

Modeling Crime diffusion and crime suppression on transportation networks: An initial report

Chao Zhang¹, Albert Xin Jiang¹, Martin B. Short², P. Jeffrey Brantingham², Milind Tambe¹

¹University of Southern California, Los Angeles, CA 90089
{zhan661, jiangx, tambe} @usc.edu

²University of California, Los Angeles, CA 90095
mbshort@math.ucla.edu, branting@ucla.edu

Abstract

In urban transportation networks, crime diffuses as criminals travel through the networks and look for illicit opportunities. It is important to first model this diffusion in order to recommend actions or patrol policies to control the diffusion of such crime. Previously, game theory has been used for such patrol policy recommendations, but these applications of game theory for security have not modeled the diffusion of crime that comes about due to criminals seeking opportunities; instead the focus has been on highly strategic adversaries that plan attacks in advance. To overcome this limitation of previous work, this paper provides the following key contributions. First, we provide a model of crime diffusion based on a quantal biased random movement (QBRM) of criminals opportunistically and repeatedly seeking targets. Within this model, criminals react to real-time information, rather than strategically planning their attack in advance. Second, we provide a game-theoretic approach to generate randomized patrol policies for controlling such diffusion.

Introduction

Crime in transportation networks is a threat to passengers. Given the structure of these networks, crime diffuses as criminals traveling by public transportation seize opportunities to commit crimes. Unlike strategic adversaries who may carefully plan to exploit security weaknesses and attack targets, criminals may opportunistically react to real-time information, which means that crime diffuses dynamically. Indeed, recent research in criminology shows that crimes are often crimes of opportunity and how offenders move and mix with their potential targets or victims is a key determinant of the structure of any crime opportunity (Brantingham and Tita 2008; Felson, Clarke, and Britain 1998).

Indeed, transportation networks play an important role in driving local crime patterns and in the diffusion of crime (Block and Block 2000; Matthews et al. 2010). Individual transit hubs that are strong crime generators export that propensity to other locations on the transit network. Such diffusive potential may be particularly strong since

a substantial portion of criminals use public transportation as their primary means of transportation (Loukaitou-Sideris, Liggett, and Iseki 2002). Transportation networks may themselves be at unique risk of crime because of the way in which they concentrate large numbers of people in time and space (Matthews et al. 2010; Taylor and Harrell 1996; Brantingham and Brantingham 1995). Within a transportation network, not all locations are at equal risk. Certain transit stations, and certain transit vehicles, may have design features that promote crime, be it poor lighting and lack of natural surveillance (Loukaitou-Sideris 1999) or environmental cues such as poor maintenance and graffiti that suggest that the facility is not well protected (Keizer, Lindenberg, and Steg 2008). Some transit locations are therefore more likely to attract offenders than others.

We take a metro rail network as a concrete example. In such a network, crimes such as thefts and snatches usually occur at nodes, such as stations or junctions where it is easy for criminals to escape. These potential crime spots are connected by trains with a fixed timetable. Crime at one node can diffuse to a far-away node without affecting its neighbors, which corresponds to the situation that criminals take the train directly to the far-away node without getting off halfway. This diffusion may be stochastic, with higher probabilities of crime at more attractive stations.

Deploying police to patrol in such transportation networks is a way to suppress crime and control its diffusion. In our example metro rail network, the police patrols throughout all stations by trains. Previous work applying game theory in a metro network has successfully generated randomized patrol schedules for police (Yin et al. 2012; Jiang et al. 2012). These works deal with highly strategic attackers who conduct full surveillance and plan their illegal acts in advance; they assume attackers cannot adjust these plans given real-time information. Another difference of that work from ours is that attackers have fixed routes. As a result, the crime does not diffuse.

There are two key contributions in this paper. The first contribution is a new model of crime diffusion. In this model, criminals visit targets based on a quantal biased random movement (QBRM), which has been used to model criminal motion previously (Short et al. 2008), instead of executing fixed routes. As a result, crime diffuses freely and the pattern of such diffusion is stochastic. In addition, rather

than planning their attack in advance, criminals opportunistically react to real-time information in the network. For example, the real-time observations criminals make on the location of police affects the QBRM model output and their probability of committing a crime.

The second contribution is a game-theoretic approach to generate randomized patrol schedules. We model the interactions between criminals and the police as a Stackelberg game, with the police acting as the leader and criminals as followers. The police must commit to her patrol strategy first and then criminals will choose targets to attack given surveillance of police’s deployment. However there are two differences with previous work in Stackelberg Security games, which have been alluded to earlier. First, criminals react to real-time information in our model as mentioned earlier, which is different from previous work. Second, after one attack, criminals can still stay in the network and find another target to attack using our QBRM model, which is modeled as crime diffusion. Our objective is to find a randomized patrol strategy for the police that optimizes her expected utility against crime diffusion. We formulate the problem as a nonlinear optimization problem on a Markov chain model. Initial numeric experiments show that police strategies computed by an off-the-shelf nonlinear solver on our optimization formulation significantly outperform the uniformly random patrol strategy.

