2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

# Investigating Organized Crime Groups: A Social Network Analysis Perspective

Mohammad A. Tayebi and Uwe Glässer
*Software Technology Lab, Computing Science*
*Simon Fraser University, B.C., Canada*
*Email: {tayebi,glaesser}@cs.sfu.ca*

***Abstract*—In this paper, we analyze co-offending networks derived from a large real-world crime dataset for the purpose of identifying organized crime structures and their constituent entities. We focus on methodical and analytical aspects in using social network analysis methods and data mining techniques. The goal of our work is to promote computational co-offending network analysis as an effective means for extracting information about criminal organizations from large real-life crime datasets, specifically police-reported crime data. We contend that it would be virtually impossible to obtain such information by using traditional crime analysis methods. For our approach we provide an experimental evaluation with promising results.**

***Keywords*-Co-offending networks; Criminal organization; Social network analysis; Community detection**

## I. Introduction

Co-offending networks link offenders who have jointly committed criminal offences. They constitute a widespread form of social networks that plays a central role in crime investigations, and has broad and important implications for the study of crime and criminal justice [8]. In fact, "understanding co-offending is central to understanding the etiology of crime and the effects of intervention strategies" [6]. 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 better understand organized crime and as an instrument in evidence-based policy development aiming at crime reduction and prevention strategies.

In this paper, we propose a new computational approach to organized crime group detection based on a social network analysis perspective. We focus on methodical and analytical aspects in utilizing social network analysis methods and data mining techniques. The main goal of this work is to promote *computational co-offending network analysis* as an effective means for extracting information about criminal organizations from large real-life crime datasets, specifically police-reported crime data. We contend that it would be virtually impossible to obtain such information by using traditional crime analysis methods. The work presented here extends our work in [2], [5], adding an experimental evaluation of the method and refining the conceptual model.

The approach described here comprises four major building blocks: *1)* formal models for crime data and co-offending network extraction that aim at bridging the conceptual gap between data level and mining level; *2)* a group detection method, an extension of the clique percolation method [3], to match a working definition of offender group; *3)* a crime assessment method which covers and formalizes common characteristics of organized crime found in the criminology literature; *4)* a group evolution model for analyzing offender group behavior over the observable life cycle of a group.

A challenging aspect is the need for a precise definition of *what* exactly constitutes a criminal organization. Striving for a general and open definition, a natural source is the criminal code, although this depends on the specific country. A baseline definition of criminal organization in the *Criminal Code of Canada* is described in [12, p. 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. ... Section 467.1(1) of the Criminal Code of Canada.*

Confronted with a bewildering diversity of characteristics in definitions of organized crime and criminal organizations, the conceptual model of organized crime appears not clearly rendered in the literature—at least not for the purpose of computational analysis. Based on extensive literature review [12], [1], [10], [11], [13], [4] and in-depth consultation with subject matter experts, we have devised a working definition. We evaluate here the efficacy of our approach on a crime dataset representing five years of police arrest-data for the Province of British Columbia, comprising several million data records, each of which refers to a reported offence. This data was made available for research purposes by Royal Canadian Mounted Police (RCMP) and retrieved from the Police Information Retrieval System (PIRS), a large database keeping information for the regions of the Province of British Columbia which are policed by the RCMP.

To the best of our knowledge, there is no comparable study on large real-world crime data. Generally, the extracted results are promising in two ways: first of all, they are consistent with established characteristics in the literature

978-0-7695-4799-2/12 $26.00 © 2012 IEEE
DOI 10.1109/ASONAM.2012.96



IEEE computer society

on organized crime groups, including their structure and dynamics. Secondly, our experiments for organized crime group detection result in 35 groups. In our case study, we conclude that 65% of these groups are possible organized crime groups with a high probability.

Section II discusses related work. Section III introduces the basic concepts and the terminology used in Section IV to propose an organized crime group detection framework. Next, Section V interprets and discusses our experimental results on the crime dataset. Section VI concludes the paper.

## II. Related Work

This section discusses relateded work on co-offending network analysis, organized crime and community detection.

