2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)

CRIMETRACER: Activity Space Based Crime Location Prediction

Mohammad A. Tayebi, Martin Ester, Uwe Glässer
School of Computing Science
Simon Fraser University, B.C., Canada
Email: {tayebi, ester, glaesser}@cs.sfu.ca

Patricia L. Brantingham
School of Criminology
Simon Fraser University, B.C., Canada
Email: pbrantin@sfu.ca

***Abstract*—Crime reduction and prevention strategies are vital for policymakers and law enforcement to face inevitable increases in urban crime rates as a side effect of the projected growth of urban population by the year 2030. Studies conclude that crime does not occur uniformly across urban landscapes but concentrates in certain areas. This phenomenon has drawn attention to spatial crime analysis, primarily focusing on *crime hotspots*, areas with disproportionally higher crime density. In this paper we present CRIMETRACER, a personalized random walk based approach to spatial crime analysis and crime location prediction outside of hotspots. We propose a probabilistic model of spatial behavior of known offenders within their *activity space.* Crime Pattern Theory concludes that offenders, rather than venture into unknown territory, frequently commit opportunistic crimes and serial violent crimes by taking advantage of opportunities they encounter in places they are most familiar with as part of their activity space. Our experiments on a large real-world crime dataset show that CRIMETRACER outperforms all other methods used for location recommendation we evaluate here.**

***Keywords*-Predictive policing; Random walk model; Activity space; Crime occurrence space; Co-offending networks**

I. INTRODUCTION

Urban population is expected to grow globally from $2.86\,bn$ in the year 2000 to $4.98\,bn$ by 2030 [1]. A major concern in many countries is spiking crime rates in urban areas along with the projected growth of urban population. Policymakers and law enforcement agencies will inevitably face enormous challenges deploying notoriously scarce resources even more efficiently to apprehend criminals, disrupt criminal networks and effectively deter crime by investing in crime prevention and reduction strategies for urban areas.

In recent years criminologists have remarked the role of spatial analysis in crime prevention [2], [3], [4], [5], [6]. Different studies conclude that crime does not occur uniformly across urban landscapes [3], [4], [6]. Crime *hotspots*, areas with high crime intensity, generate a larger percentage of criminal events [3]. Understanding why hotspots emerge in some places rather than others may be a difficult question, identifying these places is not [3], [4]. Hotspot analysis enables law enforcement to better prioritize and use their resources more effectively for crime reduction and prevention. At the same time, there is lack of attention to areas with low crime density, referred to as *coldspots*, although these areas overall account for approx. half of all urban crimes [6].

A 16-year longitudinal study of crime in Seattle, concludes that roughly half of the yearly crime incidents occur within only five to six percent of the city's road segments [6]. Given that coldspots cover a much wider area than hotspots, targeted policing is not feasible for crime prevention in these areas. Spatial distribution of crime incidents in coldspots is essential for intervention strategies. While in hotspot analysis the focus is on modeling the emergence and evolution of the hotspots, in the coldspot analysis for detecting offenders' spatial behavior we need a model which is flexible to be personalized for every single offender.

Existing studies analyzing crime distribution mostly focus on prediction models for future crime locations and time intervals [7]. Typically these studies rely exclusively on hotspot models to identify clusters of incidents in crime intensive areas. These models do not consider preferences of offenders and routine decision making for crime location selection.

Crime Pattern Theory [2] concludes that offenders, rather than venture into unknown territory, frequently commit opportunistic crimes and serial violent crimes by taking advantage of opportunities they encounter in places they are most familiar with as part of their *activity space* [2], which includes the most frequently visited places as determined by a person's daily routine activities, such as commuting patterns. The rational is that, outside of their activity space, offenders will have to hunt for criminological opportunities and more likely face uncertain or unforeseeable risks. Hence, *crime occurrence space* is with a high probability the part of activity space that intersects with the location of targets preferred by an offender.

Random walk models appropriately model how offenders encounter criminal opportunities, given that the behavior of a random walk model is basically local [8]. In [8], the authors propose a random walk based model for capturing the dynamics of hotspot formation. We present here an extended random walk model, CRIMETRACER, for generating the activity space associated with offenders living in an urban area. In CRIMETRACER, random walk process is personalized to uncover the unknown spatial behavior of every single offender.

We personalize the random walk process using co-offending information, crime trends of offenders, and also the crime event history of road segments. For the urban layout we assume a small-scale *road network* on which an offender traverses an urban area. By doing so, we gradually compute an approximation of the offender's activity space by reflecting the probability of visiting (and possibly committing crime) for each road segment of the urban area. This result is then used for predicting crime locations for individual offenders, something not addressed in crime spatial analysis to the best of our knowledge. Based on our experimental evaluation, personalization is successful for detecting crime locations in coldspots. The extended random walk model outperforms the random walk model and the other evaluated methods in terms of the recall and precision metrics.

Section 2 addresses related work, and Section 3 explains key characteristics of crime data. Section 4 describes the CRIME-TRACER model. Next, Section 5 presents our experimental evaluation and the results. Section 6 concludes the paper.

II. BACKGROUND

Criminology theories state that involvement in crime is the result of: 1) an individual's crime propensity, and 2) criminogenic features of the environment to which an individual is exposed. While propensity towards crime has long been studied, in the past few decades criminogenic features of the environment received specific attention, and it is concluded that environmental criminology plays an essential role in crime reduction and prevention tactics [9]. New research areas emerge, like crime mapping [10], geographic profiling [5] and crime forecasting [11], [12].

