

Urban City: Crime Attractors and Offender Mobility

Vincent T. Nguyen and Sohrab Ganjian [†]

Project advisors: Razvan Fetecau [‡] and Patricia L. Brantingham [§]

Abstract. This article develops a generalizable agent-based simulation model of crime. This model is a direct application of a diffusion model of residential burglary and offender mobility by Short et al. [31] to real world crime. Offender mobility is modeled through biased random movement and crime attractiveness of a location is derived from past crimes, diffusion of information and memory decay of past events. Simulations use realistic spatial and crime data for Vancouver, Canada with an error analysis performed at the macro, meso, and micro level. The experimental results are convincing, achieving Pearson Correlation values of 0.988, 0.909, and 0.732 for each level respectively using a Truncated Levy Flight mobility framework. Without a mobility framework, we were able to achieve 0.997, 0.954, and 0.645 for these same levels.

1. Introduction. Research in patterns in crime has been part of social science research and law for over 200 years [19, 26, 29]. Within the last fifty years this has become more of an interdisciplinary area of interest. Sociologists, psychologists, criminologists, lawyers, geographers, engineers, mathematicians and computing scientists currently engage in individual and jointly guided research [8, 18, 30, 31, 32, 34, 35, 40, 45]. A focus of the research is increasingly on the distribution of criminal activity in urban space. Crime is not uniformly distributed across an urban area, as it tends to cluster strongly in some places, to disperse to lower density in other places and be close to non-existent in many areas [10, 17, 41].

Crime Pattern Theory [7, 11] is the basis of research that finds that crime clusters at or near high activity nodes and the key travel or transit pathways between these nodes. These nodes and pathways help shape the activity spaces and awareness spaces of the urban space within which offenders select their targets. Stated another way, there are special nodes within cities that pull people towards them. As well, people from the same general area of a city are pulled along the same arterial routes toward these major activity nodes. Crime sites are selected at or near the high activity nodes or along the major arteries where there are suitable targets [8, 18, 32, 33].

Crime attractors and crime generators are activity nodes that attract many people. City centres with large shopping areas or large regional shopping malls with many shops, restaurants, bars, clustering of cars and major bus or metro stops are also attractive for shoplifting, theft from and of motor vehicles or bicycles or street robbery of persons waiting for public transit. Drinking establishments frequently are the location of assaults. Sometimes shopping areas also become areas to sell drugs, steal credit card information or areas where peripheral activity like noisy interactions, uttering threats, conflicting users of the space generate additional crimes and nuisance type behaviors. Schools and the shopping areas near schools, parts of parks, sports centres and general entertainment areas also generate crime.

[†]Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada (vtn5@sfu.ca, sganjian@sfu.ca).

[‡]Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada

[§]School of Criminology, Simon Fraser University, Burnaby, BC, Canada

Add in easy accessibility by motor vehicle, by foot traffic and by public transit to the areas described above and the likelihood of crime goes up [11, 22, 23, 33]. Research shows a pattern of crime occurring along high use pathways where people can stop, particularly at intersections of busy pathways and where shopping and drinking areas cluster along these pathways.

The increasing multidisciplinary work is leading to more simulation and mathematical research in the area of crime attractors and offender mobility on street networks. Along these networks, offenders spend much of their daily lives pursuing activities common to the general populace [1]. They commute to school or work, shop, attend to routine leisure activities, and engage in unique or annual activities like summer vacation. In addition to maintaining their normal routines, we find that offenders find opportunities to commit crimes along the pathways they used during their commutes [9]. Furthermore, we find areas along the street network where crime occurs at high levels disproportionate to the rest of the network, such as at crime attractors and generators [9, 11, 12]. These spatial-temporal regions of elevated crime are often referred to as crime concentrations or hot spots [1, 5, 37, 39].

Our interest in this article is to study the dynamics of these crime concentrations within an urban city. Notably, we study their evolution with respect to a city's road network. The technique by which we study the dynamics of these crime concentrations is by utilizing a mathematical model we propose in this paper. This model is developed as an advancement to a residential burglary hotspot model proposed initially by Short and colleagues [31] and several others based on it [15, 25]. These models explore crime concentrations through the perspective of mathematics and diffusion [15, 25, 30, 31]. Hence, we will refer to this series of papers as the *diffusion papers*. Diffusion in this sense is the phenomenon of spreading substances throughout a medium. In this research, concentrations of crime serve as our primary substance of interest.

Although our model is similar to that of the diffusion papers, it is inherently far more discrete mathematically. That is, our model is entirely discrete and does not employ the use of calculus. Specifically, we explore spatial-temporal criminology under the framework of network theory. The discreteness of our model is beneficial as it allows us to utilize empirical data readily available in many urban settings. As such, we are able to explore our model on a real-world street network. In this article, we present our findings using the city of Vancouver, BC, Canada. Crime records from the city are used to initialize our model on its street network. However, the techniques we explore can be generalized to any other city and street network. (see [46] for modelling on a metro rail system.)

2. Attractiveness. Crime attraction to activity nodes or major pathways between nodes can also be considered as relative attractiveness to offenders laid over the backcloth or the urban surround. Some parts of a city will have high attractiveness; other parts will not. The attractiveness can also be viewed at different levels of resolution or granularity. This level of resolution is relative and dependent on the overall scale being used, but within this relative scale it is usually divided in macro, meso and micro. In research in crime attractors, the macro level usually means a complete city or urban area and is not used since the focus is finding out the reasons for dense clusters of crime in some parts of a city and low or no crime in other parts. Meso analysis is often used to look at the difference in crime concentrations in different and distinct parts of a city like a neighborhood or a defined shopping area. Micro

analysis is tied to finding out which locations within a higher crime area have little or no crime and which have more crime [22] It is common for a street segment on a street to have many shops, but only certain types of shops experience burglary or shoplifting and others do not. For example food stores experience substantial shoplifting and flower shops do not. Bars experience assaults and threatening activity while children's clothing stores do not. In no street segment is crime attractiveness a constant value or even represent a smooth flow towards higher attractiveness or lower attractiveness. In all depends on the land use or more appropriately the use of land.

2.1. Diffusion Model of Attractiveness. Attractiveness in the diffusion papers is a quantifiable value assigned to each potential target in a given study space, while the attractiveness field is the collection of all attractiveness values over the span of the targets in the space. Specifically, let s be a given site for crime to occur at and let the quantity $A_s(t)$ be the attractiveness at this site at a given time t . The attractiveness is assumed to consist of two components:

$$(2.1) \quad A_s(t) \equiv A_s^0 + B_s(t),$$

where A_s^0 represents a static, but possibly spatially varying, component and $B_s(t)$ is the dynamic component associated with repeat and near-repeat victimization. The attractiveness field A is then, the collection of all $A_s(t)$ for all sites in a given space. This space is defined as a two-dimensional lattice in the paper by Short et al. [31] and as a one-dimension lattice in the papers by Chaturapruek et al. [15] and Pan et al. [25]. Our space of interest is an urban city and in our model, this will be a street network.

The dynamic component in the diffusion papers is based on an area's increase in vulnerability due to recent or past victimization. This phenomenon is often referred to as Repeat Near-Repeat Victimization. To replicate it, the papers make several assumptions about attractiveness: i) Once a crime has been committed at a crime site, that site is more vulnerable to another event and therefore more attractive to future crimes. ii) Attractiveness diffuses to nearby sites due to the spread of common knowledge. iii) Vulnerability decays over time due to increased protection and simply due to the fading memory of criminals as they move on to other targets.

In the diffusion papers, a value θ is given to quantify the instantaneous increase in attractiveness. Next, the awareness of this site's attractiveness spreads towards neighbouring sites to represent the effects of publicity, communications between offender groups, and general common knowledge. The spreading is the characteristic effect of diffusion, giving the models their namesake in this work. How much diffusion occurs is determined by a tightness factor $\eta \in [0, 1]$ where values close to 0 restrict diffusion and ones close to 1 allow for full diffusion. Finally, any lingering attractiveness from historical events decays as a result of memory loss in the offenders awareness of their activity spaces. The decay rate is given by the constant ω .

The dynamic component $B_s(t)$ is influenced by the feedback mechanisms based on events which occur at the site and its neighbours within a time interval of length δt . These feedback mechanisms include the parameters θ , ω , and η defined above. The dynamic attractiveness

component is updated after a time interval of length δt through the equation:

$$(2.2) \quad B_s(t + \delta t) = \left[B_s(t) + \frac{\eta \ell^2}{z} \Delta B_s(t) \right] (1 - \omega \delta t) + \theta E_s(t),$$

where z is the number of sites s' which neighbour s , ℓ is the common distance between points, $E_s(t)$ is the number of crimes which occurred at site s at time t , and Δ is the discrete spatial Laplacian operator such that

$$(2.3) \quad \Delta B_s(t) = \left(\sum_{s' \sim s} B_{s'}(t) - z B_s(t) \right) / \ell^2.$$

2.2. Diffusion Model on a Road Network. Remarkably, adapting Equation (2.2) is a straightforward process if we model our road network as a graph. In the mathematical field of graph theory, a graph is a collection of nodes, which are often used to represent objects, and edges which are the connections between nodes. A common example of a graph is a social network of friends where individuals are represented by nodes and edges between nodes symbolizes that the two individuals are friends.

To represent our road network as a graph, we let sites on the network be represented by nodes and the streets between sites as edges. Here, sites are the road network's intersections and also points in our data used to interpolate a curved street. We will denote a given site on the network by s to stay consistent with the notation in the diffusion papers.

In the setting of graph theory, s is simply an identifier and does not need to be a coordinate, despite representing spatial coordinates in our model. Therefore, the attractiveness field can be represented by the vector

$$(2.4) \quad \vec{A}(t) = [A_1(t), \dots, A_m(t)]^T,$$

where m is the number of sites in our network. Similarly, the dynamic attractiveness field $\vec{B}(t)$, that consists of all $B_s(t)$, can be represented by a vector

$$(2.5) \quad \vec{B}(t) = [B_1(t), \dots, B_m(t)]^T.$$

Therefore, we have $\vec{A}(t) = \vec{A}^0 + \vec{B}(t)$ where A^0 is the static attractiveness field over the whole network. In this research, we are primarily interested in the dynamic component and so we set $\vec{A}^0 = \vec{0}$. In the same fashion, we also define the collection of all crimes at time t , $\vec{E}(t)$, as a vector consisting of all $E_s(t)$:

$$(2.6) \quad \vec{E}(t) = [E_1(t), \dots, E_m(t)]^T.$$

One concept graph theorists use to describe the connections between nodes is the *adjacency matrix*, which we denote as \mathcal{A} . Given a graph G with n nodes, we number each node uniquely with the numbers $1, \dots, n$. The adjacency matrix \mathcal{A} is the $n \times n$ matrix with components:

$$(2.7) \quad \mathcal{A}_{ij} = \begin{cases} 1 & \text{if sites } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

Note that \mathcal{A} is a symmetric matrix, that is, $\mathcal{A}_{ij} = \mathcal{A}_{ji}$ for all i, j .

A more general concept than the adjacency matrix is the weighted adjacency matrix \mathcal{W} , where connectivity is defined by the *cost* it takes to travel between two connected nodes. These costs replace the 1's in the matrix and weights can be thought of as distances, prices, or lengths of time. We will utilize a distance weighted adjacency matrix in our model. Another important matrix we define for a graph G is its *degree matrix*, denoted by \mathcal{D} . This is an $n \times n$ diagonal matrix whose diagonal consists of the degrees of the nodes of G where a node's degree is the number of edges it is connected to.

The final step in translating the diffusion model onto a road network is incorporating a Laplacian operator. The graph's *discrete Laplacian matrix*, denoted with \mathcal{L} , is defined as the degree matrix subtracted by the adjacency matrix:

$$(2.8) \quad \mathcal{L} = \mathcal{D} - \mathcal{A}.$$

From, \mathcal{L} we define the *random walk normalized Laplacian* denoted as \mathcal{L}^{rw} (see [27]). This matrix is a modified Laplacian matrix where the elements of each row are divided by the degree corresponding to each row. Figure 1 demonstrates these matrices in an example.

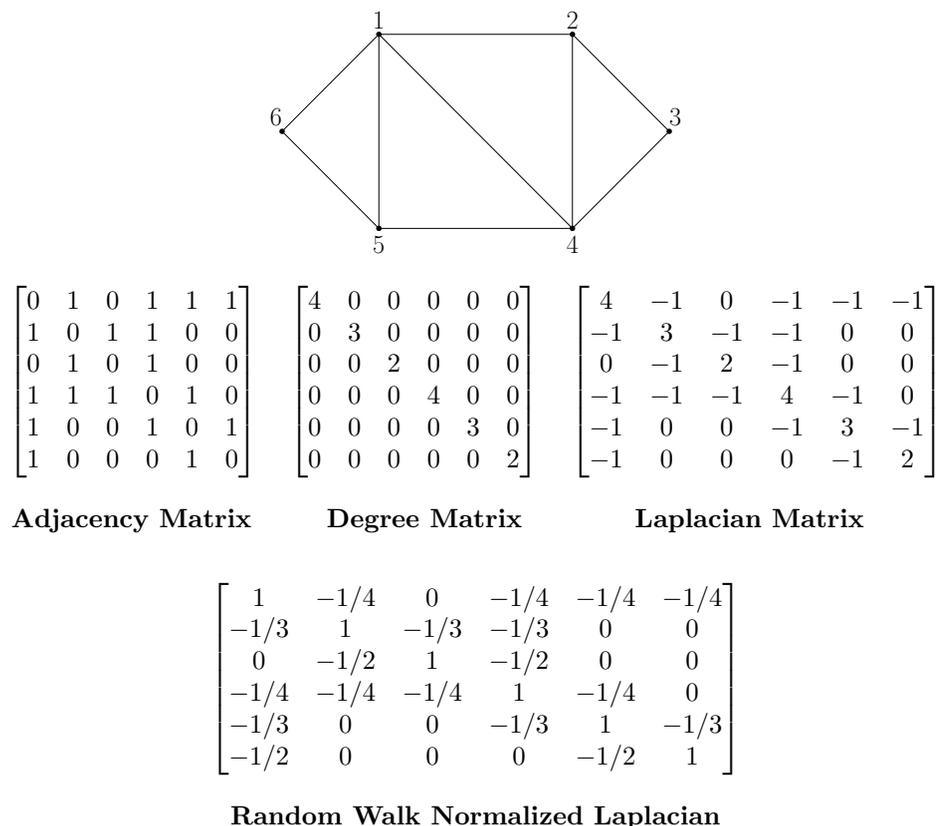


Figure 1: An example of a graph with 6 nodes and its adjacency, degree, and Laplacian matrices.

