape as a series of A-B pairs, where A is a voice sample that is passed through the codec without imposed errors, and B is the same sample with some form of imposed degradation. Two such A-B pairs were included for each test configuration. The voice sample for one of the pairs was a female reciting "The navy attacked the big task force, see the cat glaring at the scared mouse". The voice sample for the other pair was a male reciting "March the soldiers past the next hill, a cup of sugar makes sweet fudge". These sentences are from the phonetically-balanced Harvard sentence collection. Finally, six "null pairs" (A-A) were included to test the anchoring of the listeners' assessments.

6. TEST CONFIGURATIONS

The bit-error probabilities for the test configurations are summarized in Table 1 and Table 2. These error levels were derived in an informal preliminary trial to produce degradation ranging from slight to severe. The implied frame-error probability (i.e. the probability of one or more bit errors within a frame) is included in brackets. In order to keep the test trials at a manageable length, single-bit evaluations were only performed for the ordered indices (*idxp* and *idxg*), and each *idxr* and *idxsp* index was not individually tested.

Both the random error mode and the random hold-over mode were tested for all but the single-bit configurations. In addition, the "full dropout" condition (i.e. BER=0.5) was evaluated for the Markov model. A total of 158 test pairs were tested, comprised of 60 random error pairs, 16 Markov error pairs and 3 anchoring (null degradation) pairs for each voice sample.

Table 1: Bit-error probabilities for random error test configurations. The degradation was judged in an informal preliminary trial. "n" = P_{error} for each bit, and "(n)" = the implied P_{error} for each frame. Bit 1 is the least significant.

Errored Index	Degradation low	mid	high
idxg bit 1	.1	.2	.5
idxg bit 3	.05	.1	.2
idxg bit 5	.02	.05	.1
idxp bit 1	.1	.2	.5
idxp bit 3	.05	.1	.2
idxp bit 5	.02	.05	.1
idxr	.02 (.997)	.05 (≈1)	.1 (≈1)
idxg	.02 (.11)	.05 (.26)	.1 (.47)
idxp	.02 (.13)	.05 (.30)	.1 (.52)
idxpp	.02 (.26)	.05 (.47)	.1 (.74)
idxp+idxpp	.02 (.23)	.05 (.49)	.1 (.75)
idxcl+idxsp	.001 (.03)	.005 (.13)	.01 (.25)
all	.001 (.15)	.005 (.55)	.01 (.80)

Table 2: Bit-error probabilities for the "bad" state in Markov error test configurations. Here an average of 90% of the frames are error-free, and the remaining frames have the random bit-error probability designated below. The degradation was judged in an informal preliminary trial. "n" = P_{error} for each bit in a "bad" frame, and "(n)" = the implied P_{error} for each "bad" frame.

Errored Index	Degradation low	mid	high
idxcl+idxsp	.01 (.25)	.02 (.44)	.05 (.77)
all	.01 (.80)	.02 (.96)	.05 (≈1)

7. RESULTS

The means of the degradation scores for all test configurations are summarized in Tables 3 through 7. Each mean was derived from 16 judgements (i.e. 8 listeners and 2 samples per listener). The average variance was approximately 0.5 for these judgements. Based on a one-tailed Students' t, this implies that differences of greater than 0.6 are significant at the 1% level, and differences of greater than 0.4 are significant at the 5% level. The mean of the degradation for the null pairs was 4.9, indicating that the judgements were well anchored.

Table 3: Degradation MOS for random errors in bits of the gain index *idxg*. Bit 1 is the least significant.

Errored bit	Bit-error probability .02	.05	.1	.2	.5
1	---	---	4.9	5.0	4.8
3	---	4.9	4.8	4.6	---
5	3.3	2.2	1.7	---	---

Table 4: Degradation MOS for random errors in bits of the pitch-period index *idxp*. Bit 1 is the least significant.

Errored bit	Bit-error probability .02	.05	.1	.2	.5
1	---	---	4.5	4.9	4.3
3	---	4.2	3.4	3.1	---
5	4.8	3.5	2.5	---	---

The single-bit conditions summarized in Table 3 and Table 4 demonstrate the expected relationship between bit-error sensitivity and bit significance for *idxg* and *idxp*. For *idxg* there was a sudden onset of severe degradation; corruption of bit 1 or bit 3 had little effect, but corruption of bit 5 caused severe degradation. This is partially explained by the logarithmic quantization of this index. The onset of degradation was more gradual for *idxp*.

The following observations can be drawn from the single-index conditions summarized in Table 5:

— Corruption of the residual vector indices (*idxr*) caused a moderate level of degradation for the tested bit-error rates. Cross-index hold-over provided a statistically significant reduction of the bit-error sensitivity, but notable degradation was still present.

— Corruption of the gain index (*idxg*) caused severe degradation at all tested error levels. The cross-frame hold-over strategy provided a large improvement, with only moderate degradation produced by the worst error rate (BER=0.1).

— The pitch-period index (*idxp*) was relatively sensitive to bit errors. Fortunately, as with the gain index, cross-frame hold-over provided a significant improvement.

— The pitch prediction index (*idxpp*) was relatively insensitive to bit errors, and no significant improvement was obtained from cross-frame hold-over. Furthermore, there appears to be no significant interaction between *idxp* and *idxpp* in terms of the bit-error sensitivity, as corruption of both indices has approximately the same effect as corruption of *idxp* alone.

Table 5: Degradation MOS for random errors in the pitch, gain and residual indices. Data are for random bit-errors, and random index hold-over in response to such errors.

Errored Index	Random Errors Bit-error probability .02	.05	.1	Random Hold-Over Bit-error probability .02	.05	.1
idxr	3.4	2.7	2.2	4.0	3.7	2.5
idxg	2.3	1.7	1.5	4.6	3.4	3.3
idxp	2.9	2.6	1.6	4.0	3.3	2.6
idxpp	4.4	3.6	2.8	4.6	3.8	3.1
[*p+pp*]	2.9	2.6	1.9	4.3	2.8	2.4

Table 6: Degradation MOS for random errors in spectral indices (*idxcl+idxsp*) and all indices. Data are for random bit-errors, and random index hold-over in response to such errors. The "decoder hold-over mode" was used for the spectral indices.

Errored Index	Random Errors Bit-error probability .001	.005	.01	Random Hold-Over Bit-error probability .001	.005	.01
[*sp+cl*]	4.1	2.9	2.6	3.9	3.0	1.9
all	3.6	2.2	1.6	4.4	2.3	1.9

Table 7: Degradation MOS for bursty errors in spectral indices (*idxcl+idxsp*) and all indices. A Markov error model is used, where an average of 90% of the frames are error-free, and the remaining frames have the random bit-error probability designated below. Data are for random bit-errors, and random index hold-over in response to such errors. The "decoder hold-over mode" was used for the spectral indices.

Errored Index	Markov Errors Bit-error probability .01	.02	.05	.5	Markov Hold-Over Bit-error probability .01	.02	.05	.5
[*sp+cl*]	4.0	3.1	3.7	1.3	3.9	3.2	3.2	3.3
all	4.3	3.7	2.8	1.0	3.6	3.7	3.5	2.8

When one considers that the bit error rates in Table 6 and Table 7 are 10 times less than those in Table 5, it is clear that the spectral indices (*idxcl* and *idxsp*) are by far the most sensitive to bit errors. Except for the "full dropout" condition (BER = 0.5) in the Markov error simulation, the cross-frame hold-over strategy did not improve the situation, and produced a significant *degradation* at a random bit-error rate of 0.01. Thus, the hold-over strategy should only be counted on when data transmission is severely compromised. This view is supported by the "all indices" data in these Tables, as a significant improvement was only provided when severe corruption was present. The one exception (random errors at a bit-error rate of 0.001) may be due to the shortness of the speech samples, as few bit-error combinations are encountered at low error rates.

There was a wide diversity in the quality of the perceived error effects. Corruption of *idxr* caused "garbling" of the speech but did not produce an alarming disturbance. Errors in the gain index, on the other hand, tended to impose intermittent and extremely loud bursts. The spectral errors caused intermittent alarming "whoops" and "squawks", that is, the disturbances were very loud and irritating, and appeared to have an entirely inappropriate frequency content. Finally, corruption of the pitch indices had the expected effect of introducing a hoarse quality to the speech, with intermittent abnormal jumps in pitch.

8. DISCUSSION

A general conclusion of this study is that most of the effort in error control should be devoted to protection of the short-term spectral information (*idxcl* and *idxsp*), with attention also given to the most significant bits of the gain index (*idxg*) and the pitch-period index (*idxp*). The spectral parameters were followed in importance by the gain index (*idxg*), the pitch-period index (*idxp*), the residual vector indices (*idxr*) and the pitch prediction index (*idxpp*). Errors in the three least significant bits of the pitch and gain indices (*idxp* and *idxg*) had little perceived effect. Also, there is little reason to protect the pitch prediction index if the residual vector indices are left unprotected, as the degradation caused by corruption of *idxr* is relatively severe before corruption of *idxpp* becomes noticeable.

If a moderate degradation is acceptable at bit-error rates of 0.05 or more, then the practice of leaving residual vector indices unprotected is justified. A comprehensive error correction protocol requires excessive redundancy, as the RV indices comprise the majority of the bits of the message. The lack of a natural ordering for the residual vectors makes it difficult to rank order the message bits, thus ruling out bit-selective strategies. While this study indicates that some improvement can be obtained by using a cross-index hold-over strategy, this requires a coding method with sufficient power to localize the error(s) to specific indices.

It is recommended that global application of cross-frame hold-over should only be relied upon in burst error conditions, at least for the short-term spectral information. Here the main advantage is in preventing extreme and highly irritating signal disturbances. However, the hold-over strategy was beneficial for the gain index (*idxg*), the pitch-period index (*idxp*) and to a lesser extent the residual vector indices (*idxr*).

Other methods of error masking may be beneficial to augment or replace the hold-over strategy. For example, progressive muting of the output during bursts has been recommended.[1,9] Both linear and non-linear approaches can be used to derive estimates of corrupted parameters based on past history. Cross-frame hold-over is a special case of this. Other examples are linear extrapolation and median filtering.

Running estimates of the probability distribution or other statistics of parameters would be useful in accounting for context-dependent effects. The parameters used in such an analysis can be taken from any stage of the decoder. The use of "sped-up speech" in combination with automatic repeat request (ARQ) protocols has been proposed for bursty channels.[8] The bursty speech that results may be less annoying than the disturbance associated with the other strategies.

Index assignment optimization methods have been proposed for error masking.[3,7] Here a measure of the effect of an error is assumed, and the indices are assigned such that the most probable error patterns produce the smallest effects. Such strategies are attractive in that they are simply implemented and do not require added redundancy or added run-time computation. Unfortunately, a number of factors argue against their success. For example, mathematical measures of error effects have not demonstrated a good correlation with the actual perceived effect. Even if the measure is accurate, most parameters of speech are highly nonstationary, so an optimized index allocation based on a fixed statistical model may well be inferior in many conditions. Nonetheless, this approach may be beneficial in situations where other strategies are not practical, such as for protecting residual vector indices.

The small size of this study limits the general applicability of the results. We have limited ourselves to random errors imposed on short, albeit phonetically-balanced, samples of speech passed through a single VAPC codec. It is recognized that the length of the sample is undoubtedly insufficient for thorough testing of all speech contexts, particularly at low error rates. Limiting the experiment to two English-language speakers neglects numerous external factors, such as age, health, linguistic background, habitual pitch, etc.. The effects of changing the codec configuration or the input amplitude were not tested. Finally, the diversity of perceived effects caused by various bit-errors makes it potentially misleading to use a single opinion score as a basis for comparison.

It is recognized that the sound reproduction and listening environment were of a higher quality than can reasonably be expected in most applications. This method of test presentation facilitates the detection of subtle degradations and makes it easier to concentrate throughout the test. An informal verification of the presentation format was performed, where one listener repeated the test on a different day using a standard telephone handset. As expected, there was a reduced ability to detect subtle degradations over the handset, and the severely degraded samples were not as alarming. The variance of the difference between the two sets of judgements from this listener was approximately the same as the average variance of the audio-speaker-based assessments across listeners. A complete assessment of this issue requires simulation of the range of receiving apparatus and noise environments.

REFERENCES

[1] **Bryden, B., Seguin, G.E., Conan, J., Bhargava, V.L., and Brind'Amour, A.**, 1989. "Error Correction/Masking for Digital Voice Transmission Over the Land Mobile Satellite System." *IEEE Trans. on Commun.*. pp. 309-314.

[2] **CCITT,**, 1988. "Subjective Performance Assessment of Digital Processes using the Modulated Noise Reverence Unit (MNRU)." *CCITT Blue Book, Vol. V, Supplement 14*. pp. 341-360.

[3] **Chen, J.H., Davidson, G., Gersho, A., and Zeger, K.**, 1987. "Speech Coding for the Mobile Satellite Experiment." *IEEE Int. Conf. on Commun.*. pp. 756-763.

[4] **Chen, J.H. and Gersho, A.**, 1987. "Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering." *IEEE Int. Conf. on ASSP*. pp. 2185-2188.

[5] **Cox, R.V., Hagenauer, J., Seshadri, N., and Sundberg, C.E.**, 1988. "A Sub-Band Coder Designed for Combined Source and Channel Coding." *IEEE Int. Conf. on ASSP*. pp. 235-238.

[6] **Davidson, G. and Gersho, A.**, 1988. "Multiple-Stage Vector Excitation Coding of Speech Waveforms." *IEEE Int. Conf. on ASSP*. pp. 2185-2188.

[7] **DeMarca, J.R.B. and Jayant, N.S.**, 1987. "An Algorithm for Assigning Binary Indices to the Codevectors of a Multi-dimensional Quantizer." *IEEE Int. Conf. on Commun.*. pp. 1128-1132.

[8] **Lynch, J.T.**, 1987. "Evaluation of the Comprehension of Noncontinuous Sped-up Vocoded speech: A Strategy for Coiping with Fading HF Channels." *IEEE J. on Selected Areas in Commun. 5*. pp. 308-311.

[9] **McLaughlin, M.J. and Rasky, P.D.**, 1988. "Speech and Channel Coding for Digital Land-Mobile Radio." *IEEE J. on Selected Areas in Commun. 6*. pp. 332-345.

[10] **Suda, H. and Miki, T.**, 1988. "An Error Protected 16 kbit/s Voice Transmission for Land Mobile Radio Channel." *IEEE J. on Selected Areas in Commun. 6.* pp. 346-352.

[11] **Wilkinson, M.H.**, 1990. *Voice Codec Test and Evaluation Procedure VCTEP/2.* Telecom Research Laboratories; Australian Telecommunication Corporation

[12] **Yong, M., Davidson, G., and Gersho, A.**, 1988. "Encoding of LPC Spectral Parameters using Switched-Adaptive Interframe Vector Prediction." *IEEE Int. Conf. on ASSP.* pp. 402-405.

A Robust CELP Coder with Source-Dependent Channel Coding

Rafid A. Sukkar [1] **& W. Bastiaan Kleijn** [2]
AT&T Bell Laboratories
[1] 200 Park Plaza, Naperville, IL 60566
Phone: (708) 713-5281
[2] 600 Mountain Ave, Murray Hill, NJ 07974

ABSTRACT

A CELP coder utilizing Source-Dependent Channel Encoding (SDCE) for optimal channel error protection is introduced. With SDCE, each of the CELP parameters are encoded by minimizing a perceptually meaningful error criterion under prevalent channel conditions. Unlike conventional channel coding schemes, SDCE allows for optimal balance between error detection and correction. Our experimental results show that our CELP system is robust under various channel bit error rates and displays a graceful degradation in SSNR as the channel error rate increases. This is a desirable property to have in a coder since the exact channel conditions cannot usually be specified *a priori*.

I. INTRODUCTION

Significant strides have been made in improving the speech quality of Code Excited Linear Prediction (CELP), making it a viable method for many telecommunication applications where bandwidth is scarce. In many of these applications, including mobile satellite communications, the speech coding algorithm must be robust in the presence of channel errors. CELP research efforts have focused mainly on improving the speech quality, and minimizing the computational complexity. Recently, more attention has been directed toward the robustness of the algorithm in the presence of channel errors [1].

In this paper a CELP system with source-dependent channel encoding scheme is introduced, extending earlier work described in [6]. For every CELP parameter, the source-dependent channel code is obtained by minimizing an appropriate distance measure. Compared to conventional forward error protection methods, SDCE is more efficient due to several factors. First, conventional error protection codes are designed without knowledge of the source coder implying that the bits that need to be protected must be hand picked, thereby providing only a rudimentary form of source-dependent channel coding. SDCE on the other hand provides error correction/detection such that highly probable quantization levels receive more accurate correction and/or serious errors are more likely to be detected. Second, with conventional methods, error correction/detection performance is predetermined, while with SDCE an optimal trade-off between error correction and detection is obtained. Third, conventional error correction codes are designed to perform exact error correction, with associated large increase in bit rate. With SDCE, significant improvement in performance can be obtained by reducing the impact of errors rather than reducing the number of errors. Also, error sensitivity can be reduced by an arbitrary amount using fractional bit allocation.

The organization of this paper is as follows. In the next section a brief description of our CELP system is given. In Section III SDCE is applied to each of the CELP parameters individually and performance with respect to channel errors is shown. Finally, in Section IV a complete CELP system with error protection bit allocation is given. Performance and experimental results are shown.

II. CELP CODER DESCRIPTION

The CELP system used here is based on the system described in [2]. Spectral information is transmitted as 10 line spectral frequencies and updated every 30 msec. Each 30 msec. frame is divided into four subframes for LPC excitation modeling. The LPC excitation modeling consists of two codebook searchs; an adaptive codebook search for modeling the speech periodicity, and a stochastic codebook search for modeling the speech randomness. The adaptive codebook has 128 overlapping entrees consisting of samples of previous frame excitations. The stochastic codebook is also overlapping, consisting of 512 entrees of center-clipped white Gaussian noise samples. However, only even numbered entrees are allowed for transmission implying a total of 256 codewords. A summary of the CELP parameter bit allocation without error protection is given in Table 1. The effective bit rate is 4233 bits/s.

parameter	bits/subframe	bits/frame
LSF(1)	N/A	4
LSF(2)	N/A	4
LSF(3)	N/A	4
LSF(4)	N/A	4
LSF(5)	N/A	4
LSF(6)	N/A	3
LSF(7)	N/A	3
LSF(8)	N/A	3
LSF(9)	N/A	3
LSF(10)	N/A	3
adap. bk index	7	28
adap. bk gain	4	16
stoch. bk index	8	32
stoch. bk gain	4	16
Total		127

Table 1. Bit allocation without error protection

III. SOURCE-DEPENDENT CHANNEL ENCODING

The first step in designing source-dependent channel codes is to define a suitable error criterion. For the CELP parameters an ideal error criterion would be a function of the final synthetic speech quality. However, because of the computational complexity, such an error criterion is unrealistic for the combinatorial optimization required to find good channel codes. Instead, any criterion that is monotonically related to the synthetic speech quality can be used to produce similar results.

Let r_j, (j = 0,1,...J–1), be a quantized version of a given parameter, r. Let the available codewords be denoted as c_m, (m = 0,1,...,M–1), where $M \geq J$. Our goal is to find an optimal mapping, $f(c_m)$ that maps the the codeword c_m into a quantization index j ($j = f(c_m)$). This optimal mapping is obtained by minimizing an appropriate error criterion. The error criterion takes on the following general form:

$$E = \sum_{m=0}^{M-1} P(c_m) \sum_{n=0}^{M-1} P(c_n \mid c_m)\, D(r_j, r_i) \qquad (1)$$

where $P(c_m)$ is the a-priori probability that the codeword c_m is transmitted (this probability is zero for redundant codewords), and $P(c_n \mid c_m)$ is the transitional probabilities due to channel errors. The function $D(r_j, r_i)$ is a distance measure indicating the penalty for using r_i instead of r_j, where $j = f(c_m)$ and $i = f(c_n)$. The upper limit on the sum, M, is the total number of codewords.

If the error function, E, is evaluated as in Equation (1), the channel characteristics need to be defined. However, in many cases the channel characteristics are not well defined and stationarity cannot be guaranteed. Therefore, we like to modify the error criterion such that only broad assumptions are made about the channel, resulting in channel codes with performance that does not degrade significantly under varying channel conditions. A reasonable assumption to make is that the channel can alter at most a predefined number of bits in each codeword, where the most likely errors are weighted more heavily. For instance, single bits errors are weighted more heavily than double bit errors. For the derivation of channel codes in our CELP system, we assumed single bit errors only, though the method can easily be extended to cover any number of bit errors. We also assumed that all one bit errors are equally likely. The assumption of single bit errors is realistic if bit interleaving is employed and the channel performance is relatively good.

Based on the above assumptions the error criterion can now be written as

$$E = \sum_{m=0}^{M-1} P(c_m) \sum_{k=0}^{K-1} D(r_j, r_{j_k} \mid f(.)) \qquad (2)$$

where, as before, $j = f(c_m)$, and j_k is the quantization index corresponding to the codeword c_m with bit k inverted ($j_k = f(c_{m_k})$). The function $D(r_j, r_{j_k} \mid f(.))$ is the penalty function associated with replacing r_j with r_{j_k}, given a specific mapping function, f(.). Here, K is the total number of bits in each codeword. The error function, E, is minimized with respect to the mapping function, f(.). This minimization is highly non-linear requiring a simulated annealing-type procedure [3,4] to find the optimal f(.).

If redundant codewords are used, then the minimization of E can be used for both error correction and detection. For error detection, an additional fictitious level is introduced. Any of the redundant codewords can map into this fictitious quantization level. Receipt of a redundant codeword mapping into this level would indicate a transmission error, triggering the error recovery procedure. The penalty for synthesizing with this fictitious quantization level can be determined and must be used during optimization. Error correction is performed by assigning more than one codeword to map into a single quantization level. This error detection/correction SDCE scheme results in an optimal trade-off between error correction and detection. In our system the penalty function in Equation (2) depends on the CELP parameter at hand. We will now treat each parameter separately.

Line Spectral Frequencies

The penalty function for the line spectral frequencies is based on the cepstral distance measure [5]. For each LSF parameter the distance measure is defined as

$$D(LSF_j(p),LSF_{j_k}(p)\,|\,f(.)) = E[c_j{}^T c_{j_k}\,|\,f(.)] \qquad (3)$$

where p is the LSF number, c_j is the cepstral coefficient vector based on the quantized LSF's, and c_{j_k} is the cepstral coefficients vector corresponding to the the quantized LSF's with $LSF_{j_k}(p)$ replacing $LSF_j(p)$. This LSF replacement may, however, result in unrealistic LSF vectors since the monotonicity property may be lost. These cases can be thought of as error detect cases where the decoder receives an unrealistic LSF vector due to channel errors. Therefore, the strategy used in these cases should be the same as the strategy used in the decoder when unrealistic LSF vectors are received. If p is odd, then the previous frame $LSF_j(p)$ and $LSF_j(p+1)$ are substituted for the present frame $LSF_{j_k}(p)$ and $LSF_j(p+1)$, respectively. If p is even, then the previous frame $LSF_j(p)$ and $LSF_j(p-1)$ are substituted for the present frame $LSF_{j_k}(p)$ and $LSF_j(p-1)$, respectively. The monotonicity is checked again and if the resulting LSF vector is still unrealistic, then the whole LSF vector of the previous frame is used to compute the penalty function for the present frame. The expected value in Equation (3) is computed over all voiced frames in a database consisting of 24 sentences.

Parameter	NBC Mean Error	Gray Mean Error	SDCE Mean Error
LSF(1)	9.55	7.77	6.28
LSF(2)	13.62	11.16	10.41
LSF(3)	13.44	11.09	10.15
LSF(4)	13.73	10.73	10.66
LSF(5)	15.08	12.58	12.36
LSF(6)	16.89	15.26	14.65
LSF(7)	15.78	14.99	14.61
LSF(8)	13.00	11.09	10.29
LSF(9)	10.54	8.34	7.97
LSF(10)	7.92	7.27	7.07

Table 2. LSF error criterion after minimization

The error function of Equation (2) incorporating the penalty function of Equation (3) is minimized using a simulated annealing procedure [3]. With no redundant bits, the results of the minimization are given in Table 2. For comparison purposes the penalty function corresponding to the Natural Binary Code (NBC) and Gray code are also given.

Table 2 shows that SDCE consistently outperforms the other two schemes with a large improvement for LSF(1). The large improvement for LSF(1) is attributed to the fact that the the penalty functions associated with the quantization levels of LSF(1) have larger variation in dynamic range than the penalty functions of LSF(2)-LSF(10). This is typical of SDCE where serious errors are weighted more heavily than less serious errors in the optimization process.

To test SDCE on actual speech, the 24 sentence database was used to obtain a channel bit stream that was then corrupted on an LSF by LSF basis. For every other frame in the bit stream, one random bit of a given LSF codeword was inverted. The resulting Segmental Signal-to-Noise Ratio (SSNR) between the original speech and the synthetic speech over all voiced frames in the database is given in Table 3. The clear channel SSNR is 9.96 dB.

Parameter	NBC SSNR(dB)	Gray SSNR(dB)	SDCE SSNR(dB)
LSF(1)	6.58	6.96	7.35
LSF(2)	6.16	6.60	6.60
LSF(3)	7.36	7.71	7.72
LSF(4)	7.84	8.27	8.25
LSF(5)	9.01	9.15	9.20
LSF(6)	9.07	9.20	9.34
LSF(7)	9.35	9.46	9.45
LSF(8)	9.56	9.62	9.62
LSF(9)	9.80	9.83	9.82
LSF(10)	9.87	9.87	9.88

Table 3. SDCE actual speech performance after optimization

Again, Table 3 shows a significant improvement is obtained for LSF(1), while only marginal improvement to no improvement is obtained for LSF(2)-LSF(10).

To take better advantage of SDCE properties, we can consider combining two or more quantized LSF's and code them as one parameter (i.e., vector coding). The advantage of this is that by combining two or more quantized LSF's some combination of quantization levels become unrealistic due to the LSF monotonicity property. These levels, which can correspond to any fraction of a bit, can be used by

the SDCE procedure as redundant levels for error correction and detection. In the scalar case, these unrealistic levels correspond to receiving an unrealistic LSF vector, thereby providing only a rudimentary form of error detection. In the vectorized case, SDCE uses these redundant levels to strike an optimal balance between error correction and detection.

Because of the computational complexity involved in the optimization process, we chose to combine only two LSF's at a time, although the coding efficiency increases as more LSF's are combined. The results for 0-bit redundancy are tabulated in Table 4. The SSNR column in Table 4 represents the average SSNR of the synthetic speech over voiced frames in the database after inverting a single bit in a combined LSF codeword every other frame. To compare with the performance of the scalar case, we have generated in Table 5 SSNR values for the case of channel encoding each LSF individually but corrupting, every other frame, a random bit taken from the set of bits spanning the codewords of two LSF's. The results of Tables 4 and 5 indicate that the combined case gives a significant improvement for LSF(1,2), without adding extra bits, or reducing the number of valid quantization levels. These results also demonstrate the ability of SDCE to use non-integer bit redundancy for error protection.

Parameter	Bits	Quant/Redun Levels	Mean Error	SSNR dB
LSF(1,2)	8	190/66	7.07	7.12
LSF(3,4)	8	190/66	9.01	7.83
LSF(5,6)	7	96/32	11.82	9.26
LSF(7,8)	6	52/12	10.72	9.60
LSF(9,10)	6	52/12	6.42	9.87

Table 4. SDCE performance of combined LSF's with 0-bit redundancy (vectorized case).

Parameter	Bits	Quant/Redun Levels	SSNR dB
LSF(1,2)	8	256/0	6.70
LSF(3,4)	8	256/0	7.76
LSF(5,6)	7	128/0	9.18
LSF(7,8)	6	64/0	9.55
LSF(9,10)	6	64/0	9.85

Table 5. SDCE performance of the scalar LSF optimization.

Table 6 shows the performance of the combined case after adding one-bit redundancy. Comparing Tables 4 and 6, it is clear that one redundant bit results in a significant improvement in SSNR for LSF(1,2) and LSF(3,4). These results indicate that the speech quality is susceptible to errors in LSF(1)-LSF(4) and is only marginally sensitive to errors in LSF(5)-LSF(10).

Parameter	Bits	Quant/Redun Levels	Mean Error	SSNR dB
LSF(1,2)	9	190/322	4.27	7.75
LSF(3,4)	9	190/322	5.46	8.52
LSF(5,6)	8	96/160	6.16	9.50
LSF(7,8)	7	52/76	3.50	9.80
LSF(9,10)	7	52/76	3.34	9.88

Table 6. SDCE performance of combined LSF's with 1-bit redundancy (vectorized case).

Codebook gain parameters

The penalty function used for the adaptive and stochastic gain parameters is derived from the error criteria used in CELP for choosing the codebook winning indices and determining the optimal gain. The penalty function is written as

$$D(\lambda_j, \lambda_{j_k} \mid f(.)) =$$

$$E\left[10\log\left[(\lambda_{j_k} s_w - t)^T \mathbf{H}^T \mathbf{H} (\lambda_{j_k} s_w - t)\right] \mid f(.)\right], \quad (4)$$

where λ_j is the optimal quantized gain, and λ_{j_k} corresponds to the quantization level obtained by inverting bit k of the codeword assigned to λ_j. The matrix $\mathbf{H}$ is the matrix which transforms the excitation vector of CELP into its zero-state response of the inverse linear predictive filter [2]. The vector s_w is the winning entree into the codebook, and t is the target excitation vector in CELP. The expected value is carried over voiced frames in the 24 sentence database.

Table 7 shows the adaptive codebook gain performance under 1-bit channel errors for various encoding schemes. The mean penalty is measured as a mean signal-to-noise ratio defined as, $E[10\log(t^T t)] - D(\lambda_j, \lambda_{j_k} \mid f(.))$. The distribution of the adaptive gain quantization levels is highly non-uniform with values close to unity having the highest probability. The third row of Table 7 shows an example where a non-integer number of bits is used for protection. In this example the number of quantization levels is dropped from 16 to 12 by eliminating four quantization levels. With only 4

redundant levels a significant improvement is achieved at a minimal cost to the clear channel SSNR performance which dropped from 9.96 dB to 9.78 dB. This large improvement with only a small number of redundant levels is typical of SDCE and is the result of the channel code protecting the quantization levels with high probability only.

Adding a redundant bit results in a significantly higher performance. This performance is even higher than that obtained when 1-bit parity is used despite the fact that the parity bit was not subjected to bit errors. The error recovery strategy used in the error detect cases was to repeat the previous frame adaptive codebook gain.

Code	Bits	Quant/Redun Levels	Mean Penalty dB	SSNR dB
Gray	4	16/0	-0.89	2.38
SDCE	4	16/0	0.71	2.69
SDCE	4	12/4	3.59	6.54
Parity	5	16/16	3.91	7.15
SDCE	5	16/16	4.56	8.21
SDCE	6	16/48	4.95	9.37

Table 7. Adaptive codebook gain performance

The performance of the stochastic codebook gain displays the similar trends to those of the adaptive codebook gain, although the improvements over Gray code are not as dramatic. This is because of the smaller dynamic range of the stochastic codebook compared to the adaptive codebook, and the more uniform statistical distribution of the quantization levels. A complete discussion of the stochastic codebook gain performance is given in [6].

Codebook indices

The penalty function used here is similar to the one used for the gain parameters. It is defined as

$$D(\mathbf{s}_j, \mathbf{s}_{j_k} \mid f(.)) =$$

$$E\left[10\log\left[(\lambda_w \mathbf{s}_{j_k} - \mathbf{t})^T \mathbf{H}^T \mathbf{H} (\lambda_w \mathbf{s}_{j_k} - \mathbf{t})\right] \mid f(.)\right], \quad (5)$$

where $\mathbf{H}$ is define as before, and λ_w is the optimal quantized gain. The vector $\mathbf{s}_j$ is the winning codebook entree, and $\mathbf{s}_{j_k}$ correspond to the codebook entree obtained by inverting bit k of the codeword associated with $\mathbf{s}_j$.

Table 8 shows the performance results of various methods of encoding the adaptive codebook index. When redundant levels were employed with SDCE, the error detection/correction optimization resulted in mostly error detection. The error recovery strategy used in the optimization was to repeat the previous frame adaptive codebook index. The SDCE performance is slightly lower than that of the parity-bit performance since in the latter procedure the parity bit was again assumed to be immune against channel errors. Additional redundant codewords result in some improvement in performance. SDCE does provide a significant advantage if only a small number of redundant codewords are available as the third row of Table 8 indicates. The associated decrease in clear-channel SSNR performance is minimal; from 9.96 to 9.85 dB.

Code	Bits	Delays	Redun Cdwds	Mean Penalty dB	SSNR dB
Gray	7	21-148	0	-0.21	2.25
SDCE	7	21-148	0	0.25	2.14
SDCE	7	21-118	30	1.49	3.14
SDCE	8	21-148	128	2.88	4.22
Parity	8	21-148	128	2.95	4.28
SDCE	9	21-148	384	3.19	4.62

Table 8. Adaptive codebook index performance

The behavior of the penalty function of the stochastic codebook does not show regularity similar to that of the adaptive codebook index. The only structure results from the overlapping nature of the stochastic codebook. The difference between clear-channel and 1-bit error performance is smaller than that of the adaptive codebook. However, SDCE gives a relatively large improvement over the Gray code since it can take advantage of the irregular structure of the penalty functions. A complete evaluation of the stochastic codebook index performance is given in [6].

IV. CELP WITH SDCE

The error protection bit allocation for the CELP parameters were based on the results of the previous section. Table 9 shows the total bit allocation for our CELP coder. The effective channel bit rate is 4800 bits/s. All of the parameters, regardless of the number of redundant levels used were channel encoded using SDCE. The line spectral frequencies were encoded as pairs as described in Section III. Most of the redundant bits were assigned to the adaptive codebook index and gain parameters since the synthetic speech quality is very sensitive to distortion in the speech periodicity during voiced regions. The rest of the redundant bits

were assigned to the combined encoding of LSF(1) and LSF(2) (LSF(1,2)). The 24 sentence database was used to evaluate the overall coder performance by corrupting the associated CELP bit stream with errors at various rates. Table 10 displays the SSNR performance for this SDCE CELP coder computed over voiced frames in the database. For comparison, the performance of the basic 4233 bits/s coder (Table 1) using Gray code to channel encode the parameters is also shown in Table 10. The results of Table 10 show that, for SDCE CELP, there is a graceful degradation in performance as the error rate is increased from 0% to 1%. At error rates exceeding 1% the performance drops substantially because at such rates the probability of multiple bit errors per parameter is high. Since the SDCE optimization is carried over 1-bit errors, this substantial drop is expected. However, if multiple bit errors are likely, then the optimization process can be extended to cover such errors.

parameter	bits/subfrm (redun. levels)	bits/frm (redun. levels)
LSF(1,2)	N/A	9 (322)
LSF(3,4)	N/A	8 (66)
LSF(5,6)	N/A	7 (32)
LSF(7,8)	N/A	6 (12)
LSF(9,10)	N/A	6 (12)
adap. bk index	9 (384)	36
adap. bk gain	6 (48)	24
stoch. bk index	8 (0)	32
stoch. bk gain	4 (0)	16
Total		144

Table 9. CELP total bit allocation

Error Rate	Basic CELP SSNR (dB)	SDCE CELP SSNR (dB)
0%	9.96	9.96
0.1%	7.29	8.88
0.3%	5.45	7.26
0.5%	3.20	6.00
1.0%	1.42	4.10
2.0%	−0.35	1.69

Table 10. Overall CELP performance

V. CONCLUSIONS

A CELP coder utilizing source-dependent channel encoding was introduced. Unlike conventional error protection methods, SDCE allows for non-integer bit redundancy and strikes an optimal trade-off between error detection and correction. With SDCE, only broad assumptions need to be made about the channel providing, as our experimental results show, a graceful degradation in performance as the channel error rate increases. Although single bit errors were assumed throughout the paper, the extension to include multiple bit errors is straight forward. Also, more sophisticated error recovery strategies can be used in the error detect cases to further improve performance.

REFERENCES

1. R. V. Cox, W. B. Kleijn, and P. Kroon. 1989. Robust CELP coders for noisy backgrounds and noisy channels. In *Proceedings 1989 IEEE Int. Conf. Acoust. Speech Sig. Proc.* Vol. 2, pp. 739-742.

2. W. B. Kleijn, D. J. Krasinski, and R. H. Ketchum. 1988. An efficient stochastically excited linear predictive coding algorithm for high quality low bit rate transmission of speech. *Speech Communication.* Vol. 7, pp. 305-315.

3. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science.* Vol. 220, pp. 671-680.

4. A. E. Al Gamal, L. A. Hemachandra, I. Shperling, and V. K. Wei. 1987. Using simulated annealing to design good codes. *IEEE Trans. Information Theory,* Vol. 33, pp. 116-123.

5. A. H. Gray and J. D. Markel. 1976. Distance measures for speech processing. *IEEE Trans. Acoust., Speech, and Sig. Proc.* Vol. 24, pp. 380-391.

6. W. B. Kleijn. 1990. Source-Dependent Channel Coding for CELP. In *Proceedings 1990 IEEE Int. Conf. Acoust. Speech Sig. Proc.* Vol. 1, pp. 1-4.

Structured Codebook Design in CELP

W.P. LeBlanc and S.A. Mahmoud
Carleton University, Ottawa, Canada, K1S-5B6
Phone: 1–613–788–5740/5753
FAX: 1–613–788–5727
Email: wilf@sce.carleton.ca

ABSTRACT*

Codebook Excited Linear Prediction [1] is a popular analysis by synthesis technique for quantizing speech at bit rates from 4 to 16 kbps. Codebook design techniques to date have been largely based on either random (often gaussian) codebooks, or on known binary or ternary codes which efficiently map the space of (assumed white) excitation codevectors. It has been shown that by introducing symmetries into the codebook, good complexity reduction can be realized with only marginal decrease in performance. In this papers we consider codebook design algorithms for a wide range of structured codebooks.

INTRODUCTION

This paper considers CELP codebook design algorithms for a variety of structured codebooks. A structured codebook has certain properties which enable it to be searched faster than unstructured codebooks. The design algorithms are applied to CELP coders, but are sufficiently general to be applied to other distortion measures as well.

Consider the CELP analysis structure shown in Figure 1. The long term (quantized) inverse filter (with $2q+1$ non-zero taps), $B(z)$, for subframe n is given by:

$$B(z) = 1 - \sum_{k=-q}^{q} b_k z^{-(M+k)} \qquad (1)$$

and the short term (quantized) inverse filter (order p), $A(z)$, for subframe n is given by:

$$A(z) = 1 - \sum_{k=1}^{p} a_k z^{-k} \qquad (2)$$

* This work has been sponsored by the Telecommunications Research Institute of Ontario (TRIO).

The perceptual weighting filter, which attempts to obtain a larger signal to noise ratio in inter-formant regions is given by:

$$W(z) = \frac{A_p(z/\beta)}{A_p(z/\gamma)} \qquad (3)$$

where γ and β are optimized based on subjective measures, and $A_p(z)$ is the optimum unquantized inverse filter (for subframe n).

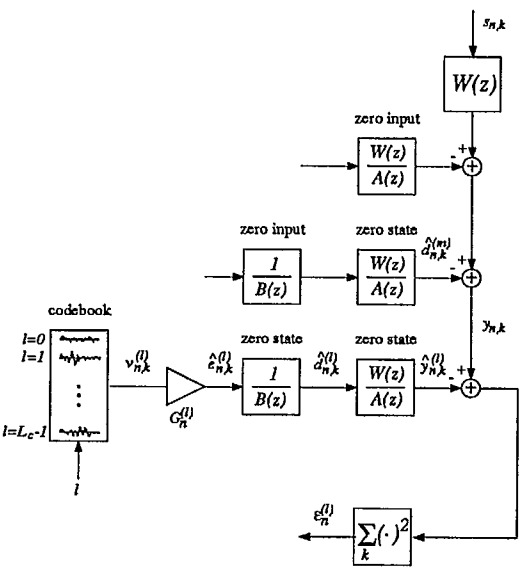


Figure 1: CELP Search Procedure. The codebook dimension, or subframe size is K_c. The index n is over all subframes, and the index k is over all elements of a particular subframe. Thus, $s_{n,k}$ is the k^{th} element of the n^{th} subframe.

Typically, $A_p(z)$ is determined to minimize the open loop residual energy, and $B(z)$ is determined (closed loop) to minimize the noise weighted error before determination of the codebook excitation (the

energy in $y_{n,k}$). The determination of these parameters and complexity reduction techniques based on the structured codebooks is beyond the scope of this paper. The interested reader is referred to [2, 3, 4, 5, 6]. Overlap is often used to reduce block coding edge effects. That is, components of the excitation vector near the end of a subframe have little effect on the current subframe, but may adversely affect future subframes. Overlap considers the influence of these elements by letting the filters ring for K_o samples after the last sample in the excitation vector.

The weighted mean squared error for a particular codebook index l over a subframe (at subframe index n) of dimension K_c with overlap K_o is given by:

$$\begin{aligned}\epsilon_n^{(l)} &= \left|\mathbf{y}_n - \hat{\mathbf{y}}_n^{(l)}\right|^2 \\ &= \mathbf{y}_n^T\mathbf{y}_n - 2G_n^{(l)}\mathbf{y}_n^T\mathbf{H}_n\mathbf{v}^{(l)} \\ &\quad + G_n^{(l)2}\mathbf{v}^{(l)T}\mathbf{H}_n^T\mathbf{H}_n\mathbf{v}^{(l)} \end{aligned} \quad (4)$$

where the $K_c + K_o$ by K_c dimensional lower triangular Toeplitz matrix $\mathbf{H}_n$ represents the zero state filtering operation (of $W(z)/A(z)$). The l^{th} excitation (column) vector $\mathbf{v}^{(l)}$ is of dimension K_c, and the (column) vectors $\mathbf{y}_n$ and $\hat{\mathbf{y}}_n^{(l)}$ $\left(= G_n^{(l)}\mathbf{H}_n\mathbf{v}^{(l)}\right)$ are of dimension $K(= K_c + K_o)$.

The codebook design algorithms are all based on the Generalized Lloyd Algorithm (GLA) [7, 8, 9] and require a sufficiently rich training sequence to design the codebook. Due to the long memory in $1/B(z)$, the algorithm is not guaranteed to converge to a local minimum. That is, the set of training vectors $\mathcal{T} = \{\mathbf{y}_n\}$ changes from one iteration to the next. The problem arises because (for simplicity) we assume the training vectors do not depend on the codevectors. Due to the long and short term predictor memory this is not the case. In practice, convergence is similar to the GLA, although the average weighted mean squared error has been observed to increase (slightly) after some iterations.

