Personalized crime location prediction

MOHAMMAD A. TAYEBI, UWE GLÄSSER, MARTIN ESTER and PATRICIA L. BRANTINGHAM

European Journal of Applied Mathematics / Volume 27 / Special Issue 03 / June 2016, pp 422 - 450
DOI: 10.1017/S0956792516000140, Published online: 28 April 2016

Euro. Jnl of Applied Mathematics (2016), *vol.* 27, *pp.* 422–450.
doi:10.1017/S0956792516000140

Personalized crime location prediction

MOHAMMAD A. TAYEBI[1], UWE GLÄSSER[1], MARTIN ESTER[1]
and PATRICIA L. BRANTINGHAM[2]

[1]*School of Computing Science, Simon Fraser University, B.C., Canada*
email: tayebi@cs.sfu.ca, glaesser@cs.sfu.ca, ester@cs.sfu.ca
[2]*School of Criminology, Simon Fraser University, B.C., Canada*
email: pbrantin@sfu.ca

(*Received 31 December 2014; revised 10 February 2016; accepted 29 February 2016*)

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[1], a personalized random walk-based approach to spatial crime analysis and crime location prediction outside of hotspots. We propose a probabilistic model of spatial behaviour of known offenders within their *activity spaces*. Crime Pattern Theory concludes that offenders, rather than venture into unknown territory, frequently select targets in or near 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.

Key words: Spatial crime analysis, Random walk model, Activity space, Crime occurrence space, Co-offending networks.

1 Introduction

The spatial analysis of crime is re-emerging in importance [7, 9, 35, 36, 38, 40, 52]. Studies find that crime does not occur uniformly or randomly across the urban landscapes [4,7,27,38,52]. Crime *hotspots*, areas with high crime intensity, generate a larger percentage of criminal events [38]. From the criminological perspective, the best known study of hotspots and coldspots is a 16-year longitudinal study of crime in Seattle, WA in the United States. It finds that roughly half of the yearly crime incidents occur within only 5% to 6% of the city's road segments [52]. Understanding why hotspots emerge in some places and not in others is a challenging question [5–7, 38]. But the concentration of interest on hotspots pulls attention away from better understanding areas with more moderate or low concentration of criminal events. These areas can be referred to as *coldspots*. Better

[1] This work is an extension of a previously published conference paper [44].

understanding coldspots is of value because these areas account for approximately half of all urban crimes [52].

In the work reported here, the term "coldspot" refers to any location not included in any hotspot. Coldspots cover a much wider area than hotspots. Better understanding the spatial distribution of crime incidents in coldspots is essential. In hotspot analysis, the focus is on modelling the emergence, evolution and stability of hotspots. Such analysis is often based on analysis of aggregate crime patterns for all offenders. Coldspot analysis, as is explored in this paper, requires modelling of individual offenders' spatial behaviour. This is fundamentally different from hotspot analysis because the concentration and rate of offending in hotspots is so high that the patterns of individual offending in these areas are averaged out since collective spatial behaviour of offenders is rather modelled at an aggregate level. What is needed are models that are flexible and can be personalized to individual offenders.

Existing models of crime distribution mostly focus on models for predicting crime locations for time intervals [15]. These studies rely heavily on modelling hotspot emergence, bifurcation and diffusion [29, 39] to identify clusters of incidents in crime intensive areas. These models frequently use concepts of crime attractiveness that pull people towards locations [10, 18, 39, 40, 42]. There is some research that explores offender spatial decision making using decision theory [4] and the tendency of offenders to commit offences near prior offences [26]. But the models tend to use one decision making process for all potential offenders.

The model presented in this paper focuses on individualized offending and target selection or target decision making with the decision rules being different for the occasional offender and the frequent repeat offender and repeat co-offender. The model is derived from Crime Pattern Theory [9] based on the assumption that offenders, rather than venture into unknown territory, frequently commit crimes in places they are most familiar with as part of their *activity space* [9]. Surrounding an activity space, an individual develops visually, and through local information, an awareness of the surrounds. An activity space is a subset of this awareness space. Activity spaces and awareness spaces change over time with movement to new home locations, new employment, the development of new shopping and entertainment areas, new friendship networks and the development of new mass transit and roads. But, as noted before, crime is relatively rare, and acceptable targets of crime or victims are likely to be found easily within an awareness space. Outside an activity or awareness space, an offender will have to consciously search for criminal opportunities and likely face higher uncertain or unforeseeable risks. *Crime occurrence space* is more likely a part of an activity space that intersects with the location of suitable targets preferred by an offender [8, 9].

Human cognition, spatial decision-making and human movements help to describe the activities of individuals—a way of thinking that has a long history in neurology, geography and psychology [2, 3, 14, 21, 27, 33, 50]. People do not move randomly across urban landscapes [22]. For the most part, they commute between a handful of routinely visited places like home, work, recreational facilities and favourite shopping centres. With each 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. Activity space has two main components: Nodes and Paths [9]. 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. Activity paths connect the nodes and represent the persons path of travel between them.

The creation of the main attractor nodes and paths are developed through normal mobility shaped by the urban backcloth or urban environment. Each individual has normal, routine pathways or commuting/mobility routes that are unique. However, the environment in which we live influences our actions and movements. Highways, streets and road networks in general guide us to our destinations like home, work place, recreation centre and business establishments. For those using public transit, the transportation routes lead us to our destinations in limited and predictable way.

In the aggregate, these individual routes overlap or intersect in time and space. These areas of overlap often have rush hours and congestion at intersections or mass transit stops associated with handling large numbers of people. These high activity locations can become crime attractors and crime generators when there are enough suitable targets in those locations. Crime attractors and generators affect directionality of offenders' movement [10, 16–18, 42]. Offenders develop routine mobility patterns and routine alternate routes. In many ways, they follow the same process in their mobility as non-offenders in the urban environment. Their spatial awareness is formed based on their destinations and transportation routes, and potential targets located in frequently visited places may attract them.

The method proposed here is designed and optimized for coldspot analysis at the level of individual offenders and significantly outperforms baseline hotspost methods for the purpose of crime location prediction in coldspots. While we include information about crimes linked to individual offenders in hotspot areas in the training process, we exclude crime location prediction for hotspot areas in the evaluation process to illustrate the advantage of the proposed method in comparison with the baseline methods.

Our method uses random walk to model how offenders encounter criminal opportunities at a local level near anchor locations in an activity space [41]. In [41], the authors propose a random walk-based model for capturing the dynamics of hotspot formation (See [41] for a Levy Flight model). We present here an extended random walk model, CrimeTracer, for generating the activity space associated with offenders living in an urban area. In CrimeTracer, the random walk process is personalized to uncover the spatial behaviour of all individual offenders.

For the urban layout, we assume a small-scale road network on which an offender moves about in an urban area. By doing so, we gradually compute an approximation of the offenders 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.

The following sections of the paper address related work, key characteristics of crime data used in the analysis, details of the CrimeTracer model, our experimental model evaluation and the results.

2 Crime location prediction

Criminology theories state that involvement in crime is the result of: (1) an individual's crime propensity, and (2) features of the environment to which an individual is exposed. While propensity towards crime has long been studied, in the past few decades features of the environment received specific attention, and it is concluded that environmental criminology plays an essential role in developing crime reduction and prevention tactics that consider individual offenders [8]. New research areas have emerged, like crime mapping [25], geographic profiling [36] and crime forecasting [23, 31].

Several studies have explored the activity space of offenders. Rossmo [36] 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. His approach is most often used for serious serial offenders and known as Geographic Profiling. He assumes that offenders will select targets and commit crimes near their home address or another major activity node or anchor point. 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/nodal location of an offender more accurately with increasing number of crimes associated to the offender.

Canter [13] 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 [19] 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 centre 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 behaviour. For instance, in [31], 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 [41], 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 behaviour of each offender is modelled 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 [47], identifying key players in co-offending networks [43], organized crime group detection [20] and co-offence prediction [45]. Given partial information about a crime incident, CrimeWalker [47] is an unsupervised method for top-k suspect recommendation, which applies a random walk method to a co-offending network. In [43], we study key player identification and removal for criminal network disruption. We also propose a framework for co-offence prediction using supervised learning [45]. In [20], 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 locations at the aggregate level, CrimeTracer models offender activity space to predict crime locations at the individual level.

Note that all of the above-mentioned methods solve related but different problems to which the experiments presented here can not be compared. The model presented in [31] only predicts the time and location of the crime in the aggregate level. For a different purpose but similar to our work, [41] uses standard random walk to model offenders' criminal behaviour. The method proposed in [36] and [13] discover offender home locations based on his crime locations. And finally, the output of the method proposed in [19] is locations which are centres of interest for committing crime. To provide a meaningful comparison, we compare CrimeTracer to different *Collaborative Filtering* methods which are used for location recommendation in location-based social networks [51, 55]. Collaborating filtering (CF) infers the user's implicit preference form the explicit opinions of similar users based on the idea that users with similar behaviour in the past will have similar behaviour in the future [32].