### A. Co-offending Networks

Several empirical studies using social network analysis methods to analyze co-offending or terrorist networks have focused on the stability of associations in such networks. Morselli [1] offers a thoughtful 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 [6] concludes that the majority of co-offending groups are unstable, and their relationships are short-lived. This is corroborated by McGloin et al. [9], who show that there is some stability in co-offending relationships over time for frequent offenders, but delinquents do in general not tend to reuse co-offenders. Reiss and Farrington [7] also found that co-offenders have many different partners, and are unlikely to commit offences with the same individuals over time. However, Reiss [6] also finds that high frequency offenders are "active recruiters to delinquent groups and can be important targets for law enforcement." Given that the above findings are based on very small datasets, 205 individuals in [7], and 5,600 individuals in [9], they may not be representative.

### B. Organized Crime

Studying prominent historical research on how organized crime developed in New York City, Block [10] concludes that "organized crime was not only more fragmented and chaotic than believed, but also it involved *webs of influence* that linked criminals with those in positions of power in the political and economic world."

Looking for a quantitative definition, van der Heijden [11], in an attempt to measure organized crime, proposes common characteristics, including: commission of serious criminal offences, pursuit of profit and/or power, a prolonged or indefinite time period, some form of discipline and control, commercial or businesslike structures, engaging in money laundering, and exerting influence on politics, the media, public administration, judicial authorities, or economy.

A major study in the Netherlands [13] finds great variations in collaborative forms and concludes that "the frameworks need not necessarily exhibit the hierarchical structure or meticulous division of labor often attributed to mafia syndicates. Intersections of social networks with a rudimentary division of labor 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 more than 150 definitions of organized crime was gathered by von Lampe [4]. This also includes comments on how to define organized crime, and definitions by prominent individuals and government agencies, such as the Federal Bureau of Investigation (FBI).

In most cases, existing definitions in the literature on organized crime concentrate on three essential perspectives for characterizing the nature of this form of crime: In the first view, organized crime is primarily about crime. Organized crime is seen as a *specific type* of criminal activity that has certain specific characteristics such as continuity in contrast to irregular criminal behavior. In the second view, organized crime is more related to the *concentration of power*, either in economic or in political structures of the society. And in the third view, the emphasis is on *organized*. That is, the important aspect of organized crime is on how offenders are connected to each other more than what they do.

### C. Community Detection in Social Networks

Algorithms for community detection in static graphs are usually looking for a 'good' partition of the nodes, implying that no node is member of more than one community. The main problem is "what does 'good' really mean?". For dealing with this problem, some quality measures have been defined that give a score to a partition: a good partition is one which maximizes this quality measure. One of the most commonly used quality measures is *modularity* [14].

In studies of how communities evolve over time, two main approaches have been used: *1)* applying temporal information directly in the community detection process, and *2)* tracking communities over a number of snapshots in time. To take into account temporal information, recently, a new type of clustering, called *evolutionary clustering*, that captures the evolutionary process of clusters in time-stamped data was introduced. Chakrabarti et al. [15] address this issue in their paper by proposing a framework called "temporal smoothness".

Another method for identifying relations between communities is constructing the networks for each time step. First, communities are identified within each of these networks, then relationships among communities on subsequent snapshots are recognized. Hence, such an algorithm operates in two steps: *1)* static community detection on each snapshot, and *2)* applying a matching function to recognize how these static communities evolve over a number of time steps.

## III. Basic 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 "lifecycle." Finally, we define the concept of *organized crime group* in terms of characteristics that discriminate possible criminal organizations from regular offender groups. The rational for the applied characteristics is to be in line with the respective definition in the *Criminal Code of Canada*.

**Crime Data** We model crime data in the form of an attributed tripartite *hypergraph* $\mathcal{H}(\mathcal{N}, \mathcal{E})$ with node set $\mathcal{N}$ and a set of hyperedges $\mathcal{E}$. 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, such as a mobile phone, a tool, a vehicle, a weapon or a bank account. A hyperedge $e$ of $\mathcal{E}$ is a non-empty subset of nodes $\{n_1, n_2, \ldots n_p\} \subseteq \mathcal{N}$ such that the following three conditions hold: $|e \cap I| = 1$, $|e \cap A| \geq 1$ and $|e \cap R| \geq 1$.