Human cognition, spatial decision-making and human movements help to describe the activities of individuals—a way of thinking derived from Quetelet [13]. People do not move randomly across urban landscapes [14]. For the most part, they commute between a handful of routinely visited places like home, work, recreational facilities, and favorite shopping centers. With each and every trip, they will get more familiar with, and gain new knowledge about, these places and everything along the way. Eventually, a person will be at ease with a place. At this point, the place becomes part of the person's activity space, illustrated in Figure 1. Activity space has two main components: Nodes and Paths [2]. The Nodes, called activity Nodes, are the locations that the person frequents, such as a work place, residence, or recreation. These are the end-points of the journey. The Paths, called activity Paths, connect the Nodes and represent the person's path of travel between them.

Several studies have explored the activity space of offenders. Rossmo [5] has developed a widely recognized method of inferring the activity space of an offender to determine the likely home location based on the person's crime locations. In a method known as *Geographic Profiling*, he assumes that offenders will select targets and commit crimes near their home address. Using this assumption, each new crime location is plotted on a map and a distance-decay function is used to calculate a probability space around each crime to denote the possible home location (and corresponding probability) of the offender. Geographic Profiling narrows down the probable home location of an offender more accurately with increasing number of crimes attributed to the offender.

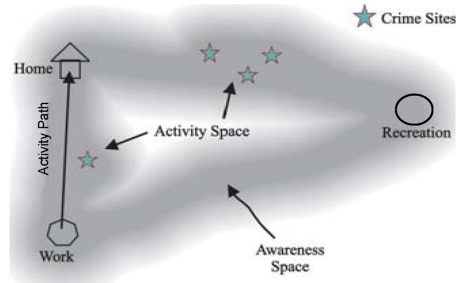


Fig. 1: Activity Space

Canter [15] splits movement patterns of offenders into commuters and marauders. Marauders use a fixed base location (home, for example) and commit their crimes around it, making geographic profiling on this type of offender possible. According to Canter, and consistent with Crime Pattern Theory, marauders derive their offending locations from spatial patterns of their non-criminal daily activities. Although commuters probably also have a consistent base location, they travel to other places to commit crimes. Such travel patterns must be taken into account, making geographic profiling much more difficult.

Frank [16] proposed an approach to infer the activity Paths of all offenders in a region based on their crime and home locations. Assuming the home location as the center of an offender's movements, the orientation of Activity Paths of each individual offender were calculated so as to determine the major directions, relative to their home location, into which they tended to move to commit crimes.

Based on criminological theories, several studies propose mathematical models of spatial and temporal characteristics of crime to predict future crimes. However, these models do not predict individual offender behavior. For instance in [12], the authors use a point-pattern-based transition density model for crime space-event prediction. This model computes the likelihood of a criminal incident occurring at a specified location based on previous incidents. In [8], the authors model the emergence and dynamics of crime hotspots. This work uses a two-dimensional lattice model for residential burglary, where each location is assigned a dynamic attractiveness value, and the behavior of each offender is modeled with a random walk process. The authors study the impact of the model parameters on hotspot formation using a computer simulation.

In our own work, we address investigative problems such as suspect investigation [17] and organized crime group detection [18]. Given partial information about a crime incident, CrimeWalker [17] is an unsupervised method for top-k suspect recommendation which applies a random walk method on a co-offending network. In [18], we present a social network analysis based approach to identify traces of possible criminal organizations in operational police records. While the main focus of existing methods is predicting crime at the aggregate level, CRIMETRACER models offender activity space to predict crime at the individual level.

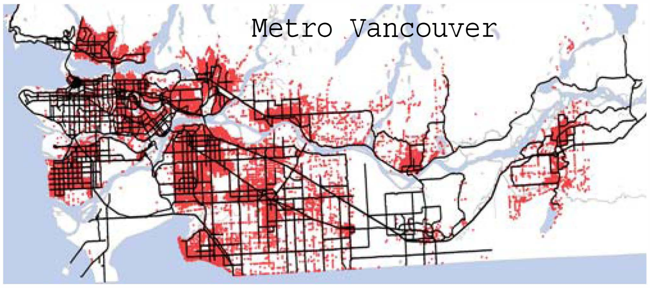


Fig. 2: Crime distribution in Metro Vancouver; red dots show crime locations, and black lines show the major roads.

Property	Value
No. of crimes	125,927
No. of offenders	189,675
No. of offenders with more than one crime	25,162
No. of co-offending links	68,577
No. of co-offenders in co-offending network	17,181
Avg. node degree in co-offending network	4
No. of road segments	64,108
Avg. Crime Per Road Segment	2

TABLE I: Statistical properties of the used dataset

Note that all of the above methods solve related but different problem so that our experiments cannot compare with them. The model presented in [12] only predicts the time and location of the crime in the aggregate level. For a different purpose but similar to our work [8] uses standard random walk to model offenders criminal behavior. In our evaluation we consider the standard random walk model. The method proposed in [5] and [15] discover offender home locations based on his crime locations. And finally the output of the method proposed in [16] is locations which are centers of interest for committing crime. However we compare CRIME-TRACER to different *Collaborative Filtering* methods which are used for location recommendation in location-based social networks [19], [20]. Collaborating filtering (CF) infers users' implicit preference form the explicit opinions of similar users based on the idea that users with similar behavior in the past will have similar behavior in the future [21].

III. CHARACTERISTICS OF CRIME DATA

Crime data mining, as analytic tool, has enormous potential for law enforcement, criminal intelligence agencies, and beyond, to facilitate crime investigations by increasing efficiency and reducing mistakes. On the other hand, access to and sharing of crime data is subject to many restrictions and can even be a national security concern because of the highly sensitive nature and related personal information.