3 Data model and characteristics

In this section, we describe a unified model of crime data [11,20,43,45,46,48] for specifying in a concise and unambiguous way properties of interest in the analysis of criminal networks and their constituent entities. This model aims at bridging the conceptual gap between data level, mining level and interpretation level, and facilitates the separation of the abstract description of crime data from the details of data mining and analysis.

3.1 Data model

Intuitively, the crime data model is the starting point for the extraction of networks of offenders who have committed crimes together, called *co-offending networks*. Connections between offenders play a central factor in the analysis of crime. We also explain here the abstract representation of urban environments in terms of graph structures defined on segmentations of road networks. We then link road segments to known offenders and their crimes.

3.1.1 *Crime data*

Formally, we represent the logical organization of crime data and information associated with a crime dataset $\mathscr{C}$ as a finite graph structure in the form of an attributed tripartite *hypergraph* $\mathscr{H}(\mathscr{N},\mathscr{E})$ with a set of nodes $\mathscr{N}$ and a set of hyperedges $\mathscr{E}$. The set $\mathscr{N}$ is composed of three disjoint subsets, $A = \{a_1, a_2, \ldots, a_q\}$, $I = \{i_1, i_2, \ldots, i_r\}$ and $R = \{r_1, r_2, \ldots, r_s\}$, respectively representing *actors* like offenders, suspects, victims, witnesses and bystanders; *incidents* referring to reported crime events; and *resources* used in a

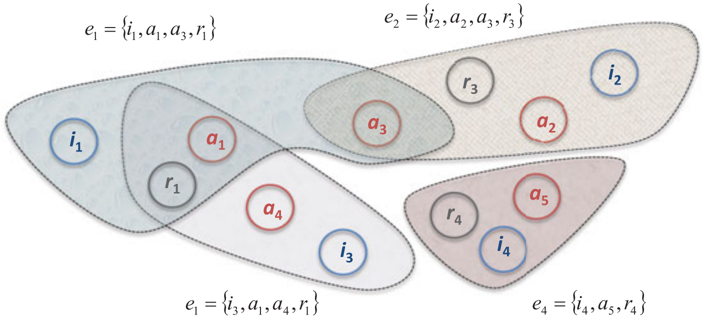


FIGURE 1. Hypergraph $\mathscr{H}$ (without attributes) for a simple crime data model $\mathscr{C}$.

crime[2], such as weapons, tools, mobile phones, vehicles and bank accounts. Generic actors serve as placeholders if a person's identity remains unclear, say an unrecognized offender who evaded apprehension. Whenever no specific resource can be identified or has been reported, the distinguished element "*unknown*" is used as a placeholder.

A hyperedge e of $\mathscr{E}$ is a non-empty subset of nodes $\{n_1, n_2, \ldots n_p\} \subseteq \mathscr{N}$ such that the following three conditions hold: $|e \cap I| = 1$, $|e \cap A| \geqslant 1$ and $|e \cap R| \geqslant 1$. For any $e, e' \in \mathscr{E}$ with $e \cap I = e' \cap I$, it follows that $e = e'$. Intuitively, a hyperedge e of $\mathscr{H}$ associates a set of actors $\{a_{i_1}, a_{i_2}, \ldots, a_{i_j}\} \subseteq A$ and a set of resources $\{r_{i_1}, r_{i_2}, \ldots, r_{i_l}\} \subseteq R$ with a crime incident $i_k \in I$, where $e = \{i_k, a_{i_1}, a_{i_2}, \ldots, a_{i_j}, r_{i_1}, r_{i_2}, \ldots, r_{i_l}\}$ as illustrated in Figure 1.

Finally, with each node $n \in \mathscr{N}$, we associate some non-empty list of attributes characterizing the entity represented by n. Attributes of actors, for instance, include the name, address and contact details, and the criminal profile information of known offenders; while attributes of incidents include the crime type, the time of the incident, longitude and latitude coordinates of the crime location and the role of each person identified in connection with the incident, among various other types of data and information.

Co-offending Network. Co-offending networks constitute a widespread form of social networks that is of considerable interest in crime investigations and in the study of crime. As Reiss contends in [34], "understanding co-offending is central to understanding the etiology of crime and the effects of intervention strategies". We explain the extraction of a co-offending network from crime data in detail in our previous work [11].

3.1.2 *Urban environment*

Intuitively, a road network can be decomposed into *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

[2] Resources used in a crime do often provide essential clues in criminal investigations. For uniformity of representation, we assume that R includes a distinguished element *nil* referring to situations in which no specific resource can be identified.

and l_k are connected, $\{l_j, l_k\} \in Q$, if they have an adjacent intersection in common. 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.

Road features. A vector $\bar{y}_j$ denotes the features of the road l_j including road length d_j, and *road attractiveness* features vector $\bar{a}_j$. Further, $\bar{a}_j$ 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 neighbours of road l_j in the road network. $\Delta \subset L$ denotes a set of roads with the highest crime rate, called crime *hotspots*. D_{l_j,l_k} is the shortest path distance of road l_j from hotspot $l_k \in \Delta$, and f_j denotes the total number of crimes at road l_j.

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 u_i was 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 kth entry of $\bar{x}_i$ corresponds to the number of crimes of type k committed by offender u_i.

3.2 Data characteristics

Crime data mining, as an analytic tool, has enormous potential for law enforcement, criminal intelligence agencies, and beyond, to facilitate crime investigations by increasing efficiency and reducing mistakes. At the same time, access to and sharing of crime data is subject to strictly enforced restrictions arising from security and privacy needs because of the highly sensitive nature and related personal information.

3.2.1 *BC crime data*

As a result of a research memorandum of understanding between ICURS[3] and "E" Division of Royal Canadian Mounted Police and the Ministry of Public Safety and Solicitor General, 5 years of real-world crime data for the regions of the Province of British Columbia policed by the Royal Canadian Mounted Police has been made available for research purposes.

The British Columbia police arrest dataset used in this study amounts to approx. 4.4 million crime records, each referring to a single incident reported between July 2001 and June 2006. The information provided for each incident includes all persons associated with a crime, such as offenders, victims, witnesses and bystanders—overall, 39 different subject (person) groups. In our experiments, for offenders, we consider all subjects in four main categories: *charged, chargeable, charge recommended or suspect*. Being in one of these categories in fact means that the police were serious enough about a subject's involvement in a crime as to warrant calling them "offenders".

For the study presented here, we concentrate on the use a subset of this dataset which includes all crimes in Metro Vancouver, B.C. (total population: over 2.4 million), where different regions are connected through a road network composed of 64,108 road

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

Table 1. *Statistical properties of the dataset used in this study*

Property	Value
Number of crimes	125,927
Number of offenders	189,675
Number of offenders with more than one crime	25,162
Number of co-offending links	68,577
Number of co-offenders in co-offending network	17,181
Average node degree in co-offending network	4
Number of road segments	64,108
Average crime per road segment	2

segments with an average length 0.2 km. Table 1 shows a statistics for the used crime dataset.

3.2.2 *Characteristics of crime locations*

Figures 2(a) and (b) respectively show the average Euclidean distance between home location to crime location and the average Euclidean distance between crime locations for all offenders in the dataset. 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. In fact, 41% of the offenders who committed more than one crime have more than one home location.

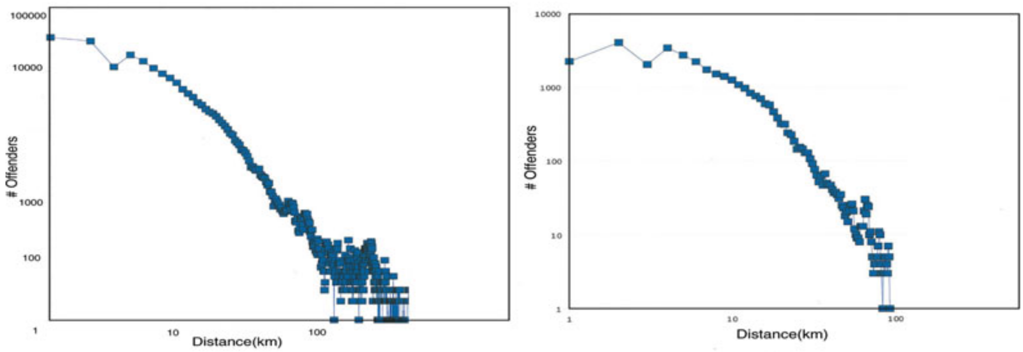
The dataset differentiates more than 1,000 crime types, with half of them occurring only a few times. For three well-defined categories of personal crime (like assault), property crime (like break & enter) and drug crime, as expected, the property crime category has the largest average home location 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 specialize in at least one category, and they keep their crime trend for a while.

4 CrimeTracer model

In this section, we present CrimeTracer, our proposed crime spatial analysis model, starting with the problem characterization.