**Co-offending Network** A co-offending network is derived in several steps from a crime data model (see [2] for details) and comprises one or more connected components consisting of offender nodes, where nodes are connected for offenders 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)* $\in \mathbb{N}$. Assuming $k$ offenders and $m$ crime events ($k, m > 1$), we define a $k \times m$ matrix $M$, such that $m_{uv} = 1$, if offender $o_u$ is involved in event $i_v$, and "0" otherwise. A co-offending network is thus a $k \times k$ matrix $N = MM^T$.

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

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

**Offender Group** An offender group comprises three or more 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 $n$ offender groups in the co-offending network at time period $t$.

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

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

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

**Serious Offender Group** An offender groups who's overall criminal activity at time period $t$ shows a high degree of serious criminal offences is called serious offender group and is referred to by $S_i^t$.

**Organized Crime Group** In theory, the two concepts of organized crime group and 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 organized crime group and offender group is not always clear-cut and can be challenging. To qualify as criminal organization, a necessary (but not sufficient) condition is the commission of serious offence motivated by material benefit. While the meaning of 'serious offence' can be clearly defined in terns 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 $m$ organized crime groups in the co-offending network at time period $t$.

**Organized Crime Group Evolution Trace** 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 organized crime groups over $n$ consecutive time periods that shows the dynamic transformation, or evolution, of the organized crime group $O_a^t$ since time period $t$.

## IV. Organized Crime Group Detection

Community detection is a prominent research topic in social networks. The nature of organized crime groups, however, is different from other types of groups like friendship or co-authorship groups. Organized crime groups are usually well established with group membership being defined explicitly and strictly. Unlike friendship or co-authorship communities, offender groups as well are characterized by member relationships that are more systematic and organized to achieve material benefit. Therefore, detecting organized crime groups calls for a stricter definition of community.

Based on fundamental discussions in the criminology literature, one can summarize the important characteristics of organized crime groups as: *1)* These groups have at least three members and can be categorized as centralized

**Algorithm 1** Organized Crime Group Detection

**Input:** (1) Crime event dataset
(2) Crime seriousness index
(3) Activity and criminality thresholds $\alpha$, $\beta$
**Output:** Organized crime groups $O_1^t, O_2^t, \ldots, O_m^t$

/* Data Preparation */
**For each** set of crime incidents in the interval $[t_1, t_2]$
  Extract the co-offending network
  Detect offender groups $C_1^t, C_2^t, \ldots, C_n^t$
  **For each** offender group $C_i^t \in C^t$
    Compute the group *activity* $\Theta_i^{t_1,t_2}$
    Compute the group *criminality* $\Phi_i^t$
  Identify possible organized crime groups
  **For each** possible organized crime $O_i^t$
    Assess overall group *material benefit*
 Apply the evolution trace model on $O_1^t, O_2^t, \ldots, O_m^t$

or distributed or hierarchical groups. Regardless of this classification, our focus is on offender groups for which the density of their intra-group collaborations is higher than the density of intergroup collaborations; *2)* Organized crime groups are characterized by a distribution of roles and different degrees of agency amongst individuals, where groups can overlap and may have common members; *3)* These groups commit serious crimes with the perspective of gaining material benefit; 4) Their activity is more continuous compared to regular offender groups.

For the purpose of organized crime group detection, in each time snapshot of a co-offending network the following tasks are carried out in consecutive steps: (1) Discover offender groups in the current network; (2) Compute the activity and criminality of these groups in the time period between the current network and the previous network based on the offences that were committed by their members; (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 groups evolutionary trace for the current time period. In the following each of these steps is explained in more detail.

### A. Offender Group Detection

In the first step of the proposed method, offender groups are built up from $k$-cliques. A group consists of adjacent $k$-cliques, sharing at least $k-1$ nodes with each other. Since an offender group should have at least three members, we assume $k = 3$. Each clique uniquely belongs to one community, but cliques within different communities may share nodes. Hence, we have overlapping groups with common members. For each offender group $C_i$, these members are assigned as their kernels $K(C_i)$. Kernels are the main members of an offender group and are completely involved in the group activities. In the second step, neighbor nodes connected directly to the kernels are added to the offender groups. These nodes are called *peripheries*. Peripheries of an offender group $C_i$ are denoted by $P(C_i)$.

### B. Organized Crime Group Detection

Activity and criminality of an offender group are two key characteristics toward understanding the group structure. Below we present how these two measures are computed.