As a result of a research memorandum of understanding between ICURS[1] and "E" Division of Royal Canadian Mounted Police (RCMP) and the Ministry of Public Safety and Solicitor General, five years of real-world crime data was made available for research purposes. This data was retrieved from the Police Information Retrieval System (PIRS), a large database keeping police records for the regions of the Province of British Columbia (BC) which are policed by the RCMP.

For a time period of five years (2001-2006), the BC police arrest dataset amounts to approx. 4.4 million crime records, one for each reported crime incident. This includes all persons associated with a crime, such as offenders (from complainant to charged), victims, witnesses and bystanders, overall, 39 different subject (person) groups. In our experiments, we consider all subjects in four main categories: *charged, chargeable, charge recommended or suspect*. Being in one of these categories means that the police were serious enough about a subjects involvement in a crime as to warrant calling them 'offenders'. For the study presented here, we concentrate on the crimes in Metro Vancouver, where different regions are connected through a road network composed of 64,108 road segments with an average length $0.2\,km$. Table I shows the statistics of the used crime dataset. Figure 2 shows the spatial distribution of crimes in Metro Vancouver.

Figures 3a and 3b illustrate the distribution function of crime incidents per offender and per road segment. Both distributions have heavy-tailed pattern. 83% of the offenders committed only one crime, while less than 1% of the offenders committed 10 or more crimes. Further, 38% of the road segments are linked to at least one crime and 9% are linked to ten or more crimes. Half of all the crimes occurred in only 1% of all road segments, and a total of 25% in only 100 road segments.

Figures 4a and 4b respectively show the average home location to crime location distance and the average distance between crime locations for an offender. The average home to crime location distance of 80%, 63% and 40% of all offenders is less than 10 km, 5 km and 2 km, respectively. And the average crime location distance of 73%, 52% and 26% of all offenders is less than 10 km, 5 km and 2 km, respectively. One can assume that frequent offenders are generally mobile and have several home locations identified in their records. 41% of the offenders who committed more than one crime have more than one home location.

The dataset differentiates more than 1,000 crime types, with half of them occurring only a few times. For three well defined categories of personal crime (like assault), property crime and drug crime, as expected, the property crime category has the largest average home location to crime location distance. For half of repeat offenders at least half of their crimes belong to only one category, meaning that half of the repeat offenders are experts in at least one crime category and they keep their crime trend for a while.

IV. PROBLEM CHARACTERIZATION

In this section we present the probabilistic activity space problem and define the notation.

[1]The Institute for Canadian Urban Research Studies (ICURS) is a university research center at Simon Fraser University.

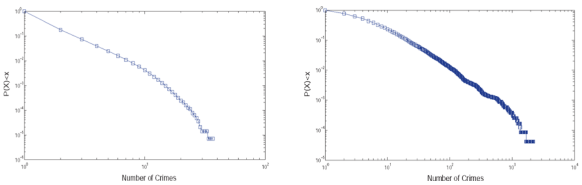


Fig. 3: Distribution function: crimes per offender and per road segment (a) Crimes per offender; (b) Crimes per road segment

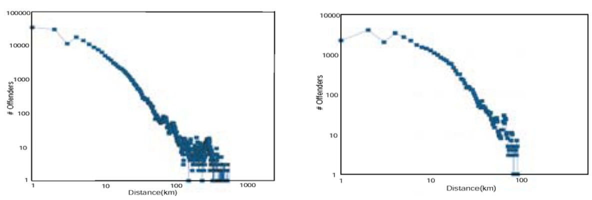


Fig. 4: Avg. distance (a) home-crime; (b) crime-crime

A. Notation

Intuitively, a road network consists of *road segments*, each of which starts and ends at an intersection. We use the *dual* representation where the role of roads and intersections is reversed. All physical locations along the same road segment are mapped to the same node. Formally, a road network is an undirected graph $R(L, Q)$, where L is a set of nodes, each representing a single road segment. Road segments l_j and l_k are connected, $\{l_j, l_k\} \in Q$, if they have a common intersection.

Let V be a set of offenders and C be a set of crimes. Each crime $c_i \in C$ involves a non-empty subset of criminal offenders $X \subseteq V$. With each crime incident we associate a type of crime, a date and time when the crime occurred as well as longitude and latitude coordinates of the crime location and home location of all offenders known to be involved in the crime. Crime locations within a studied geographic boundary are mapped to the closest road segment. Henceforth, the term "road" is used to refer to a road segment.

$\bar{y}_j$ is a vector denoting the features of the road l_j including road length d_j, and *road attractiveness* features vector $\bar{a}_j$. $\bar{a}_j$ is a vector of size m where the value of the kth entry of $\bar{a}_j$ corresponds to the total number of crimes of type k committed previously at l_j. Π_j denotes the set of neighbors of road l_j in the road network.

Anchor locations. L_i is the set of roads at which offender u_i has been observed, including all of his known home and crime locations. $f_{i,j}$ and $t_{i,j}$ respectively denote the frequency and the last time being u_i at anchor location l_j. *Offender trend* is given by a vector $\bar{x}_i$ of size m which indicates the crime trend of u_i as extracted from his criminal history. That is, the value of the k^{th} entry of $\bar{x}_i$ corresponds to the number of crimes of type k committed by offender u_i.

A *co-offending network* $G(V, E)$ is an undirected graph. Each node represents a known offender $u_i \in V$. Offenders u and v are connected, $u_i, u_k \in V$ and $(u_i, u_k) \in E$, if they are known to have committed one or more offences together, and are not connected otherwise. A co-offending network is derived from a crime dataset referring to reported crime incidents over a time period [22]. Γ_i denotes the set of neighbors of offender u_i in the co-offending network.

