Public Safety Canada | Sécurité publique Canada

BUILDING A SAFE AND RESILIENT CANADA

Estimating Possible Criminal Organizations from Co-offending Data

November 2012

Canada

Estimating Possible Criminal Organizations from Co-offending Data

By

Uwe Glässer
Simon Fraser University

Mohammad A. Tayebi
Simon Fraser University

Patricia L. Brantingham
Simon Fraser University

Paul J. Brantingham
Simon Fraser University

Prepared for

PUBLIC SAFETY CANADA

The views expressed herein are those of the authors and do not necessarily reflect those of Public Safety Canada.

Report No. 29, 2012

Cat. No.: PS14-7/2012E-PDF
ISBN No.: 978-1-100-21386-6

Contents

Acknowledgements 4
Executive Summary 5
1 Introduction 7
1.1 Analytic Approach 8
2 Related Work 10
3 Crime Dataset 12
3.1 Dataset Overview 12
3.2 Co-offending Network 13
4 Definitions 15
4.1 Crime Data Model 15
4.2 Co-offending Network Model 15
4.3 Co-offender Group Structures 17
5 Criminal Organization Detection 19
5.1 Co-offender Group Characteristics 20
5.2 Hard Constraint Approach 28
5.3 Soft Constraint Approach 33
6 Concluding Remarks 36
References 38

Figures

Figure 1: Crime occurence frequency by crime types 13
Figure 2: Average number of offenders for each crime type 14
Figure 3: Second largest connected component in the co-offending network 14
Figure 4: Tripartite graph derived from the sample dataset 16
Figure 5: Co-offending network extracted from sample event dataset 17
Figure 6: Number of co-offender groups using different minimum clique size 21
Figure 7: Number of observed co-offender groups by different activity thresholds 22
Figure 8: Number of co-offender groups active over a range of time snapshots 23
Figure 9: Number of co-offender groups by the total number of observed time snapshot in which a group was considered active 24
Figure 10: Size of co-offender groups, active co-offender groups 25
Figure 11: Group committed crime frequency 26
Figure 12: Co-offender groups evolution trace 24
Figure 13: Number of shared members for the overlapping groups by the number of pairs of overlapped groups 28
Figure 14: Number of co-offender groups in respect to the proportion of committed serious crimes 30
Figure 15: Number of co-offender groups in respect to the proportion of committed serious crimes 31
Figure 16: Number of co-offender groups in respect to the number of serious crimes they committed 32
Figure 17: Number of active co-offender groups in respect to the number of serious crimes they committed over their observable life cycle 32
Figure 18: Life cycle duration of possible criminal organizations 33
Figure 19: Number of co-offender groups in respect to differrent criminality thresholds 34
Figure 20: Number of active co-offender groups in respect to differrent criminality thresholds 35

Acknowledgements

We thank Austin Lawrence and Michael Saunders, Public Safety Canada, Research and National Coordination, Organized Crime Division, for their constructive feedback and many detailed and insightful comments on draft versions of this report that have substantially contributed to this final version.

Executive Summary

A method of data mining regular police records to identify possible criminal organizations has been developed. Between 2001 and 2006, offending related to 236 possible criminal organizations was reported to RCMP "E" Division, with 39 of the groups being particularly serious.

This study combined computational mathematical analysis, social network analysis methods, and data mining techniques in a unique way to automatically identify traces of possible criminal organizations in operational police records.

Under Canadian law organized crime groups, such as gangs, are termed "criminal organizations." The minimum requirements characterizing a criminal organization are that it consists of three or more people; that there is the commission of a serious criminal offence that can result in a material benefit; and that group offending happen more than once.

The dataset that was used in the study was extracted from the Police Information and Retrieval System of RCMP "E" Division. (RCMP "E" Division covers most of British Columbia, not including some urban areas in the Lower Mainland like Vancouver and the Victoria area.) The massive dataset of more than 4 million records covered all reported offences and all persons associated with a crime, from complaint to charge, from mid-2001 to mid-2006, for the policing jurisdiction.

Using social network analysis methods, the research first identified groups of people that the police-reported data indicated had co-offended with one another. (A "co-offence" is when one or more offenders are associated with a crime incident.) The level of activity, seriousness of criminality, and material benefit associated with the offending for these co-offending groups was then calculated and compared between years. Two different methods were used to determine the level of criminality for a co-offending group. This identified which co-offending groups demonstrated the minimum characteristics of a possible criminal organization and which demonstrated the characteristics of a particularly serious criminal organization. The research then examined how group membership and the structure of these groups changed over time.

The analysis identified more than 18,000 groupings of co-offenders in the crimes that came to police attention. Of these 18,000 groupings, approximately 300 groups were active over a period of time. Of the 300 groups active over a period of time, 236 committed at least one serious offence. These 236 groups represent possible criminal organizations, as they met the minimum quantitative criteria under law for a criminal organization. When only co-offending groups that were active over a period of time which consistently committed crimes that were of above average seriousness were considered, 39 possible criminal organizations of particular seriousness were identified.

Most of the more serious criminal organizations that were identified were also very active over a number of years, indicating their greater stability and intensity of offending compared to the other possible criminal organizations. Similarly, if a group was more criminally active, its members were more likely to have committed serious crimes.

Most of the possible criminal organizations were quite small, with an average core group's size being between six and seven individuals. The particularly serious possible criminal organizations had an even smaller average size, of just less than five. The less serious possible criminal organizations tended to have more peripheral members and a less tightly connected core group.

This type of analysis may eventually provide a useful tool for operational policing in the real-time identification of individuals possibly associated with a criminal organization, as well as serve as an alternative source of information in intelligence gathering and verification.

Intelligence and further criminal analysis is required to properly use this type of information in the investigation of, and reporting on, organized crime because there are a number of caveats regarding these possible criminal organizations that have been identified. Further work would be required to determine if the possible criminal organizations identified were component parts of larger organized crime groups. It is possible that not all individuals in a criminal organization are included in the identified networks because only police-reported crime information was analyzed. Individuals operating in the background or who are more able to escape police interventions, who may be more likely to direct the activities of others, would not be captured with this type of methodology.

1 Introduction

This report explores a new computational approach used by the Institute of Canadian Urban Research Studies (ICURS) at Simon Fraser University to detect traces of criminal organizations in large crime datasets and characterize their criminal activities over longer time periods. The applied methodology and analytic approach use co-offending network analysis, combining social network analysis methods and data mining techniques. Data mining, with appropriate research, can become a complementary decision support tool to be used both by operational criminal justice analysts and by policy makers in framing enforcement operations and in framing crime control policy.

Co-offending networks constitute a widespread form of social network that play a central role in crime investigations, and have broad implications for the study of crime and criminal justice (McGloin and Nguyen 2011). In fact, "understanding co-offending is central to understanding the etiology of crime and the effects of intervention strategies" (Reiss 1988). With increasing academic and societal awareness of the importance of social networks, law enforcement and intelligence agencies have long realized the potential of co-offending network analysis to provide a better understanding of organized crime and to serve as an instrument in evidence-based policy development aimed at crime suppression and prevention strategies.

The concepts presented here are new. Most current offender network analysis either starts with small datasets or develops the network from expert identification of criminal organizations built by linking to an initially small group of individuals. Expert identified criminal organizations usually have large amounts of information about individuals, but the total number of people in the datasets is limited by the necessary scale of effort needed to identify and add separate pieces of information. The goal of this work is to illustrate the potential effectiveness of computational co-offending network analysis as a practical means for extracting information about possible criminal organizations from large real-life crime datasets, specifically police-reported crime data. It should be noted that co-offending network analysis is a user tool and could be applied to any dataset that identifies links between people.

The goal of this research is to test a data mining tool for computationally intensive co-offender network analysis. A key aspect is the design of a coherent and consistent framework for defining the problem scope and analytic methods to determine structures in co-offending networks that correspond to the prescribed characteristics of "criminal organization" in the *Criminal Code of Canada* as described in the International Centre for the Prevention of Crime (ICPC) 2010 report "International Report on Crime Prevention and Community Safety: Trends and Perspectives" (ICPC 2010, 49):

> In Canada a criminal organization is a group, however organized that: (a) is composed of three or more persons in or outside Canada; and (b) has as one of its main purposes or main activities the facilitation or commission of one or more serious offences, that, if committed, would likely result in the direct or indirect receipt of a material benefit, including a financial benefit, by the group or by any one of the persons who constitute the group. The definition further specifies that it excludes a group of three or more

> persons that has formed randomly for the immediate commission of a single offence. Section 467.1(1) of the *Criminal Code*.

Building on existing work by members of the research team, Tayebi and Glässer (2011), Tayebi, Glässer and Brantingham (2011) and Brantingham et al. (2011), we follow a constructive approach was followed that uses mathematical models of crime data and criminal activity as its underlying semantic foundation. This way, the meaning of central concepts related to criminal organization and organized crime can be gradually refined and transformed into systematic computational representations as required for encoding this meaning in mining algorithms and social network analysis methods. The algorithm used in this study identifies all possible co-offending groups, from young teens to older co-offending chronic offenders to serious active co-offenders. Future analysis with additional information should make it possible to narrow the scale of the co-offending networks, moving closer to known criminal organizations.

Special emphasis is put on robustness and scalability in the development of algorithmic methods to ensure their potential applicability and reusability to more comprehensive datasets outside of this project. All descriptive studies for this project use an anonymized crime dataset comprising five years of RCMP "E" Division data from a dataset of official recorded crime in the Province of British Columbia from August 1, 2001 to July 31, 2006. This data was retrieved from the RCMP's Police Information Retrieval System (PIRS), a large database system keeping information for the cities, towns and regions of the Province of British Columbia which are policed by the RCMP. This dataset is known as the "BC crime dataset."

This research is exploratory in nature and aims at producing results that may provide new insights into the size and characteristics of organized crime. To this end, our report presents descriptions of possible serious, active criminal organizations derived from analysis of a much larger set of co-offending networks and describes the volume and intensity of offences committed by these identified networks. Our analysis towards detection of possible criminal organizations starts with extracting the co-offending network from more than four million crime records, which results in a co-offending network with more than 150,000 offenders. In the next step, we detect more than 18,000 co-offender groups. From this set we identify 313 active offender groups that had criminal activity for longer time periods (on average more than 700 days). We study offences committed by active offender groups in order to decide if they could be considered possible criminal organizations. For this purpose we define the concept of group criminality and compute this measure using two different approaches. Based on the findings resulting from our analysis of the five years of crime data, we conclude that there are from 39 to 236 groups of co-offenders identifiable in the BC crime dataset that are both active and serious and might thus be considered possible criminal organizations according to the definition in the *Criminal Code of Canada*. The wide variability in the estimated number of possible criminal organization groups observable in police-recorded crime is related to various analytic approaches that can be taken to interpret the dataset, which will be explored in this paper.

1.1 Analytic Approach

Conceptually the analytic approach applied in this study to analyse the BC crime dataset comprises three major phases: data preparation, co-offending network extraction, and co-offending network analysis. This outline describes the overall logical organization of the approach in common terms, explaining the basic idea rather than technical details.

After an initial data preparation phase covering all activities prior to analyzing the data, such as data selection and data pre-processing, further explained in Brantingham et al. (2011), a comprehensive co-offending network is extracted in several steps from the BC crime dataset using the network extraction method described in Tayebi and Glässer (2011). A sample co-offending network component from this network is illustrated in Section 3.2 of this report.

For the purpose of detecting possible criminal organizations in co-offending networks and for analysing how those organizations and, more generally, co-offender groups evolve over time, regarding aspects like group membership and co-offending behaviour, we first partition the dataset, and also the extracted co-offending network, into five consecutive time intervals, called *snapshots*, each of which spans a time period of 12 months. The implicit assumption is that over a time period of 12 months a co-offending network is mainly characterized by static properties, whereas for longer time periods the dynamics of the network plays a more important factor.[1]

Subsequently, for each snapshot of a co-offending network, the following computational tasks are carried out in consecutive steps: *(1)* Discover co-offender groups in the current co-offending network; *(2)* Compute the activity and criminality of the co-offender groups detected in the time period between the current network and the previous network based on the offences that were committed by members of these groups; *(3)* Assess the material benefit associated with each of the offences considered in Step 2; *(4)* Identify those groups that qualify as possible criminal organizations; *(5)* Update the group's evolutionary trace (how group membership and co-offending network structure has evolved) for the current time period.

In the first step of the above algorithmic method, co-offender groups are built up from k-cliques[2]. It is assumed that each group is composed of either a single k-clique or several adjacent k-cliques, sharing at least $k-1$ nodes with each other. Since any co-offender group that is considered a criminal organization under Canadian law must have at least three members, we assume k=3. Each clique uniquely belongs to one community, but cliques within different communities may share members. Hence, we may have overlapping groups. For each co-offender group, all members identified as described above together form the *kernel* of that group. Kernels refer to the main members of a co-offender group and are completely involved in the group activities. In the second step, neighbour nodes connected directly to the kernels are added to the co-offender groups. These nodes are called *peripheries*.

---

[1] One may argue that the choice of snapshots considered here, both with respect to the time frame and also how they partiton the dataset, is only one possible option, and shorter or longer time periods (even snapshots with variable duration depending on observed events, a.k.a. "soft clustering") may be considered as well. This can be done by extending the work presented here in a straightforward way. However, the same analytic approach could be applied to any other partitioning of the dataset, with increasing computational effort for more fine granular snapshots.

[2] A k-clique is a complete subgraph of the co-offending network of size k, where k refers to the number of nodes. We are using the mathematical terminology of graph theory while in some social network literature *nodes* are sometimes described as persons or actors and *edges* are sometimes described as ties or connections. See Borgatti, et al., 2009; Wasserman & Faust, 1994.