Criminality of an offender group $C_i$ at time step $t$, denoted by $\Phi_t(C_i)$, is defined as:

$$\Phi(C_i) = \sum_{k=1}^{k=n} \frac{\varphi_{i_k}}{n} \tag{2}$$

where $\varphi_{i_k}$ indicates the seriousness of an offence $i_k$ that is committed by members of group $C_i$ at time step $t$.

Let $i_1, i_2 \ldots, i_n$ be the offences in which members of $C_i$ were involved at time step $t$. Activity of the offender group $C_i$ at time $t_1$ respect to time $t_2$, denoted by $\Theta_{t_1,t_2}(C_i)$, is computed as follows:

$$\Theta_{t_1,t_2}(C_i) = \frac{|R_{t_1}(C_i)|}{|R_{t_2}(C_i)|} \tag{3}$$

where $|R_{t_1}(C_i)|$ and $|R_{t_2}(C_i)|$ denote the number of binary relationships (co-offences) within offender group $C_i$ at time steps $t_1$ and $t_2$, respectively.

In order to determine whether a detected offender group qualifies as a possible organized crime group, activity and criminality of the group are considered. For this purpose, we define two thresholds $\alpha$*-activity* and $\beta$*-criminality*. A given offender group $C_i$ is considered as active group $A_i$, if $\Theta(C_i) > \alpha$, and as serious group, if $\Phi(C_i) > \beta$. We consider an offender group a possible organized crime group, if it is serious and active. Meaningful values for $\alpha$ and $\beta$ are to be determined experimentally. Algorithm 1 outlines the pseudo-code of this approach.

### C. Organized Crime Group Evolution Model

Like other communities, organized crime groups typically evolve over time. An organized crime group may grow by admitting new members, shrink by losing members, split into smaller groups, or a new group may form by merging existing groups. Therefore, we devise a model that addresses all these aspects of organized crime group evolution.

The model needs to determine which group at previous time has evolved into which group at current time. Five phenomena can occur for a group in a single snapshot: a community may survive, split, merge, emerge or cease [16]. For this purpose we introduce a matching function

$$match : \mathcal{G} \times 2^{\mathcal{G}} \to \mathcal{G}$$

where $\mathcal{G}$ denotes a set of groups and $2^{\mathcal{G}}$ is the powerset of $\mathcal{G}$. For a given organized crime group $O_i^t$ and set of organized crime groups $\mathcal{G}^{t+1}$, let $match(O_i^t, \mathcal{G})$ yield the group $O_i^{t+1}$ such that this group has the largest intersection with $O_i^t$ above a given threshold $\lambda$, as formally defined below.

$$match(O_i^t, \mathcal{G}) = O_j^{t+1} \quad \text{with} \quad \forall\ O_k^{t+1} : O_k^{t+1} \in \mathcal{G} \ \wedge$$
$$overlap(O_i^t, O_j^{t+1}) \geq overlap(O_i^t, O_k^{t+1}) \ \wedge$$
$$overlap(O_i^t, O_j^{t+1}) > \lambda \quad (4)$$

where, for two organized crime groups $O, O' \in \mathcal{G}$, we define

$$overlap(O, O') = min(\frac{|O \cap O'|}{|O|}, \frac{|O \cap O'|}{|O'|}) \quad (5)$$

Using the matching function, we apply the following rules for tracking the evolution of organized crime groups:

- $O_i^t$ survives in the next time slot as $O_j^{t+1}$, if $O_j^{t+1} = match(O_i^t, \mathcal{G}^{t+1})$ and for each $O_k^t \neq O_i^t$, $O_j^{t+1} \neq match(O_k^t, \mathcal{G}^{t+1})$.
- $O_i^t$ 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 defined threshold.
- $O_i^t$ merges with some other groups into $O_j^{t+1}$, if $O_j^{t+1} = match(O_i^t, \mathcal{G}^{t+1})$ and $\exists\ O_k^t \neq O_i^t$: $O_j^{t+1} = match(O_k^t, \mathcal{G}^{t+1})$.
- $O_i^t$ ceases, if none of the above scenarios happened.
- $O_j^{t+1}$ emerges, if $\forall\ O_i^t$: $O_j^{t+1} \neq match(O_i^t, \mathcal{G}^{t+1})$.