B. Problem Definition

Given a crime dataset C, an offender u_i associated with C, and an underlying road network $R(L, Q)$, the goal is to learn the *activity space distribution* F for u_i on R. That is, for any road $l_j \in L$, the activity space distribution $F(u_i, l_j)$ states the probability that l_j is part of the activity space of u_i, and thus the likelihood for offender u_i committing a crime at road l_j:

$$F(u_i, l_j) \longrightarrow [0, 1] \text{ with } \sum_{j=1}^{|L|} F(i, j) = 1 \tag{1}$$

V. CRIMETRACER MODEL

A *random walk* over a graph is a stochastic process in which the initial state is known and the next state is decided using a transition probability matrix that identifies the probability of moving from a node to another node of the graph. Under a certain condition the random walk process converges to a stationary distribution which assigns an importance value to each node of the graph.

The random walk method satisfies the locality aspect of crimes which states that offenders do not attempt to move far from their anchor locations. But it has some shortcomings that we aim to address in our proposed model, CRIMETRACER.

The CRIMETRACER model consists of three important components: an offender, the road network including all locations where the offender commits crime and the co-offending network that connects offenders. Starting from an anchor location, the offender explores the city through the underlying road network. At each road he decides whether to proceed to a neighboring road or return to one of his anchor locations. The random walk process continues until it converges to the steady state which reflects the probability of visiting a road by the offender. This probability can be relevant to the offender's exposure to a crime opportunity.

For learning the activity space of an offender we need to understand his daily lives and routines, but in the crime dataset generally we miss the Paths completely and the Nodes partially (refer to Section II), which is a major challenge. To address these challenges, we improve our incomplete knowledge about offenders with the available information in the dataset. The set of anchor locations of each offender is extended by adding his co-offenders' anchor locations, denoted as *main anchor locations*. This extension is motivated by the assumption friends in the co-offending network are likely to share the same location.

For each offender using a Gaussian model we also define his *intermediate anchor locations* as the roads closest to the set of main anchor locations. An offender starts his random walk either from a main anchor location or from an intermediate anchor location.

Since we are not aware of offenders' trajectories in his journey to crime we bias offenders movement to the locations with higher chance of crime commitment. Therefore the transition probabilities on the road network depend on the road features value, i.e. the number of crimes committed in each road and the road length. Different crime categories, such as drug related and property related crimes, reflect the crime opportunity of each road. We argue that different crime categories have different weights in offenders decision making toward finding crime opportunity. Using the *supervised random walks* method [23] we learn the importance of these features, and use them in computing transition probabilities of the random walk. The road length feature guides an offender in his random walk to select the next road when all neighbors of the current road are empty of crime.

The offender in his random walk stops in a road which provides opportunity for his crime trend and commit a crime. This is detected by comparing the offender crime trend and road attractiveness. Below we describe different elements of the proposed model in more details.

A. Random Walk Process

For each single offender, we perform a series of random walks on the road network $R(L, Q)$ to detect the probability of being at a road. In each random walk the offender starts his exploration from one of his anchor locations, traversing the road network to locate a criminal opportunity.

For offender u_i, the random walk process starts from one of his anchor locations with predefined probabilities as described in Section V-B. At each step k of the random walk, the offender is at a certain road l_j and makes one of two possible decisions:

- with probability α he decides to return to an anchor location, and start looking for a criminal opportunity another time. He can decide about the anchor location in two ways:
  - with probability β he decides to return to a main anchor location $l \in \mathcal{L}_i$.
  - with probability $1-\beta$ he returns to an intermediate anchor location $l \in \mathcal{I}_i$.
- with probability $1-\alpha$ he continues his attempt for finding a criminal opportunity.

If he continues his random walk then he has two options:

- with probability $\theta(u_i, l_j, k)$ stop the random walk, which means the offender commits a crime at road l_j.
- with probability $1-\theta(u_i, l_j, k)$ continue the random walk, moving to another road segment which is a direct neighbor of l_j.

To continue the random walk at road l_j, we need to select one of its direct neighbor roads Π_{l_j}. The probability of selecting road segment l_k in the next step is defined as:

$$P(l_j \to l_k) = \frac{\phi_{\bar{w}}(l_k)}{\sum\limits_{l_p \in \pi_{l_j}} \phi_{\bar{w}}(l_p)} \tag{2}$$

The function $\phi_{\bar{w}}(l_j)$ is computed based on the road features:

$$\phi_{\bar{w}}(l_j) = \sum_{k=1}^{m+1} w_k \times y_{j,k} \tag{3}$$

where $\bar{y}_{j,k}$ is the value of kth feature of the road l_j, and w_k is the corresponding weight of the feature k. The probability of being at road l_r at step $k+1$ given that the walker was at road l_j at step k is:

$$\begin{aligned} P(X_{l,k+1} = l_r | X_{l,k} = l_j) &= \\ (1-\alpha)(1-\theta_{l_j,k}) \times P(l_j \to l_r) &= \\ (1-\alpha)(1-\theta_{l_j,k}) \times \frac{\phi_{\bar{w}}(l_r)}{\sum\limits_{l_s \in \Pi_{l_j}} \phi_{\bar{w}}(l_s)} \end{aligned} \tag{4}$$

We need to terminate the overall process when we have done enough random walks. We terminate the random walks when $||F^{m+1}|| - ||F^m|| \leq \epsilon$ where $F^m = \begin{pmatrix} F(u_i, l_1) \\ \vdots \\ F(u_i, l_{|L|}) \end{pmatrix}$ is the results for u_i after m random walks. For some offenders the random walks do not converge in which we terminate the overall process when $m > 10000$.