Activity and criminality of a co-offender group are two key characteristics toward understanding group structure. Activity of a co-offender group is equal to the division of the number of co-offences committed by members of the group at the current time snapshot by the number of co-offences committed by the group at the previous time snapshot. Criminality of a co-offender group is equal to the division of the summation of seriousness of all the crimes of a group by the number of these crime incidents. Activity and criminality of each detected co-offender group are analysed to determine whether or not a group is a possible criminal organization, where we consider a co-offender group a possible criminal organization, if its activity and criminality exceed defined thresholds. This distinction is based on the working definition of the concept of criminal organization according to which a criminal group is considered a possible criminal organization, if the group activity is continuous and its members are involved in one or more serious crimes that imply direct or indirect receipt of a material benefit. In practice, we do not expect to observe frequent collaboration of each pair of members of a possible criminal organization in any given time period.

In general, arrest data can provide only partial knowledge of real-world crime and does not cover every instance and detail of criminal activities. This is the idea behind the definition of activity measures. Identifying the minimum threshold for activity and criminality measures experimentally is a challenging problem that ultimately exceeds the scope of this project as it requires even larger data sets covering longer time periods and more extensive experiments. See Section 5 for more details.

For detecting more active and dangerous possible criminal organizations, one can gradually increase the activity and criminally measures. We expect to determine meaningful threshold values for activity and criminality based on comprehensive experiments exceeding what can be done within the scope of a single study.

2 Related Work

With academic and societal awareness of the importance of social networks increasing, law enforcement agencies and intelligence agencies have come to realize the value of detailed knowledge of criminal or co-offending networks. Groups and organizations operating within such networks to engage in conspiracies, terrorist activities and crimes like drug trafficking typically operate in a concealed fashion, trying to hide their illegal activities and often their associates or associations. In analysing such activities, investigations do not only focus on individual suspects but also attempt to uncover criminal groups. Arguably, it is desirable to identify criminal networks in large volume data sources readily available to investigators, such as police arrest data and court data, and study these data using social network analysis methods. Social network analysis can also provide useful information about individuals as well. For instance, investigators may identify key players and subject them to closer inspection. Knowledge about co-offending network structures provides a basis for law enforcement and intelligence agencies to make strategic or tactical decisions. This section briefly reviews related studies in co-offending network analysis in general, and then home in on research relevant to locating central actors in co-offending networks.

Several empirical studies that use social network analysis methods to analyse co-offending or terrorist networks focus on stability of associations in such networks. Morselli (2009) offers a thoughtful general insight into 'criminal organizational systems' from a criminal network perspective and applies social network analysis to a number of case studies of criminal groups and organizations. Reiss (1988) concludes that the majority of co-offending groups are unstable, and their relationships are short-lived. This is corroborated by McGloin et al. (2008), who show that there is some stability in co-offending relationships over time for frequent offenders, but delinquents do not, in general, tend to reuse co-offenders. Reiss and Farrington (1991) also found that co-offenders have many different partners, and are unlikely to commit crimes with the same individuals over time. However, Reiss (1988) also states that high frequency offenders are "active recruiters to delinquent groups and can be important targets for law enforcement." It should be noted that the findings of these works are based on very small datasets: 205 individuals in (Reiss and Farrington, 1991), and 5600 individuals in (McGloin et al. 2008), and may therefore not be representative.

Confronted with a bewildering diversity of characteristics referred to in existing definitions of organized crime and criminal organizations, the conceptual model itself appears not clearly depicted in the literature—at least not for the purpose considered here. Looking for a quantitative definition, in an attempt to measure organized crime, van der Heijden (1996) proposes a number of common characteristics:

1. Collaboration of more than two people;[3]
2. Commission of serious criminal offences (suspected);
3. Determined by the pursuit of profit and/or power;
4. Each having their own appointed tasks;
5. For a prolonged or indefinite period of time;
6. Using some form of discipline and control;
7. Operating across borders;
8. Using violence or other means suitable for intimidation;
9. Using commercial or businesslike structures;
10. Engaged in money laundering;
11. Exerting influence on politics, the media, public administration, judicial authorities, or economy.

According to van der Heijden (1996), for any criminal group to be categorized as an criminal organization it needs to have at least six of the above characteristics, where Items 1, 2, and 3 are obligatory, thus requiring the addition of three more characteristics.

[3] In contrast, the *Criminal Code of Canada* requires a minimum of three persons to form a criminal organization.

A major study in the Netherlands (Fijnaut et al. 1998) mentions great variations in collaborative forms of organized crime and concludes that "the frameworks need not necessarily exhibit the hierarchical structure or meticulous division of labour often attributed to mafia syndicates. Intersections of social networks with a rudimentary division of labour have also been included as groups in the sub-report on the role of Dutch criminal groups, where they are referred to as cliques. As is demonstrated . . . there can be sizable differences in the cooperation patterns within these cliques and between the cliques and larger networks of people they work with on an incidental basis."

An impressive collection of definitions of organized crime specific for various countries, comprising more than individual 150 entries in total, has been gathered by von Lampe (2012). In addition, this collection also includes comments on how to define organized crime, and definitions by prominent individuals and government agencies, for instance, such as the Federal Bureau of Investigation (FBI). Not included though are definitions of the term 'criminal organization.' Given the abstract nature and informal language of these definitions, it is not clear at all how and to what extend one may utilize this resource for defining organized crime in precise computational and/or mathematical terms.

3 Crime Dataset

As a result of a research memorandum of understanding between ICURS and RCMP "E" Division and the Ministry of the Attorney General of British Columbia, five years of historic real-world crime data was made available for research purposes. This section provides some basic characteristics of the BC crime dataset and the co-offending network extracted from the dataset.

3.1 Dataset Overview

The BC crime dataset contains information about all reported offences (translating into approximately 4.4 million records) and all persons, including offenders, victims, witnesses, among others, associated with a crime from complainant to charged. In total, there are 39 different subject (person) groups. For any given crime incident, every related subject has up to three different status fields, stating the subject's "role" in this incident.

In total, about 1,000 different crime categories (coded) are defined in the dataset, each referring to a specific type of criminal event or violation of a code with a penalty. For extracting the co-offending network, we consider all of the codes (except for traffic related offences). In case of offences with more than one offender, only about 100 offence types have an occurrence frequency greater than 100, and only 30 offence types have an occurrence frequency greater than 1,000. Figure 1 and Figure 2 below respectively show the occurrence frequency and average number of involved co-offenders for each type of offence, where the type of offence is uniquely identified by a number.

In our experiments, we consider only those subjects from the dataset that appear as suspect, charge recommended, chargeable or charged. Being in one of these categories means that information was available to classify an individual as more than someone against whom a complaint was made. For the purposes of this experiment these were called "offenders." The data did not contain information about prosecutions and convictions.

3.2 Co-offending Network

The extracted co-offending network comprises approximately 150,000 nodes (or individuals) and approximately 600,000 edges (or connections between the individuals). The average node degree[4] is four and the maximum degree is 525; that is, the average number of connections with other offenders with whom an individual collaborates ranges from a minimum of four to a maximum of 525. About 50% of all the nodes have degree one, meaning these offenders have committed co-offences with only one other offender during the time period considered in this analysis. The largest connected network component (that is, a subset of nodes in which any two nodes are connected by a path consisting of a sequence of edges connecting nodes from this subset) links about 18% of all the nodes together, which is fairly big for this kind of network. Figure 3 illustrates an example of a connected component of a co-offending network.

For the experiments, we have divided the dataset into five chronological *snapshots*, each of them representing a time period of 12 months. Excluding all traffic related ones, the total number of reported offences in each of the five time periods is respectively: 151,866; 475,288; 521,727; 533,404; and 954,029. Considering only offences with more than one offender reduces these numbers to: 9,943; 18,819; 18,350; 16,939; and 20,111. For crime incidents with multiple offenders, we extract the respective co-offending network for each of the five snapshots. Figure 3 shows an illustrative example of the second largest connected component in the co-offending network derived from the BC crime dataset. This component comprises 137 offender nodes, female ones are marked by bigger nodes, and male ones are marked by smaller nodes.

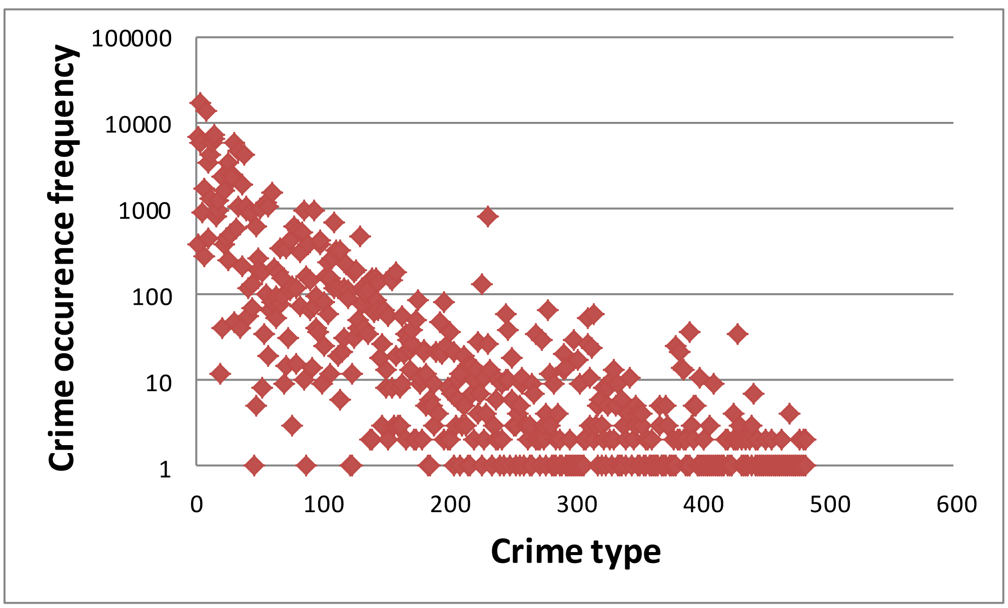

Figure 1: Crime occurence frequency by crime types

[4] In graph theory, the degree of a node (or vertex) is the number of edges incident to the node.