The optimum codebook is defined as that which minimizes Equation 4 over the whole training sequence. We minimize:

$$\begin{aligned}\bar{\epsilon} &= \sum_{n=0}^{N-1}\epsilon_n^{(l_n)} \\ &= \sum_{n=0}^{N-1}\Big(\mathbf{y}_n^T\mathbf{y}_n - 2G_n\mathbf{y}_n^T\mathbf{H}_n\mathbf{v}^{(l_n)} \\ &\quad + G_n^2\mathbf{v}^{(l_n)T}\mathbf{H}_n^T\mathbf{H}_n\mathbf{v}^{(l_n)}\Big) \end{aligned} \quad (5)$$

The index n is over all training vectors ($\mathbf{y}_n$), l_n is the optimum codebook index for training vector (or subframe) n, $\mathbf{v}^{(l_n)}$ is the optimum codevector (for subframe n) and G_n is the optimum gain for codevector $\mathbf{v}^{(l_n)}$ ($G_n = G_n^{(l_n)}$). The codebook design techniques are all based on minimization of Equation 5. All design techniques assume training vector $\mathbf{y}_n$ is not a function of the current, or past codevectors.

In Section 2 we consider general codebook design. The codebook is given by L_c distinct K_c dimensional codevectors. This section also considers codebooks in which the codevectors have many zero elements.

GENERAL CODEBOOK DESIGN

We now discuss techniques whereby near optimal codebooks may be design for general, or sparse codebooks. The technique is based on a vector quantizer design algorithm using the noise weighted mean squared error distortion measure. Due to the influence of previous codevectors on future codevectors (via the long term predictor memory), only suboptimal codebooks may be designed, (the error is not guaranteed to decrease continually to a local optimum). In practice, the average distortion usually decreases until a *local optimum* is found, then oscillates slowly in the vicinity of that local optimum.

Unstructured Codebooks

The goal is to minimize Equation 5 over all possible codebooks of size L_c and dimension K_c. Given a training sequence of N speech vectors $\mathcal{S} = \{\mathbf{s}_n\}$, and an initial codebook $\mathcal{C}^{(0)} = \{\mathbf{v}^{(l)}\}$, we analyze the vectors using the CELP structure to obtain the training set $\mathcal{T}^{(0)} = \{\mathbf{y}_n\}$. Essentially, we use the initial codebook to partition the training sequence ($\mathcal{T}^{(0)}$) into L_c cells, or regions $\mathcal{R}^{(j)}$ according to the nearest neighbour search, and compute new centroids (or codevectors) for the regions. Cell j is comprised of those subframes which have $\mathbf{v}^{(l_n)} = \mathbf{v}^{(j)}$ (the optimum codebook index at time n is j). Equation 5 can then be split up into L_c terms, one term for each particular cell:

$$\begin{aligned}\bar{\epsilon} = &\sum_{n\in\mathcal{R}^{(0)}}\left|\mathbf{y}_n - G_n\mathbf{H}_n\mathbf{v}^{(0)}\right|^2 + \\ &\sum_{n\in\mathcal{R}^{(1)}}\left|\mathbf{y}_n - G_n\mathbf{H}_n\mathbf{v}^{(1)}\right|^2 + \ldots + \\ &\sum_{n\in\mathcal{R}^{(L_c-1)}}\left|\mathbf{y}_n - G_n\mathbf{H}_n\mathbf{v}^{(L-1)}\right|^2 \end{aligned} \quad (6)$$

where the summation indicates summation over the region in which all codevectors are identical. Minimization of Equation 6 is equivalent to minimizing each term, since a particular codevector only influences the summation in its region. Furthermore, in

each region (j), $v^{(l_n)}$ does not depend on n (since $j = l_n$. Thus we minimize (with respect to $v^{(j)}$):

$$\epsilon^{(j)} = \sum_{n \in \mathcal{R}^{(j)}} \left(\mathbf{y}_n^T \mathbf{y}_n - 2G_n \mathbf{y}_n^T \mathbf{H}_n \mathbf{v}^{(j)} + G_n^2 \mathbf{v}^{(j)T} \mathbf{H}_n^T \mathbf{H}_n \mathbf{v}^{(j)} \right) \quad (7)$$

for each region j, $0 \leq j < L_c$. Since $v^{(l_n)} = v^{(j)}$ is a constant for each region and does not depend on the index n, we may write:

$$\begin{aligned} \epsilon^{(j)} &= \sum_{n \in \mathcal{R}^{(j)}} \mathbf{y}_n^T \mathbf{y}_n - 2 \left(\sum_{n \in \mathcal{R}^{(j)}} G_n \mathbf{y}_n^T \mathbf{H}_n \right) \mathbf{v}^{(j)} \\ &\quad + \mathbf{v}^{(j)T} \left(\sum_{n \in \mathcal{R}^{(j)}} G_n^2 \mathbf{H}_n^T \mathbf{H}_n \right) \mathbf{v}^{(j)} \\ &= \bar{\sigma}_y^{(j)2} - 2\bar{\mathbf{c}}^{(j)} \mathbf{v}^{(j)} + \mathbf{v}^{(j)T} \bar{\mathbf{R}}^{(j)} \mathbf{v}^{(j)} \end{aligned} \quad (8)$$

where

$$\bar{\sigma}_y^{(j)2} = \sum_{n \in \mathcal{R}^{(j)}} \mathbf{y}_n^T \mathbf{y}_n \quad (9)$$

$$\bar{\mathbf{c}}^{(j)} = \sum_{n \in \mathcal{R}^{(j)}} G_n \mathbf{H}_n^T \mathbf{y}_n \quad (10)$$

and:

$$\bar{\mathbf{R}}^{(j)} = \sum_{n \in \mathcal{R}^{(j)}} G_n^2 \mathbf{H}_n^T \mathbf{H}_n \quad (11)$$

It can easily be shown (differentiate with respect to $\mathbf{v}^{(j)}$), that to minimize equation 8 we choose:

$$\mathbf{v}^{(j)} = \left(\bar{\mathbf{R}}^{(j)} \right)^{-1} \bar{\mathbf{c}}^{(j)} \quad (12)$$

which can be efficiently accomplished by using Choleski decomposition. This is performed over all j, $0 \leq j < L_c$.

We will now have a new codebook ($\mathcal{C}^{(1)}$), which can be used in the CELP analysis structure to obtain the training set $\mathcal{T}^{(1)}$. Unlike typical VQ design techniques, the training set $\mathcal{T}^{(1)}$ will not be the same as $\mathcal{T}^{(0)}$. The above design algorithm is just a simple extension of the GLA for a CELP type distortion based on the above assumptions.

Sparse Codebook Design

To design sparse codebooks, we essentially want to minimize Equation 8 (for each j), given the constraint that there are a large number of zero values in the codevectors. We use the multipulse sequential approach (for complexity reasons), and first compute the optimum pulse location and gain (assuming one non-zero value in the codevector) to minimize Equation 8.

We then iteratively add another pulse location, and so on, until we have the desired number of non-zero pulses in the codevector. After each iteration, the pulse amplitudes are re-optimized.

To minimize:

$$\epsilon^{(j)} = \bar{\sigma}_y^{(j)2} - 2\bar{\mathbf{c}}^{(j)T} \mathbf{v}^{(j)} + \mathbf{v}^{(j)T} \bar{\mathbf{R}}^{(j)} \mathbf{v}^{(j)} \quad (13)$$

for the first pulse position (k_0) and amplitude ($v_{k_0}^{(j)}$) we minimize:

$$\epsilon^{(j,1)} = \bar{\sigma}_y^{(j)2} - 2\bar{c}_{k_0}^{(j)} v_{k_0}^{(j)} + v_{k_0}^{(j)} \bar{R}_{k_0 k_0}^{(j)} v_{k_0}^{(j)} \quad (14)$$

which has solution (for a particular position k_0):

$$v_{k_0}^{(j)} = \frac{\bar{c}_{k_0}^{(j)}}{\bar{R}_{k_0 k_0}^{(j)}} \quad (15)$$

The first position is computed by trying all locations, and choosing that which minimizes Equation 14.

Assuming the first pulse location is fixed, the second location is chosen to minimize:

$$\begin{aligned} \epsilon^{(j,2)} = \bar{\sigma}_y^{(j)2} &- 2\bar{c}_{k_0}^{(j)} v_{k_0}^{(j)} - 2\bar{c}_{k_1}^{(j)} v_{k_1}^{(j)} + v_{k_0}^{(j)} \bar{R}_{k_0 k_0}^{(j)} v_{k_0}^{(j)} \\ &+ v_{k_1}^{(j)} \bar{R}_{k_1 k_1}^{(j)} v_{k_1}^{(j)} + 2v_{k_1}^{(j)} \bar{R}_{k_1 k_0}^{(j)} v_{k_0}^{(j)} \end{aligned} \quad (16)$$

If $v_{k_0}^{(j)}$ is not to be modified (as part of the search, for complexity reasons), then:

$$v_{k_1}^{(j)} = \frac{\bar{c}_{k_1}^{(j)} - v_{k_0}^{(j)} \bar{R}_{k_1 k_0}^{(j)}}{\bar{R}_{k_1 k_1}^{(j)}} \quad (17)$$

and:

$$\epsilon^{(j,2)} = \bar{\sigma}_y^2 - \frac{\left(\bar{c}_{k_1}^{(j)} - v_{k_0}^{(j)} \bar{R}_{k_1 k_0}^{(j)} \right)^2}{\bar{R}_{k_1 k_1}^{(j)}} \quad (18)$$

The mean squared error is minimized by maximizing the square of the second term. At the end of the search for the second pulse position, the amplitudes of the first and second pulse positions can be optimized by minimizing Equation 16 with respect to the unknown amplitudes $v_{k_0}^{(j)}$ and $v_{k_1}^{(j)}$.

In general, the n^{th} pulse position is given by computing the minimum over all pulse locations (k_n) of:

$$\epsilon^{(j,n)} = \bar{\sigma}_y^2 - \frac{\left(\bar{c}_{k_n}^{(j)} - \sum_{i=0}^{n-1} v_{k_i}^{(j)} \bar{R}_{k_n k_i}^{(j)} \right)^2}{\bar{R}_{k_n k_n}^{(j)}} \quad (19)$$

and the pulse amplitudes are optimized by solving (by Choleski Decomposition):

$$\tilde{\mathbf{v}}^{(j)} = \left(\tilde{\mathbf{R}}^{(j)}\right)^{-1} \tilde{c}^{(j)} \tag{20}$$

where:

$$\tilde{\mathbf{v}}^{(j)} = \begin{bmatrix} v_{k_0}^{(j)} \\ v_{k_1}^{(j)} \\ \vdots \\ v_{k_n}^{(j)} \end{bmatrix}, \quad \tilde{\mathbf{c}}^{(j)} = \begin{bmatrix} \bar{c}_{k_0}^{(j)} \\ \bar{c}_{k_1}^{(j)} \\ \vdots \\ \bar{c}_{k_n}^{(j)} \end{bmatrix} \tag{21}$$

and:

$$\tilde{\mathbf{R}}^{(j)} = \begin{bmatrix} \bar{R}_{k_0k_0}^{(j)} & \bar{R}_{k_0k_1}^{(j)} & \cdots & \bar{R}_{k_0k_n}^{(j)} \\ \bar{R}_{k_1k_0}^{(j)} & \bar{R}_{k_1k_1}^{(j)} & \cdots & \bar{R}_{k_1k_n}^{(j)} \\ \vdots & \vdots & \ddots & \bar{R}_{k_{n-1}k_n}^{(j)} \\ \bar{R}_{k_nk_0}^{(j)} & \bar{R}_{k_nk_1}^{(j)} & \cdots & \bar{R}_{k_nk_n}^{(j)} \end{bmatrix} \tag{22}$$

The sequential multipulse search procedure is inherently suboptimal since it does not try all combinations of pulse positions. However, a full search technique is prohibitively complex. Rather than keeping a single pulse position from stage to stage however, it is clearly better to keep M_l *survivors* after each stage (for the sequential optimization described above, $M_l = 1$).

SHIFT SYMMETRIC CODEBOOKS

A shift symmetric codebook is defined as a codebook in which a single codevector has all but t elements in common with the next codevector. The l^{th} codevector can be written $\mathbf{v}^{(l)} = \mathbf{S}^{(l)}C$ where C is a $K_c + t(L_c - 1)$ dimensional column vector (the codebook) and $\mathbf{S}^{(l)}$ is a K_c by $K_c + t(L_c - 1)$ dimensional *shifting matrix* with ones on the tl^{th} (upper) diagonal, and zeros elsewhere:

$$\mathbf{S}^{(l)} = [\mathbf{0}_{K_c,tl} | \mathbf{I}_{K_c,K_c} | \mathbf{0}_{K_c,L_c-tl}] \tag{23}$$

where $\mathbf{0}_{K_c,n}$ is a K_c by n matrix of zeros, and $\mathbf{I}_{K_c,K_c}$ is the K_c by K_c identity matrix. With $t = K_c$ we obtain the general codebook discussed above.

Shift symmetric codebooks present a problem since elements from a single codevector are included in possibly many other codevectors. Thus, the design algorithm must reflect this property. A modification to the Vector Quantization (VQ) design algorithm was utilized to account for the shift symmetric codebooks. We have:

$$\bar{\epsilon} = \sum_{n=0}^{N-1} \left(\mathbf{y}_n - G_n \mathbf{H}_n \mathbf{v}^{(l_n)}\right)^2 \tag{24}$$

Again we assume we have an initial codebook, but rather than partitioning the codebook into L_c cells or regions using the nearest neighbour, minimum distortion search criteria, we simply substitute $\mathbf{v}^{(l)} = \mathbf{S}^{(l)}C$ into Equation 24 which yields:

$$\begin{aligned} \bar{\epsilon} &= \sum_{n=0}^{N-1} \left(\mathbf{y}_n - G_n \mathbf{H}_n \mathbf{S}^{(l_n)} C\right)^2 \\ &= \bar{\sigma}_y^2 - 2\bar{\mathbf{c}}^T C + C^T \bar{\mathbf{R}} C \end{aligned} \tag{25}$$

where:

$$\bar{\sigma}_y^{(j)2} = \sum_{n=0}^{N-1} \mathbf{y}_n^T \mathbf{y}_n \tag{26}$$

$$\bar{\mathbf{c}} = \sum_{n=0}^{N-1} G_n \mathbf{S}^{(l_n)T} \mathbf{H}_n^T \mathbf{y}_n \tag{27}$$

(a $K_c + t(L_c - 1)$ dimensional column vector) and:

$$\bar{\mathrm{R}} = \sum_{n=0}^{N-1} G_n^2 \mathbf{S}^{(l_n)T} \mathbf{H}_n^T \mathbf{H}_n \mathbf{S}^{(l_n)} \tag{28}$$

(a square $K_c + t(L_c - 1)$ dimensional band matrix).

The codebook is thus given by $C = \bar{\mathbf{R}}^{-1}\bar{\mathbf{c}}$ which, again, can be efficiently computed using Choleski Decomposition. Further storage and computational savings can be realized by using the fact that $\bar{\mathbf{R}}$ is a band matrix. Computation of Equation 28 and 27 can be greatly simplified by exploiting the structure in the shifting matrix.

Sparse shift symmetric codebooks can be designed by applying a multipulse procedure to Equation 25, as was done with general sparse codebooks.

VSELP CODEBOOK DESIGN

Let $L_c = 2^M$, where M is the number of bits in the codebook index. The VSELP excitation can be given by $\mathbf{v}^{(l)} = \tilde{C}\mathbf{b}^{(l)}$ where $\tilde{C}$ is the VSELP codebook (a K_c by M dimensional matrix), and $\mathbf{b}^{(l)}$ (an M dimensional column vector with elements ± 1) is the l^{th} codeword. Alternatively, yet equivalently, the excitation can be written as $\mathbf{v}^{(l)} = \mathbf{B}^{(l)}C$ where C is a K_cM dimensional column vector (containing the stacked columns of $\tilde{C}$) and $\mathbf{B}^{(l)}$ is a K_c by K_cM dimensional Toeplitz matrix, with the first row having elements $b_k^{(l)}$ in positions B_{0,kK_c}

Over the training sequence, we may write:

$$\bar{\epsilon} = \sum_{n=0}^{N-1} \left(\mathbf{y}_n - G_n \mathbf{H}_n \mathbf{v}^{(l_n)}\right)^2 \tag{29}$$

Substituting $\mathbf{v}^{(l)} = \mathbf{B}^{(l)}C$ into 29 leads to:

$$\bar{\epsilon} = \sum_{n=0}^{N-1} \left(\mathbf{y}_n - G_n \mathbf{H}_n \mathbf{B}^{(l_n)} C\right)^2$$
$$= \bar{\sigma}_y^2 - 2\bar{\mathbf{c}}^T C + C^T \bar{\mathbf{R}} C \quad (30)$$

where:

$$\bar{\sigma}_y^{(j)2} = \sum_{n=0}^{N-1} \mathbf{y}_n^T \mathbf{y}_n \quad (31)$$

$$\bar{\mathbf{c}} = \sum_{n=0}^{N-1} G_n \mathbf{B}^{(l_n)T} \mathbf{H}_n^T \mathbf{y}_n \quad (32)$$

(a $K_c M$ dimensional column vector) and:

$$\bar{\mathrm{R}} = \sum_{n=0}^{N-1} G_n^2 \mathbf{B}^{(l_n)T} \mathbf{H}_n^T \mathbf{H}_n \mathbf{B}^{(l_n)} \quad (33)$$

is a $K_c M$ by $K_c M$ dimensional matrix.

The VSELP (stacked) codebook C is computed by solving (again by Choleski Decomposition):

$$C = \bar{\mathrm{R}}^{-1} \mathbf{c} \quad (34)$$

Computation of Equation 32 and 33 can be greatly simplified by exploiting the structure of $\mathbf{B}^{(l_n)}$.

RESULTS

In this section we present results of computer simulations conducted on a 10 minute speech database and a 30 second speech database. Codebooks were trained on the large database and the performance was computed on both databases. Objective measures of performance included the segmental signal to noise ratio defined by:

$$\mathrm{SEGSNR} = \frac{1}{N} \sum_n 10 \log_{10} \left(\frac{|\mathbf{s}_n|^2}{|\mathbf{s}_n - \hat{\mathbf{s}}_n|^2} \right) \quad (35)$$

where

$$\hat{\mathbf{s}}_n \quad (36)$$

is the synthesiszed (20 msec) speech vector and the noise weighted signal to noise ratio defined by:

$$\mathrm{NWSNR} = \frac{\sum_n \mathbf{s}^T \mathbf{s}}{\bar{\epsilon}} \quad (37)$$

In our examples, the CELP coder used subframe dimensions of 40 samples, 2 samples of overlap (which was determined to be near optimal), and frame sizes of 160 samples. The inverse filter ($A(z)$) was determined at the frame rate using the autocorrelation method and quantized using interframe vector linear prediction of the line spectrum pairs followed by scalar quantization of the error [10]. The long term predictor was optimized closed loop to minimize the closed loop weighted mean squared error. The pitch period was constrained to be in the range from 21 to 148 samples. The general codebooks used the autocorrelation method discussed in [2] (which does contain certain approximations). Our experiments with shift symmetric codebooks considered $t = 1$ only, (and no approximations were used). The design of the sparse codebooks used the tree searched multipulse search procedure outlined above, with $M_l = 128$. The sparse shift symmetric codebooks had more than 90% zero samples (52 non-zero samples in a 512 level codebook).

Table 1 displays the performance of random gaussian codebooks for various codebook sizes (L_c).

Codebook Size (bits)	NWSNR (SEGSNR)
7	14.39 (17.13) dB
8	14.84 (17.75) dB
9	15.24 (18.21) dB
10	15.59 (18.66) dB
11	16.10 (19.26) dB

Table 1: Performance of random gaussian codebooks of various sizes (30 second database). The values are accurate (with 95% confidence) to within 0.1 dB.

By comparison, a 9 bit random gaussian shift symmetric codebook obtained a noise weighted SNR of 15.05 dB (SEGSNR=18.11 dB) and a 9 bit random VSELP codebook obtained a NWSNR of 13.92 dB (SEGSNR=16.83 dB). Again the values are accurate (with 95% confidence) to within 0.1 dB.

Trained 9 bit general codebooks, sparse shift symmetric, and VSELP codebooks (using the design techniques discussed above) obtained performance both inside and outside of the training sequence as shown in Tables 2 and 3.

Outside the training sequence the performance (NWSNR) of sparse shift symmetric codebooks is within 0.2 dB of the general codebooks which is within the 95% confidence intervals. Inside the training sequence the performance of the general codebook is approximately 0.7 dB better than the sparse shift symmetric codebooks. Imposing structure limits the performance inside the training sequence but has little effect outside the training sequence in this

Codebook	NWSNR (SEGSNR)
General	14.67 (19.40) dB
Sparse Shift Symmetric	13.95 (18.84) dB
VSELP	13.43 (18.54) dB

Table 2: Performance of trained general, sparse shift symmetric, and VSELP codebooks inside the training sequence (10 minute database). The values are accurate (with 95% confidence) to within 0.1 dB.

Codebook	NWSNR (SEGSNR)
General	15.75 (18.79) dB
Sparse Shift Symmetric	15.60 (18.48) dB
VSELP	14.72 (17.84) dB

Table 3: Performance of trained general, sparse shift symmetric, and VSELP codebooks outside the training sequence (30 second database). The values are accurate (with 95% confidence) to within 0.1 dB.

instance. The VSELP codebooks appear to have too much structure and performance suffers by more than 0.8 dB both inside and outside the training sequence. However, the performance of VSELP improves by more then 0.8 dB after codebook design (outside the training sequence).

The general and sparse shift symmetric trained 9 bit codebooks have objective performance virtually equivalent to the untrained 10 bit random codebooks. Thus, for equivalent objective performance half the number of levels in the codebook are required, resulting in a lower data rate and a lower complexity.

CONCLUSIONS

This paper considered the CELP codebook design problem for a variety of structured codebooks. It was determined that a savings of one bit per vector could be realized with virtually no decrease in the objective measures while decreasing complexity by a factor of two. For fixed codebook sizes, improvements of more than 0.5 dB were observed with no increase in computational complexity. It was observed that Vector Sum Excited Linear Prediction had too much structure, and performance was noticeably inferior to the general or sparse shift symmetric codebooks.

The structured codebook design techniques are relatively simple, and only require Choleski Decomposition or a relatively straightforward multipulse algorithm. The design algorithms were applied to CELP coders, but are sufficiently general to be applied to other distortion measures as well.

REFERENCES

[1] M. Schroeder and B. Atal, "Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 1985.

[2] I. Trancoso and B. Atal, "Efficient Procedures for Finding the Optimum Innovation in Stochastic Coders," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1986.

[3] G. Davidson and A. Gersho, "Complexity Reduction Methods for Vector Excitation Coding," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1986.

[4] G. Davidson, M. Yong, and A. Gersho, "Real-Time Vector Excitation Coding of Speech at 4800 bps," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1987.

[5] W. Kleijn, D. Krasinski, and R. Ketchum, "Improved Speech Quality and Efficient Vector Quantization in SELP," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 1, pp. 155–158, April 1988.

[6] I. Gerson and M. Jasiuk, "Vector Sum Excited Linear Prediction (VSELP)," *IEEE Speech Coding Workshop*, 1989.

[7] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. IT-28, pp. 129–137, March 1982.

[8] Y. Linde, A. Buzo, and R. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications*, vol. COM-28, pp. 84–95, January 1980.

[9] J. Makhoul, S. Roucos, and H. Gish, "Vector Quantization in Speech Coding," *Proceedings of the IEEE*, vol. 73, pp. 1551–1588, November 1985.

[10] M. Yong, G. Davidson, and A. Gersho, "Encoding of LPC Spectral Parameters Using Switched-Adaptive Interframe Vector Prediction," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 1, pp. 402–405, April 1988.

A 4800 bps CELP Vocoder with an Improved Excitation

Hisham Hassanein, André Brind'Amour and Karen Bryden
Department of Communications
Communications Research Centre
3701 Carling Ave., P.O.Box 11490, Station H
Ottawa, Ont., CANADA. K2H 8S2.
Phone : 613-998-2462
Fax : 613-990-7987

ABSTRACT

The Stochastic or Code-Excited Linear Predictive Coder (CELP) is among the promising candidates for producing good quality speech at low bit rates. However, the speech quality produced suffers from perceived roughness. Many researchers have used pole-zero postfilters to mask the roughness at the output of the synthesis filter. Although the postfilters are effective in masking the noise at low bit rates, they produce spectral distortions. In this paper[1], it is proposed to improve the speech quality by introducing two modifications to the fixed stochastic codebook. In the first modification, the stochastic codebook is used only when the long-term corrrelations are low. Otherwise a pulse-like codebook is selected. In the second modification, the selected codebook output is weighted using an adaptive spectral shaping procedure. These two modifications have been incorporated in a 4800 bps CELP coder and have resulted in a perceptually improved vocoded speech.

1. INTRODUCTION

Code-Excited Linear Predictive Coding (CELP) was first introduced by Atal and Schroeder in 1984 [1]. This algorithm represented a breakthrough for achieving good quality speech at rates below 4800 bps. Its major drawback was its computational complexity, which was prohibitive for real-time applications. Since then, many speech researchers, recognizing the tremendous potential of this algorithm, experimented with different methods for simplifying the algorithm. At the same time, the DSP chips became more powerful and floating point processors were introduced. These activities have resulted in the implementation of CELP (or similar algorithms developed later like Vector Adaptive Predictive Coders (VAPC) [2] or Vector Sum Excited Linear Predictive Coder (VSELP) [3]) using a single DSP chip. In this paper, a brief description of a *reference* CELP coder is described, followed by two modifications proposed for the stochastic codebook. The modified CELP algorithm has been implemented on a single TMS320C25 DSP chip and operates in full-duplex at the rate of 4800 bps.

2. REFERENCE CODER DESCRIPTION

The reference coder's synthesizer is depicted in Fig. 1. The excitation to the synthesizer e(n) given by (1) is the linear combination of a vector from the stochastic codebook x(K+n) and a vector from the adaptive codebook e(n-L)

$$e(n) = Gx(K+n) + \beta e(n-L), \quad n=0,59 \qquad (1)$$

where K and L are the optimal indices of the stochastic and adaptive codebooks respectively, and G and β the gains of the vectors from the respective codebooks. During voiced sounds, the lag coefficient β is close to unity. In this case most of the contribution to the excitation comes from the adaptive codebook (which represents the excitation history) and the stochastic codebook entry acts only as a correcting term. For lower bit rates the stochastic codebook contribution may cause noisy synthetic speech for two reasons. First, the LPC model with a limited number of coefficients (10) fails to entirely remove all the short-term correlations from the speech input signal. The result is an intelligible residual with a non-flat spectrum. This may suggest that the stochastic codebook may be inadequate, particularly for a

[1]This project was sponsered by DCEM/DRDCS, code number 0417U.

small codebook size (60 entries in our implementation) giving rise to rough synthetic speech. The second reason for the synthetic speech roughness may be attributed to the use of the stochastic codebook during voiced frames. In order to mask the noise, the reference coder uses a formant and lag postfilters [7] at the output of the synthesizer, in order to compensate for the two defficiencies listed above. The postfilters were reported to be effective in masking the noise but have the disadvantage of creating spectral distortions, particularly when the vocoders are used in tandem. In order to improve the synthesized speech quality without introducing the inherent spectral distortions produced by the postfilters, two modifications to the stochastic codebook are introduced, namely the use of a pulse-like codebook instead of the stochastic codebook during segments with long-term correlation coefficients larger than a certain threshold, and the spectral weighting of the output of the codebooks. These two modifications are described in the following sections.

3. VOICED/UNVOICED CODEBOOKS

It was observed that the synthetic speech produced by Pitch-Excited LPC coders sounds smoother (although less natural) than the reference coder. This may be attributed to the inadequacy of the stochastic codebook during voiced sounds. In order to smooth the CELP synthetic speech, we tried to imitate the excitation model of the Pitch-Excited coder. Thus, instead of using the stochastic codebook during voiced subframes, a pulse-like codebook is used. The number of pulses in the codebook is a function of the lag (which has been determined earlier in the analysis process) and the subframe size. The number of the codebook indices is chosen to be equal to the subframe size. For lag values larger than the subframe size, each codebook vector contains only one non-zero pulse positioned at a location equal to the codebook index. For lag values less than the subframe size, the first S-L+1 vectors, where S is the subframe size and L is the lag, contain two pulses each (with equal amplitudes) separated by the lag value and the location of the first pulse is equal to the codebook index. The remaining vectors contain only one non-zero pulse. The dual codebook is shown in Fig. 2, where if $|\beta|$ is larger than a certain threshold T, (T>0), the speech is considered voiced, and consequently the pulse-like codebook is chosen. It is to be noted that the threshold is valid for values of $|\beta|$ close to unity, where speech is voiced, as well as large values of $|\beta|$ which indicates voicing onset. Otherwise, the stochastic codebook is used. The output of either codebook is frequency weighted as described in the following section.

4. WEIGHTING OF THE DUAL CODEBOOK

As mentioned earlier, the residual signal exhibits some short-term correlations which have not been successfully removed by LPC analysis. Intuitively, it makes sense to spectrally weight the dual codebook adaptively so that its spectrum will be as close as possible to that of the residual (after removing the long-term correlations). Several researchers have tried to apply spectral shaping to the excitation. In [4], an all-pole spectral shaping was applied to the excitation of a Pitch-Excited Linear Predictive Coder. In [5] a pole-zero weighting filter was used to shape the excitation of a CELP coder. The result was a perceived quality comparable to that obtained with the postfilter without the inherent distortion introduced by the latter. We have chosen a filter similar to [4]. The dual codebook is adaptively weighted every frame by

$$W(z) = \frac{1}{1+F\sum_{i=1}^{10} a_i z^{-i}} \qquad (2)$$

where the a_i's are the LPC filter coefficients and F is a modulus reduction factor given by

$$F = \alpha\prod_{i=1}^{10}(1 - k_i^2), \quad 0<=\alpha<=1 \qquad (3)$$

In (3), the k_i's are the reflection coefficients and α is a scaling factor. When the LPC prediction is efficient as is the case for front vowels, murmurs and nasals, the residual has a relatively flat spectrum and F is small. In this case W(z) acts as an all-pass filter and the dual codebook is minimally weighted. For speech sounds where the LPC prediction is not as good, F is relatively large and W(z) gives the dual codebook a spectral shape similar to that of the residual.

4. EXPERIMENTAL RESULTS

The two modifications described above have been implemented in the TMS320C25 code. The objective measure chosen to quantify the results is the segmental Signal to Weighted Noise ratio (SWNR) defined as

$$SWNR = \frac{\sum_{n=0}^{N-1} s^2(n)}{\sum_{n=0}^{N-1} e_w{}^2(n)} \qquad (4)$$

where s(n) is the original input speech signal, N is the frame size in samples, and $e_w(n)$ is the weighted error signal. The latter is obtained by frequency weighting the difference between the input and synthesized speech signals, by the conventional CELP weighting filter CW(z) defined by

$$CW(z) = \frac{1 + \sum_{i=1}^{10} a_i z^{-i}}{1 + \sum_{i=1}^{10} \gamma^i a_i z^{-i}}, \ \gamma = 0.75 \qquad (5)$$

Informal listening tests using the real-time hardware were performed on a variety of input speech samples. Both of the modifications introduced above produced synthetic speech which may be described as cleaner and fuller than that produced by the reference coder. However, this improvement was not translated into a significant improvement in the average segmental SWNR. Over a limited database of 999 frames, the improvement was only 0.27 dB in the case of the weighted codebook and 0.24 dB in the case of the dual codebook. However, the improvement seems to be localized and went as high as 5 dB in the case of the dual codebook and 7 dB in the case of the weighted codebook. The combined use of the two modifications resulted in increases as high as 10 dB in some frames. In Fig. 3a, 3 seconds of input speech from a male speaker are shown and the segmental SWNR of the corresponding synthesized speech is shown in Fig. 3b. The difference in segmental SWNR between the synthesized speech for each modification and the reference signal are shown in Figs. 3c and 3d. In can be noted from Fig. 3d that the largest increases in SWNR between the coder using the dual codebook and the reference coder tend to occur at the beginning of transitions.

5. CONCLUSIONS

In this paper, two modifications to the stochastic codebook were introduced, namely the adaptive frequency weighting of the stochastic codebook and the use of a dual codebook. This has resulted in an improved speech quality of the vocoder. Although these modifications resulted in an increase of up to 10 dB, the improvements seem to be localized and were not translated into a significant increase in the average segmental SWNR.

REFERENCES

1. Atal B.S. and M.R. Schroeder 1984. Stochastic Coding of Speech at Very Low Bit Rates. *Proc. of ICC, Amsterdam*, pp. 1610-1613.

2. Chen J.H. and A. Gersho 1987. Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering. *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, p. 51.3.

3. Gerson I.A. and M.A. Jasiuk 1990. Vector Sum Excited Linear Prediction (VSELP) Speech Coding at 8 kbps. *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, pp. 461-464.

4. Kang G.S. and S.S. Everett 1985. Improvement of the Excitation Source in the Narrow-Band Linear Prediction Vocoder. *IEEE Trans. Acoust., Speech and Signal Processing*, ASSP-33(2).

5. Kroon P. and B.S. Atal 1988. Strategies for Improving the Performance of CELP Coders at Low Bit Rates. *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, pp. 151-154.

6. Kroon P. and E.F. Deprettere Feb 1988. A Class of Analysis-by-Synthesis Predictive Coders or High Quality Speech Coding at Rates between 4.8 and 16 kb/s. *IEEE J. on Selected Areas in Communications*, SAC-5.

7. Kroon P. and B.S. Atal 1987. Quantization Procedures for the Excitation in CELP coders. *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, pp. 1649-1652.

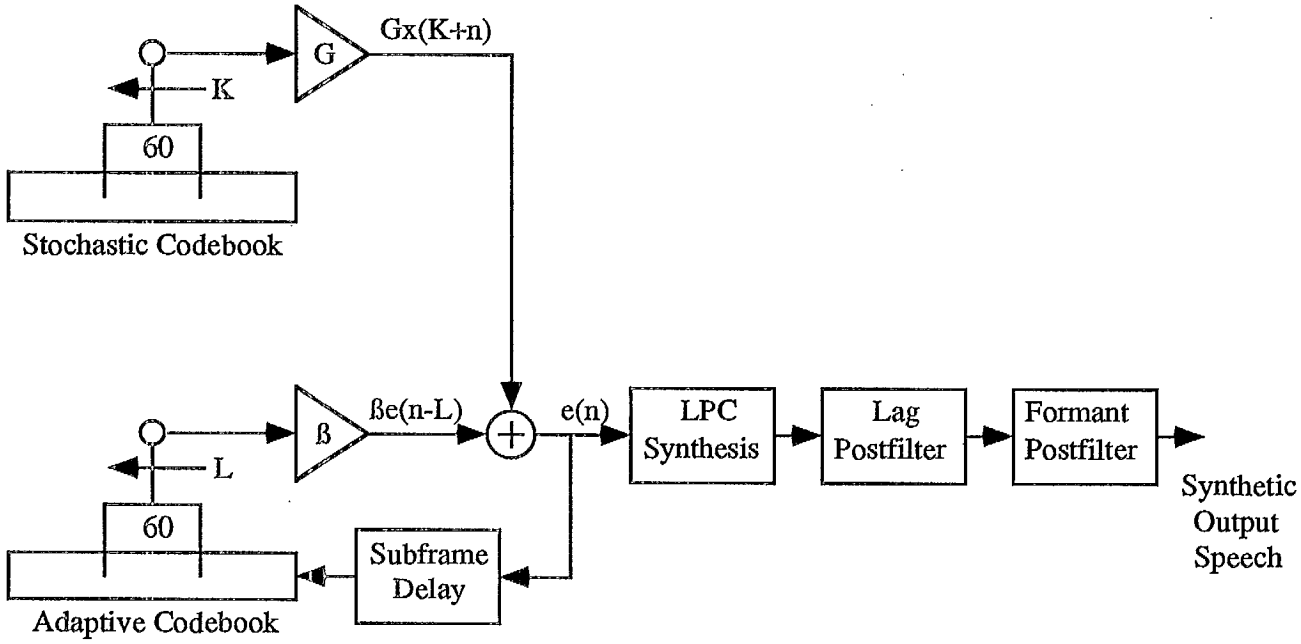


Figure 1. Synthesis structure of the reference coder.

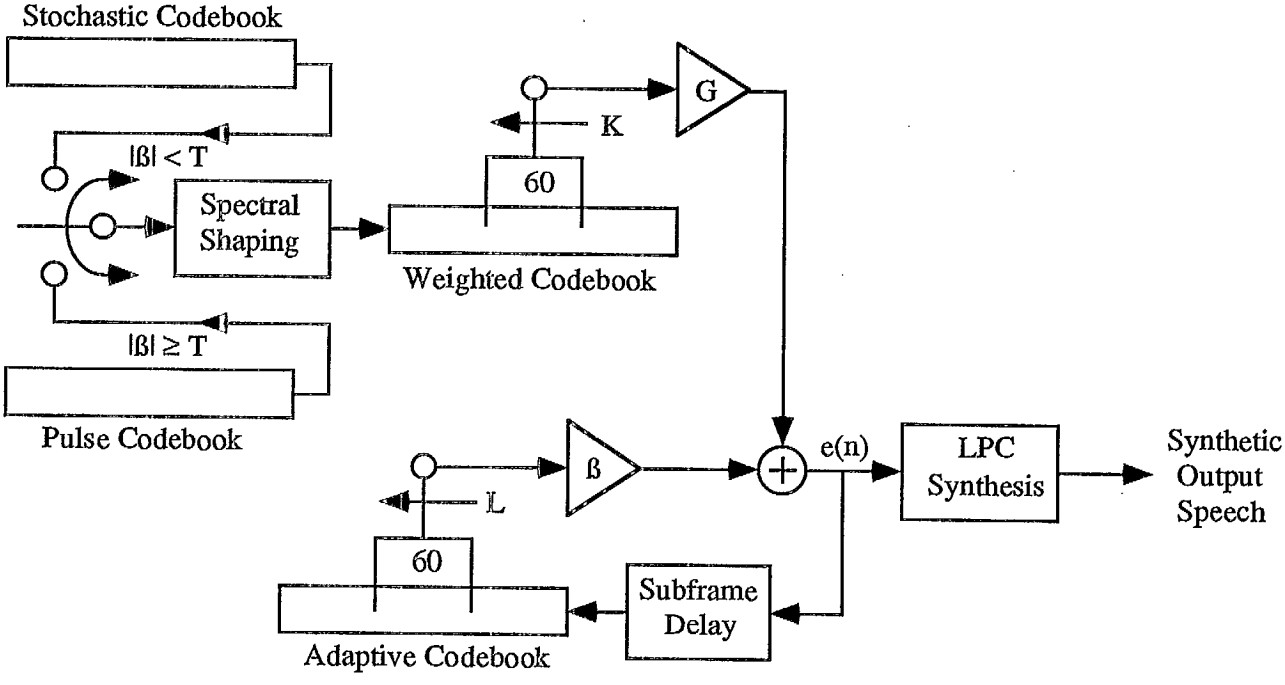


Figure 2. Modified synthesis structure.

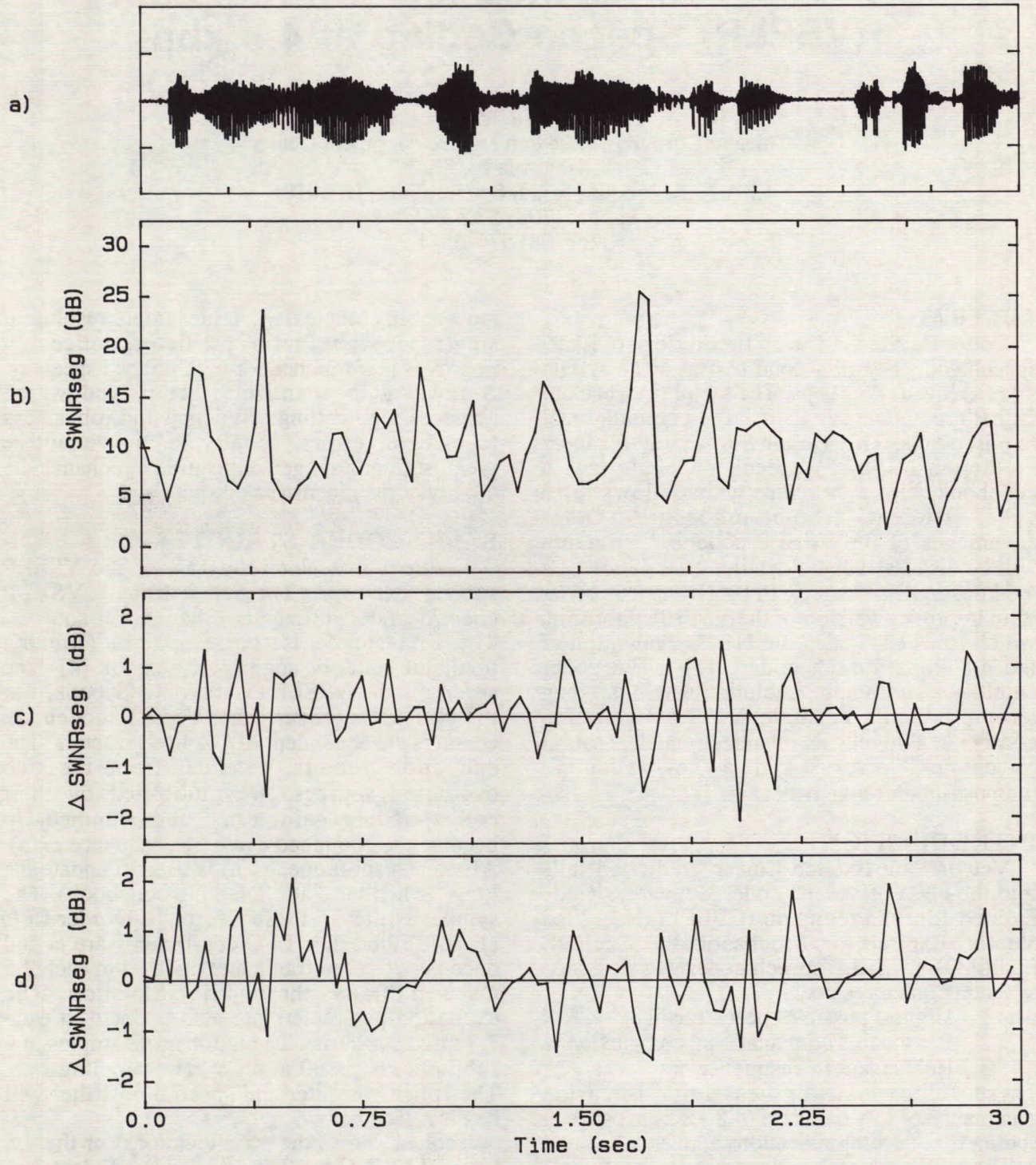


Figure 3. Illustration of SWNR improvements. a) Input speech signal, b) SWNR of the reference coder, c) difference in SWNR between the coder with a weighted codebook and the reference coder, d) difference in SWNR between the coder with a dual codebook and the reference coder.