Related Work

There has been research on a wide range of topics related to controlling diffusion in networks. One line of work considers game-theoretic models of controlling contagion in networks. These are games between defenders and attackers where the attacker attempts to maximize its diffusion influence over the network while the defender tries to minimize this influence. Algorithms have been proposed to approximately solve such games under different models of diffusion, including (Tsai, Nguyen, and Tambe 2012) for the Independent Cascade Model (ICM) and (Nguyen et al. 2012) for the Linear Threshold Model (LTM). In these contagion games, the two players can only select a number of initial seed nodes and their influence diffuses automatically. Such models are thus not applicable to model the diffusive behavior of opportunistic criminals.

Another line of research uses recent advances in criminology on opportunistic criminal behavior to describe crime diffusion in networks. (Short et al. 2008) applied a biased random walk model for house burglary, and (Zipkin, Short, and Bertozzi 2013) analyzed the effect of the police on controlling the crime diffusion in the house burglary domain. In their works, criminals have no knowledge of the overall strategy of the police, and their behavior is only affected by their observation of the current police allocation in their immediate neighborhood. Also in (Zipkin, Short, and Bertozzi 2013), police behave in a similarly reactionary way, allocating their resources in an instantaneously optimal way in response to the current crime risk distribution rather than optimizing over the time horizon and within a transportation network.

The motions of both criminals and police in (Zipkin, Short, and Bertozzi 2013) also vary significantly from those in the current work. Each instance of a criminal’s motion in (Zipkin, Short, and Bertozzi 2013) may only be between adjacent locations, after which the nearby police allocation is observed anew and another movement can be made, leading to highly localized diffusion of criminals. In contrast, criminals in the current work may make “large” directed movements over the transportation network between distant locations, as they see fit, before updating their beliefs and moving again, leading to much less localized crime. Furthermore, in (Zipkin, Short, and Bertozzi 2013) there is no notion of the “movement” of police - rather, the distribution of police is chosen to be instantaneously optimal, with no regard for the mechanics of exactly how the allocation may transform from one timestep to the next.

Game theoretic approaches have been successfully applied to security domains for generating randomized patrol strategies against strategic adversaries, e.g., (Tsai et al. 2009) generated schedules for the Federal Air Marshals to protect flights; (Shieh et al. 2012) generated schedules for the US Coast Guard to protect ports; and (Yin et al. 2012; Jiang et al. 2012) generated schedules for Los Angeles Sheriff Department to conduct fare checking on the LA Metro network. We have discussed our differences with that literature in the Introduction Section.

Our approach combines and generalizes the randomized patrolling model of previous security applications and the criminology-based random-walk diffusion model of (Short et al. 2008): now police can move inside the network in a randomized fashion, and the criminals are opportunistic and can diffuse throughout the network.

Problem Settings

In this section, we describe the problem setting of crime diffusion in a metro rail network. By convention, the police are referenced as “she” and criminal as “he”.

The metro rail network

The metro rail network consists of multiple stations along a straight line and trains traveling in both directions through this line. For simplicity, we assume that the distances between any two neighboring stations are the same. The trains follow a fixed timetable. We divide time into time steps of equal duration, so that it takes 1 time step for trains to go from one station to its neighbors. We furthermore assume that the trains can only arrive at or leave stations at the beginning of these time steps. Denote by N the number of stations. There can be at most $2(N - 1)$ trains running simultaneously. Stations and trains running between stations are collectively referred to as *places*. We denote the train from station i to its neighboring station j (i.e., $|i - j| = 1$) as $i \rightarrow j$. Figure 1 shows such a metro rail network structure with three stations from left to right.

There are two phases in one time step, which are shown in Figure 2. The first phase, $P1$, is called the *decision phase*, which is instantaneous. In this phase, trains are at stations and passengers can make decisions. There are at most three

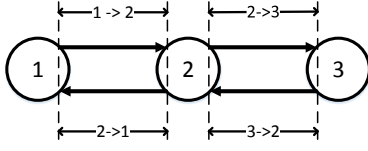


Figure 1: The metro network

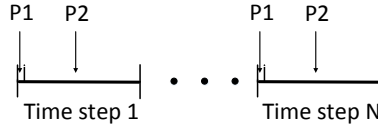


Figure 2: Two phases in time steps

Place	1	1->2	2->3	3	3->2	2
Time step	1	2	3	4	5	6
Strike	1			2		3

Figure 3: Example of strikes

decisions if the passenger is at a station: *Go left* means he will get on the train heading to the left; *Go right* means he will get on the train heading to the right; *Stay* means he stays at the current station. The decisions that the passenger can make if he is on a train are similar. If his train is heading to the left, *Go left* means he will stay on the train; *Stay* means he will get off the train and stay at the station; *Go right* means he will get off his train and immediately get on the train heading in the opposite direction. The second phase, *P2*, is called the *action phase*. In this phase, trains are in motion from one station to another.