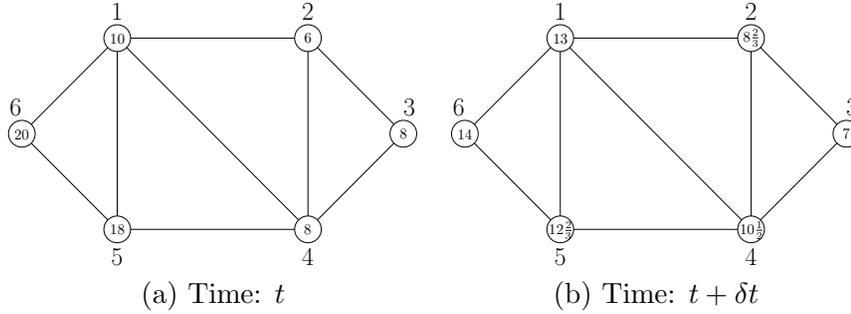
To apply 2.2 to a street network, let s be an arbitrary site with n -many neighbours. After a time increment of δt , the site's attractiveness spreads out evenly to all of its neighbours. That is, $B_s(t)$ loses $B_s(t)$ units of attractiveness while each of its neighbours receives $B_s(t)/n$ units after a δt timestep. Moreover, $B_s(t)$ increases by what its neighbours contributed after the given time step. This is done using the *random walk normalized Laplacian* \mathcal{L}^{rw} , using the update formula (see [27]):

$$B_s(t + \delta t) = B_s(t) - \mathcal{L}_s^{rw} \cdot B(t),$$

where $B_s(t)$ is subtracted by the dot product of \mathcal{L}_s^{rw} , the s^{th} row of the \mathcal{L}^{rw} matrix, and $\vec{B}(t)$. Furthermore, this generalizes to an operation which updates the entire system, using the formula:

$$(2.9) \quad \vec{B}(t + \delta t) = \vec{B}(t) - \mathcal{L}^{rw} \vec{B}(t).$$

In Figure 2 we illustrate Equation (2.9) with an example, using the graph from Figure 1 with an arbitrary initial state of substance allocation.



$$\vec{B}(t + \delta t) = \vec{B}(t) - \mathcal{L}^{rw} \vec{B}(t) = \begin{bmatrix} 10 \\ 6 \\ 8 \\ 8 \\ 18 \\ 20 \end{bmatrix} - \begin{bmatrix} 1 & -1/4 & 0 & -1/4 & -1/4 & -1/4 \\ -1/3 & 1 & -1/3 & -1/3 & 0 & 0 \\ 0 & -1/2 & 1 & -1/2 & 0 & 0 \\ -1/4 & -1/4 & -1/4 & 1 & -1/4 & 0 \\ -1/3 & 0 & 0 & -1/3 & 1 & -1/3 \\ -1/2 & 0 & 0 & 0 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 6 \\ 8 \\ 8 \\ 18 \\ 20 \end{bmatrix} = \begin{bmatrix} 13 \\ 8\frac{2}{3} \\ 7 \\ 10\frac{2}{3} \\ 12\frac{2}{3} \\ 14 \end{bmatrix}$$

Figure 2: Diffusion on a graph with 6 nodes. (a) The graph from Figure 1 with an initial allocation of attractiveness. (b) The attractiveness allocation after one time step δt .

Thus, using Equation (2.6) in place of $E_s(t)$, and Equation (2.9) in place of the diffusion operation in Equation (2.2), we derive:

$$(2.10) \quad \vec{B}(t + \delta t) = \left[\vec{B}(t) - \eta \mathcal{L}^{rw} \vec{B}(t) \right] (1 - \omega \delta t) + \theta \vec{E}(t).$$

3. Offender Mobility. In addition to the attractiveness field, the models in the diffusion papers have criminal agents which move throughout the city searching for opportunities to commit crime. During a given time interval between t and $t + \delta t$, a criminal agent is able to

perform one of two actions: commit a crime, or move to a new site in the city. Committing a crime during the time interval is a random event given by the probability

$$(3.1) \quad p_s(t) = 1 - e^{-A_s(t)\delta t}.$$

This probability follows the standard Poisson process in which the expected number of criminal events during the time interval δt is $A_s(t)\delta t$. Once an agent has committed a crime, they are removed from the space. This can be thought of as an offender laying low or being caught by police enforcement after their crime.

Under the circumstance that the offender chooses not to commit a crime, they will pursue a crime at a different site and will need to move. Their movement behaviour is random but biased towards areas of higher attractiveness. The behaviour of this movement is treated differently in each of the diffusion papers. In our research, we will test Biased Random Walks, Biased Truncated Levy Flights and two hybrid biased mobility models, which incorporate Biased Random Walks and Biased Truncated Levy Flights as they will be explained in subsequent sections.

3.1. Biased Random Walks. The offenders in Short et al. [31] treat the movement as Biased Random Walks (RW¹). In RWs, agents are only able to move to neighbouring sites, however, this movement is weighted towards neighbours with higher attractiveness. In Short's model, the distance between neighbouring sites is equidistant fixed at a constant length ℓ to reflect reasonable movement speed of criminals in a given time interval of length δt . The probability of an agent moving from a site s to a neighbouring site n is given as

$$(3.2) \quad q_{s \rightarrow n}(t) = \frac{A_n(t)}{\sum_{s' \sim s} A_{s'}(t)},$$

where the notation $s' \sim s$ indicates all of the sites neighbouring site s . In our model, neighbouring sites are not equally spaced apart due to the constraints of the real world. However, our RWs will still be treated as such and therefore analogous to Short's implementation.

By nature, RWs limit the range in which an agent can travel. This has its benefits computationally due to the low number of search cases and in some sense it corresponds to real world scenarios. It is widely acknowledged that offenders find most of their opportunities for crime in local areas surrounding locations they routinely visit such as work, places of recreation, and especially their home area [2, 3, 6]. These local areas are known as awareness spaces [11]. Offenders also are less motivated to travel far away from their homes to commit crimes and so there is a tenancy to stay close to home. Therefore, if we want the model to be representative of reality, agents in the model need to stay close to their awareness spaces when searching for targets.

3.2. Biased Truncated Levy Flights. However, it is difficult to recreate agents' home, work, and locations of recreation without the suitable data. This sensitive data is private and is not possible to obtain for an entire population. This means that in a simulation model of crime that lacks this sensitive data, agents need to be able to appropriately move large

¹It is understood that all mobility frameworks are biased and so any abbreviations of the introduced frameworks will not include a B to reference their biased nature.

distances towards their awareness spaces. Chaturapruek et al. [15] and Pan et al. [25] address this issue by giving their agents the opportunity to take large jumps through Biased Levy Flights and Biased Truncated Levy Flights, respectively.



(a) Biased Random Walk (RW)

(b) Biased Truncated Levy Flight (TLF)

Figure 3: Biased Random Mobility Models

Levy flights are a random motion in which agents take steps of lengths which obey a fat tailed Pareto inverse power law probability distribution [28]. Pareto inverse power law distributions are characterized by specific exponents α , the Pareto index, and are given by the equation:

$$(3.3) \quad P(x) = \begin{cases} \frac{\alpha x_M^\alpha}{x^{\alpha+1}} & x \geq x_M \\ 0 & x < x_M, \end{cases}$$

where x is the length of the step and x_M is the minimum possible positive value of x . In the context of crime, x is the distance from site s to another site k in the city. Distances in the city can be measured in several metrics such as the Euclidean Metric (or the distance as the crow flies), and street distances. Hence, in general, the distance between site s and site k is given by $d(s, k)$ under the given metric of the model.

The purpose of x_M in equation (3.3) is to avoid the distribution blowing up as $x \rightarrow 0$. Furthermore, it supports $x_M \rightarrow \infty$, so that integrating across gives $\int P(x)dx = 1$ to justify its use as a probability density function. In our model, agents always move away from their starting current site and all neighbouring sites are at least 10m away using our data. Hence, $x > x_M, \forall x$ in our street network. Therefore we may ignore the second case in Equation (3.3) and x_M can be any number smaller than 10. For simplicity, we will let $x_M = \frac{1}{\alpha}$ given that $\alpha > \frac{1}{10}$ so that the numerator will equal to 1. Setting $\mu = \alpha + 1$, we obtain:

$$(3.4) \quad P(x) = \frac{1}{x^\mu}.$$

Similar to RWs, Levy flights are biased towards areas of high attractiveness in our model.

We are then, in fact, interested in Biased Levy Flights (LF) where the probability of an agent to move from site s to site k , where $s \neq k$, is given by

$$(3.5) \quad P(s \rightarrow k) = \frac{w_{s \rightarrow k}}{\sum_{k' \neq s} w_{s \rightarrow k'}},$$

where $w_{s \rightarrow k}$ is the relative weight of a criminal moving from site s to site k , given by

$$(3.6) \quad w_{s \rightarrow k} = \frac{A_k}{d(s, k)^\mu}.$$

An unrealistic downside of the traditional probability distribution for LFs is that it is possible for agents to travel infinite distances. This is undesirable in realistic simulations for human mobility especially with the boundaries of a city. In light of this, the weighted probability defined in Equation (3.5) restricts LFs to distances within the city. This is because areas outside of the constraints of a city have zero attractiveness contributing no weight in the probability sample space. To further restrict movement, one can limit the maximum distance a step size can be using Biased Truncated Levy Flights (TLF). The key difference between TLFs and LFs is in how the weighting is given. In TLFs, the relative weight $w_{s \rightarrow k}$ of a criminal moving from site s to site k is given by

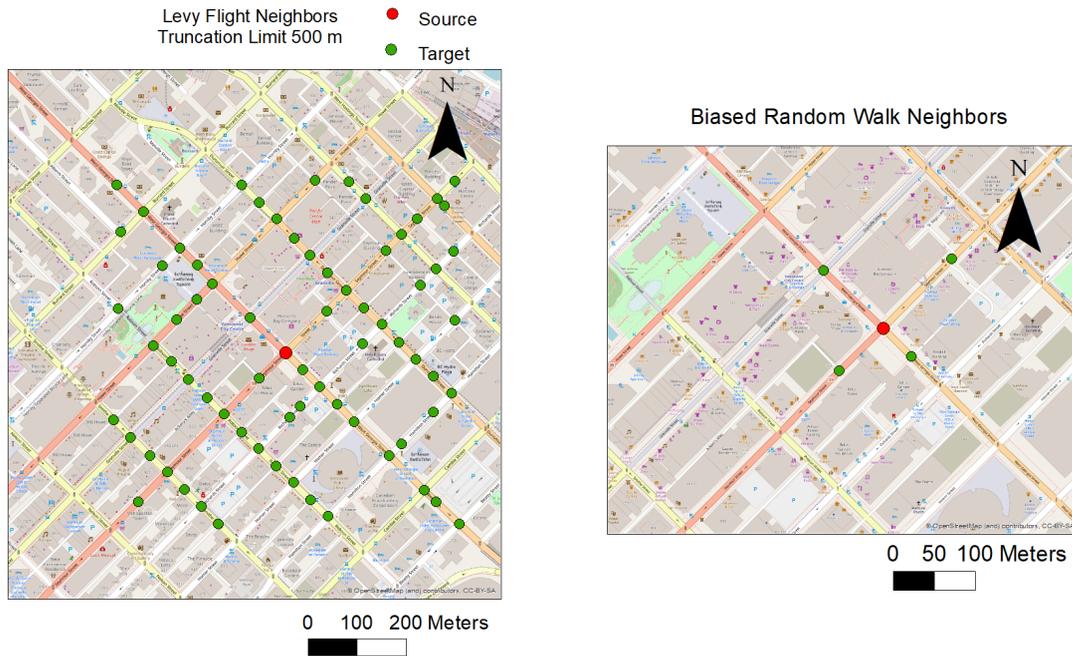
$$(3.7) \quad w_{s \rightarrow k} = \begin{cases} \frac{A_k}{d(s, k)^\mu} & \text{if } 0 < d(s, k) \leq L \\ 0 & \text{otherwise,} \end{cases},$$

where L is the maximum distance, or truncation, of the TLF.

3.3. Hybrid Biased Mobility Models. In addition to RWs and TLFs, we also introduce 3 hybrid models which incorporate elements from both of these two standards: Biased Random Walk then Biased Truncated Levy Flight (RW_TLF), Biased Truncated Levy Flight then Biased Random Walk (TLF_RW), and Biased Truncated Levy Walks (TLW). As per their namesakes, the first two hybrids alternate our previously introduce standards. RW_TLF, the first of these two hybrid models, constrain the agents to alternate between RWs for a limited number of steps and one step of TLF before going back to a RW movement for the same number of steps. The converse is the case for TLF_RW. We also introduce the TLW which is similar to the TLF in which offenders are able to take jumps to other nodes, however, in the TLW the offender will check each node along their path to the other node for potential targets.