Given a crime dataset $\mathscr{C}$, an offender u_i and road network $R(L, Q)$ associated with $\mathscr{C}$, the goal is to learn the *activity space distribution* F using the random walk model for u_i on R. That is, for each road $l_j \in L$, $F(i, j)$ 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 is

$$F(i, j) \longrightarrow [0, 1] \text{ with } \sum_{j=1}^{|L|} F(i, j) = 1. \tag{4.1}$$



(a) (b)

FIGURE 2. Avg. distance (a) home-crime locations; (b) crime-crime locations.

By learning the activity space distribution of individual criminal offenders, we obtain a probabilistic model of offender activity space that can be used for personalized prediction of future crime locations of the offender. The assumption is that the richer and more detailed the offender profile is, the more accurate is the probabilistic activity space model, and also the prediction of future crime locations. This probabilistic view of activity space means that there is no sharp boundary between activity space and awareness space, which directly corresponds to the intuitive understanding of the concept of activity space in criminology.

4.1 Model description

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 certain conditions, the random walk process converges to a stationary distribution [24], which assigns an importance value to each node of the graph.

The random walk method can be modified in a way that satisfies the locality aspect of crimes, which states that offenders do not attempt to move far from their anchor locations. For instance, the random walk method works locally if the likelihood of terminating the walk increases with the distance from the anchor locations. But it has some shortcomings that we aim to address in the CRIMETRACER model.

The CRIMETRACER model consists of three important components: an offender, the road network, including all locations where the offender committed 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 neighbouring 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 life and routines, but in the crime dataset generally we miss the Paths completely and the Nodes partially (refer to Section 1), 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. The extended set is called *main anchor locations*, denoted by $\mathscr{L}_i$ for offender *i*. This extension is motivated by the assumption that friends in the co-offending network are likely to share the same location.

For each offender, using a Gaussian model, we define his *intermediate anchor locations* as the roads closest to the set of his main anchor locations, denoted by $\mathscr{I}_i$ for offender *i*. An offender starts his random walk either from a main anchor location or from an intermediate anchor location.

Given that the actual trajectories in an offender's journey to crime are not known, the model guides offender movements in directions with a higher chance of committing a crime. This is done by taking into account two different aspects that influence offender movement directionality in computing the transition probabilities in a random walk. The first aspect refers to road characteristics in terms of road feature values: the number of crimes committed on this road for each different crime category and the road length. The second aspect refers to the personal preferences of each offender for certain types of crime, as stated in the offender profile, as a driving factor in the decision process when encountering a crime opportunity. Whenever none of the neighbours of the current road promise any crime opportunities of interest, road length is the single determining factor. Using the *supervised random walks* method [1], we learn the importance of these features and exploit them in computing the transition probability matrix for the random walk.

The second proposed approach for learning the movement directionality of an offender uses the concept of crime generators and attractors [10]. These are two types of locations where crimes tend to cluster. Assuming that offenders are drawn in directions leading toward criminal attractors, we assign a higher probability to roads leading toward crime hotspots—much like gravity centres affecting the random walk.

The random walk always stops in a road which provides an opportunity for committing a crime, depending on both road characteristics and offender crime preferences. Below we describe different elements of the proposed model in detail.

4.2 Random walk process

For each single offender, we perform a series of random walks on the road network $R(L,Q)$. 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 4.3. 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 not look for a criminal opportunity this time, choosing an anchor location in one of two ways:
  - with probability β, he decides to return to a main anchor location $l \in \mathscr{L}_i$.
  - with probability $1-\beta$, he returns to an intermediate anchor location $l \in \mathscr{I}_i$.

- with probability $1-\alpha$, he continues looking for a crime opportunity.

If he continues his random walk, then he has two options in each step of the walk:

- 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 which is a direct neighbour of l_j.

To continue the random walk at road l_j, we select a direct neighbour road from Π_{l_j}. The function ϕ is used in computing the transition probability from a road segment to one of its neighbour road segments. The way ϕ is computed is described in the next section. The probability of selecting road segment l_r in the next step is defined as

$$P(l_j \to l_r) = \frac{\phi(l_r)}{\sum\limits_{l_p \in \pi_{l_j}} \phi(l_p)}. \tag{4.2}$$

The probability of being at road l_r at step $k+1$ given that the offender was at road l_j at step k is

$$P(X_{u_i,k+1} = l_r | X_{u_i,k} = l_j) = (1-\alpha)(1-\theta_{l_j,k}) \times P(l_j \to l_r) = (1-\alpha)(1-\theta_{l_j,k}) \times \frac{\phi(l_r)}{\sum\limits_{l_s \in \Pi_{l_j}} \phi(l_s)}. \tag{4.3}$$

$X_{u_i,k}$ is the random variable for u_i being at road l_r in step k.

We terminate the random walks when $||F^{m+1}|| - ||F^m|| \leqslant \epsilon$, 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 case we terminate the overall process at $m > 10,000$.

4.3 Starting probabilities

CrimeTracer distinguishes two different types of starting nodes:

- Main anchor locations are all anchor locations of a single offender and his co-offenders: $\mathscr{L}_i = L_i \cup \{l_j : l_j \in L_v, v \in \Gamma_u\}$. Co-offending links are important since they are the reasons for many spatial effects related to crime [37]. It is concluded that offenders who are socially close, are spatially close too [46]. The rationale is that offenders who have collaborated in the past likely may have shared information on anchor locations in their activity space, an aspect that possibly affects their choice of future crime locations. In computing the starting probability of each anchor location, the two primary factors are the frequency and the last time an offender visited an anchor location. The probability that offender u_i starts his random walk from l_j thus is

$$S(i,j) = \frac{f_{i,j} \times e^{\frac{-(t-t_{i,j})}{\rho}}}{\sum\limits_{l_k \in \mathscr{L}_i} f_{i,k} \times e^{\frac{-(t-t_{i,k})}{\rho}}}, \tag{4.4}$$

where t is the current time, and ρ is the parameter controlling the effect of the timing.

- Intermediate anchor locations are the closest locations to main anchor locations. Human mobility models use Gaussian distribution to analyse human movement around a particular point such as home or work location [12, 22]. 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 centre, and the probability of u_i being located in a road is modelled with a Gaussian distribution. Given road l the probability of u_i residing at l is computed as follows:

$$S(i,l) = \sum_{l_j \in \mathscr{L}_i} \frac{f_{i,j}}{\sum\limits_{l_k \in \mathscr{L}_i} f_{i,k}} \frac{\mathscr{N}(l|\mu_{l_j}, \Sigma_{l_j})}{\sum\limits_{l_k \in \mathscr{L}_i} \mathscr{N}(l|\mu_{l_k}, \Sigma_{l_k})}. \tag{4.5}$$

  Here, l is a road which does not belong to the set of main anchor locations. $\mathscr{N}(l|\mu_{l_j}, \Sigma_{l_j})$ is a Gaussian distribution for visiting a road when u_i is at anchor location l_j, with μ_{l_j} and Σ_{l_j} as mean and covariance. We consider 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 an intermediate anchor location are added to the set $\mathscr{I}_i$ as additional starting nodes besides the main anchor locations.

4.4 Movement directionality

As discussed in Section 1, directionality of offender movement plays an important role in activity space formation. We propose here two approaches to determining movement directionality. The first approach learns the weights of the features that determine the probability of selecting a road among all neighbour roads in a random walk process. The second approach leads an offender in the direction that gets him closer to the crime hotspots.

4.4.1 *Hotspots influence*

In this approach, the transition probability is computed based on proximity of a road to the crime hotspots and the importance of each crime hotspot, which is proportional to the number of crimes committed there. The function $\phi(l_j)$ is used in computing the transition probability (refer to Section 4.2) of moving offender u_i from l_k to l_j:

$$\phi(l_j) = \sum_{n=1}^{|\Delta|} f_n \times \frac{1}{D_{j,n}}, \tag{4.6}$$

where f_n is the number of crimes committed at l_n. $D_{j,n}$ is the distance of road l_j from the hotspot $l_n \in \Delta$, which is equal to the length of shortest path between two roads on the road network.

4.4.2 *Learning road feature weights*

Road feature weights $\bar{w}$ are used to compute the transition probabilities. The function $\phi_{\bar{w}}(l_j)$ is computed based on the road features

$$\phi_{\bar{w}}(l_j) = \sum_{k=1}^{m+1} w_k \times y_{j,k}, \tag{4.7}$$

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.

We use the same idea used in the supervised random walks method [1] for link prediction in social networks. This method guides the random walk toward the preferred target nodes by utilizing node and edge attributes.

Each offender in a random walk starting from his home location 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 $Z = \{z_1, z_2, \ldots, z_m\}$. The probability of visiting a node p_d is influenced by the road transition probabilities. And the transition probabilities are dependent on the road features weight. Now, we sway an offender starting from node s so as to visit destination nodes $d_i \in D$ more often than non-destination nodes $z_i \in Z$ by formulating the following optimization problem:

$$\min_{\bar{w}} F(\bar{w}) = \|\bar{w}\|^2 + \lambda \sum_{d \in D, z \in Z} loss(p_z - p_d), \tag{4.8}$$

where λ is the regularization parameter, and $loss$ is a predefined loss function for penalizing the cases in which the stationary probability of a non-destination node p_z is higher than the stationary probability of a destination node p_d. In our experiments, we use the Wilcoxon–Mann–Whitney loss function [54].