The police

There are two kinds of passengers in this metro rail network. The first kind are the police, who patrol throughout the network. For the initial model in this paper we assume that there is a single police unit. This police unit conducts randomized patrol using a *Markov strategy* π , which specifies for each place a probability distribution over the available actions at that place. Specifically, for each place m denote by l_m, r_m, s_m the probabilities of actions *Go left*, *Go right* and *Stay*, respectively; note that $l_m + r_m + s_m = 1$ for each place.

Example 1: A simple scenario with 3 stations (1, 2, 3) and 4 trains (1 \rightarrow 2, 2 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 2) is given in Figure 1. A sample Markov strategy π is:

$$\begin{aligned}
 s_1 &= 0.5, r_1 = 0.5; \\
 l_2 &= 0.3, s_2 = 0.4, r_3 = 0.3; \\
 l_3 &= 0.2, s_3 = 0.8; \\
 l_{1 \rightarrow 2} &= 0.1, s_{1 \rightarrow 2} = 0.7, r_{1 \rightarrow 2} = 0.2; \\
 s_{2 \rightarrow 1} &= 0.8, r_{2 \rightarrow 1} = 0.2; \\
 l_{2 \rightarrow 3} &= 0.3, s_{2 \rightarrow 3} = 0.7; \\
 l_{3 \rightarrow 2} &= 0.4, s_{3 \rightarrow 2} = 0.1, r_{3 \rightarrow 2} = 0.5;
 \end{aligned}$$

$s_1 = 0.5, r_1 = 0.5$ means that if the police is at station 1, she will have 0.5 probability to choose action *Stay* and 0.5 probability to choose action *Go right* at next decision phase; $l_{1 \rightarrow 2} = 0.1, s_{1 \rightarrow 2} = 0.7, r_{1 \rightarrow 2} = 0.2$ means that if the police is at the train from station 1 to station 2, she will have probability 0.1 to choose action *Go left*, probability 0.7 to choose action *Stay*, and probability 0.2 to choose action *Go right* at next decision phase.

Criminals

The second kind of passengers are criminals. Criminals travel throughout the network with a station-based plan. In the decision phase, a criminal currently located at a station will choose the next target station as described below; if he is on a train, he can only execute the plan made earlier at his previous station, and does not have the option to change course.

Criminals can commit crimes during their time in the metro network. We ignore the possibility of crimes during decision phases because these phases are considered instantaneous. In action phases, crimes can only be committed by criminals located at stations, where it is easy for them to escape.

We call the subset of time steps in which a given criminal is at a station during the action phase *Strikes*. Figure 3 is a toy example that shows the relationship between the time steps and strikes. For criminals with different station sequences, their strikes are different. Also, we assume that at each strike, and at the end of the action phase, the criminal exits the metro system with a constant probability α .

Criminals are opportunistic and may not conduct long time surveillance of the police. However, for simplicity, we assume in this work that they have perfect knowledge of the police unit's Markov strategy (i.e., the probabilities of police actions at each place) and can compute the stationary distribution of the police unit over the places that is induced by the Markov strategy. Furthermore, once arrived at a station, a criminal will observe whether or not the police unit is present at that station. Based on this knowledge, the criminal will decide whether to commit crime at that station during the action phase, and then where to go at the decision phase of the next time step. Each criminal is *opportunistic* in the sense that his decision to commit crime is based on whether the police unit is present at his current station, along with the availability of crime opportunities at the station. When choosing the next location for a Strike, the criminal tends to move towards stations that he believes have a higher expected value. His belief about the expected values of a station will in turn depend on his belief about the presence of police at that station upon his arrival, the general availability of crime opportunities at the station, as well as the cost of time spent in traveling to that station. We consider *boundedly rational* criminals, who do not always choose the station with the highest value, but instead do random move-

N	The total number of stations
π	Police's Markov strategy
$E(i)$	The expected value of station i for criminals
$V_p(k)$	The expected value for police at strike number k
T_p	Transition matrix for police
p_c	Criminal's probability vector of committing a crime for each state
\vec{c}^t	Police coverage vector
\vec{c}^s	Police stationary coverage vector
p	Criminal's probability distribution of target station for next strike.
$\vec{c}_{b,t}$	Criminal's belief coverage after t steps from current time step
λ	The rationality of criminal
α	The exit rate of criminal
T_s	Transition matrix for the Markov chain model
\vec{X}_k	The state distribution of Markov chain at strike k
R_p	Police's revenue vector for each state
Obj	Police's objective

Table 1: Notation

ments biased toward stations of high expected utility.

Diffusion Model

In this section, we propose a model for crime diffusion in a metro network, which describes the criminal's behavior in the presence of police patrols. The interactions between the police and the criminal are modeled as a Markov chain, where each state is a combination of the criminal's station and the police's place. We then formulate the problem of computing a maximum-revenue Markov patrol strategy for the police as a non-linear optimization problem. The notation used in this paper is described in Table 1.