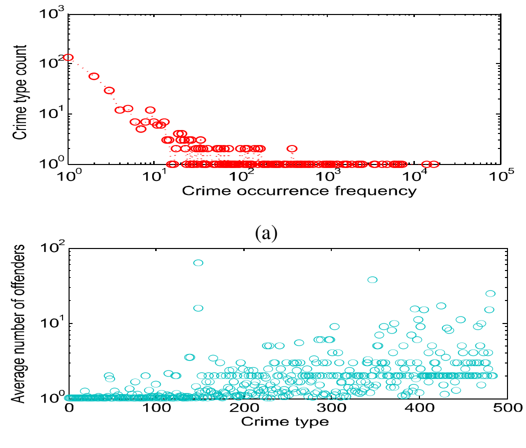
These rules are intuitive and easy to observe in the life cycle of groups, but they are not yet rigorous enough. The main problem lies in defining the threshold $\lambda$. This threshold needs to be determined based on experimentation and observation, for instance, by learning from existing histories for real-world organized crime groups.

## V. Experimental Results

In this section, we experimentally evaluate the proposed method on a large real-world crime dataset. We first describe basic features of the crime dataset used in our study before discussing different experiments and the obtained results.

### A. Crime Dataset

For the time period from 1/8/2001-31/7/2006, the crime dataset contains information about all reported offences (≈4.4 million) and all persons (offenders, victims, witnesses, etc.) associated with a crime incident (≈9 million referring to about 4 million unique individuals), 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. Out of 4 million subjects in the dataset, 250,492, 255,302, 190,406, and 228,792 respectively appear at least once as charged, chargeable, charge recommended or suspect. In our experiments, we restrict on the subjects in these four categories. Being in one of these categories means that the police were serious enough about the subjects involvement in a crime as to warrant calling them 'offenders'.




(a)

(b)

Figure 1: (a) Crime occurrence frequency; (b) Average number of offenders for each crime type

The extracted co-offending network comprises ≈150,000 nodes and ≈600,000 edges. The average node degree is four and the maximum degree is 525. About 50% of all the nodes have degree one, meaning these offenders have committed co-offences with only one offender in their criminal life. The largest connected network component links ≈18% of all the nodes together, which is fairly big for this kind of network.

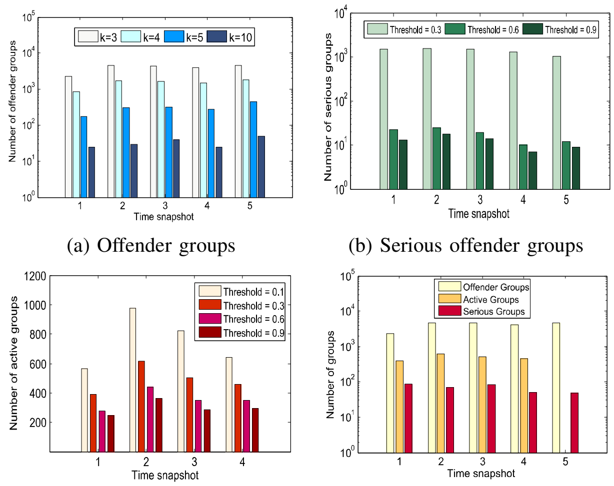
For the experiments, we have divided the dataset into five chronological snapshots, each of them representing 12 months. Considering only offences with more than one offender reduces these numbers to 9,943, 18,819, 18,350, 16,939 and 20,111. For the crime incidents with multiple offenders, we extract the respective co-offending network for each of the five snapshots.

Table I: 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 |

### B. Crime Seriousness

In total, about 1,000 different crime codes are defined in the dataset, each referring to a specific type of offence. For extracting the co-offending networks, we consider all of the codes (except for traffic related offences). In case of offences with more than one offender, only about 100 crime types have an occurrence frequency greater than 100, and only 30 crime types have an occurrence frequency greater than $1,000$. Figure 1a and Figure 1b respectively show the occurrence frequency and average number of involved offenders for each of these crime types.




(a) Offender groups (b) Serious offender groups

(c) Active offender groups (d) All groups

Figure 2: Number of different types of offender groups

For calculating offender group criminality, we apply the RCMP crime seriousness index, called OSR index. 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 indicator of the seriousness of offences. Table I shows a small sample of the OSR crime seriousness hierarchy and corresponding seriousness values.