4.5 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 the 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:

$$\theta(u_i, l_j, k) = Sim(i, j) \times \frac{1}{1 + e^{\frac{-k}{2}}}. \tag{4.9}$$

The stopping probability is inversely proportional to the step number k. $Sim(i, j)$ denotes the cosine similarity of crime trend of u_i and the road attractiveness of the road l_j. Cosine similarity defines the similarity between two vectors as the angle between these vectors:

$$Sim(i, j) = \frac{\bar{x}_i.\bar{a}_j}{|\bar{x}_i||\bar{a}_j|}. \tag{4.10}$$

5 Experimental evaluation

In this section, we present our experimental design, the comparison partners, and the results.

5.1 Experimental design

For each offender, we order his crime events chronologically based on their time. Then, we split these events into a training set and a test set. The first 80% of the crimes are used for training the model which predicts the offender activity space. The remaining 20% of crimes are used for testing the model. We consider only offenders with at least two different crimes 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 4.4.2 is not included in the evaluation to prevent biasing CrimeTracer.

After learning the offender activity space in the training phase, the trained model is applied in the test phase to predict future crime locations. To do so, the top-N roads with the highest probability are suggested as the most probable places for an offender to commit future crimes.

As discussed above, the focus of this work is modeling offenders' spatial behaviour in the coldspots. Thus, in our experiments, we exclude the top 100 roads with the highest crime numbers, the hotspots. The number of crimes in these hotspot roads is 100 to 1,100 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 prediction, we measure the number of crimes committed by an offender in his testing dataset among the top-N predicted locations. If a crime location in an offender's test set is also among the top-N predicted locations, that crime location is considered to be correctly *predicted*. Three accuracy measures, *precision*, *recall* and *utility*, are used as evaluation metrics:

- Recall computes the ratio of the number of correctly predicted crime locations (true positives) to the number of crime locations of the offender in the test set (true positives + false negatives).

- Precision computes the ratio of the number of correctly predicted crime locations (true positives) to the number of all predictions N (true positives + false positives).
- 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.

5.2 Comparison partners

In this section, we introduce different versions of CrimeTracer and the comparison partners methods used in our performance evaluation.

For evaluating the CrimeTracer performance, we test the two different movement directionality approaches and the following types of locations included in the activity space of offenders. For every offender locations are categorized into three groups: (a) *Known locations* that includes home and crime locations of the offender. (b) *Derived locations* which are locations shared with co-offenders and intermediate anchor locations. These locations are derived from observed information in the crime dataset. (c) *Unknown locations* that includes any location which is not a known or derived location.

For a deeper understanding of CrimeTracer performance and the role of each of above-mentioned location types, we consider three approaches: (1) In the first approach (denoted by U), we include only unknown locations in the activity space of an offender and consequently in the crime location prediction. (2) In the second approach (denoted by D), we include only unknown and derived locations in the activity space of an offender. (3) In the last approach (denoted by A), all locations are considered.

Two different movement directionality methods are introduced in Section 4.4: hotspot influence (denoted by H) and learning road feature weights (denoted by F). For each of these CrimeTracer versions, we consider the three above-mentioned evaluation approaches. For instance, CrimeTracer-HU denotes CrimeTracer using the hotspot influence method (H) for movement directionality that includes only unknown locations (U) in the predicted locations.

As discussed in Section 2, there is no related work that solves the problem of personalized crime location prediction. However, we use the following methods which are equivalent to state-of-the-art methods for location recommendation [51]:

Random Walk. This is the standard random walk with restart method (RWR) [49].

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 the offender's anchor locations. Here, distance denotes the length of the shortest path between two roads on the road network.

Offender-based CF. The intuition behind the offender-based CF approach is that offenders who had similar behaviour in the past will have similar behaviour in the future.

Let $b_{ij} = 1$ if $l_j \in \mathscr{L}_i$, and $b_{ij} = 0$ if $l_j \notin \mathscr{L}_i$. Now, $F(i, j)$ is the probability of a crime committed in road l_j by u_i:

$$F(i,j) = \frac{\sum_{u_k \in V \wedge k \neq i} Sim(i,k).b_{k,j}}{\sum_{u_k \in V \wedge k \neq i} Sim(i,k)}, \tag{5.1}$$

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

$$Sim(i,k) = \frac{\sum_{l_j \in L} b_{i,j}.b_{k,j}}{\sqrt{\sum_{l_j \in L} b_{i,j}^2}\sqrt{\sum_{l_j \in L} b_{k,j}^2}}. \tag{5.2}$$

Location-based CF. In location-based CF, we consider the similarity of locations instead of the similarity of offenders:

$$F(i,j) = \frac{\sum_{l_k \in L \wedge k \neq j} Sim(j,k).b_{i,k}}{\sum_{l_k \in L \wedge k \neq j} Sim(j,k)}, \tag{5.3}$$

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

$$Sim(j,k) = \frac{\sum_{u_i \in V} b_{i,j}.b_{i,k}}{\sqrt{\sum_{u_i \in V} b_{i,j}^2}\sqrt{\sum_{u_i \in V} b_{i,k}^2}}. \tag{5.4}$$

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 a crime being committed in road l_j by u_i as follows:

$$F(i,j) = \frac{\sum_{u_k \in \Gamma_i} Sim(i,k).b_{k,j}}{\sum_{u_k \in \Gamma_i} Sim(i,k)}. \tag{5.5}$$

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

$$Sim(i,k) = \frac{|\Gamma_i \cap \Gamma_k|}{|\Gamma_i \cup \Gamma_k|} + \frac{|\mathscr{L}_i \cap \mathscr{L}_k|}{|\mathscr{L}_i \cup \mathscr{L}_k|}. \tag{5.6}$$

5.3 Results

5.3.1 CrimeTracer *scenarios*

Figures 3–5 show performance of six different versions of CrimeTracer including CrimeTracer-HU CrimeTracer-HD, CrimeTracer-HA, CrimeTracer-FU, CrimeTracer-FD and CrimeTracer-FA in terms of recall, precision and utility measures.

With regard to the type of locations included in the prediction process, as expected

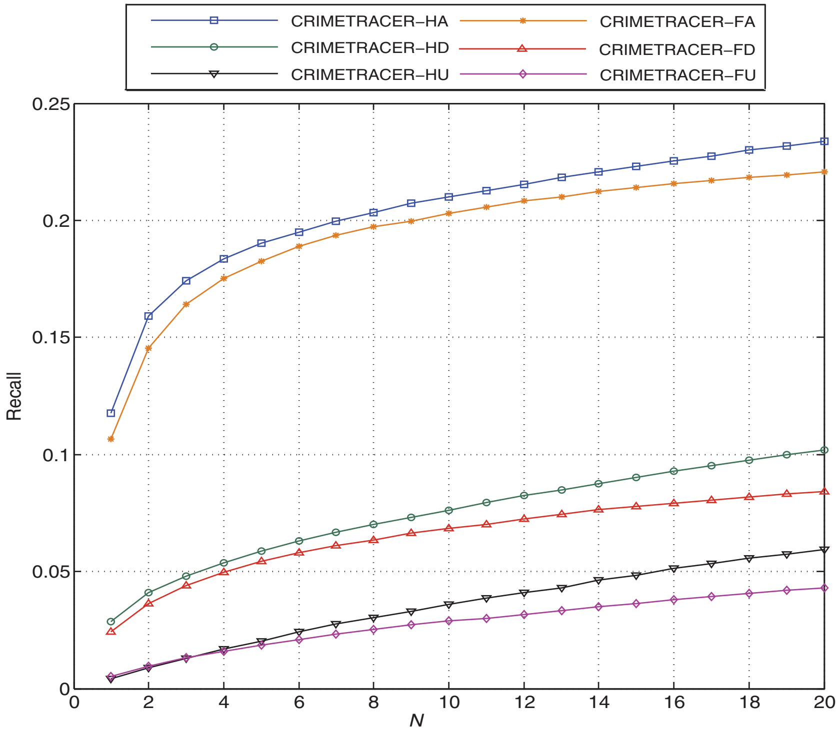


FIGURE 3. Recall of different versions of CRIMETRACER for different values of N.

CRIMETRACER-HA and CRIMETRACER-FA have the best performance, and CRIMETRACER-HU and CRIMETRACER-FU have the worst performance. The recall of CRIMETRACER-HA, CRIMETRACER-HD and CRIMETRACER-HU for $N = 20$ is 23.4%, 10.2% and 5.9%, respectively. Derived and known locations increase the recall by 4.3% and 13.1%, respectively. We observe a similar result when comparing the performance of CRIMETRACER-FA, CRIMETRACER-FD and CRIMETRACER-FU.