Police Patrol Model

Recall that the police's Markov strategy is a mapping from the police's places to distributions over actions. There are in total $3N - 2$ places, including N stations and $2N - 2$ trains. Given the Markov strategy, the police's movement is a Markov chain over the places. We call the transition matrix of this Markov chain the *police transition matrix* T_p . T_p is a $(3N - 2) \times (3N - 2)$ matrix whose entry at column i , row j specifies the probability of the police going from place i to place j in one time step. We choose a basis such that, in this matrix, line $3i - 2$ ($i \in 1, \dots, N$) represents the station i ; line $3i - 1$ represents the train from station i to station $i + 1$; line $3i$ represents the train from station $i + 1$ to i .

Example 2: Given the same scenario in Example 1, where there are 3 stations in total, we have the following police transition matrix.

$$T_p = \begin{pmatrix} s_1 & 0 & s_{2 \rightarrow 1} & 0 & 0 & 0 & 0 \\ r_1 & 0 & r_{2 \rightarrow 1} & 0 & 0 & 0 & 0 \\ 0 & l_{1 \rightarrow 2} & 0 & l_2 & 0 & l_{3 \rightarrow 2} & 0 \\ 0 & s_{1 \rightarrow 2} & 0 & s_2 & 0 & s_{3 \rightarrow 2} & 0 \\ 0 & r_{1 \rightarrow 2} & 0 & r_2 & 0 & r_{3 \rightarrow 2} & 0 \\ 0 & 0 & 0 & 0 & l_{2 \rightarrow 3} & 0 & l_3 \\ 0 & 0 & 0 & 0 & s_{2 \rightarrow 3} & 0 & s_3 \end{pmatrix} \quad (1)$$

Using this matrix and given a probability distribution of police's places for the current step, we can calculate the probability distribution over police's places for the next time step (or any time step in the future). We denote the probability distribution of police's places as $\vec{c} = (c_1, c_{1 \rightarrow 2}, c_{2 \rightarrow 1}, c_2, \dots, c_{N \rightarrow (N-1)}, c_N)$

$$\vec{c}^{t+1} = T_p \cdot \vec{c}^t \quad (2)$$

Assuming the transition matrix is ergodic, there exists a stationary coverage of each place \vec{c}^s , which is one of the eigenvectors of T_p with eigenvalue 1. We assume T_p has only one eigenvalue equal to 1, in which case there exists a unique stationary coverage. The stationary coverage can be computed in polynomial time by matrix inversion.

Crime Diffusion

Crime diffuses throughout the transportation network as the criminals move around and opportunistically commit crimes. We start by focusing on a single criminal. Before further analysis, we first enumerate the inputs for this crime diffusion model.

1. *Input 1* : The current station of the criminal.
2. *Input 2* : The Markov strategy of police.
3. *Input 3* : The stationary coverage of police for each location.
4. *Input 4* : A boolean value for the relationship between the criminal and the police in terms of their places: 1 if they are at the same station; 0 if not.
5. *Input 5* : The *attractiveness* of stations \vec{Att} , which influences the criminal's probabilities of committing crimes as described below.

We can use inputs 1, 4, and 5 to calculate the criminal's probability of committing a crime in the action phase at the current time step. There are two factors that affect the probability of the criminal committing a crime. First, the place relationship between the police and the criminal at this time step. If they are at the same station, the probability that the criminal will *attempt* to commit crime is p_s ; else it is p_d . Second, the attractiveness of the criminal's current station. The concept of attractiveness is from the literature on opportunistic crime, and measures the availability of crime opportunities, i.e., how "easy" it is to commit a crime at different stations. Hence, it measures the probability of successfully committing a crime at a station assuming that the criminal attempts to commit crimes there in the first place. Thus, the crime probability is the product of two terms: if the criminal is at station i and the police is not at that station at this time

step, he will commit a crime with probability $p_d \cdot Att(i)$; otherwise, the criminal and the police are both at station i at this time step, and he will commit a crime with probability $p_s \cdot Att(i)$. To simplify the problem, we assume for the remainder of this work that $p_s = 0$ and $p_d = 1$. That is, if the criminal and police are at different stations, then the criminal will certainly attempt to strike, and if they are at the same station, the criminal will certainly not attempt to strike. Formally, then, let $p_c(i)$ be the probability of successfully committing a crime if the criminal is at station i , then

$$p_c(i) = \begin{cases} 0 & \text{if police is at station } i \\ Att(i) & \text{otherwise.} \end{cases} \quad (3)$$

Example 3: We still use the scenario in Example 1, where there are 3 stations (1, 2, 3). Suppose the attractiveness of station 1 is 0, station 2 is 0.5, and station 3 is 1, and that the criminal is at station 1. If the police is at station 2, $p_c(1) = Att(1) = 0$; $p_c(2) = 0$ (since $p_s = 0$); $p_c(3) = Att(3) = 1$.