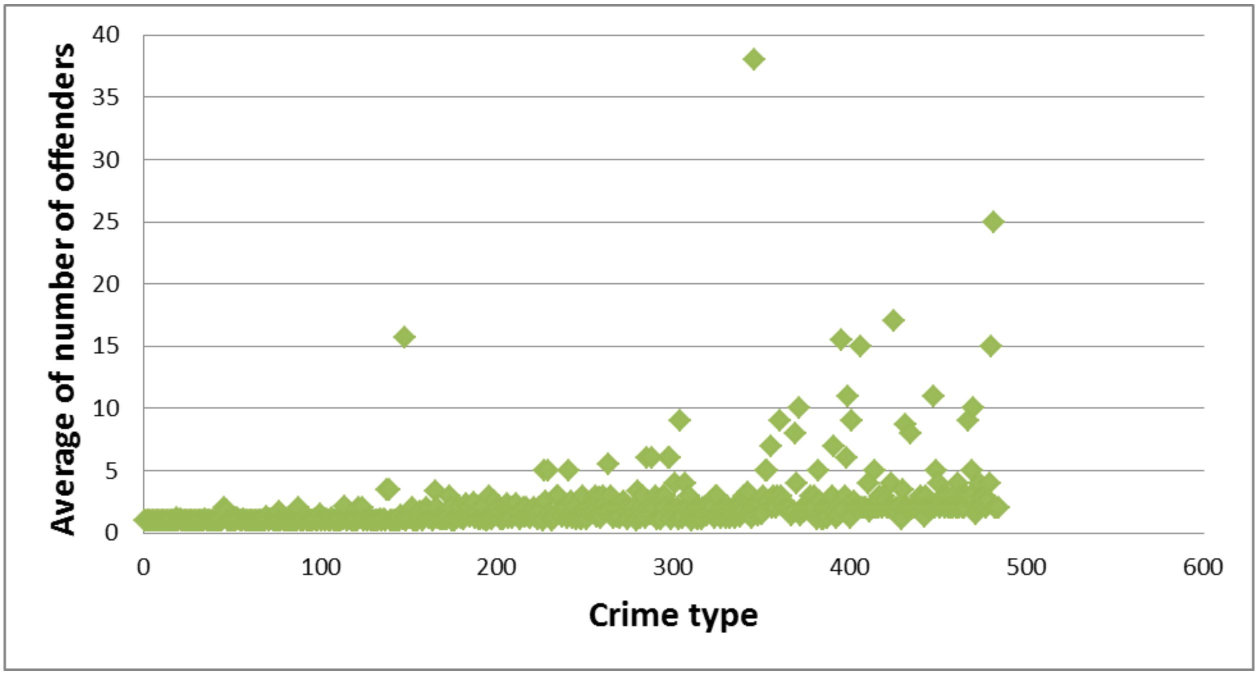

Figure 2: Average number of offenders for each crime type

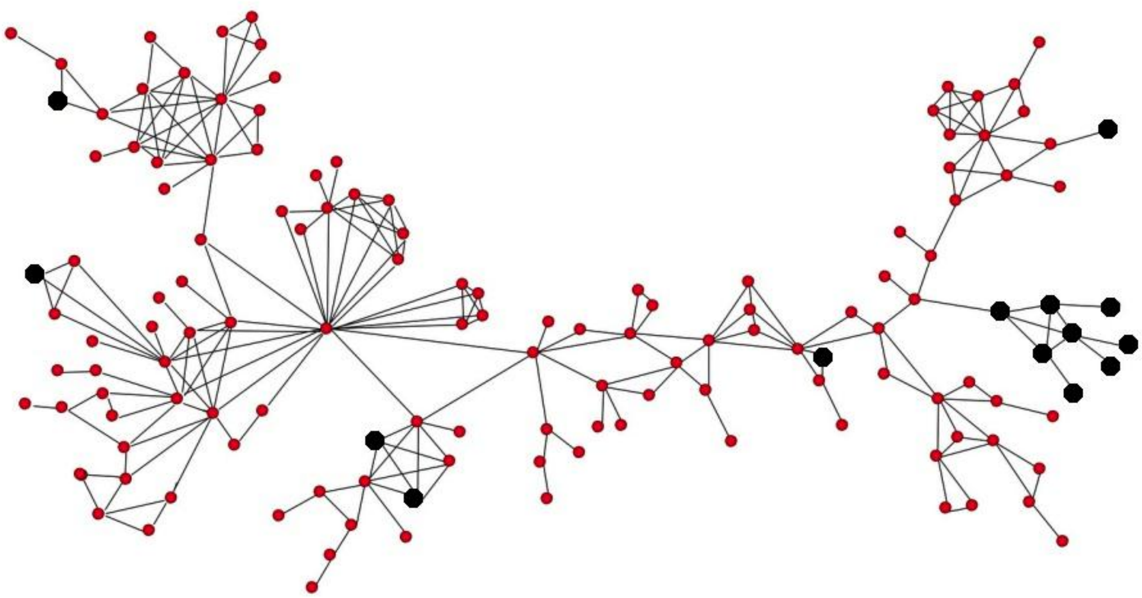

Figure 3: Second largest connected component in the co-offending network

4 Definitions

This section introduces the basic concepts and definitions used in the following sections. We start with a formal *crime data model* serving as semantic framework for defining in a concise and unambiguous way how a co-offending network is derived from a crime dataset and how we analyze this network for the purpose of identifying criminal network structures and their constituent entities. We further define the concept of *offender group* as a basic substructure of a co-offending network and describe the analytic method for tracing how offender groups evolve over their "life cycle." Finally, we define the concept of *criminal organization* in terms of characteristics that discriminate possible criminal organizations from regular co-offender groups. The rational for the applied characteristics is to be in line with the definition for criminal organization in the *Criminal Code of Canada*.

4.1 Crime Data Model

We model crime data in the form of an attributed tripartite *hypergraph* $H(N,E)$ with node set N and a set of hyperedges E. The nodes are partitioned into three subsets, $A=\{a_1,a_2,\ldots,a_q\}$, $I=\{i_1,i_2,\ldots,i_r\}$ and $R=\{r_1,r_2,\ldots,r_s\}$, representing *actors*, such as offenders, victims, witnesses, suspects and bystanders; crime *incidents*, each referring to a reported offence of a certain crime type; and *resources* used in committing an offence,[5] such as a mobile phone, a tool, a vehicle, a weapon or a bank account. A hyperedge e of E is a non-empty subset of nodes $\{n_1,n_2,\ldots,n_p\}\subseteq N$ such that the following three conditions hold: $|e\cap I|=1$, $|e\cap A|\geq 1$ and $|e\cap R|\geq 1$.

4.2 Co-offending Network Model

A co-offending network is derived in several steps from a crime data model, as explained in (Tayebi and Glässer 2011), and comprises one or more connected components consisting of co-offender nodes, where the nodes are connected for all co-offenders u,v who have committed crimes together. The number of co-offences committed by co-offenders u,v is indicated by a value *strength* associated with link $l=\{u,v\}$, where $strength(l)\in N$. Assuming k co-offenders and m crime events (k,m>1), we define a $k\times m$ matrix M, such that $m_{uv}=1$, if co-offender o_u is involved in event i_v, and "0" otherwise. A co-offending network is thus a $k\times k$ matrix $N=MM^T$.

$$n_{u,v}=\sum_{x=1}^{k} n_{ux}n_{xv}$$

Starting with the general co-offending network model as a basis, one can derive more specific co-offending network substructures in a straightforward way simply by restricting the type of offences being considered. For instance, one may only consider drug trafficking incidents.

---

[5] Resources do often provide essential clues in criminal investigations. For simplicity, we assume here that R includes a distinguished element *nil* referring to situations in which no specific resource can be identified.

For the purpose of illustrating the construction of a co-offending network from a crime dataset, we use a fictitious dataset as an example. Table 1 lists an example of the type of relationships that should be found in the analysis, while Figure 4 shows the tripartite graph extracted from this data. The data used in this example is completely unrelated to the real crime dataset being analysed.

Offenders, events and resources nodes are respectively marked in the colours brown, blue and peach.

Table 1: Sample event dataset

Event	**Offenders**	**Event Location**	**Event Time**	**Resources**	**Crime Type**
E_1	A,B,C	L_1	2/3/2006	R_1	P_4
E_2	C,D	L_2	4/5/2006	R_2	P_1
E_3	A,C	L_3	5/16/2006		P_2
E_4	A,B,C,E	L_4	6/11/2006	R_1	P_4
E_5	F	L_5	9/8/2006		P_3

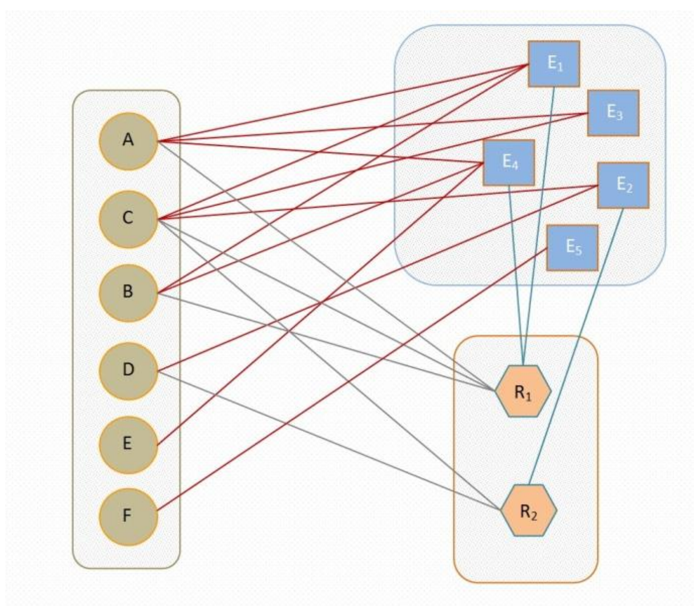

Figure 4: Tripartite graph derived from the sample dataset

A co-offending network consists of one or more connected components of offenders who have committed offences together. More specifically, a co-offending network is viewed as an affiliation network in which there is a link between each pair of offenders involved in the same crime event.

This aspect is illustrated in Figure 5 for the co-offending network of the sample data from Table 1. For instance, from event e_1, which has 3 offenders *A*, *B* and *C*, three co-offending links {*A*,*B*}, {*A*,*C*} and {*B*,*C*} are extracted. We do not consider singleton offenders such as *F* as co-offending network member. The singleton nodes are representing offenders who have never made a connection with another offender in the co-offending network.

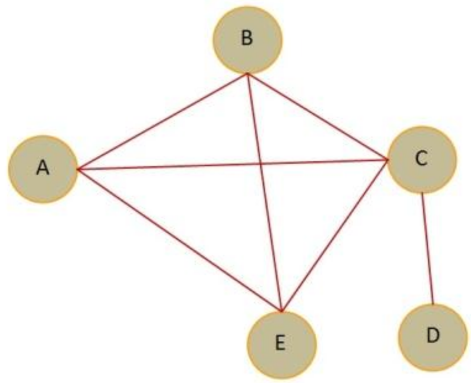

Figure 5: Co-offending network extracted from sample event dataset

4.3 Co-offender Group Structures

We define here in a step-by-step manner the basic concepts for distinguishing different types of offender group structures and their characteristic attributes as substructures of a co-offending network.

Co-offender Group A co-offender group comprises three or more co-offenders who collaborate in committing offences. This does not mean that each and every group member participates (in an active role) in all offences committed. These groups are not necessarily formed as the result of a predefined plan and also they need not be active continuously. Their members have generally local clustering within larger loosely connected networks, thus constituting a small group with varying degrees of connection to other larger groups. In our model, $C_1^t, C_2^t, \ldots, C_n^t$ refer to *n* co-offender groups in the co-offending network at time period *t*.

Group Activity For co-offender group C_i^t, the *activity* $\tau_i^{t_1,t_2}$ indicates how frequently members of this group have committed offences during time period t_1 compared to time period t_2.

Group Criminality Group criminality Φ_i^t represents a measure for the degree of *seriousness* of offences committed by members of co-offender group C_i^t during time period *t*.

Active Co-offender Group Active co-offender groups have a history of continued criminal activity over some longer time period. $A_i^{t_1,t_2}$ represents an active co-offender group that is active at time period t_1 and is still active at time period t_2.

Serious Co-offender Group A co-offender group whose overall criminal activity at time period t shows a high degree of serious criminal offences is called a *serious co-offender group* and is referred to by S_i^t.

Criminal Organization In theory, the two concepts of criminal organization and co-offender group differ in at least three basic aspects: 1) group scale and motivation; 2) time interval of collaboration; and 3) type of criminal activity. In practice, however, the distinction between a criminal organization and a co-offender group is not always clear-cut and can be challenging. To qualify as an organized criminal group, a necessary (but not sufficient) condition is the commission of a serious offence motivated by material benefit. While the meaning of 'serious offence' can be clearly defined in terms of offences classified as indictable or hybrid offence or statute serious offence in the *Criminal Code / Controlled Drugs and Substances Act,* the meaning of material benefit may be interpreted in a narrow or in a broader sense. In our model, $O_1^t, O_2^t, \ldots, O_m^t$ refer to m criminal organizations in the co-offending network at time period t.

Criminal Organization Evolution Trace Criminal organizations, similar to any other form of community, typically evolve over time. A criminal organization may grow by admitting new members, shrink by losing members, split into two or more groups, or a new group may form by merging two or more existing groups. Therefore, we need to devise a model that can express these dynamic aspects of how a possible criminal organization can evolve over a number of consecutive time periods. The model needs to determine which group at previous time has evolved into which group at current time. Five phenomena can be observed in a single snapshot: a group may survive, split, merge, emerge or cease. An evolution trace $E(O_a^t)$ is a sequence $O_a^t, O_{a_1}^{t+1}, O_{a_2}^{t+2}, \ldots, O_{a_n}^{t+n}$ of related criminal organizations over n consecutive time periods that shows the dynamic transformation, or evolution, of the criminal organization O_a^t since time period t.

As noted, five phenomena can be observed in a single snapshot: a group may survive, split, merge, emerge or cease. For this purpose we use a matching function

$$match: G \times 2^G \rightarrow G$$

where G denotes a set of groups and 2^G denotes the powerset of G. The meaning of this function is that, for a given possible criminal organization O_i^t and set of possible criminal organizations G^{t+1}, we let $match(O_i^t, G)$ yield the group O_i^{t+1}, such that this group has the largest intersection with O_i^t, where this intersection is above a given threshold λ. Formally this is defined as

$$match(O_i^t, G) = O_j^{t+1} \text{ with } \forall\ O_k^{t+1}: O_k^{t+1} \in G \ \wedge \mathrm{overlap}(O_i^t, O_j^{t+1}) \geq \mathrm{overlap}(O_i^t, O_k^{t+1}) \ \wedge \ \mathrm{overlap}(O_i^t, O_j^{t+1}) > \lambda$$

where for two possible criminal organizations $O,O' \in G$ we define

$$overlap(O,O') = min\left(\frac{|O \cap O'|}{|O|}, \frac{|O \cap O'|}{|O'|}\right)$$

which, given two different groups, returns the minimum of the division of the intersection of these two groups by the size of each of them respectively. Using the matching function, we apply the following rules for tracking the evolution of criminal organizations:

- O_i^t *survives* in the next time period as O_j^{t+1}, if $O_j^{t+1} = match(O_i^t, G^{t+1})$ and for each $O_k^t \neq O_i^t$, $O_j^{t+1} \neq match(O_k^t, G^{t+1})$. In other words, a criminal organization *survives* in the next time period, if there exists a group that has sufficient intersection with the group in question.
- An criminal organization *splits* into two or more separate groups, if there is sufficient overlap between each of the resulting groups and the original one, and also the union of the resulting groups exceeds a predefined minimum threshold. Formally, we describe this phenomenon as: O_i^t splits into groups $O_1^{t+1}, O_2^{t+1}, \ldots, O_n^{t+1}$, if there is enough overlap between each of these splitted groups and O_i^t, and also $(O_1^{t+1} \cup O_2^{t+1} \cup \ldots \cup O_n^{t+1}) \cap O_i^t$ is above a predefined minimum threshold.
- O_i^t *merges* with some other group into O_j^{t+1}, if $O_j^{t+1} = match(O_i^t, G^{t+1})$ and $\exists\ O_k^t \neq O_i^t$: $O_j^{t+1} = match(O_k^t, G^{t+1})$. This means, a criminal organization *merges* with another group into a new criminal organization, if the original group and at least one other group each map to the same group in the next time period.
- O_i^t *ceases*, if none of the above scenarios happened.
- O_i^t *emerges*, if $\forall\ O_i^{t-1} : O_j^t \neq match(O_i^{t-1}, G^t)$, meaning that, if a group does not map to any group existing in the previous time period, it is considered a newly emerged group.