An important question is which of these scenarios should be used in a real-world application of CRIMETRACER. According to criminological theories such as exact-repeat/near repeat event [28] and broken window theory [53], known locations of offenders are always likely places to commit a new crime. The results presented in this section also support this idea. In a real-world application, known locations may be included in the predicted locations automatically. One may conclude that CRIMETRACER-HD and CRIMETRACER-FD are more appropriate versions of CRIMETRACER for a real-world application.

Considering the two movement directionality approaches, both versions of CRIMETRACER achieve higher performance compared to the standard random walk approach. CRIMETRACER-HD compared to CRIMETRACER-FD and CRIMETRACER-HU compared to CRIMETRACER-FU have higher recall, precision and utility. CRIMETRACER-HA compared to CRIMETRACER-FA has higher recall and utility for all values of N, but their precision values are almost identical for $N \geqslant 6$. We conclude that the hotspot influence approach outperforms the other method, showing the great impact of crime attractors and generators in committing a new crime by an offender.

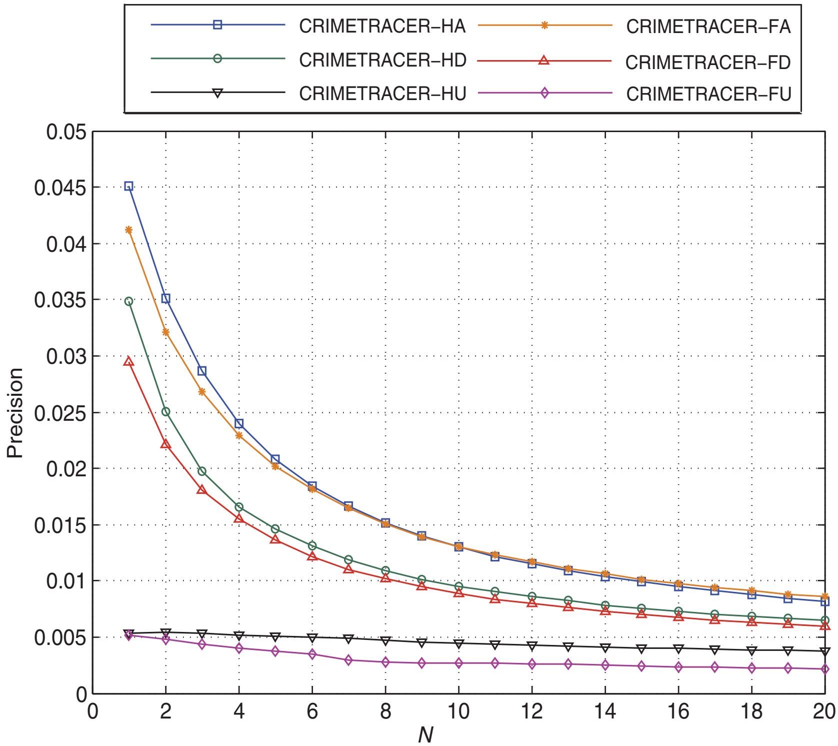


FIGURE 4. Precision of different versions of CRIMETRACER for different values of N.

5.3.2 *Comparison partners*

Figures 6–8 show the overall performance of the different evaluated methods in terms of recall, precision and utility. To compare CRIMETRACER against the baseline methods, we use only the best performing versions CRIMETRACER-HD and CRIMETRACER-HU. Both of these methods consistently outperform all baseline methods for all values of N with regard to all evaluation metrics. The baseline methods use the same experimental design as CRIMETRACER-HD, but we also test CRIMETRACER-HU in the comparison to show that even in this case of a more restricted scenario, CRIMETRACER still outperforms the baseline methods.

DS (Proximity) obtains the lowest precision and recall values. Despite the well-studied theory of the relationship between crime commitment and distance from anchor locations, this result shows that this approach is not effective for personalized crime prediction. Among the CF-based approaches, Offender-based CF has the poorest performance. Location-based CF achieves better recall, but SCF (Co-offending-based CF) achieves higher precision. It is interesting to observe that location similarity contributes more to the accuracy of crime location prediction than offender similarity. One can conclude that SCF uses more reliable but limited information for predicting the offenders activity space. The recall of HS (Hotspots) improves with increasing N, but this method naturally is strong in predicting crimes in hotspots and not in coldspots.

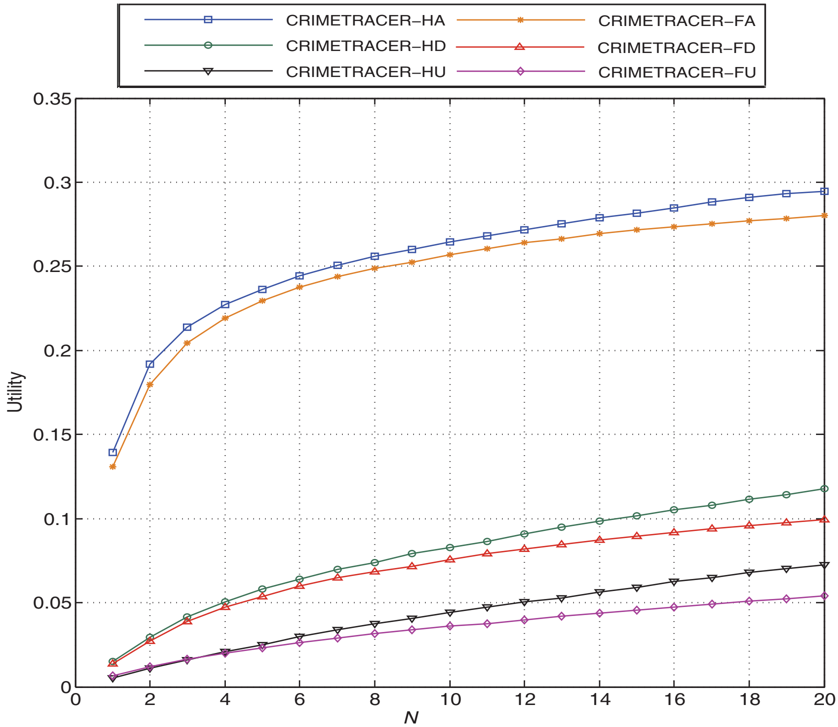


FIGURE 5. Utility of different versions of CRIMETRACER for different values of N.

Predicting even one crime location of each offender is very important for the critical task of crime prevention. As for the other two evaluation metrics, both versions of CRIMETRACER outperform the baseline methods in terms of utility. The utility of CRIMETRACER-HU and CRIMETRACER-HD is 1.3% and 1.5%, respectively larger than their recall ($N = 20$), making no significant difference. One reason for this effect is that half of the offenders committed only two crimes, and we can predict only one crime location for them, meaning that for these offenders the recall and utility values are the same.

There has long been interest in the behaviour of repeat offenders since controlling these groups of offenders can reduce the overall crime level significantly. Figures 9–11 depict the performance of the different methods for offenders with different numbers of crimes. We expect more successful activity space learning for offenders who have committed more crimes, and for whom we have more information. We observe such a trend for CRIMETRACER-FU, where the average recall for offenders who committed only two crimes is about 4% while this value increases by 3% for offenders who committed 10 or more crimes, as well as for RWR (Random walk) and SCF (Co-offending-based CF).

Interestingly, the hotspot influence approach causes a significant increase in recall of non-repeat offenders (the biggest group of offenders). Comparing CRIMETRACER-HU to CRIMETRACER-FU, the recall increases by 2% for this group of offenders, while the recall for repeat offenders is almost equal for these two methods. On the other hand, while

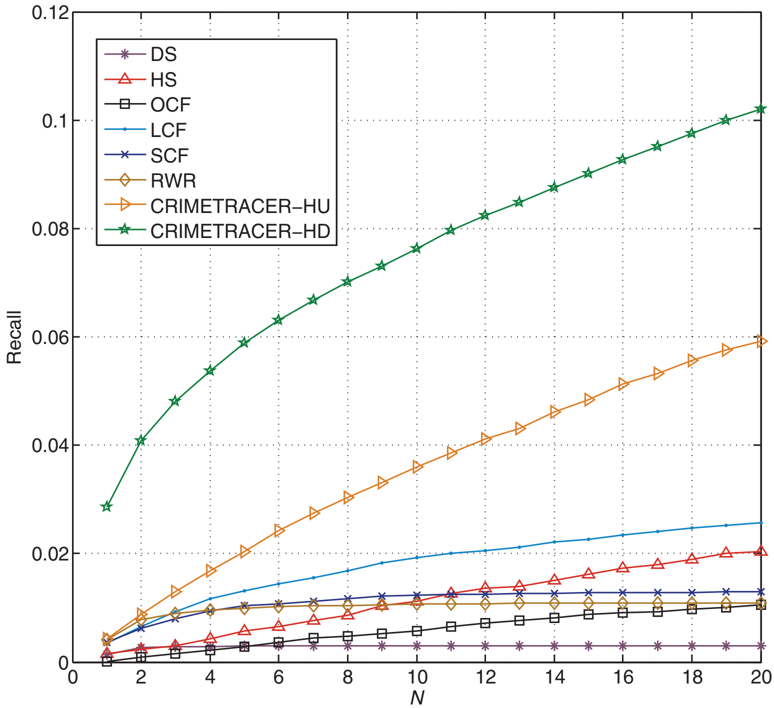


FIGURE 6. Recall for different values of N for the methods DS (Proximity), HS (Hotspots), OCF (Offender-based CF), LCF (Location-based CF), SCF (Co-offending-based CF), RWR (Random Walk), CrimeTracer-HU and CrimeTracer-HD.

for CrimeTracer-FU the recall of repeat offenders is 3% higher than the recall of non-repeat offenders, this difference is only 1% for CrimeTracer-HU. Thus, the directionality movement approach influenced by hotspot locations contributes more to the recall of non-repeat offenders than to the recall of repeat offenders.