The criminal will use inputs 1, 2, 3, and 4 to update his belief state of the police's place. In theory, the criminal could base his belief on the entire history of his observations. However, he would then need to carry out complicated calculations and this conflicts with our assumption of boundedly-rational criminals. Instead, we use a *memory wipe* model of the criminal, which means that the criminal forgets what he observed in previous time steps and only remembers his current observation. If the criminal is at station i and the police is also here, he will set his current belief for the coverage of police to $\vec{c}_{b,0}$ as $(0, 0, \dots, 1, 0, \dots, 0)^T$, where row $3i - 2$, which stands for the coverage of station i , is 1 and others are 0; otherwise, the criminal is at station i while the police is not, and $\vec{c}_{b,0}$ will be set as $\left(\frac{\vec{c}_s(1)}{1 - \vec{c}_s(3i - 2)}, \frac{\vec{c}_s(2)}{1 - \vec{c}_s(3i - 2)}, \dots, 0, \frac{\vec{c}_s(3i - 1)}{1 - \vec{c}_s(3i - 2)}, \dots, \frac{\vec{c}_s(3N - 2)}{1 - \vec{c}_s(3i - 2)} \right)^T$, where row $3i - 2$ is 0 and others rows are proportional to the corresponding rows in the stationary distribution \vec{c}_s . The future belief state of the police's places can be calculated using the current belief state and the police transition matrix:

$$\vec{c}_{b,t} = T_p^t \cdot \vec{c}_{b,0}, \quad (4)$$

where $\vec{c}_{b,t}$ denotes the criminal's belief state of police's place in t time steps.

After we have the belief state for future **time steps**, we can calculate the *expected value* for each station at the next **strike**. The expected value for one station is the revenue that the criminal expects to get if he chooses this station as the target for the next strike. There are two factors contributing to this revenue. The first part is the probability of a successful crime; the second part is the expected value for one successful crime at the next strike. If the criminal is at station i for the current strike and chooses station j for the next strike, there are $|i - j| + 1$ time steps between these two

Input: i : the criminal's station; T_p : the transition matrix for police; \vec{c}_s : the police's stationary coverage for all places; *relation*: the place relationship between the criminal and the police; \vec{Att} : the attractiveness of the stations; N : the number of stations; u_c : the utility of one crime; λ : the rationality of the criminal

Output: $\vec{p}(\cdot|i, \vec{c}_{b,0})$: The criminal's probability distribution for next target

```

1 Initial  $\vec{c}_{b,0}$  with a  $1 \times 3N - 2$  zero vector ;
2 if relation = 1 then
3    $\vec{c}_{b,0}(3i - 2) = 1$ ;
4 end
5 if relation = 0 then
6   for  $j \in Station$  do
7      $\vec{c}_{b,0}(j) = \frac{\vec{c}_s(j)}{1 - \vec{c}_s(3i - 2)}$ 
8   end
9    $\vec{c}_{b,0}(3i - 2) = 0$ 
10 end
11 for  $j \in Station$  do
12    $t = |i - j| + 1$ ;
13    $\vec{c}_{b,t} = T_p^t \cdot \vec{c}_{b,0}$ ;
14    $E(j|i, \vec{c}_{b,0}) = (1 - \vec{c}_{b,t}(j))Att(j) \cdot u_c / t$ ;
15 end
16 for  $j \in Station$  do
17    $p(j|i, \vec{c}_{b,0}) = \frac{(E(j|i, \vec{c}_{b,0}))^\lambda}{\sum_{h=1}^N (E(h|i, \vec{c}_{b,0}))^\lambda}$ ;
18 end
19 return  $\vec{p}(\cdot|i, \vec{c}_{b,0})$ ;

```

Algorithm 1: QBRM ALGORITHM

strikes. Hence, the probability of a successful crime occurring during a strike at station j is

$$p_c(j) = (1 - \vec{c}_{b,|i-j|+1}(j))Att(j).$$

The second part is the value the criminal assigns for one successful crime at the next strike. For each successful crime, the real utility the criminal gains is u_c . However, we assume that the value the criminal places on a successful crime is inversely proportional to the time he spent during the strike. The time associated with the next strike being located at j is $(|i - j| + 1)$. Considering these two factors, the value placed on one successful crime at the next strike is $\frac{u_c}{|i-j|+1}$. Taking this all together, the expected value for choosing station j as the next strike if the criminal is currently at station i is

$$E(j|i, \vec{c}_{b,0}) = \frac{(1 - \vec{c}_{b,|i-j|+1}(j))Att(j) \cdot u_c}{(|i - j| + 1)}. \quad (5)$$

The criminal is boundedly rational: instead of always picking the station with highest expected value, his movement is a random movement process that is biased toward stations of high expected values. Given the expected value for each station $E(\cdot|i, \vec{c}_{b,0})$, we determine the probability

distribution for each of them being chosen as the next strike, $\vec{p}(\cdot|i, \vec{c}_{b,0})$ through the equation

$$p(j|i, \vec{c}_{b,0}) = \frac{(E(j|i, \vec{c}_{b,0}))^\lambda}{\sum_{h=1}^N (E(h|i, \vec{c}_{b,0}))^\lambda}, \quad (6)$$

where $\lambda > 0$ is a parameter that describes the criminal's level of rationality. This is an instance of the *quantal response* model of boundedly rational behavior (Luce 1959; McFadden 1974; McKelvey and Palfrey 1995). The criminological and game-theoretic effects described earlier are incorporated into the model by the expected value. A mathematical description of the criminal's movement is given in Algorithm 1.