Vector Sum Excited Linear Prediction (VSELP) Speech Coding at 4.8 kbps

Ira A. Gerson and Mark A. Jasiuk
Chicago Corporate Research and Development Center
Motorola Inc.
1301 E. Algonquin Road, Schaumburg, IL 60196
Phone: (708) 576-3893
Fax: (708) 576-0541

ABSTRACT

Code Excited Linear Prediction (CELP) speech coders exhibit good performance at data rates as low as 4800 bps. The major drawback to CELP type coders is their large computational requirements. The Vector Sum Excited Linear Prediction (VSELP) speech coder utilizes a codebook with a structure which allows for a very efficient search procedure. Other advantages of the VSELP codebook structure will be discussed and a detailed description of a 4.8 kbps VSELP coder will be given. This coder is an improved version of the VSELP algorithm, which finished first in the NSA's evaluation of the 4.8 kbps speech coders [1]. The coder employs a subsample resolution single tap long term predictor, a single VSELP excitation codebook, a novel gain quantizer which is robust to channel errors, and a new adaptive pre/postfilter arrangement.

INTRODUCTION

Vector Sum Excited Linear Prediction falls into the class of speech coders known as Code Excited Linear Prediction (CELP) (also called Vector Excited or Stochastically Excited) [2,4,6]. The VSELP speech coder was designed to accomplish three goals:

1. Highest possible speech quality
2. Reasonable computational complexity
3. Robustness to channel errors

These three goals are essential for wide acceptance of low data rate (4.8 - 8 kbps) speech coding for telecommunications applications.

The VSELP speech coder achieves these goals through efficient utilization of a structured excitation codebook. The structured codebook reduces computational complexity and increases robustness to channel errors [1,3]. A single optimized VSELP excitation codebook is used to achieve high speech quality while maintaining reasonable complexity. A subsample resolution single tap long term predictor noticeably improves performance for high pitched speakers. A novel gain quantizer is employed which achieves high coding efficiency and robustness to channel errors. Finally, a new adaptive pre/post filter arrangement is used to enhance the quality of the reconstructed speech .

BASIC CODER STRUCTURE

Figure 1 is a block diagram of the VSELP speech decoder. The 4.8 kbps VSELP coder/decoder utilizes two excitation sources. The first source is the long term ("pitch") predictor state, or adaptive codebook [4]. The second is the VSELP excitation codebook. For the 4.8 kbps coder, the VSELP codebook contains the equivalent of 1024 codevectors. The excitation vectors, selected from the two excitation sources, are multiplied by their corresponding gain terms and summed, to become the combined excitation sequence ex(n). After each subframe, ex(n) is used to update the long term filter state (adaptive codebook). The synthesis filter is a direct form 10th order LPC all-pole filter. The LPC coefficients are coded once per 30 msec frame and updated in each 7.5 msec subframe through interpolation. The excitation parameters are also updated in each 7.5 msec subframe. The number of samples in a subframe, N, is 60 at an 8 kHz sampling rate. The "pitch" prefilter and spectral postfilter will be discussed below.

Table 1 shows the bit allocations for the 4.8 kbps VSELP coder. The 10 LPC coefficients are coded using scalar quantization of the reflection coefficients. An energy term, $R_q(0)$, which represents the average speech energy per frame is also coded once per frame. To accommodate noninteger values of the long term predictor delay, eight bits are used to code L. A polyphase

FIR interpolating filter generates the excitation vectors for noninteger delays [5]. The two excitation gains are vector quantized to 7 bits (GS-P0 code) per subframe.

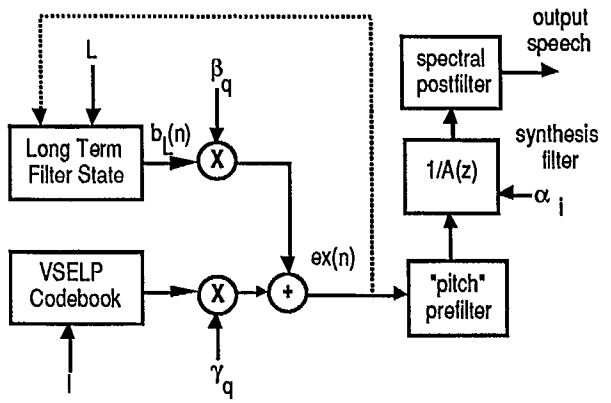


Figure 1 – VSELP Speech Decoder

PARAMETER	BITS/SUBFRAME	BITS/FRAME
LPC coefficients		37
energy - $R_q(0)$		5
excitation code - I	10	40
lag - L	8	32
GS-P0 code	7	28
<synch, parity>		2
TOTAL	25	144

Table 1 – Bit Allocations for 4.8 kbps

VSELP CODEBOOK STRUCTURE

The coder uses a single VSELP excitation codebook, which contains 2^M codevectors. These are constructed from a set of M basis vectors, where M = 10 for the 4.8 kbps coder. Defining $v_m(n)$ as the m^{th} basis vector and $u_i(n)$ as the i^{th} codevector, each from the VSELP codebook, then:

$$u_i(n) = \sum_{m=1}^{M} \theta_{im} v_m(n) \qquad (1)$$

where $0 \le i \le 2^M-1$ and $0 \le n \le N-1$.

In other words, each codevector in the codebook is constructed as a linear combination of the M basis vectors. The linear combinations are defined by the θ parameters. θ_{im} is defined as:

$\theta_{im} = +1$ if bit m of codeword i = 1

$\theta_{im} = -1$ if bit m of codeword i = 0

Note that if we complement all the bits in codeword i, the corresponding codevector is the negative of codevector i. Therefore, for every codevector, its negative is also a codevector in the codebook. These pairs are called complementary codevectors since the corresponding codewords are complements of each other.

The excitation codewords for the VSELP coder are more robust to bit errors than the excitation codewords for random codebooks. A single bit error in a VSELP codeword changes the sign of only one of the basis vectors. The resulting codevector is still similar to the desired codevector.

SELECTION OF EXCITATION VECTORS

Figure 2 is a block diagram which shows the process used to select the two codebook indices L and I. These excitation parameters are computed every subframe.

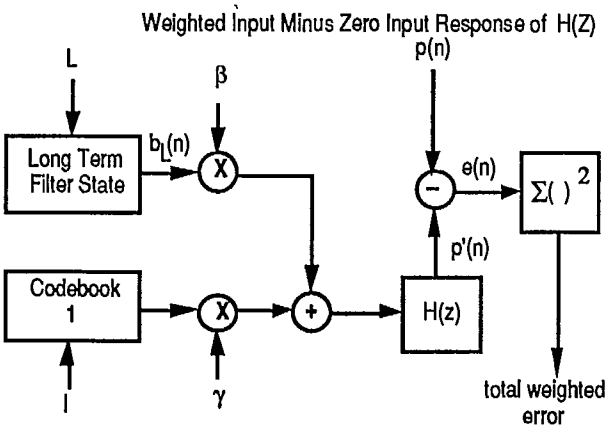


Figure 2 – Excitation Codeword Selection

H(z) is the bandwidth expanded synthesis filter, $H(z) = 1/A(z/\lambda)$, where λ is the noise weighting factor. Signal p(n) is the perceptually weighted (with noise weighting factor λ) input speech for the subframe with the zero input response of bandwidth expanded synthesis filter (H(z)) subtracted out [6].

The two excitation vectors are selected sequentially, one from each of the two excitation codebooks (adaptive codebook and VSELP

codebook). Each codebook search attempts to find the codevector which minimizes the total weighted error.

Although the codevectors are chosen sequentially, the gain of excitation vector chosen from the adaptive codebook is left "floating". The adaptive codebook is searched first, assuming gain γ is zero. The VSELP codebook is searched next, with optimal γ and β used for each codevector being evaluated. This joint optimization can be achieved by orthogonalizing each weighted (filtered) codevector to the weighted excitation vector selected from the adaptive codebook, prior to the VSELP codebook search. While this task seems impractical in general, for VSELP codebook it reduces to orthogonalizing only the weighted basis vectors.

The adaptive codebook is searched first for an index L which minimizes:

$$E'_L = \sum_{n=0}^{N-1} \left(p(n) - \beta' \, b'_L(n)\right)^2 \qquad (2)$$

where $b'_L(n)$ is the zero state response of H(z) to $b_L(n)$ and where β' is optimal for each codebook index L.

To search the VSELP codebook, the zero state response of each codevector to H(z) must be computed. From the definition of the VSELP codebook (1), filtered codevector $f_i(n)$ can be expressed as:

$$f_i(n) = \sum_{m=1}^{M} \theta_{im} \, q_m(n) \qquad (3)$$

where $q_m(n)$ is the zero state response of H(z) to basis vector $v_m(n)$, for $0 \le n \le N-1$.

The orthogonalized filtered codevectors can now be expressed as:

$$f'_i(n) = \sum_{m=1}^{M} \theta_{im} \, q'_m(n) \qquad (4)$$

for $0 \le i \le 2^M-1$ and $0 \le n \le N-1$. Thus $q'_m(n)$ is $q_m(n)$ with the component correlated to $b'_L(n)$ removed. The codebook search procedure now finds the codeword i which minimizes:

$$E'_i = \sum_{n=0}^{N-1} \left(p(n) - \gamma' \, f'_i(n)\right)^2 \qquad (5)$$

where γ' is optimal for each codevector i. Once we have filtered and orthogonalized the basis vectors, the VSELP codebook search is initiated. Defining:

$$C_i = \sum_{n=0}^{N-1} f'_i(n) \, p(n) \qquad (6)$$

and

$$G_i = \sum_{n=0}^{N-1} \left(f'_i(n)\right)^2 \qquad (7)$$

then the codevector which maximizes:

$$\frac{(C_i)^2}{G_i} \qquad (8)$$

is chosen. The search procedure evaluates (8) for each codevector. Using properties of the VSELP codebook structure, the computations required for computing C_i and G_i can be greatly simplified. Defining:

$$R_m = 2\sum_{n=0}^{N-1} q'_m(n) \, p(n) \qquad 1 \le m \le M \qquad (9)$$

and

$$D_{mj} = 4\sum_{n=0}^{N-1} q'_m(n) \, q'_j(n) \qquad 1 \le m \le j \le M \qquad (10)$$

C_i can be expressed as:

$$C_i = \frac{1}{2}\sum_{m=1}^{M} \theta_{im} R_m \qquad (11)$$

and G_i is given by:

$$G_i = \frac{1}{2}\sum_{j=2}^{M}\sum_{m=1}^{j-1} \theta_{im}\,\theta_{ij}\,D_{mj} + \frac{1}{4}\sum_{j=1}^{M} D_{jj} \qquad (12)$$

Assuming that codeword u differs from codeword i in only one bit position, say position v such that $\theta_{uv} = -\theta_{iv}$ and $\theta_{um} = \theta_{im}$ for $m \ne v$ then:

$$C_u = C_i + \theta_{uv} R_v \qquad (13)$$

and

$$G_u = G_i + \sum_{j=1}^{v-1} \theta_{uj}\,\theta_{uv}\,D_{jv} + \sum_{j=v+1}^{M} \theta_{uj}\,\theta_{uv}\,D_{vj} \qquad (14)$$

If the codebook search is structured such that each successive codeword evaluated differs from the previous codeword in only one bit position, then (13) and (14) can be used to update C_i and

G_i in a very efficient manner. Sequencing of the codewords in this manner is accomplished using a binary Gray code.

Note that complementary codewords will have equivalent values for (8). Therefore only half of the codevectors need to be evaluated. Once the codevector which maximizes (8) is found, the sign of the corresponding C_i will determine whether the selected codevector or its negative will yield a positive gain. If C_i is positive then i is the selected codeword; if C_i is negative then the one's complement of i is selected as the codeword.

QUANTIZATION OF EXCITATION GAINS

The quantization of the excitation gains consists of two stages. The first stage codes the average speech energy once per frame. The quantized value of this energy, $R_q(0)$, is specified with five bits, using 2 dB quantization steps for 64 dB of dynamic range. In the second stage, a GS-P0 code is selected every subframe. This code, when taken in conjunction with $R_q(0)$ and the state of the speech decoder, determines the excitation gains for the subframe. The selection of the GS-P0 code takes place after the two excitation vectors, L and I, have been chosen.

The following definitions are used to determine the GS-P0 code. The combined excitation function, ex(n), is given by:

$$ex(n) = \beta\, c_0(n) + \gamma\, c_1(n) \qquad 0 \le n \le N-1 \qquad (15)$$

where:

$c_0(n)$ is the long term prediction vector, $b_L(n)$

$c_1(n)$ is the codevector selected from the VSELP codebook, $u_I(n)$

The energy in each excitation vector is given by:

$$R_x(k) = \sum_{n=0}^{N-1} c^2_k(n) \qquad k = 0,1 \qquad (16)$$

Let RS be the approximate residual energy at a given subframe. RS is a function of N, $R_q(0)$, and the normalized prediction gain of the LPC filter. It is defined by:

$$RS = N\, R_q(0) \prod_{i=1}^{N_p} (1-r_i^2) \qquad (17)$$

where r_i is the i^{th} reflection coefficient corresponding to the set of direct form filter coefficients (α_i's) for the subframe. GS, the energy offset , is a coded parameter which refines the estimated value of RS. R, the approximate total subframe excitation energy, is defined as:

$$R = GS\ RS \qquad (18)$$

P0, the approximate energy contribution of the long term prediction vector as a fraction of the total excitation energy at a subframe, is defined to be:

$$P0 = \frac{\beta^2 R_x(0)}{R} \qquad \text{where } 0 \le P0 \le 1 \qquad (19)$$

Thus β and γ are replaced by two new parameters: GS and P0. The transformations relating β and γ to GS and P0 are given by:

$$\beta = \sqrt{\frac{RS\ GS\ P0}{R_x(0)}} \qquad (20)$$

$$\gamma = \sqrt{\frac{RS\ GS\ (1-P0)}{R_x(1)}} \qquad (21)$$

The GS-P0 pair is vector quantized using a codebook of 128 vectors. The codebook was designed using the LBG algorithm [7], using the normalized weighted error as the distortion criterion. Figure 3 shows the distribution of the GS-P0 codebook vectors.

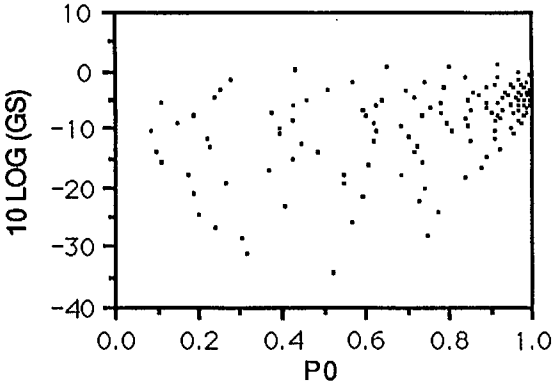


Figure 3 – P0 vs GS in dB for gain codebook

The vector, which minimizes the total weighted error energy for the subframe, is chosen from the GS-P0 codebook. The codebook search procedure requires only five multiply-accumulates per vector evaluation.

This technique of quantizing the gains has a number of advantages. First, the coding is efficient. The coding of the energy once per

frame solves the dynamic range issue. The gain quantization performs equally well at all signal levels within the range of the $R_q(0)$ quantizer. With the average energy factored out, the two gains can be vector quantized efficiently. In minimizing the weighted error, the vector quantizer takes into account the correlation between the two weighted excitation vectors. Second, the values of GS and P0 are well behaved as can be seen in Figure 3. Whereas the optimal value for β, the adaptive codebook gain, can occasionally get very large, P0 is bounded by 0 and 1. Error propagation effects are also greatly reduced by this quantization scheme. Since the energies in the excitation vectors are used to normalize the excitation gains, previous channel errors affecting the energy in the adaptive codebook vector have very little effect on the decoded speech energy. Channel errors in the LPC coefficients are also automatically compensated for at the decoder in calculating the excitation gains. In fact as long as the code for the average frame energy, $R_q(0)$, is received correctly, the speech energy at the decoder will not be much greater than the desired energy (see Figure 3 for range of GS) and no "blasting" will occur.

OPTIMIZATION OF BASIS VECTORS

The basis vectors defining the VSELP codebook are optimized over a training database. The optimization criterion is the minimization of the total normalized weighted error. The normalized weighted error for each subframe can be expressed as a function of individual samples of each of the 10 basis vectors from the VSELP excitation codebook, given I, $b_L(n)$, p(n), the excitation gains, and the impulse response of H(z) for each subframe of the training data. The optimal basis vectors are computed by solving the 600 (10 basis vectors, 60 samples per vector) simultaneous equations which result from taking the partial derivatives of the total normalized weighted error function with respect to each sample of each basis vector and setting them equal to zero. Since the coder subframes are not independent, this procedure is iterated in a closed loop fashion. Figure 4 shows the improvement in weighted segmental SNR for each iteration.

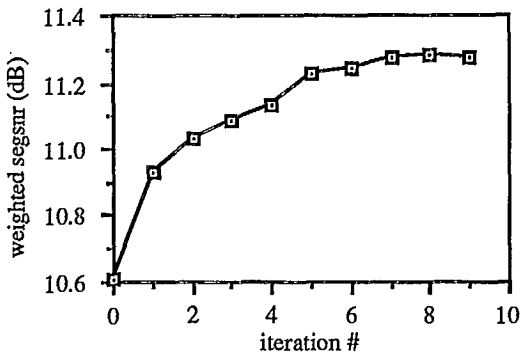


Figure 4 – Basis Vector Optimization

Initially the basis vectors are populated with random Gaussian sequences (iteration 0) which yields a weighted segmental SNR of 10.61 dB. The weighted segmental SNR increases to 11.28 dB after nine iterations. The subjective quality improvement due to the optimization of the basis vectors is significant. The objective as well as subjective improvements are retained for speech data outside the training data base.

ADAPTIVE PRE AND POSTFILTERING

The speech decoder creates the combined excitation signal, ex(n), from the long term filter state and the VSELP excitation codebook. The combined excitation is then processed by an adaptive "pitch" prefilter to enhance the periodicity of the excitation signal (see Figure 1). Following the adaptive pitch prefilter, the prefiltered excitation is applied to the LPC synthesis filter. After reconstructing the speech signal with the synthesis filter, an adaptive spectral postfilter is applied to further enhance the quality of the reconstructed speech. The pitch prefilter transfer function used is given by:

$$H_p(z) = \frac{1}{1 - \xi z^{-L}} \qquad (22)$$

where $\xi = \varepsilon \, \mathrm{Min}[\, \beta, \sqrt{P0}\,]$ and $\varepsilon = 0.4$ (23)

Note that the periodicity enhancement is performed on the synthetic residual in contrast to pitch postfiltering which performs the enhancement on the synthesized speech waveform [8]. This significantly reduces artifacts in the reconstructed speech due to waveform discontinuities which pitch postfiltering sometimes introduces. For

noninteger values of L, a polyphase FIR filter is used to compute the fractionally delayed excitation samples. Finally to ensure unity power gain between the input and the output of the pitch prefilter, a gain scale factor is calculated to scale the pitch prefiltered excitation prior to applying it the LPC synthesis filter.

The form of the adaptive spectral postfilter used is:

$$H_S(z) = \frac{1 - \sum_{i=1}^{10} \eta_i z^{-i}}{1 - \sum_{i=1}^{10} \nu^i \alpha_i z^{-i}} \qquad 0 \le \nu \le 1 \quad (24)$$

where the α_i's are the coefficients of the synthesis filter. To derive the numerator, the $\nu^i \alpha_i$ coefficients are converted to the autocorrelation domain (the autocorrelation of the impulse response of the all pole filter corresponding to the denominator of (24) is calculated for lags 0 through 10). A binomial window is then applied to the autocorrelation sequence [9] and the numerator polynomial coefficients are calculated from the modified autocorrelation sequence via the Levinson recursion. This postfilter is similar to that proposed by Gersho and Chen [10]. However, the use of the autocorrelation domain windowing results in a frequency response for the numerator that tracks the general shape and slope of the denominator's frequency response more closely. To increase postfiltered speech "brightness", an additional first order filter is used of the form:

$$H_B(z) = 1 - u z^{-1} \qquad (25)$$

The following postfilter parameter values are used: ν=.8, B_{eq}=1200 Hz, u=0.4. Note that B_{eq} is the bandwidth expansion factor which specifies the degree of smoothing which is performed on the denominator to generate the numerator.

As in the case of the pitch prefilter, a method of automatic gain control is needed to ensure unity gain through the spectral postfilter. A scale factor is computed for the subframe in the same manner as was done for the pitch prefilter. In the case of the spectral postfilter, this scale factor is not used directly. To avoid discontinuities in the output waveform, the scale factor is passed through a first order low pass filter before being applied to the postfilter output.

CONCLUSIONS

A high quality 4.8 kbps speech coder has been described. The complexity of the coder is low enough so that it can be implemented on a single DSP device such as the Motorola DSP56000. In addition to its very high speech quality, the parameters are inherently robust to channel errors and can be error protected very efficiently.

REFERENCES

[1] D. P. Kemp, R. A. Sueda and T. E. Tremain, "An Evaluation of 4800 bps Voice Coders", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, May 1989.

[2] M. R. Schroeder and B. S. Atal, "Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, pp. 937-940, March 1985.

[3] I. Gerson and M. Jasiuk, "Vector Sum Excited Linear Prediction (VSELP)", *IEEE Workshop on Speech Coding for Telecommunications*, pp. 66-68, September 1989.

[4] W. B. Kleijn, D. J. Krasinski and R. H. Ketchun, "Improved Speech Quality and Efficient Vector Quantization in SELP", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, pp. 155-158, April 1988.

[5] P. Kroon and B.S. Atal, "Pitch Predictors with High Temporal Resolution", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, April 1990.

[6] G. Davidson and A. Gersho, "Complexity Reduction Methods for Vector Excitation Coding", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, pp. 3055-3058, May 1986.

[7] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Trans. Comm.*, vol. COM-28, pp. 84-95, Jan. 1980.

[8] P. Kroon and E. F. Deprettere, "A Class of Analysis-by-Synthesis Predictive Coders for High Quality Speech Coding at Rates Between 4.8 and 16 kbits/s", *IEEE J. Select. Areas Commun.*, vol. SAC-6, No. 2, February 1988.

[9] Y. Tohkura, F. Itakura and S. Hashimoto, "Spectral Smoothing Technique in PARCOR Speech Analysis-Synthesis", *IEEE Trans. Acoustics Speech and Signal Processing*, vol. ASSP-26, pp. 587-596, Dec. 1978.

[10] J. Chen and A. Gersho, "Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering", *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing*, pp. 51.3.1-51.3.4, April 1987.

A Low-Delay 8 Kb/s Backward-Adaptive CELP Coder

L.G. Neumeyer, W.P. LeBlanc and S.A. Mahmoud
Dept. of Systems and Computer Engineering
Carleton University
Ottawa, Ontario, K1S 5B6, Canada
Phone: 613-788-5753
Fax: 613-788-5727

ABSTRACT*

Code-excited linear prediction coding is an efficient technique for compressing speech sequences. Communications quality of speech can be obtained at bit rates below 8 Kb/s. However, relatively large coding delays are necessary to buffer the input speech in order to perform the LPC analysis. In this paper we introduce a low–delay 8Kb/s CELP coder in which the short-term predictor is based on past synthesized speech. A new distortion measure that improves the tracking of the formant filter is discussed. Formal listening tests showed that the performance of the backward-adaptive coder is almost as good as the conventional CELP coder.

INTRODUCTION

Recent advances in linear prediction coding have made it possible to achieve communications quality of speech at bit rates below 8 Kb/s. Practical real-time implementations are possible due to efficient algorithms based on Code-Excited Linear Prediction (CELP) [1]. In these coders, the residual is vector quantized using an analysis-by-synthesis search procedure. The excitation vector (or codevector) is chosen from a large codebook. All the codevectors are passed through the synthesis filters and compared with the original speech vector. The index of the codevector that minimizes an objective distortion measure between original and quantized speech is sent through the channel. The parameters of the synthesis filters (gain, long- and short-term LPC coefficients, and pitch lag) are sent to the decoder as side information. Gain, pitch lag, and long-term predictor coefficients can be optimized using closed-loop procedures. Ideally, the formant filter could also be optimized in a closed-loop procedure but this would lead to a mathematically untractable set of non-linear equations [2]. Therefore, the short-term predictor coefficients are calculated using an open-loop solution based on the original speech. In order to obtain a reliable linear prediction filter, approximately 20 ms of speech samples are buffered. The one-way delay of the coder, although highly dependent on real-time implementations, could be as high as 60 ms. The delay could be reduced by using only past speech (no buffering). However, the linear prediction analysis would be unreliable, resulting in poor speech quality. This problem can be overcome by updating the LPC parameters at a higher rate. This would require more bits/sample, thereby increasing either the total bit rate or the distortion.

In this paper, we present an 8 Kb/s CELP coder in which the short-term linear prediction parameters are updated in a backward-adaptive manner. That is, the linear prediction analysis is performed on past synthesized speech which is available, assuming no transmission errors, at both ends of the channel. Therefore, the LPC parameters are not sent through the channel and can be updated at high rates, even in a sample-by-sample basis. Speech quality is as good as in the conventional (or forward-adaptive) coder. A new distortion measure is introduced to prevent predictor mistracking.

A diagram of the encoder is shown in figure

* This work has been sponsored by the Telecommunication Research Institute of Ontario (TRIO).

1. The synthesis filters are separated into their zero-input and zero-state components. The minimum pitch is constrained to be always greater than the block size. Therefore the transfer function of the zero-state pitch synthesis filter is unity. The weighting filter is moved from its original location (filtering the error between original and synthesized speech) to both of its input branches.

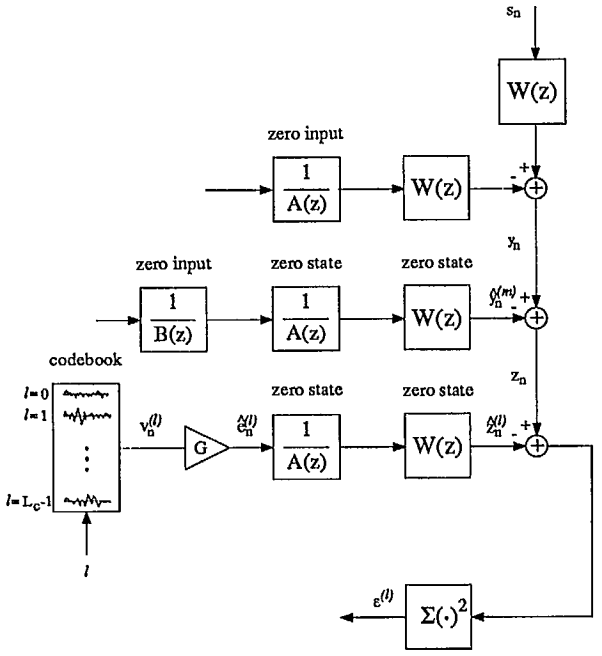


Figure 1 CELP Encoder

CODER DESIGN

Backward adaptation

In pure backward-adaptive 16 Kb/s CELP coders [3] [4] only the excitation vector index is sent through the channel. The rest of the parameters are computed in a backward-adaptive mode. The vector dimension (block of samples) is 4–5 samples and the one-way delay is below 2 ms. Unfortunately, as the bit rate decreases the quantization effects become more pronounced, leading to poor filter tracking and to further distortion of the original speech. As a result, in the BA-CELP coder only the short-term predictor is computed in a backward-adaptive manner. Pitch filter parameters and gain are optimized in closed-loop procedures and sent through the channel as side information. The three-tap pitch filter plays an important role, not only in the fine structure but also in the shape of the spectrum of the reconstructed speech.

In our BA-CELP coder, all the parameters are updated at the end of each block of samples. The bit allocation scheme is shown in table 1. The vector dimension is 26 samples and the sampling rate is 8 Khz. Consequently, the total delay (typically 4 times the vector dimension) is around 13 ms. The short-term LPC analysis is performed using the modified covariance method. The length of the frame is four times the vector dimension. Note that the frames are highly overlapped and relatively short. This is necessary to improve the tracking of the adaptive filter, specially when rapid changes of the spectrum occur. The autocorrelation method proved to be inefficient in this application. This is because the windowing process weights the error in the middle of the frames higher than at the edge of the frames. As a result, spectral match is poor in regions of rapid spectrum variations.

Parameter	bits/block	Kbits/sec
Formant filter	0	0.0
Pitch filter	5	1.5
Pitch lag	7	2.2
Gain	5	1.5
Excitation vec.	9	2.8
Total	26	8.0

Table 1 Bit allocation and corresponding bit rate. The vector dimension is 26 samples and the sampling rate is 8 Khz.

Perceptual weighting filter

Psychoacoustical studies show that the human auditory system can tolerate more errors in the formants of the speech spectrum than in the valleys. Therefore, we can obtain a more subjective distortion measure by weighting the spectrum of the error. Regions of the error spectrum that correspond to valleys between formants in the speech spectrum are de-emphasized and regions corresponding to the formants are emphasized. Using a weighting function $W(z)$ we can

write,

$$\epsilon_w = \frac{1}{2\pi}\int_0^{2\pi} \left|S\left(e^{j\omega}\right) - \hat{S}\left(e^{j\omega}\right)\right|^2 W\left(e^{j\omega}\right) d\omega \quad (1)$$

where ϵ_w is the noise-weighted mean-squared error (NWMSE). A general weighting filter is discussed in [3] and [5].

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)}, \qquad 0 < \gamma_2 < \gamma_1 \leq 1 \quad (2)$$

A good choice for the parameters is γ_1=*0.9* and γ_2=*0.4*. Note that in conventional CELP coders only one LPC analysis is necessary for the synthesis and weighting filters. Conversely, the backward-adaptive approach "requires" two separate LPC analyses. One based on reconstructed speech for the synthesis filter and the other one based on the original speech for the weighting filter.

Mixed distortion

Further improvement in filter tracking can be achieved by using a *mixed distortion measure* in the excitation vector search procedure. The proposed mixed distortion combines mean-squared error with a subjectively meaningful LPC distortion measure.

Log-likelihood ratio distortion measure.

In linear prediction theory, the minimum residual energy for a particular speech frame is given by

$$\alpha = r_0 - \mathbf{a}^T\mathbf{r} \quad (3)$$

where $\mathbf{r}$ is the correlation vector, r_0 is the energy of the segment and $\mathbf{a}$ is the optimum LPC coefficient vector. If the same frame is passed through a non-optimum inverse filter then the residual energy β must be greater than α,

$$\beta = r_0 - 2\hat{\mathbf{a}}^T\mathbf{r} + \hat{\mathbf{a}}^T\mathbf{R}\hat{\mathbf{a}} \geq \alpha \quad (4)$$

where $\mathbf{R}$ is the correlation matrix. Equality holds when $\mathbf{a} = \hat{\mathbf{a}}$. The log-likelihood ratio (LLR) distortion measure is defined as

$$d_{LLR}\left(A(z), \hat{A}(z)\right) = \log\left(\frac{\beta}{\alpha}\right) \quad (5)$$

which is equivalent to the difference of the logarithmic prediction gains. The LLR distortion measure has proved to be subjectively meaningful [6][7].

Figure 2 shows the filtering operation. The two input sequences are *s(n)* and *ŝ(n)*. The corresponding p^{th} order inverse filters are *A(z)* and *Â(z)*. When one of the input sequences, say *s(n)* is passed through the filters, the resulting residual energies are α and β. A different distortion would be obtained by using *ŝ(n)* instead of *s(n)* as the input sequence. This difference shows the asymmetric nature of the likelihood ratio.

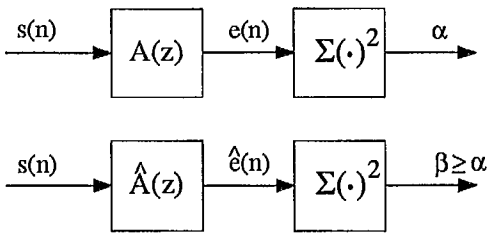


Figure 2 Computation of the residual energies in the log-likelihood ratio distortion measure.

Search procedure. The optimum excitation vector is searched in two sequential steps. First, a conventional search algorithm finds the best n_c excitation vectors that minimizes the NWMSE. The best n_c candidates are used in the second stage in order to minimize the mixed distortion measure.

The convolution of the l^{th} codevector $\mathbf{v}(l)$ with the impulse response of the weighted synthesis filter can be written in matrix form as,

$$\hat{\mathbf{z}}(l) = G\mathbf{H}\mathbf{v}(l) \quad (6)$$

where G is the gain, $\mathbf{H}$ is a lower triangular toeplitz matrix containing the impulse response in its first column and K is the vector dimension,

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{K-1} & h_{K-2} & h_{K-3} & \cdots & h_0 \end{bmatrix} \quad (7)$$

The NWMSE is given by,

$$\epsilon_w(l) = \|\mathbf{z} - \hat{\mathbf{z}}(l)\|^2$$
$$= \|\mathbf{z}\|^2 - 2G\mathbf{z}^T\mathbf{H}\mathbf{v}(l) + G^2\mathbf{v}^T(l)\,\mathbf{H}^T\mathbf{H}\mathbf{v}(l) \quad (8)$$

Taking the derivative with respect to G we get the minimum NWMSE and the optimum gain for the l^{th} codevector,

$$\frac{\partial \epsilon_w}{\partial G} = -2\mathbf{z}^T\mathbf{H}\mathbf{v}(l) + 2G\mathbf{v}^T(l)\mathbf{H}^T\mathbf{H}\mathbf{v}(l) = 0$$
$$\Rightarrow \quad G_{opt}(l) = \frac{\mathbf{z}^T\mathbf{H}\mathbf{v}(l)}{\mathbf{v}^T(l)\mathbf{H}^T\mathbf{H}\mathbf{v}(l)}$$
$$\Rightarrow \quad \epsilon_{w_{min}}(l) = \|\mathbf{z}\|^2 - \frac{\left(\mathbf{z}^T\mathbf{H}\mathbf{v}(l)\right)^2}{\mathbf{v}^T(l)\mathbf{H}^T\mathbf{H}\mathbf{v}(l)} \quad (9)$$

In order to find the n_c optimum excitation vectors out of the L-level codebook it is necessary to maximize the second term of the minimum error:

$$find \quad l_{opt} \Rightarrow \max_{l=0\ldots L-1} \frac{\left(\mathbf{z}^T\mathbf{H}\mathbf{v}(l)\right)^2}{\mathbf{v}^T(l)\mathbf{H}^T\mathbf{H}\mathbf{v}(l)} \quad (10)$$

The computational complexity of the search is reduced by using a shift-symmetric codebook [8].

In the second stage of the search, the optimum codevector is chosen out of the n_c candidates. The objective is to choose a codevector that minimizes the distortion between the original LPC model *A(z)* and the backward-adaptive LPC model $\hat{A}(z)$ one vector into the future. The original LPC model has already been computed for the weighting filter. To calculate the corresponding $\hat{A}(z)$ for each candidate, we compute the next block of speech samples and perform the LPC analysis on the updated synthesized speech sequence. The mixed distortion is defined as,

$$d_{mix}^{(i)} = d_{LLR}^{(i)}\left(A(z), \hat{A}(z)\right) + \eta \log \frac{\epsilon_w^{(i)}}{\epsilon_w^{(min)}}$$
$$i = 1 \ldots n_c \quad (11)$$

where $\epsilon_w^{(min)}$ is the minimum NWMSE of the candidates and η is a parameter to be optimized in subjective tests. In equation 11, as η goes to infinity the mixed distortion measure becomes equivalent to the NWMSE. On the other hand, as η approaches zero the LPC distortion of future frames decreases at the cost of accuracy in the current block of samples.

SIMULATION RESULTS

Computer simulations results were obtained for the BA-CELP coder and for a conventional forward-adaptive version. The conventional 8 Kb/s CELP coder computed the LPC analysis on 20 ms of buffered speech. The autocorrelation method was used to calculate the LPC coefficients which were transformed to linear-spectrum pairs and quantized. For the BA-CELP coder we used the mixed distortion parameters η=*1* and n_c=*16*. For these values, the NWMSE was greater than the minimum in 20% of the speech blocks. The shift-symmetric excitation codebook was optimized using a 10–minute speech database.

Formal listening tests were conducted following the CCITT recommendations in [9]. The stimulus material contained six different sentences spoken by different males and females. Twenty listeners evaluated speech quality under five different conditions, 2 coders and 3 references. The reference conditions consisted of the original speech (PCM 64 Kb/s) and speech corrupted with random noise which has amplitude proportional to the instantaneous signal amplitude. The distorted speech is specified according to the modulated noise reference unit (MNRU) [10]. Mean opinion scores and 95% confidence intervals are shown in table 2. According to our results, speech quality in the forward-adaptive coder is only 0.1 points in the MOS scale better than the BA-CELP coder.

Condition	Mean	Error
PCM 64 Kb/s	4.24	0.15
Forward	3.42	0.20
Backward	3.33	0.19
MNRU 20 dB	2.39	0.18
MNRU 15 dB	1.68	0.17

Table 2 Mean opinion score test.
Mean and 95% confidence intervals.

Figure 3 shows noise-weighted signal-to-noise ratio as a function of η for a 30 second segment of speech. The dashed line represents the NWSNR for the conventional search (no LLR distortion). Observe that for values of η between 0.2 and 3 the global NWSNR is greater than for the regular search NWSNR. This shows how the global NWMSE was reduced by using the suboptimal (in a NWMSE sense) mixed distortion measure. In other words, an increase in the error of one block of samples helps in filter tracking and therefore improves the overall performance of the coder. Figure 4 shows the log-likelihood ratio distortion measure for different values of η.

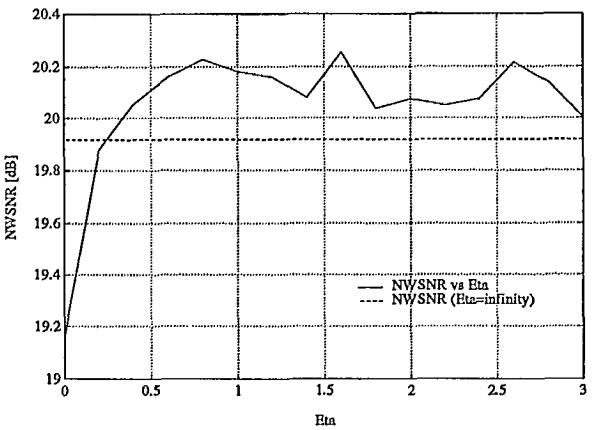


Figure 3 Noise-weighted signal-to-noise ratio versus η.

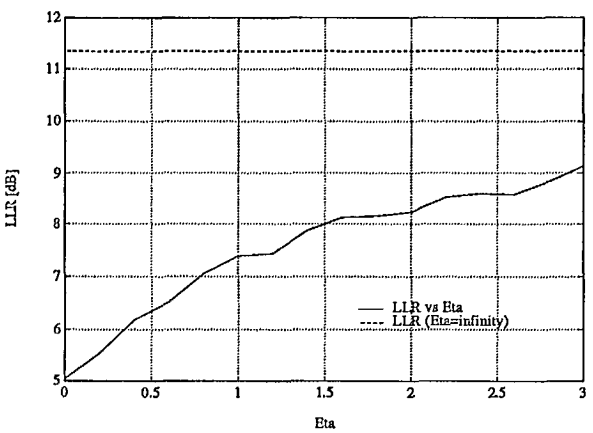


Figure 4 Log-likelihood ratio distortion vs η.

CONCLUDING REMARKS

In this paper we discussed how a delay of approximately 13 ms is achieved in the BA-CELP. Based on subjective MOS tests, speech quality has been found to be comparable to that of conventional forward-adaptive CELP coders. However, several LPC analyses are necessary to compute the mixed distortion measure. The number of candidates in the search procedure determines the computational complexity of the coder. Further reductions in complexity may be possible by using recursive LPC algorithms instead of block algorithms. To further reduce the delay, future investigation would include backward-adaptation of the remaining parameters.

REFERENCES

[1] M. Schroeder and B. Atal, "Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 937–940, 1985.

[2] P. Kabal, J. Moncet, and C. Chu, "Synthesis Filter Optimization and Coding: Applications to CELP," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 147–150, 1988.

[3] J. Chen, "A Robust Low-Delay CELP Speech Coder," *IEEE Global Communications Conference*, pp. 1237–1241, 1989.

[4] V. Cuperman, A. Gersho, R. Pettigrew, J. Shynk, and J. Yao, "Backward Adaptation for Low Delay Vector Excitation Coding of Speech at 16 Kbit/s," *IEEE Global Communications Conference*, pp. 1242–1246, 1989.

[5] J. Chen and A. Gersho, "Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2185–2188, 1987.

[6] A. Gray and J. Markel, "Distance Measures for Speech Processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-24, no. 5, pp. 380–391, October 1976.

[7] B. Juang, "On Using the Itakura-Saito Measures for Speech Coder Performance Evaluation," *AT&T Bell Laboratories Technical Journal*, vol. 63, no. 8, pp. 1477–1498, October 1984.

[8] W. Kleijn, D. Krasinski, and R. Ketchum, "Improved Speech Quality and Efficient Vector Quantization in SELP," *IEEE Inter-*

national Conference on Acoustics, Speech and Signal Processing, pp. 155–158, 1988.

[9] CCITT Supplement No. 14 to Recommendation P.81, Blue Book, *Subjective Performance Assessment of Digital Processes Using the Modulated Noise Reference Unit*, 1988.

[10] CCITT Supplement No. 14 to Recommendation P.81, Blue Book, *Modulated Noise Reference Unit (MNRU)*, 1988.

A Variable Rate Speech Compressor for Mobile Applications

S. Yeldener A. M. Kondoz B. G. Evans

Department of Electronics and Electrical Engineering
University of Surrey
Guildford, Surrey, GU2 5XH.

ABSTRACT

One of the most promising speech coder at the bit rate of 9.6 - 4.8 kbits/s is CELP [1]. CELP has been dominating 9.6 to 4.8 kbits/s region during the past 3 to 4 years, its set back however, is its expensive implementation. As an alternative to CELP, we have developed the Base-Band CELP (CELP-BB) [2] which produced good quality speech comparable to CELP and a single chip implementable complexity as reported in [3]. We have also improved its robustness to tolerate errors up to 1.0% and maintain intelligibility up to 5.0% and more [4]. Although, CELP-BB produces good quality speech at around 4.8 kbits/s, it has a fundamental problem when updating the pitch filter memory. We proposed a sub-optimal solution to this problem in this paper. Below 4.8 kbits/s, however, CELP-BB suffers from noticeable quantization noise as result of the large vector dimensions used. In this paper, therefore, we also report on efficient representation of speech below 4.8 kbits/s by introducing Sinusoidal Transform Coding (STC) [5] to represent the LPC excitation which is called Sine Wave Excited LPC (SWELP). In this case natural sounding good quality synthetic speech is obtained at around 2.4 kbits/s.