3.4. Mobility Frameworks Adapted to a Street Network. The mobility frameworks based on the Levy Flights are dependent on the distances between sites in space. To implement these frameworks, we must utilize a distance metric which is based on street distances. This is straightforward using the network's distance metric and Dijkstra's algorithm. Using both, we compute a large matrix of distances between all points using Dijkstra's algorithm which we call the Dijkstra Distance Matrix. Although this initial calculation is computationally expensive, it can be reused for different truncation limits with minimal computations. With the Dijkstra Distance Matrix, we can find all the sites within a neighbourhood of radius, or truncation limit, $L > 0$ around any given site s . We note that such a neighbourhood is non-circular and is more akin to a square, as illustrated in Figure 4a.

Traditionally, RWs are a special case of TLF given a significantly small truncation limit $L > 0$ and equidistant sites in space. However, the topographical challenges of a city presents difficulties in creating constant spacing between different nodes. Street blocks cannot be divided universally into equidistant street segments and as such, it is impossible to perform true RWs on a street network. Our implementation of RWs still utilizes distances between neighbours in the same fashion as TLF, however, we ignore the truncation limit L as it is sufficient to simply look at a site’s neighbours, as presented in Figure 4b . Alternatively, one can simply use the adjacency matrix and have equal probability for neighbouring nodes as an approach to implementing RWs.



(a) Neighbourhood of radius 500m.

(b) Neighbourhood with immediate neighbours

Figure 4: Neighbourhood for different mobility frameworks.

4. Methodology. We now proceed with an explanation of our methodology. As mentioned above, our model is based off of the frameworks described by the diffusion papers [15, 25, 30, 31]. Similarly to those models, we need to initialize our attractiveness fields and run simulations on it. More specifically we need to build this attractiveness field on the network that is representative of the real world, which requires initializing it with empirical data.

4.1. Data. To be as representative of the real world as possible, we incorporate Vancouver’s GIS (geographic information system) data into our model. Although we present results specific to the city, our methodology can be used for other cities provided access to the ap-

appropriate data resources. The data used in our model consists of Vancouver’s street network, city and neighbourhood boundaries, population data, and crime data.

Vancouver’s topography is unique as it is a peninsula with several public parks. As such, its road network does not form a perfect grid. Instead, its roads come in varying shapes such as several curved roads and one major road running on the diagonal from its downtown area to other local municipalities. We obtain Vancouver’s road network and boundary data through OSMnx [4], a Python package which queries data from OpenStreetMap², where 15,149 unique sites from Vancouver’s road network were extracted to form the foundation of our graph. These sites are street intersections and points in between these intersections used to interpolate curved streets. The average distance between these sites is 75.6m. Figure 5a shows the sites located in Vancouver downtown.

To supplement this road network, we also incorporate the city’s neighbourhood boundary and population data. Vancouver consists of 22 city defined neighbourhoods used for administrative purposes. The boundary data associated with these neighbourhood was obtained from Vancouver’s Open Data Catalogue³ and will be used in our Macro Analysis of the model. In addition to these neighbourhoods, Statistics Canada partitions Vancouver into 996 smaller neighbourhoods known as Dissemination Areas (DA). Dissemination Areas are constrained by the rules that they are consistent in population size and follow the natural boundaries of Vancouver’s street network. Of course, population size varies from DA to DA along with the number of sites associated to it. What is particularly nice about these discrepancies is that we are able to use it to calculate a population density associated with any particular site. For a given site in a DA, the population density associated with it is the population count of the DA divided by the number of nodes in the given DA. In our model, we use Vancouver’s population density in order to generate offenders on the network proportional to the given population. We present a map of Vancouver’s population density in Figure 5c below.

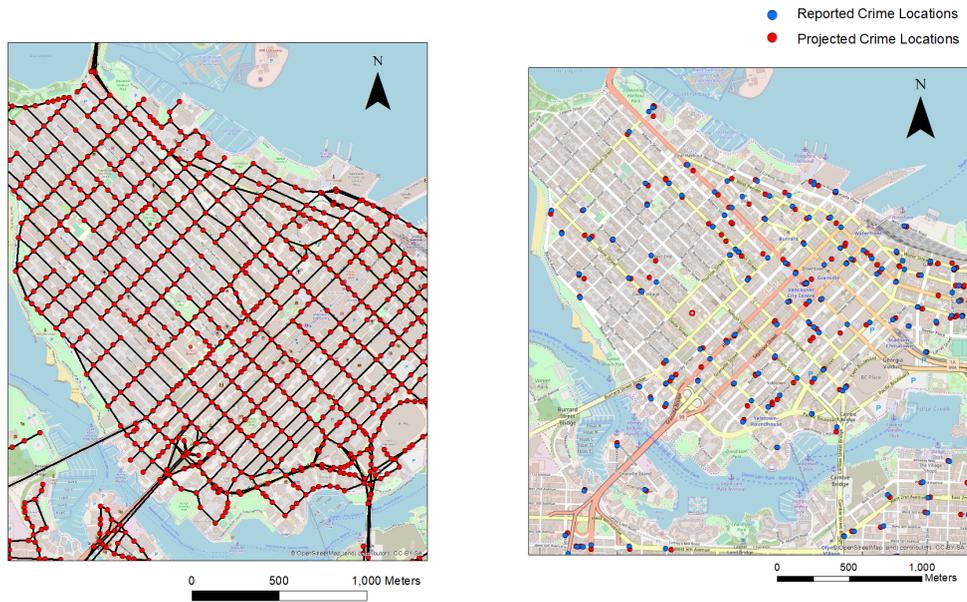
Finally, we require Vancouver’s empirical crime data to initialize our attractiveness backcloth using Equation (2.10) and to test our model against. The crime data used was collected by the Vancouver Police Department [38] covering the time period January 2014 to January 2020. This data consists of various types of crimes but are only representative of reported offences. The offences were represented with point data projected using the Plate Carrée projection, which is the same projection used in the road network. This criminal data was then added to vertices on the graph using points which were closest to where the crimes were committed, as seen in Figure 5b.

4.2. Initializing the Attractiveness Field with Empirical Data. We use Equation (2.10) to recursively initialize our attractiveness field for our simulations. We set $\vec{B}(0)$ to be 0 everywhere on the street network and $\delta t = 1$ to represent a single day. Setting $t = 0$ on January 1st, 2014, we use the crime data to find the locations of crime on this date and update $\vec{E}(t)$ accordingly. That is, if a crime is recorded at site s in the data, $E_s(t)$ is incremented by 1. Therefore, the attractiveness field on January 2nd is updated using $\vec{E}(0)$, to get:

$$(4.1) \quad \vec{B}(1) = \theta \vec{E}(0)$$

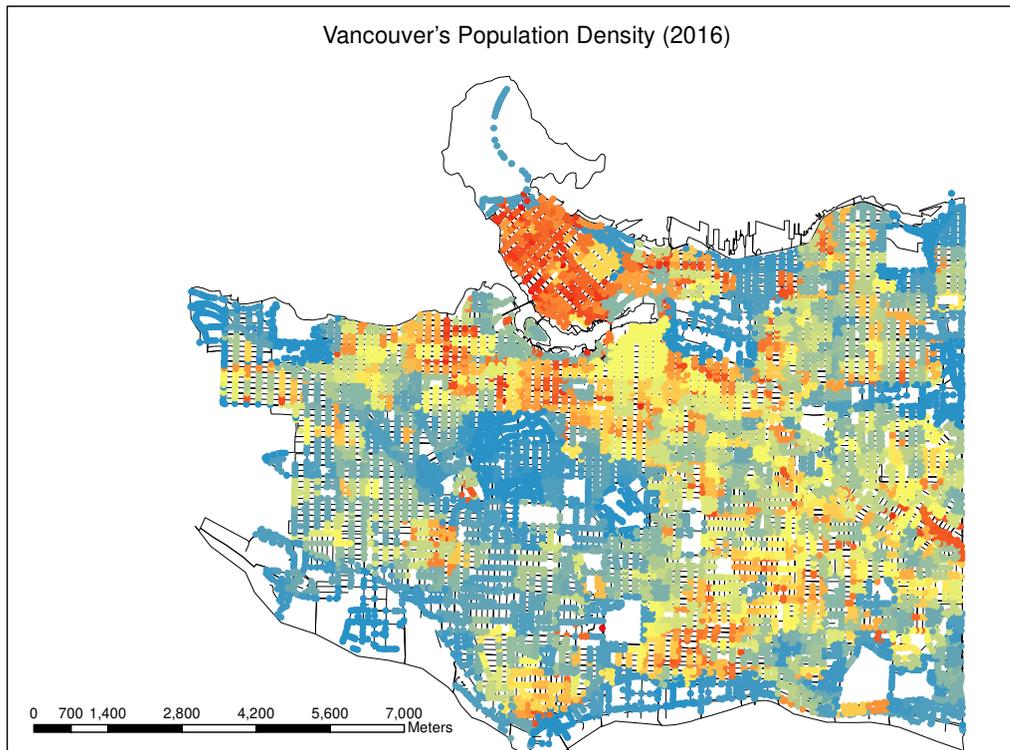
²OpenStreetMap <https://www.openstreetmap.org/>

³Vancouver Open Data Catalogue <https://opendata.vancouver.ca/pages/home/>



(a) Sites on Vancouver's street network.

(b) Matching crimes to closest site on street network.



(c) Vancouver's Population Density in 2016

Figure 5: Vancouver's data used in the model.

Using the crime data from January 2nd, we similarly get $\vec{E}(1)$ and then,

$$(4.2) \quad \vec{B}(2) = \left[\vec{B}(1) - \eta L^{rw} \vec{B}(1) \right] (1 - \omega) + \theta \vec{E}(1).$$

This process is continued for each subsequent day until we reach the end of our initializing data. The final attractiveness field is then used as a seed in which we run our simulations. In our model, we initialize our seed until December 31st, 2019. Thus, initialization is done through a sequence $B(1) \dots B(\text{Dec 31 2019})$. We stop the initializing here as we wish to compare the results of our simulations to the real world, leaving us a month of data left to use. Note that this initializing can be continued further, but given the unpredictable nature of the COVID-19 pandemic, we did not wish to go further into crimes committed in 2020.

4.3. Parameter Setup. The final step in creating the attractiveness field is to set up values for the parameters η , ω , and θ (the diffusivity constant, decay rate, and exciting constants respectively), as they are key components in Equation (2.10). In this introduction of the model, we propose an experimental approach for finding optimal parameters. The experimental technique we present requires the groupings of different parameter assignments and comparing each groupings' results with empirical data. These experiments are run using the time-step process of the model which we explain in the subsequent section.

The decision to choose these values comes as a consequence of different optimizations. For example, η and ω are parameters with domain restrictions $\eta \in (0, 1]$ and $\omega \in [0, 1/\delta t]$. As such, the values for these parameters are chosen to convey different orders of magnitude. That is, in our testing we will be using $\omega = 0.01$ and 0.1 and $\eta = 0.01, 0.1$, and 1 as they differ in orders of magnitude of 10. As θ is the only parameter which increases the attractiveness field, we should set it so that it is optimal to the empirical data. Table 1 contains all the parameter values we consider, including the variable diffusivity values, which will be explained below.

To determine a starting point for θ based on our other parameters, the expected number of crimes in our simulation should correspond to the expected number of crimes at a given site in the real world. From Equation (3.1), we have that the expected number of crimes to occur at a given site s is $A_s(t)\delta t$. When looking at the crime in Vancouver for the month of January 2020, the average number of crimes over all sites was 0.21. In initial trials with the attractiveness field, we found that $\theta = 1$ achieved similar expected values with respect to other parameter choices. That is, upon calculating the expected value across all the sites, the average was close to 0.21. This is not to say this is the optimal value for θ but simply a starting point for our tests. To further examine the effects of θ , we will also test values with orders of magnitude of 10, smaller than $\theta = 1$. We found that values of θ larger than 1 resulted in our model being extremely biased towards a few sites and that crime was not remotely realistic.

A final technical choice of parameterization in our methodology is creating variable diffusivity which is spatially dependent. In the previous studies on the subject, all the parameters in Equation (2.10) were chosen as constant values [15, 25, 30, 31]. In particular, the attractiveness field's diffusivity η , which determines the effect of a crime on its neighboring nodes, was also constant at every site. In reality, the effects of the spread of common knowledge varies greatly and this value should reflect this dynamic property. This is especially true in different neighbourhoods and areas of the city and so, diffusivity is spatially dependent. In

this research we propose the use of a diffusivity that varies from neighbourhood to neighbourhood. Our approach in creating dynamic diffusivity is by using empirical recordings of crime and the different neighbourhoods of the city.

4.4. Variable Diffusivity. To create variable diffusivity we consider a neighbourhood of a city which we will refer to as a cluster. Let c be a given cluster and using recorded instances of crime in c , let η^c be the corresponding diffusivity value for that cluster defined as:

$$(4.3) \quad \eta^c = \frac{\text{nodes affected by crimes in } c}{\text{total number of nodes in } c}.$$

Let the variable diffusivity η_v be the collection of $\{\eta^c\}$, where c ranges from 1 to N , the total number of clusters. By choosing η_v in this manner, we get higher values of diffusivity in clusters where crimes occur more frequently, and lower values in neighbourhoods where crimes occur in fewer nodes. In practice, over a long period of time we see that most areas of the city are affected by crimes. Thus, Equation (4.3) should be used for a relatively short time interval. We apply equation (4.3) to one month of data, December 2019.

It is worth noting that that number of clusters one uses and the methodology to define said clusters can vary drastically. Cities can define clusters based on social, economic, or topographical factors or even based on population density. It is often the case that the number of clusters is dependent on these reasons but it can also be quite arbitrary.