Optimization Formulation in a Markov Chain

We have now fully described the behavior of the police and the criminals in the metro system; however it is not immediately obvious how to formulate the police's problem of finding an optimal strategy as a computational problem. In this section we give such a formulation, by modeling the game between the police and the criminal as a finite state Markov chain. As the interactions between the police and criminals are independent across different criminals, without loss of generality we can focus on the interaction between police and a representative criminal.

Each state of this Markov chain is a combination of the criminal's station and the police's place. Here we only consider the situations where the criminal is at a station as states because the criminal only makes decisions at stations. Since there are N stations and $3N - 2$ places, the total number of states is $N(3N - 2)$. Because only the criminal's stations are in the Markov chain states, transitions for these states are based on *strikes* rather than *time steps*.

The transition matrix for this Markov chain, denoted as T_s , can be calculated by combining the police patrol model and the crime diffusion model. However, note that the crime diffusion model is determined by the police's patrol strategy, because it models the interactions between the criminal and the police. Therefore the police's Markov strategy π fully determines T_s .

Each element in T_s represents the transition probability from one state to another. For further analysis, we pick the element $p_{S1 \rightarrow S2}$ that represents the transition probability from $S1$ to $S2$. In $S1$, the criminal is at station $i \in \{1, 2, \dots, N\}$ while the police is at location $m \in \{1, 2, \dots, 3N - 2\}$ and in $S2$, the criminal is at station j while the police is at location n . We need two steps to calculate this transition probability.

The first step is to calculate the transition probability of the criminal from $S1$ to $S2$. Given $S1$, $S2$, and the police's Markov strategy π , we have all the inputs in the crime diffusion model. Algorithm 1 uses these inputs to achieve the belief state updating for the criminal. It then generates expected utility of each station for the criminal. Finally, it gives us the probability for the criminal to choose j for the next strike, $p(j|i, \vec{c}_{b,0})$, in terms of π .

The second step is to calculate the transition probability of the police from $S1$ to $S2$. The number of time steps needed

to transition from $S1$ to $S2$ is $|i-j|+1$ so we should consider the probability distribution of the police's place in $|i-j|+1$ time steps. Denote by $\vec{c}_{|i-j|+1}(n)$ the probability that the police is at place n in $|i-j|+1$ time steps, we have

$$\vec{c}_{|i-j|+1}(n) = (T_p^{|i-j|+1} \cdot \vec{c}_0)(n), \quad (7)$$

where \vec{c}_0 is the basis vector corresponding to the current location, which is 1 for row m and 0 for others.

After we find the transition probabilities for both the criminal and the police, we can calculate $p_{S1 \rightarrow S2}$ as the product of these two,

$$p_{S1 \rightarrow S2} = p(j|i, \vec{c}_{b,0}) \cdot \vec{c}_{|i-j|+1}(n). \quad (8)$$

Because $S1$ and $S2$ can be any states in the Markov chain, we can represent all the elements in the transition matrix in terms of the police's Markov strategy.

Given this Markov chain model, we can calculate the utility for the police at each strike. For each successful crime, police receive utility $u_p < 0$; i.e., the successful occurrence of crimes is a bad thing for police. If there is no crime or the crime attempt fails, the police receive utility 0. At each state in the Markov chain, we can find the expected number of successful crimes and therefore evaluate the police's expected utility. This is not a zero-sum game because we do not consider the time factor in police's expected utility.

Example 4: *Given the same scenario as Example 1, where there are 3 stations in total, we will enumerate the police's expected utility for some state. In $S1$, where the criminal is at station 1 while the police is also at station 1, the police's expected utility is $r_p = 0$, because the criminal will not commit a crime in this instance (recall p_s is assumed 0). In $S2$, where the criminal is at station 2 while the police is at station 1, the police's expect utility is $r_p = Att(2) \cdot u_p$.*

We can use a $1 \times N(3N - 2)$ vector R_p to represent the utility for all states. R_p is a pre-generated vector based on u_p and Att only. $V_p(k)$, the police's expected utility during strike number k (note that the strike number k is **not** necessarily the same as the strike location), is

$$V_p(k) = R_p \cdot X_k, \quad (9)$$

where X_k is the state distribution at strike number k . Using the transition matrix T_s and the exit rate for the criminal after each strike α , we can calculate X_k from the initial state distribution X_1 :

$$X_k = ((1 - \alpha) \cdot T_s)^{k-1} X_1. \quad (10)$$

The objective of the police is the total expected utility, summed over all the strikes:

$$\begin{aligned} Obj &= \lim_{K \rightarrow \infty} \sum_{k=0}^K V_p(k+1) \\ &= \lim_{K \rightarrow \infty} \sum_{k=0}^K R_p \cdot ((1 - \alpha) \cdot T_s)^k X_1 \\ &= R_p \cdot (I - (1 - \alpha)T_s)^{-1} X_1, \end{aligned} \quad (11)$$

where I is the identity matrix. For the last equality we make use of the geometric sum formula and the fact that the largest