1. INTRODUCTION

The type of speech compression technique is very important for maritime and land mobile satellite communication systems. For these systems, the resources are very limited interms of the very small transceiver terminals requiring larger satellite power, and the very restricted bandwidth currently available. This especially applies to the land mobile satellite service which currently has only 4 MHz allocated on primary service transmission. For such services to be economic, they must employ very narrow bandwidth per channel. The competition is with analogue systems that employ Amplitude Companded Single Side Band (ACSSB) and achieve reasonable performance at C / N_0's of around 50 dB-Hz in 5 kHz transmitted bandwidth. Now in order to be competitive and to use modulation schemes that will not cause excessive distortion over the difficult land mobile propagation channel, digital speech coding of around 4.8 kbits/s or less is required. The performance must be better than the analogue contender in worst case degraded channels and the speech quality must be acceptable enough to be connected on to PSTN transmission.

With Land mobile satellite systems (INMARSAT, AVSAT, MSAT, etc.) proposing to operate voice services soon, the race is on to produce digital speech coders that can meet all the requirements. Until now Analysis By Synthesis (ABS) schemes such as CELP have only achieved these qualities down to 6 kbits/s. Its quality at 4.8 kbits/s can also be made adequate by post filtering the recovered speech. However, below 4.8 kbits/s, most of these schemes suffer from noticeable quantization noise as result of the large vector dimensions used. Another major disadvantage with these schemes is that they are not really practical for real time implementation owning to enormous computational demand. For the land mobile service we are looking for a speech coder whose cost is a fraction of the mobile terminal, thus we need a scheme which is simple to implement in a single DSP chip. The CELP-BB [2] satisfies these requirements, however below 4.8 kbits/s it suffers from noticeable quantization noise as mentioned earlier. In this paper we present results of a coder called Sine Wave Excited LPC (SWELP) which we believe is

capable of meeting all the requirements and thus is a serious contender for future land mobile satellite systems.

2. OVERVIEW OF CELP-BB SYSTEM

In CELP-BB system, a block of speech, $S(n)$ is first analized and 10 LPC coefficients are computed. These are transformed into Line Spectral Frequencies (LSF) and then quantized [4]. The quantized LSF are inverse transformed into LPC parameters which are then used to form an inverse filter to derive the LPC residual signal. The LPC residual signal is then divided into sub-blocks, each of which is filtered by the weighting filter (smoothing filter) separately. Filtered sub-blocks are split into a number of sequences equal to the decimation factor. These sequences are compared in terms of their energies and the one with the highest energy is selected for transmission. The position of the selected sequence in each sub-block is also transmitted to the decoder in order to place the sequence in the correct location in High Frequency Regeneration (HFR). The selected decimated signal is then treated as a continuous signal which now contains decimation factor times less samples than the original. This continuous signal is then quantized via a restricted ABS procedure operating around merely a pitch synthesis filter. First the pitch filter delay and gain of the continuous signal are computed by cross correlation with the past decoded samples in the pitch synthesis filter. Using these parameters in a pitch synthesis filter the memory response of the filter is computed and subtracted from the decimated signal to form the reference signal. Gaussian code-book sequences are then searched one by one to match the output response of the pitch synthesis filter with no memory, to the reference signal. The index of the optimum sequence, together with its scale value are transmitted to the decoder. At the decoder, selected code-book sequences are scaled up by their scale factors and passed through the pitch synthesis filter in order to recover the continuous decimated base-band signal. The recovered signal is then sub segmented and shifted to the correct positions with zeros inserted in between the samples, to form the LPC filter excitation sequence. Using this sequence the LPC synthesis filter is excited to recover the output speech.

2.1 The Performance Of CELP-BB

Although the encoder of CELP-BB appears very similar to CELP [1], it in fact offers one very significant advantage, namely simplicity. As the ABS procedure operates on the base-band, the vector dimensions are reduced by a factor of decimation. Also the ABS procedure is restricted around the pitch synthesis filter and noise shaping filter. This eliminates considerable amount of convolution computations required by the LPC filter as in standard CELP. These two differences contribute to the enormous savings in computations. Although, CELP-BB is much simpler coding scheme, its speech quality remains to be comparable to CELP [1].

Although, CELP-BB produces good quality speech at around 4.8 kbits/s with a single chip implementable complexity [3], it has a fundamental problem updating the pitch filter memory. The worst case situation arises, when the first and the last (third) decimated sequences are selected in two consequative sub-blocks as shown in figure 2.1.

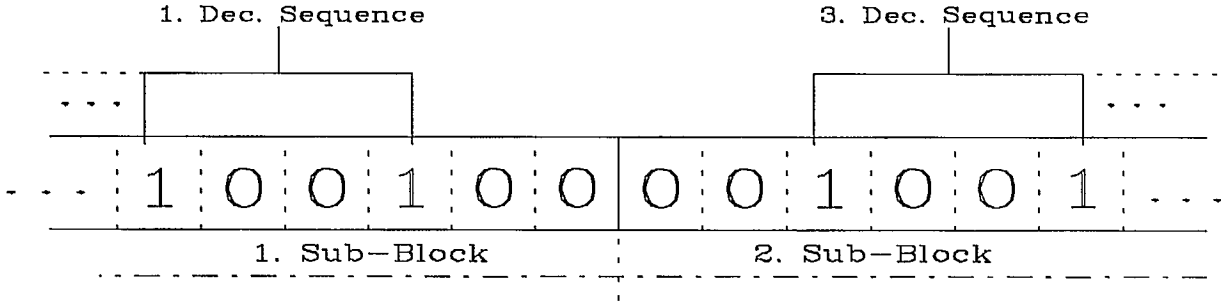


Figure 2.1 The representation of selected sequences in worst case.

In this figure, 1's represent the selected samples of each sub-block having the highest energy

comparing with the other sequences which are represented by 0's. In this case, The first and the third decimated sequences are selected for transmission for the first sub-block and second sub-block respectively. When the selected sequences are treated as a continuous signal which each consists of decimation factor times less samples than the original, in the worst case as shown in figure 2.1, an offset time occurs in the third sequence of the second sub-block. Since the pitch filter operates on the decimated sequence, this situation disturbs the pitch structure. Therefore, a method has to be found to update the pitch filter memory efficiently and accurately. One solution for this problem is to interpolate the decimated sequences in the pitch filter memory. With this strategy, pitch delay can be calculated by maximizing the equation 2.1 as,

$$E(p)=\sum_{i=1}^{N/D} h(iD+p)\,r(i) \qquad (2.1)$$

where N is the sub-block size, D is the decimation factor, p represents the pitch delay which ranges from 20 to 147 samples, $h(i)$ is the interpolated samples of pitch filter memory and $r(i)$ is the selected decimated sequence for the updated sub-block. In this case, although, the performance improvement is dependent on the accuracy of interpolation, we have noticed partial improvement with a simple interpolation schemes.

Below 4.8 kbits/s, however, CELP-BB suffers from noticeable quantization noise, back ground noise and hence roughness as result of the large vector dimensions used. The reason for this is that during vector quantization of excitation, all of the components are matched as a single vector. This results in poor matching performance for larger vector dimensions. Therefore, we have introduced the Sinusoidal Transform Coding (STC) [5] to represent the LPC excitation. Whit this it is possible to code speech below 4.8 kbits/s with a minimum quality loss. This system is called Sine Wave Excited LPC (SWELP) The results of this system is addressed in next sections.

3. SWELP CODING SYSTEM

In the speech production model, the speech waveform, $S(n)$ is assumed to be the output of passing a glottal excitation waveform, $r(n)$ through a linear time-varying system with impulse response $h(n)$ representing the characteristics of the vocal tract. Mathematically, this can be written as,

$$S(n)=\sum_{m=0}^{n} r(m)h(n-m) \qquad (3.1)$$

The excitation will be represented by a sum of sine waves of arbitrary amplitudes, frequencies and phases,

$$r_m(n)=\sum_{k=1}^{L_m} A_k^m\,Cos\,(n\omega_k^m+\Phi_k^m) \qquad (3.2)$$

where L_m is the number of sine-waves and A_k^m, ω_k^m and Φ_k^m are the amplitude, frequency and phase respectively for the k^{th} sine-wave components at the m^{th} frame.

A block diagram of the SWELP Analysis/Synthesis system is given in figure 3.1 which operates as follows:

3.1 Analysis

A block of speech, $S(n)$ is first LPC analized to obtain the LPC coefficients. These coefficients are then quantized. The quantized coefficients are used to form an inverse filter to derive the LPC excitation in sub-blocks, $r(n)$

$$r(n)=S(n)-\sum_{k=1}^{p} a_k S(n-k) \qquad (3.3)$$

where a_k's are the LPC filter coefficients and p is the order of the filter. The spectrum of LPC excitation is then analized using a 512 point FFT and a hamming window having a minimum width of 2.5 times the average pitch for accurate peak estimation. The spectrum, $R(\omega_k)$ can be computed as,

$$R(\omega_k)=\sum_{n=0}^{N-1} r(n)W(n)e^{-jn\omega_k} \qquad (3.4)$$

where $W(n)$ is a hamming window and ω_k's are the discrete frequencies ($\omega_k=\frac{2\pi k}{N}$).

The number of peaks L_m is typically about 40 - 50 over a 4 kHz range. The maximum number of peaks that can be specified is limited by a threshold that is also function of the measured average fundamental. In general, the performance can be affected by the choice of this

threshold only when too few peaks were allowed. The locations of the largest peaks are estimated by simply searching for a change of slope from positive to negative in the uniformly spaced samples of the short-time Fourier transform magnitude, ($|R(\omega_k)|$). The amplitude and phase components (modula 2π) of the sine waves are given by the appropriate samples of the high resolution FFT corresponding to $R(\omega_k)$ at the chosen frequency locations.

3.2 Synthesis

The LPC coefficients and the set of amplitudes, frequencies and phases which are estimated in the encoder, are transmitted to the decoder. Received set of amplitudes, frequencies and phases are used to generate the sine waves for each tone. The generated sine waves are then added together to form the LPC excitation, $\hat{r}(n)$ using equation 3.2. The final quantized speech, $\hat{S}(n)$ is then obtained by passing the recovered LPC excitation through the LPC filter,

$$\hat{S}(n)=\hat{r}(n)+\sum_{k=1}^{p} a_k \hat{S}(n-k) \qquad (3.5)$$

Due to the time-varying nature of the parameters, however, this straightforward approach leads to discontinuities at the frame boundaries, if this approach is directly applied to speech as in [5]. Therefore, a method was found which smoothly interpolates the parameters measured from one block to those that are obtained in the next. Although, recovered speech in [5] is interpolated, this interpolation procedure introduces some back ground noise and block edge effects in the recovered speech. In SWELP system, we used the LPC filter to interpolate the sine wave components. This way all the discontinuities were eliminated from the recovered speech with the cost of coding the LPC coefficients.

4. BIT RATE REDUCTION STRATEGIES

Since the parameters of the SWELP system are the LPC coefficients, amplitudes, frequencies and phases of the underlying sine waves, and since for a typical low-pitched speaker there can be as many as 80 sine waves in a 4 kHz speech bandwidth, it is not possible to code all of the parameters directly. Therefore, an important focus of this work has been on techniques for efficient coding of the model parameters. The first step in reducing the size of the parameter set to be coded, was to develop a pitch extraction algorithm, which leads to a harmonic set of sine waves. The computed harmonic set is perceptual best fit to the measured sine waves. With this strategy, coding of individual sine wave frequencies is avoided. A new set of sine wave amplitudes and phases is then obtained by sampling an amplitude and phase envelop at the pitch harmonics.

In addition to the development of pitch extraction algorithm which led to a harmonic set of sine waves, a predictive model for the phases of sine waves was developed. The model given in [6] is quite accurate during steady voiced speech, while during unvoiced speech, it is poor, resulting in phase excitations that appeared to be random values within $[-\pi,\pi]$. For this purpose, we have developed another phase prediction model which works in both voiced and unvoiced speech regions. Since the speech coder is independent of v/uv decision, another parameter which is called error compensation component, has to be coded and transmitted to the receiver. The recovered LPC excitation was then presented as,

$$\hat{r}(n)=\sum_{k=0}^{L_m} A_k^m \cos((n-n_o)k\,\omega_o+\Phi_k) \quad (4.1)$$

where $\Phi_k=(-\Phi_o)^{k+1}$ is the phase and frequency error compensation component, n_o is the onset time of the pitch pulse [6] and ω_o is the fundamental frequency. Experiments showed that during steady voiced speech, Φ_o was automatically selected as zero, on the other hand, during unvoiced speech ${\Phi_o}^{k+1}$ took a value between $[-\pi,\pi]$. Since the amplitudes of the LPC excitation sine waves are more or less flat, a good criterion to use is the minimum mean-squared error for seeking the optimum values of n_o and Φ_o.

A block diagram of the complete analysis/synthesis system is given in figure 3.1. A non-real time floating point simulation was developed in order to determine the effectiveness of the proposed approach in modeling real speech. In SWELP system, the LPC and Spectrum analysis took place on block by block and sub-block by sub-block basis respectively. In LPC analysis using 30 ms (240 sample) block intervals (each consists of 2 sub-blocks), 7 or 8 LPC coefficients was found to be sufficient for smoothly interpolating the sine wave

components. A 512 point DFT using a 20-22 ms Hamming window was found to be sufficient for accurate peak estimation for each sub-block of LPC excitation. The overall bit rate of SWELP system is chiefly determined by allocating a certain number of bits for the LPC coefficients and sine wave components for each block of speech. We therefore, feel that by allocating more sine wave components to the excitation representation, the overall bit rate can be varied from 2.4 to 4.8 kbits/s or higher. This also varies the quality of speech. Thus, this scheme can easily be operated in a variable rate environment if required. A 2.4 kbits/s SWELP was simulated using vector quantization for both LPC coefficients and amplitude components of sine waves. The bit allocation for the coder implementation is shown in table 4.1, and the waveforms of the original and decoded speech are shown in figure 4.1. In this case, 2.4 kbits/s SWELP system produces natural sounding good quality synthetic speech.

Parameter	Bits Per Frame	Bit Rate
LPC Coeff.	13	500.0
Fund. Freq.	14	400.0
Phases (n_o,Φ_o)	18	600.0
Amplitudes	27	900.0
Overall	72	2400.0

Table 4.1 Bit allocations for 2.4 kbits/s SWELP coder.

5. CONCLUSION

In this paper, the advantages of CELP-BB, its problems and solutions were examined. We saw that the interpolation of the decimated sequences in the pitch filter memory improved the subjective performance of CELP-BB system. Below 4.8 kbits/s, however, it seemed that efficient representation of the model parameters for good quality speech is very difficult. Therefore, we presented the results of SWELP system which is used to represent the LPC excitation at very low bit rates (around 2.4 kbits/s) and produced good quality speech. The strategies for bit rate reduction in transmission parameters were described. Depending on the detailed bit allocation rules, operation at rates from 2.4 to 9.6 kbits/s can be obtained with the variation of speech quality. Thus, this scheme can easily be operated in a variable rate environment if required. Finally, a 2.4 kbits/s SWELP system was simulated and bit allocation of its parameters has been given as an example. The speech produced by this system was found very intelligible and natural sounding. Currently, the complete analysis/synthesis blocks of the SWELP scheme is being modified to form an Analysis By Synthesis (ABS) select procedure for reduced set of amplitudes, frequencies and phases. Results of this modification will be published later.

6. REFERENCES

[1] M. R. Schroeder, B. S. Atal "Code Excited Linear Prediction (CELP): High quality speech at very low bit rates" Proc. of ICASSP-85 pp.937-940.

[2] A. M. Kondoz, B. G. Evans "CELP base-band coder for high quality speech coding at 9.6 to 2.4 kbits/s" Proc. of ICASSP-88 pp.159-162.

[3] G. H. Asjadi, A. M. Kondoz, B. G. Evans "A real-time implemented 8 kbits/s CELP base-band coder" 7. Fase Symposium on speech, Edinburgh, 1988, pp.1039-1042.

[4] S. A. Atungsiri, A. M. Kondoz B. G. Evans "Robust 4.8 kbits/s CELP-BB speech coder for satellite-land mobile communications ", First European Conf. on Satellite Communications, September, 1989, W. Germany.

[5] R. J. McAulay and T. F. Quatieri "Speech analysis/synthesis based on a sinusoidal representation" IEEE trans. ASSP-34, pp.744-754 (August 1986).

[6] R. J. McAulay and T. F. Quatieri "Phase modeling and its application to Sinusoidal Transform Coding" IEEE proc. Int. conf. on ASSP, Tokyo, Japan, April 1986.

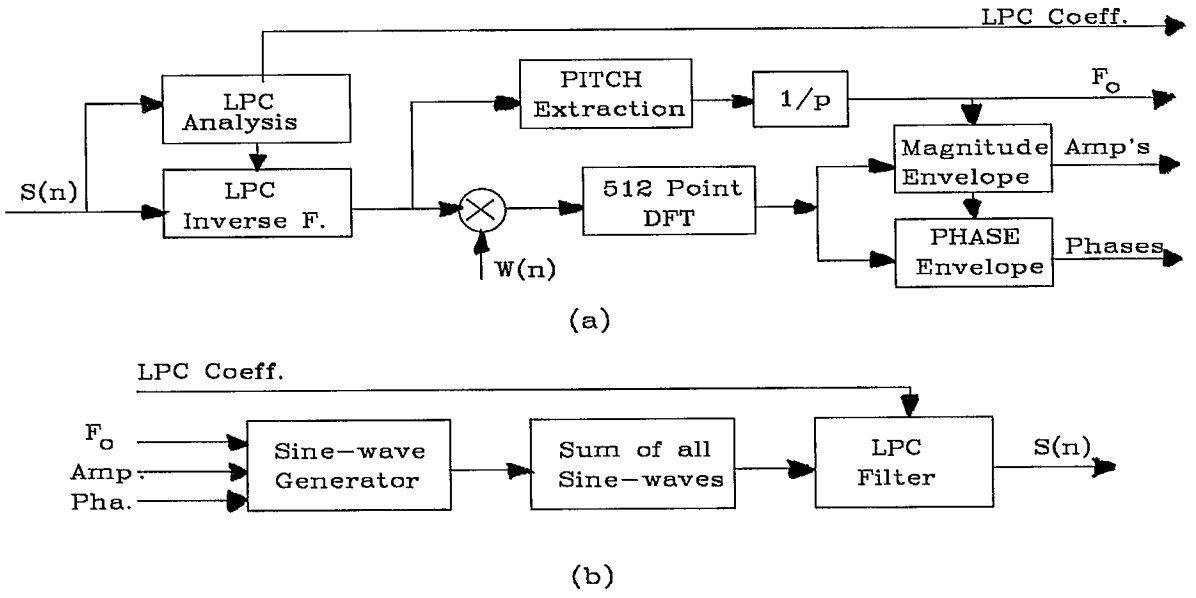


Figure 3.1 The block diagram of SWELP system a) Analysis b) Synthesis.

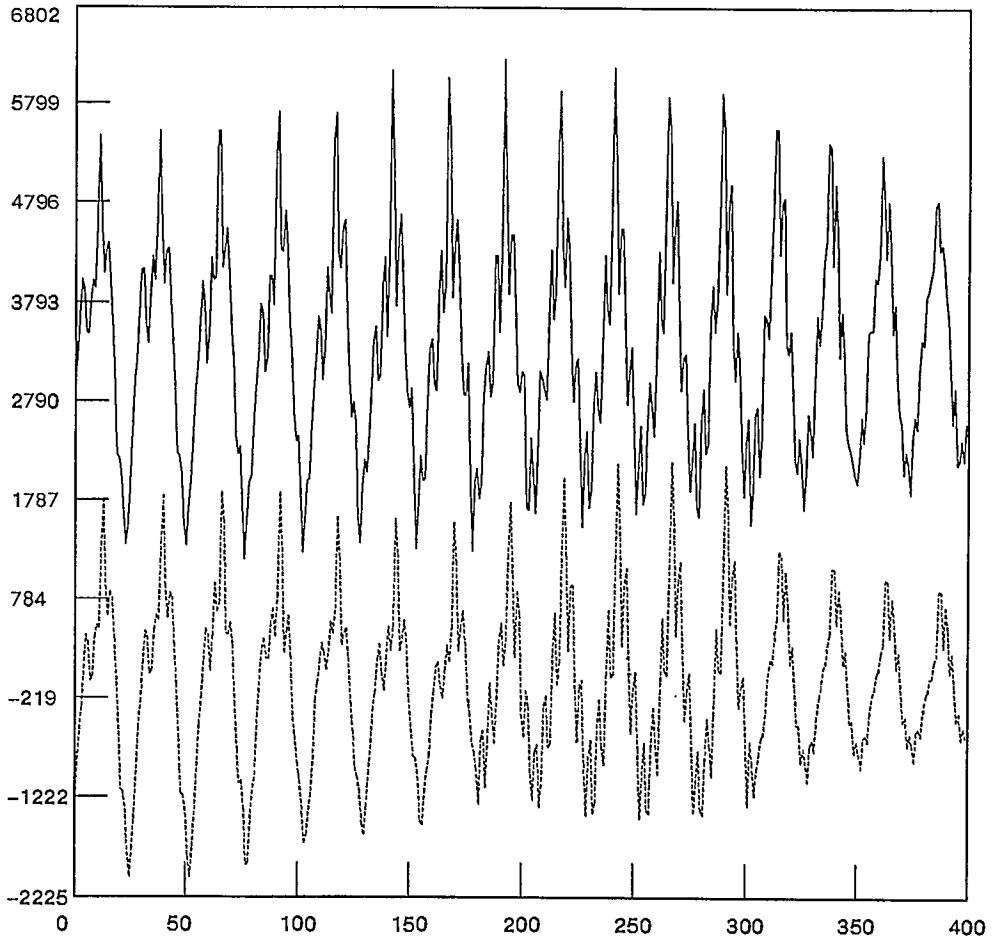


Figure 4.1 Typical waveforms of a) Original speech b) Recovered speech.

Evaluation of Voice Codecs for the Australian Mobile Satellite System

Tony Bundrock, Mal Wilkinson,
Telecom Australia Research Laboratories,
PO Box 249, Clayton, Vic.,
AUSTRALIA 3168
PHONE: +61 3 541 6421
FAX: +61 3 543 3339

ABSTRACT

This paper describes the evaluation procedure to choose a low bit rate voice coding algorithm for the Australian land mobile satellite system. The procedure is designed to assess both the inherent quality of the codec under 'normal' conditions and its robustness under various 'severe' conditions. For the assessment, 'normal' conditions were chosen to be a random bit error rate with added background acoustic noise and the 'severe' condition is designed to represent burst error conditions when the mobile satellite channel suffers fading due to roadside vegetation.

The assessment is divided into two phases. In the first phase, a reduced set of conditions is used to determine a short list of candidate codecs for more extensive testing in the second phase. The first phase conditions include quality and robustness and codecs are ranked with a 60:40 weighting on the two.In the second phase, the short-listed codecs are assessed over a range of input voice levels, BER's, background noise conditions, and burst error distributions. Assessment is by subjective rating on a five level opinion scale and all results are then used to derive a weighted Mean Opinion Score M_{total} using appropriate weights for each of the test conditions.

INTRODUCTION

The proposed Australian MOBILESAT system which is being developed by the Australian Government organisations, Aussat and Telecom Australia, is designed to provide circuit switched mobile voice/data services and packet data messaging services using the Aussat B series of satellites due to be launched in 1991/92. The economics and the practicality of this system dictate that the bit rate for the voice codec must be a minimum so that an individual voice channel will occupy minimum bandwidth and also require a minimum of satellite power. However, the quality of the voice channel must be adequate for interconnection to the PSTN and this must be maintained for the range of fading conditions expected on a mobile satellite channel.

There have been various low bit rate voice coding algorithms proposed for a mobile satellite system[1] but there has been no consensus on the best for this particular application. Therefore, to select the voice codec to be used for the MOBILESAT system, Aussat and Telecom Australia invited interested parties to submit codecs for an evaluation process to determine the one best suited to the specific requirements of a mobile satellite system. A document[2] providing details of the evaluation procedure was prepared and made available to all interested organisations. In late February of this year eight codecs were submitted to the Telecom Australia Research Laboratories, Melbourne, Australia, where the evaluation was conducted.

EVALUATION PROCEDURE

The voice codec evaluation and selection procedure is to be conducted in three phases. In the first two phases, a digital error generator test set is used to simulate the land mobile satellite channel. This test set enables evaluation of the inherent performance of the codec separate from the modem performance because it enables a baseband connection between coder and decoder. The test set is used to introduce random and burst error patterns. During the third phase, a QPSK modem and an analog mobile satellite channel simulator is used to carry out additional evaluation of the joint performance of the codec and modem. The QPSK modem also provides soft decision decoding outputs and differential encoding.

The first phase is designed for a quick evaluation of all codecs to determine a short list for a more extensive program in the second phase. During this first phase, speech quality and robustness is subjectively rated using a five point opinion scale. Quality is assessed by two conditions; (a) a random bit error rate of 10^{-3}, and, (b) an acoustic SNR of 15dB with zero channel error rate. Robustness is also assessed by 2 conditions; (a) a burst error pattern (described below), and, (b) a random bit error rate of 10^{-2}. A final rating is derived by determining a weighted average of the four conditions with a weighting of 60:40 in favour of the quality conditions.

The second phase assesses codecs over a wider range of conditions testing the effects of various random BERs, burst error patterns, voice input level and acoustic background noise. A final overall rating is then determined by using various weightings for each condition and the codecs are ranked accordingly. The codecs are also assessed for complexity, transmission delay, tone handling capability and tandem operation. Spot checks are also made on the performance with carbon microphones.

The source material for the evaluation consists of Harvard sentence pairs spoken by two male and two female English speaking talkers with clearly different voice characteristics. The recordings are made with a linear microphone with spectrum weighting according to the Intermediate Reference System (IRS) specified in CCITT Recommendation P.48. All source material was normalised to the same level prior to playing through test codecs and all output is also adjusted to the same level so that a constant listening level is maintained.

The recorded test speech samples for the various codecs are randomised prior to the listening tests. The listening takes place in a quiet room using an Australian 200 series handset. The listening level is calibrated so that the average speech level of -10 dBPa is obtained with an IEC artificial ear. The assessment is based on a five point Absolute Category Rating (ACR) method with a scale of; (a) Excellent, (b) Good, (c) Fair, (d) Poor, and, (e) Bad, which are scored 5 to 1 for analysis purposes.

Reference conditions for the evaluation are obtained by using a Modulated Noise Reference Unit (MNRU) test unit. Eight MNRU dBQ values are used as reference conditions and these are randomised with the codec samples for the listening tests.

All test material was prepared using a computer controlled test fixture. The computer ensures all codecs are tested under identical conditions. For instance, the identical burst error pattern is applied for each codec and the passage of acoustic background noise is also identical. A standard PCM I/O interface was used between the codec and the source and test recordings.

BURST ERROR MODEL

In order to realistically test the codecs, a suitable channel model must be assumed which is able to reproduce the error phenomena expected to occur on the channel. Errors are produced on the channel by thermal noise when there is clear line of sight to the satellite and in this case the channel can be modelled as a memoryless Binary Symmetric Channel. In addition, the occurrence of roadside vegetation will attenuate the signal and produce error bursts. The length and severity of these bursts will depend on the fade duration and depth.

To reproduce the error phenomena in the codec test set a Markov model of the channel is used. In this model, the channel is described by a number of states each with a different BER. Transition probabilities are determined for each state and the channel may be viewed as one which switches back and forth between several BSCs, each with an associated BER. The Markov parameters were chosen from data which have been measured for various channels by the Telecom Australia Research Laboratories[3]. During the testing, burst error models for light and heavily shadowed channels are used.

INMARSAT

While Aussat and Telecom Australia were preparing for the evaluation and selection of a voice codec for the MOBILESAT system, Inmarsat were pursuing a very similar course for the Inmarsat-M system. Late in 1989 Inmarsat issued an RFT for an organisation to conduct an evaluation procedure of codec on its behalf. With a goal of producing a common standard for a codec for mobile satellite systems, Australia and Inmarsat discussed the two evaluation procedures and made changes to both so that comparisons can be readily made. Telecom Australia Research Laboratories also submitted a bid to conduct the Inmarsat evaluation and were awarded a contract for this purpose in January 1990. With some minor variations, the procedure described here is the same for the Inmarsat evaluation and the codecs submitted for evaluation are the same for both.

RESULTS

Eight codecs were submitted to the Telecom Australia Research Laboratories for evaluation. British Telecom also provided a codec using the 9.6 kbit/s algorithm chosen for the Inmarsat aeronautical service to provide a reference to the results obtained for the lower bit rate codecs. At this time, phase one of the evaluation has been completed and five codecs selected for the extensive phase two testing. The results are shown in Figures 1 and 2. Figure 1 shows the results obtained for a 10^{-3} BER averaged over all four speakers. The error bars shown are the 95% confidence intervals. From Figure 1, we see a fairly close grouping with most codecs rating a MOS of better than 3 (Fair). For this test codec 3 ranks highest although, at the 95% confidence level, it is not statistically separate from codecs 1, 2, 4, 5 and 7. Figure 2 shows the results obtained when a burst error pattern corresponding to moderately severe shadowing (85% tree coverage) is applied. In this case, codec 3 is clearly superior. For the phase 2 evaluation, codecs 1 to 5 have been chosen. Figure 3 shows the MOS results obtained for the MNRU reference.

ACKNOWLEDGEMENTS

The contributions of the following are gratefully acknowledged: Mushfiq Rahman for the development of the Markov channel model and for planning the experimental sequencing ; Alistair Urie for developing the computer controlled test fixture and his general contributions on codecs; Dan Cerchi, Adrian Martinus, and Steve Beyer for organising the source and test material; and Barry Thomas for conducting the subjective testing.

REFERENCES

1. **Rafferty, W.** 1988. Speech Compression Session, *Proceedings of the Mobile Satellite Conference*. JPL 88-9, (Pasedena, California; The Jet Propulsion Laboratory)

2. **Wilkinson, M. H.** 1990. *Voice Codec Test and Evaluation Procedure*. VCTEP/2 Telecom Research Laboratories, Australian Telecommunications Corporation, Melbourne Australia.

3. **Bundrock, A. J., and Harvey, R. A.** 1989 Propagation Measurements for an Australian Land Mobile Satellite System. *Australian Telecommunication Research,* Vol 23, No 1.

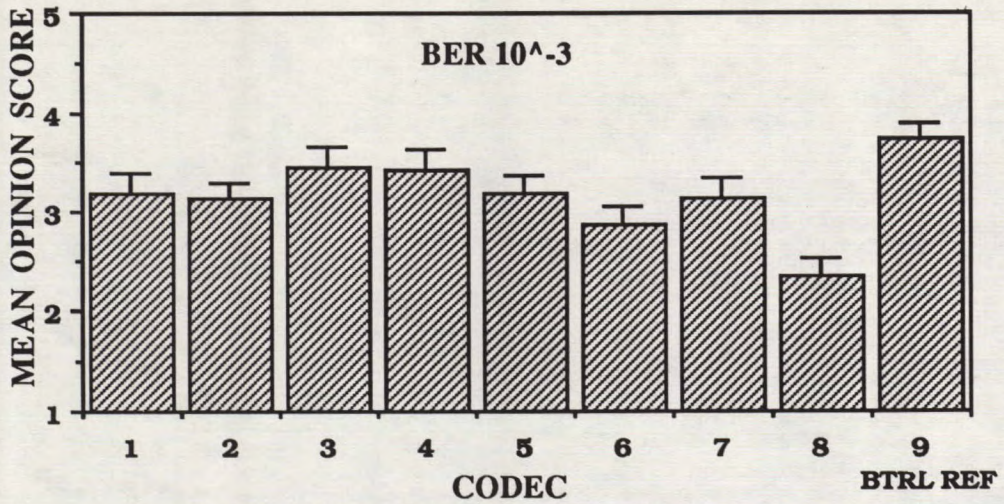


Fig. 1 Mean Opinion Score for Codecs

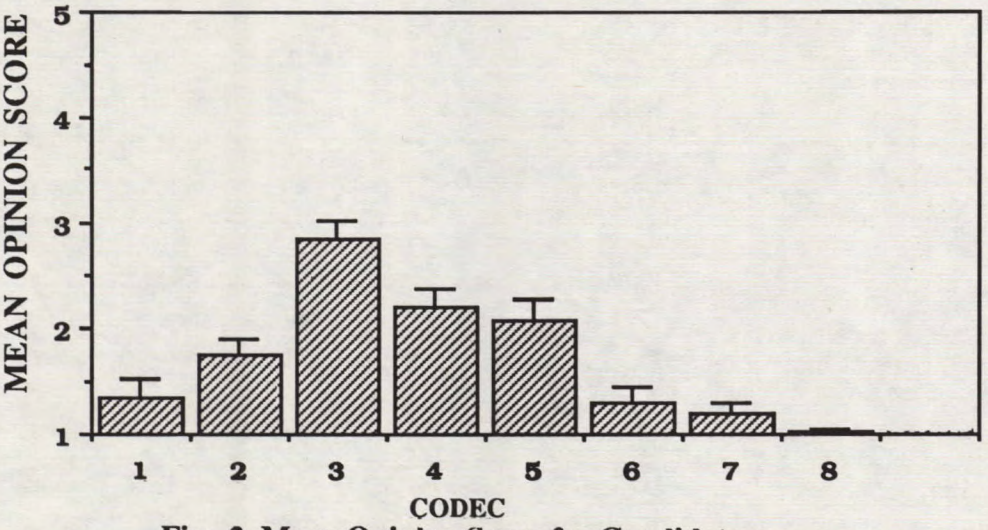


Fig. 2 Mean Opinion Score for Candidate Codecs with Burst Errors

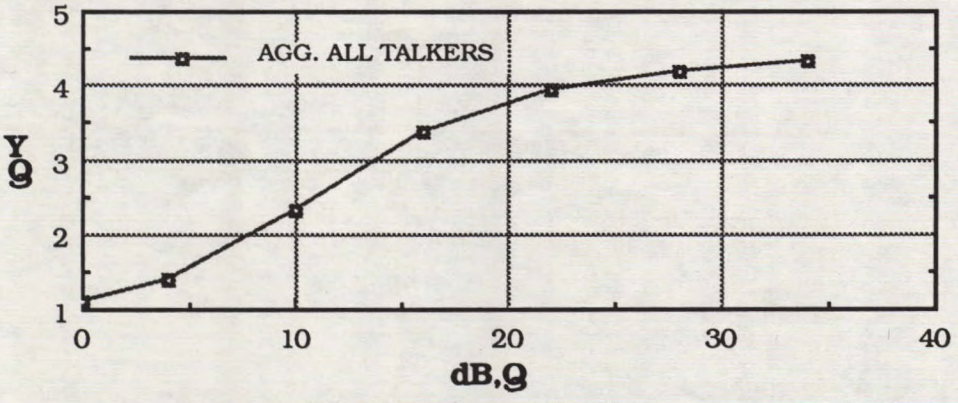


Fig. 3 MNRU Results for Test Listeners

Session 16
User Requirements

Session Chairman - *Jerry Freibaum,* Senior Advisor to PSSC, USA
Session Organizer - *Allister Pedersen,* DOC

MSAT an Alternative Choice?

K.E.Fagan and C.Elmer
Telesat Mobile Inc.
P.O. Box 7800 Ottawa, Ontario
K1L 8E4, Canada
Phone: (613) 746-5601
Fax: (613) 746-2277

ABSTRACT

The paper describes a number of potential applications of MSAT that utilize the unique properties of this transportation mechanism. Emphasis is placed on the market introduction strategy for the North American system.

INTRODUCTION

Early in the 1970's the use of a Mobile satellite concept was proposed utilizing Spot Beams and Demand assignment Multiple Access technology. However it was not until the 1980's that technology enabled the concept to be taken into a viable business scenario that would allow the launch and management of dedicated satellites for the purposes of Mobile Satellite communication.

As with any potential new service the wheels of progress can move very slowly and it has been only in the past two years that the necessary conditions have been achieved to be able to propose the concepts to the investment communities. Summarized those necessary conditions included :-

a) The availability of spectrum and the agreements necessary to use it for Mobile Satellite services,

b) The degree of cooperation required to allow North America wide roaming,

c) Identified demand for such a service offering

d) Advances in technology to enable a viable business plan in terms of a reasonable cost to the end user.

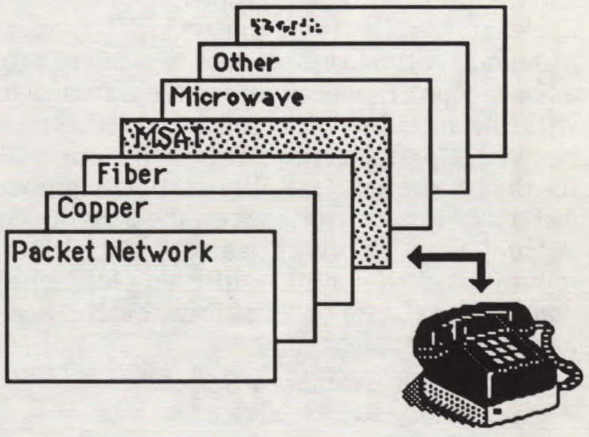


Fig 1
MSAT as an alternate long-line choice

Introducing a new service is a function of the end-user demanding solutions to a perceived problem. In Canada the need for reliable mobile communications in the non-urban areas has been a prime driving force for the MSAT developments. Similarly rural fixed communication has never enjoyed the same ease of use nor the potential for service as that enjoyed by the urban subscribers, and pressure from many rural customers has forced telephone companies to seek innovative solutions to this problem. This has been seen in the past few years in the ILS (Individual Line Service) programs of many of the telephone companies. This, plus the deregulation of many of the traditional telco services has placed enormous pressures on the service providers. MSAT technology (in a non-mobile application) is seen as a way to help relieve those pressures by providing remote, single-channel, low traffic service.

Service Offerings

The easiest way to introduce a new service is to directly emulate existing services, and to do it in such a fashion that the subscriber is unaware of the specific transport mechanism being used. Many would advocate that this strategy removes the visibility of the provider and therefore a competitive advantage. However if MSAT is to be recognized as truly an alternate long-line choice then it must emulate and be compatible with the existing services initially and gain competitive advantage by providing those service in a more cost effective or efficient fashion for specific applications.

The Mobile Satellite proposed for use by TMI (Telesat Mobile Inc.) and AMSC (American Mobile Satellite Corp.) is such that it will provide a "bent-pipe" function only at the Satellite itself, with the intelligence being provided through a Network Control Center. This architecture allows for the maximum flexibility in the assignment of different "Bearer" services on the satellite and to define new services as opportunities are identified. In keeping with the philosophy of emulating existing services four "bearer" services have been identified.

a) Mobile Telephony Service
b) Mobile Radio Service
c) Mobile Data Service
d) Aeronautical Services

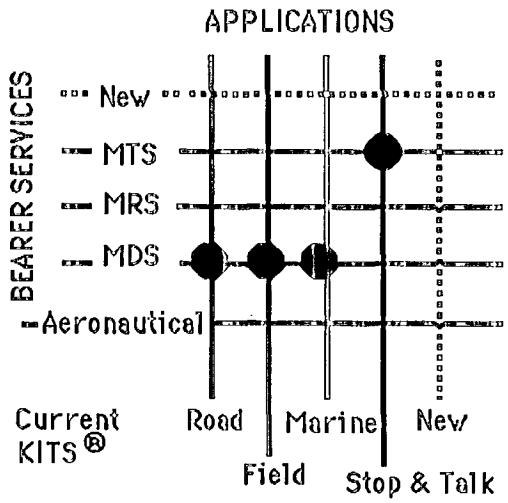


Fig 2
MSAT Bearer service philosophy

Due to the specialized nature of the Aeronautical services for Commercial applications little time will be spent discussing it in this paper. It is sufficient to say that these services can and will be provided on MSAT through the approved agencies.

It is Telesat Mobiles Inc. intentions not to be a manufacturer of hardware, and as such has no vested interest in the recommendation of a specific set of hardware or end-user services. To enable manufacturers to benefit from the MSAT opportunity TMI will make available the specifications of the interfaces to the network with the encouragement for Distributors and resellers to design their own applications to meet their specific customers requirements. These specifications are the intent of a recently released Request for Proposal that will define the final system achitecture for the Ground Segment

The proposed architecture of the MSAT Network is beyond the scope of this paper but is covered in other papers presented in other sessions at this conference.

APPLICATIONS

If one considers the capability of MSAT in both its coverage area and from the fact that it is designed to be used with small inexpensive Mobile Earth Terminals (METS), the scope of potential usage is only limited by the imagination of the intended users. Early applications that have been designed include the following:-

Voice calling. Communication to and from mobile or transportable earth terminals.

Vehicle Location. Automatic Vehicle Location (AVL) will allow a Central dispatcher to monitor the position of a fleet of vehicles, with periodic updates. This will be accomplished using a navigation device on the vehicle,early systems will use Loran-C although other systems such as GPS.are not precluded This can be coupled with other navigation services to allow for optimum routing of a vehicle.

Two-way general messaging. Messages can be sent between a MET and public or private data networks. An electronic mail box would be a value added service to store messages if a MET is unavailable.

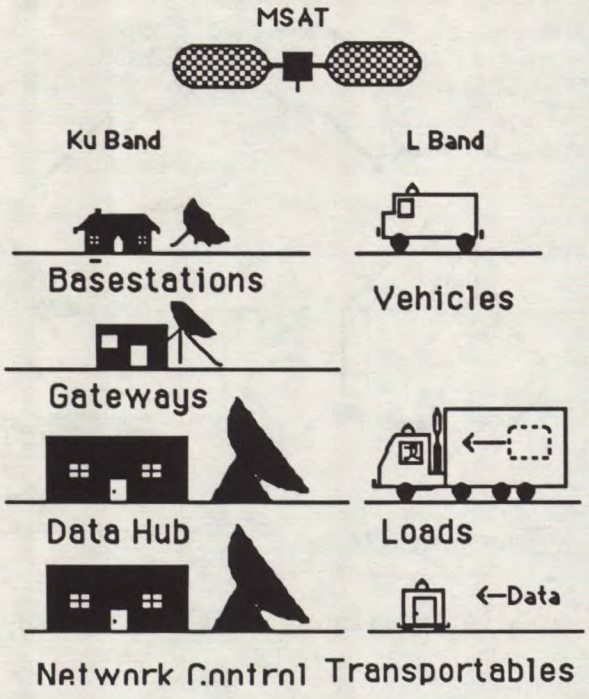


Fig 3
MSAT Vision

File transfer. This would provide for the MET to exchange relatively long files with the Data Hubs. It will be an economic tradeoff as to whether a file should go as packets on the MDS or demand a circuit via the MRS/MTS.

Interactive data. In order to allow for interactive data sessions, the MDS will have the capability of guaranteeing a specified response time. This will be done either on a priority basis or by assigning a TDMA slot for the MET.