In this paper we are interested in the effects of variable diffusivity on a large scale across the entire system. This choice is in part due to the novelty of this approach and our desire to understand the broad effects of this strategy. Hence, we use the 22 city defined neighbourhoods which relates to our macro scale analysis of our model.

We also wish to understand variable diffusivity relative to our constant values. In our experiments, we use $\eta = 1, 0.1$, and 0.01 . To ensure similarity with these values, we will explore two additional sets of variable diffusivity $\eta_{0.1}$ and $\eta_{0.01}$ which takes our original set η_v and normalizes each cluster specific diffusivity values so that the global mean is equal to 0.1 and 0.01 respectively. That is,

$$\eta_{0.1}^c = \frac{\eta^c}{\text{mean}(\eta^c)} \times 0.1$$

$$\eta_{0.01}^c = \frac{\eta^c}{\text{mean}(\eta^c)} \times 0.01.$$

The variable diffusivities $\eta_{0.1}$ and $\eta_{0.01}$ are the collections of $\{\eta_{0.1}^c\}$ and $\{\eta_{0.01}^c\}$, respectively. Note that we cannot normalize η_v so that the global mean is 1 as that would create diffusivity values larger than 1.

We recognize that it is not always applicable to use city defined clusters for macro, meso, and micro analysis or for creating these variable diffusivity sets. It is our intent in this article is to present a model which can be generalized to other cities, and not all cities have neighbourhood and district boundary data readily available. To remedy this, we propose using the K-means clustering algorithm to define clusters in the city.

4.5. K-means Clustering. K-means Clustering is an algorithm that partitions spatial data into k clusters [21]. These k clusters are determined in such a way that spatial variance is minimized. This type of clustering has many applications in statistical data analysis, pattern recognition, and geography.

In this research, clustering is used to determine different neighbourhoods which appear in the urban landscape. It is possible that these neighbourhoods are geographically dependent to those neighbourhoods outlined by city planners but more often than not they are dependent on the cultural and economic practices of its inhabitants. Finally, there is speculation that district boundaries are drawn to benefit different political groups in a process known as Gerrymandering. K-means can avoid many of these dependencies for a neutral look at the city but if one is interested in their effects, this algorithm can be applied based on those factors as well. Ultimately, we wish to study the spatial nature of crime and through studying clusters which are defined spatially and free from political or cultural bias, we are able to appropriately study the diffusive dynamics of crime through a city.

One of the main benefits of clustering in this way is the flexibility and universality it provides. Some cities like Vancouver have neighbourhood or district boundary data readily available but this is not always the case. Hence, a model dependent on this data cannot be generalized easily. Fortunately, through K-means and other similar clustering algorithms one does not need rely on this data. Thus, in a model which aims to be generalizable to all cities, the universal nature of clustering algorithms is far more practical.

We will present findings for clusters defined spatially through the K-means algorithm without any special considerations to additional data, such as demographics and geographic location. In our initial trials, results using spatially group clusters had negligible differences from ones using predefined clusters. Furthermore, we continued to use 22 and 996 clusters for the city of Vancouver. These clusters will be used for our macro and meso analysis respectively and our micro analysis will be done on a per site basis. In the absence of the number of predefined clusters by the city, one can apply an elbow analysis to systematically find an optimal number of clusters. In our research, choosing the same number of clusters as the number of clusters defined by the city, allowed us to compare our results with those of predefined clusters. The choice of number of clusters is an area that could be explored further in future research.

4.6. Simulation Model.

4.6.1. Simulation Process. Using the seeds resulted from initializing the attractiveness field with empirical data, we model crimes using an agent based simulation process. These simulations follow a simple time-step process. Let $n_s(t)$ be the number of offenders at site s at time t . Then, the offenders throughout the whole system is described by the vector

$$(4.4) \quad \vec{n} = [n_1 \dots n_m]^T,$$

where we recall that m denotes the number of sites on the network. The vectors $\vec{n}(t)$ and $\vec{E}(t)$ (see Equation (2.6)) evolve according to the following 4 step process. We present a visual summary of this four step process in Figure 6.

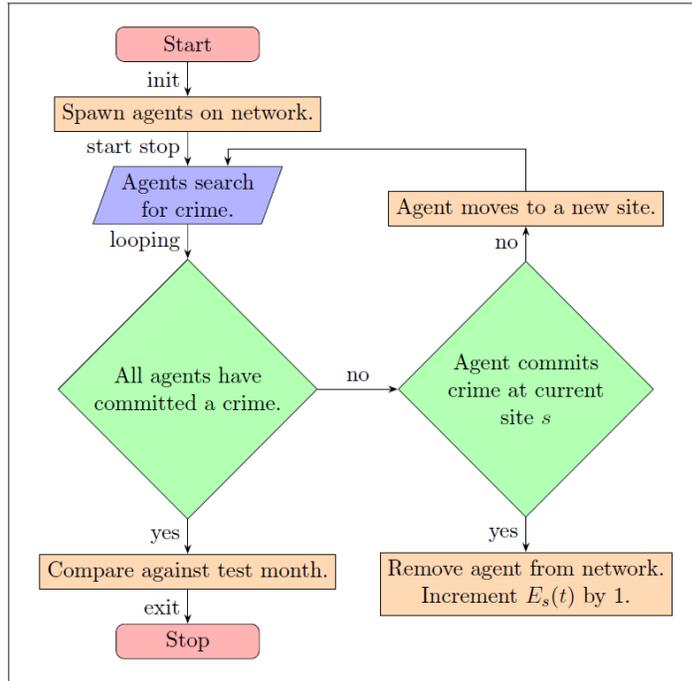


Figure 6: Flowchart summarizing the model

Step 1. Offenders are added randomly to the network, weighted by the city’s population density. The number of offenders added has a one-to-one correspondence with the recorded number of crimes for the given month.

Step 2. Each criminal on the network makes a decision to commit a crime at their given site with the probability defined in Equation (3.1).

Step 3. If an agent decides to commit a crime at their current site s , they will be immediately removed from the system and $n_s(t)$ will be decremented by 1. This can be thought of the agent fleeing the site of the crime. If a crime is committed at site s , we increment $E_s(t)$ by 1. If an agent chooses not to commit a crime, they will move to another site according to the mobility framework which is being tested.

Step 4. After every criminal agent has committed a crime, we record the crime locations and compare them to the empirical data of the test month.

We note that in the work of Short et. al., [31], the authors added offenders to the network based on a constant rate Γ . In their model, criminals are removed upon committing a crime but are replaced at a rate Γ at each location. Our approach is to assign N criminals (N being the number of crimes committed over the test month) to the network and have them search until they have committed a crime. For the purposes of introducing this model, this is sufficient as it enables us to directly test the simulation. We require a one-to-one correspondence in order to appropriately evaluate our results as we are interested in the behaviour of the offenders. Without this injective correspondence, our results would be inconsistent especially in its Monte Carlo setting.

However, this is not without flaws. For instance, we do not make predictions as we will not know how many crimes there will be in the future. This can be overcome using probability density functions which give an estimate for the number of crimes which will occur. We also initialize the attractiveness field daily while simulating crimes over a month. This means that all the crimes could be committed in a single day in any given day within the month. This is a limitation of the data available. Testing over a single day would be ideal but given our setup, our evaluation would be misleading. With so few crimes occurring on a single day and over 15,000 sites to evaluate, our error would be virtually non-existent resulting in poor analysis. One can also perform these tests over the span of a year but this would provide too many crimes would be computationally expensive. Thus, a one month test was reasonable to introduce this model.

4.6.2. Evaluation Methods. Our simulations test the validity of the model, the effectiveness of different parameter groupings, and the mobility models. We use a Monte Carlo methodology that test different parameter groupings and mobility frameworks multiple times. The parameter groupings are shown in Table 1. For each such grouping, an attractiveness field is generated using data from January 2014 to the end of December 2019. We also test every mobility framework shown in Table 2 in each grouping.

Parameter	Values
η	0.01, 0.1, 1, $\eta_v, \eta_{0.01}, \eta_{0.1}$
ω	0.01, 0.1
θ	0.1, 0.5, 1

Table 1: Parameter Values

Mobility Model	Maximum Steps in RW	Truncation Limit (in meters)
RW		
TLF		500, 750, 1000, 2500, 5000, 10000
TLW		500, 750, 1000, 2500, 5000, 10000
RW_TLF	5, 10, 20	500, 750, 1000, 2500, 5000, 10000
TLF_RW	5, 10, 20	500, 750, 1000, 2500, 5000, 10000

Table 2: Different Mobility Models Tested

Each grouping is run for 100 trials and the final state of the trials is compared to the recorded crime events for January 2020. The purpose of this comparison is to determine how representative of the real-world our model is. We assign 2,440 agents to the road network to match a one-to-one correspondence of agents to recorded instances of crime.

We have chosen January 2020 as the test month, since it was the last month before the impacts of Covid-19 became noticeable, and changed the crime patterns [24]. More specifically, the pandemic has changed the mobility patterns of criminals as well. The intent of our research is not to examine the effects of the pandemic on crimes and the mobility of criminals; however, we acknowledge our current mobility framework may not be applicable during the times when mobility has been limited.

4.6.3. Evaluation Metrics. Our method in determining whether our simulation model can reasonably predict crime is through a cross-validation strategy. Cross-validation involves the use of a training dataset used to initialize our attractiveness backcloth and a validation dataset to compare our simulations against. These are referred to as the training dataset and the validation dataset respectively. In this work, the training dataset is the set of all of Vancouver’s crime data from January 2003 to December 2019 and the validation dataset is January 2020. To compare our simulations to the validation dataset, we need to develop metrics to measure our model’s performance. For our purposes, we propose 2 criteria for our metrics: scope and accessibility.

In our metric, scope refers to analyzing error at macro, meso, and micro levels. An ideal model should be able to preserve the crime rates at different neighbourhood scales. For instance, if we find that 30% of crimes occur in the downtown area of a city and a model predicts only 5% in the same area then we can conclude that the model performs poorly. Additionally, if the model can predict crime well at the scale of an intersection but fails to do so in its surrounding neighbourhood we would not be able to say that the model performs well.

Therefore, we present the simulation results at three different scales. The first one is at the per-node level; this is the smallest (micro) scale. The other two scales (meso and macro scales, respectively) are determined by the 22 neighborhoods and the 996 dissemination areas defined by the City of Vancouver. Of course, this is if one has this data readily available to do their analysis. The alternative would be to use K-means clustering in order to create clusters so that we can do analysis at the meso and macro scales. In fact, in our results we found little difference between performing our scope level error analysis using predefined boundaries and K-means defined ones. As such, we will only be presenting the results from the K-means defined boundaries.

At these different scales, we are interested in determining the accuracy and likeness of our simulation to the validation set. To compute the accuracy of our model, we consider using either Root Mean Square Error (RMSE) or Mean Absolute Error (MAE). Both these error metrics calculate errors where crime is the unit, and are the most prolific metrics when describing average model-performance error [13, 14, 44]. Of the two metrics we have opted to use MAE due to its unambiguous nature. Willmott et al. [44] argue that MAE directly describes a model’s average error while RMSE creates more variable error magnitudes as it is based off of the sum of squares. Brassington [13] suggest that RMSE may be a better metric than MAE given that we assume that our error is Gaussian. But due to the spatial and temporal nature of crime, we cannot assume that the error in the implementation of our model is Gaussian.

Thus, we will use MAE as our error metric at each of our 3 scales. Given N clusters in a given scope, let V_i and S_i ($i = 1, \dots, N$) be the numbers of empirical and simulated crimes in the i^{th} cluster, respectively. MAE is then defined as

$$(4.5) \quad \frac{1}{N} \sum_{i=1}^N |V_i - S_i|.$$

The second evaluation metric we utilize is the Pearson Correlation Coefficient, denoted by ρ . This evaluation metric ρ tells us how correlated our simulated set is to our validation

set. Denote by $V = \{V_1, \dots, V_N\}$ and $S = \{S_1, \dots, S_N\}$ the collections of all empirical and simulated crimes in the N clusters, respectively. Also, let \bar{V} and \bar{S} be the mean of V and S , respectively. Then, the correlation coefficient ρ is defined as

$$\rho = \frac{Cov(V, S)}{\sigma_V \times \sigma_S},$$

where $Cov(V, S)$ is the covariance between V and S given by

$$Cov(V, S) = \frac{1}{N} \sum_{i=1}^N (V_i - \bar{V})(S_i - \bar{S}),$$

and σ_V , σ_S are the standard deviations of the crime events, given by

$$\sigma_V = \sqrt{\frac{\sum_{i=1}^N (V_i - \bar{V})^2}{N}}, \quad \sigma_S = \sqrt{\frac{\sum_{i=1}^N (S_i - \bar{S})^2}{N}}.$$

Typically ρ takes values between -1 and 1 and values closer to either end correspond to greater correlation. In our analysis, ρ will indicate how correlated our experimental results are in relation to its empirical counterpart. We will take $\rho > 0.68$ to be a strong relationship in our analysis as is commonplace in statistics [36].

5. Experimental Results. We now present the findings of our experiments. Our aim in this discussion is to demonstrate what effects each of our experimental parameter has in our implementation of the diffusion model of crime on a real world road network and its corresponding crime [15, 25, 30, 31]. The results shown will be based on our evaluation metrics: the Pearson Correlation Coefficient and the Mean Absolute Error.

The values of the evaluation metrics are presented over the macro, meso, and micro scales. At each scale, we also show their respective minimum (min), mean, median, and maximum (max) values over all parameter groupings. The best value for each metric at each scale is bolded. The most optimal values for the min, mean, median, and max categories are highlighted. This is to help us gauge the consistency of the parameters.