B. Starting Probabilities

In the CRIMETRACER model for each offender we have two different types of starting nodes:

- Main anchor locations. Main anchor locations are all of offender's anchor locations and his co-offenders' anchor locations: $\mathcal{L}_i = L_i \cup \{l_j : l_j \in L_v, v \in \Gamma_u\}$. Co-offending links are important since they are the reasons for many spatial effects related to crime [24]. It is concluded that offenders who are socially close they are spatially close too [25]. We share co-offenders' anchor locations with this intuition that offenders who collaborated together previously will share their location in the future too. In computing the starting probability of each anchor location two primary factors are frequency and time of offender being in an anchor location. The probability that offender u_i starts his random walk from l_j is:

$$S(i,j) = \frac{f_{i,j} \times e^{\frac{-(t-t_{i,j})}{\rho}}}{\sum\limits_{l_k \in \mathcal{L}_i} f_{i,k} \times e^{\frac{-(t-t_{i,k})}{\rho}}} \tag{5}$$

  where t is the current time and ρ is the parameter for controlling the effect of the time.
- Intermediate anchor locations. In the absence of information about offenders Paths we derive intermediate anchor locations which are the closest to the set of main anchor

locations. Human mobility models use Gaussian distribution to analyze human movement around a particular point such as home or work location [26], [27]. We assume that offender movement around his main anchor locations follows a Gaussian distribution. Each main anchor location of offender u_i is used as the center, and the probability of being u_i in a road is modeled with a Gaussian distribution. Given road l the probability of being u_i at l is computed as following:

$$S(i,l) = \sum_{l_j \in \mathcal{L}_i} \frac{f_{i,j}}{\sum_{l_k \in \mathcal{L}_i} f_{i,k}} \frac{\mathcal{N}(l|\mu_{l_j}, \Sigma_{l_j})}{\sum_{l_k \in \mathcal{L}_i} \mathcal{N}(l|\mu_{l_k}, \Sigma_{l_k})} \quad (6)$$

Here l is a road which does not belong to the set of main anchor locations. $\mathcal{N}(l|\mu_{l_j}, \Sigma_{l_j})$ is the probability distribution function of Gaussian distribution, with μ_{l_j} and Σ_{l_j} as mean and covariance, of visiting a road when u_i is at anchor location l_j. We affect the normalized activity frequency of u_i at l_j, meaning that a main anchor location with higher activity frequency has higher importance. For offender u_i the roads with the highest probability of being intermediate anchor location are added to the set $\mathcal{I}_i$ as augmented starting nodes beside the main anchor locations.

C. Learning Road Feature Weights

In this section we discuss how we learn road feature weights $\bar{w}$ which are used to compute the transition probabilities. We use the same idea used in the supervised random walks method [23] for link prediction in social networks. This method guides the random walks towards the preferred target nodes by utilizing node and edge attributes.

Each offender in a random walk staring from his home location he reaches a crime location. In the training data for each offender we have a series of crime journeys, meaning that for a source node s we have a set of destination nodes $D = \{d_1, d_2, \ldots, d_n\}$, and a set of non-destination nodes $Q = \{q_1, q_2, \ldots, q_m\}$. The probability of visiting a node p_d is influenced by the road transition probabilities. And the transition probabilities are dependent to the road features weight. Now we tend to give a bias to the offender that starting from s he visits destination nodes $d_i \in D$ more than other non-destination nodes $q_i \in Q$, formulating the following optimization problem:

$$\min_{\bar{w}} F(\bar{w}) = \|\bar{w}\|^2 + \lambda \sum_{d \in D, q \in Q} loss(p_q - p_d) \quad (7)$$

where λ is the regularization parameter, and $loss$ is a predefined loss function for penalizing the cases that the stationary probability of a non-destination node p_q is higher than the stationary probability of a destination node p_d.

D. Stopping Criteria

The probability of stopping the random walk for an offender at a given road corresponds to the probability of this offender committing a crime in that road segment. Two factors influence the stopping probability of offender u_i in the road l_j. The first one relates to the similarity of the crime trend of offender u_i and the criminal attractiveness of road l_j, where higher similarity means a higher chance that u_i's random walk stops at l_j. The second factor is the distance of l_j from starting point measured in the number of steps from the starting point. To satisfy the locality aspect of crimes the probability of continuing the random walk should decrease while getting farther from the starting point:

$$\theta(u_i, l_j, k) = Sim(i,j) \times \frac{1}{1 + e^{\frac{-k}{2}}} \quad (8)$$

where Sim(i,j) denotes the cosine similarity of crime trend of u_i and the road attractiveness of the road l_j. The stopping probability is inversely proportional to the step number k:

$$Sim(i,j) = \frac{\bar{x}_i . \bar{a}_j}{|\bar{x}_i||\bar{a}_j|} \quad (9)$$

VI. Experimental Evaluation

In this section we describe our experiments and their results.