5 Criminal Organization Detection

This section describes the experimental evaluation of the proposed analytic method for criminal organization detection on the BC crime dataset. A basic principle in the characterization of "criminal organization" defined in the *Criminal Code* is "the facilitation or commission of one or more serious offences, that, if committed, would likely result in the direct or indirect receipt of a material benefit ... by the group or by any one of the persons who constitute the group." Consequently, the specific type of committed offences plays a crucial role in determining whether or not an identified crime group is considered a possible criminal organization. In abstract computational terms this aspect effectively constrains the search space to be analyzed by the applied data mining algorithms, such that all offences that do not qualify as a serious offence with associated material benefit are disregarded. In general, knowledge discovery in databases and data mining is ultimately restricted to the information encoded in the underlying datasets. This experiment was run on historic anonymized data with limitations in quality (referring to attributes such as completeness, consistency and noise) as in all historic police datasets where verification against the 'grand truth' is often not feasible or even possible so one has to work with the data as it is. In practice, data mining algorithms, when used with current and enhanced crime datasets, provide decision support for exploring hand-entered data.

The first step of the analysis, referred to as the "hard constraint approach," forms the core of the project since it most closely matches the *Criminal Code* definition of criminal organization. In the hard constraint approach all crimes are categorized into two classes: serious crimes with material benefit, and non-serious crimes or serious crimes that do not appear to provide the offender group with a material benefit. The second approach, called the "soft constraint approach" explores how the number and characteristics of possible criminal organizations identified by data mining is affected when the definition of seriousness and material benefit is allowed to vary and when time "snapshots" are used to measure the stability of networks. The soft constraint approach is exploratory. Future research will explore other ways to vary rules used in interpreting police-defined crime categories and in adding additional non-offence data that show supportive links to identified criminal organizations. It should be noted again as it was on p. 3 that the addition of non-offence data, such as age, is the type of information from other data sets that could be used to bring this type of tool more in line with the functional meaning of organized crime in operations or in policy development.

In both approaches, the actual analysis (as encoded in the applied algorithms) is the same for the co-offending network, co-offender group detection, and active co-offender group detection. In each of the approaches, we use a corresponding crime seriousness index to calculate the criminality of each group for all offences to be included (as determined by the threshold) in which at least one member of the group was involved. Based on the criteria defined, for each of the approaches, we detect serious co-offender groups. Finally, in the last step, co-offender groups that are active and serious are identified as possible criminal organizations.

5.1 Co-offender Group Characteristics

In this part we explore the characteristics of the extracted co-offender groups and active co-offender groups. The crime data is partitioned into the following five time snapshots, each of which represents a 12 month time interval: mid-2001 to mid-2002; mid-2002 to mid-2003; mid-2003 to mid-2004; mid-2004 to mid-2005; and mid-2005 to mid-2006.

Figure 6 shows the number of co-offender groups for different clique sizes *k*. As expected, the number of co-offender groups decreases with increasing the clique size. All experiments discussed below are based on clique size *k* equal or greater than 3, the minimum group size required for a criminal organization under the *Criminal Code*.

As a simple example, we assume a co-offender group C_1^2 consisting of seven members with 10 co-offending links detected at time snapshot $t = 2$. We follow the behaviour of this group at time snapshots $t = 3$, $t = 4$ and $t = 5$, and further assume that in the snapshots 3-5 we have observed

2, 5, and 9 co-offences respectively among members of group C_1^2. Then, the computed activity for these snapshots would be $\tau_1^{2,3} = 0.2$, $\tau_1^{2,4} = 0.5$ and $\tau_1^{2,5} = 0.9$. Activity measures the relative activity of a co-offending group as observed in a particular time snapshot, compared to the number of distinct co-offences of that same co-offending group in the snapshot this group was first detected. Threshold values are introduced as a flexible means to adjust the level of observed activity above which co-offending behaviour is actually taken into account for the analysis. Implicitly, higher levels of activity observed over a number of consecutive time snapshots suggest a higher degree of stability of a co-offending group. Figure 7 illustrates the number of co-offender groups in each time snapshot for different *activity thresholds* α, where k equals the number of people in the co-offender group. An *activity threshold α* describes the percentage of the structure of the co-offending group that remains unchanged between time snapshots.

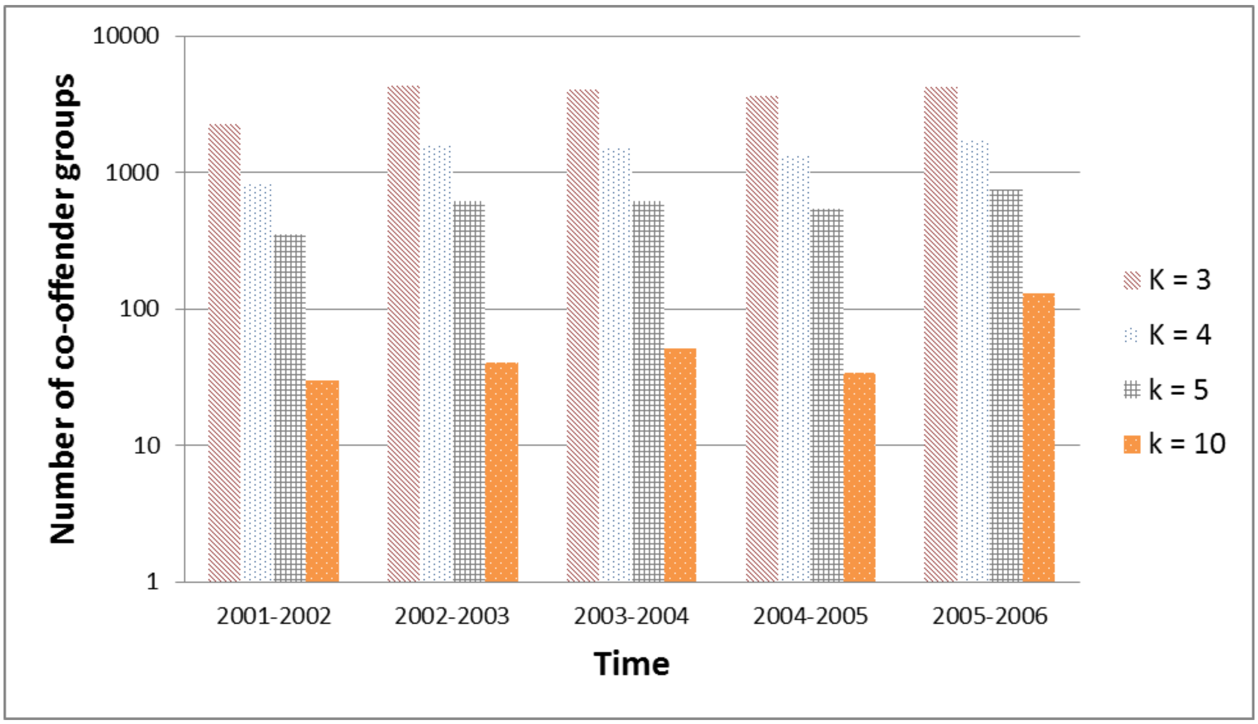

Figure 6: Number of co-offender groups using different minimum clique size

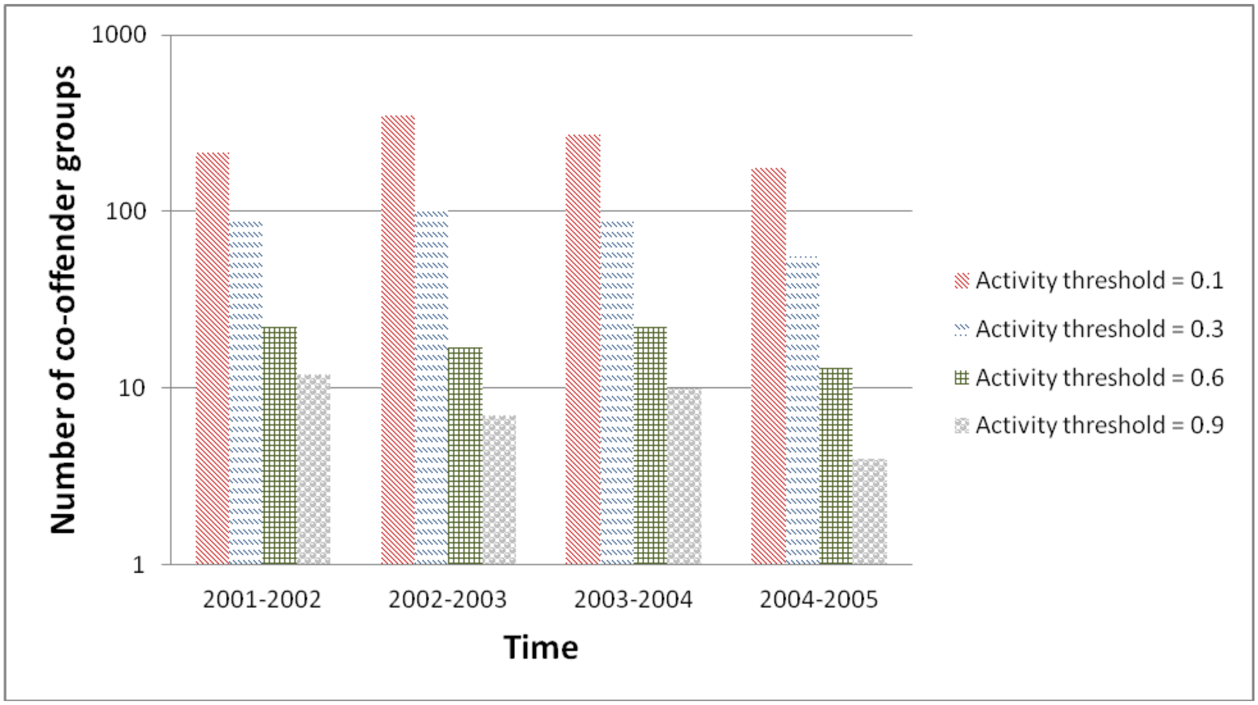

Figure 7: Number of observed co-offender groups by different activity thresholds

Even with 60% of the structure of the co-offending group remaining intact between snapshots (α=0.6), about one percent of all the co-offender groups remain in the list of active co-offender groups, meaning that some co-offender groups keep their collaboration intact and unchanged over longer time periods.

Active co-offender groups can be further characterized as continuously active over several consecutive snapshots or as sporadic, with their activity occurring at irregular time intervals with inactive intermediate snapshots.[6] In assessing the continuity of activities of co-offender groups, we study their criminal activity over several time snapshots, applying activity thresholds. For instance, assume the following scenario. Group C_1^3 was detected at snapshot $t = 3$, and this group has four members and four co-offending links. No activities were observed in time snapshot $t = 4$. At snapshot $t = 5$, three co-offences were observed. For any activity threshold equal to or less than 0.75, we can consider the time snapshot difference for group C_1^3 equal to two.

[6] We are aware of the possibility that apparent sporadic activity could result from group activities not coming to the attention of the police during a given time period rather than from actual lack of criminal activity during that time period.

Figure 8 shows the number of co-offender groups observed over time periods with one, two, three, and four years difference. The important point here is that with increasing time difference the number of observed groups decreases exponentially. Even with very low values for α only few groups can be observed over four snapshots, and with high values for α no group can be observed over four snapshots. However, one can also see that from one snapshot to the next, one continued group activity is more common, even for higher values of α. This finding supports the theory of short-time collaborations of most co-offender groups.

According to Albanese (2004), many criminal organizations are short-lived and comprised of offenders with desired skills who form temporary networks to take advantage of a crime opportunity. Albanese 2004 mentions that these groups often dissolve after exploiting the opportunity, looking for new chances which may need other combinations of skills.

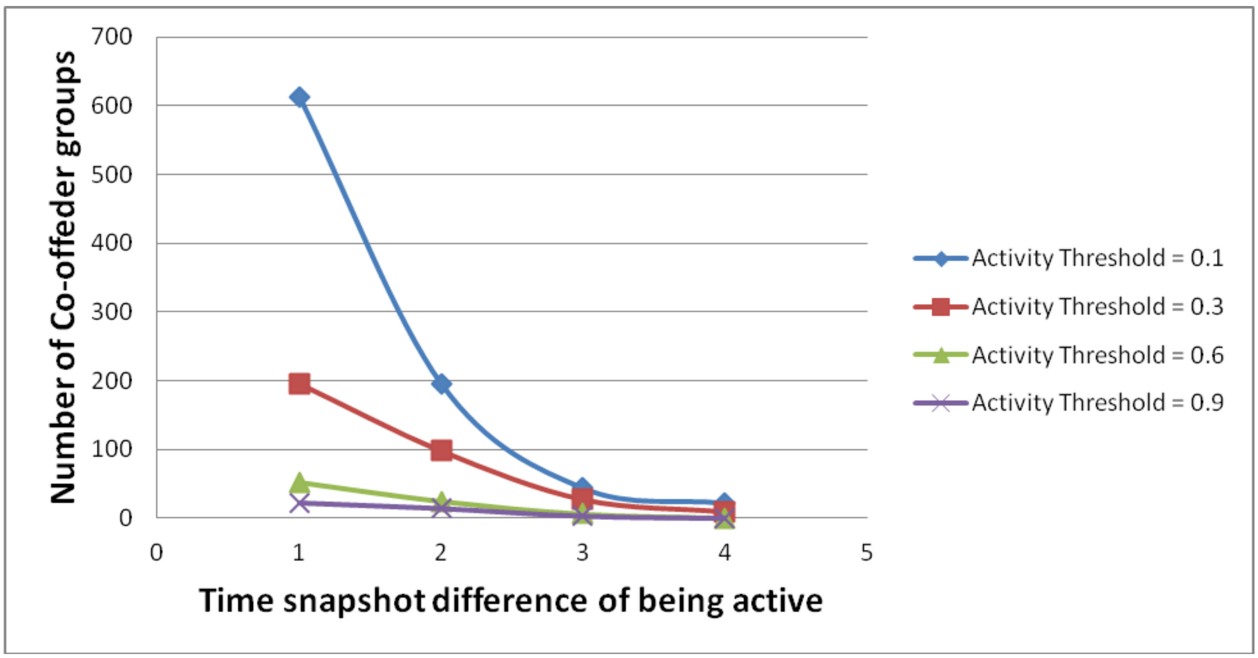

Figure 8: Number of co-offender groups active over a range of time snapshots (which may include inactive snapshots other than the first and the last snapshot in that range)