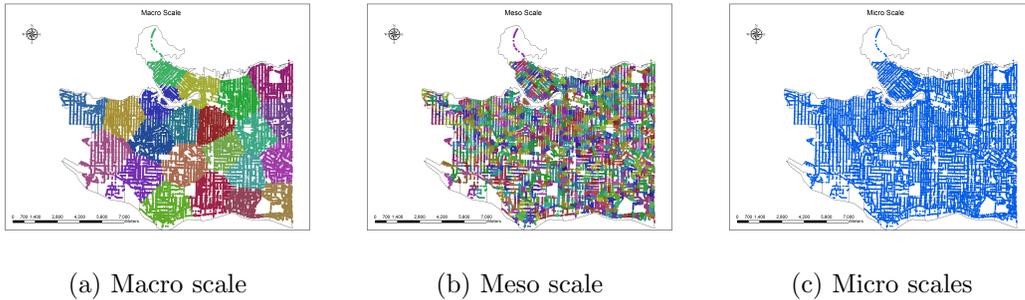


Figure 7: Different Scopes used in Error Analysis.

5.1. Baseline Control for Comparison. To better gauge the effectiveness of our model, we require a control experiment as a point of comparison. This control serves to see whether our simulations fit the data and the realistic nature of crime in Vancouver. To create this baseline, we will need to generate crimes at victimized sites and assess them against our empirical data using our evaluation metrics. There are multiple ways to generate the crimes. These include:

1. Choosing crimes uniformly at random;
2. Choosing crimes weighted by the population density; or
3. Choosing crimes weighted by our existing parameterized attractiveness fields.

In our tests, we found that the first option was terrible as it achieved correlation values around 0 across the macro, meso, and micro levels as expected. The second option was only marginally more appropriate as it achieved correlation values of 0.61, 0.42, and around 0 at the macro, meso, and micro levels respectively. This would suggest that there is some general relationship with crime rates in each of Vancouver’s neighbourhoods that does not apply directly to their crime sites. The most critical option is by far the third method. Through this method, we achieved impressive results using the parameterization $\theta = 0.1, \omega = 0.1$, and $\eta = 0.01$ as seen in Table 3.

	Macro	Meso	Micro
Pearson Correlation Coefficient	0.997	0.954	0.645
Mean Absolute Error	10.238	1.721	0.206

Table 3: Baseline to compare simulation results. Victimized sites were chosen weighted by our existing parameterized attractiveness fields and no mobility framework. These sites were then compared to our empirical data using our Pearson Correlation Coefficient and Mean Absolute Error evaluation metrics. The parameterization of this baseline was done using $\theta = 0.1, \omega = 0.1$, and $\eta = 0.01$.

We can see that using the diffusion model alone is effective in predicting crime as we achieve near perfect correlations at the macro and meso levels. We also achieve a 0.645 correlation at the micro level which is considered very high within the social sciences. Despite this, we are also interested in studying the mobility of criminals in an urban setting. As such, our simulations will employ agents who search for crimes over our network. These agents will not be generated in the same fashion as this baseline but will instead be placed on the network based on the population density. From these start locations, they will move towards areas where they believe they can commit crimes successfully. We argue that this is more realistic as agents do not have perfect knowledge of the city to begin with and that offenders must live somewhere. Ideally we would have the agents start at known criminal home locations but this data is inaccessible. Regardless, the results from Table 3 will serve as our baseline for comparison.

5.2. Effects of Parameters.

5.2.1. Effects of Near-Repeat Victimization. We begin by discussing the constants ω and θ which have a direct influence on the Near-repeat victimization phenomenon (NRV) as described in the diffusion papers [15, 25, 30, 31]. The constant θ is the value associated with an increase in a site’s vulnerability due to a recent victimization and ω is the decay relating

to increased protection, intervention, or simply forgetting about the site.

	Pearsons Correlation Coefficient				Mean Absolute Error			
	Macro				Macro			
	min	mean	median	max	min	mean	median	max
0.1	0.531	0.739	0.674	0.993	19.676	63.447	70.426	73.027
0.5	0.567	0.693	0.631	0.987	20.355	67.427	72.040	73.294
1	0.571	0.682	0.624	0.985	21.412	68.464	72.281	73.341

	Meso				Meso			
	min	mean	median	max	min	mean	median	max
0.1	0.278	0.466	0.386	0.911	1.590	2.604	2.609	3.071
0.5	0.259	0.397	0.328	0.866	1.775	2.666	2.687	3.037
1	0.252	0.381	0.310	0.832	1.896	2.675	2.707	3.005

	Micro				Micro			
	min	mean	median	max	min	mean	median	max
0.1	0.051	0.239	0.197	0.793	0.188	0.258	0.262	0.283
0.5	0.048	0.167	0.138	0.680	0.212	0.266	0.270	0.283
1	0.045	0.149	0.120	0.602	0.218	0.268	0.271	0.284

Table 4: Evaluation metrics across macro, meso, and micro scales for parameter θ . We observe that $\theta = 0.1$ was optimal over each scope.

In Table 4 we observe that our evaluation metrics are most optimal when $\theta = 0.1$. This corresponds with the baseline we see in Table 3. This trend occurs across the macro, meso, and micro scales. This suggests that the agents in our model are quite sensitive to the value of θ . We believe this is due to overt aggregation to highly victimize sites. With higher θ , agents would aggregate to areas which have been recently victimized while with smaller values we see agents clustering closer to Vancouver’s historical hot spots. We found that larger θ values corresponded with greater spread throughout the city but this was not indicative of the true nature of hot spots in the city and therefore the criminals themselves.

For our decay coefficient ω , we similarly observe optimal values in our evaluation metrics when ω is small. In Table 5 the most consistent value of ω across our scales and metrics is when $\omega = 0.01$ as it relates to the highest correlation coefficients and the smallest MAEs. Just as we saw with θ , this is consistent with our baseline in Table 3. However, this optimally is not uniform with a discrepancy at our micro scale for our correlation coefficient. We also see that this optimality is inconsistent when judging the spread of our metrics across the min, mean, median, and max categories. Despite this, the majority of our optimal values in the evaluation metrics is when $\omega = 0.01$, especially at larger scales. This suggests a long lasting awareness of the road network’s attractiveness backcloth among our agents. This indicates that this attractiveness backcloth is perpetuated greatly influencing future crimes. Note that $\omega = 0.01$ is not the same optimal value as our baseline which was $\omega = 0.1$.

5.2.2. Diffusivity. In contrast to our previous analysis, the results for our values of η are slightly different in the sense that in the case of θ and ω the maximum of Pearsons Correlation

Pearsons Correlation Coefficient					Mean Absolute Error			
Macro					Macro			
	min	mean	median	max	min	mean	median	max
0.01	0.531	0.679	0.624	0.993	19.676	68.054	72.107	73.339
0.1	0.571	0.730	0.663	0.988	20.350	64.839	70.927	73.341
Meso					Meso			
	min	mean	median	max	min	mean	median	max
0.01	0.257	0.395	0.332	0.911	1.590	2.534	2.607	2.735
0.1	0.252	0.434	0.351	0.890	2.066	2.763	2.848	3.071
Micro					Micro			
	min	mean	median	max	min	mean	median	max
0.01	0.045	0.137	0.099	0.732	0.188	0.268	0.272	0.284
0.1	0.060	0.233	0.192	0.793	0.212	0.260	0.265	0.281

Table 5: Evaluation metrics across macro, meso, and micro scales for parameter ω . $\omega = 0.01$ achieved the highest Pearson value and lowest Mean Absolute Error over each scope consistently. This was generally consistent over each grouping $\omega = 1$ was included in as seen in the results for the min, mean, median, and max. However $\omega = 0.1$ did perform well for the Pearson metric at a micro level. Despite this, $\omega = 0.01$ generally performed better.

Coefficient and the minimum of Mean Absolute Error correspond to the same value, with the exception of ω at the micro level. We observe from Table 6 in the case of η , only the maximum of Pearsons Correlation Coefficient and the minimum Mean Absolute Error match at the micro level. This could be explained by noting the mean and similarly the median at macro, and to a lesser degree at meso, are close to each other. For instance, the difference between the highest and lowest mean value of Pearson Correlation Coefficient at the macro level is 0.008 whereas for θ and ω it is 0.057 and 0.051, respectively. What this implies is that various values of η have less effect in the current model at macro and meso level.

The reason for this lack of sensitivity to different values of diffusivity at larger scales may depend on the clustering method. Our clustering method, due to reasons concerned with generalizations, does not cluster hot spots and indeed highly active nodes, from the crime point of view, may end up in clusters with less active nodes. This implies the occurrence of a crime, regardless of where it happens, has little effect to its neighboring sites at larger scales, and overall diffusivity does not play an important role at larger scales.

However, at the micro level, extreme values of η give a completely different picture. The difference between the highest and lowest mean value of Pearsons Correlation Coefficient is 0.138. Moreover, we see lower values of η perform better. This shows that occurrence of a crime has less influence on neighboring sites. We find that $\eta = 0.01$ is the most consistent parameterization for η which corresponds to our baseline in Table 3.

5.3. Influence of Mobility Models. The attractiveness field is not our only focus in this study. The field influences the criminal agents in our model but without a mechanism for them to traverse the road network, then crime would be stagnant leaving us nothing to work

	Pearsons Correlation Coefficient				Mean Absolute Error			
	Macro				Macro			
	min	mean	median	max	min	mean	median	max
$\eta_{0.01}$	0.567	0.704	0.639	0.989	19.676	66.735	71.697	73.301
$\eta_{0.1}$	0.559	0.709	0.644	0.992	20.350	66.473	71.598	73.205
η_v	0.546	0.709	0.644	0.992	20.355	66.449	71.661	73.204
0.01	0.560	0.703	0.640	0.988	20.422	66.892	71.776	73.338
0.1	0.558	0.703	0.637	0.991	21.125	66.584	71.771	73.341
1	0.531	0.701	0.627	0.993	21.412	65.544	71.505	73.339
	Meso				Meso			
	min	mean	median	max	min	mean	median	max
$\eta_{0.01}$	0.258	0.427	0.352	0.902	1.623	2.642	2.669	3.040
$\eta_{0.1}$	0.259	0.409	0.336	0.891	1.633	2.653	2.673	3.050
η_v	0.257	0.406	0.333	0.886	1.652	2.656	2.678	3.041
0.01	0.263	0.432	0.356	0.909	1.590	2.642	2.667	3.045
0.1	0.260	0.419	0.340	0.911	1.590	2.649	2.672	3.071
1	0.252	0.394	0.315	0.846	1.853	2.648	2.690	3.008
	Micro				Micro			
	min	mean	median	max	min	mean	median	max
$\eta_{0.01}$	0.058	0.232	0.197	0.784	0.198	0.257	0.262	0.277
$\eta_{0.1}$	0.049	0.167	0.136	0.650	0.217	0.266	0.271	0.282
η_v	0.047	0.159	0.129	0.616	0.219	0.267	0.272	0.282
0.01	0.060	0.251	0.206	0.793	0.188	0.256	0.260	0.276
0.1	0.048	0.190	0.156	0.745	0.214	0.265	0.269	0.282
1	0.045	0.113	0.083	0.464	0.231	0.274	0.279	0.284

Table 6: Evaluation metrics across macro, meso, and micro scales for parameter η . Our results for η varied in our tests. At the meso and micro levels the Pearson and Mean Absolute Error values for $\eta = 0.01$ was generally the most consistent. Any differences in ranking were negligible. Generally speaking, smaller values of η performed better than higher values even accounting for variable diffusivity. As such, we conclude that $\eta = 0.1$ is the most consistent.

with. The mechanisms for movement are the mobility frameworks RW and TLF along with their hybrid models RW_TLF and TLF_RW. In this section, we discuss their performance, effects of their associated parameters, and some surprising consequences of each framework.

5.3.1. Performance of Mobility Frameworks. In Figure 8, we present our findings for the evaluation metrics of walk types at different scales. We see that the RW is the poorest performing mobility framework with low pearson values and high MAE values. In contrast, TLFs performed the best with high pearson values and low MAE values. In light of this, the differences in variances between the frameworks are extreme. For RW, we see minimal variance with errors differing by at most 10 in contrast to the variance in Levy Flights with differences as large as 100 when looking at our macro error. This variance is easy to understand as RW

offenders are limited in where they can travel to, in contrast to the large radius allowed in a Levy Flight. Despite this variance, we see that over the macro and meso scales, 50% of TLF and the hybrid models perform better than all of the results for RW. Thus, we conclude that there is a nontrivial advantage for our agents when they are given a large sample space to travel in and the ability to jump from node to node.

5.3.2. Effects of Truncation Limits. The variance of the TLF is unsurprisingly a consequence of the frameworks' greater search radius. The mechanism which creates this large radius is the truncation limit L . Thus, with large L , one would expect great variance. As expected, we find this result in our experiments as presented in Figure 9.

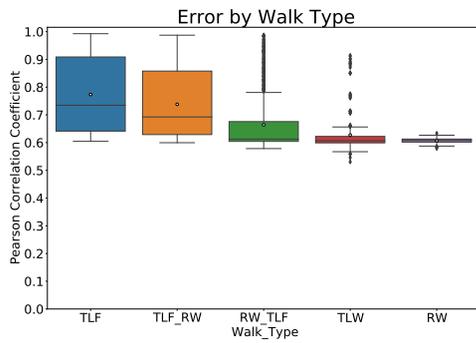
What is also evident in Figure 9 is how our error improves with large L . Our best error clearly corresponds with a truncation limit $L = 10000\text{m}$ which was our experimental maximum. We attribute this to our agents having greater knowledge of the hot spots in our model and therefore aggregating to these attractive areas. Extending the truncation limit to the furthest distances in the city would then create agents with perfect knowledge of hot spots in the city, suggesting improved errors. This would be unrealistic within the boundaries of a city but may be applicable if our experiments included Vancouver's neighbouring municipalities. This would be an extremely interesting avenue to explore in further research.