### C. *Groups Characteristics*

In the extracted co-offending networks, one for each of the five time periods (snapshots), we identify offender groups, active offender groups, serious offender groups, and also possible organized crime groups.

Figure 2a shows the number of offender groups using different clique sizes $k$. As expected, the number of offender groups decreases by increasing the clique size. All the experiments discussed below are based on clique size $k = 3$.

Figure 2c shows the number of active groups in each time period for different *activity thresholds* $\alpha$. Interestingly, with $\alpha = 0.9$ still half of the offender groups remains in the list, meaning that some offender groups keep their collaboration completely intact and unchanged over longer time periods.

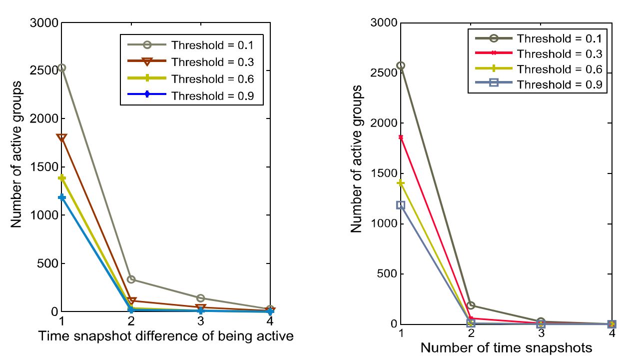
Active offender groups can be further characterized as continuously active over several consecutive snapshots or their activity is occurring at irregular intervals with inactive snapshots. Figure 3a shows the number of active groups observed over time periods with one, two, three, and four years difference. The important point here is that with increasing time difference the number of observed groups decreases exponentially. Even with very low values for $\alpha$ only few groups can be observed over four snapshots, and with high values for $\alpha$ 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 $\alpha$. This finding supports the theory of short-time collaborations of offender groups. According to Albanese [17], many organized crime groups are short-lived and comprised of offenders with desired skills who form temporary networks to take advantage of a crime opportunity. Albanese [17] mentions that these groups often dissolve after exploiting the opportunity, looking for new chances which may need other combinations of skills.

Another important aspect of active offender groups is the number of snapshots they were being active. This is illustrated in Figure 3b. Even with small $\alpha$, we do not have any offender group active in all time snapshots. With median $\alpha$, we observe only few offender groups which were active in three time snapshots. This shows that, due to reasons like incarceration or changing crime-committing tactics and trends, offender groups generally do not maintain their criminal activity for a long time period.

Figure 2b illustrates the number of serious offender groups for different criminality thresholds $\beta$. About 35% of all offender groups pass the threshold $\beta = 0.5$, which means a larger percentage of the offender groups commit minor crimes, which is intuitive. Finally, $\beta = 0.5$ results in 2% of the groups which implies that only very few offender groups are involved in serious crimes.

For considering a group active, we apply the activity threshold $\alpha = 0.3$, and for considering a group serious, we use the criminality threshold $\beta = 0.8$. All further discussions are based on these settings. Figure 2d shows the number of offender groups, active groups, and serious groups in each of the time steps.

Figure 4b shows the frequency of committed offences per group. A clear point is that serious groups commit less offences compared to offender groups and active groups.




(a) (b)

Figure 3: (a) Number of offender groups active in two time snapshots in respect to time difference; (b) Number of active offender groups by the number of observed time snapshot

### D. Groups Size and Cores and Peripheries Role

The work in [17] concludes that most organized crime groups are quite small. Our study corroborates this result. Figure 4a provides the size distribution for offender groups, active offender groups, and serious crime groups.

Average group size for both offender groups and active offender groups is about six, whereas it is only four for serious groups. 63% of the members of offender groups and active groups belong to the kernel. For serious groups this value increases to 92%. When comparing active offender groups to offender groups, a larger percentage of active offender groups has periphery members, and the average number of periphery members is slightly greater, which shows that the periphery members play a more important role in the structure of active offender groups. But the percentage of serious groups having periphery members compared to the other two group types is significantly smaller. This result implies that collaborations regarding serious crimes happen around the closest circle of offenders.