While we do not observe a significant increase in recall of repeat offenders compared to non-repeat offenders for either of the CrimeTracer versions, we observe such a trend in the precision measure. Another interesting observation is that for SCF using co-offending information causes a significant performance gain for repeat offenders who have higher co-offending rates.

Non-repeat offenders are the majority of offenders, and in this study half of the offenders used for the evaluation committed only two crimes. As shown in Figures 9–11, for non-repeat offenders CrimeTracer-HU and CrimeTracer-HD outperform the baseline methods by large margins. We notice that location-based CF also works 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, while SCF is not accurate for beginners, 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.

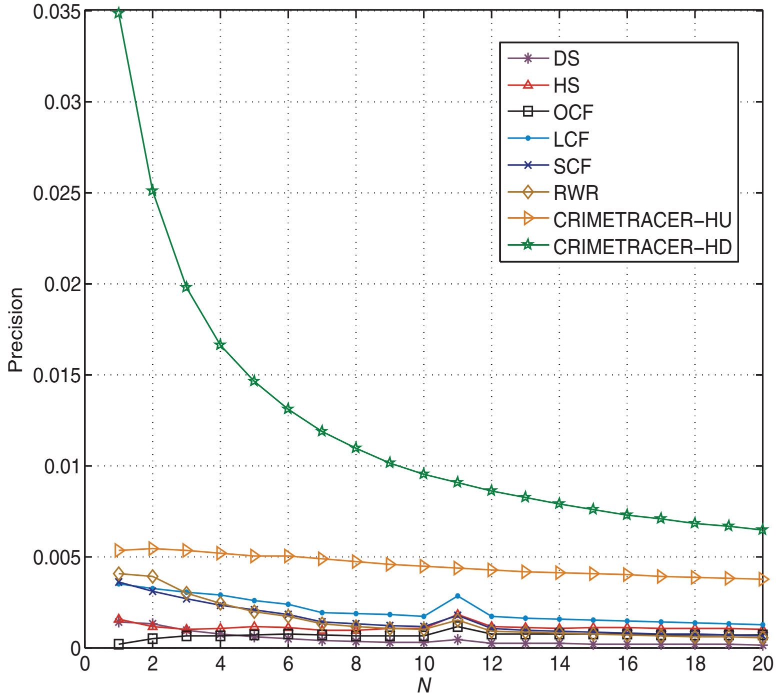


FIGURE 7. Precision for different values of N for the methods DS (Proximity), HS (Hotspots), OCF (Offender-based CF), LCF (Location-based CF), SCF (Co-offending-based CF), RWR (Random Walk), CrimeTracer-HU and CrimeTracer-HD.

5.3.3 CrimeTracer *elements*

We studied the contribution of different components of CrimeTracer to its performance. Compared to the standard RWR, CrimeTracer incorporates additional anchor locations (co-offending information and intermediate anchor locations), movement directionality and stopping criteria. We added these components separately to RWR to determine their individual contribution. Table 2 shows the results. The strongest component is the stopping criteria and the weakest is the learning of road feature weights. The main idea behind the stopping criteria is to stop the random walk of an offender in a road where the crime history is similar to the offender crime trend. However, combining all components in CrimeTracer-HD achieves the best result and improves the performance of RWR significantly in terms of all evaluation metrics. We include the performance of other versions of CrimeTracer in Table 2 to be able to compare the performance of different versions of CrimeTracer more exactly.

We note that the overall performance of CrimeTracer is comparable to the performance of state-of-the-art methods for location recommendation [30, 51], where the information about users' spatial patterns is much denser than 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 [15], roads are the natural domain for many policing activities, and a more realistic

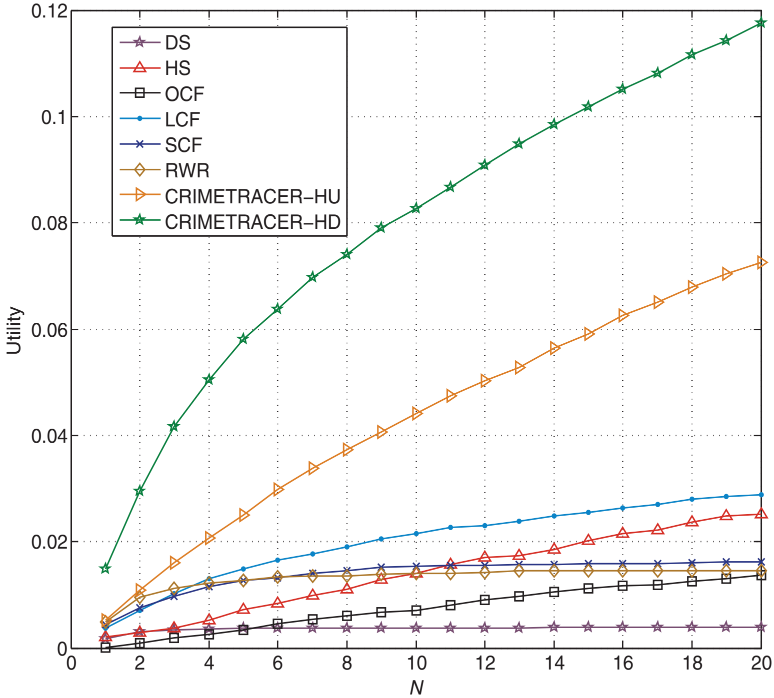


FIGURE 8. Utility for different values of N for the methods DS (Proximity), HS (Hotspots), OCF (Offender-based CF), LCF (Location-based CF), SCF (Co-offending-based CF), RWR (Random Walk), CRIMETRACER-HU and CRIMETRACER-HD.

Table 2. *Contribution of different elements of* CRIMETRACER *to its performance* ($N = 20$)

Method	Recall	Precision	Utility
RWR	0.011	0.004	0.014
RWR + Road features weight	0.013	0.003	0.017
RWR + Hotspot influence	0.015	0.003	0.016
RWR + Additional anchor locations	0.019	0.001	0.024
RWR + Stopping criteria	0.036	0.003	0.045
CRIMETRACER-HU	0.059	0.006	0.073
CRIMETRACER-FU	0.043	0.004	0.054
CRIMETRACER-HD	0.102	0.007	0.118
CRIMETRACER-FD	0.084	0.006	0.010
CRIMETRACER-HA	0.23	0.008	0.30
CRIMETRACER-FA	0.22	0.008	0.28

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.

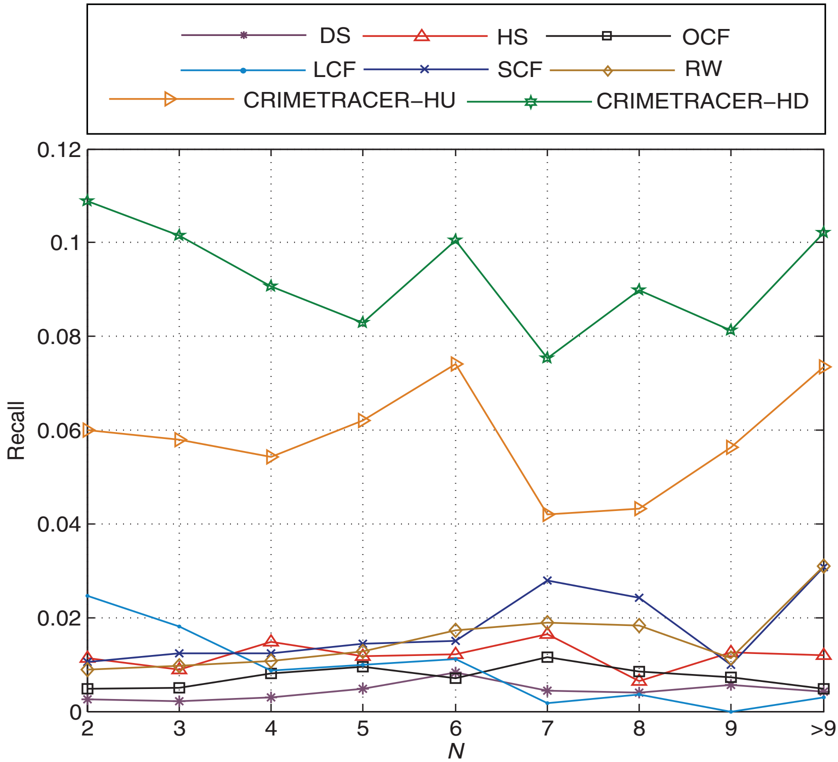


FIGURE 9. Recall for the offenders with different number of crimes ($N = 20$).