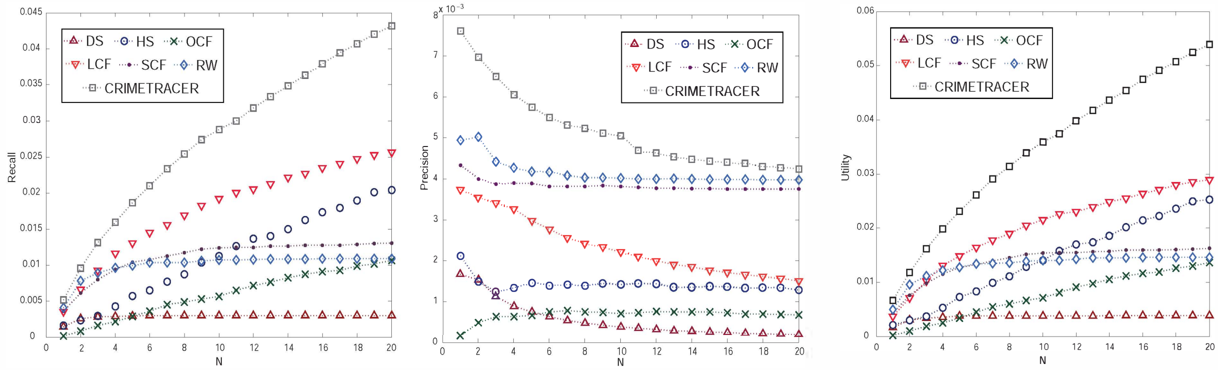
A. Experiential Design

For each offender first we order his crime events chronologically based on their time. Then we split these records to training and test set. The first 80% of the crimes are used for training the model which generates the offender activity space. The crime locations of the remaining 20% are used for testing the model. In the evaluation process we only consider offenders with at least two different crime locations which includes about 10% of the offenders in the crime dataset. We note that the training data used for learning road features as described in Section V-C is not included in the evaluation to prevent biasing CrimeTracer.

Learning offender activity space should result in predicting his future crime location more successfully. We need to emphasize that the test set of each offender only includes new locations and we exclude the home and crime locations in the training set. After learning the offender activity space the top-N roads with the highest probability are suggested as the most probable places for offender to commit his next crime.

As discussed the focus of this work is modeling offender behaviors in the coldspots. Thus, in our experiments we exclude the top 100 roads with the highest crime number, the hotspots. The number of crimes in these hotspot roads is 100 to 1100 times greater than the average number of crimes in a road. In the evaluation we distinguish two groups of offenders: repeat offenders with 10 or more crimes and non-repeat offenders with less than 10 crimes.

To evaluate the accuracy of activity space detection, we measure the number of crimes committed by each offender in his testing dataset among the top N predicted locations. For an offender if crime location in his test set is also in the learned activity space, that crime location is correctly *predicted*. Three measures, *precision*, *recall* and *utility*, are used for this purpose:



(a) Recall (b) Precision (c) Utility

Fig. 5: Performance for different numbers of predictions

- Recall computes the ratio of the number of correctly predicted crime locations (true positive) to the number of crime locations of the offender in the test set (true positive + false negative).
- Precision computes the ratio of correctly predicted crime locations (true positive) to the number of all predictions N (true positive + false positive).
- Utility computes the percentage of offenders with at least one correctly predicted crime location.

Recall and precision are averaged across all offenders to determine the overall performance for different values of N. In computing the precision value for an offender, if the activity space contains $M < N$ roads, we use M instead of N.

B. Evaluated Methods

As discussed in Section II, there is no related work that solves the same problem of offender activity space problem. However we use the below methods which are equivalent to sate-of-the-art methods for location recommendation [20]:

Random Walk. This is the standard random walk with restart method (RWR) [28].

Hotspots. Using the basic hotspot approach (HS), roads are ranked based on the number of crimes in that road.

Proximity. In the proximity approach (DS) we rank the roads based on their distance from offenders' anchor locations. Here distance denotes the length of the path between two roads on the road network.

Offender-based CF. The intuition behind the offender-based CF approach (OCF) is that offenders who had similar behavior in the past will have similar behavior in the future. Let $b_{ij} = 1$ if $l_j \in \mathcal{L}_i$, and $b_{ij} = 0$ if $l_j \notin \mathcal{L}_i$. Now $F(i, j)$ is the probability of crime committing in road l_j by u_i:

$$F(i,j) = \frac{\sum_{u_k \in V \wedge k \neq i} Sim(i,k).b_{k,j}}{\sum_{u_k \in V \wedge k \neq i} Sim(i,k)} \quad (10)$$

where $Sim(i, k)$ denotes the cosine similarity measure between offenders u_i and u_k:

$$Sim(i,k) = \frac{\sum_{l_j \in L} b_{i,j}.b_{k,j}}{\sqrt{\sum_{l_j \in L} b_{i,j}^2}\sqrt{\sum_{l_j \in L} b_{k,j}^2}} \quad (11)$$

Location-based CF. In location-based CF (LCF) we consider the similarity of locations instead of offenders' similarity:

$$F(i,j) = \frac{\sum_{l_k \in L \wedge k \neq j} Sim(j,k).b_{i,k}}{\sum_{l_k \in L \wedge k \neq j} Sim(j,k)} \quad (12)$$

where $Sim(j, k)$ is the cosine similarity measure between roads l_j and l_k:

$$Sim(j,k) = \frac{\sum_{u_i \in V} b_{i,j}.b_{i,k}}{\sqrt{\sum_{u_i \in V} b_{i,j}^2}\sqrt{\sum_{u_i \in V} b_{i,k}^2}} \quad (13)$$

Co-offending-based CF. Co-offenders can share their information about criminal opportunities, and take advantage of this information in committing a new crime. Co-offending-based CF (SCF) computes the probability of crime committing in road l_j by u_i as following:

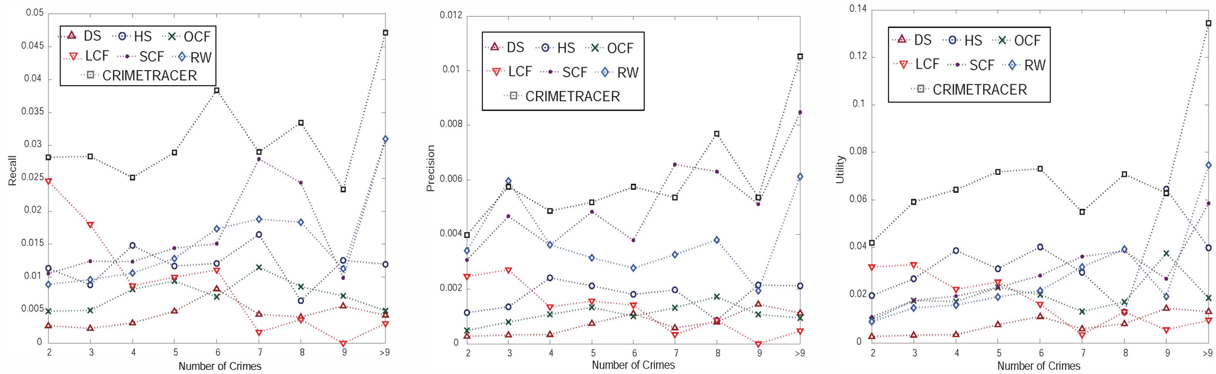
$$F(i,j) = \frac{\sum_{u_k \in \Gamma_i} Sim(i,k).b_{k,j}}{\sum_{u_k \in \Gamma_i} Sim(i,k)} \quad (14)$$