The maximum number of kernel and periphery members in active offender groups, compared to offender groups, is significantly smaller. Likewise, this is the case when comparing serious groups with the other two group types. 7% of the offender groups have more than 10 and only 0.7% have more than 50 members. In the active offender group set, 10% and 0.7% of them have more than 10 and 50 members, respectively. For serious groups, the number of groups with more than 10 members is less than one percent and there is no group with more than 50 members.

Figure 4c presents the distribution of size of overlap for offender groups and for active 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 offender groups means that many of 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 only little observable overlap, which again confirms their completely different structure compared to offender groups and active groups.

### E. Offender Group Evolution

Given the limited observable time span, it is difficult to quantify the whole lifecycle of offender groups, not knowing their previous history in the first time step and their future history in the last time step. Figure 4d 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, and a value greater than 0.2 and smaller than 0.3 for split and merged,

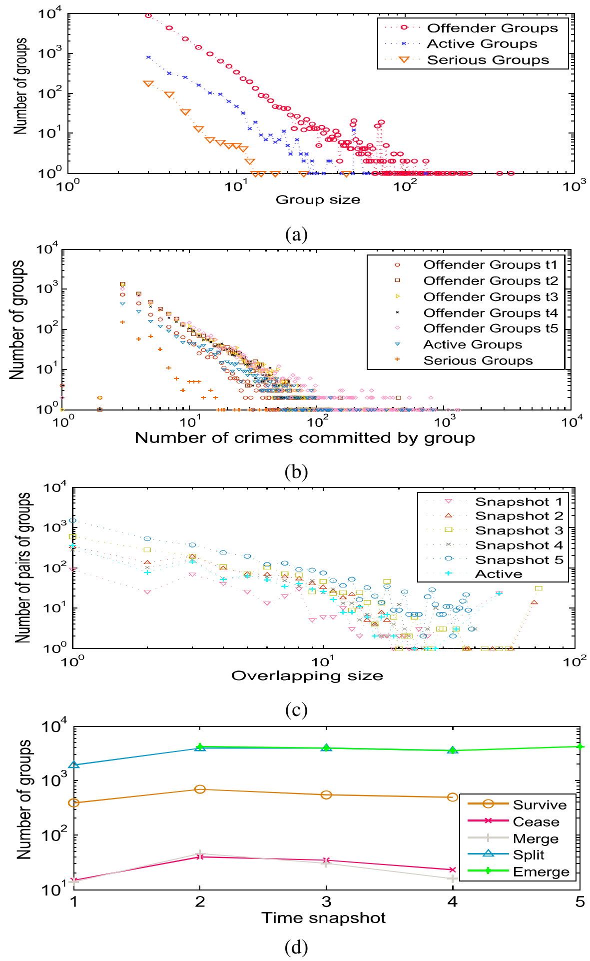



Figure 4: (a) Size of different types of groups; (b) Number of offender groups in respect to the number of crimes they committed; (c) Number of shared members for the overlapping groups by the number of pairs of groups; (d) Offender groups evolution trace

respectively. Groups with matching thresholds smaller than 0.2 are considered ceased groups. On average, about 14% of all offender groups survive, but split and merge events occur rarely, less than 1% of the groups. About 88% of the offender groups are considered ceased, since we do not observe their activity in the next time step, and 90% of all groups are newly emerged ones.

### F. Organized Crime Groups

Totally there are 35 groups that are both active and serious and thus are considered possible organized crime groups. Interestingly, most of these groups have very high activity, which shows the close relationship among their group members. The average size in this set of groups is 4.3, which is much smaller than the size of serious and active groups. This point supports the theory that with increasing group criminality offender group size decreases. Having only $0.4$ periphery members on average also states that in possible organized crime groups, the kernel members are not eager to collaborate with offenders outside of the group's kernel.

The final step of our experiments studies the crimes committed by possible organized crime groups. We concluded that for seven groups their offences were limited to sexual assault or criminal harassment—offences that may not associate material benefit. Six groups only committed murder-related crimes, again making a decision about material benefit more difficult, although they may be very severe organized crime groups. The other 22 groups committed a wide range of serious crimes, including kidnapping, drug trafficking, robbery with firearms, and so on. These groups are with high probability organized crime groups and their members fulfill their mission under this criminal group structure.