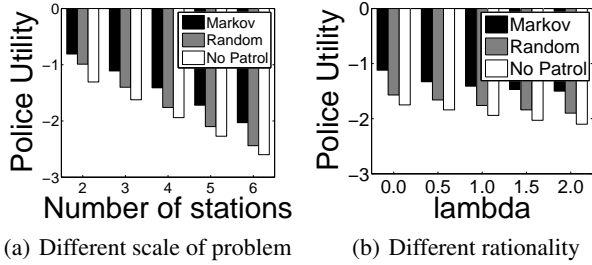


Figure 4: Experimental result

eigenvalue of T_s is 1, and therefore $I - (1 - \alpha)T_s$ is nonsingular for $0 < \alpha < 1$.

The only unknown factor in the objective is the transition matrix T_s , which can be expressed in terms of the police’s Markov strategy π via Equations (4), (5), (6), (7), and (8). We have thus formulated the police’s problem of finding the optimal Markov strategy to commit to as a nonlinear optimization problem, specifically to choose π to maximize Obj (that is, minimize the total amount of crime in the network).

Experiments

The optimization problem (11) is non-linear due to the criminal’s QBRM (6) and the multiplication of the police’s transition matrix. We solve this non-linear optimization using the `FindMaximum` function in Mathematica. This function automatically chooses a local-optimal solver for the non-linear problem. Possible solvers include Conjugate Gradient, Principal Axis, Levenberg Marquardt, Newton, Quasi Newton and Interior Point. In this initial report, we evaluate the solution quality for this Markov chain model against crime diffusion.

Data Sets

For our experiments we used a metro network that is similar to the toy network we introduced in the Problem Setting section, which is a simplification of real-world situations. In our metro network, there were a number of stations N . The stations in the network were along a straight line. The distances between any two neighboring stations are the same and each train needed one time step to go from one station to its neighbor. There are always two trains on a station heading to different directions in the decision phase of any time step. The initial distribution of police places was set as her stationary coverage. For simplicity, we assume in police’s Markov strategy, the distribution over actions at a train is the same as that at the station where the train will reach at the decision phase of next time step. This is feasible because the available actions in both situations are the same. We set the attractiveness of station i as a function of i , $Att(i) = 0.05(i + 1)$. Here, because we are dealing with small scale problems, we assume $i < 19$; as a result, $Att(i) < 1$, which is consistent with the definition of the attractiveness of stations. The police’s utility of a successful crime was $u_p = -1$. In our experiments, we assumed the distribution of criminal’s initial stations was uniform over all the stations. We set the exit

rate of the criminal $\alpha = 0.1$. Finally, we chose the revenue of a successful crime to be $u_c = 1$; however, any $u_c > 0$ will give the same results, as the factor u_c^λ can be factored out of both the numerator and denominator of Eq.(6), so that the criminal’s behavior is actually independent of this parameter.

We applied two different patrol strategies for police. The first is the uniformly random strategy, which means the police will randomly choose her action from all the available actions at current place with the same probability. For any given stations i except the two end stations, $l_i = r_i = s_i = 1/3$. For the end stations, the probabilities of moving and staying are both $1/2$. The second strategy applied is the strategy returned from the non-linear solver. When solving the non-linear program in Mathematica, we set the starting point as the uniformly random strategy.

Results

In our first set of experiments, we fixed $\lambda = 1$ and varied the number of stations N . Figure 4(a) shows the police’s objective, which is the total utility of police over all the strikes. Since our nonlinear solver is only guaranteed to return a local maximum of the objective function, this is only the lower bound for the globally optimum solution. As we can see, the police’s utility decreased as the number of stations increased. Unsurprisingly, police patrol can significantly increase police’s utility against crime diffusion and the Markov strategy from non-linear programming outperformed the random strategy for any number of stations. At the same time, the ratio between the utility we get from the optimal Markov strategy and the random strategy was stable. The ratio for 2 stations is 0.82; for 3 stations, it is 0.79; for 4 stations, it is 0.80; for 5 stations, it is 0.82; and for 6 stations, it is 0.83. The police’s utility decreased nearly linearly with the increase of the number of stations, which is to be expected for a constant number of police. The theoretic lower bound of the objective Obj_m occurs in the situation where the criminal commits a crime at each strike before he exits the system. Given this behavior, $Obj_m = u_p + (1 - \alpha) \cdot u_p + \dots + (1 - \alpha)^N \cdot u_p + \dots = u_p \cdot \frac{1}{\alpha} = -9$. The police’s utilities in small scale problems were much higher than Obj_m , which meant that the patrol could significantly control crime diffusion in these cases.