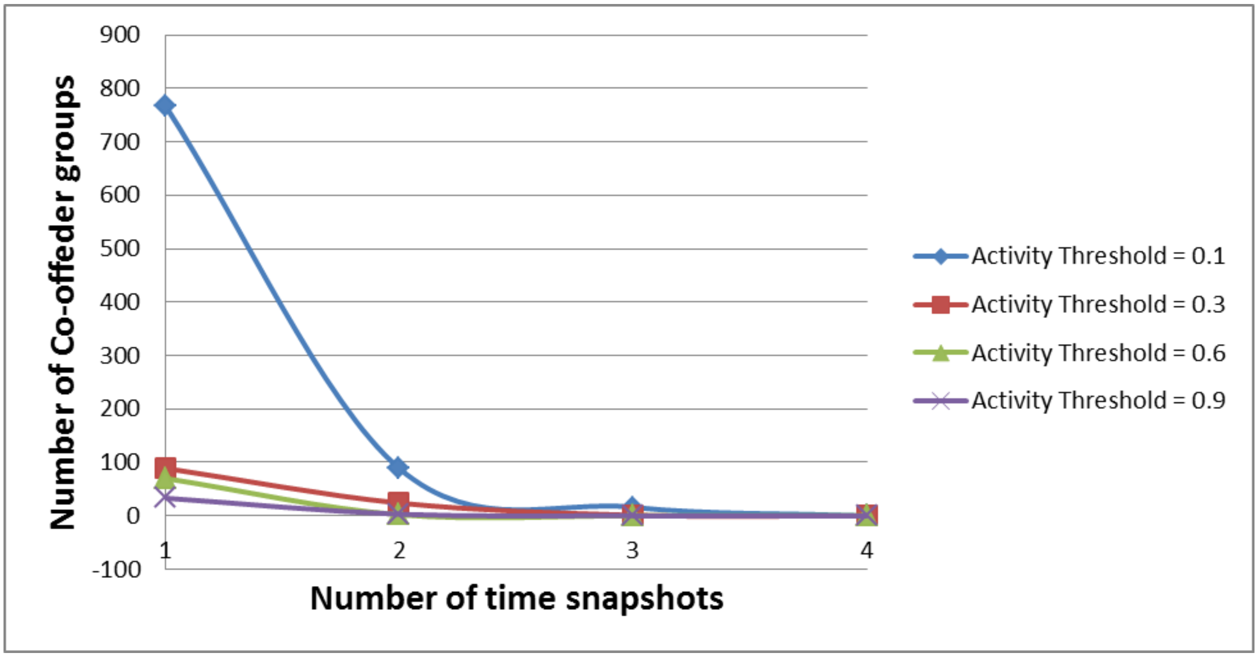

Figure 9: Number of co-offender groups by the total number of observed time snapshots in which a group was considered active

Another important aspect of active co-offender groups is the number of snapshots in which a group is being active. Assume a co-offender group C_1^1 having 4 members and 5 co-offending links detected at time snapshot $t=1$. We have observed 1, 0, 4, and 5 co-offences, respectively, among members of this group in the time snapshots from t=2 to t=5. Activity of this group for each of these snapshots would be respectively $\tau_1^{1,2}=0.2$, $\tau_1^{1,3}=0$, $\tau_1^{1,4}=0.8$ and $\tau_1^{1,5}=1$. Then we conclude that co-offender group C_1^1 was active at three snapshots for any activity threshold equal to or less than 0.2. Statistics of this phenomenon is illustrated in Figure 9. Even with small *activity thresholds* α, we do not have any co-offender group active in all time snapshots. With median α, we observe only a few co-offender groups which were active in three time snapshots. This may indicate that, due to reasons such as incarceration or changing crime-committing tactics and trends, co-offender groups generally do not maintain their co-offending activity for a long time period.

For considering a group active, we apply the activity threshold α=0.3. This means that a group is considered active if it maintains at least 30 percent of its structure unchanged in the next time snapshot.

5.1.1CO-OFFENDER GROUP SIZE

Albanese (2004) concludes that most criminal organizations are quite small. Our study corroborates this result. Figure 10 provides the size distribution for known co-offender groups and active co-offender groups, and Figure 11 shows the frequency of committed offences per group. Most groups committed less than 10 offences, but there are a few groups that committed even more than 100 offences during their life-cycle period.

Average group size for co-offender groups is 4.2 and for active co-offender groups it is about 6.5. When comparing active co-offender groups to offender groups, a larger percentage of active offender groups have periphery members, and the average number of periphery members is greater, which may indicate that the periphery members play a more important role in the criminal activities of active co-offender groups. The maximum number of kernel and periphery members in active co-offender groups, compared to co-offender groups, is significantly smaller. Seven percent of the co-offender groups have more than 10 and only 0.1% have more than 50 members. In the active co-offender group set, 8% of them have more than 10 members, and there is no group with more than 50 members. Of course, the size of these groups would likely be larger if all offences (not just known offences) were available.

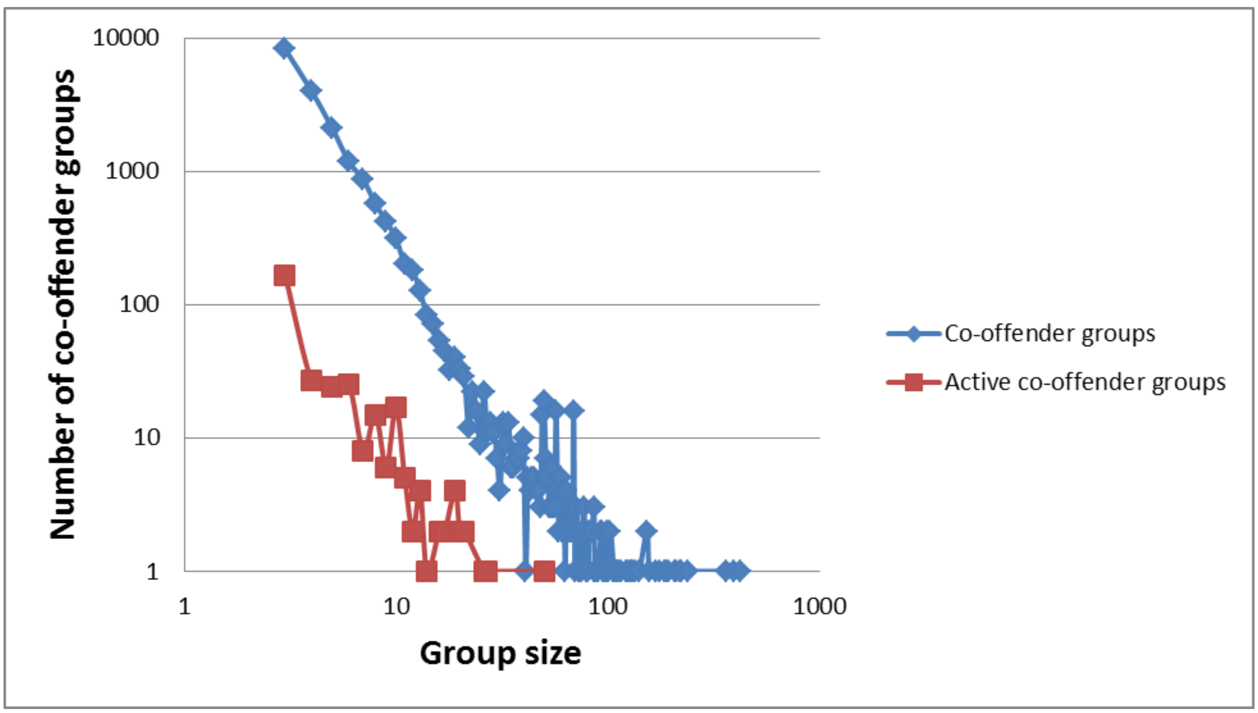

Figure 10: Size of co-offender groups, active co-offender groups

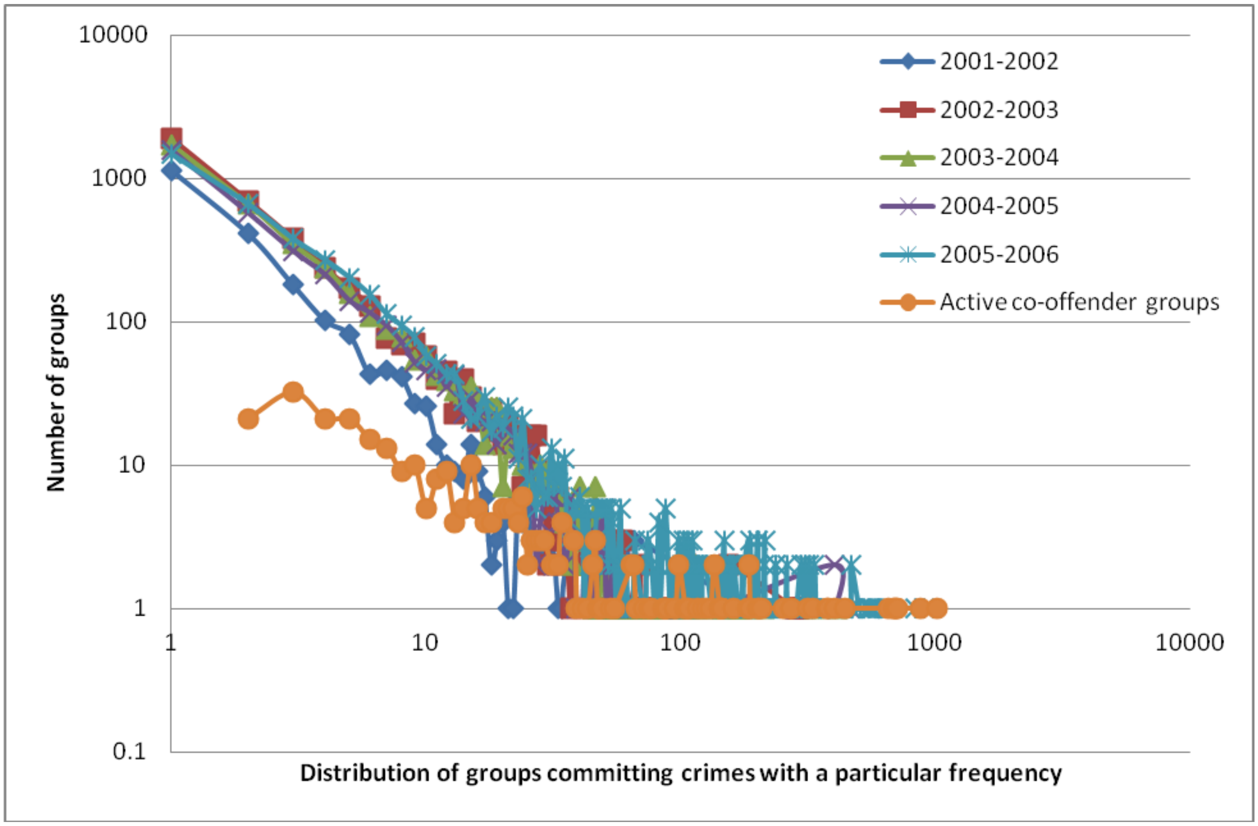

Figure 11: Group committed crime frequency

5.1.2CO-OFFENDER GROUP EVOLUTION

Co-offender groups, similar to any other form of social community, typically evolve over time. A co-offender group may grow by admitting new members, shrink by losing members, split into two or more groups, or a new group may form by merging two or more existing groups. Given the limited observable time span, it is difficult to quantify the whole life cycle of co-offender groups, not knowing their history previous to the first time step and their future history past the last time step. Figure 12 shows the statistics of different evolution scenarios in the five studied snapshots. For the matching function, the threshold value 0.3 applies for considering a group as survived (that is, it continues to exist in any subsequent time snapshot) and a value greater than 0.2 and smaller than 0.3 for split and merged, respectively. Groups with matching thresholds smaller than 0.2 are considered as ceased groups (that is, not visible in any subsequent time snapshot). Over the 5 years of data, about 4% of all co-offender groups survive, but split and merge events occur rarely, less than 1% of the groups. About 96% of the co-offender groups are considered ceased, since we do not observe their activity in the next time step, and 95% of all groups are newly emerged ones.