## VI. Conclusions

Controlling crime necessitates investigation of criminal networks, organized crime groups, criminal organizations and their illegal activities, constituting a notorious problem for law enforcement and criminal justice. We propose here a computational co-offending network analysis approach for detecting organized crime groups and evaluate the proposed methods by experiments on a large real-world crime dataset. Our experimental results show that although criminal group activity does not occur as routinely as other social activities, which is intuitive, there is continuous criminal collaboration inside crime groups. But for most of the groups such activities do not persist over longer time periods. Our experiments also show that active offender groups typically have more peripheral members in contrast to serious groups which tend to have less peripheral members. This finding suggests that serious groups operate primarily from inside their kernel.

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 $20,000$ offender groups, including about $1,800$ active groups. Eventually our experiment results in $35$ possible organized crime groups, with $63\%$ of them, based on the type of committed offences, most probably constitute organized crime groups.

Our analytic approach provides important insights into the ways in which co-offending networks shape and affect criminal behavior. 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 give the orders, may not be visible in the data. For obtaining a more holistic picture of criminal organizations, one may combine police-reported crime data with data from intelligence agencies. Further, the approach taken here primarily concentrates on organized crime groups with dense member relationships, which is not always the case, especially not for certain forms or criminal networks.

In our future work, we will explore ways to determine and visualize the organizational structure of organized crime groups, differentiating between hierarchical organization and decentralized organization, and also analyze how these structures evolve over time. We will also continue our close collaboration with major law enforcement agencies.

## References


[1] C. Morselli, Inside Criminal Networks. Studies of Organized Crime, Vol. 8, Springer, 2009.

[2] M. A. Tayebi and U. Glässer. Organized Crime Structures in Co-offending Networks. Proc. Intl. Conf. on Social Computing and its Applications, Sydney, Australia, Dec. 2011.

[3] G. Palla, I. Derenyi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature 435, 2005.

[4] Klaus von Lampe, Definitions of Organized Crime. Last visited, Feb. 2012 [Online]. Available: http://www.organized-crime.de/organizedcrimedefinitions.htm.

[5] P. L. Brantingham, M. Ester, R. Frank, U. Glässer and M. A. Tayebi, Co-offending Network Mining. In U. Kock Wiil (ed.), Counterterrorism and Open Source Intelligence, Lecture Notes in Social Networks, Vol. 2: 211-239, Springer, 2011.

[6] A. J. Reiss, Co-offending and criminal careers. Crime and Justice: A Review of Research, 1988.

[7] A. J. Reis, D. P. Farrington, Advancing knowledge about co-offending: Results from a prospective longitudinal survey of London males. J. Criminal Law & Criminology 82 (2), 1991.

[8] J. M. McGloin and H. Nguyen. The Importance of Studying Co-Offending Networks for Criminological Theory and Policy. In Proc. Third Annual Illicit Networks Workshop, Montréal, Québec, October 2011.

[9] J. M. McGloin et al., Investigating the stability of co-offending and co-offenders among a sample of youthful offenders. Criminology 46 (1), 2008.

[10] Alan A. Block, East SideWest Side: Organizing Crime in New York City, 19301950, 2nd ed., New Brunswick, NJ: Transaction 1994.

[11] T. Van Der Heijden, Measuring Organized Crime in Western Europe. In Milan, Pagon (eds.) Policing in Central and Eastern Europe: Comparing First Hand Knowledge with Experience from the West, Slovenia: College of Police and Security Studies, 1996.

[12] International Report on Crime Prevention and Community Safety: Trends and Perspectives, 2010. International Centre for the Prevention of Crime, Montreal, Québec, Canada, 2010.

[13] C. Fijnaut, F. Bovenkerk, G. Bruinsma and H. van de Bunt, Organized Crime in the Netherlands, The Hague: Kluwer Law International, 1998.

[14] M. E. J. Newman. Fast algorithm for detecting community structure in networks, Phys. Rev. E 69, 066133, 2004.

[15] D. Chakrabarti, R. Kumar, and A. Tomkins. Evolutionary clustering. In Proc. of the 12th ACM SIGKDD Conf., 2006.

[16] M. Spiliopoulou et al. Monic: modeling and monitoring cluster transitions. Proc. 12th ACM SIGKDD Intl. Conf. on Knowledge discovery and data mining, ACM New York, NY, USA, 2006.

[17] J. Albanese. North American Organised Crime. Global Crime vol. 6 no 1, 2004.