5.3.3. Journey to Crime. One final interesting area of study in our model is the computational and behavioural patterns of our agents in the different mobility frameworks. The question which arises from studying scavenging agents is how far the agent must travel before finding an optimal target. Furthermore, how many computational steps must the simulation go through for each offender for the optimal target to be found.

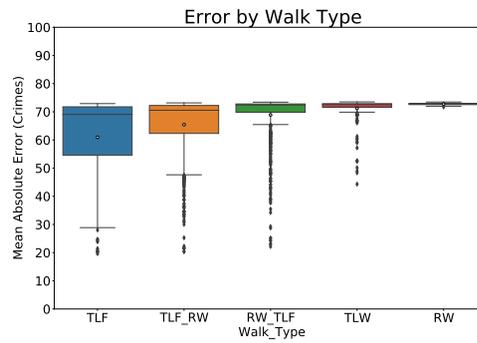
To investigate the differences in how the different mobility frameworks influence our scavenging behaviour, we plot in Figure 10 the top 100 simulations (in terms of MAE) of each framework against their Average Distance Travelled and the Average Number of Steps to Crime. In Figures 10(a), (c), and (e), the mobility framework which resulted in the least errors was the TLF whose agents travelled an average distance of 5940m. In general, agents with this framework found produce the lowest errors when their average distance travelled was between 4000-8000m. The hybrid models also perform similarly well with optimal ranges being between 2000-6000m.

These results show that the agents are efficient at finding sites to victimize. In Figures 10(b), (d), (f), it is evident that offenders with frameworks based on the TLF were able to find their desired target within 20 steps. This number of steps is also inflated due to our RW step minimums for the hybrid models. The frameworks based around the TLF had drastically fewer steps taken than RW based models while having smaller errors. This suggests is that the TLF is very efficient in finding optimal targets in an agent based simulation likely due to it ignoring intermediate steps. However, it is unrealistic to expect that a real criminal would take a single step to arrive at their target as some searching would be anticipated. Hence, the greater number of steps in the TLW model as opposed to the TLF.

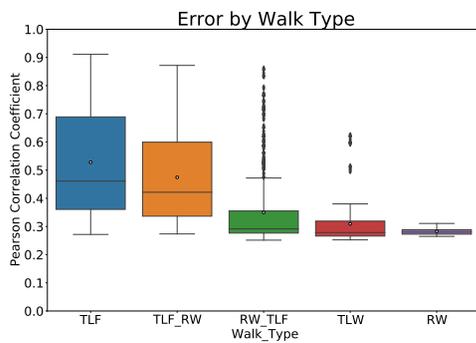
5.4. Summary of Most Consistent Parameters. To summarize our discussion, we present our results from combining our most consistent parameters. The parameters used are $\theta = 0.1$, $\omega = 0.01$, $\eta = 0.01$, Walk Type = TLF, $L = 10000$ as they performed well across each



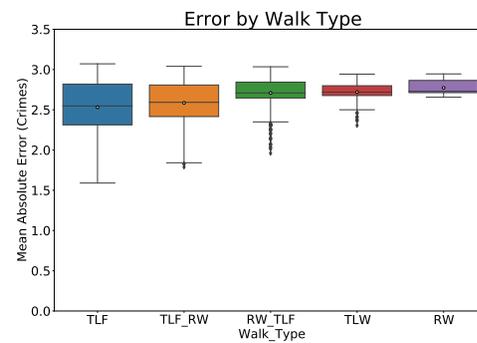
(a) Pearson's Coefficient at macro scale.



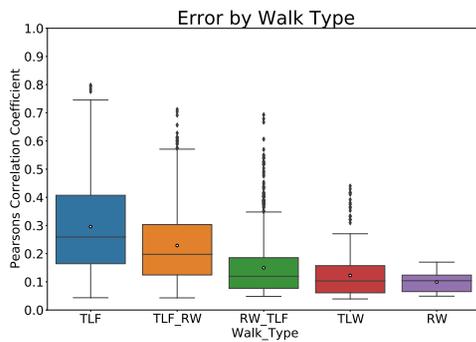
(b) Mean Absolute Error at macro scale.



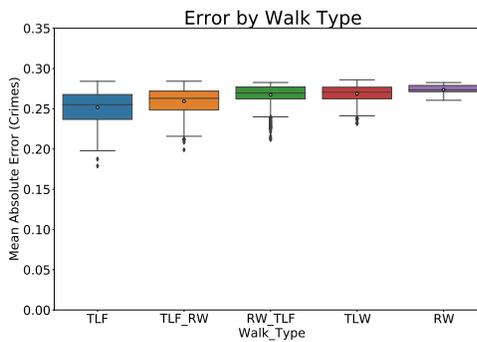
(c) Pearson's Coefficient at meso scale.



(d) Mean Absolute Error at meso scale.

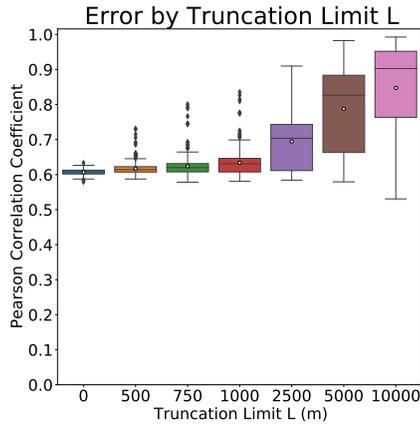


(e) Pearson's Coefficient at micro scale.

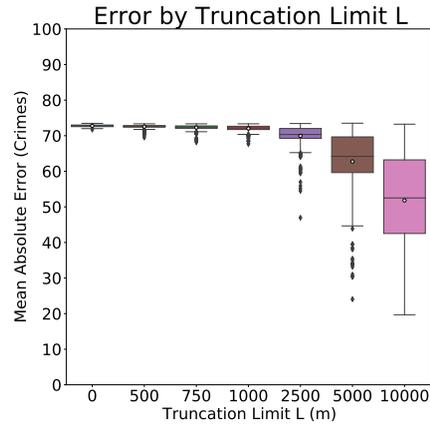


(f) Mean Absolute Error at micro scale.

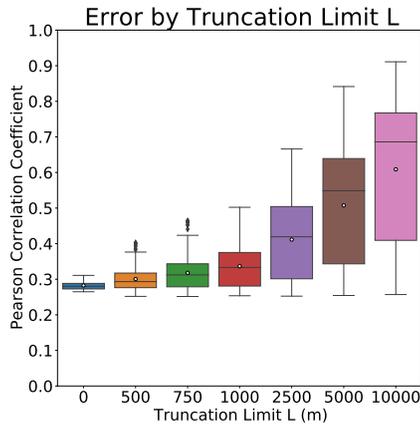
Figure 8: Evaluation metrics of walk types at different scales.



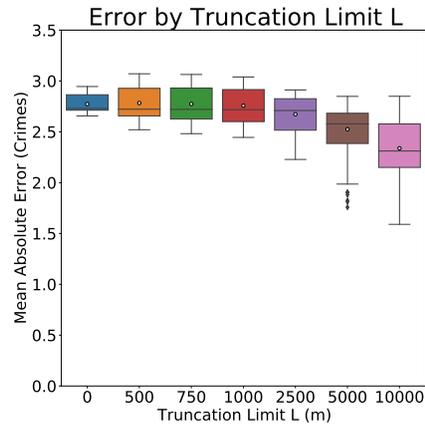
(a) Macro Pearson's Coefficient



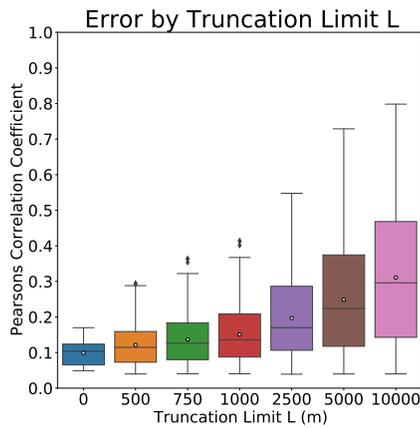
(b) Macro Mean Absolute Error



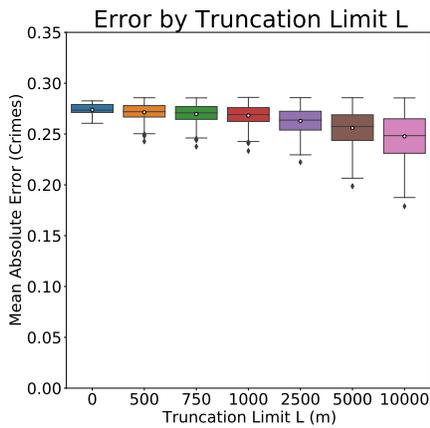
(c) Meso Pearson's Coefficient



(d) Meso Mean Absolute Error



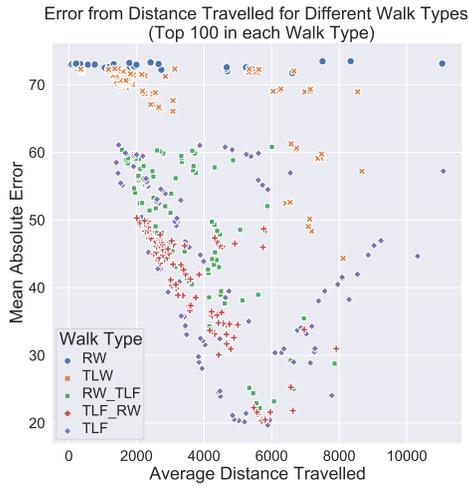
(e) Micro Pearson's Coefficient



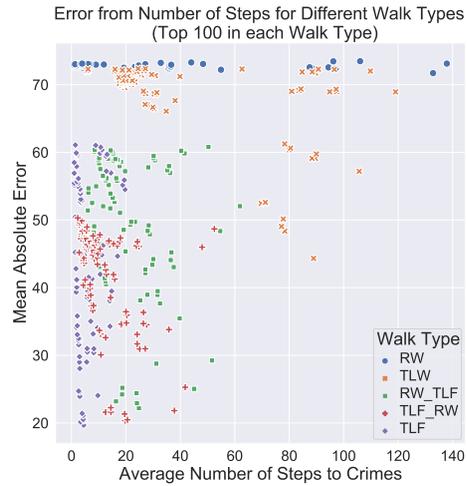
(f) Micro Mean Absolute Error

Figure 9: Evaluation metrics of truncation limits at different scales

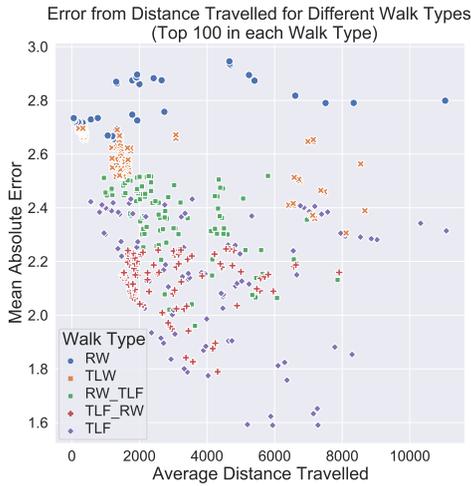
URBAN CITY: CRIME ATTRACTORS AND OFFENDER MOBILITY



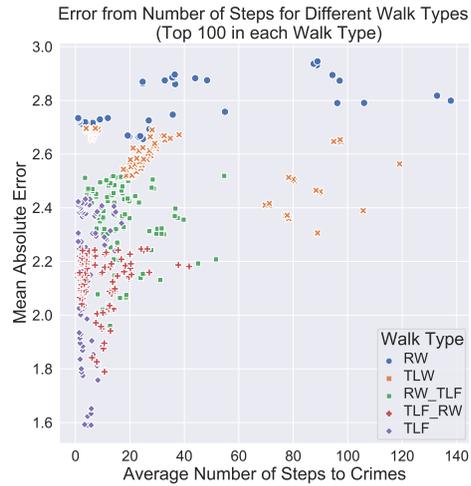
(a) Macro



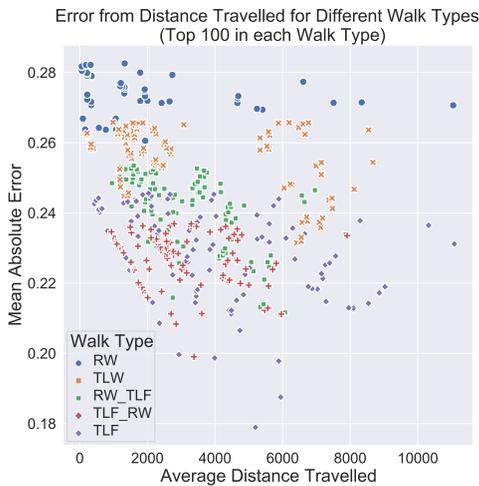
(b) Macro



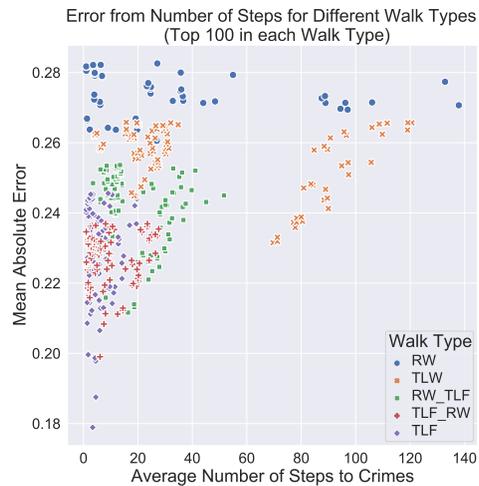
(c) Meso



(d) Meso



(e) Micro



(f) Micro

Figure 10: Journey to crime statistics for each walk type.

scope and metric. The values of the evaluation metrics for this parameter set are presented in Table 7. Figures 11 - 13 presents sample crime plots for the baseline, the mobility simulations, and the original real world data.