Random walk restart probabilities α and β are two important parameters that can influence the CRIMETRACER performance significantly. Small values of α let an offender move to farther locations to find criminal opportunities, while the higher values of α restricts the offenders movements around his anchor locations. On the other hand, higher values of β mean a higher chance that the offender starts his random walk from one of his intermediate anchor locations. We run a series of experiments to find the parameter values that maximize the CRIMETRACER performance. We experimentally vary values of α and β between zero and one to determine the optimal parameter values. The results presented in this paper are based on $\alpha = 0.3$ and $\beta = 0.3$, meaning that the offender will restart with probability of 0.3 the random walk and get back to one his anchor locations, and with probability of 0.7 he will move to the next road segment. In many applications of the random walk method, it is common to assign a higher value for the restart probability. However, here $\alpha = 0.3$ shows that restricting the offender to traverse only locations around his anchor locations does not result in the best CRIMETRACER performance. $\beta = 0.3$ means, that when the offender restarts the random walk, with probability of 0.3 he will return to one of his main anchor locations, and with probability of 0.7 he will return to one of his intermediate anchor locations. Although it would be great to learn the parameters that optimize the performance of the method, instead of providing them as input, this is not feasible because random walk methods are typically

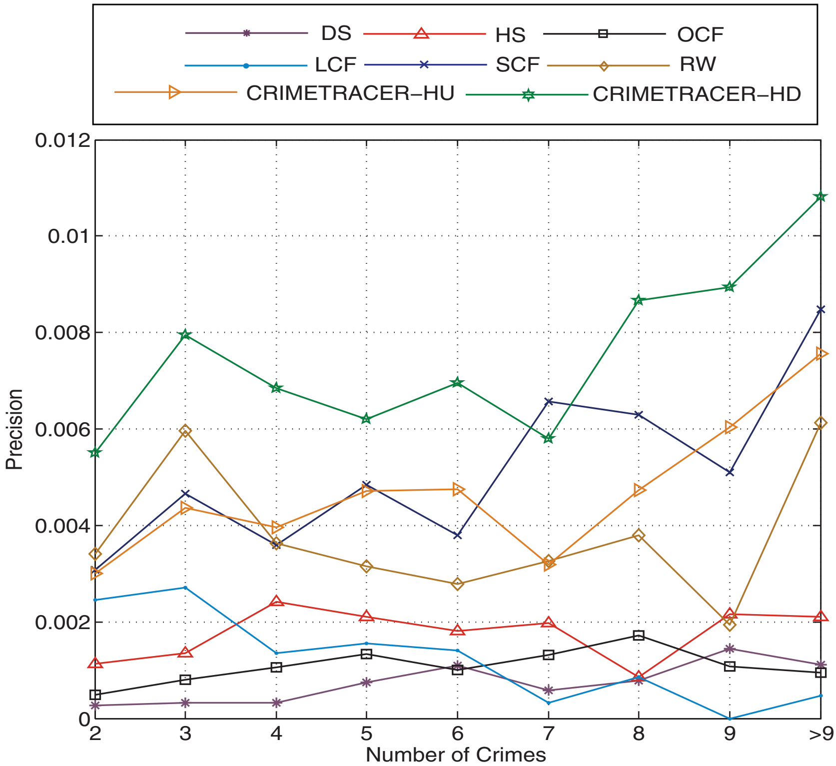


Figure 10. Precision for the offenders with different number of crimes ($N = 20$).

fully unsupervised and do not learn their parameters. However, parameter learning for restart probablity is beyond the scope of the paper, but we consider it as a future work.

Another important parameter on where an offender commits crime in the future is the temporal distribution of crimes he committed in the past. Intuitively, the time of crimes which have been committed more recently impact the prediction of future crime locations more strongly. However, we are not aware of any criminological theories which establish this concept in a tangible way. Ultimately, one would have to consider both spatial and temporal distributions of these crimes to find the right balance in determining the importance of these offences for predicting future crime locations. What's more, an exact model to solve this problem in a real-world setting should consider these aspects not only at the level of individual offences but also at the aggregate level. This is a complex question which to answer needs deeper research that is beyond the scope of this paper. In this paper, we refer to this control parameter in Equation (4.4) as ρ. In our experiments, we split the time into one-month intervals. Then, we use the month in which a crime incident occurred as input. In our experimental setting, we use $\rho = 25$ which regulates the time effect in Equation (4.4) by assigning values between 0.1 and 1 to this factor.

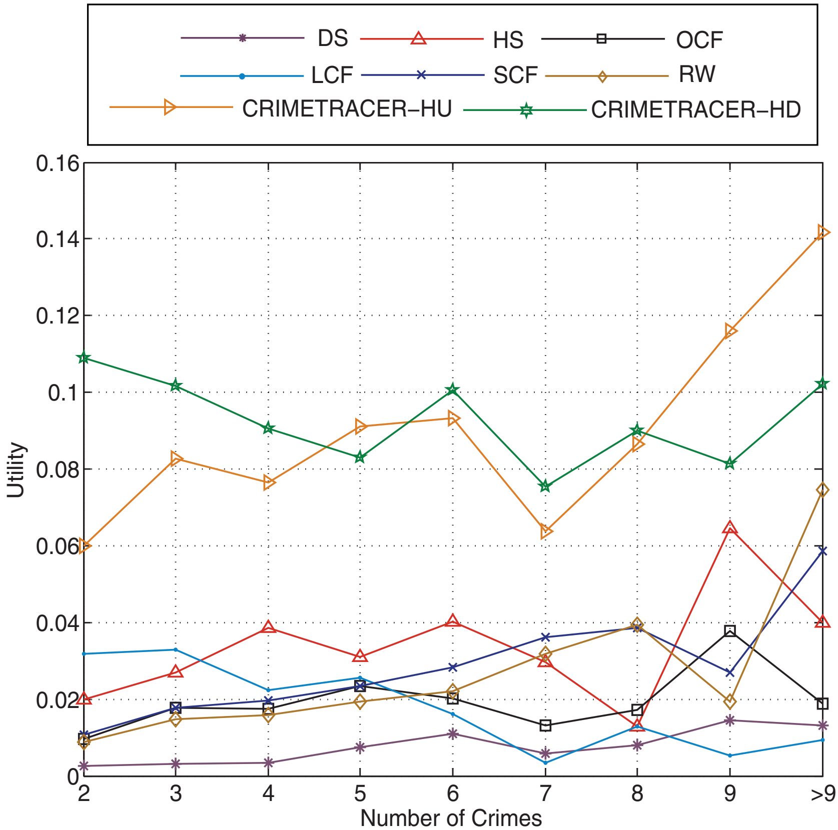


FIGURE 11. Utility for the offenders with different number of crimes ($N = 20$).

6 Conclusions

Through linking data mining algorithms with social network analysis, advanced crime data analysis methods can provide a scientific foundation for analysing spatsial decision-making of criminal offenders and their social standing as is necessary for developing effective crime reduction and prevention strategies.

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 behavioural patterns. CRIMETRACER uses a personalized random walk to derive a *probabilistic activity space* model for known offenders based on facts from their criminal history as documented in an offender profile. 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. As expected, the chance of having co-offending links is higher for 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 of visiting and time spent at these locations. Exploring this aspect in more depth remains future work.

We believe that the ideas presented here can inspire new research trends in social network analysis and data mining with useful applications for 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 and the Natural Sciences and Engineering Research Council of Canada for their continued support of this research.

References

[1] Backstrom, L. & Leskovec, J. (2011) Supervised random walks: Predicting and recommending links in social networks. In: *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM'11)*, Hong Kong, China, pp. 635–644.

[2] Bernasco, W. (2006) Co-offending and the choice of target areas in burglary. *J. Investigative Psychol. Offender Profiling* **3**(3), 139–155.

[3] Bernasco, W. (2008) Them again? Same-offender involvement in repeat and near repeat burglaries. *Eur. J. Criminology* **5**(4), 411–431.

[4] Bernasco, W. (2010) Modeling micro-level crime location choice: Application of the discrete choice framework to crime at places. *J. Quant. Criminology* **26**(1), 113–138.

[5] Bernasco, W. & Block, R. (2009) Where offenders choose to attack: A discrete choice model of robberies in chicago. *Criminology* **47**(1), 93–130.

[6] Bernasco, W. & Nieuwbeerta, P. (2005) How do residential burglars select target areas? A new approach to the analysis of criminal location choice. *Br. J. Criminology* **45**(3), 296–315.

[7] Braga, A. A. (2001) The effects of hot spots policing on crime. *Ann. Am. Acad. Political Soc. Sci.* **578**(1), 104–125.

[8] Brantingham, P. J. & Brantingham, P. L. (1981) *Environmental Criminology*, Sage Publications, Beverly Hills, CA, USA.

[9] Brantingham, P. L. & Brantingham, P. J. (1993) Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *J. Environ. Psychol.* **13**(1), 3–28.