Monitoring and control. Supervisory Control and Data Acquisition (SCADA) will allow for monitoring and control METs on a real-time basis.

Virtual circuits. A MET will be able to request a circuit of variable throughput in order to establish a 'virtual circuit'. This would give a user a TDMA slot into which they can put whatever data they wish. From the Data Hub to a MET, the user can be allocated one of the Time Division Multiplex slots on a demand-assigned basis.

The above are Generic capabilities of the MSAT achitecture however if we start now to combine these requirements with other peripheral units the potential is compounded:-

Load Tracking. .... A MET used on a transport vehicle tells you where the wheels are!! if we interface with the MET the capability to append auxiliary identification tags then specific loads can now be tracked. Already there has been great interest shown in the tracking of hazardous goods, High value cargo's etc.

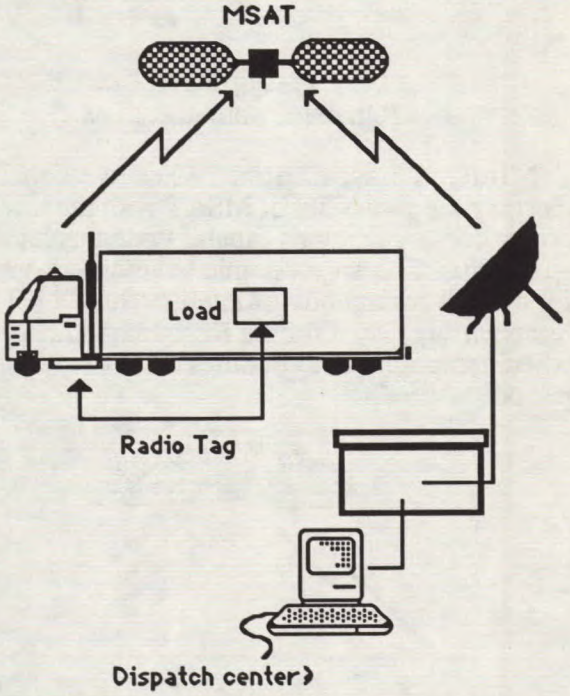


Fig 4
Secondary data acquisition

Extending this capability in a slightly modified form, will allow for the capability of continent wide tracking of hazardous or high value goods.

Remote Communications. Provision of quality communications into areas of the Canadian geography where a traditional Land-line solution is uneconomic economic is both possible and attractive. Since the MSAT service is distance independent communication to the more densely populated areas is very economic.

Facsimile Transmission. The explosion of Facsimile usage over the past five years has made the Group 3 Fax a "normal" office requirement. In keeping with this the MSAT Earth Terminals are being designed to accommodate this service.

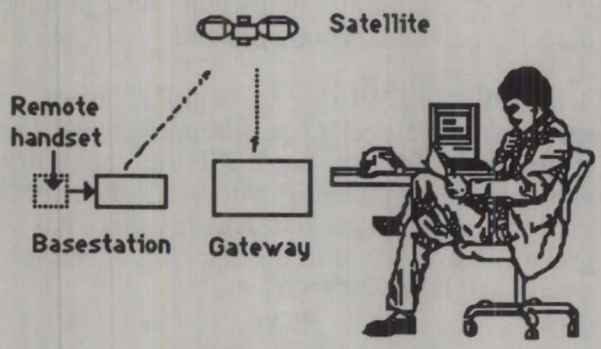


Fig 5
Remote Communications

MiniCell basestation. When we consider merging the capability of MSAT with the low power communications capability of a technology similar to CT-2, an economic solution is possible for a small community of interest. Such a solution has been suggested for the Resource industry where temporary communities can exist for short periods of time.

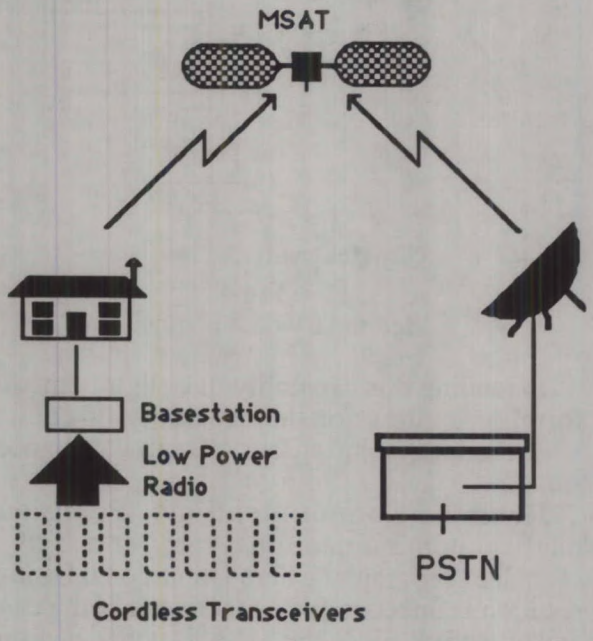


Fig 6
MSAT as a Community of interest base station

Emergency Services The transportability of the MSAT mobile Earth terminals is attracting a lot of interest in the Emergency measures organizations. With the TMI equipment temporary communication can be rapidly set up in the most inhospitable of areas.

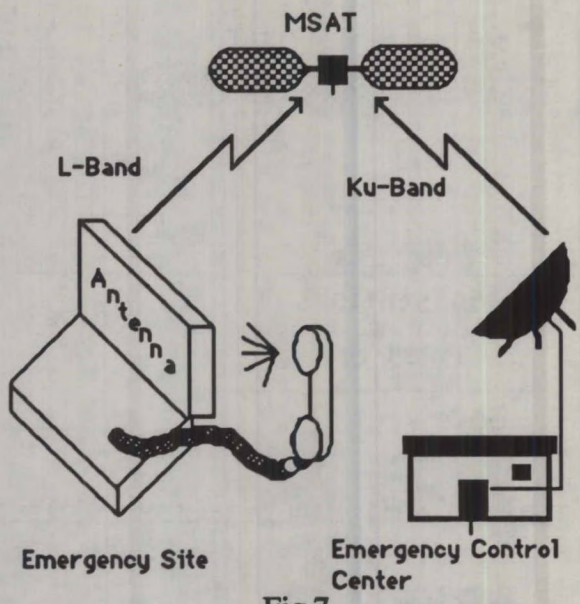


Fig 7
MSAT as an Emergency Communications facility

Conclusions

It is difficult in a short paper to address all of the capabilities of the system offered by TMI and AMSC. Hopefully this paper has indicated some of the interest being shown by the potential users and some of the solutions being innovativly applied using the MSAT facilities.

User Applications Unique to Mobile Satellites

David Castiel
American Mobile Satellite Corporation
1233 20th Street N.W.
Suite 301
Washington, D.C. 20036
Phone: 202-331-5858
FAX: 202-331-5861

As AMSC enters the market with its mobile satellite services, it faces a sophisticated user group that has already experimented with a wide range of communications services, including cellular radio and Ku-band satellite messaging.

AMSC's challenge is to define applications unique to the capabilities of its dedicated L-band satellite and consistent with the provisions outlined in its FCC license.

Through a carefully researched approach to its three main markets -- aeronautical, land-mobile, and maritime -- AMSC is discovering a wellspring of interest in corporate and general aviation, trucking companies, pipeline monitoring and control companies, maritime management firms, telecommunications companies, and government agencies.

This paper will provide a general overview of AMSC's FCC licence and corporate history, and discuss the specific applications unique to each user group.

Today's conference is a testament to the fact that there is tremendous interest within the scientific and governmental communities in mobile satellite services (MSS). For the American Mobile Satellite Corporation (AMSC), a company on the cutting edge of this technology, a primary challenge is to succinctly define user applications so that the mass market will embrace this new technology.

Mass market acceptance of MSS should be of great interest to the government and scientific communities because of the accompanying decreases in mobile earth terminal, space, and ground segment charges that a large consumer and business user base will make possible.

As AMSC enters the MSS market, we face a relatively sophisticated user base that has already been exposed to satellite technology's mobile applications, albeit via systems that are not pure MSS in the L-Band. The Geostar and Qualcomm systems, for example, together have approximately 20,000 Ku-band land mobile satellite terminals on order. Qualcomm is the market leader in applications of satellite positioning and data messaging to the land-mobile sector, with 7,000 units deployed, primarily to large trucking companies, and Geostar follows with 2,000 units. These units are relatively low priced at a cost of under $4,000 per unit. AMSC's mobile earth terminals, which will be available in November for beta testing, will be priced similarly.

Qualcomm and Geostar have achieved these numbers in just over two years, whereas the INMARSAT system needed 13 years to reach a deployment level of 10,000 mobile, primarily maritime, units, at a substantially higher cost of about $30,000 to $40,000.

As AMSC prepares to offer its low speed data messaging and position determination services in the fourth quarter of this year, our challenge is to define unique, value-added services that can serve our three principal markets: land-mobile, which includes trucking, private automobile, oil well drilling, pipeline management and government applications; maritime, which includes ships as varied as Coast Guard cutters, large and small yachts, and inland waterway barges; and aeronautical, which includes services for operational, administrative, and passenger communications.

Before looking at an overview of how AMSC will address these challenges, a review of AMSC's corporate history would be appropriate.

The mobile satellite technology that will be utilized in the AMSC system is an outgrowth of years of NASA experimentation with mobile satellite technologies, dating back to the 1970's.

NASA took the lead in identifying a market for MSS and later petitioned the Federal Communications Commission for frequency allocations. Industry interest in the U.S. was so intense that within a two year period, from 1983 to 1985, twelve applications for MSS service were filed at the FCC.

The FCC found that in order to use the limited MSS spectrum most efficiently, the applicants should together form a consortium.

Eight of the original twelve applicants agreed with this principle, and in 1988 they formed the American Mobile Satellite Consortium, now known as the American Mobile Satellite Corporation. In 1989 AMSC received its FCC license authorizing it to function as a common carrier for the provision of domestic mobile satellite services to air, land, and maritime users.

AMSC functions in a unique environment. We at AMSC share the excitement and challenges common to many new companies, while at the same time we benefit from a corporate parentage of established, major telecommunications companies, namely Hughes Communications Mobile Satellite Services, Inc.; McCaw Space Technologies, Inc., a subsidiary of McCaw Cellular; MTEL Space Technologies Corp., the owner of the Skytel nationwide satellite paging system; Mobile Satellite Corporation; North American Mobile Satellite, Inc., whose parent company is Millicom Cellular; Satellite Mobile Telephone Company, Skylink Corporation, and Transit Communications, Inc. It's worth noting the large cellular and paging companies which are backing the development of L-Band mobile satellite services.

AMSC's FCC license authorizes us to operate three spacecraft to provide all domestic mobile satellite services -- land, sea, and air -- to users in the continental United States, Hawaii, Alaska, Puerto Rico, the Virgin Islands, and 200 miles of coastal waters. The AMSC license obligates AMSC to construct and launch a mobile satellite system that will serve as a national communications resource

for the United States of America.

Clearly, mobile satellite services offer universal coverage while moving. Not only will AMSC provide complete coverage of the U.S., even in rural and mountainous areas, but through an agreement with Telesat Mobile Inc., the Canadian operator of a domestic mobile satellite service, users will receive seamless service while on the go between Canada and the U.S. In addition, we are negotiating an agreement with the Mexican government to provide land mobile services. A seamless hand-off system will interconnect domestic aviation using the AMSC system and international aviation.

The AMSC satellites will be high capacity birds, with central load sharing and an effective EIRP much higher than in other mobile satellites. While the exact parameters of our satellite are still being defined, we anticipate that AMSC's MSAT 1 spacecraft will have a nominal power of 55 dBW.

AMSC will optimize usage of our space segment by offering service to a wide variety of user groups. We have defined several major applications that we are now actively marketing to our three major markets. In the land-mobile sector, these applications include emergency services, car theft prevention, interstate trucking, news gathering, oil exploration, supervisory control and data acquisition, and environmental monitoring. Aeronautical users will be able to benefit from services for passenger communication and cockpit data such as airline operational and administrative information. Many of the same applications common to the land-mobile market -- fleet management, passenger and employee communications, and data transfer -- will be used in the maritime arena. The applications for this technology will continue to multiply as it is implemented, and users discover new solutions for recurring communications problems.

AMSC voice services feature toll quality, communications quality, and emergency services. A variety of data services facilitate the transmission of position reports and "canned" messages, short messages, maps, facsimiles, and forms.

A unique aspect of the AMSC system vis a vis other systems is our ability to establish private networks with private base stations and gateways to the public switched telephone network, public data networks, and other networks. It is through a private network that Rockwell International, our first service provider, will provide mobile satellite services to enhance an existing service, known as the Tripmaster system, that Rockwell markets to the trucking industry.

Data applications will be the hallmark of AMSC's pre-launch services, which we will introduce in the fourth quarter of this year. Services will feature 2-way data messaging at speeds of 300 bps from the mobile earth terminal to the satellite and 600 bps from the satellite to the mobile earth terminal, Loran-C for position determination, and a variety of transmission options, including scheduled, unscheduled, and variable length.

The long-haul trucking industry is one of our main markets for data services. Data messaging applications in the trucking industry include fleet management systems, expert systems for dispatch and other operations, customer

service, real-time billing, real-time crediting to drivers, and emergency road services. Data messaging and position location services also will be available after the late 1993 launch of our satellite, which will provide two-way voice and high speed data services. We will tailor our services and support systems, such as network control and earth station facilities to the individual needs of each customer. Our four primary service packages are full period channel, shared channel, virtual circuit, and packet network. The full period channel service provides customers with dedicated channels, and the customers themselves provide the earth station and network control center. With shared channel service, AMSC provides customers with virtual network service, fixed period circuits, and demand capacity. The customer provides the earth station. Virtual circuit service provides full voice capabilities and interactive data at speeds up to 9600 bps on a call per minute basis. Our packet network service provides standard 128 byte packets and broadcast messages for mobile response.

AMSC's voice services will not directly compete with cellular services, but will supplement cellular coverage in areas where it is not readily accessible, such as rural areas and remote industrial sites. Our North America-wide voice coverage will enable a car phone's coverage range to leap from a local cell across the continent, forming a seamless, continental cell. AMSC's provision of both voice and data services also will allow long-haul transportation companies, such as trucking, bus, and railroad operators to access a wide range of services not available via conventional cellular phones.

Through AMSC's private networks, users have direct access to satellite capacity for voice and data communications. Via this direct access system, they can operate a network of one or more private base stations which operate separately from the public mobile network and feature interconnection to the public switched telephone network. This system allows our customers to take advantage of economies of scale available through management of their own networks while at the same time benefiting from access to our space segment.
According to our FCC license, AMSC is obliged to provide priority and pre-emptive access to our space segment for aviation flight safety. Our services will feature an interface to FAA facilities for voice and data communications, and compatibility with Arinc 741, as it is adopted. Communications with government and other aircraft will facilitate air traffic control, operations, administrative, and passenger communications.

Government interest in our services is strong. In addition to government aviation applications, government agencies will use our services to improve emergency search and rescue operations; to enhance communications during sensitive law enforcement actions; to monitor dams and reservoirs; and for Coast Guard operations.

Under the terms of our launch agreement with NASA, certain government agencies will receive free experimental capacity on our satellite for a fixed period of time. NASA and AMSC sponsored a government seminar in March, at which nearly 38 agencies met to discuss their applications needs. We're witnessing a tremendous outpouring of interest in our services.

We also hope to integrate our services into the new FTS 2000 phone system that the U.S. government is implementing.

To summarize, I would like to reiterate that insofar as the data market is concerned, we see AMSC's unique service offerings as stemming from our ability to establish private networks that offer entrepreneurial service providers the opportunity to bundle their services with ours, thereby offering a new services. AMSC is embarking on a value-added program to work with service providers in the aeronautical, SCADA, maritime, and trucking markets.

With regard to the voice market, we see AMSC as a service that will complement and supplement cellular service, reaching areas even more remote than the 428 Rural Statistical Areas that are slated to receive cellular services.

The applications that I have outlined for you today represent a true revolution in the way North Americans will communicate and the way our businesses will be run. This is just the beginning. As we continue to tell our story to users throughout our country, we are finding new, niche market applications that we had never before envisioned. The market for AMSC's services is fast-growing and ever-changing, and I look forward to keeping you apprised of our successes.

MSAT: A Booster for Land Based Mobile Radiocommunication Networks?

Guy Boulay
Bell Canada Enterprises Mobile Inc. (BCE Mobile)
7350 TransCanada Highway
Ville St-Laurent, Québec, Canada H4T 1A3
Phone: (514) 345-9440
FAX: (514) 345-9434

ABSTRACT

This paper intends to demonstrate that the foreseen phenomenal growth of mobilesat services will impact positively existing terrestrial mobile radio services. Mobilesat systems will not displace the existing terrestrial market in the near future, partly due to the high cost of their terminal units and associated airtime, but also due to some technical limitations, such as lack of spectrum efficiency and high susceptibility to shadowing. However the ubiquity of mobilesat services will open new markets to terrestrial radio technologies: the latter is expected to be the most economical way of extending locally mobilesat services to many users. Conversely, Mobilesat systems will be used to extend the capabilities of terrestrial radio systems in areas where the former cannot be implemented cost efficiently. It is believed that terrestrial mobile networks operators using these service extension capabilities will have a competitive advantage over those who do not.

Overall, it is expected that emerging mobilesat services, far from being a threat to terrestrial radio systems, will rather provide these with numerous opportunities of incrementing their base market.

FOREWORD

This paper reflects the opinion of a Canadian operator of terrestrial mobile networks (cellular, paging, mobile data and private networks) regarding the potential impact of the North American MSAT system on the land based mobile radio industry. This operator is also involved in the distribution of mobilesat products and services.

INTRODUCTION

Between now and 1994, 12 satellites dedicated to mobile satellite communications will be launched world-wide. The associated investments are in the order of $US 5,000,000,000 [1]. Forecasts indicate that the actual level of 20,000 satellite mobile units in service at the end of 1989 will increase to 80,000 units in 1993, and to 2,000,000 units in 1995, over the entire globe [2]. By then, the full range of mobilesat voice and data services will be available, namely via the North American MSAT system. The highly flexible architecture of such a system will allow it to emulate virtually any type of terrestrial mobile service: cellular, paging, mobile data, two-way radio, etc. Unlike these services, MSAT provides a complete and seamless North American coverage, even in the most remote areas. What will be the net effect of this definite advantage over the apparently limited land based networks?

Rather than displacing a significant share of the existing market, mobilesat communications are likely to impact positively the penetration of their terrestrial counterparts. Hence, this paper proposes to demonstrate the anticipated impact of the mobilesat systems on the existing terrestrial mobile radio services..

A COMPARISON BETWEEN MOBILESAT AND LAND BASED SYSTEMS

Terminal Costs. Any new technology or system is generally more expensive at its introduction, largely because it lacks the economy of scale associated to mass production. The Mobilesat technology will not be any different. Moreover, the lack of world-wide standardization amongst mobilesat systems will further delay any savings for close to a decade, as a universal mobilesat standard is not foreseen in the near future. Each family of fairly complex mobilesat terminals is and will be produced in much smaller quantities than their terrestrial equivalents. For instance,

close to 7 million cellular terminals will be put in service between now and the end of 1993 in the United States alone [3,4]. This is much more than the total of all types of mobilesat terminals that will be sold in the entire world during the same period of time. Similar relative order of magnitudes would hold true for SMR's and private systems.

Airtime Costs. In addition to the retail price of the terminal, the tariff associated with mobilesat airtime is likely to remain higher than the terrestrial systems. This is partly due again to the lack of world-wide standards, causing each system to incur large up-front non-recurring development expenses. The high price is mainly occasioned by the very high start-up costs inherent to any satellite system.

$CDN 3,000 to $4,000 is a typical projected figure used for planning purposes of a voice mobilesat terminal. In comparison, many cellular phones sell already for less than $CDN 500. Typical planning figures for airtime mentioned by mobilesat operators vary from $CDN 1.00 to $2.00 a minute for voice, compared to about $CDN 0.50 for cellular usage in Canada. Similar comparisons also hold true for data services.

Susceptibility to Shadowing. Land based mobile systems generally benefit from a more generous link budget margin than mobilesat systems. Satellite power is expensive and limited: for a fixed total satellite power, doubling the power per channel (3 dB increase) implies that the total satellite capacity (and airtime revenues) would be reduced by a factor of two. For instance, a satellite that could provide service to 100,000 mobile users with a fade margin of 3 dB could service only 50,000 users with a fade margin of 6 dB. That is why these systems are designed for relatively low fade margins (typically 3-10 dB). In addition, mobilesats generally operate in the 1.5 GHz band, and are thereby more affected by the shadowing effect of nearby objects than mobiles operating in terrestrial bands (150 to 900 MHz). These two factors are responsible for the expected high susceptibility of mobilesat to shadowing. This could actually force a mobile unit to remain stationary during voice communications in shadowed areas (roads with adjacent trees, downtown and suburban areas, hilly areas, etc.). Similarly, mobile data users would experience a non-negligible reduction of their throughput when moving in the same areas.

Spectrum Efficiency. Despite the very narrow 5 kHz channel spacing planned for the North American MSAT, satellite systems are far less spectrum efficient than their terrestrial counterparts: even with a spot-beam technology, a specific frequency can barely be reused once over the North American continent, at L band. On the other hand, cellular systems can reuse a specific frequency hundreds of times over that same area, as can other terrestrial systems, although not as efficiently as cellular. RF spectrum is a limited resource: the system able to carry more calls per MHz for a specific territory is technically capable of achieving a larger market share. This definite advantage of land based systems over mobilesat is proportional to the global call carrying capacity of each system over a defined territory. A similar rationale applies to data transmission.

Propagation Delay. Satellite delay, well known to those who call frequently overseas, is likely to incite some mobile users to use voice terrestrial systems, where available. Data transmission is not affected by satellite delay, except for some rare real time remote control applications.

Mobilesat and Their Terrestrial Counterparts: Two Distinct Markets

We believe that the higher retail prices of mobilesat communications and some of the inherent technical constraints described above are likely to keep most of these terminals in areas where there are no terrestrial alternatives. The net impact of mobilesat on existing land based mobile networks is thus likely to be minimal.

One noticeable exception to this statement is the niche market of long haul vehicles (trucks, trains, airplanes): these users need continuous coverage, and seamless operation, which terrestrial systems cannot adequately provide at the moment. It is mainly in this niche market that the non-maritime mobilesat operations are actually growing. However, the continuous expansion of land based systems and their eventual mutual interconnections is likely to generate serious competition to mobilesat systems in this specific area, in the near future.

SYNERGY BETWEEN TERRESTRIAL AND SATELLITE SERVICES

Despite the constraints and limitations mentioned above, the growth of mobilesat communi-

cations is predicted to be phenomenal by many observers: mobilesat terminals will be found all over the world within the decade. Person to person communication will be possible anywhere in the world.

The phenomenal growth of cellular systems in the recent years caused a positive impact on private mobile systems and SMR's: cellular systems taught people that they could stay in touch at any time. Many users then chose non-cellular systems to fill this recently discovered need of communications.. Similarly, the new non-geographically limited mobilesat communications services will reenforce the image of "keep in touch everywhere". This is likely to impact positively the existing terrestrial mobile radio systems, as it becomes more commonplace to remain in contact with the office, the residence, etc.

Mobilesat is a key ingredient in implementing the concept of the universal personal communicator, this telephone-like unit that will be used anywhere in the world, accessing networks (including mobilesat) on a least cost and optimal performance basis. The ubiquity of mobilesat terminals will shortly be the best promoter of that concept, which will benefit both land based and satellite systems.

New Terrestrial Markets Opened by Mobilesat

Mobilesat services will bring communications capabilities to areas that could not be reached before. In these areas, terrestrial services will permit a local extension of satellite services to numerous users needing to occasionally reach remote individuals out of their local coverage The following paragraphs outline a few examples.

Remote Mobilesat Telephone Systems. Like a cellular system, Mobilesat services will allow users to access directly the Public Switched Telephone Network (PSTN). Many remote users (small communities, fishing and hunting camps, mines installations, etc.) could share a fixed mobilesat link(s) by using terrestrial wireless technology. Fig. 1 suggests a possible mix of a few services for this application: VHF links, "CT-2" like telephones and a wireless PBX (CT-2 is the new generation of intelligent Cordless Telephones (CT)). A mini cellular switch could also be used locally to provide service to cellular telephones in the area.(Fig. 2).

These two mobilesat service extensions have the following advantages over the traditional wired solutions: the end users would get a completely *mobile* local and toll telephone service, within the local coverage area. All they need to carry is their wireless handset (or portable cellular phone).

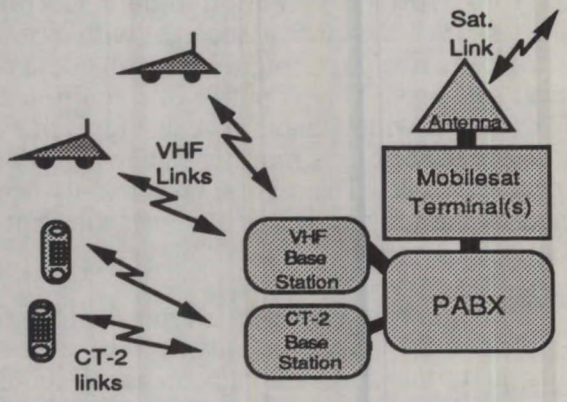


Fig. 1. Remote Mobile Telephone Configuration A

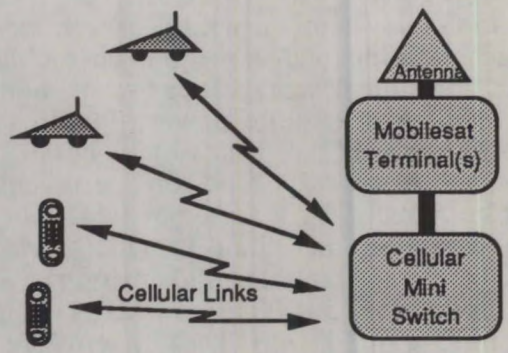


Fig. 2 Remote Mobile Telephone Configuration B

Furthermore, the portability of mobilesat terminals and their low cost (in comparison to other types of satellite communication) make them very well suited for temporary services (seasonal hunting and fishing, large construction projects, etc.)

Remote Mobilesat Radio Systems. Remote mobile radio systems can be interconnected to a remote dispatcher via a mobilesat link (Fig. 3), thus giving each vehicle a link with its head office, in addition to its existing link with other vehicles. Economics are likely to prove that it is less expensive to provide local coverage with a terrestrial system, rather than equipping each mobile with a mobilesat terminal.

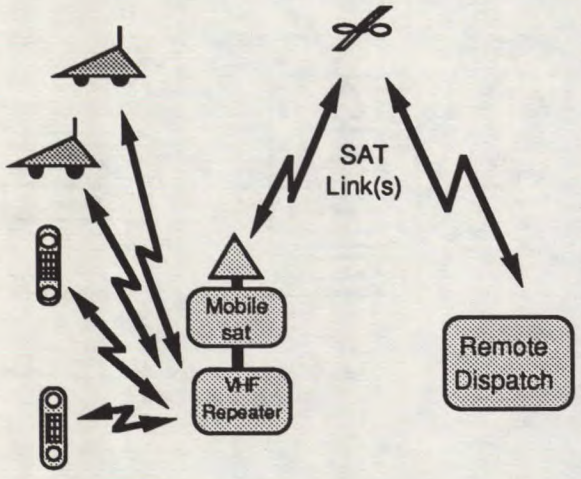


Fig. 3 Remote Dispatch using Mobilesat Terminal

The local range of mobilesat units can be extended by using vehicular repeaters, thereby allowing the mobile user to access his system when away from the vehicle.

Remote Mobilesat Data Systems. Data transmission is one of the most promising areas of mobilesat. The portable and inexpensive mobilesat terminal will start a new era in gathering data from countless locations and in controlling them remotely. Here again, a terrestrial mobile data system might prove a more economical alternative in linking locally various points to a mobilesat terminal (Fig. 4)

Using Mobilesat for Extending Existing Terrestrial Networks

So far, we have shown how terrestrial radio systems can extend remote mobilesat systems. But mobilesat terminals can also be used to extend terrestrial radio systems in areas where the latter cannot be implemented cost efficiently .

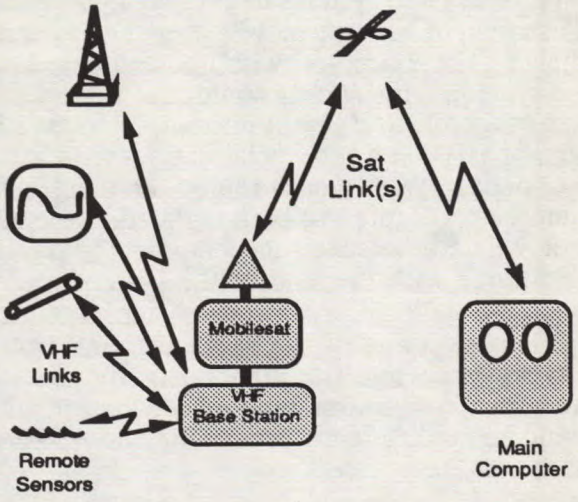


Fig. 4 Remote Sensing using Mobilesat Data Terminal

Extending Cellular Systems. Mobilesat Telephone Service will essentially provide the same capabilities as a cellular system. This service can be used to extend the coverage provided by actual cellular operators. The two systems will be technically incompatible, at least initially, and the competitive advantage of this extension will reside in the marketing strategy used to jointly market these services: one stop shopping, unique billing for both services, etc. Then, eventually, some mobile units will be capable of accessing both systems, and the respective network controller of each system will allow these units to hand off from satellite to terrestrial coverage, and vice-versa, on a least cost and optimal performance basis. From the user viewpoint, this will translate into a truly North American mobile telephone coverage, with the operational benefits of both systems: excellent communications in downtown and shadowed areas, at a lower price, in cellular areas; service everywhere else within satellite coverage.

Extending Land Based Mobile Systems. SMR's and private radio systems address the needs of the multiuser type of communica-

tions: taxi fleets, law enforcement agencies, public utilities, etc. Because mobilesat systems will have the capacity of emulating current terrestrial systems, they will be capable of extending SMR's and private radio systems in a way that is virtually transparent to both the mobile user and to the dispatcher. For example: with his land based mobile system, a dispatcher could reach a national fleet of mobile users, split into sub-fleets on a regional basis (on a provincial basis, for instance, in Canada). Within each region mobilesat units could be configured to be integrated to the sub-fleet, where it becomes uneconomical to provide the service with the landmobile system (in less densely populated areas). Both mobilesat and land mobile units of the same sub-fleet would then be accessed simultaneously by the dispatcher. From an operational viewpoint, there would be no difference between the units, except for the coverage areas.

Another interesting application is the combined use of both satellite and terrestrial mobile services to provide emergency communications. Terrestrial systems can be seriously affected by major disasters (fallen towers, damaged landlines, etc.). However, there are generally sufficient mobile units that survive the event to maintain good communications, if they had access to a repeater... For these eventualities, frequency agile repeaters, teamed with mobilesat links, could be mounted in emergency vehicles or in mini-shelters that can be carried by helicopter. A mobilesat link could be established between this repeater and the rest of the world, thus allowing the emergency rescue team to be coordinated by a central coordinator. (Fig. 5).

Extending Mobile Data Systems. Along the same line of thought, terrestrial mobile systems can be used to provide coverage where they can be cost efficient; mobilesat could provide coverage elsewhere. The mobile units would all be interconnected together at the application level, making possible transparent exchanges of data between any given mobile and/or fixed point in the system.

Eventually, some mobile units will be integrated at the system level: they will be able to access their destination using either land or satellite based systems, on a least cost and performance basis.

Mobile data services will boom within the next few years; we believe that terrestrial operators who will provide extended coverage by using a mobilesat system will definitely have a competitive advantage over other service providers.

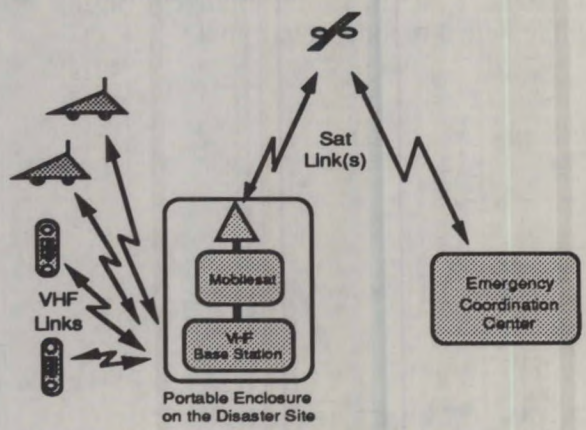


Fig. 5 Emergency Communications with an MSAT Link

CONCLUSION

As illustrated in the previous sections, we believe that mobilesat services are unlikely to displace a significant share of the existing land based mobile services. They will rather generate new market opportunities, either by making possible new applications, or by extending the capabilities of terrestrial systems on a continent wide basis.

Both mobilesat and land based radio systems must be coupled to realize the full level of global service that neither can achieve unilaterally. This needed cooperation is the first step towards the "Universal Communicator" described above.

REFERENCES

1. **Pemberton, J.** 1990. Mobile Satellite Services. *Gartner Group Inc* pp. 2-1 and 2-7.
2. **Pemberton, J.** 1990. Mobile Satellite Services. *Gartner Group Inc* pp. 2-1
3. Internal report provided to BCE Mobile by *Economic and Management Consultants Inc. (EMCI, Inc.)*
4. Cellular Industry Report, February 1990, *CTIA*, p. 13.

Canadian MSAT Field Trial Program User Requirements

Allister Pedersen
Manager, MSAT Trials Planning
Communications Research Centre
Department of Communications, Canada
P.O. Box 11490, Stn. H
Ottawa, Ontario, Canada K2H 8S2
Phone (613) 998-2011 Fax: (613) 998-5355

ABSTRACT

A wide range of mobile satellite service offerings will be available in late 1993 with the launch of Canada's first satellite devoted almost exclusively to mobile and transportable services. Early mobile satellite service offerings and field trials will be available through leased satellite capacity. During the last 7 years the Department of Communications has been meeting with potential MSAT (Mobile SATelllite) users in government and the private sector as part of a $20M Communications Trials Program. User trials will be conducted using leased capacity as well as capacity on Canada's MSAT satellite.

This paper will discuss user requirements identified under the Communications Trials Program. Land, marine, aeronautical and fixed applications will be described from the perspective of the end-users. Emphasis will be placed on field trials being accomplished using leased capacity such as the marine data trial being implemented by Ultimateast Data Communications, trials utilizing transportable briefcase terminals and additional field trials being considered for implementation with the TMI Mobile Data Service.

The pre-MSAT trials that will be conducted using leased capacity are only a limited sample of the overall end-user requirements that have been identified to date. The paper will conclude with a discussion of additional end-user applications and provide a summary of the user benefits of mobile satellite communications.

BACKGROUND

In support of its mandate to foster the development of improved telecommunications the Canadian Department of Communications (DOC) has held extensive meetings with mobile radio users to identify end-user requirements that are not currently satisfied by terrestrial systems. DOC initially took the lead in the promotion of mobile satellite (MSAT) services, conducting R&D and supporting market studies. Telesat Canada, and subsequently Telesat Mobile Inc. (TMI), eventually took over the lead for the implementation of MSAT when commercial viability was established and radio spectrum made available.

A very beneficial dialogue was established from the outset with end-users who provided substantial input critical to

establishing commercial viability and the need for government support in certain areas. This ongoing dialogue between DOC, Telesat and end-users over the years resulted in the definition and revision of various services to the current stage where TMI are well-positioned to respond to a very broad range of end-user requirements which are outlined below.

END-USER REQUIREMENTS

Existing mobile radio users in Canada identified many limitations to terrestrially-based mobile services. The following end-user requirements were developed in line with potential improvements that could be realized with satellite-based systems.

Geographical Coverage. Many Canadian-based end-users have identified a requirement to have mobile communications over the entire Canadian territory (including 200 miles offshore) and in many cases (e.g. truckers, railcar owners) all of North America. The current system concept for a regional North America-wide service [1,2] that provides for "roaming" anywhere in North America will meet the requirements of most Canadian users except those operating in the Northern Arctic. Coverage beyond 80 degrees North latitude is not possible with MSAT, a geostationary satellite. Polar-orbiting satellites, are required for these very high-latitude applications.

Equipment Configuration. End-users have identified a requirement for the following land-mobile configurations; trucks, cars, busses, trains and battery-powered units for individual railcars. Maritime mobile equipment is required for government vessels (Coast Guard, Fisheries), ferry fleets, fishing vessels and fixed and floating buoys (battery-powered). Aeronautical mobile applications range from large passenger aircraft requiring full cockpit and passenger cabin voice and data services to small single-engine bush planes with a requirement for "flight-following" and fleet management. As well as traditional mobile communications a significant MSAT market has been identified for a wide-range of SCADA (Supervisory Control and Data Acquisition) applications serving fixed platforms such as natural gas pumping stations, river flow gauges, avalanche monitors and lightning location detectors used in forest fire monitoring systems. A fifth broad category of equipment configuration identified by end-users was for transportable/portable, manpack and/or equipment described as "small hand-held radios providing wide-area communications coverage", the latter being along the lines of what is anticipated to be available as "personal communications" devices early in the 21st century. Man-pack communications are required, for example, by forest-fire fighters. Personal communications and geographic position-ing equipment (weighing less than a kilogram) is required by mineral claim inspectors and geologists travelling by aircraft, canoe and eventually on foot in the remote areas of Canada.

Telecommunication Services. The services required by end-users include voice, data, facsimile, and image transmission. The demand for voice ranges from half-duplex "field quality" (better than HF single sideband) for mobile radio applications to full duplex toll quality for mobile telephone. There is a major requirement for secure (encrypted) voice communications by government and private sector users such as those in the resource development sectors (oil and gas, mining, fishing).

Data transmission is required by a large group of end-users and this requirement ranges from once-per-day transmissions of a few bytes of information suitable for a packet switched service to full duplex circuit-switched data at 1200 bps and

higher for interconnection with corporate databases and electronic mail services. In the last few years there has been a much stronger interest in facsimile. The transmission of weather maps, updates to electronic marine charts and ice-flow information has also been proposed for MSAT.

In addition to 2-way voice and data communications, end-users have more recently responded very positively to various position location services such as LORAN C and GPS which, when coupled with 2-way messaging, provide an excellent overall fleet management service. The position location accuracies available from LORAN C (500 metres) meet the needs of most end-users such as long-haul truckers. Some end-users are looking forward to the much higher position accuracies available from GPS (Global Positioning System) [3] when it becomes fully available in the mid 1990's. The use of differential GPS with accuracies of several centimetres has been proposed for monitoring the relative movement of adjacent sections of pipelines in, for example, mountainous areas subject to slides. The indication of a significant movement of adjacent segments of pipe beyond a given threshold would be sent by an event-triggered MSAT transmission .

Network Services and Capabilities. End-users are looking for complete end-to-end telecommunication services from a wide range of value-added service providers with knowledge of their specific requirements. Opportunities exist for entrepreneurs to provide much more than the actual telecommunications links between various mobiles/fixed platforms and their operations centres. Fleet managers and other users such as pipeline companies require either complete systems consisting of hardware and fleet management software systems or a communications system that can interface with existing corporate systems.

Users have identified requirements for fleet management systems with capabilities of communicating simultaneously in a broadcast mode to all mobiles or user-determined subfleet. Operators monitoring fixed platforms require the capability to access information by polling all platforms, addressing individual platforms or through "event-triggered" activations. The actual approach to the collection, presentation and storage of information is more important to end-users than how the information was obtained (e.g. by satellite vs. terrestrial link).

Because the full range of voice services will not be available until the launch of Canada's MSAT in 1993 Canadian end-users will be forced to examine the early mobile data services. The requirement for end-users to consider something other than the standard type of mobile voice offerings will probably in retrospect appear as beneficial to some end-users as outlined below.

FIELD TRIAL APPLICATIONS-MDS

Following is a description of some of the DOC field trial program applications that are being implemented (or are under consideration) through the TMI Mobile Data Service.

School Bus Fleet Management

The end-user has a requirement to provide "safety" communications between schools and school busses travelling on very long lightly-travelled routes subject to severe weather and temperatures lower than -40 degrees Celsius. Bus drivers need the capability to contact the school (or a dispatcher) to advise of breakdowns or accidents. Although the original field trial proposal called for voice communications, particularly during emergencies, this is one application that may very effectively be handled by the TMI

RoadKIT service, until such time as the full range of voice and data services are available in 1993. RoadKIT could provide a school with 15 minute updates on the location of busses operating on remote routes automatically without the requirement for driver intervention or a "dispatcher" constantly monitoring incoming radio calls. The driver would have the opportunity to send short messages indicating the nature of any breakdowns or delays. A more extreme scenario might involve a bus sliding off a highway and overturning in a blizzard. Radio equipment in this situation, even if it is an existing terrestrial system, may be unusable. In this latter case, with RoadKIT, the dispatch software could raise an alarm when regularly scheduled position reports are not received from busses.

Marine Data Trial

Under an Agreement with Sea Link Ltd. and an associated contract with Ultimateast Data Communications Ltd. of St. John's Newfoundland the DOC is conducting a marine data trial involving the participation of 5 end-user organizations. The objective is to conduct a marine data service market trial that will provide fleet management services for 3 government agencies (Canadian Coast Guard, Department of Fisheries and Oceans [DFO], Marine Atlantic coastal ferries) and 2 fishing companies (Fishery Products International, National Sea).

Shipboard Equipment. The radio transceiver being used for the trials is manufactured by Canadian Astronautics Ltd. [4] For messaging services the shipboard end-user interface will consist of one of the following; PC (MS-DOS personal computer), laptop computer or Gandalf Systems Group Ltd. mobile data terminal.

Shore-based Equipment. A Sea Link/TMI Marine Network (MNET) will provide interconnection between the vessels, the Telesat Mobile hub, Sea Link, end-user PC's or mainframe computers and various other networks such as public electronic mail services.

Services. The primary services defined include the automatic transmission of LORAN C position reports (requested by all field trial participants) and various types of 2-way messaging. DFO and Coast Guard messaging will include electronic mail and situation reports for fisheries patrol and search and rescue vessels. Marine Atlantic messaging will consist of EMAIL, arrival/departure information and cargo manifest. Fishing company messaging will include a "pro forma" ground fish catch log, digital weather station reporting and data related to temperature directed fishing.

Oceanographic Trial

The Canadian land mass is bordered by 3 of the world's oceans which have a major impact on the Canadian climate and contribute to the economy through fishing and recreational pursuits. The Arctic and Atlantic Oceans are subject to large iceflows that create hazards for fishing, transportation and offshore oil platforms. DOC is developing a field trial with the Bedford Institute of Oceanography and Seimac Ltd. of Dartmouth Nova Scotia in support of satellite-based monitoring of tethered and floating ocean data collection buoys. This oceanographic trial would involve "real-time" monitoring of weather and sea conditions for ice-flow prediction purposes and be a pioneer trial to a wide range of potential applications including the monitoring of remote navigational aids as well as research buoys deployed for oceanographic research.