	Macro	Meso	Micro
Pearson Correlation Coefficient	0.988	0.909	0.732
Mean Absolute Error	20.422	1.590	0.188

Table 7: Evaluation Metrics for Most Consistent Parameters. The most consistent parameters are $\theta = 0.1, \omega = 0.01, \eta = 0.01$, Walk Type = TLF, $L = 10000$.

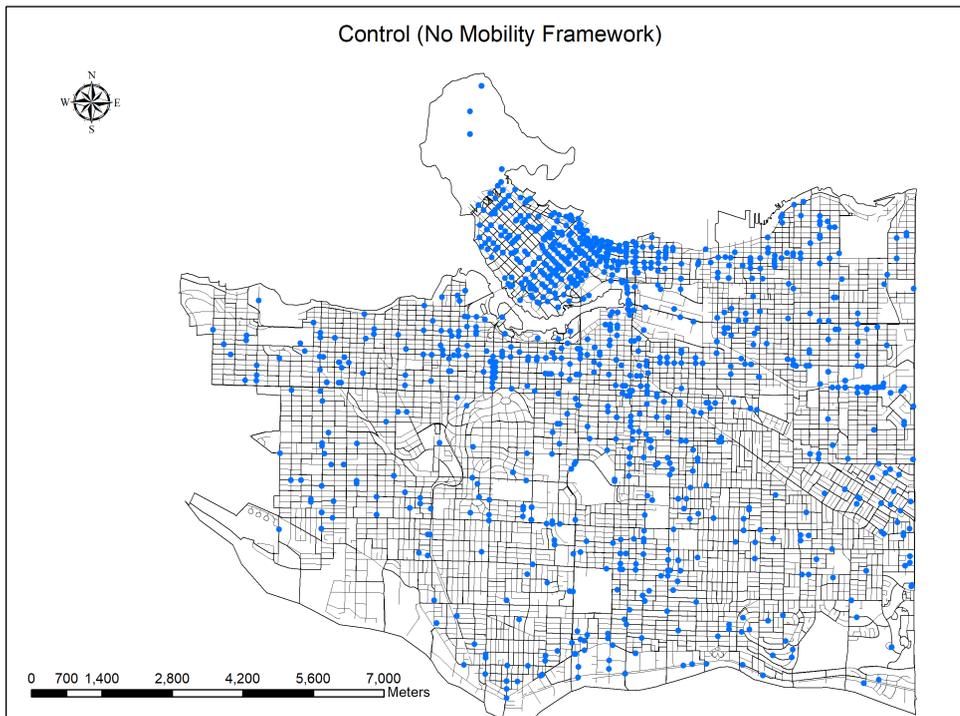


Figure 11: Plot of crime from control comparison. Generated using attractiveness field using $\theta = 0.1, \omega = 0.1$, and $\eta = 0.01$.

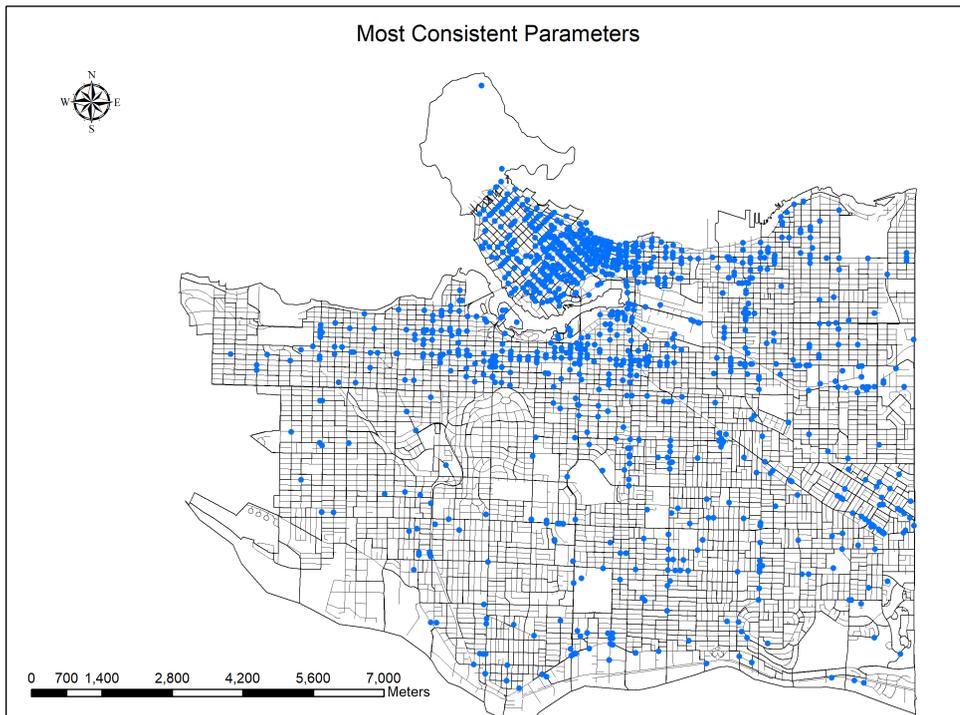


Figure 12: Simulation using $\theta = 0.1, \omega = 0.01, \eta = 0.01$, Walk Type = TLF, $L = 10000$

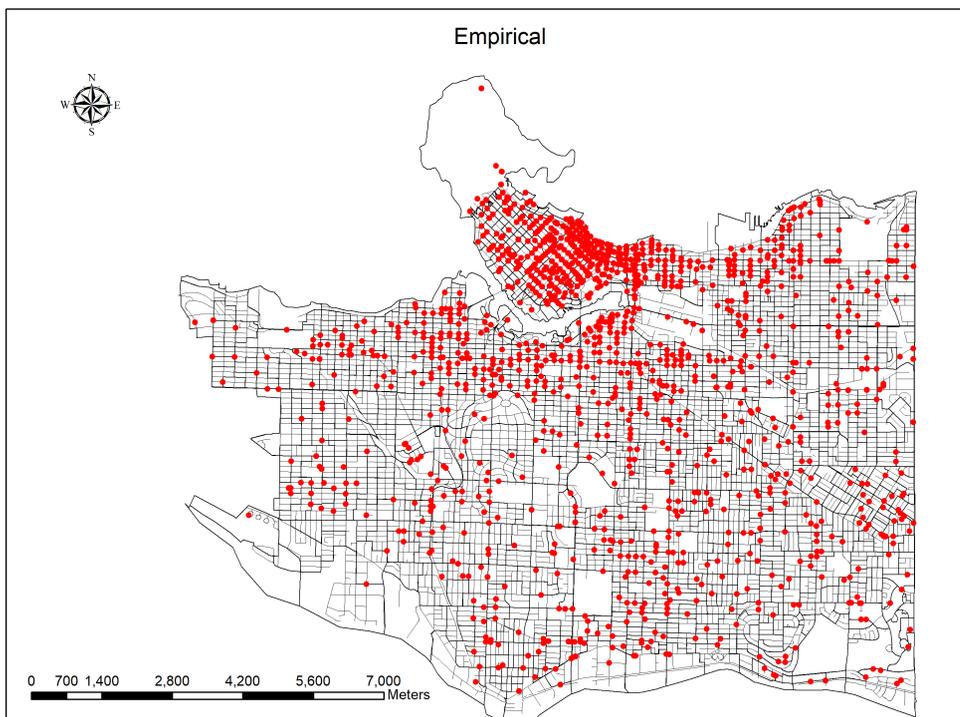


Figure 13: Empirical Crime Map for January 2020

The maps presented in Figures 11 - 13 show significant similarities. First, there is strong centrality towards Vancouver’s downtown core in the northern central peninsula of the city which is a spatial phenomenon of crime seen in other cities [16]. Moreover, we see strong crime patterns along Broadway, the long horizontal street along the coastline and below the peninsula, and in South Vancouver which are well known hot spots in the city. However, a key difference we observe in the maps is the tightness of crime clustering in the simulated map as opposed to the greater spread in the empirical map. There is one cluster which appears on Broadway in the baseline (Figure 11) that does not appear in the experimental (Figure 12) and empirical (Figure 13) maps. We attribute this primarily to our small diffusion parameter and believe that further investigations into this variable are needed in future research.

Despite these similarities, the simulation does not overwhelmingly out perform the baseline. Tables 3 and 7 show an even split for the evaluation metrics. For instance, the baseline is dominant at the macro level while the simulation is superior at the micro level. At the meso level, the baseline has greater correlation while the simulation has a lower Mean Absolute Error.

To better understand these differences, we present several summary statistics in Tables 8 and 9. These tables show the Absolute Errors of a single sample of our baseline ($\theta = 0.1, \omega = 0.1$, and $\eta = 0.01$) and our experimental simulation ($\theta = 0.1, \omega = 0.01, \eta = 0.01$, Walk Type = TLF, $L = 10000$). They include values for our min, mean, max, and standard deviation as well as values for the first, second, and third quartiles. Each of these values helps us analyze the spread of our Mean Absolute Error. Numbers which are close together among the quartiles indicate a tighter spread. This is further reflected if the standard deviation is small.

Baseline

	Min	Q1	Q2	Q3	Mean	Max	STD
Macro	1	4	6.5	15.25	9.91	31	7.94
Meso	0	0	1	2	1.79	50	2.94
Micro	0	0	0	0	0.21	39	0.80

Table 8: A sample statistical summary of the Absolute Error for each scope for our control results.

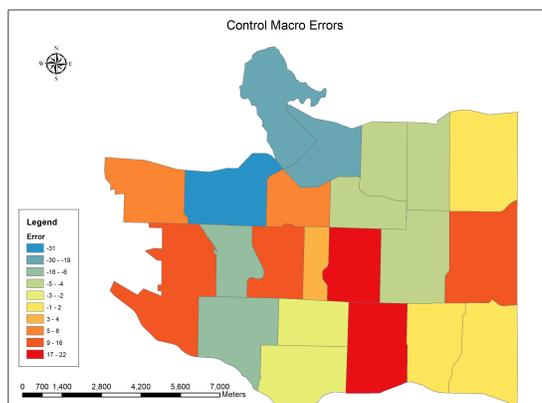
Experimental

	Min	Q1	Q2	Q3	Mean	Max	STD
Macro	1	7.5	14	31.75	20.36	75	19.13
Meso	0	0	1	2	1.68	43	3.43
Micro	0	0	0	0	0.19	28	0.75

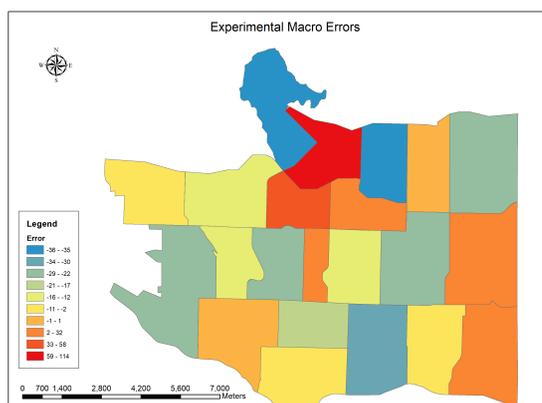
Table 9: A sample statistical summary of the Absolute Error for each scope for our experimental results.

It is evident that the spread of the baseline’s macro level errors is far tighter and lower than the experimental results. The maximum error of the baseline is twice as small as that of the experimental results and so is the standard deviation. Figure 14 visualizes the macro errors. The errors in Vancouver’s downtown area dominate in the experimental results. It is evident that in the experimental results, the model over predicts crime in this area due to its

high crime rates. Given that at the meso and micro levels, the maximum error of the baseline is greater than the experiment, we believe that this over prediction does obtain more accurate numbers at these finer levels.



(a) Macro



(b) Macro

Figure 14: A comparison of the Macro Absolute Error for the Control (14a) (no mobility framework) and Experimental (14b) (with mobility framework) simulated crimes. Both use the most consistent parameters $\theta = 0.1, \omega = 0.01, \eta = 0.01$. The experimental result uses Walk Type = TLF with $L = 10000$.

One can argue either way for the two methods which affirms Short's diffusion model [31]. We see the merits for both approaches but the simulation results with the mobility framework are more realistic, despite having higher macro errors. In reality criminals do not have perfect knowledge of a city's crime nor do they start their searches close to where they will commit their crimes. They start their journey at home and find their opportunities throughout the day [7, 11]. Hence the necessity of including the mobility framework for a realistic representation of crime. Moreover, this inclusion helps us gain insight into how appropriate Levy Flights are at replicating human mobility in general. Therefore, the simulation results are sufficient

in fitting real world crime and the applicable data.

6. Conclusion. In this article, we proposed a new framework for modelling an attractiveness backcloth within the topographical boundaries of a city's road network. We discussed the importance of framing the structure of the road network in the setting of network theory. Additionally, a predictive agent-based model of crime was introduced which was built upon the foundations set by the attractiveness framework of the crime concentrations. The results of this agent-based model were functionally authentic (see Figures 11 - 13) and in agreement with crime pattern theory.

Our experiments suggest the existence of a network-wide attractiveness backcloth based on common knowledge of past crime events. We found that this backcloth can be described by increased local attractiveness due to increase vulnerability, local decay due to time, and spread due to common knowledge where θ , ω , and η can describe these factors mathematically. In our experiments, we found optimal values of these variables to be $\theta = 0.1$, $\omega = 0.01$, and $\eta = 0.01$.