$Sim(i, k)$ is the geo-social influence between u_i and u_k:

$$Sim(i,k) = \frac{|\Gamma_i \cap \Gamma_k|}{|\Gamma_i \cup \Gamma_k|} + \frac{|\mathcal{L}_i \cap \mathcal{L}_k|}{|\mathcal{L}_i \cup \mathcal{L}_k|} \quad (15)$$

C. Results

Figure 5 shows the overall performance of the different evaluated methods in terms of recall, precision and utility. Our proposed method, CrimeTracer, always outperform all other methods for all values of N in terms of recall, precision and utility metrics, where N is the number of roads with



(a) Recall (b) Precision (c) Utility

Fig. 6: Performance for the offenders with different number of crimes (N=20)

highest likelihood of being in an offender activity space.

DS obtains the lowest precision and recall values. Despite the well-studied theory, the relation between crime commitment and distance from anchor locations, this result shows this approach is not effective for crime prediction. Among the CF based approaches OCF has the weakest result. LCF works better in terms of recall, but SCF has higher precision. It is interesting that location similarity can contribute more than offender similarity in crime prediction. One can conclude that SCF uses more reliable but limited information for detecting offenders activity space. The recall of HS improves with increasing N, but this method naturally is strong in predicting crimes in hotspots and not in coldspots.

Predicting even one crime location of each offender is very important for the critical task of crime prevention. As for the other two measures, CRIMETRACER outperforms the other methods in terms of utility. The utility is only 1% larger than the recall ($N = 20$), making no significant difference. One reason is that half of the offenders committed only two crimes, and we can predict only one location for them, meaning that for these offenders the recall and utility results are the same. But for the repeat offenders the utility is more than two times greater than the recall, 6% compared to 13%.

There has long been interest in the behavior of repeat offenders since controlling these groups of offenders can reduce the overall crime level significantly. Figure 6 depicts the performance of the different methods for offenders with different number of crimes. We expect more successful activity space learning for the offenders who have committed more crimes, and for whom we have more information. We can observe such trend only in CRIMETRACER, RWR and SCF results, and not for the other methods. For instance in CRIMETRACER the average recall for offenders who committed only two crimes is about 4% while this value increases to 7% for offenders who committed 10 or more crimes. Using co-offending information in SCF causes a significance improve for repeat offenders who have higher co-offending rates.

Non-repeat offenders are the majority of offenders. For instance in this study half of the offenders used for the

Method	Recall	Precision	Utility
RWR	0.011	0.004	0.014
RWR + Road features weight	0.013	0.003	0.017
RWR + Additional anchor points	0.019	0.001	0.024
RWR + Stopping criteria	0.036	0.003	0.045
CRIMETRACER	0.043	0.004	0.054

TABLE II: Contribution of different elements of CRIMETRACER to its performance ($N = 20$)

evaluation committed only two crimes. As shown in Figure 6 for non-repeat offenders CRIMETRACER outperforms the other methods. As depicted LCF works very well for offenders who committed only two crimes. This interesting result shows that beginner offenders tend to commit crimes in common locations. On the other hand, we see while SCF is not strong for beginners but with increasing crime numbers its performance increases significantly. This means that being more experienced in crime boosts the number of co-offenders and consequently the chance of sharing criminal opportunities.

We studied the contribution of different elements of CRIMETRACER to its performance. CRIMETRACER has three additional elements compared to the standard random walk with restart including additional anchor locations(co-offending information and intermediate anchor location), learning feature weights and stopping criteria. We added these elements separately to RWR to determine their contribution. Table II shows the results. The strongest element is stopping criteria and the weakest is learning road feature weights. The main idea behind the stopping criteria is stopping random walk of an offender in a road where crime history is similar to the offender crime trend. However combining all elements together in CRIMETRACER achieves the best result and improves the recall and utility of RWR by a factor of 4.

The overall performance of CRIMETRACER is comparable to the state of the art of the location recommendation methods [20], [29] in which the information about users spatial patterns is much denser compared to the available information about offenders. One may criticize that in location recommendation the exact locations are predicted while in CRIMETRACER only roads are predicted as offender activity space. However, as discussed in [7] roads are the natural domain for many policing activities, and a more realistic urban element for predicting a crime than the exact latitude and longitude. In addition, the road network we use in our study is in the micro scale with the average road length of 0.2 km.

VII. CONCLUSIONS

Advanced crime data analysis methods, linking data mining algorithms with social network analysis, can provide a scientific foundation for developing effective strategies by analyzing spatial decision-making of offenders and their social standing.