DOC Radio Inspector Trial

As part of its mandate to manage the

radio spectrum, the DOC has a network of offices across Canada and an associated fleet of specially equipped vehicles to investigate radio interference problems, conduct inspections and monitor the radio spectrum. A trial will be undertaken in Northern Ontario to improve communications between DOC offices and radio inspectors in their vehicles. The general features of TMI's Mobile Data Services will be evaluated in an effort to increase efficiency, provide safety communications to drivers, and provide better response to complaints of harmful radio interference especially where safety radio services are affected. This particular "fleet management" application is typical of many that will be undertaken with various federal and provincial agencies requiring wide-area mobile communications services.

SCADA Trials

In support of several federal and provincial government agencies as well as the private sector, the federal government is supporting trials related to SCADA (Supervisory, Control and Data Acquisition) applications. Support is being finalized for the development of a transportable terminal that will operate from a rechargeable battery which can be charged from solar cells or other means for a long period of time without maintenance.

Discussions are underway with Munro Engineering of Calgary Alberta regarding field trials for clients in the oil and gas industry for applications that can supported with the Canadian Astronautics terminal.

Railcar Location Service

A field trial to support the implementation of a railcar location service employing a modified Canadian Astronautics terminal is under discussion with a value-added service provider and rail tankcar users.

FIELD TRIAL APPLICATIONS-TRANSPORTABLE VOICE TERMINALS

In addition to the Mobile Data Services being implemented commercially in 1990 by TMI, Teleglobe Canada will be implementing commercial voice services for 2 Ontario Air Ambulances [5] and a limited number of SkyWave Electronics briefcase terminals. [6] Field trials using 10 briefcase terminals will be conducted with more than a dozen end-users including Emergency Preparedness Canada, Canadian Coast Guard, Fisheries and Oceans, CBC Radio and Television Networks, 4 provincial governments, CTV Television and INCO Gold Mgt. Inc.

Because the satellite being used for the voice trials is only about one-tenth as powerful as MSAT, the trials relate primarily to transportable applications that allow the use of fixed directional antennas with higher gains than available from current mobile antennas. An adaptive array antenna developed at the DOC Communications Research Centre [7] will be used for a land-mobile data trial and a maritime-mobile Canadian Coast Guard trial for voice and data applications.

BENEFITS

A highly-motivated group of future end-users of mobile satellite services has identified many potential applications and provided significant input in support of the definition of services to be offered on MSAT when it is launched in 1993, and other early-entry services provided with leased satellite capacity. Benefits to MSAT end-users resulting from the applications identified will include nation-wide coverage; increased efficiency, productivity, and customer satisfaction; increased reliability and quality of communication; confidentiality and improved safety.

REFERENCES

1. **Wachira, M.** 1990. Domestic Mobile Satellite Systems in North America. IMSC'90 Conference Proceedings Ottawa Canada

2. **Sward, D.** 1990. Mobile Data Services IMSC'90 Conference Proceedings Ottawa Canada

3. **Wells** et al. 1986 Guide to GPS Positioning Canadian GPS Associates Univ. of New Brunswick Graphic Services Fredericton, N.B.

4. **Sutherland, C.A.** 1990 A Satellite Data Terminal for Land Mobile Use IMSC '90 Conference Proceedings Ottawa Canada

5. **Butterworth, J.S.** 1988 Satellite Communications Experiment for the Ontario Air Ambulance Service Fourth International Conference on Satellite Systems for Mobile Communications and Navigation London England

6. **Rossiter, P. Reveler, D.** 1990 A Lightweight Mobile Voice/data Terminal IMSC '90 Conference Proceedings Ottawa Canada

7. **Milne. R.** An Adaptive Array Antenna for Mobile Satellite Communications IMSC '90 Conference Proceedings Ottawa Canada

Secure Voice for Mobile Satellite Applications

Arvydas Vaisnys, Jeff Berner
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91109
Phone: 818-354-6219
FAX: 818-393-4643

ABSTRACT

This paper describes the initial system studies being performed at JPL on secure voice for mobile satellite applications. Some options are examined for adapting existing STU III secure telephone equipment for use over a digital mobile satellite link, as well as for the evolution of a dedicated secure voice mobile earth terminal (MET). The work has included some laboratory and field testing of prototype equipment.

The work is part of an ongoing study at JPL for the National Communications System (NCS) on the use of mobile satellites for emergency communications. The purpose of the overall task is to identify and enable the technologies which will allow the NCS to utilize mobile satellite services for its National Security Emergency Preparedness (NSEP) communications needs. Various other government agencies will also contribute to a mobile satellite user base, and for some of these, secure communications will be an essential feature.

BACKGROUND

There is a trend toward digital implementation of voice in most mobile communications services. This includes the United States versions of Cellular Telephone, Mobile Radio, the land mobile satellite service (LMSS), as well as European developed systems such as CT2, and the various international mobile satellite services (MSS). While not necessarily the case for first generation systems, the data rate for future digital voice systems appears to be converging toward a 4.8 kbps compromise between quality and bandwidth requirements. A leading candidate for a voice encoding standard in the U.S. is the 4800 bps Code Excited Linear Predictive (CELP) voice coder defined in Proposed Federal Standard 1016.

Secure voice capability would be a useful addition to most mobile, small terminal, communications services in applications such as disaster communications and others where privacy may be important. Coincidentally, the mobile earth terminals (MET's) for all these services contain very similar functional blocks, which works in favor of a common secure voice interface. The options to be considered are whether the secure voice capability should be added as an external interface to the MET, or built into the MET.

MET FUNCTIONAL ELEMENTS

A generic MET contains the following elements: a voice codec, a coder/decoder for forward error correction (FEC), a modem, and a transceiver, as shown in block diagram form in Figure 1. In addition to these basic functions there is usually circuitry for call set-up, and an interface for external digital data, most likely at point A on the block diagram.

STU III FUNCTIONAL ELEMENTS

The STU-III (Secure Telephone Unit III) is a communications terminal unit that connects to the Public Switched Telephone Network (PSTN) and allows the user to make secure telephone calls.

In the non-secure mode of operation, a STU III operates as an ordinary telephone, using analog voice and all the signalling of plain old telephone service (POTS).

In the secure mode, the voice signal is digitized and compressed to a rate of 2.4 kbps using a LPC-10 algorithm. The compressed voice signal is then encrypted and is output on a subcarrier in V.26 format. During the secure call set-up there is also some signalling and echo canceler control using 2100 Hz tones, as well as digital hand shaking.

A functional block diagram of the present generation STU III is shown in Figure 2. It should be noted that the STU III has a key element of a MET, the voice codec, but that certain signal combinations, such as digitized clear voice and the encrypted voice data stream at baseband, are not available at the output.

The next generation STU III's, which will become available in late 1990, will include 4.8 kbps voice using the CELP voice coder defined in Proposed Federal Standard 1016. In the secure mode the output signal will be in V.32 format.

STU III INTERFACING

To interface the present generation STU III to a digital link, therefore, requires that its analog signals be digitized externally and the V.26 signal be restored to baseband by means of another V.26 modem. This can be accomplished by means of an adapter at both the mobile and the earth station ends of the link. Companies such as Electrospace, Inc., of Richardson, Texas, Stanford Telecommunications, Inc. of Santa Clara, California, and some of the STU III vendors have worked on the development of such an adapter.

Three options exist for interfacing the next generation STU equipment to a digital link. These options are for the MET end of the link. In all cases an adapter will still be needed at the satellite base station to provide the conversion between the digital data stream and the analog signals required by the PSTN.

One option is to develop a similar adapter, capable of operating at 4.8 kbps. Such an adapter would provide essentially identical functions at either end of the link.

The second is to develop a STU III version that provides the digital encrypted voice signal as well as digitized secure call set-up signals as direct outputs. This would provide a simpler interface to the MET. One question to be resolved is whether the STU III or MET voice codec would be used during clear voice operation.

The third option is to develop an integrated STU III MET. The secure call set-up and encryption functions could be provided as an optional module to a common MET.

THE LMSS CHANNEL

In clear line-of-sight conditions, the LMSS channel can be impaired by multipath signal components and vehicle motion induced Doppler. The most serious problem, however, is caused by signal blockage due to roadside obstacles. An example time plot of satellite signal strength data, measured in a fairly benign rural environment, is shown in Figure 3. The signal dropouts are caused by trees, telephone poles and other obstacles, and can last significant fractions of a second at normal driving speeds. Signal loss can easily be greater than 10 dB and providing link margin to overcome this phenomenon under all conditions is impractical.

OPERATION OVER THE LMSS LINK

Links over this type of channel are generally protected with some type of FEC coding, including time interleaving to break up the signal dropout induced error bursts prior to decoding.

The amount of time interleaving that is practical in a duplex voice link, however, cannot protect the link from most of these dropouts. It is therefore important that the system be able to resynchronize rapidly. In more difficult environments such as suburban and urban locations, operational problems such as complete loss of link may result. For STU III operation, the signal dropout

problem may result in frequently dropping out of the secure mode. Reestablishing the secure mode must be initiated manually and can become burdensome if needed often enough.

Link error protection and signal dropout mitigation are thus very important, especially while operating a secure link. This however can only be done to a limited extent. Certain environments will still be difficult to operate in, and this must be understood by potential users of the system.

ERROR PROTECTION

Error protection can be applied in layers. The MET FEC coder will apply an outer code to the composite data stream. The more critical signals, such as those involved in the call set-up process can be further encoded. Further protection of the critical bits of the voice codec output is also planned in some systems. For example, the proposed CELP codec employs an error correcting Hamming code within the voice frame.

In a non-secure voice MET there will be two basic phases of operation; the call set-up and the conversation phase. Call set-up may take place end to end over a dedicated request channel, or additional routing through the PSTN may be allowed by dialing over the assigned voice channel. Dial tones must of course be converted to digital codes, but they are generated so slowly that they can be reliably encoded.

In the case of a secure MET using a STU III, signal structure is even more complex because there are four additional signalling phases. These consist of a status phase, a crypto variable exchange phase, a crypto synchronization data exchange phase, and finally, the secure communications phase. These are described briefly below. More detailed information on the structure of these signals is contained in [1] and [2].

The status phase is used to set up the secure link. The secure call is initiated by one of the users pressing the "secure" button on the STU-III. That unit is now known as the initiator unit. The initiator sends a 2100 Hz tone (the 2100 Hz tone disables the echo suppression and echo cancelers on the telephone line). After some hand shaking, the initiator turns off the 2100 Hz tone. The two units then send signal patterns that are used to train the modem echo cancelers and the modem equalizers. This completes the first stage.

The crypto variable stage exchanges status and FIREFLY protected messages (FIREFLY is the encryption scheme). This is done to establish a cryptographic variable that is used by the STU-III's. The STU-III's perform several checks on the bits, including a parity check. If the data fails any of the checks, the secure call is terminated; these bits are not error correction coded. Next, the STU-III's exchange Random Component Cipher (RCC) messages. The RCC is FIREFLY encrypted and BCH encoded. Upon receipt, the RCC is BCH decoded and decrypted, and processed. If the message fails any of the processing parity checks, the call is terminated. Otherwise, the call proceeds to the next phase.

The STU-III's next exchange Crypto Synchronization (CS) messages. This occurs both at call set up and any time the call is interrupted. The CS is used to coordinate the selection of the secure mode, to align each STU-III's receive key generator to the other unit's key generator, and to coordinate the start of the secure transmission. Once the units are in sync, secure communications can begin. This is indicated by the STU-III "scrolling" messages about the CIK classification across the LCD display and then displaying the message "SECURE".

Once the link is secure, the user can end the call by either hanging up, which disconnects the link between the STU-III's, or he can push the "Non-Secure" button, which returns the units to non-secure communication. If the units drop out of sync, the CS messages are sent again.

As was mentioned previously, the outer coder will use time interleaving in combination with some type of error correcting code. The JPL MSAT project has chosen a trellis code for this application. Golay codes have been studied for this application in Land Mobile Radio [3]. The

type of coding and modulation to be specified by the American Mobile Satellite Corporation (AMSC) and other LMSS system providers has not yet been specified.

MSAT MET DESIGN

The MSAT MET was designed with a combination of coding and modulation which provides robust performance over a LMSS channel, yet minimizes bandwidth requirements. The FEC and modulation functions are implemented within the modem and use time interleaving and combined rate 2/3 trellis coded modulation (TCM)/eight phase differentially encoded phase shift keying (8DPSK). This allows a 4.8 kbps data stream to utilize a 5 kHz wide channel.

The time interleaving helps break up the error bursts caused by signal drop-outs. Another feature of the modem is the use of differentially coherent demodulation. This and a very stable clock results in instantaneous recovery after a signal dropout. The ability of the symbol clock to coast through long signal dropouts greatly diminishes the need for interleaver and voice codec resynchronization.

PROTOTYPE SYSTEM TESTING

A prototype system was tested in mid 1989 by JPL, using MSAT equipment and a developmental model Digital Transmission Interface (DTI) supplied by Electrospace, Inc.

The test set-up consisted of a pair of STU III's, DTI's, and MSAT MET's (with the voice codecs bypassed). The STU III output data rate was doubled to 4.8 kbps by bit stuffing in order to match the data rate of the MSAT equipment. No additional error protection, other than provided by the MSAT system, was implemented due to schedule constraints.

Testing was accomplished in two stages, first in the laboratory, then in the field in Australia using the Japanese experimental satellite ETS V. Field testing in Australia was done in conjunction with a scheduled MSAT-X field test.

Testing in the laboratory was accomplished using a satellite channel simulator which can be used to add Gaussian noise, produce signal fading by introducing multipath signals, and simulate other signal impairments. The actual measured satellite signal strength data of Figure 3 was also used to drive the simulator and ensure that the system would work in the field.

Minimum signal to noise ratio conditions for secure call set-up and secure call maintenance were investigated. The secure call set-up phase was more sensitive to channel impairments. Once a secure call was established, the signal to noise ratio could be lowered and the link was more tolerant of impairments such as signal dropouts.

In Australia, a secure link was set up and maintained over a satellite link which was operating with very little margin and was affected by multipath and blockage by roadside obstacles. Details of the testing are available in [4].

SUMMARY AND CONCLUSIONS

Several options exist for incorporating secure voice into LMSS MET's. Ideally, STU III functions should be incorporated directly into LMSS MET equipment. Special attention must be paid to error protection of the signalling that is part of secure call set-up and operation.

The LMSS channel is very difficult to quantify because its characteristics can vary so much from region to region. A great deal of propagation data has been gathered by both the JPL MSAT Project and NASA Propagation Program. It is available to anyone working in the area of mobile satellite MET design. Use of this data will allow the testing of proposed modulation and error protection techniques under realistic operating conditions.

ACKNOWLEDGEMENT

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. *FSVS-210, Revision E, FSVS Signaling Plan-Interoperable Modes*, 26 February 1988.

2. *FSVS-220, Revision B, FSVS Terminal Performance Specification*, 26 February 1988.

3. **Rahikka, D. J., Tremain, T. E., Welch, V. C., Campbell, Jr., J. P.,** CELP Coding for Land Mobile Radio Applications, *ICASSP'90*, Albuquerque, New Mexico, April 3-6, 1990.

4. **Berner, J., and Vaisnys, A.,** *Evaluation of Secure Communications Equipment Over a Mobile Satellite Link*, November 30, 1989, Internal JPL Report.

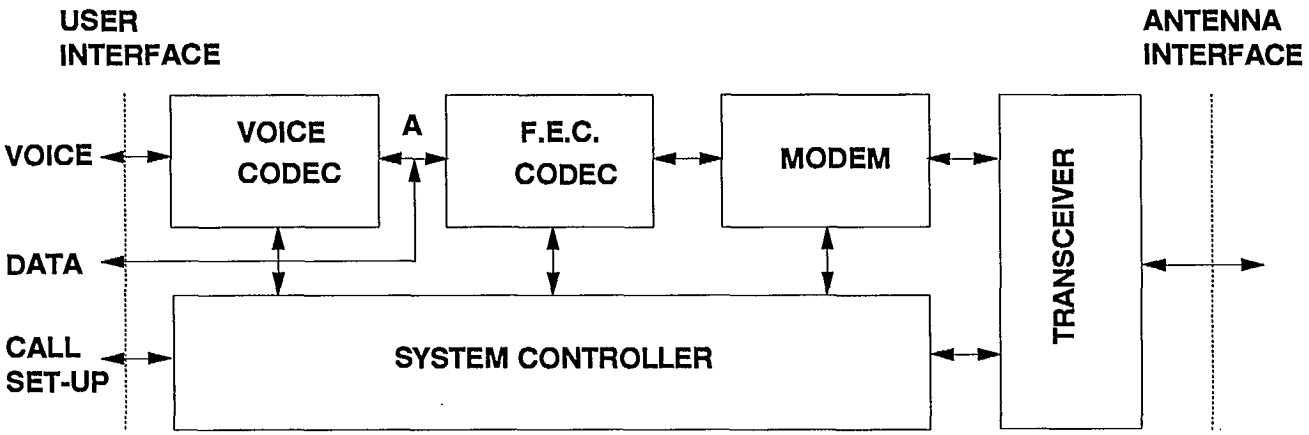


Figure 1. MET functional block diagram

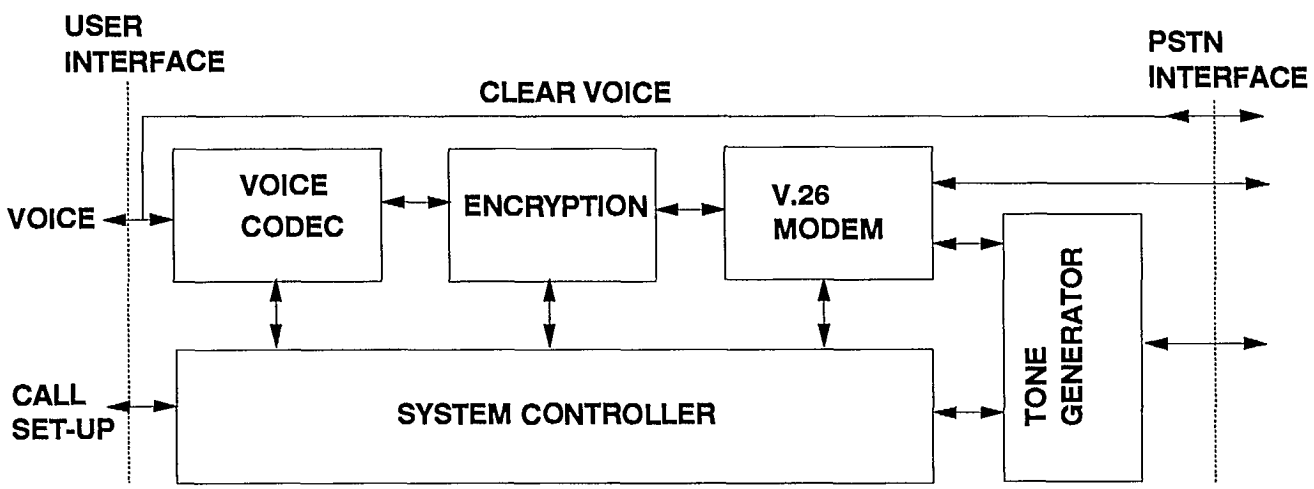


Figure 2. STU III functional block diagram

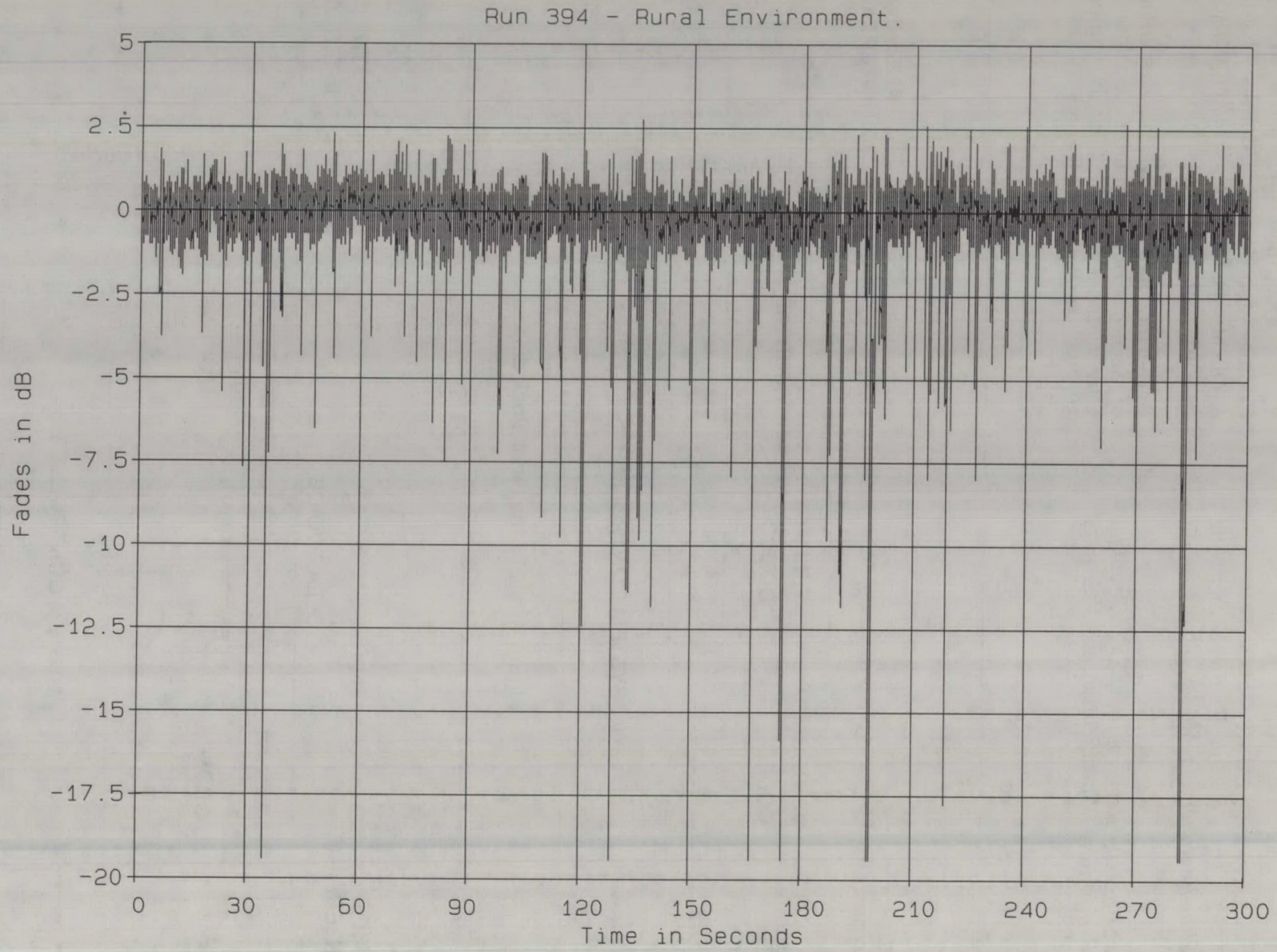


Figure 3. ETS V Satellite signal strength measurement.

USDA Forest Service Mobile Satellite Communications Applications

John R. Warren
USDA Forest Service
3905 Vista Ave.
Boise, ID 83705
208-389-2653

ABSTRACT

The U. S. Department of Agriculture (USDA) Forest Service (FS) manages 191 million acres of public-owned land in the 156 National Forests of the USA. Much of this land is in remote locations with poor or no commercial communications services. Normal communications from Ranger stations and other offices are typically handled by the telephone companies. For internal working communications a Forest will use land mobile or handheld radios with appropriate base and repeater stations. Handheld radios are also used. There are usually some places within a Forest that does not have adequate coverage due to line-of-sight limits or other reasons.

When wildfire or other emergencies occur, independent radio communications systems are set up on separate frequencies so that the emergency communications will not interfere with daily operations in the area. Communications outside the immediate area of the emergency are not always adequate or available because of the remote locations where the command post may be set up.

The primary FS infrared (IR) line scanning systems presently rely on physical delivery of IR images on film for subsequent geo-location of the fire information onto maps. The FS is procuring two airborne IR line scanning systems (called Firefly) which will do digital signal processing on-board the aircraft to locate, by latitude and longitude, the fire and hot areas. This information may then be transmitted over a voice-grade radio-frequency (RF) link to the Incident Command Post (ICP). To avoid the need for the aircraft to establish line-of-sight with the ICP, and maintain it until completion of the transmission, the mobile communications satellite system will be used. This will permit the aircraft to transmit the pertinent location data while en-route to the next fire, saving valuable flight and pilot time.

Other potential uses for the mobile communications satellite may include search and rescue (SAR), automatic in-flight locating and reporting, dispatch orders, and en-route resource and situation status reporting.

INTRODUCTION

The USDA Forest Service (FS) and the Dept. of Interior spend an average of over $200,000,000 annually in the suppression of wildland fires. In extreme years the costs go up. In 1988, the suppression costs in the Greater Yellowstone Area alone was around $120,000,000. The number of personnel on large fires may vary from 300 to over 2000. Again, in the 1988 Yellowstone fires, more than 25,000 people worked there with a peak of 9500 and 117 aircraft. With the hazards and logistical complexities associated with

large wildland fires, the communications systems become very important. Fire Camps and ICP's are often located in remote areas. It is necessary to provide eating, sleeping, and sanitation arrangements, as well as transportation of people, equipment, and supplies often during rapidly changing situations.

Communications regarding the availability of resources (people, aircraft, ground vehicles, tools), and the situation status (fire location, behavior, predicted spread rate) are essential for effective and efficient suppression actions. The FS has used thermal infrared (3-5 and 8-12 micrometer bands) in airborne IR line scanning systems for about 25 years for fire management efforts. The primary problems related to this use are the availability, when needed, and the timeliness of interpreted, fire locations onto maps. Mobile satellite communications coupled with on-board digital signal processing, advanced navigation systems, and stored digital terrain elevation data are expected to improve the efficiency of the IR systems over the next 5 years.

INFRARED INTELLIGENCE

Areas with large fires often are completely "smoked in". That is, there is a smoke pall over a large area which reduces the visibility from aircraft to ground to near zero. Thermal infrared systems can detect hot areas like fire perimeters and fire spots through the smoke in most cases. The IR signals can be processed on board, resulting in a film display of terrain features and hot areas. The film must then be physically delivered, usually from the nearest landing strip, and a manual interpretation performed, to locate the fire on the maps of the area. The nature of the line scanners operating over uneven, rugged terrain results in a horizontal scale that is constantly changing from line to line or even within a single line scan.

The FS has successfully transmitted both IR video, and standard RS-170 video with IR and terrain features from the aircraft to the ground since 1974 (1). In 1983, the feasibility of further geo-referencing IR information on board, and transmitting same via satellites was studied (2). The methods established in the feasibility study are very similar to what is now being developed for the FS Firefly system, which is scheduled for operational use in 1993. The standard IR line scanner signals will be digitized and processed along with Global Positioning System (GPS) data, on-board gyro attitude sensors, and stored digital elevation data to provided a latitude-longitude location tag for all parts of an area exceeding a selected threshold value. This will permit the transmission of the fire location data over a voice grade link. Previous methods have successfully transmitted the IR video or images, but required wide RF bandwidths to reduce the transmission time down to a reasonable number. Also, the transmission and reception equipment were considerably more complex and expensive than what is required for voice-grade or lower data rate transmissions.

The new methods will permit the IR aircraft to "map" a fire and proceed on to the next fire area without the need for first flying to the landing strip nearest to the ICP, landing, delivering the film strips, and then taking off for the next fire. This will improve the amount of time spent collecting fire data to the amount of time spent delivering the data. In fact, there will be essentially no time lost in the delivery, because the aircraft will be on its way to the next destination while that is taking place via the mobile communications satellite.

Location accuracy and relative location of the fire perimeter or spots to firebreaks may not be as good as what can be found from the film. But

advances in forward looking infrared (FLIR), coupled with GPS or Loran C, can augment the large area, strategic information provided by the line scanners. The FLIR based units could also be equipped with mobile satellite communications capability which would also speed up their delivery of tactical IR information to the Fire Staff.

OTHER FIRE COMMUNICATIONS

There are other communications needs associated with the management of large fires or complexes of fires. The need for "routine" communications associated with the management of large numbers of people and suppression resources was mentioned. Fires are often in remote areas, and usually there is little or no time for a normal build-up of communications systems. Whatever is used must be highly portable and be able to talk to or network with commercial telephone systems at some place. Stringing telephone lines has been done many times in the past, but a totally mobile wireless method is much preferred. Besides some locations are too remote or inaccessible to consider the older common methods. Lines and in some cases even microwave or radio repeater tower sites may also be in jeopardy of the fire.

With numerous hand crews at diverse locations around the fire line, retardant aircraft, dozers, and engines deployed, it is likely that automatic location systems based on satellite communications may be used in the future. The locations of all units could be displayed on a color monitor and messages processed to show progress or emergencies. As wind or other conditions change, the changes and latest fire spread predictions would be available to all the crews and others out on the line in a matter of minutes. Satellite communications would let messages be sent to only the selected units or to all units simultaneously.

Another advantage to mobile satellite communications would be the "instant" set up time, compared to other methods, including other satellite systems. The fire staff could be on-line while in their vehicles en route to the fire scene.

FIRE OPERATIONS

Wildfire is quite unlike some other natural disasters such as earthquakes, tornadoes, hurricanes, or even floods. There is usually very little that can be done to control the other events. Steps can be taken to minimize the damage and loss of life that may be incurred with those, but once they start they are pretty much out of control. Wildfire is sometimes uncontrollable, too, when driven by high winds, but often it can be controlled or managed to a large extent. Such management necessarily requires a large amount of timely and pertinent data, as well as good judgment, and prediction models. Plans can be made, but are subject to the availability of the various types of resources that are available. The uses of mobile satellite communications systems offer ways to provide better data for better management resulting in lower suppression costs and natural resource losses.

CONCLUSION

The airborne IR signal processing system being developed will require the use of mobile satellite communications to achieve its full capability and improvement in delivery timeliness of processed IR data to the Fire Staff. There are numerous other beneficial uses, both during wildland fire management operations or in daily routine tasks, which will also benefit from the availability of reliable communications from remote areas.

REFERENCES

1. Wilson, Ralph A. and John Warren, "Airborne Infrared Forest Fire Surveillance--A Chronology of USDA Forest Service Research and

Development", General Technical Report INT-115, 1981.

2. McLeod, Ronald G., T.Z. Martin, and John Warren, "A Feasibility Study: Forest Fire Advanced System Technology (FFAST)", JPL Publication 83-57, 1983.

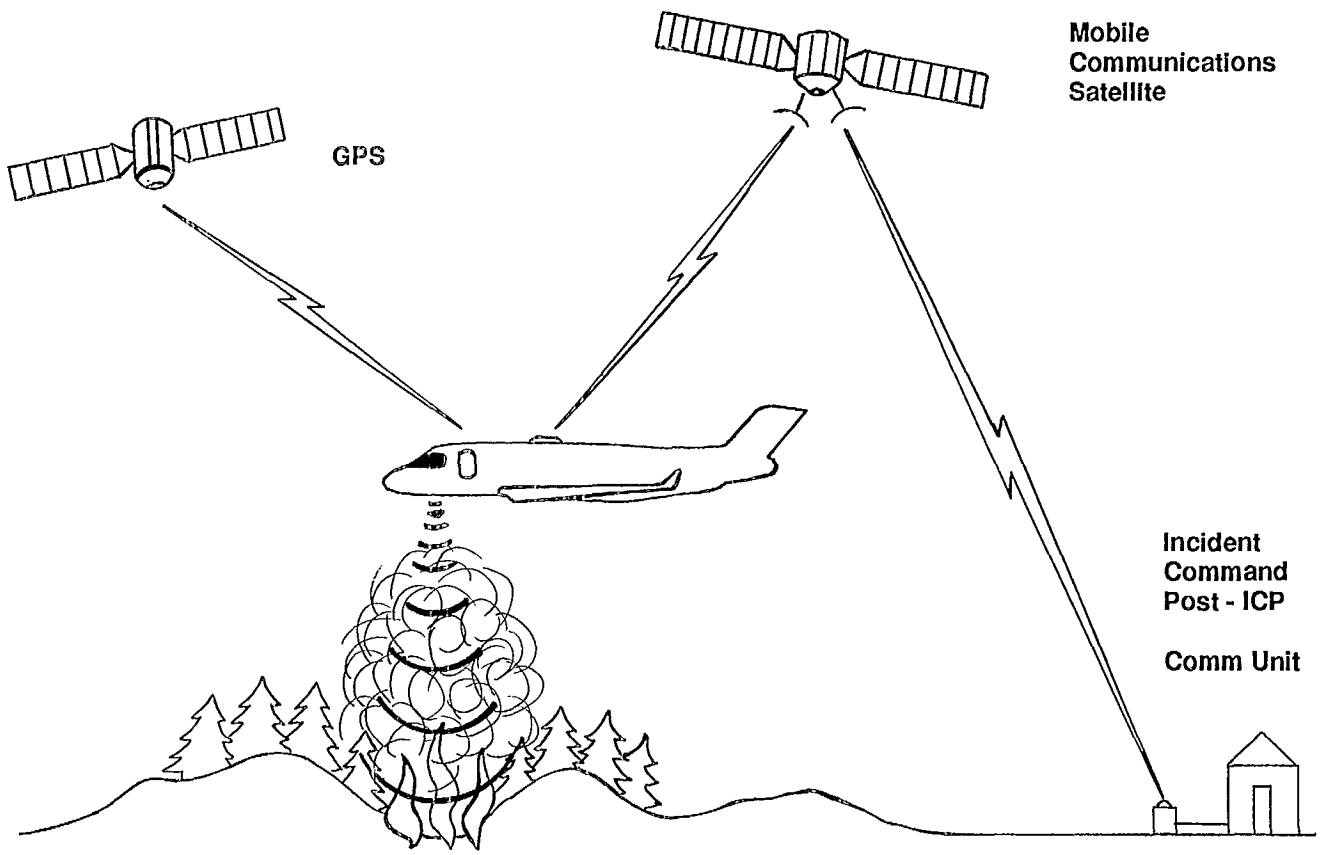


Fig 1 Firefly Concept

Mobile Satellite Services for Public Safety, Disaster Mitigation and Disaster Medicine

Jerry Freibaum

Formerly: Chief, Policy and Regulatory Studies, NASA Communications Division
Telecommunications Consultant
6535 Elgin Lane
Bethesda, Maryland 20817
Phone: 301-320-5550
FAX: 301-229-9694

ABSTRACT

Between 1967 and 1987 nearly three million lives were lost and property damage of $25-$100 billion resulted from natural disasters that adversely affected more than 829 million people. The social and economic impacts have been staggering and are expected to grow more serious as a result of changing demographic factors.

Recognizing the global consequences of these events the United Nations has designated the 1990's as "The International Decade for Natural Disaster Reduction." The major role that Mobile Satellite Services (MSS) can play in the "International Decade" is discussed .

MSS was not available for disaster relief operations during the recent Loma Prieta/San Francisco earthquake. However, the results of a review of the performance of seven other communication services with respect to public sector operations during and shortly after the earthquake are described. The services surveyed were: public and private telephone, mobile radio telephone, non-cellular mobile radio, broadcast media, CB radio, ham radio and government and non-government satellite systems.

The application of MSS to disaster medicine, particularly with respect to the Armenian earthquake is also discussed.

INTRODUCTION

The stakes are high. The social and economic impacts of natural disasters have been and will continue to be enormous. Disasters respect no boundaries or economic classes. The distance from the epicenter of an earthquake is not necessarily an indicator of the full damage potential. A destructive earthquake in Tokyo, Japan can, for example, be economically devastating to the United States(1). A big quake has hit Tokyo roughly every 70 years for four centuries: 1633, 1703, 1782, 1853, 1923, and ---- 1993?---.

The world wide seriousness of the problem is the reason the U.N. established the "International Decade for Natural Disaster Reduction" beginning in 1990.(2,3) Its goal is to promote cooperative efforts between nations to reduce the ravages of all types of natural hazards. Individual countries, including the United States, have declared their own "National Decades" consistent with the international program.

Satellite communications and the application of remote sensing technology can play major roles in reducing the devastating effects of natural disasters by dramatically improving hazard and risk assessment, disaster preparedness, early warning and onset and post disaster relief operations.(4,5)

For most emergency response users, the cost of satellite terminals and service charges have been too expensive. Now, new advances in communications technologies, the integration of satellite and terrestrial services, mass production of user terminals and market forces should dramatically reduce costs and alleviate the problems of affordability, universal access and coverage.

DISASTER MITIGATION - LOMA PRIETA / SAN FRANCISCO EARTHQUAKE

Background

Loma Prieta was the epicenter of the earthquake that affected the Santa Cruz/San Francisco corridor on October 17, 1989. It was considered by many to

be only a "rehearsal" for the "real thing".

Within moments of the major tremor, communications difficulties common to most major, rapid onset, disasters were experienced during the first few hours. The application of satellite communications for disaster management during the earthquake was conspicuous by its absence.

A review was conducted by PSSC, with a grant from AMSC, to provide information on the performance of seven different communications services used by public service and emergency response organizations during and shortly after the earthquake(6). These organizations included: health care facilities, fire departments, public utilities and transportation, police, the Red Cross and public schools.

Emphasis was placed on identifying communications failure modes and vulnerabilities. It should be noted that, in general, communications for most critical situations was adequate except for the first few hours.

The quake effects were unusual in that structural damage to communications facilities was relatively light. Had the earthquake occurred a few hours earlier, lasted about ten seconds longer and occurred several miles further north, casualties and damage to structures, including communications facilities, would have been dramatically higher.

Sampling of Findings, Observations and Conclusions(6)

Telephone outages. The adverse effects of the breakdown of the public telephone system during the first few hours was a dominant theme. Outages were primarily due to overloaded circuits and lasted anywhere from several hours to two weeks. Dial tones were affected by power outages and overloads.

Receivers were knocked off their mounts exacerbating the problem.

Long distance calls into the area were particularly difficult. "Essential service", which usually exempts emergency-services from local telephone company "line-load control", did not get priority service on long lines. 911 service was inoperative in many areas and, in some cases, for several days.

Structural damage to communications facilities was relatively minor and network infrastructures appeared to remain intact.

Cellular mobile radio telephone. Cellular companies loaned 2000 portables to many public safety organizations. Most who used the units and service appeared pleased with the results.

Many organizations indicated that the purchase of cellular radios would be included in their future budgets.

Cellular phones became the "work horse" in the field. They were also used extensively to solve "... the interoperability problem of incompatible radio frequencies among cooperating agencies...."

Despite general enthusiasm for the cellular phone there was concern by most that cellular towers were vulnerable to structural damage and that the network was also vulnerable to overloading. The cell site at the Emergency Operations Center did, in fact, overload regularly. A few sites were damaged and several experienced overload outages. Portable sites were brought in quickly where needed.

Facsimile/Data Links. The use of facsimile "exploded" whenever the phone lines worked. This tied up valuable voice circuits for lengthy transmissions, sometimes as many as 10 to 15 pages. Interest was expressed in using some sort of burst mode transmission to reduce the length of transmission time.

There is a need for expanded use of data links and information networking for linking personal computers, FAX machines, information management systems, etc..

Power supplies. Power supply problems were common and generally serious. Many could have been avoided.

Mobile to mobile. This capability was considered very important. In several cases, where base stations were not functioning, vehicles were used as base stations. Coordination within and between organizations (eg. fire/police) was carried out via mobile to mobile operations.

Smoke Detectors. The inability to reach San Francisco by phone created the potential for major fire damage and loss of life since some and possibly

many of San Francisco's smoke detectors are monitored in locations outside of San Francisco including Chicago. The location of an alarm in San Francisco that is monitored in Chicago would normally be phoned from Chicago to the nearest fire station in San Francisco. This problem is under study by the Fire Department.

Remote monitoring of burglar alarms may present a similar problem. However, this was not verified.

Vulnerability of public switched networks. Public switched networks appear to have emerged relatively intact and in good shape. However, they are growing more vulnerable to outages, according to a controversial conclusion of the National Research Council(7).

Special data, position location and remote sensing needs. Chemical spills were observed in some school laboratories. This danger could have some very serious data and remote sensing implications with respect to hazardous materials transfer and storage. Given a more serious quake, problems of fractured storage tanks, leaky gas lines and damaged transporters all become problems of immediate detection, location, assessment and emergency response.

Highway status (damage and traffic information) becomes critical for effective emergency response measures. This information would also be critical for managing the transport of medical, water, food and other relief supplies.

Transportation. Only the BART System was reviewed. However a complete review of all public transportation operations would be useful since the transport of search and rescue teams, relief supplies, heavy equipment, evacuees, etc. is almost solely dependent on their functioning.

Additional concerns are tunnel communications, prisoner transport, law and drug enforcement operations. Encrypted or scrambled communications are needed for these functions.

Accurate and Timely Information. The inability to obtain and disseminate useful,accurate, real time information, that is not out of context, on casualties, damage, and resources needed was identified as a major problem by most of the respondents.

Human factors. Attitudes, habits, skills and resistance to change of field and supervisory personnel make it very difficult to introduce new communications systems or devices during a crisis. Therefore, any new equipment, satellite or terrestrial, must be simple and look like, sound like, and feel like a telephone or existing mobile radio.

Back-up communications planning. The public telephone and mobile radio are, by far, the most important modes of communications for emergency response. Steps are being taken to "harden" these systems to minimize disaster damage. Back-up plans seem to be emphasizing cellular and caches of mobile radios both cellular and non-cellular.
Other back-up modes being considered are: HF radio, satellite communications, redundancy and alternate routing.

Since terrestrial back-up is being emphasized major outages can be expected for everyone, given the occurrence of a quake larger than Loma Prieta.

Satellite role. There is a general, almost universal, awareness that satellite systems can play a major role in disaster management. It is firmly believed that, ultimately, it may be the only means of retaining or immediately reconstituting communications during a major disaster.

This awareness, however, does not seem to be translated into implementation plans at the local level. The State level appears to be exploring the concept of integrating satellite technology into their communications networks but probably still has a long way to go before an operational system is realized. Planning for satellite systems integration seems to be furthest developed at the Federal level.

The California earthquake was a typical example of a situation in which an integrated satellite/terrestrial capability would have been invaluable. It could have linked the State with local communities, provided coverage for emergency medical services, provided long distance service for priority traffic, off-loaded non-emergency traffic from public safety channels and provided the much needed additional communications capacity all of the local communications systems critically needed. Vital communications could have been reconstituted within minutes instead of hours or days

Other areas of special interest not covered in this review. The roles and performance of marine communications in the Bay area, particularly if a quake of much larger magnitude were to hit the Bay

area;. the failure potential of air traffic control communications in or near the airports; and the use and effectiveness of Tsunami or Tidal Wave warning systems for California's coastal areas.