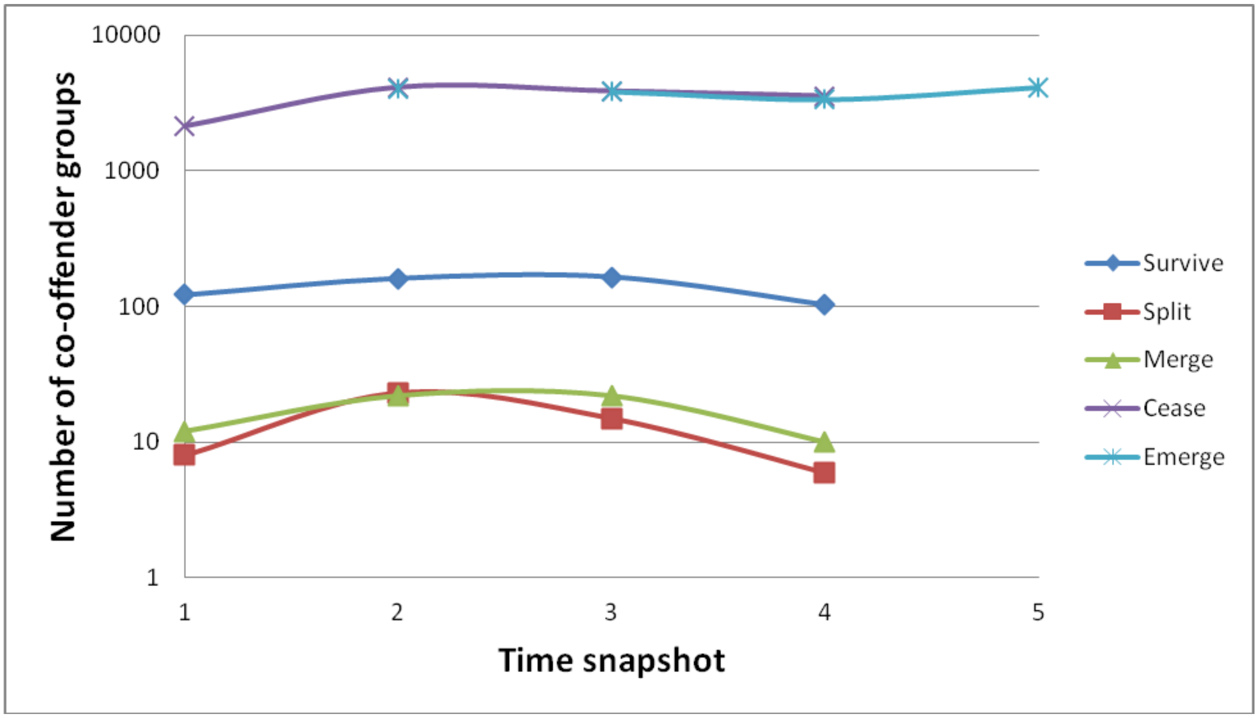

Figure 12: Co-offender groups evolution trace

5.1.3CO-OFFENDER GROUP OVERLAPPING

Figure 13 presents the distribution of size of overlap for co-offender groups and for active co-offender groups. For both group types the result is fairly similar. We see higher numbers for smaller sizes of overlap, which was predictable due to the applied method which is designed based on a strict definition of communities in the networks. Using a less strict definition of co-offender groups means that many of the currently overlapping groups merge into larger groups. In some cases, we observe several pairs of groups with more overlap. This is also because the applied method even differentiates between groups that have common periphery members but completely different kernel members. Regarding serious groups, there is little observable overlap, which again confirms their completely different structure compared to co-offender groups and active co-offender groups.

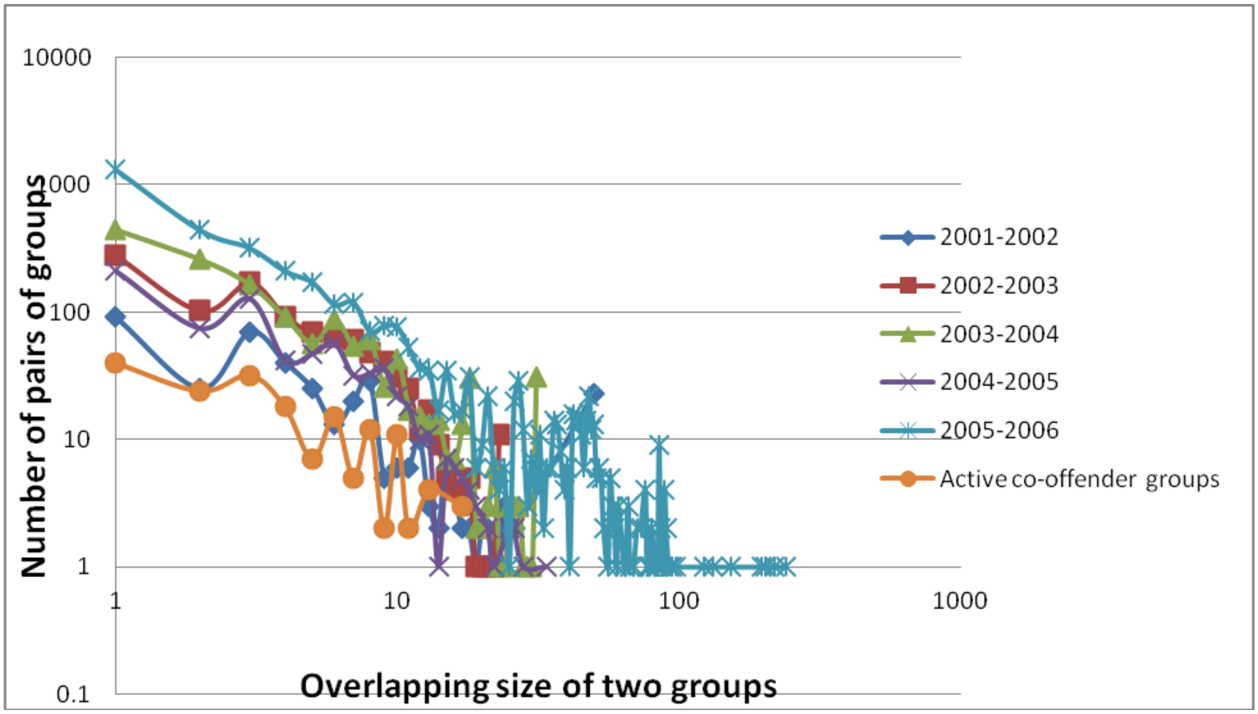

Figure 13: Number of shared members for the overlapping co-offender and active co-offender groups by the number of pairs of overlapped individuals

5.2 Hard Constraint Approach

The "hard constraint approach" is the most basic test to determine possible criminal organizations in the data set since the approach most closely matches the *Criminal Code* definition of criminal organization. In the hard constraint approach all crimes are categorized into a binary classification of the two classes: 1) serious crimes with material benefit; and 1) non-serious crimes or serious crimes that do not appear to provide the offender group with a material benefit.

The research team worked from a list of serious offences prepared by Public Safety Canada (PS) officials. The list of 'serious material benefit crimes' likely associated with criminal organization offending was developed as follows (Saunders 2012):

> First, the list of all offences in the *Criminal Code* was taken as the preliminary set. The criteria for seriousness—offences subject to indictment and having a penalty of five or more years imprisonment—was then applied to the full set of *Criminal Code* offences, with all offences not meeting this requirement removed from the set. As some offences (commonly referred to as hybrid offences) may proceed either by indictment or by summary process, all offences where indictment was an option were included as in the list of section 467.1 "serious crimes".

The set of offences this decision rule generated was then further examined to extract those offences with a "material benefit." Under section 467.1 and current case law, "material benefit" could include an act that results in an intangible gain such as to one's criminal reputation as well as financial gain. The law on point is still evolving so the current list of material benefit crimes represents an approximation of material benefit offences based on examination of literature and other information related to organized criminality in Europe and the US. Essentially, the list includes offences for which at least one known case of a direct material benefit (e.g., cash payment) or intangible benefit (e.g., increases one's criminal reputation) exists in either of the two named jurisdictions.

The PIRS data set used in this study identifies a very large set of offence category types. The PIRS list is a classification of calls for police services and contains information on offences and violations of many federal and provincial statutes as well as violations of the *Criminal Code*. The PIRS criminal event categories match the list of 'serious material benefit crimes' to a substantial but incomplete extent: 112 out of 192 'serious material benefit crime' categories have a matching PIRS crime category. 'Serious material benefit crime' offences that lacked a PIRS category match could not be used in this study because we could not attribute any of the offences under study to those 'serious material benefit crime' categories. All PIRS offences that do not qualify as a serious offence resulting in a material benefit are disregarded in our analysis of the crimes committed by a co-offender group.

It should be noted that this study tests the feasibility of a new analysis technique for exploring networks in criminal organizations using a historical data base. Future work could use current data rather than historic data. With current data it would be feasible to determine how police crime categories are translated into *Criminal Code* sections in the process of laying charges and trying cases by linking police records with court records. This is not possible at this time in British Columbia with historic data. Future research could also make use of future modifications in 'serious material benefit crime' classifications or in Canadian Centre for Justice Statistics (CCJS) data collection methods. As this study shows, there is value in improving analysis techniques concurrent with advancements in data collection and changes in information classification schemes.

The analysis presented in this section uses a binary classification of offence types according to whether or not an offence constitutes a serious offence that results in potential material benefit for a co-offender group committing the offence. All offences that do not qualify as a serious offence resulting in material benefit are disregarded in the analysis of the crimes committed by a co-offender group. This analysis used a binary classification scheme that used the categories of offences in the historic PIRS data and divided these offences as closely as possible to a list of serious crimes resulting in material benefits used by Public Safety Canada (Saunders 2012). The 'serious material benefit crime' list includes crimes defined under relevant *Criminal Codes* which correspond directly with offences prosecuted in Canadian courts. The PIRS list is a classification of calls for police services and contains some categories that are not provided in detail under the *Criminal Code.* As well the PIRS categories do not contain all of the categories in the Public Safety Canada list. But the classification used does divide police classification by seriousness resulting in material benefit.

The following presentation and discussion of results of the analysis is strictly based on this binary classification of the hard constraint approach. We can identify serious groups based on two aspects: 1) the ratio of committed serious offences to total committed crimes, and 2) the number of committed serious offences. In the first approach, a co-offender group is considered serious if *P*% of all offences committed by this group is serious in the above sense. In the second approach, we consider a group serious if two or more members of this group were involved in more than *N* serious offences (where *P%* and *N* refer to adjustable threshold values for the percentage and number of serious offences committed by a group, introduced for the purpose of controlling the analysis). Further to the overall approach taken here, another possible way would be to calculate the ratio of serious crimes to the number of individuals in the group.

In the first approach to detecting possible criminal organizations, we consider the ratio of a group's committed serious crimes to total committed crimes. Figure 14 and Figure 15 respectively show the number of co-offender groups and active co-offender groups in respect to the percentage of committed serious crimes. Considering different percentage thresholds, *P*=30, *P*=60 and *P*=90, respectively 25%, 10% and 8% of the co-offender groups and 33%, 9% and 4% of the active co-offender groups remain in the list.

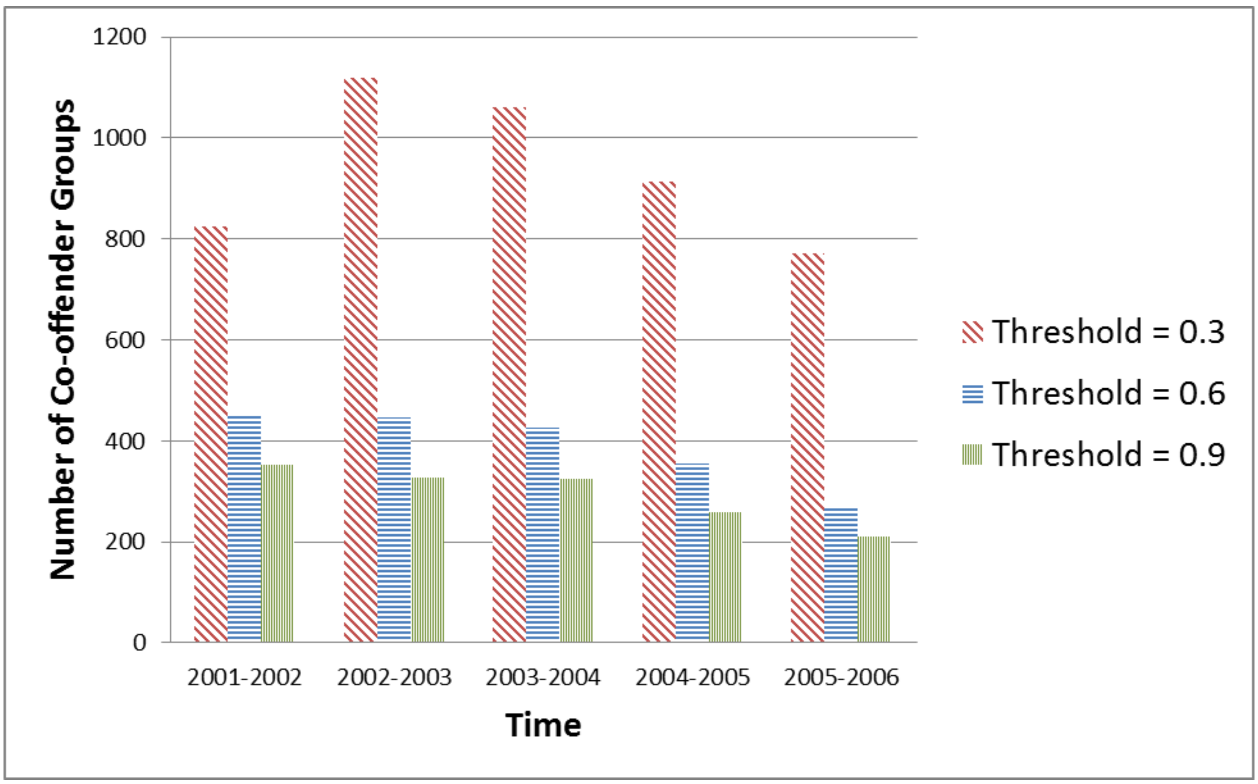

Figure 14: Number of co-offender groups in respect to the proportion of committed serious crimes