On this attractiveness backcloth we implemented an agent-based simulation model of crime to better understand how criminal agents move on the network and to predict crime. This agent-based model simulated criminal movement using biased mobility frameworks towards areas of high attractiveness in the backcloth. In the experiments, we found that the most efficient framework was the Truncated Levy Flight. The Truncated Levy Flight enables our offenders to take large jumps to areas of high attractiveness while being restricted by a Truncation Limit L and decreased probability with greater distances. We found that increasing this Truncation Limit improved our results with our best results occurring for $L = 10000\text{m}$. However, this attractiveness backcloth performed well independent of an agent-based simulation model affirming the work of Short et. al. [31].

To test our results, we also propose a method to determine model accuracy based on macro, meso, and micro scales. These scales were based on municipal and federal defined neighbourhoods for the City of Vancouver. To generalize these scales we demonstrated that one could use the K-means clustering algorithm so that this type of analysis could be used on any city street network.

The results show the importance of better understanding offender mobility within an urban space and the power of crime attractors. The model worked well using past crimes at a site as a proxy for crime attractiveness and diffusion of attractiveness on a street network. This shows much promise in future research that looks at the pushes and pulls of crime attractiveness or crime attractors in urban space.

7. Future Research. The simulations and results presented in this model are unique to the City of Vancouver. One interesting exploration would be to see how our techniques and model could be applied to other cities as we have developed it to be generalizable. Additionally, in reality offenders and offender groups have different levels of motivation, factors that influence their decision processes and crime actions that are triggered in different situational contexts [20, 42, 43]. Dividing agents into social networks as well as distinguishing individuals with varying types of motivation and degrees of influence, would be an interesting direction for future research.

Currently, offender generation and movement are limited in the current version of our

model. We speculate that introducing more realistic offender generation and movement information would improve results in our model. Data such as offender home density, transit networks, and road speeds or travelling times are all areas to be explored in future research. Introducing modes of transportation (foot, bicycle, car, or mass transit) available to each individual would especially be interesting. We suspect that various means of transportation will impact results. For example, rapid transit lines allow for including Biased Random Waypoint Movement (BRWM) between different access points within our model. BRWM is a means of movement would add value to the model.

Future research on optimizations to the existing model are also lucrative. One of the most important aspects of this research is the choice of network that we use for this model. The nodes in the network are not equidistant and on average they are about 75 meter apart. In a network where nodes are less far apart, offenders have more opportunities to commit crimes. However, there is a trade-off between the two approaches. We have been able to compute the full Street Network Matrix for our network, but this is an extremely expensive computational task. It is not recommended to make these computations for larger networks. The benefit of a smaller network is being able to run more simulations in a timely manner. This allows us to experiment more parameters, and ultimately understand each parameter better. Having said this, we believe it might be possible to optimize this process for larger networks and the effects of the size of the network could be a separate topic for research.

Additionally, in previous researches on the topic, a constant value has been used for the whole network. This means, for instance, the same coefficient of diffusivity has been used in the hot spots as well as less active areas. In our research we experimented spatially varying coefficient for diffusivity. We believe this is an idea that could be further studied; not only for one parameter, but also for other parameters as well. For example, one could determine a clustering based on past crimes instead of spatially. Similarly, one could study the effects of temporal varying coefficients in addition to spatial coefficients. One can also consider a more complex model for agent behaviour. This can be done through adding agents with different histories of offending such as high repeat offenders and occasional offenders.

This article presents a new simulation model for predicting crime patterns in an urban setting using the street network as the base for limiting where offenders can move about the city. The actual street network, together with a truncated Levy Flight, attractiveness of potential targets and diffusion of attractiveness helped produce a rewarding simulation model. The current model, although simple, was very strong given that it is a starting point. We expect that future research will improve the model and increase the power of network analysis in the study of crime patterns.

Acknowledgments. We would like to acknowledge the support of our research supervisors Dr. Razvan Fetecau and Dr. Patricia J. Brantingham. Additionally, we would like to thank the Simon Fraser University Undergraduate Student Research Awards (USRA), the Natural Sciences and Engineering Research Council of Canada (NSERC), and SSHRC Insight Development Grant through Thompson Rivers University for funding this project.

REFERENCES

- [1] M. A. ANDRESEN AND P. J. BRANTINGHAM, *The place of environmental criminology within criminological thought*, in *Classics in Environmental Criminology*, Simon Fraser University Publications, 2010, pp. 5–28.
- [2] W. BERNASCO, *Them again?*, *European Journal of Criminology*, 5 (2008), pp. 411–431.
- [3] W. BERNASCO AND F. LUYKX, *Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighborhoods*, *Criminology*, 41 (2003), pp. 981–1002.
- [4] G. BOEING, *Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks*, *Computers, Environment and Urban Systems*, 65 (2017), p. 126–139, <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>.
- [5] R. BOIVIN AND S. N. DE MELO, *The concentration of crime at place in montreal and toronto*, *Canadian Journal of Criminology and Criminal Justice*, 61 (2019), pp. 46–65.
- [6] P. J. BRANTINGHAM AND P. L. BRANTINGHAM, *Patterns in crime*, Macmillan New York, 1984.
- [7] P. J. BRANTINGHAM, P. L. BRANTINGHAM, AND M. A. ANDRESEN, *The geometry of crime and crime pattern theory*, *Environmental criminology and crime analysis*, 2 (2017).
- [8] P. J. BRANTINGHAM, P. L. BRANTINGHAM, J. SONG, AND V. SPICER, *Crime hot spots, crime corridors and the journey to crime: An expanded theoretical model of the generation of crime concentrations*, in *Geographies of behavioural health, crime, and disorder*, Springer, 2020, pp. 61–86.
- [9] P. L. BRANTINGHAM, *Crime Pattern Theory*, SAGE Publications, Inc., 2010.
- [10] P. L. BRANTINGHAM AND P. J. BRANTINGHAM, *Notes on the geometry of crime*, in *Classics in Environmental Criminology*, Simon Fraser University Publications, 1981, pp. 231–255.
- [11] P. L. BRANTINGHAM AND P. J. BRANTINGHAM, *Environment, routine and situation: Toward a pattern theory of crime*, *Advances in criminological theory*, 5 (1993), pp. 259–294.
- [12] P. L. BRANTINGHAM AND P. J. BRANTINGHAM, *Criminality of place*, *European Journal on Criminal Policy and Research*, 3 (1995), pp. 5–26.
- [13] G. BRASSINGTON, *Mean absolute error and root mean square error: which is the better metric for assessing model performance?*, in *EGU General Assembly Conference Abstracts*, 2017, p. 3574.
- [14] T. CHAI AND R. R. DRAXLER, *Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature*, *Geoscientific model development*, 7 (2014), pp. 1247–1250.
- [15] S. CHATURAPRUEK, J. BRESLAU, D. YAZDI, T. KOLOKOLNIKOV, AND S. G. MCCALLA, *Crime modeling with lévy flights*, *SIAM Journal on Applied Mathematics*, 73 (2013), pp. 1703–1720.
- [16] P. CRUCITTI, V. LATORA, AND S. PORTA, *Centrality in networks of urban streets*, *Chaos: an interdisciplinary journal of nonlinear science*, 16 (2006), p. 015113.
- [17] A. S. CURMAN, M. A. ANDRESEN, AND P. J. BRANTINGHAM, *Crime and place: A longitudinal examination of street segment patterns in vancouver, bc*, *Journal of Quantitative Criminology*, 31 (2015), pp. 127–147.
- [18] R. FRANK, V. DABBAGHIAN, A. REID, S. SINGH, J. CINNAMON, AND P. BRANTINGHAM, *Power of criminal attractors: Modeling the pull of activity nodes*, *Journal of Artificial Societies and Social Simulation*, 14 (2011), p. 6.
- [19] A.-M. GUERRY, *Essai sur la statistique morale de la France*, Crochard, 1833.
- [20] D.-H. HAAR AND P.-O. H. WIKSTRÖM, *Crime propensity, criminogenic exposure and violent scenario responses: Testing situational action theory in regression and rasch models*, *European Journal of Applied Mathematics*, 21 (2010), pp. 307–323.
- [21] T. KANUNGO, D. M. MOUNT, N. S. NETANYAHU, C. D. PIATKO, R. SILVERMAN, AND A. Y. WU, *An efficient k-means clustering algorithm: Analysis and implementation*, *IEEE transactions on pattern analysis and machine intelligence*, 24 (2002), pp. 881–892.
- [22] J. B. KINNEY, P. L. BRANTINGHAM, K. WUSCHKE, M. G. KIRK, AND P. J. BRANTINGHAM, *Crime attractors, generators and detractors: Land use and urban crime opportunities*, *Built environment*, 34 (2008), pp. 62–74.
- [23] A. LOUKAITOU-SIDERIS, R. LIGGETT, AND H. ISEKI, *The geography of transit crime: Documentation and evaluation of crime incidence on and around the green line stations in los angeles*, *Journal of Planning Education and Research*, 22 (2002), pp. 135–151.
- [24] A. E. NIVETTE, R. ZAHNOW, R. AGUILAR, A. AHVEN, S. AMRAM, B. ARIEL, M. J. A. BURBANO,

- R. ASTOLFI, D. BAIER, H.-M. BARK, AND ET AL., *A global analysis of the impact of covid-19 stay-at-home restrictions on crime*, Jun 2021, <https://www.nature.com/articles/s41562-021-01139-z>.
- [25] C. PAN, B. LI, C. WANG, Y. ZHANG, N. GELDNER, L. WANG, AND A. L. BERTOZZI, *Crime modeling with truncated lévy flights for residential burglary models*, *Mathematical Models and Methods in Applied Sciences*, 28 (2018), pp. 1857–1880.
- [26] L. A. QUETELET, *A treatise on man and the development of his faculties (a facsimile reproduction of the english translation of 1842 with an introduction by solomon diamond)*, (1842).
- [27] A. P. RIASCOS AND J. L. MATEOS, *Fractional dynamics on networks: Emergence of anomalous diffusion and lévy flights*, *Physical Review E*, 90 (2014), p. 032809.
- [28] N. SCAFETTA, *Understanding the complexity of the levy-walk nature of human mobility with a multi-scale cost/benefit model*, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21 (2011), p. 043106.
- [29] C. R. SHAW, *Delinquency areas.*, (1929).
- [30] M. B. SHORT, P. J. BRANTINGHAM, A. L. BERTOZZI, AND G. E. TITA, *Dissipation and displacement of hotspots in reaction-diffusion models of crime*, *Proceedings of the National Academy of Sciences*, 107 (2010), pp. 3961–3965.
- [31] 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 (2008), pp. 1249–1267.
- [32] G. SONG, W. BERNASCO, L. LIU, L. XIAO, S. ZHOU, AND W. LIAO, *Crime feeds on legal activities: Daily mobility flows help to explain thieves' target location choices*, *Journal of Quantitative Criminology*, 35 (2019), pp. 831–854.
- [33] V. SPICER, J. SONG, P. BRANTINGHAM, A. PARK, AND M. A. ANDRESEN, *Street profile analysis: A new method for mapping crime on major roadways*, *Applied Geography*, 69 (2016), pp. 65–74.
- [34] M. A. TAYEBI, M. ESTER, U. GLÄSSER, AND P. L. BRANTINGHAM, *Crimetracer: Activity space based crime location prediction*, in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, IEEE, 2014, pp. 472–480.
- [35] M. A. TAYEBI, U. GLÄSSER, M. ESTER, AND P. L. BRANTINGHAM, *Personalized crime location prediction*, *European Journal of Applied Mathematics*, 27 (2016), pp. 422–450.
- [36] R. TAYLOR, *Interpretation of the correlation coefficient: a basic review*, *Journal of diagnostic medical sonography*, 6 (1990), pp. 35–39.
- [37] M. S. TILLYER AND R. J. WALTER, *Busy businesses and busy contexts: The distribution and sources of crime at commercial properties*, *Journal of Research in Crime and Delinquency*, 56 (2019), pp. 816–850.
- [38] VANCOUVER POLICE DEPARTMENT, *Crime data download*, 2020, <https://geodash.vpd.ca/opendata/>.
- [39] D. WEISBURD, *The law of crime concentration and the criminology of place*, *Criminology*, 53 (2015), pp. 133–157.
- [40] D. WEISBURD, W. BERNASCO, AND G. BRUINSMA, *Putting crime in its place*, Springer, 2008.
- [41] D. WEISBURD, E. R. GROFF, AND S.-M. YANG, *The criminology of place: Street segments and our understanding of the crime problem*, Oxford University Press, 2012.
- [42] P.-O. H. WIKSTRÖM, *Why crime happens: A situational action theory*, in *Analytical Sociology*, John Wiley & Sons, Ltd, Mar. 2014, pp. 71–94.
- [43] P.-O. H. WIKSTRÖM AND K. TREIBER, *Social disadvantage and crime*, *American Behavioral Scientist*, 60 (2016), pp. 1232–1259.
- [44] C. J. WILLMOTT AND K. MATSUURA, *Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance*, *Climate research*, 30 (2005), pp. 79–82.
- [45] H. YU, L. LIU, B. YANG, AND M. LAN, *Crime prediction with historical crime and movement data of potential offenders using a spatio-temporal cokriging method*, *ISPRS International Journal of Geo-Information*, 9 (2020), p. 732.
- [46] C. ZHANG, A. X. JIANG, M. B. SHORT, P. J. BRANTINGHAM, AND M. TAMBE, *Modeling crime diffusion and crime suppression on transportation networks: An initial report.*, in *AAAI Fall Symposia*, 2013.