DISASTER MEDICINE - SPACE BRIDGE - ARMENIAN EARTHQUAKE

Background

The Armenian earthquake, which occurred in early December 1989, caused more than 150,000 casualties as well as enormous, wide spread, destruction. The economic impact to the Soviet Union has been estimated at 20 to 40 billion dollars.

Under the auspices of the US/USSR Joint Working Group on Space Biology and Medicine, NASA's Communications and Life Sciences Divisions initiated, funded and implemented a "Telemedicine Spacebridge" which provided satellite video, voice, FAX and telex links between Soviet and US medical teams for treatment of Armenian earthquake victims(8,9).

Initial satellite communications links consisted of one-way video, two-way audio, data, fax and telex. It was the product of a cooperative effort on the part of the COMSAT Corporation, INTELSAT, AT&T, Satellite Transmission and Reception Specialists and NASA's Goddard Space Flight Center in the US, and the Soviet Ministries of Post and Telecommunications and Health.

The US physicians and specialists that participated in the spacebridge project provided consultative support to more than 200 Soviet physicians, primarily in the areas of reconstructive surgery, rehabilitation and psychiatric care for post-traumatic stress disorder. The four month project was considered highly successful by the participants and they recommended that the concept be continued and expanded.

Participating US centers were: (a) The Uniformed Services University of the Health Sciences in Bethesda, MD, (b) The University of Maryland Institute of Emergency Medical Services Systems,Baltimore, MD, (c) The University of Texas Health Science Center, Houston, Texas, and (d) LDS Hospital and the University of Utah, Salt Lake City, UT. The Public Service Satellite Consortium (PSSC) provided telecommunications applications support to NASA and the USSR Ministry of Health.

In June, the Soviet Ministry of Health urgently requested technical assistance from NASA and PSSC for Soviet medical experts treating victims of a gas explosion near the city of Ufa in the Ural region. Twelve hundred casualties resulted when sparks from a passenger train ignited gas escaping from a ruptured line nearby.

NASA and PSSC responded in real time by extending the Space Bridge to include black-and-white slow scan video and voice communications between Ufa and Yerevan, Armenia for retransmission through the space bridge to the US Hospitals which then provided consultation to burn specialists in Ufa. The slow scan equipment was donated by Colorado Video, Inc. in CO.

It is significant to note that the application of slow scan and audio technology in Ufa was constrained to using only existing voice circuits and infrastructure. This permitted the rapid deployment of low-cost interactive equipment to a region whose existing communications network could not accommodate full-motion video. MSS can provide this kind of low cost, real time deployment and capability.

Several participants, including PSSC, are involved in proposals to continue and expand the spacebridge concept on a worldwide basis.

A CHALLENGE

More than 20 years of development, user experiments, regulatory proceedings, political processes and countless studies have been completed. This "new" technology will finally be integrated into useful economically viable systems worldwide serving the public and private sectors.

Will public sector and government institutions and funding sources recognize this opportunity? Will they take the actions needed to enable them to apply this technology in ways described in this and other papers presented at this and other related conferences?(10,11)

Will service and hardware providers facilitate the use of the Mobile Satellite Service by the public sector? Will the providers remember that the U.S. Government and public sector elements facilitated

the commercialization of the MSS and that some reciprocity might be of benefit to all.

This is an era of increasing world-wide social, economic, technological and political inter-dependencies. "... The Decade is an opportunity for the world community, in a spirit of global cooperation, to use the considerable existing scientific and technical knowledge to alleviate human suffering and enhance economic security."(3)

REFERENCES

1. Lewis, M.. *How a Tokyo Earthquake Could Devastate Wall Street.* June 1989. Manhattan,inc. pages 69-79

2. National Research Council. *Confronting Natural Disasters: An International Decade for Natural Hazard Reduction.* 1987. National Academy Press. Washington, D.C.

3. Kreimer, A. and M. Zador (eds.). *Colloquium on Disasters, Sustainability and Development: A Look to the 1990s,* Environmental Paper No. 23. December 1989 (Washington, DC: The World Bank Policy Planning and Research Staff). Note: This is an internal document.

4. Scott, J. C. and J. Freibaum. (PSSC) *Mobile Satellite and Remote Sensing Technologies: Tools For Disaster Mitigation in the Next Decade.* Published in BOSTID Developments, National Research Council. Fall 1988

5. Public Service Satellite Consortium. Testimony on: *The Application of New Space Technology to Reduce the Adverse Effects of Natural Disasters and Dramatically Improve Disaster Medicine and Rural Health Care.* before the Senate Subcommittee on Science, Technology and Space. October 19, 1989.

6. Freibaum, J., *Review of the Effectiveness of Communications During and Shortly After The Loma Prieta, California Earthquake.* February 1990. Public Service Satellite Consortium through a grant from the American Mobile Satellite Corp.. Washington, D.C.

7. National Research Council. *Growing Vulnerability of the Public Switched Networks: Implications for National Security Emergency Preparedness.* 1989 National Academy Press, Washington D.C.

8. U.S. - USSR Joint Working Group on Space Biology and Medicine. *U.S - U.S.S.R. Telemedicine Consultation Spacebridge to Armenia and UFA.* Presented at Third Joint Working Group Meeting December 1-9,1989. Washington, D.C.

9. Scott, J.C., Freibaum, J.(PSSC). *Facilitating The Growth of International Emergency Medical Assistance Via Satellite.* January 14-17, 1990. Proceedings of the 1990 Pacific Telecommunications Council. Honolulu, Hawaii.

10. Office of the United Nations Disaster Relief(UNDRO) (assisted by the Public Service Satellite Consortium). *International Conference on Disaster Communications.* April 1990. Geneva, Switzerland.

11. Public Service Satellite Consortium (PSSC) (organizer). *Symposium on Disaster Communications.* April, 1989. Sponsored by American Mobile Satellite Corp., COMSAT, GEOSTAR, INMARSAT, McCaw Cellular Communications, NASA, and QUALCOMM. Ontario, California.

European User Trial of Paging by Satellite

R E Fudge and C J Fenton
British Telecom,
Mobile Communications, Mobile House,
Euston, LONDON NW1 2DW United Kingdom
Phone +44-71-388-4222

1. INTRODUCTION

Satellite communication to land mobiles is still in its infancy due to the combined limitations of economic satellite erp and the abilitiy to accommodate suitable antennas on vehicles. The latter factor is a particular constraint for two-way transmission, since small antennas not only give a poor link budget but also cause 'pollution' of the geostationary orbit. If however, only one-way transmission is considered then it becomes obvious that currently available satellite performance and the use of an acceptably sized antenna would be able to support a low speed data service.

Ideally such a data service would have its modulation and coding constructed in a way that made optimum use for the satellite channel characteristics. If such a modulation/coding scheme were to be implemented then users would have to bear the cost of production and development of such specialised terminals.

British Telecom conceived the idea of adapting their existing paging service, together with the use of existing terrestrial pagers, to yield a one-way data (ie paging) satellite service to mobiles. This was recognised as not providing the absolute best use of the channel, but the degredation was small and economically tolerable.(1)

A series of design studies, and, technical evaluations were carried out which lead to the conduct of an user trial and this is reported on here.

2. SYSTEM CONFIGURATION

The system design was based on the maximum use of the existing British Telecom terrestrial paging system and the use of normal terrestrial paging terminals.

The present paging network consists of 450 base sites. These radiate in the VHF region (153 MHz) and give an extensive coverage in the UK, reaching 98% of the population. A variety of services are offered from simple tone page, through numeric page, to full alpha-numeric message transmission of 90 characters. Standard pagers are used based on POCSAG protocols (CCIR Radiopaging Code No 1). Originally, these operated at 512 bits/sec and more recently at 1200 bit/sec. Paging users elect to take service from a number of geographic zones. An initiating call can be made either directly to the control system (PACE) in the case of tone pagers, or to a bureau for numeric or alpha-numeric paging.

The satellite service used the same protocols and modulation. The paging network was configured with an extra zone, termed 'satellite zone', and the network signals from this were fed to the Goonhilly Earth Station. Here they were modulated in the normal way on to a terrestrial paging carrier and this was up-converted to the satellite uplink frequency. The satellite down link at L-band

was received by the tubular quadrifiliar helix mobile antenna (omni directional) and then down converted to the standard 153 MHz paging frequency. This signal was inductively coupled into a normal 'message master' pager having a 90 character display. The network configuration is shown in figure 1.

The pager and its coupling were mounted in the cab of a vehicle, remote from the down converter if desired, and the housing of the pager also contained a printer which derived its information from the pager and provided a hard copy of received messages.

This configuration had a number of user advantages. The pager owner could use his pager in the terrestrial mode when desired, eg by clipping to his belt or carrying in pocket or handbag and in this mode he would obtain the normal paging service in the UK with its extensive coverage in urban areas, into buildings and into vehicles. When in his vehicle outside the normal coverage areas he could place the pager in the vehicle adapter. This then gave him coverage of all other areas of UK, and moreover allowed the vehicle to roam freely over the whole of Europe. In addition the call originator could use the normal mode of message initiation; either by speaking to the message bureau or accessing the paging network via the public data networks. The paging number remains the same whether in terrestrial or satellite mode and the caller does not need to know the location of the user.

3. OUTLINE OF TECHNICAL RESULTS

The system concept was first evaluated by way of technical trials in order to have an understanding of the nature of the channel and of the service which might be provided. This was a necessary first step prior to the user trials, and served to both prove the system and ensure that potential users would have the benefits of our prior experience.

These technical trials have been reported on in detail.(2) In summary they showed that full message fidelity was always achieved when stationary in open areas. Motion in itself had no visible effect, but the shadowing caused by trees, over-bridges and dense urban areas was noticable in generating message errors. However these were less than might have been expected:- in urban areas 65% of messages were received with no errors and 90% were received with some corruption. No variations due to satellite elevation were detected on runs conducted in the UK.

The overall system performed well and it was considered that a viable user trial could be mounted; the service would be angled towards long-distance truckers who are most likely to be on open freeways rather than in urban areas.

4. CONDUCT OF USER TRIALS

The trial had the objectives of learning the marketing requirements, and of assessing the ability of the system to provide a service. It also provided potential customers with an introduction to the service and the engineers with an assessment of installation problems. Fig 2 shows an installation with the antanna mounted on the back of the truck cab.

To achieve these objectives users were selected who operated widely a) in the UK, b) within another European country, c) across national boundaries including the use of ferries. Fig 3 shows a map of the area of operation. Reports were obtained from the fleet operators, the fleet controllers, and of course from the drivers. Table 1 shows the parameters of the trial.

5. RESULTS OF TRIAL

No significant problems were encountered during vehicle fitting, this was carried out in depots in the UK and in The Netherlands. Figure 3 shows a rear view of the cab of an articulated unit with antenna showing above the roof.

Overall the system performed well over the 4 months with no network, or satellite segment, problems. The vehicles were in extensive use for at least 5 days per week. It was judged that some 60% of messages were received correctly, and this should be reviewed against a background of some resistance to the use of the system by drivers. They had some feeling that the installation of the system in 'their' cab called into question their competence and ability to judge situations for themselves. It was very much welcomed by the car user (a salesman, for the products transported by the trucks), by the fleet controllers, and by the fleet managers.

There was a desire to have acknowledgement of correct receipt of messages, especially if these contained safety information - this use was seen as a great benefit. Particular problems were encountered with customs officers at borders who regard the installation as probably contravening the radio/licensing regulations of that country. This happened despite the fact that the apparatus was for receiving only and that the driver carried a card showing the approval and agreement of all authorities concerned to allow this experimental equipment to be used (The CEPT Card).

6. CONCLUSIONS

The user trial of paging by satellite was successful. It demonstrated that services could be provided over a very wide geographical area to low priced terminals.

Many lessons were learned in, perhaps unexpected, areas. These included the need for extensive liaison with all users involved, especially the drivers, to ensure they understood the potential benefits. There was a significant desire for a return acknowledgement channel or even a return data channel. Above all there is a need to ensure that the equipment can be taken across European borders and legitimately used in all European countries.

The next step in a marketing assessment would be to consider the impact of two-way data messaging such as INMARSAT-C.

References

1. **McClure, I.** *Lift Off for the Space Page,* British Telecom Journal, Vol8 No.4.

2. **Fenton, C J & Fudge, R E.** 1989 *Evaluational of Radiopaging Via Satellite.* IEE Land Mobile Radio Conference, Warwick.

TABLE 1 USER TRIAL PARAMETERS

Trial duration	4 months
Vehicles used	7 trucks, 1 car
Installation time	2 hours
System availability	7 days per week 24 hours per day
Access from call originator	Telex
Countries covered	UK, Netherlands, France, Belgium, West Germany
Total number of messages	2,000
Message success rate	60% overall

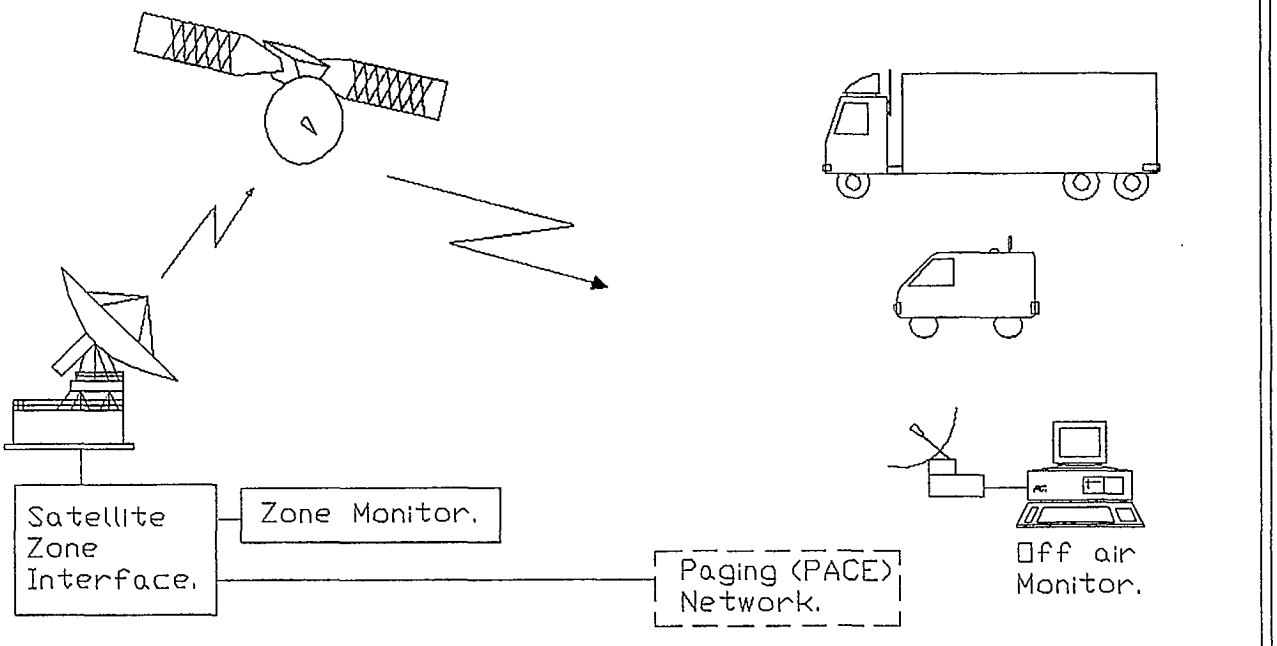


Figure 1 radiopaging via Satellite, Network Configuration.

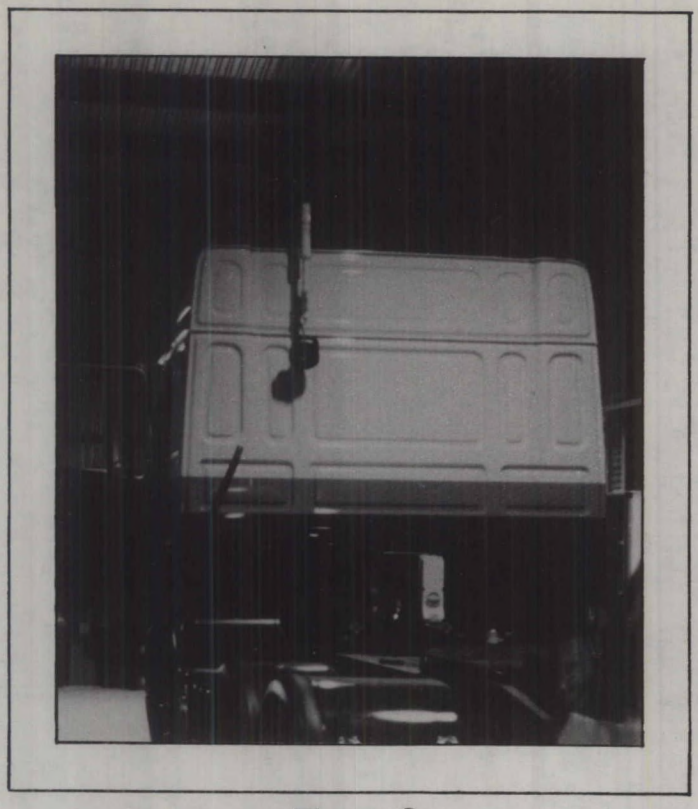


Figure 2

Figure 3

Appendix Index

An Aircraft Earth Station for General Aviation

R. Matyas, J. Broughton, Canadian Astronautics Limited,
1050 Morrison Drive, Ottawa, Ontario K2H 8K7,
Phone: (613) 820-8280, FAX (613) 820-8314

R. Lyons, S. Spenler, SkyWave Electronics Limited, 300 March Road,
Suite 304, Kanata, Ontario K2K 2E2, Phone: (613) 592-0908,
FAX (613) 592-2104

J. Rigley, Communications Research Centre, 3701 Carling Avenue,
P.O. Box 11490, Station "H", Ottawa, Ontario K2H 8S2,
Phone: (613) 991-9309

ABSTRACT

While the focus has been international commercial air traffic, an opportunity exists to provide satellite communications to smaller aircraft. For these users equipment cost and weight critically impact the decision to install satellite communications equipment. Less apparent to the operator is the need for a system infrastructure that will be supported both regionally and internationally and that is compatible with the ground segment being installed for commercial aeronautical satellite communications.

This paper describes a system concept as well as a low cost terminal that are intended to satisfy the small aircraft market.

INTRODUCTION

The provision of aeronautical communications by satellite is gaining increasing acceptance for long range commercial aircraft. International consortia of service providers are currently procuring ground earth stations (GES') in order to offer operational service via the INMARSAT system; some are presently conducting field trials and providing limited service in conjunction with aircraft earth station (AES) manufacturers and airlines.

The provision of voice and data communications capability for smaller aircraft is of considerable interest in view of the very large number of such aircraft in use. Recent studies indicate a potential market of 11,500 AES' worldwide. It is anticipated that these AES' will differ in significant respects from those being developed for commercial airlines.

Canadian Astronautics Limited (CAL) and SkyWave Electronics Limited (SkyWave) are presently developing a low cost, compact, single channel AES. This equipment will allow pilots and passengers to communicate with the worldwide public or private telephone and data networks through services offered by Teleglobe Canada and its international partners in France and Australia, and others.

Topics covered in this paper include compatibility with the INMARSAT aeronautical system, packaging philosophy, effective deployment of subsystems in the aircraft, AES architecture, and field trial objectives.

SERVICES

Aeronautical communications using geostationary satellites

overcomes the line of sight limitations of VHF radio and the unreliability of long range HF radio. Global coverage except at the poles will be in place early in 1991.

International (Inmarsat) and regional (TMI,AMSC, AUSSAT) mobile satellite system operators will provide access to the public switched telephone and data networks through their GES'. Accordingly, telephony services now provided in an office can be extended to an aircraft. These services will support the following [1,2].

Passenger Correspondence

- Voice communications via standard telephone handset.
- Data transmission via laptop personal computer.
- Facsimile transmission via portable fax unit.

Cockpit Communications

- Voice communications via headset.
- Pro-forma and coded messages via customized interface unit.

SYSTEM ARCHITECTURE

Aero Interoperability

The General Aviation (GA) service is based on standards by Inmarsat [3], ARINC [4] and ICAO which have application to both international and domestic commercial aeronautical satellite networks. Figure 1 depicts the GA system architecture.

Both Commercial and GA services will adopt Inmarsat-compatible Access/Control/Signalling (ACS) channels. GA circuit-mode voice channels can be expected to evolve with improvements in lower rate vocoder and modulation technologies providing lower cost service and AES equipment.

Voice Communications

The Inmarsat SDM specifies a 9600 bps vocoder followed by a 21 kbps Aviation-QPSK modulator with rate- 1/2 error correction coding (FEC) for C-channel voice communications. Background bit error rate monitoring is performed to permit BER measurements and dynamic power control information to be exchanged via the C-channel sub-band. This sub-band includes some additional capacity for low rate user data and system messages.

In the GES-to-AES direction, the A-QPSK carrier is voice-activated to conserve satellite power during periods of speech inactivity. On the reverse C-channel, the modulator uses continuous-mode A-QPSK with the power level dynamically adjusted under command from the GES.

For the GA service it is expected that a standard will be developed that uses 4800 bps digitally encoded voice. Pending a decision, the GA AES discussed here will allow for the use of either 9600 bps voice or a more bandwidth efficient modulation - ACSSB. ACSSB voice has been field proven for the Ontario Air Ambulance Service [2].

AES DESIGN APPROACH

The challenge in designing a satellite terminal for General Aviation is to produce a unit whose cost, weight and size are suitable for smaller aircraft while at the same time retaining

complete compatibility with the commercial aviation ground segment and Inmarsat standards. This has been accomplished by the following simplifications to the ARINC 741 baseline [4].

- Use of a single telephone channel with switched user interfaces.
- Reduction of the number of functional units relative to an ARINC-741 configuration by integrating (1) satellite data, beam steering and RF units, and (2) the LNA and diplexer with the antenna. This will reduce both equipment and installation costs.
- Use of a single top-mount antenna to minimize the number of fuselage modifications.
- A level of built-in test capabilities commensurate with expected aircraft usage.
- Designing each unit for 28 VDC power which is generally available on smaller aircraft, rather than AC power.
- Incorporating integral unit thermal cooling rather than relying on rack cooling which may not be available or practical in a small aircraft.
- Billing by aircraft registration eliminates the need for credit card reading equipment.
- Reduction and simplification of inter-unit connections and interfaces to other avionics.

AES CONFIGURATION

The GA AES comprises four distinct functional elements: a user interface, a transceiver, a high power amplifier and an antenna. The AES configuration is shown in Figure 2. These units are described as follows.

User Interface

The User Interface (UI) supports voice/data set connections, provides visual indication of operational status and incorporates features to support factory integration and test, and maintenance. The UI is placed in an aircraft where appropriate, allowing local or remote user access.

Transceiver

The transceiver incorporates a channel unit, an IF/RF unit including frequency reference and an antenna controller.

The channel unit includes a communications controller, a modem and a baseband to IF frequency converter. The communications controller implements sub-network and link layer protocols, converts asynchronous user data to synchronous form, performs low level system diagnostics, controls the modem and services commands including supervisory local/remote functions from the user interface and P-channel data. The modem implements P-, R-, T- and C-channel protocols and transfers signal units to/from the link level software in the communications controller. The modem also programs the synthesizers in the frequency converter. The baseband to IF frequency converter independently synthesizes transmit and receive frequencies in 2.5 kHz steps over the allocated uplink (34 MHz) and downlink (29 MHz) bandwidths. It also performs AFC and Doppler frequency correction under control of the controller. The relative P-channel frequency offset is used to

precompensate for the Doppler frequency shift on the transmitted carrier.

The IF/RF unit performs independent block frequency conversion for each of the transmit and receive chains between the intermediate and L-band frequencies. A high stability ($\pm 3 \times 10^{-7}$), low phase noise 6 MHz reference oscillator is also included.

The antenna controller supports open- and closed- loop antenna steering and includes the antenna drivers. Open loop control involves computing satellite direction based on geo-stationary satellite position and aircraft attitude and position data provided on an ARINC-429 bus from an onboard navigation system such as an INS or GPS receiver. Closed loop operation involves a wide area scan of the directional antenna for satellite acquisition followed by dithering of antenna pointing during tracking mode. A signal quality estimate derived in the modem from the received signal is the basis for closed loop beam control.

High Power Amplifier

The HPA is a high power Class A amplifier capable of supporting modulations requiring linear amplification as well as future multicarrier operation if needed. The unit provides a 48 dBm 1 - dB compression point. High output power capability of the HPA allows for flexible deployment in a wide range of aircraft accounting for losses due to long cable runs. A modular design provides for lower outputs if needed.

Antenna

The top mount antenna subsystem described more fully in [5] incorporates the antenna, low noise amplifier and diplexer and provides full azimuth coverage (without gaps) as well as elevation coverage from 5°-90°.

Prototype Terminal

The prototype terminal characteristics are included in Tables 1 and 2. The prototype C-channel is expected to use ACSSB for voice with 9.6 kbps vocoded speech provided in production units. The prototype will be available for field trials with a test hub in the fall of 1990.

ACKNOWLEDGEMENT

This development has been supported by Supply and Services Canada, Transport Canada and Communications Canada. The cooperation of Teleglobe Canada is also gratefully acknowledged.

REFERENCES

1. J. Rigley, "Aeronautical Mobile Satellite Service: An Overview", Proceedings Canadian Satellite User's Conference, Ottawa, 1989.

2. R. Huck, "Satcoms and general aviation" in Never Beyond Reach, B. Gallagher ed., INMARSAT, London, 1989, pp. 164-169.

3. "Aeronautical System Definition Manual", Inmarsat, London, England.

4. "Aviation Satellite Communications System - ARINC Characteristic 741-1, Parts I and II", Aeronautical Radio, Inc., Annapolis, Maryland.

5. P. Strickland, "Low Cost Electronically Steered Phased Array for General Aviation", this conference proceedings.

Table 1
GA AES Characteristics

Access, Control, Signalling and Data Channels	P-, R- and T- channels per Inmarsat Aeronautical SDM. BER < 10^{-5} with FEC
Voice Channel	Full duplex ACSSB (prototype): 2.5 Mean Opinion Score Digital vocoding (future) Voiceband FAX A-BPSK data (prototype)
Channel Bandwidth	P,R,T channels: 2.5 kHz C channel: 5 kHz
Frequency Range	Receive: 1530 - 1559 MHz Transmit: 1626.5 - 1660.5 MHz
EIRP (dBW)	R/T-Channels: 13.2 dBw (600 bps) 16.2 dBW (1200 bps) ACSSB C-channel: 20.3 dBW (average)
Antenna Installation/Coverage	Top-Mount (360° AZ;5-90° EL)
Antenna Steering	Open loop (baseline)
G/T	-13 dB/K
Environmental	RTCA DO/160B

Table 2
Configuration

Assembly	Size	Weight
Antenna/Radome	1.4m x 0.2m x 0.35m	12 kg
Transceiver	8 MCU avionics enclosure	10.5 kg
HPA	8 MCU avionics enclosure	9.8 kg
Cockpit User Interface	7.6cm x 14.6cm x 17.8cm	1.5 kg

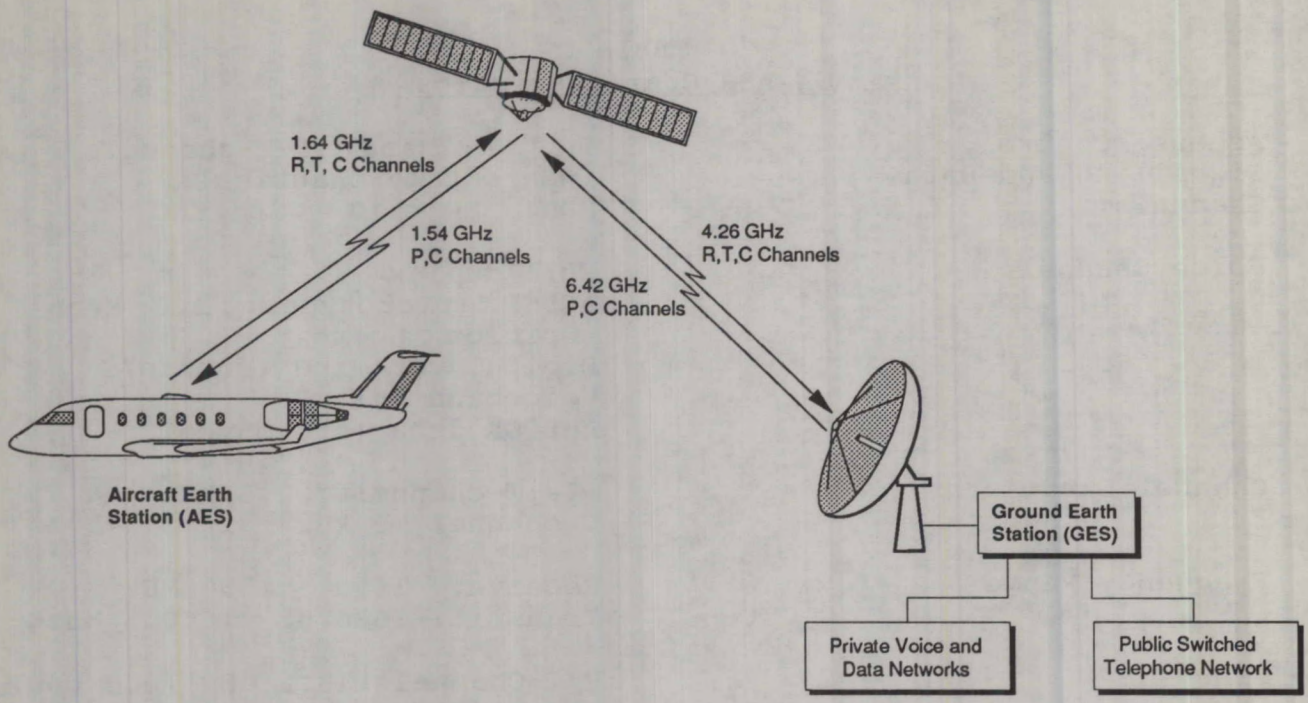


Figure 1 GA System Architecture

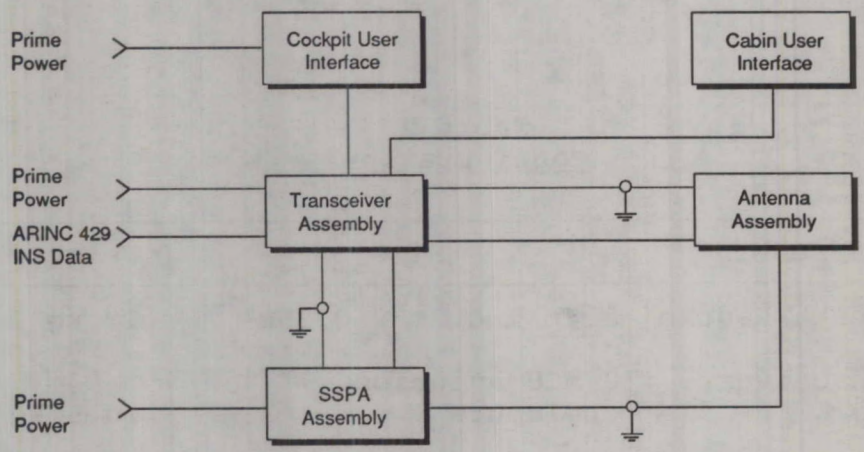


Figure 2 AES Configuration

Frequency Stabilization for Mobile Satellite Terminals via LORAN

Gregory J. Ernst, Steven M. Kee, and Robert C. Marquart
Hughes Network Systems, Inc.
Germantown, Maryland 20876

ABSTRACT

Digital satellite communication systems require careful management of frequency stability. Historically, frequency stability has been accomplished by continuously powered, high cost, high performance reference oscillators. Today's low cost mobile satellite communication equipment must operate under wide ranging environmental conditions, stabilize quickly after application of power, and provide adequate performance margin to overcome RF link impairments unique to the land mobile environment. Methods for frequency stabilization in land mobile applications must meet these objectives without incurring excessive performance degradation. A frequency stabilization scheme utilizing the Loran (Long range navigation) system is presented.

INTRODUCTION

Position determination and reporting are requirements in many mobile satellite terminal applications. Presently, the Loran-C system is a widely available and a low cost means of providing position determination. Many mobile satellite terminals include an integrated Loran-C Receiver for this purpose. A high stability frequency reference can be derived from the received Loran-C signal. Application of this technique to transmit-only mobile satellite terminals is particularly useful.

As shown in Figure 1, the Loran-C Receiver determines the frequency offset between a voltage controlled oscillator reference and the received Loran-C signal. Periodically, the mobile satellite terminal baseband processor system interrogates the Loran-C Receiver to retrieve this frequency offset. After data processing, the baseband processor corrects the frequency offset by tuning the voltage controlled oscillator via a digital to analog converter. This negative feedback control loop removes the frequency offset effects due to both temperature and time.

The implementation of this control loop incorporates algorithms to ensure reliable frequency stabilization. The Loran-C Receiver computes the frequency offset and derives a received signal quality indicator. The baseband processor ignores frequency offsets accompanied with a poor signal quality indication. Otherwise, the frequency offset is processed to establish an updated frequency correction. The most recent frequency correction is always stored in nonvolatile memory so that upon power application, a stabilized frequency can be obtained promptly. During periods of extended Loran-C outages, the stability is maintained because the voltage controlled oscillator is temperature compensated.

REQUIREMENTS

The mobile satellite terminal must provide long-term and short-term frequency stability under the following conditions:

- Extreme operating temperatures
- Rapid recovery after power cycling
- Impaired Loran-C reception
- Aging effects over time and temperature

- Automatic operation without user intervention

In addition, the stabilized oscillator reference provides the critical clock timing and phase noise performance for the mobile satellite terminal.

CONCEPTUAL MODEL

Figure 2 shows a conceptual diagram of the mobile satellite terminal frequency stabilization scheme using the Loran system as the reference source. A quartz crystal voltage controlled oscillator (VCXO) is used as the frequency reference source within the mobile satellite terminal. This VCXO has stability characteristics as described in Figure 3. As shown in the Time Domain Characteristics of the VCXO, the quartz crystal oscillator by itself is sufficient to provide the short-term stability. However, the long-term frequency drift can become large when uncompensated. Utilizing the Loran system for frequency stabilization eliminates the effects of long-term frequency drift.

The Loran-C Receiver utilizes the 100 kHz Loran signal as the reference input, $r(t)$, to the frequency stabilization control loop. Embedded within the position determination process is a frequency comparison algorithm which calculates the frequency offset between the VCXO controlled output, $c(t)$, and the reference input. The frequency offset is represented by $e(t)$.

The digital controller provides the overall operational control for the frequency stabilization process. The input to the digital controller is the output of a sampled analog to digital conversion process, $e(kT)$. The sampling period T is defined by the baseband processor interrogation interval of the Loran-C Receiver. The sampled frequency offset data is processed by the digital controller to produce a frequency correction, $m(kT)$. The digital to analog converter develops an analog tuning voltage to the VCXO based upon the frequency correction.

The VCXO is temperature compensated and not ovenized; therefore, it can quickly stabilize in frequency when power is applied. The frequency characterization of the VCXO shown in Figure 3 forms the basis to determine the control loop parameters. In addition, the VCXO frequency tuning range must be sufficient to tune to the desired center frequency for temperature variations and design life of the mobile satellite terminal. When the Loran system is not available, the oscillator must be frequency compensated for temperature variations so that mobile satellite terminal operation is not disrupted. The long-term aging effects can be easily maintained to less than 0.1×10^{-6} with only monthly Loran availability. Temperature stability without Loran availability is determined by the design of the VCXO but can be less than 0.1×10^{-6} with a state-of-the-art implementation.

IMPLEMENTATION

The implementation of the frequency stabilization control loop has been developed using a 16 MHz VCXO. The output of the 16 MHz VCXO provides the reference frequency for critical timing including position determination and carrier generation. The VCXO tuning voltage is the output of an 8-bit digital to analog converter (DAC). The initial value for the DAC is stored in nonvolatile memory and is applied as the frequency correction upon application of power and initialization of the mobile satellite terminal. The DAC value is updated as required to compensate for the frequency drift of the VCXO.

The frequency offset from the Loran-C Receiver is derived from the Loran system timing as well as the average of all received Loran signals. This averaging minimizes the effects of Doppler induced errors due to the geographic separation of Loran system transmitting stations. The frequency offset provides a resolution of 62.4×10^{-3} Hz per bit.

After initialization of the mobile satellite terminal, the baseband processor waits for a received signal quality indication from the Loran-C Receiver. During periods of poor signal quality, the frequency correction is not updated. For extended periods of poor signal quality as a result of Loran-C impairments (e.g., interference, shadowing, areas of inadequate coverage), the frequency stability of the mobile satellite terminal is provided by the VCXO.

The frequency stabilization loop operates continuously once there is good received signal quality. The algorithms within the baseband processor correct the VCXO frequency at a rate that is sufficient to track out the effects of temperature and aging.

The baseband processor examines the frequency offset, and increments or decrements the frequency correction as required. The resolution of the frequency correction through the DAC is nominally 0.05×10^{-6} per bit. The frequency correction is updated whenever the frequency offset is equal to or greater than 0.1×10^{-6}. When the loop has been adjusted to a frequency offset of less than 0.1×10^{-6}, the baseband processor compares the current frequency correction to the value stored in nonvolatile memory. The nonvolatile memory is updated if the difference between these two values is greater than 0.5×10^{-6}.

TEST RESULTS

Figure 4 shows the measured response of the frequency stabilization loop to a $+1.00 \times 10^{-6}$ change of the VCXO frequency. Initially, the VCXO frequency was $+0.17 \times 10^{-6}$ from the nominal 16.00 MHz. A step change to $+1.17 \times 10^{-6}$ from the nominal 16.00 MHz was applied to the VCXO. The frequency offset response attained 63% of its final value in 50 seconds. In 180 seconds, the frequency offset had attained 95% of its final value. This test was performed for -1.00×10^{-6} change to the VCXO frequency and the results were comparable to the positive change in VCXO frequency. Both of these tests were repeated through successive trials, temperature cycling, and power cycling. Again, the results were consistent for a 1.00×10^{-6} change in VCXO frequency.

Incremental changes of the VCXO frequency were tested up to the $\pm 3.00 \times 10^{-6}$ acquisition limit of the Loran-C Receiver. The results showed that for a $+3.00 \times 10^{-6}$ step, 95% of the frequency offset was attained in 44 minutes on the average. A -3.00×10^{-6} step resulted in an average of 42 minutes to attain 95% of the frequency offset.

Testing of the frequency stabilization control loop also included parametrics, acquisition and tracking, firmware integration, and error handling. Parametrics and acqusition and tracking tests were also conducted over temperature and time.

Extensive field tests of the Hughes Network Systems SkyRider™ Mobile Satellite Terminal have demonstrated frequency stability performance of better than 0.03×10^{-6} over temperature and time with continuous Loran coverage.

CONCLUSION

An implementation scheme for obtaining a stabilized frequency reference for land mobile applications has been presented. A combination of long-term frequency drift compensation via Loran system and short-term frequency stability characteristics of a crystal oscillator meets the requirements for an accurate and stable mobile satellite terminal frequency source. This frequency stabilization technique is particularly applicable to transmit only or random access type mobile satellite terminals.