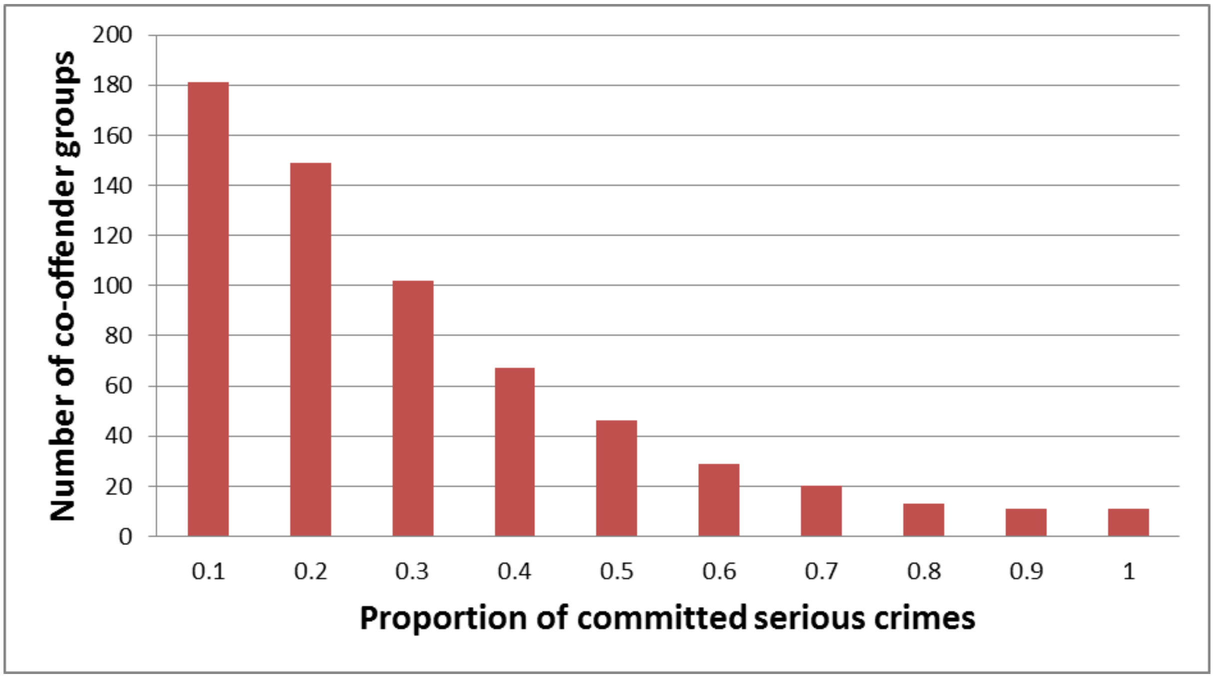

Figure 15: Number of co-offender groups in respect to the proportion of committed serious crimes

Figure 16 and Figure 17 respectively show the number of co-offender groups and active co-offender groups in respect to the number of serious crimes these groups have committed over their life cycle. In the co-offender group set, about 59% did not commit any serious crime; 91% of them were involved in less than five serious crimes and only 0.02% of the co-offender groups committed more than 10 serious crimes. The average number of serious crimes per co-offender group is 1.2. In total, 25% of the active co-offender groups did not commit any serious crime, while 73% of these groups committed less than five serious crimes during their life cycle and only 0.09% of the active groups committed more than 10 serious crimes. On average, each active co-offender group has committed 3.7 serious offences. These results show that, compared to the co-offender groups, active co-offender groups more frequently commit serious offences.

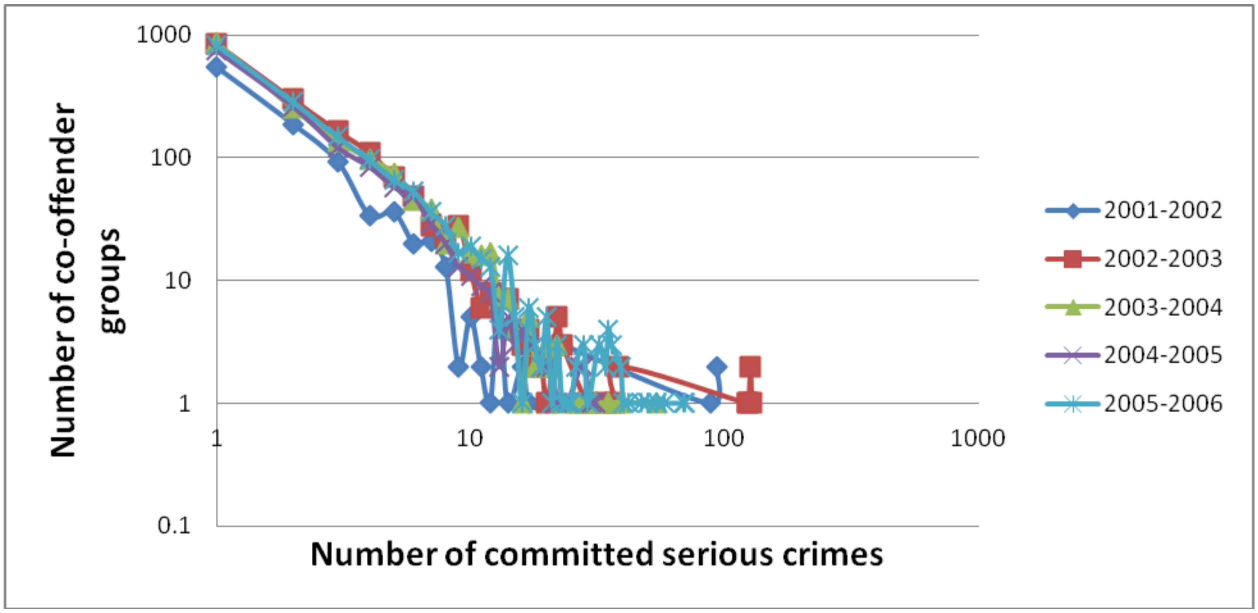

Figure 16: Number of co-offender groups in respect to the number of serious crimes they committed

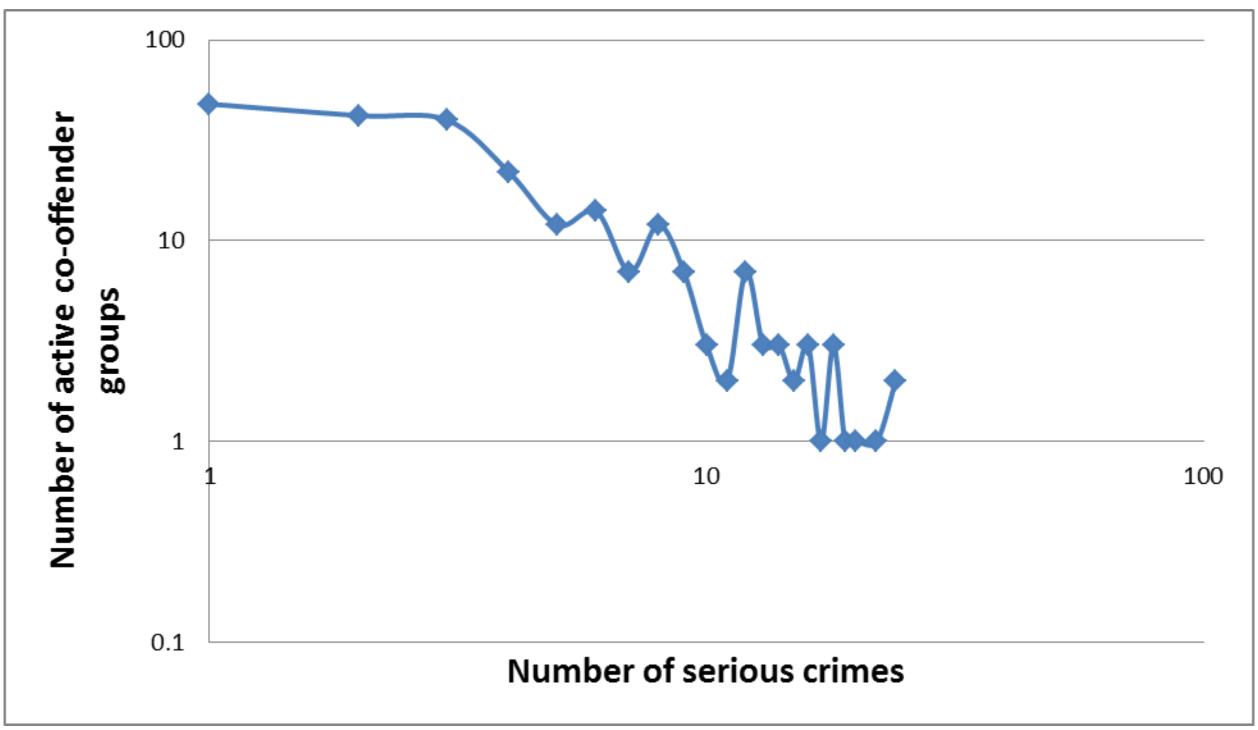

Figure 17: Number of active co-offender groups in respect to the number of serious crimes they committed over their observable life cycle

In the hard constraint approach, based on the definition of organized crime in the *Criminal Code,* we consider any active offender group that committed at least one serious offence. In this case, out of 313 active offender groups, we end up with a list of 236 groups. These 236 groups meet the minimum analytic threshold to be considered possible criminal organizations. In total, 49 of these groups are involved in only one serious offence and one of them committed 24 serious offences.

Figure 18 shows the distribution of life cycle duration (in number of days) of the possible criminal organizations in the hard constraint approach. The starting point of the life cycle of a possible criminal organization is the time point of the first observable offence committed by at least one member of the group in the time snapshot where the group emerged. The end time is the time point of the last observable offence committed by at least one member of the group in the last time snapshot where the group was active. The average life cycle duration is 773 days (close to three years). There are 21 groups with life cycle time duration of less than one year. Two hundred and fifteen groups had a life cycle duration of 1 up to 2 years; 93 groups had a life cycle of 2 up to 3 years; 30 had a life cycle of 3 up to 4 years; and 4 had a life cycle of 4 to 5 years which is the maximum possible in this test dataset.

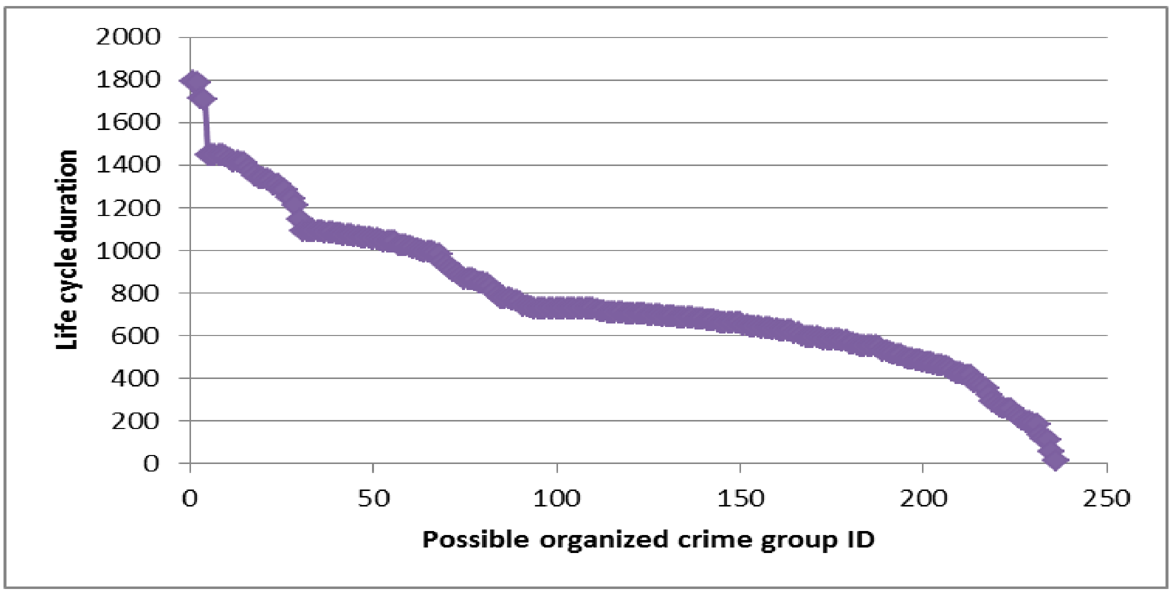

Figure 18: Life cycle duration of possible criminal organizations

5.3 Soft Constraint Approach

For calculating co-offender group criminality, we apply the RCMP crime seriousness index as delineated in the Operational Statistics Reporting System (OSR). This index uses a seriousness hierarchy with 151 groups, where each crime type belongs to one of these groups. For each crime type in the dataset, the corresponding seriousness group level is scaled linearly, and these normalized values are interpreted as indicators of the seriousness of offences. Table 2 shows a small sample of the OSR crime seriousness hierarchy and corresponding seriousness values.

Table 2: Crime seriousness hierarchy and values (sample)

Crime Type	Hierarchy Level	Seriousness
Murder 1st Degree	1	1
Abduction of Person Under 14	18	0.89
Production of Heroin	41	0.74
Break and Enter, Residence	58	0.62
Theft of Automobile	75	0.52
Theft over $5000 - Bicycles	83	0.46

Figure 19 and Figure 20 illustrate the number of co-offender groups in respect to different criminality thresholds β, where β is equal to the summation of seriousness of offences committed by members of a co-offending group divided by the total number of those offences. About 30% of all co-offender groups pass the threshold β=0.6, which means a larger percentage of the co-offender groups commit minor crimes, which is intuitive. Finally, β=0.8 identifies less than 6% of the groups, which implies that a small percentage of co-offender groups are consistently involved in serious crimes.