To study the effect of criminal’s rationality to the police’s objective, we fixed the station number to 4 and varied λ . Figure 4(b) shows the police’s objective when pitted against a criminal with λ from 0 to 2. As we can see, the non-linear programming Markov strategy outperformed the random strategy significantly for any λ . As the criminal’s rationality increased, the police’s utility decreased. This was because the criminal would commit more successful crimes when he became more strategic.

The last result that we mention is the run time of solving the optimization. As we were dealing with small scale problems in this paper, the run time was not a big concern. Using Mathematica on a standard 2.4 GHz machine with 8 GB memory, the longest runtime was less than 2 minutes, which was for solving the scenario with 6 stations.

Summary

In this paper we presented a novel crime diffusion model for opportunistic and boundedly rational criminals. We also provided a game-theoretic approach for controlling such diffusion via randomized police patrols. Unlike previous Stackelberg game models for security, in which the attackers strategically plan their attack in advance, the criminals in our model would react to real-time information. We provided a novel formulation of the resulting problem of scheduling police patrols as a nonlinear optimization problem on a Markov chain. The experimental results showed that the strategy that we generated outperformed the uniformly random strategies in different scales of problem instances and with different types of criminals. Scalability remains a major challenge: the current algorithm is only fit for small-scale instances of the problem. In the future, we plan to find more efficient algorithms to solve real-world scale problem instances; and one possible approach is to approximate this non-linear crime diffusion model with something computationally more tractable.

Acknowledgement

This research is supported by MURI grant W911NF-11-1-0332.

References

- Block, R., and Block, C. R. 2000. The bronx and chicago: Street robbery in the environs of rapid transit stations.
- Brantingham, P., and Brantingham, P. 1995. Criminality of place. *European Journal on Criminal Policy and Research* 3(3):5–26.
- Brantingham, P. J., and Tita, G. 2008. Offender mobility and crime pattern formation from first principles. *Artificial crime analysis systems: using computer simulations and geographic information systems* 193–208.
- Felson, M.; Clarke, R. V. G.; and Britain, G. 1998. *Opportunity makes the thief: Practical theory for crime prevention*, volume 98. Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate.
- Jiang, A. X.; Yin, Z.; Zhang, C.; Tambe, M.; and Kraus, S. 2012. Game-theoretic randomization for security patrolling with dynamic execution uncertainty.
- Keizer, K.; Lindenberg, S.; and Steg, L. 2008. The spreading of disorder. *Science* 322(5908):1681–1685.
- Loukaitou-Sideris, A.; Liggett, R.; and Iseki, H. 2002. 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(2):135–151.
- Loukaitou-Sideris, A. 1999. Hot spots of bus stop crime: The importance of environmental attributes. *Journal of the American Planning association* 65(4):395–411.
- Luce, R. D. 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley.
- Matthews, S. A.; Yang, T.-C.; Hayslett, K. L.; and Ruback, R. B. 2010. Built environment and property crime in seattle, 1998–2000: A bayesian analysis. *Environment and planning. A* 42(6):1403.
- McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers of Econometrics* 105–142.
- McKelvey, R. D., and Palfrey, T. R. 1995. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10(1):6–38.
- Nguyen, T. H.; Tsai, J.; Jiang, A.; Bowering, E.; Maheswaran, R.; and Tambe, M. 2012. Security games on social networks. In *2012 AAAI Fall Symposium Series*.
- Shieh, E.; An, B.; Yang, R.; Tambe, M.; Baldwin, C.; DiRenzo, J.; Maule, B.; and Meyer, G. 2012. Protect: A deployed game theoretic system to protect the ports of the united states. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 13–20. International Foundation for Autonomous Agents and Multiagent Systems.
- Short, M. B.; D’Orsogna, M. R.; Pasour, V. B.; Tita, G. E.; Brantingham, P. J.; Bertozzi, A. L.; and Chayes, L. B. 2008. A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences* 18(supp01):1249–1267.
- Taylor, R. B., and Harrell, A. 1996. *Physical environment and crime*. US Department of Justice, Office of Justice Programs, National Institute of Justice.
- Tsai, J.; Kiekintveld, C.; Ordonez, F.; Tambe, M.; and Rathi, S. 2009. Iris-a tool for strategic security allocation in transportation networks.
- Tsai, J.; Nguyen, T. H.; and Tambe, M. 2012. Security games for controlling contagion. In *Twenty-Sixth National Conference in Artificial Intelligence*.
- Yin, Z.; Jiang, A.; Johnson, M.; Tambe, M.; Kiekintveld, C.; Leyton-Brown, K.; Sandholm, T.; and Sullivan, J. 2012. Trusts: Scheduling randomized patrols for fare inspection in transit systems. In *Proc. of the 24th Conference on Innovative Applications of Artificial Intelligence, IAAI*.
- Zipkin, J. R.; Short, M. B.; and Bertozzi, A. L. 2013. Cops on the dots in a mathematical model of urban crime and police response.