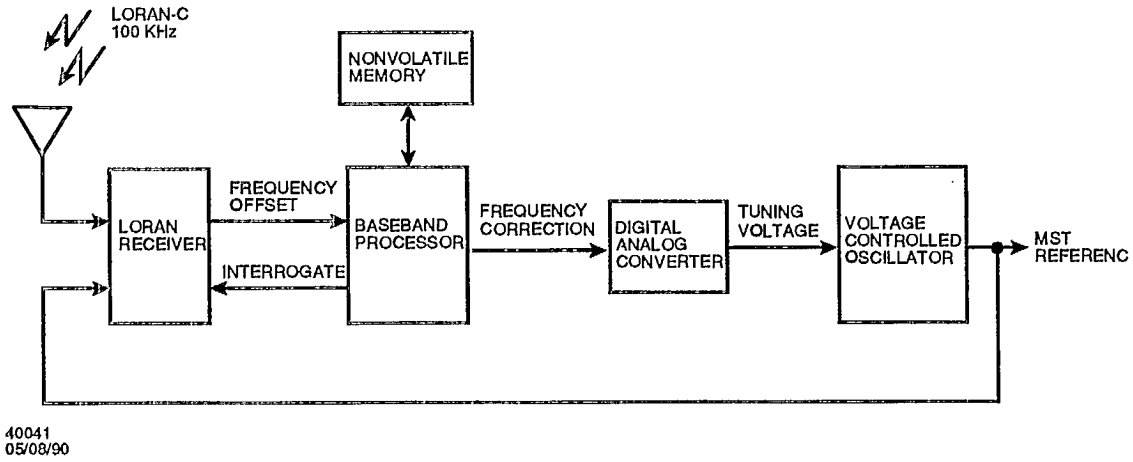
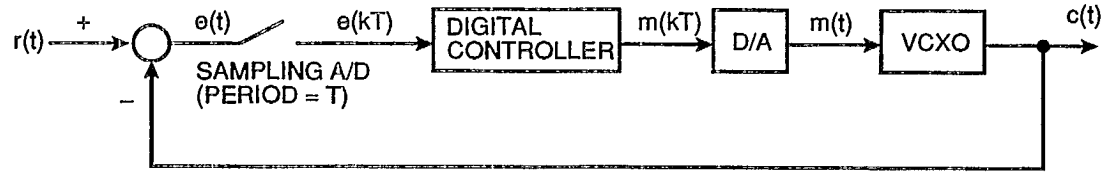


Fig. 1. Functional Diagram



T = SAMPLING PERIOD

r(t) = REFERENCE INPUT

c(t) = CONTROLLED OUTPUT

e(t) = r(t) – c(t), ACTUATING SIGNAL

m(t) = MANIPULATED SIGNAL

40042
05/08/90

Fig. 2. Conceptual Model

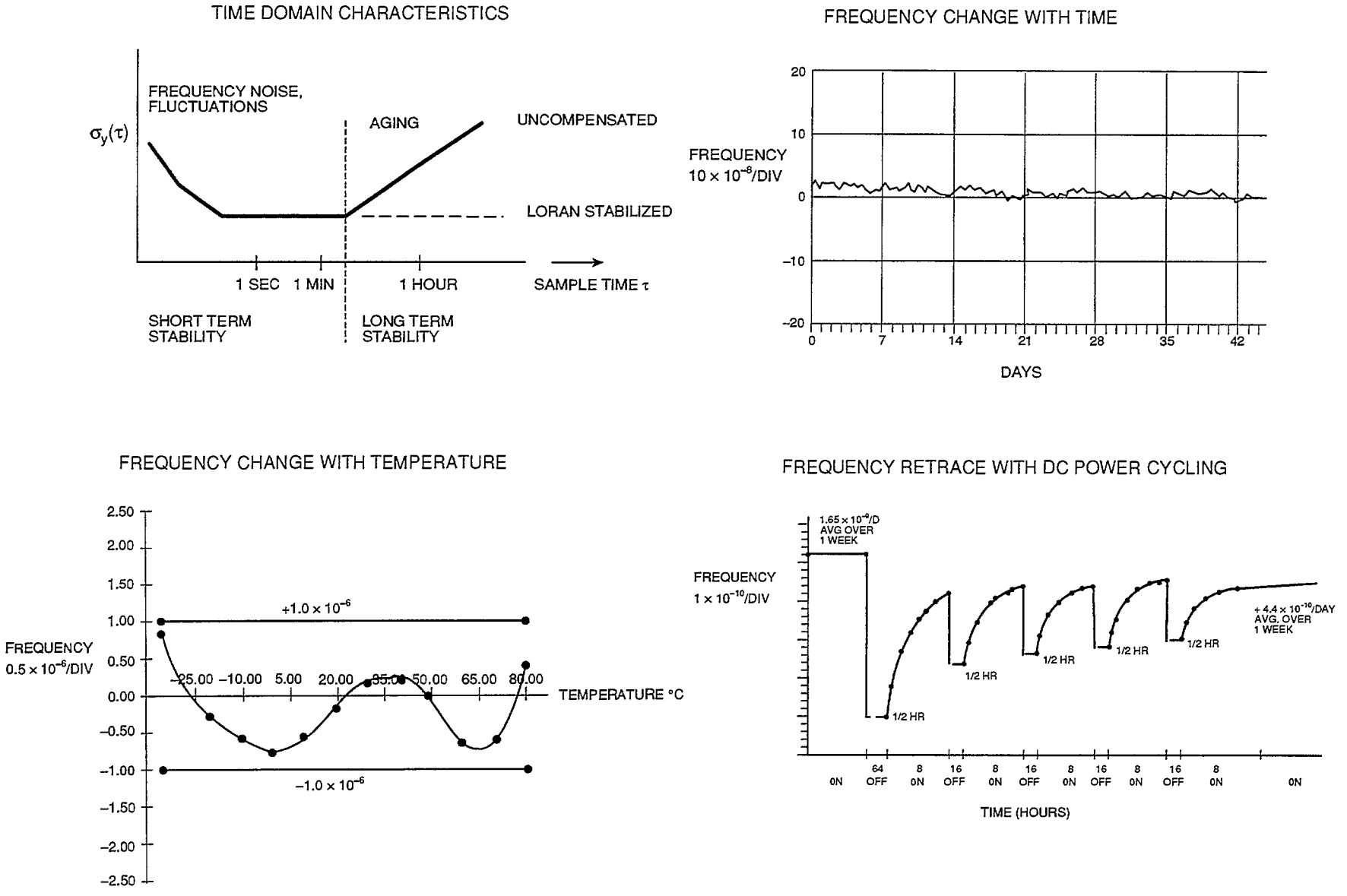


Fig. 3. VCXO Characteristics

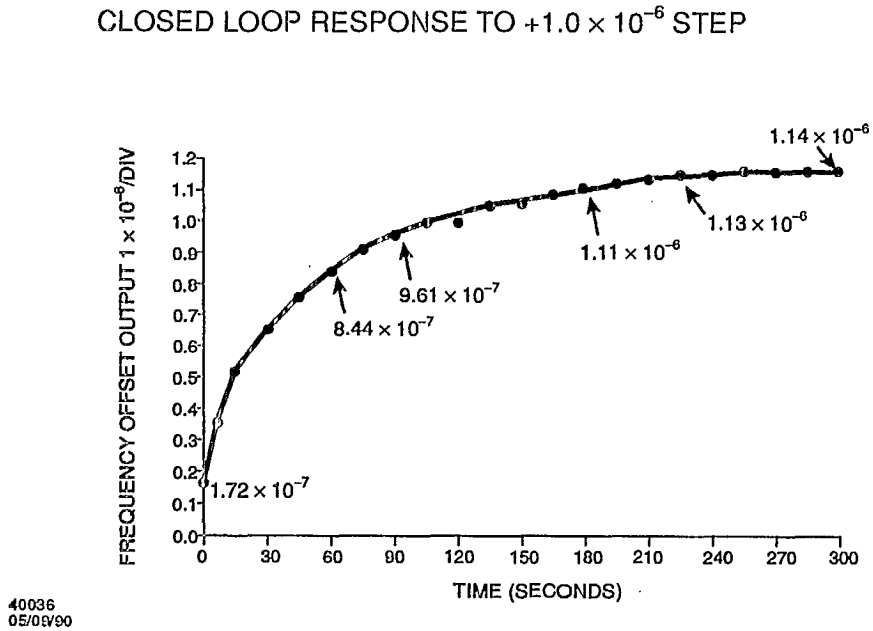


Fig. 4. Closed Loop Response

Satellite Mobile Data Service for Canada

Glenn R. Egan and David J. Sward
Telesat Mobile Inc.
333 River Road
Tower A, 14th Floor
Ottawa, Ontario
K1L 8B9
Telephone: (613) 746-5601
Facsimile: (613) 746-2277

ABSTRACT

Telesat Mobile began commercial operations on December 15, 1988 with the signing of a shareholder agreement between Telesat Canada, Canadian Pacific Ltd. and the C. Itoh Group of Japan. TMI's mission is to construct and operate a commercial mobile satellite system in Canada. This will be done in two distinct phases.

First, TMI will introduce mobile data services in June, 1990. A contract has been awarded to Canadian Astronautics Ltd. supported by Gandalf Systems Ltd. and Hughes Network Systems for the supply of hub equipment and 3,000 mobile data terminals. Over-the-satellite tests began in February, 1990. The mobile data service will provide full two-way digital messaging, automatic vehicle location and fleet management services.

The second phase is to construct, launch and make operational the MSAT satellite and associated network control facilities. This paper will focus on Phase I, i.e. the implementation of a mobile data service in Canada. in addition to a technical description, the paper will provide information on markets and applications.

BACKGROUND

The Canadian Mobile Satellite system (MSAT) is approaching completion after almost ten years of planning. The Federal Department of Communications pioneered the technical and commercial development of MSAT from 1983 to 1986. Telesat Canada worked closely with the DOC during this period. Telesat Mobile Inc. (TMI) was later formed in December, 1988 to fulfill the mandate of implementing mobile satellite services in Canada.

TMI's business plan is based on providing a North American service through joint operations with a licensed American mobile satellite operator. In May, 1989, the Federal Communications Commission authorized the American Mobile Satellite Corporation (AMSC) to provide mobile satellite service in the United States. In April, 1990, TMI and AMSC formally entered into a Joint Operating Agreement which provides a solid foundation for implementing a North American system.

TMI's major shareholders are Telesat Canada, Canadian Pacific Ltd. and the C. Itoh Group of Japan. Bell Canada Enterprises Mobile (BCEM) also has a financial interest in TMI and will distribute MSAT products and services.

TMI plans to introduce mobile satellite services in two phases. Phase I services will be introduced during the construction period of the MSAT satellite i.e. 1990 - 1993. These services include a mobile data service for introduction in June, 1990 and a "Stop and Talk" voice service planned for introduction in June, 1991. Phase II services will commence when MSAT is commissioned in late 1993 and consist of a full range of integrated voice and data services under the portfolio name "mobile ISDN".

The mobile data service provides two-way messaging, automatic vehicle location and fleet management. A contract was awarded to Canadian Astronautics Ltd. in October, 1988 to provide hub messaging equipment and vehicular terminals. Gandalf Systems Ltd. and Hughes Network Systems are major sub-contractors to this program.

Satellite capacity has been leased through Teleglobe Canada through Teleglobe Canada to facilitate this service. An INMARSAT satellite has been repositioned to 106 W to provide coverage of Canada and the U.S. Current plans are to operate the mobile data service on this satellite commencing late June, 1990. The satellite coverage provided from the 106 W position is shown in Fig. 1.

MOBILE DATA SERVICE

The mobile data system provides solutions on an end-to-end basis. It is a combination of two elements: a satellite-based, packet-switched network and network access systems.

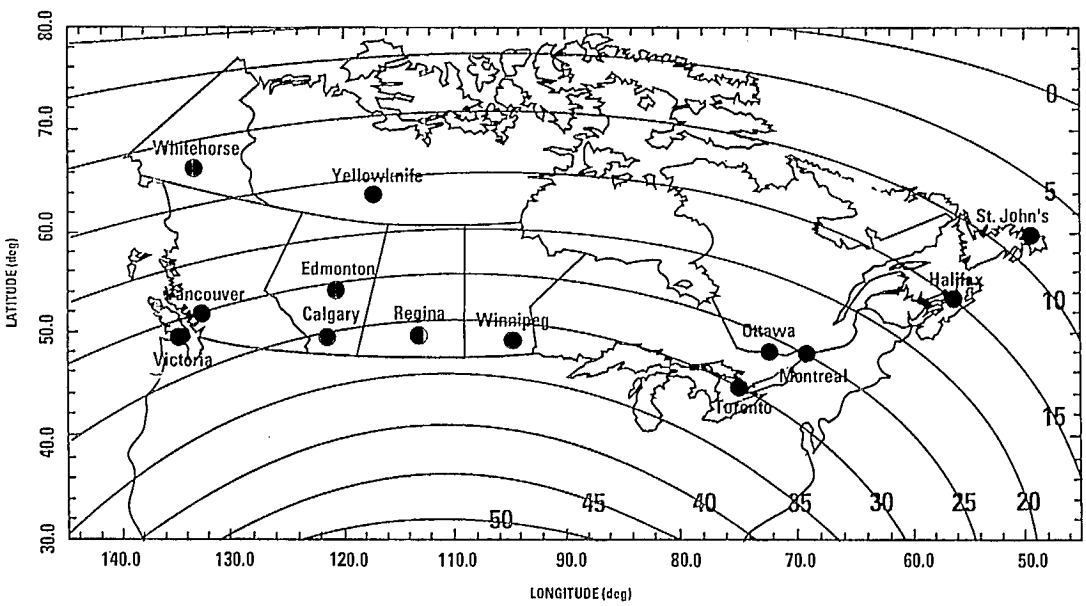
The satellite-based, packet-switched network provides data throughput, route diversity, flow control of multiple network access systems, and error detection with error correction and data retransmission algorithms. The network performs functions similar to most packet-switched networks that provide for reliable, error-free transmission. (See Fig. 2.)

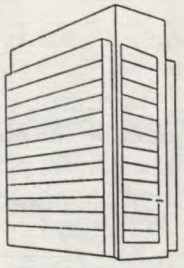
Additional value is offered to users of the network through the flexible network management system.

The network management and control system supports the network as a whole as well as individual fleets. Issues such as network access authorization, radio frequency assignments and overall system fault detection, isolation and correction are managed exclusively by Telesat Mobile network operations staff.

Each element of the network is redundant and fault-tolerant design ensures continuous operation. Many aspects of network management and

Fig. 1 Coverage from 106° W Longitude

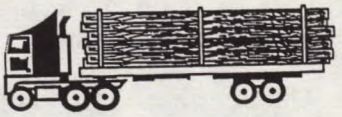




Customer Central
Network Access Systems

- Dumb terminal
- Single user DOS PC
- Multi-user multi-tasking UNIX PC
- Sub-network controller
- Graphics subsystems

Satellite
Packet-switched
Network



Mobile/Remote
Network Access Systems

- Hand-held driver terminal
- DOS Laptop/PC
- UNIX Laptop/PC
- Single-port PAD
- Multi-port PAD

Fig. 2 Mobile System Data Connectivity

control relate to the operation of a particular customer fleet, and the network through partitioning and secure network access procedures allow individual customers to directly manage their own fleet.

Prior to the launch of MSAT, L-band space segment is at a premium. Telesat Mobile has secured long term leases through Teleglobe to support our program. Efficient allocation of the space segment is critical to providing uniform quality service to several thousand mobiles. This is achieved through segmenting data according to its traffic profile and utilizing multiple satellite channels, each optimized to suit a particular type of traffic. The network architecture supports regularly scheduled reporting, demand access short messaging and demand access file transfer.

Inbound traffic is routed via three distinct channel types. Regularly scheduled information such as vehicle locations or remotely sensed data can be transmitted on one of the position reporting channels. Since all transmissions are pre-assigned, extremely high efficiencies are attainable on these channels.

Randomly scheduled short messages are transmitted on one of the general messaging channels. The satellite link protocol utilizes a combination of random access slotted aloha and pre-assigned TDMA to transmit messages. Each frame on the inbound channel is portioned into a random access and pre-assigned segment. This moveable "fence" separating the two access schemes allows for system tuning based on actual traffic patterns.

According to the total message length, the message is divided by the mobile unit into a number of packets. The first packet is transmitted in the random access portion of the frame and the remainder of the packets is transmitted in the pre-assigned portion. Information embedded in the outbound channel informs all mobiles of the upcoming slot reservations and assignments. Demand access file transfer is provided on dedicated circuit-switched channels that allow for continuous packet transmission.

Outbound messages are carried on a TDM carrier. The carrier utilizes variable packet lengths in each 639 information byte frame. This assures maximum utilization of the high power forward carrier. The outbound carrier transmission rate is 1200 sps. The mobile units employ two speed modems to support both 600 sps and 1200 sps transmission rates. (See Fig. 3.)

Network access systems are provided at the mobile and organization's central facility. These systems define the real functionality of the service and provide

the user with transparent access to the packet-switched network. Since the interface specifications to the network are open, a broad array of network access systems are evolving.

The interface to the mobile portion of the packet switched network consists of the two RS-232 ports. Devices are currently available to support simple message creation, editing, storage, transmission and receipt in a hand-held terminal. Laptop computers and single user PC's are also available. High end multi-user platforms with extensive networking capability are also provided to support more sophisticated applications. The access devices utilize DOS and UNIX operating systems to allow the migration of existing software applications to the remote or mobile domain. This allows users to continue to use familiar programs when they move from the office to the mobile environment.

Additional access devices are available to support unmanned sites such as data collection platforms or monitor and control stations. Ranging from single port to multi-port programmable, these devices offer the flexibility required to support a broad array of applications.

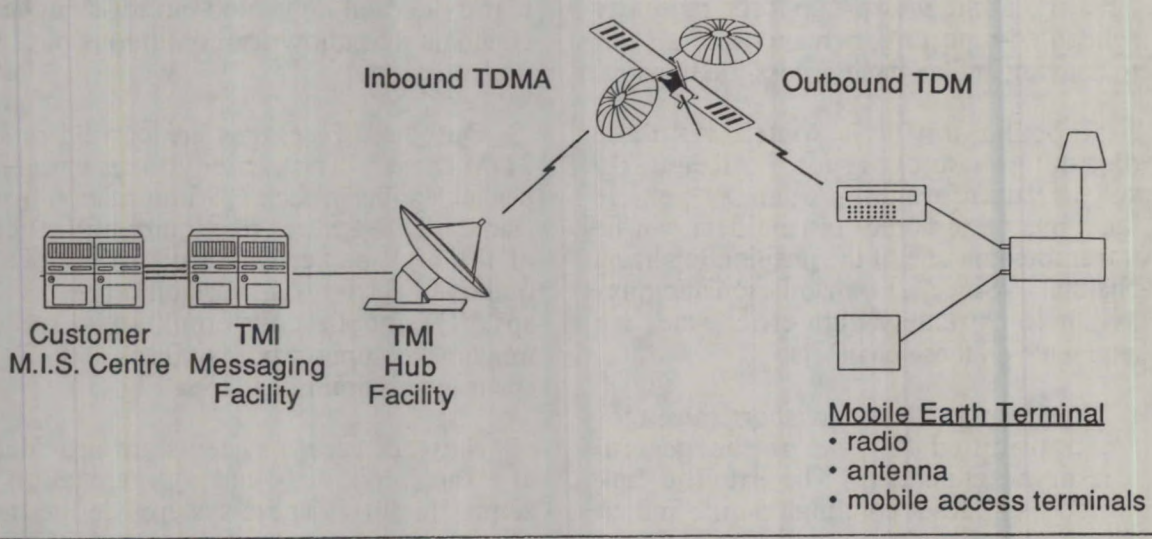
At the customer's central office, network access is also supported by a variety of systems. The most elementary of these consists of a dumb terminal driven by the network that provides full text messaging and printing capability. Single user DOS-based PC's and multi-user UNIX-based micro-systems are also available that support full text and data messaging and powerful graphics displays.

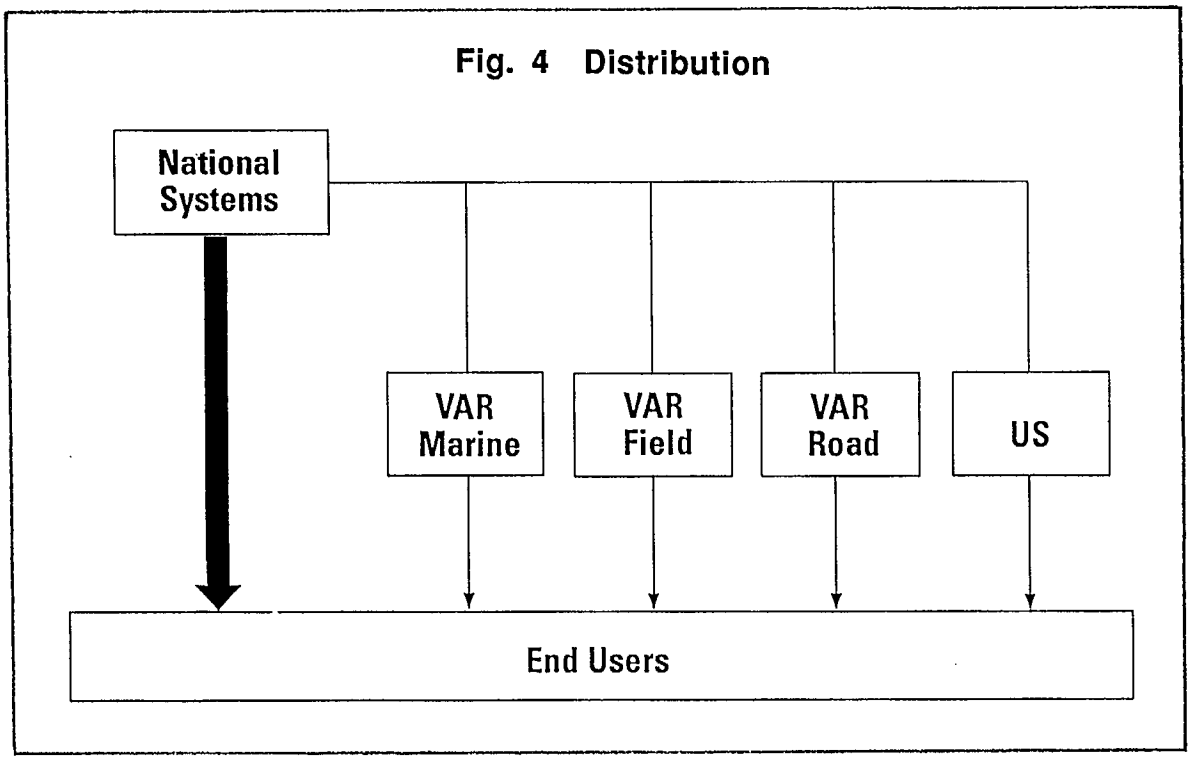
Graphics systems may be modified by customers to suit their own applications and can be displayed on standard monitors, high resolution monitors or full screen projection systems. For customers with existing unique applications, interface specifications for message and file transfer are provided, and integration support is available.

DISTRIBUTION

TMI is in the process of establishing a sales and support network for MDS through agreements with value-added resalers (VARs). VARs will provide a local presence for sales and maintenance functions. VARs' responsibilities may be segmented by service type i.e. Road KIT, Field KIT, or by geographic territory. VARs also contribute significantly as the name suggests to the "value" of the service

Fig. 3 Mobile Data System Elements





by providing specified access software or custom MET interface devices. TMI has entered into VAR agreements with Sea Link Ltd. of St. John's, Nfld. and Munro Engineering Ltd. of Calgary, Alta. (See Fig. 4.)

CONCLUSION

The Phase I Mobile Data Service is an important first step for Telesat Mobile. MDS meets the needs of the transportation, marine and SCADA markets by providing wide area, reliable communications and fleet management services. MDS will allow TMI to establish a market presence prior to the launch of MSAT and provide an initial base of users for transfer to MSAT when commissioned in 1993. MDS is the first mobile satellite service in TMI's eventual full service portfolio of Mobile ISDN.

OSI-Compatible Protocols for Mobile-Satellite Communications: The AMSS Experience

Michael Moher
Communications Research Centre
Ottawa, Canada

ABSTRACT. *The protocol structure of the international aeronautical mobile satellite service (AMSS) is reviewed with emphasis on those aspects of protocol performance, validation, and conformance which are peculiar to mobile services. This is in part an analysis of what can be learned from the AMSS experience with protocols which is relevant to the design of other mobile-satellite data networks, e.g., land-mobile.*

1.0 INTRODUCTION

This paper reviews material presented to the Aeronautical Mobile Satellite Services (AMSS) Panel of the International Civil Aviation Organization (ICAO) regarding the protocols for the planned international aeronautical-mobile data service. Particular emphasis is placed on the datalink protocols and some of the work being done at the Communications Research Centre (CRC) in this area.

2.0 OVERVIEW OF AMSS AND OSI

One underlying premise of the AMSS is that it is constructed in manner which is consistent with Open Systems Interconnect (OSI) principles and in particular that it uses the seven-layer protocol stack illustrated in Figure 1. The work of the AMSS panel is concerned with the bottom three layers of this model: the physical, the datalink and the network layers. However, the AMSS is viewed as representing a satellite subnetwork within a much larger Aeronautical Telecommunications Network (ATN), which includes numerous other subnetworks among which are VHF datalink networks, radar communication networks and fixed data networks. One conception of the ATN is illustrated in Figure 2. The prime motivation for adopting the OSI model as a guide in developing the AMSS is this view of AMSS as part of a much larger ATN.

Among the attractions of the OSI model are the belief that it leads to more interoperable systems both intranetwork and internetwork, and the standard protocols available for the different protocol layers: protocols which have been to some extent validated and thoroughly tested. A further attraction of this approach is that the communication system is kept independent of the application, implying a much easier maintenance and upgrading of application software.

The drawback of the OSI model is that it is inherently designed for medium and wideband data networks. In mobile systems, power constraints imply that data rates range from hundreds of bits to a few kilobits per second, which is one to two orders of magnitude lower than the typical minimum data rate one would find in fixed data networks. Consequently, the overhead due to the use of seven layers of protocols can use proportionately more of an already scarce resource. Minimizing the message length at the application layer only eases this problem slightly, because at some point the contribution of the protocol overhead to the total message length becomes dominant.

The single example of Automatic Dependent Surveillance (ADS) messages illustrates many of these aspects of an OSI system which are important in mobile applications. ADS reports are aircraft position information derived from on-board navigational aids and are a potentially important air

traffic control application of AMSS. At the application layer most of the redundancy is removed from these messages leaving a standard message length of 88 bits. However, after passing through the top five OSI layers one calculation [1] shows that the message length presented to the data link layer is 248 bits; the majority of this additional overhead is due to the 128-bit Network Service Access Point (NSAP) address standard for the ATN (128 bits) [2]. However, the greatest overhead (most of which is not attributable to OSI) occurs when this message is converted to the 1368 bits which are transmitted over the channel. This latter figure includes the bits needed for a burst preamble and synchronization, and also includes the rate 1/2 coding applied to all information bits. On the other hand, designing to OSI principles means that changes in the content or the length of the ADS report can be made without requiring any change in the delivery system. The interconnectability offered by the ATN means that the ADS report can be automatically delivered over the best of a number of communications alternatives, e.g. satellite, VHF data link, or possibly Mode S surveillance radar.

This example clearly illustrates the important consideration that must be given to the protocols in the design of a mobile data network where there is limited bandwidth available at the physical layer. Opportunities for improving performance exist at all the different layers. For example, at the network layer an NSAP address of 128 bits[2], in principle, allows over 10^{38} destinations to be directly addressed, which is an extremely large amount of flexibility. It would be very advantageous if the network-datalink protocol convergence function limited the amount of flexibility through the use of some default addressing or other forms of address compression.

On the other hand there may be areas where the potential improvements are limited. For example, at the physical layer the very nature of the mobile network implies a necessity for some form of random access strategy and its inherent inefficiency. Ideally one would minimize the use of a random access scheme once the mobile is logged on the system, performing subsequent accesses using some form of controlled access. However, in an OSI

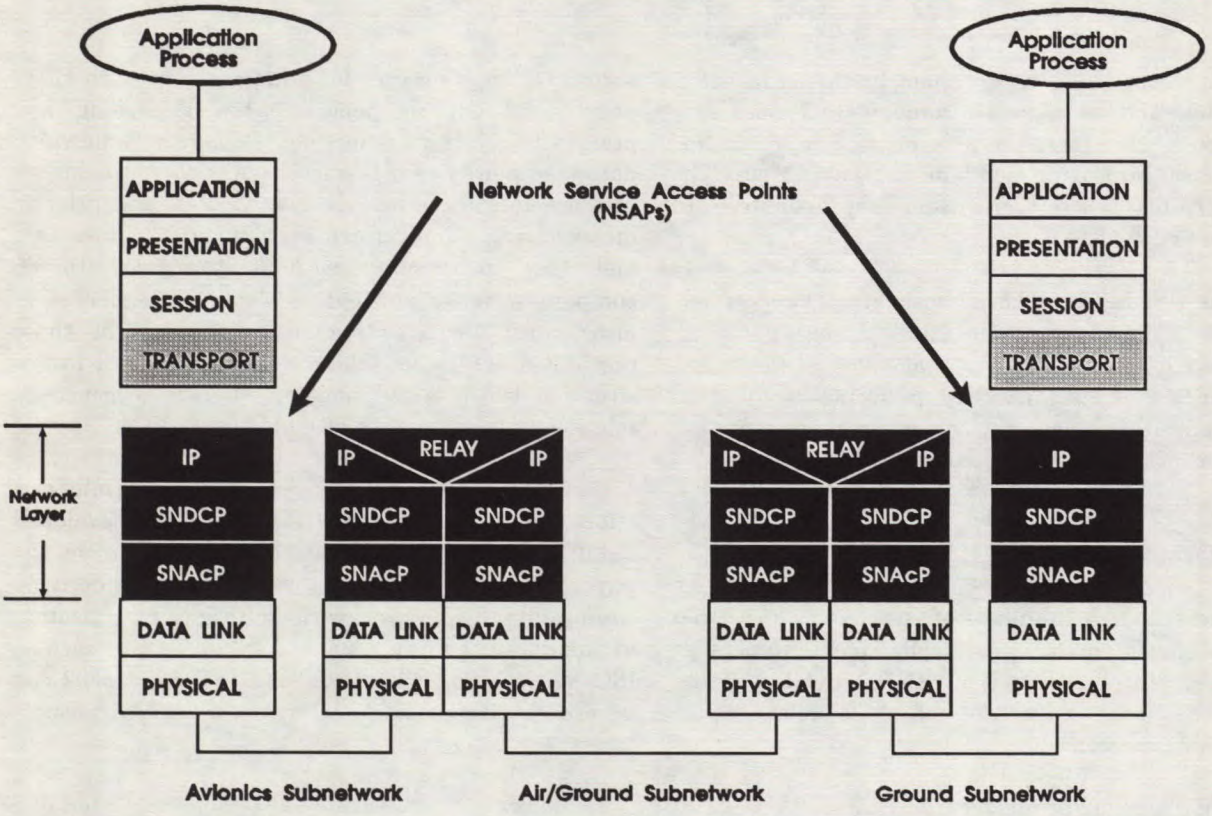


Figure 1. The ATN protocol architecture[2]. (Legend: IP - Internetwork Protocol; SNDCP - Subnetwork Dependent Convergence Protocol; SNAcP - Subnetwork Access Protocol)

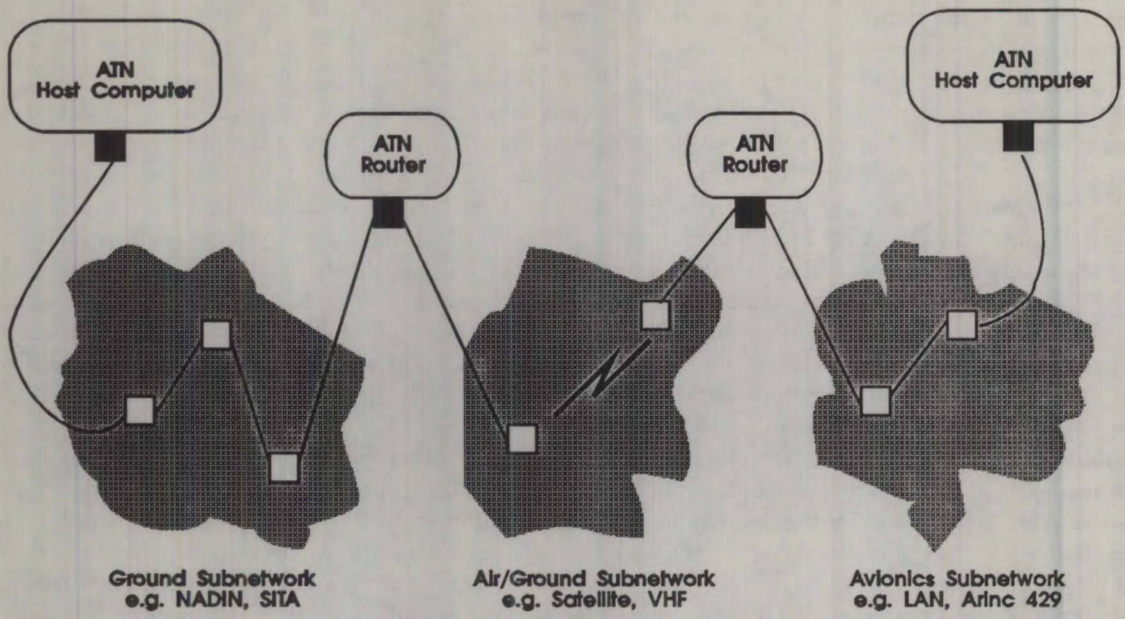


Figure 2. Conception of ATN composed of a number of interconnected subnetworks [2].

system the communications cannot be tailored to the application and as a result some inefficiencies are irreducible. However, there is ongoing research in the area of improving random access schemes [3], and other areas where the transmission media can be used more efficiently.

The OSI layer which is most directly concerned with the physical communications is the datalink layer, and it is at this layer that there is the most opportunity to optimize the performance of the limited physical resources. This is the layer that the remainder of this paper will concentrate on.

3.0 DATALINK PROTOCOLS

There are a number of services which the datalink layer must provide[4], and from the datalink protocol proposed by INMARSAT for the AMSS the most important of these relevant to mobile systems are:

— priority queuing of messages,
— message segmentation and re-assembly, and
— error detection and re-transmission.

The combination of priorities and message segmentation, breaking long messages up into short packets of uniform length, prevent a long low priority message from hogging the narrow bandwidth datalink and causing excesssive delay to higher priority messages; and it also allows low priority messages to be interrupted by high priority messages and then to resume without the need to be completely re-transmitted. Message segmentation also allows the selective re-transmission of those portions of a message which were corrupted or lost, a situation which is not uncommon due the generally relatively poorer quality of mobile data links.

In the case of AMSS, considerable time and effort has been spent by INMARSAT to develop a datalink protocol which attempts to optimize the physical resources available. The underlying assumption being that with sufficient optimization at the datalink layer, standard ISO protocols such as ISO 8208 (X.25), ISO 8473, and ISO 8073/8602 can be used at the network, internetwork, and transport layers, respectively[2].

However, even within OSI there are interlayer conflicts. For example, the priorities used by the datalink layer are derived from the standard priorities attached to all aeronautical communications by the Radio Regulations of the

CCIR. However, priority is not a standard Quality of Service (QOS) parameter for an OSI datalink layer, in particular, it is not a standard QOS parameter passed by the standard X.25 network protocol. Priorities could be passed as part of the facilities field in the X.25 protocol, but this reduces to some extent the benefits of using a standard network protocol.

A detailed analysis of the datalink protocol for the AMSS will not be performed here but we will concentrate on some system level concerns about the protocol which are shared with land-mobile networks.

4.0 ASPECTS OF DATALINK PROTOCOLS

4.1 Performance

The main performance criteria placed on a data communications system are average message delay, maximum message delay, and reliability, as a function of message priority. The verdict is still out on the what performance is expected to be provided by the AMSS. One conclusion is clear, initial service will be slow because of the low data rates available and the large overheads associated with the data. However, this is acceptable in initial AMSS which serves oceanic and low-density areas [5] where the response time is, for the most part, not critical. However, there have been suggestions that AMSS may be used for air traffic control applications in en route areas, where the required response time is significantly shorter. In this latter case a detailed simulation of the protocols will be necessary to determine if the performance requirements can be met.

4.2 Protocol specification and validation

As with the introduction of any new protocol there is a need to carefully specify and validate the protocol to insure that it performs the required functions, and that it does so efficiently and without error. Protocol validation is not a new problem, but it becomes increasingly important problem as the size and complexity of a communication network grows. Although protocol validation is not a new problem, there does not appear to be a well defined solution.

The approach taken with the AMSS datalink protocol is to specify it using the standardized formal description language SDL. Languages of this type are designed to allow the user to express all the details needed to specify an implementation[6]. In that sense they are not a minimum description of the protocol required to prove correctness and insure interoperability. While these languages are in a sense more complete, it has been our experience that at times they can also be ambiguous.

Among the several approaches that can be taken to the problem of validating a protocol are the following [6]:

— formal verification methods
— implementation of the protocol and testing the implementation,
— simulation of the protocol, and
— in a few cases, it is possible to construct automated design procedures that can be proven to produce correct designs, but this an area of current research.

Formal verification methods refer to the specification of the protocol as a transition system or the equivalent of communicating finite state machines. Then the state space of such system is exhaustively searched for undesired properties such as unmatched communication events, deadlocks, and infinite loops. Although such an analysis can be automated, the state space of such a system may grow so large with the number of messages and the number of machines present that such a verification becomes infeasible, although protocol validation for systems containing up 10^7 states have been proposed in the literature[7]. For example, the X.25 protocol has been partially verified by such an analysis.

By choosing to specify the AMSS datalink protocol in SDL, simulation/implementation appears to be the only the available method of protocol validation, until a transition system description is available. There are several levels on which the protocols can be validated by simulation/implementation. Current work being performed at CRC is validating the performance of the datalink protocols using a minimum subset of the system, one AES and one GES, together with simulated channel conditions. Validation of the protocols in a more complete system is a desirable second step in this process, but it is not clear if this can be done without either performing an in-service validation, or simplifying the simulation and possibly missing some of the protocol interactions.

4.3 Conformance testing

After a protocol has been validated, there remains the problem of insuring that all manufacturers conform to the standard. In theory, one would like to access to all interfaces between the different OSI layers, however, cost and manufacturer's design usually make this impractical. Furthermore, the goal of a standard is interoperability of different manufacturers' equipments rather than explicit specification of an implementation. This has the consequence that there are a limited number of standard test points available for testing protocol conformance. For example in the ICAO standard for AMSS, the only test points available for testing conformance are the interfaces at the network layer and the interface at the physical layer, that is, the signal-in-space as shown in Figure 3. As a consequence, the approach to testing equipment conformance will in general be very different from validating the protocol. In the latter, one can isolate the different OSI layers and validate each independently; while with the former, end-to-end conformance may be the only test available. Since the number of variations can increase exponentially with the number of layers combined, the latter may perform a far greater testing problem, and the only approach would appear to be insertion of the equipment in a simulation test suite where a equipment conformance can be tested over a wide range of standard scenarios[8]. The one saving grace is that this type of test only needs to done once for each implementation, and that there are a limited number of manufacturers.

A potential future problem is the correction of problems found in the protocol after the system has gone into service. There may be questions as to whether implementations must go through a formal test procedure with each upgrade or whether in-service testing can be sufficient. Another consideration is that, because of its size and the number of users, the complete system will not be upgraded simultaneously and thus each upgrade should be backwardly compatible.

5.0 CONCLUSIONS

In this paper several observations have been made about the development of the AMSS communication protocols and their implementation, emphasizing those areas which are relevant to other mobile systems. The OSI approach offers great flexibility and interconnectability to data communications with the penalty of significant protocol overhead. The implication is that in a narrowband system, such as mobile-satellite networks, great care should be taken to minimize these overheads at all layers. At the datalink layer, in particular, there is the opportunity to optimize the use of the available physical resources. However, it is clear that protocol validation should be performed as early as possible in the design process, and that each subsequent design change should be validated; in that way, the cost of correcting errors is minimized.

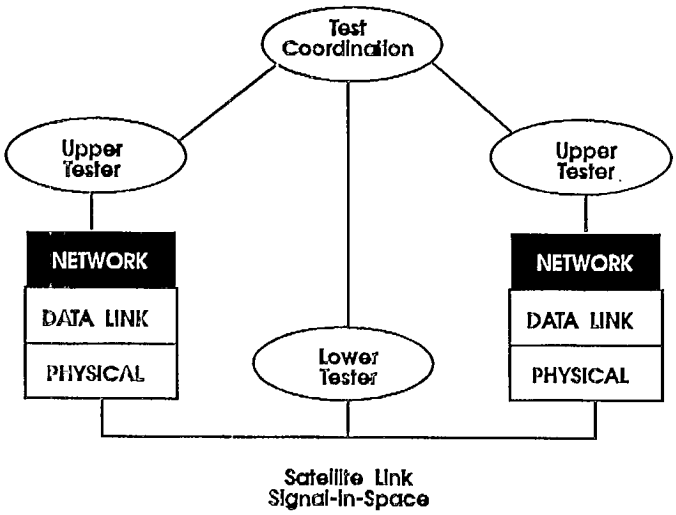


Figure 3. Illustration of test points available for protocol conformance test.

REFERENCES

[1] V.Foose et al, "Preliminary estimates of data link transmission overhead for ADS messages", WP-36, *3rd Panel Meeting of ICAO AMSS*, February 1990.

[2] F.Colliver Jr., "The Aeronautical Telecommunications Network", WP-37, *3rd Panel Meeting of ICAO AMSS*, February 1990.

[3] S.N.Crozier, "Sloppy Slotted ALOHA", *IMSC'90*, Ottawa, Canada, June 1990.

[4] "Reference Model for Open Systems Interconnection for CCITT Applications", *CCITT Rec. X.200*, 1984.

[5] Report of the Special Committee on Future Air Navigation Systems of ICAO, *FANS/4 Report*, Montreal, Canada, May 1988.

[6] B.Pehrson, "Protocol Verification for OSI", *Computer Networks and ISDN Systems*, vol. 18, pp.185-201, 1990.

[7] G.J.Holzmann, "On Limits and Possibilities of Automated Protocol Analysis", *IFIP Int. Conf. on Protocol Spec., Testing and Verif. VII*, Zurich (North-Holland, Amsterdam, 1988).

[8] R.J.Linn, "Conformance Testing for OSI Protocols", *Computer Networks and ISDN Systems*, vol.18, pp.203-219, 1990.

Author Index

Mobile Satellite Users and the ITU 1992 World Administrative Radio Conference: Planning for the Future

Lawrence M. Palmer
NTIA/Dept. of Commerce
Washington D.C. 20230 U.S.A.
Tel 202/377-1304 Fax 202/377-1865

ABSTRACT

The ITU has scheduled a World Administrative Radio Conference to be held in the first quarter of 1992 in Spain to deal with allocation matters. It is anticipated that a review of the mobile-satellite frequency allocations at 1.5/1.6 GHz will be made at this conference.

The frequency bands at 1.5/1.6 GHz have been reviewed by several ITU conferences in the past 20 years. Changes were made in the allocations as the bands lay fallow. Changes were also made at the 1987 Mobile WARC, not because of non-use, but because of the planned activity underway on three continents to proceed with development and implementation of satellite systems to service mobile customers.

The situation has now changed as systems have been licensed in North America. Interim systems are being planned. Full scale operational systems are just around the corner. As we all know, however, domestic allocation decisions have been made in the United States and Canada that are different from the ITU international table of frequency allocations.

Assuming that the 1992 WARC will address the allocations at 1.5/1.6 GHz, what are the options open to users? Do the international frequency spectrum allocations need to be changed? And if so, in what fashion? Will the proponents of change be able to influence their respective national positions to reflect such a need? What are the regulatory issues that will need to be examined in order to permit international systems?

This paper will give a review of the allocations history leading up to the 1987 WARC, and the decisions that were made there. It will examine the different frequency allocations that now exist. Emphasis will be given to showing the options, both from a regulatory and frequency allocation perspective, that may be available to the various parties. The paper then concludes with a presentation of how well the options may succeed if those views are presented to the ITU 1992 WARC.

INTRODUCTION

Last June, the International Telecommunication Union (ITU) Plenipotentiary Conference, meeting in Nice, France, scheduled a number of future ITU radio conferences. Of criticality to the mobile community is the world administrative radio conference (WARC) to be held in Spain in the first quarter of 1992, to deal with frequency allocation matters. The mandate for this radio spectrum conference will be decided by the ITU Administrative Council, that in fact, is meeting in part at the same time as the second International Mobile Satellite Conference here in Ottawa. It is expected that the 1992 world conference will have as part of its mandate, a review of the mobile-satellite frequency allocations at 1.5/1.6 GHz, and the whole of the spectrum from 1 to 3 GHz, particularly in light of Resolution No. 208 from the 1987 Mobile WARC. That resolution recommended that a conference be held to consider revising certain parts of the Table of Frequency Allocations in Article 8 of the ITU Radio Regulations in the approximate range 1-3 and other relevant provisions of the Radio Regulations with a view to providing the necessary spectrum for the mobile and mobile-satellite services.

ALLOCATION HISTORY

The frequency bands for use by the mobile satellite community can be traced back to the 1959 ITU conference. This conference adopted a Recommendation relating to the need to convene an Extraordinary Administrative Radio Conference (EARC), to allocate frequency bands for space radiocommunication purposes. The subsequent 1963 EARC allocated spectrum around 1600 MHz to space telemetering and to the aeronautical mobile-satellite service. The follow-on 1971 WARC modified the 1963 decision, allocating specific frequency bands to the maritime and aeronautical mobile-satellite services. This conference allocated two small sub-bands common to both services subject to prior operational coordination. Additionally, this allocation permitted communication with stations on land.

Since the bands were substantially not being used, the 1979 WARC also modified the allocations. While keeping the two radio mobile services separate, the WARC modified the allocations by slightly adjusting the band limits. The 1979 WARC also provided for joint aeronautical, maritime, and land mobile use of a generic mobile-satellite allocation limited to distress and safety operations.

So why were the allocations modified on a regular basis from the first days of satellite up through the broad allocation conference in 1979? Simply put--the bands lay essentially fallow. There were no services to displace or operations to disrupt. This non-use changed, however, with the implementation of the Marisat system in the early

INTERNATIONAL MOBILE SATELLITE
CONFERENCE (2nd : 1990 : Ottawa)
--Proceedings of the...

MOBILE SATELLITE

CONFERENCE

IMSC'90

Printed in Canada