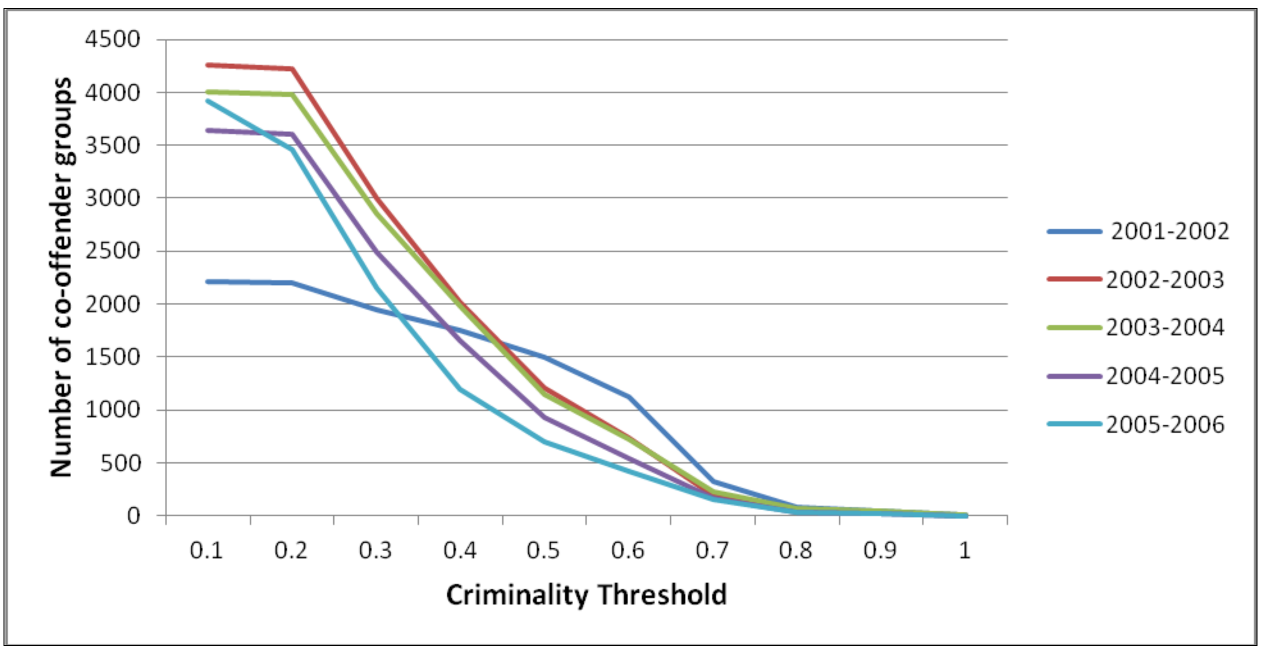

Figure 19: Number of co-offender groups in respect to differrent criminality thresholds

Finally, the possible criminal organizations are extracted from active offender groups. Applying activity threshold 0.3, a total of 313 groups are considered active offender groups. Figure 20 shows the number of active offender groups for different criminality thresholds. From 313 active offender groups 89, 39, 18, 8 and 5 groups result for criminality thresholds 0.5, 0.6, 0.7, 0.8, and 0.9, respectively. There is no active group having criminality equal to 1.

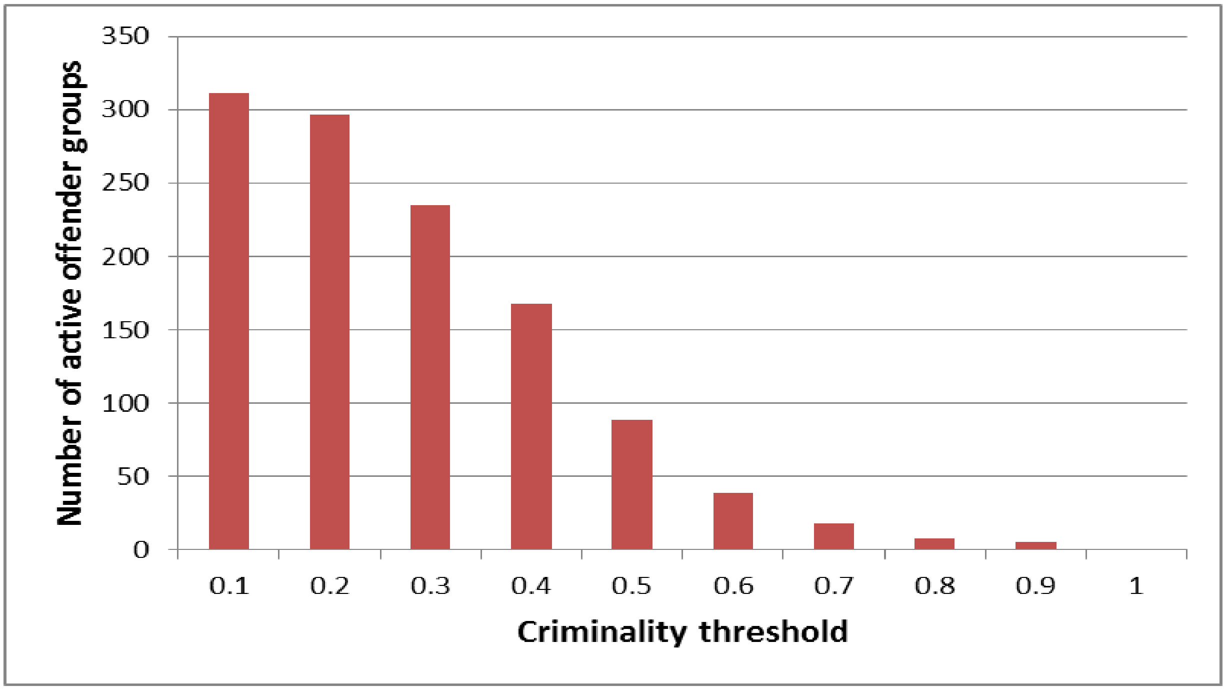

Figure 20: Number of active co-offender groups in respect to different criminality thresholds

An important question in the soft constraint approach is about the applied criminality thresholds for detecting possible criminal organizations. For this purpose we use the average seriousness of all serious offences (defined in the 'serious material benefit crime' list), which is 0.6. In total, 39 groups have criminality higher than 0.6 which are considered possible criminal organizations.

As an alternative approach for computing criminality of groups, we also test the use of the crimes' seriousness as defined in the Crime Severity Index (CSI) list. The weights used in the CSI were developed by the Canadian Centre for Justice Statistics, Statistics Canada and the Canadian law enforcement community to empirically compare the relative seriousness of criminal offending. In the CSI, offences are weighted by the relative seriousness by which they are treated by Canadian courts. Our experiments find that applying the CSI to analyze group criminalities using our present method results in findings that are much skewed and we do not see a distributed group criminality as for the other approach (OSR crimes seriousness list). The reason for this is that in the OSR list (mapped to the OSR list) crimes are assigned a seriousness value between 1 and 151. But in the CSI list the maximum and minimum values are 7041.45 and 1.16, and 90% of the crimes seriousness has a value less than 500. Therefore the group criminality results in a skewed-left distribution, with very few active offender groups identified as serious offender groups. Future research will need to develop an efficient approach to normalize crimes seriousness to get more distributed group criminality.

In the 5 years of the crime data, using the soft constraint approach, there are 39 groups that are both active and serious and thus are considered possible criminal organizations. Interestingly, most of these groups have high activity, which shows the close relationship among their group members. The average size in this set of groups is 4.7, which is much smaller than the size of active co-offender groups. This point supports the theory that with increasing group criminality co-offender group size decreases. Having less number of possible criminal organizations with periphery members compared to active co-offender groups also implies that in possible criminal organizations, the kernel members are not eager to collaborate with co-offenders outside of the group's core (kernel).

6 Concluding Remarks

Controlling crime necessitates the investigation of criminal networks, criminal organizations and their illegal activities, constituting a serious undertaking for law enforcement and the criminal justice system. We propose here a computational co-offending network analysis approach for detecting possible criminal organizations. We evaluate the proposed methods by examination of a large real-world crime dataset. Our examination shows that although criminal group activity does not occur as routinely as other criminal activities, which is intuitive, there is continuous criminal collaboration inside crime groups. But for most of the groups such co-offending behaviour does not persist over longer time periods. Our study also shows that active co-offender groups typically have more peripheral members in contrast to serious groups which tend to have fewer peripheral members and a tightly connected kernel. This finding suggests that serious groups operate primarily from inside their core membership.

Starting from a crime dataset with 4.4 million records and a co-offending network with 150,000 actors, we were able to detect more than 18,000 co-offender groups, including more than 300 active groups. Using the hard constraint approach, our study identifies 236 possible criminal organizations which committed one or more than one serious offences over the observed time frame. In addition, our study yields 39 possible criminal organizations using the soft constraint approach.

Our analytic approach provides potentially important insights into the ways in which co-offending networks shape and affect criminal behaviour. Albeit, it should be noted that co-offending networks do not necessarily identify all individuals of an organization, simply because those operating in the background, who often direct the activities of others, may not be visible in the data. Another possibility is that some of these "possible criminal organizations" represent particular functional components of larger criminal organizations that do not appear directly in police-reported crime. For obtaining a more holistic picture of criminal organizations, one must combine police-reported crime data with data from intelligence agencies. This type of analysis can identify possible individuals or offenders for further investigation or as part of a disruption strategy. Further, the approach taken here primarily concentrates on criminal organizations with dense member relationships, which is not always the case, especially not for certain forms of criminal networks.

Future work can explore alternative ways to determine and visualize the organizational structure of criminal organizations by developing new rules for identifying potential criminal organizations through data mining. As mentioned before, a major advance could flow from undertaking research with current police data that is linked to current court data. This could make it possible to use a definition firmly based on the *Criminal Code of Canada* by developing a probabilistic concordance of crimes with police categories of calls for police service. It would also be possible to merge data sets of associations between persons that are not based on co-offending. Data mining, with appropriate research, could become a complementary tool used both by operational criminal justice analysts and by policy makers in framing enforcement operations and in framing crime control policy.

References

Albanese, Jay. “North American Organised Crime.” *Global Crime* vol. 6 no 1, 2004.

Borgatti, Stephen P., Ajay Mehra, Daniel J. Brassand, and Giuseppe Labianca. "Network Analysis in the Social Sciences." *Science* 323 (5916): 892–895, 2009.

Brantingham, Patricia L, et al. “Co-offending Network Mining.” In U. Kock Wiil (Ed.), *Counterterrorism and Open Source Intelligence, Lecture Notes in Social Networks*, Vol. 2, Springer 2011, pages 73–102.

Fijnaut, Cyrille, Frank Bovbenkerk, Gerben Bruinsma, and Henk van de Bunt. “Organized Crime in the Netherlands.” *The Hague: Kluwer Law International*, 1998.

ICPC. *International Report on Crime Prevention and Community Safety: Trends and Perspectives, 2010*. International Centre for the Prevention of Crime, Montreal, Quebec, Canada, 2010.

McGloin, Jean Marie and Holly Nguyen. “The Importance of Studying Co-Offending Networks for Criminological Theory and Policy.” In *Proc. Third Annual Illicit Networks Workshop*, Montreal, Quebec, October 2011.

McGloin, Jean Mari, Christopher J. Sullivan, Alex R. Piquero, and Sarah Bacon. “Investigating the stability of co-offending and co-offenders among a sample of youthful offenders.” *Criminology* 46 (1), 2008.

McMillan, David W., and David M. Chavis. “Sense of community: A definition and theory.” *Journal of Community Psychology*, 14(1):6–23, 1986.

Morselli, Carlo. “Inside Criminal Networks. Studies of Organized Crime.” Vol. 8, *Springer*, 2009.

Reiss, Albert J. “Co-offending and criminal careers.” *Crime and Justice: A Review of Research*, 1988.

Reiss, Albert J., and David P.Farrington. “Advancing knowledge about co-offending: Results from a prospective longitudinal survey of London males.” *Journal of Criminal Law and Criminology* 82 (2), 1991.

Saunders, Michael. Personal communication with *Public Safety Canada*, May, 2012.

Statistics Canada. *Measuring Crime in Canada: Introducing the Crime Severity Index and Improvements to the Uniform Crime Reporting Survey*. Canadian Centre for Justice Statistics, Statistics Canada, 2009. Accessed on August 28, 2012 from http://www.statcan.gc.ca/pub/85-004-x/85-004-x2009001-eng.pdf.

Tayebi, Mohammad A., and Uwe Glässer. "Organized Crime Structures in Co-offending Networks." *Proc. 2011 Intl. Conference on Social Computing and its Applications*, Sydney, Australia, December 2011.

Tayebi, Mohammad A., Uwe Glässer, and Patricia L. Brantingham. "Organized Crime Detection in Co-offending Networks. " Proc. *Third Annual Illicit Networks Workshop*, Montréal, Québec, October 2011.

Morselli, Carlo, Thomas Gabor, and John Kiedrowski. *The Factors That Shape Organized Crime*. Ottawa, ON: Public Safety Canada, 2010.

van der Heijden, Toon. "Measuring Organized Crime in Western Europe." In Milan, Pagon (eds.) *Policing in Central and Eastern Europe: Comparing First Hand Knowledge with Experience from the We*st, Slovenia: College of Police and Security Studies, 1996.

von Lampe, Klaus. "Definitions of Organized Crime." Accessed on February 10, 2012 from http://www.organized-crime.de/organizedcrimedefinitions.htm.

Wasserman, Stanley and Katherine Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.