[10] Brantingham, P. L. & Brantingham, P. J. (1995) Criminality of place. *Eur. J. Criminal Policy Res.* **3**(3), 5–26.

[11] Brantingham, P. L., Ester, M., Frank, R., Glässer, U. & Tayebi, M. A. (2011) Co-offending network mining. In: U. K. Wiil (editor), *Counterterrorism and Open Source Intelligence*, Springer, Vienna, pp. 73–102.

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

[13] Canter, D. V. & Gregory, A. (1994) Identifying the residential location of rapists. *J. Forensic Sci. Soc.* **34**(3), 169–175.

[14] Chrastil, E. R. (2013) Neural evidence supports a novel framework for spatial navigation. *Psychonomic Bull. Rev.* **20**(2), 208–227.

[15] Davies, T. P. & Bishop, S. R. (2013) Modelling patterns of burglary on street networks. *Crime Sci.* **2**(1), Article 10.

[16] Frank, R., Andresen, M. A. & Brantingham, P. L. (2012) Criminal directionality and the structure of urban form. *J. Environ. Psychol.* **32**(1), 37–42.

[17] Frank, R., Andresen, M. A. & Brantingham, P. L. (2013) Visualizing the directional bias in property crime incidents for five canadian municipalities. *Can. Geogr./Le Géographe Can.* **57**(1), 31–42.

[18] Frank, R., Andresen, M. A., Cheng, C. & Brantingham, P. L. (2011) Finding criminal attractors based on offenders' directionality of crimes. In: *Proceedings of the 2011 European Intelligence and Security Informatics Conference (EISIC'11)*, Athens, Greece, pp. 86–93.

[19] Frank, R. & Kinney, B. (2012) How many ways do offenders travel – evaluating the activity paths of offenders. In: *Proceedings of the 2012 European Intelligence and Security Informatics Conference (EISIC'12)*, Odense, Denmark, pp. 99–106.

[20] Glässer, U., Taybei, M. A., Brantingham, P. L. & Brantingham, P. J. (2012) Estimating possible criminal organizations from co-offending data. Public Safety Canada, Ottawa, Canada, http://publications.gc.ca/collections/collection_2012/sp-ps/PS14-7-2012-eng.pdf.

[21] Golledge, R. G. (1981) Misconceptions, misinterpretations, and misrepresentations of behavioral approaches in human geography. *Environ. Plan. A* **13**(11), 1325–44.

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

[23] Gorr, W. & Harries, R. (2003) Introduction to crime forecasting. *Int. J. Forecast.* **19**(4), 551–555.

[24] Grimmett, G. & Stirzaker, D. (1989) *Probability and Random Processes*, Oxford University Press, Oxford, UK.

[25] Harries, K. (1999) Mapping crime principle and practice. *U.S. Department of Justice, Office of Justice Programs*, National Institute of Justice, Washington, DC, USA, https://www.ncjrs.gov/pdffiles1/nij/178919.pdf.

[26] Johnson, S. D. (2008) Repeat burglary victimisation: A tale of two theories. *J. Exp. Criminology* **4**(3), 215–240.

[27] Johnson, S. D. (2010) A brief history of the analysis of crime concentration. *Eur. J. Appl. Math.* **21**(4–5), 349–370.

[28] Johnson, S. D., Bernasco, W., Bowers, K. J., Elffers, H., Ratcliffe, J., Rengert, G. & Townsley, M. (2007) Space–time patterns of risk: A cross national assessment of residential burglary victimization. *J. Quant. Criminology* **23**(3), 201–219.

[29] Kolokolnikov, T., Ward, M. & Wei, J. (2014) The stability of steady-state hot-spot patterns for a reaction-diffusion model of urban crime. *Discrete Continuous Dyn. Syst.* **19**(5), 1373–1410.

[30] Liu, B., Fu, Y., Yao, Z. & Xiong, H. (2013) Learning geographical preferences for point-of-interest recommendation. In: *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, Illinois, USA, pp. 1043–1051.

[31] Liu, H. & Brown, D. E. (2003) Criminal incident prediction using a point-pattern-based density model. *Int. J. Forecast.* **19**(4), 603–622.

[32] Miller, B. N., Konstan, J. A. & Riedl, J. (2004) Pocketlens: Toward a personal recommender system. *ACM Trans. Inform. Syst. (TOIS)* **22**(3), 437–476.

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

[34] Reiss Jr., A. J. (1988) Co-offending and criminal careers. *Crime Justice* **10**, 117–170.

[35] Rodriguez, N. & Bertozzi, A. (2010) Local existence and uniqueness of solutions to a pde model for criminal behavior. *Math. Models Methods Appl. Sci.* **20**(supp01), 1425–1457.

[36] ROSSMO, D. K. (2000) *Geographic Profiling*, CRC Press.

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

[38] SHERMAN, L. W., GARTIN, P. R. & BUERGER, M. E. (1989) Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* **27**(1), 27–56.

[39] SHORT, M. B., BERTOZZI, A. L. &BRANTINGHAM, P. J. (2010) Nonlinear patterns in urban crime: Hotspots, bifurcations, and suppression. *SIAM J. Appl. Dyn. Syst.* **9**(2), 462–483.

[40] SHORT, M. B., BRANTINGHAM, P. J., BERTOZZI, A. L. & TITA, G. E. (2010) Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proc. Nat. Acad. Sci.* **107**(9), 3961–3965.

[41] SHORT, M. B., D'ORSOGNA, M. R., PASOUR, V. B., TITA, G. E., BRANTINGHAM, P. J., BERTOZZI, A. L. & CHAYES, L. B. (2008) A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.* **18**(supp01), 1249–1267.

[42] SONG, J., SPICER, V., BRANTINGHAM, P. L. & FRANK, R. (2013) Crime ridges: Exploring the relationship between crime attractors and offender movement. In: *Proceedings of the 2013 European Intelligence And Security Informatics Conference (EISIC'13)*, IEEE, Uppsala, Sweden, pp. 75–82.

[43] TAYEBI, M. A., BAKKER, L., GLÄSSER, U. & DABBAGHIAN, V. (2011) Locating central actors in co-offending networks. In: *Proceedings of the 2011 International Conference on Advances in Social Networks Analysis and Mining (ASONAM'11)*, Kaohsiung, Taiwan, pp. 171–179.

[44] TAYEBI, M. A., ESTER, M., GLÄSSER, U. & BRANTINGHAM, P. L. (2014) CRIMETRACER: Activity space based crime location prediction. In: *Proceedings of the 2014 International Conference on Advances in Social Networks Analysis and Mining (ASONAM'14)*, Beijing, China, pp. 472–480.

[45] TAYEBI, M. A., ESTER, M., GLÄSSER, U. & BRANTINGHAM, P. L. (2014) Spatially embedded co-offence prediction using supervised learning. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'14)*, New York, New York, USA, pp. 1789–1798.

[46] TAYEBI, M. A., FRANK, R. & GLÄSSER, U. (2012) Understanding the link between social and spatial distance in the crime world. In: *Proceedings of the 20nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS'12)*, Redondo Beach, California, USA, pp. 550–553.

[47] TAYEBI, M. A., JAMALI, M., ESTER, M., GLÄSSER, U. & FRANK, R. (2011) CRIMEWALKER: a recommendation model for suspect investigation. In: *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys'11)*, Chicago, Illinois, USA, pp. 173–180.

[48] TAYEBI, M. A. & GLÄSSER, U. (2012) Investigating organized crime groups: A social network analysis perspective. In: *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM'12)*, Istanbul, Turkey, pp. 565–572.

[49] TONG, H., FALOUTSOS, C. & PAN, J. (2006) Fast random walk with restart and its applications. In *Proceedings of the 6th International Conference on Data Mining (ICDM'06)*, Hong Kong, China, pp. 613–622.

[50] TOWNSLEY, M. & SIDEBOTTOM, A. (2010) All offenders are equal but some are more equal than others: Variation in journeys to crime between offenders. *Criminology* **48**(3), 897–917.

[51] WANG, H., TERROVITIS, M. & MAMOULIS, N. (2013) Location recommendation in location-based social networks using user check-in data. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS'13)*, San Francisco, California, USA, pp. 374–383.

[52] WEISBURD, D. L., GROFF, E. R. & YANG, S. (2012) *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*, Oxford University Press.

[53] WILSON, J. Q. & KELLING, G. L. (1982) Broken windows and police and neighborhood safety. *Atlantic Monthly* **24**(9), 29–38.

[54] Yan, L., Dodier, R. H., Mozer, M. & Wolniewicz, R. H. (2003) Optimizing classifier performance via an approximation to the wilcoxon-mann-whitney statistic. In: *Proceedings of the 20th International Conference on Machine Learning (ICML'03)*, Washington, DC, USA, pp. 848–855.

[55] Ye, M., Yin, P. & Lee, W. (2010) Location recommendation for location-based social networks. In: *Proceedings of the 18th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS'10)*, San Jose, California, USA, pp. 458–461.