Modeling activity space of individual offenders is one of the most difficult problems in human mobility modeling because of limited available information on offenders and their dynamically changing complex behavioral patterns. CRIMETRACER uses a personalized random walk to derive a *probabilistic activity space* model for known offenders based on given facts. We evaluate our algorithm by data mining operational police records from crimes in Metro Vancouver within a 5-year time period. We are not aware of any similar work for modeling offender activity space and, hence, compare the proposed approach with location recommendation methods. CRIMETRACER outperforms all other evaluated methods tested here. It boosts the prediction performance of the repeat offenders, compared to the non-repeat offenders, by using co-offending information, where the chance of having co-offending links is higher for the repeat offenders.

All elements used in CRIMETRACER, which are additional to the standard random walk model, contribute to the performance of this method. Still, there is room for further improvement. Given the importance of anchor locations to start the random walk, we take into account the frequency and time of being at these locations. Exploring this aspect in more depth remains future work. Crime attractors [30], main areas where crimes tend to concentrate, are key factors in committing crime. Detecting such areas and using them in CRIMETRACER is another direction for future work to further improve the method performance.

We believe the ideas presented here can inspire new research trends in social network analysis and data mining with useful applications for predictive policing, criminal investigations and criminal intelligence in the endeavor to combat crime.

ACKNOWLEDGEMENTS

We are thankful to RCMP "E" Division and BC Ministry for Public Safety and Solicitor General for making this research possible by providing Simon Fraser University with crime data from their Police Information Retrieval System. We further wish to thank the ICURS institute at Simon Fraser University for their support of this research. Finally, we also like to thank the anonymous reviewers for their helpful comments and suggestions.

REFERENCES

[1] United nations settlements programme, the state of the world's cities 2004/2005: Glabalization and urban culture. London: Earthscan, 2004.

[2] P. L. Brantingham and P. J. Brantingham. Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, 13(1):3–28, 1993.

[3] L. W. Sherman, P. R. Gartin, and M. E. Buerger. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56, 1989.

[4] A. A. Braga. The effects of hot spots policing on crime. *The ANNALS of the American Academy of Political and Social Science*, 578(1):104–125, 2001.

[5] D.K. Rossmo. *Geographic Profiling*. CRC Press, 2000.

[6] E. R. Weisburd, D. L.and Groff and Yang S. *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. Oxford University Press, 2012.

[7] Davies T. P. and S. R. Bishop. Modelling patterns of burglary on street networks. *Crime Science*, 2(1), 2013.

[8] M. B. Short, M. R. D'orsogna, V. B. Pasour, G. E. Tita, P. J. Brantingham, A. L. Bertozzi, and L. B. Chayes. A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, 18(supp01):1249–1267, 2008.

[9] P. J. Brantingham and P. L. Brantingham. *Environmental criminology*. Sage Publications, 1981.

[10] K. Harries. Mapping crime principle and practice. *Washington D.C, U.S. Department of Justice Office of Justice Programs*, 1999.

[11] W. Gorr and R. Harries. Introduction to crime forecasting. *International Journal of Forecasting*, 19(4):551–555, 2003.

[12] H. Liu and D. E. Brown. Criminal incident prediction using a point-pattern-based density model. *International journal of forecasting*, 19(4):603–622, 2003.

[13] L. A. J. Quetelet. *A Treatise on Man and the Development of His Faculties*. Edinburgh: W. and R. Chambers, 1842.

[14] J. Glyde. Localities of crime in suffolk. *Journal of the Statistical Society of London*, pages 102–106, 1856.

[15] D. V. Canter and A. Gregory. Identifying the residential location of rapists. *Journal of the Forensic Science Society*, 34(3):169–175, 1994.

[16] R. Frank and B. Kinney. How many ways do offenders travel – evaluating the activity paths of offenders. In *Proc. EISI'12*, 2012.

[17] M. A. Tayebi, M. Jamali, M. Ester, U. Glässer, and R. Frank. Crimewalker: a recommendation model for suspect investigation. In *Proc. RecSys'11*, 2011.

[18] U. Glässer, M. A. Taybei, P. L. Brantingham, and P. J. Brantingham. Estimating possible criminal organizations from co-offending data. *Public Safety Canada*, 2012.

[19] M. Ye, P. Yin, and W. Lee. Location recommendation for location-based social networks. In *Proc. SIGSPATIAL GIS'10*, 2010.

[20] H. Wang, M. Terrovitis, and N. Mamoulis. Location recommendation in location-based social networks using user check-in data. In *Proc. SIGSPATIAL GIS'13*, 2013.

[21] B. N. Miller, J. A. Konstan, and J. Riedl. Pocketlens: Toward a personal recommender system. *ACM Transactions on Information Systems (TOIS)*, 22(3), 2004.

[22] P. L. Brantingham, M. Ester, R. Frank, U. Glässer, and M. A. Tayebi. Co-offending network mining. *Counterterrorism and Open Source Intelligence*, pages 73–102, 2011.

[23] L. Backstrom and J. Leskovec. Supervised random walks: predicting and recommending links in social networks. In *Proc. WSDM'11*, 2011.

[24] D. R. Schaefer. Youth co-offending networks: An investigation of social and spatial effects. *Social networks*, 34(1):141–149, 2012.

[25] M. A. Tayebi, R. Frank, and U. Glässer. Understanding the link between social and spatial distance in the crime world. In *Proc. SIGSPATIAL GIS'12*, 2012.

[26] M. C. Gonzalez, C. A. Hidalgo, and A. Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

[27] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439(7075):462–465, 2006.

[28] H. Tong, C. Faloutsos, and J. Pan. Fast random walk with restart and its applications. In *Proc. ICDM'6*, 2006.

[29] B. Liu, Y. Fu, Z. Yao, and H. Xiong. Learning geographical preferences for point-of-interest recommendation. In *Proc. KDD '13*, 2013.

[30] P. L. Brantingham and P. J. Brantingham. Criminality of place. *European Journal on Criminal Policy and Research*, 3(3), 1995.