

Opportunistic Security Game: An Initial Report

Chao Zhang¹, Albert Xin Jiang¹, Martin B. Short², P. Jeffrey Brantingham³, and Milind Tambe¹

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

² Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, USA,
mbshort@math.gatech.edu,

³ University of California, Los Angeles, CA 90095, USA
pjb@anthro.ucla.edu

Abstract. This paper introduces a new game-theoretic framework and algorithms for addressing opportunistic crime. Stackelberg Security Game (SSG), focused on highly strategic and resourceful adversaries, has become an important computational framework within multiagent systems. Unfortunately, SSG is ill-suited as a framework for handling opportunistic crimes, which are committed by criminals who are less strategic in planning attacks and more flexible in executing them than SSG assumes. Yet, opportunistic crime is what is commonly seen in most urban settings. We therefore introduce Opportunistic Security Game (OSG), a computational framework to recommend deployment strategies for defenders to control opportunistic crimes. Our first contribution in OSG is a novel model for opportunistic adversaries, who (i) opportunistically and repeatedly seek targets; (ii) react to real-time information at execution time rather than planning attacks in advance; and (iii) have limited observation of defender strategies. Our second contribution to OSG is a new exact algorithm EOSG to optimize defender strategies given our opportunistic adversaries. Our third contribution is the development of a fast heuristic algorithm to solve large-scale OSG problems, exploiting a compact representation. We use urban transportation systems as a critical motivating domain, and provide detailed experimental results based on a real-world system.

Keywords: Game theory, Security games, Optimization

1 Introduction

Stackelberg Security Game (SSG), an important computational framework within multiagent systems [1, 2], enables security resource allocations against highly strategic and capable adversaries who conduct careful surveillance and plan attacks. While there are undoubtedly such highly capable adversaries [1], they likely comprise only a small proportion of the overall set of adversaries in the urban security domain.

This paper focuses on the majority of adversaries for urban security: criminals who have little planning or surveillance before attacking [3, 4]. These adversaries capitalize on locally opportunities and react to real-time information. Unfortunately, SSG is ill-suited to model such criminals, as it attributes too much planning and little execution

flexibility to adversaries. Therefore, based on modern criminological theory [3–5], this paper introduces Opportunistic Security Game (OSG), a new computational framework for addressing opportunistic crime. OSG fundamentally differs from SSG, and opens the door to new research at the intersection of computational game theory and criminology.

This paper provides three key contributions in introducing OSG. First, OSG includes a model of opportunistic criminals who exhibit a stochastic pattern of movement to search for crime opportunities. This movement, modeled as Quantal Biased Random Movement, based on a previous model of criminal motion [6], is quite different from a fixed route that an adversary is assumed to pursue in SSG. In OSG, criminals react to real-time information about rather than committing to a single plan [7, 5]. Additionally, OSG applies anchoring-bias [8] to model criminals’ limited knowledge. Our second contribution is a new algorithm EOSG to generate patrol schedules that optimize the defender’s expected utility against this new kind of adversary. OSG is similar to SSG: the defender must commit to her patrol strategy first, after which the criminals will choose targets to attack. However, there are two main differences. First, in OSG, criminals react to real-time information. Second, after any attempt, criminals in OSG can remain in the system and search for another opportunity. The third contribution is a fast algorithm, Compact Opportunistic Security game states (COPS), to solve large scale OSG problems.

In motivating this work, we focus on crime in urban transportation systems as it is an important challenge across the world. Indeed, transportation systems are at unique risk of crime because of the way in which they concentrate large numbers of people in time and space [9–11]. The challenge in controlling this crime in transportation systems can be seen as an OSG: police patrol to control crime while criminals hunt for crime opportunities. Criminals are known to travel based on their knowledge of crime opportunities [12, 13], usually committing crimes such as thefts and snatches at stations – where it is easy to escape if necessary [14]. These opportunistic criminals avoid targets if security presence is observed there [15].

2 Motivating Domain

Crime in urban transportation systems — including buses, trams and metro trains – is a critically important challenge in cities across the world. Transportation systems play an important role in driving local crime patterns, and may themselves be at unique risk of crime because of the way in which they concentrate large numbers of people in time and space [9–11].

Given that a substantial portion of criminals use public transit as a primary mode of transportation [14], it is reasonable to infer that a significant number of criminals spend extended periods of time traveling through the transportation system seeking to commit crimes. Deploying police to patrol is then one way to control such crime in a transportation system. Thus, whereas the police move within the transportation system in an attempt to dissuade crime, criminals move within this same system based on their limited knowledge of the police strategy and their knowledge of how crime opportunities may be distributed among target locations. These opportunistic criminals

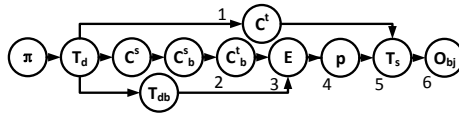


Fig. 1. Flow chart of OSG

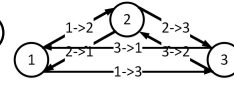


Fig. 2. The metro net work

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

Fig. 3. Example of strikes

may continue to hunt for crime opportunities by, for example, traveling to a different target location (station) when a security officer is observed at their current location.

In this work, we consider a metro rail system (e.g., the LA metro rail system) as a specific example system and modeling objective. We consider two major components in a metro rail system, stations and trains. As shown in crime data from the LA metro system, criminals usually commit crimes such as thefts and snatches at stations – where it is easy to escape if necessary.

3 OSG Framework

In OSG, the defender (“she”) – assisted by our algorithms – is modeled to be perfectly rational. The criminal (“he”) is modeled with cognitive biases. Figure 1 illustrates the OSG flowchart, with the numbers near variables referring to equations in this section; useful notation is shown in Table 1.

π	Defender’s Markov strategy	\mathbf{c}_b^s	Criminal’s belief of \mathbf{c}^s
T_d	Defender transition matrix	\mathbf{c}_b^t	Criminal’s belief of \mathbf{c}^t
\mathbf{c}^s	Defender stationary coverage	T_d^t	Criminal’s belief of T_d
\mathbf{c}^t	Defender coverage vector at time step t	E	The expected value of targets for criminals
T_s	Transition matrix for the OSG Markov chain	p	Criminal’s next strike probability distribution
Obj	Defender’s objective		

Table 1. Notation

3.1 Player decision and actions in OSG

We consider as a generic example a metro rail system and its two major components, stations and trains, which we collectively refer to as locations. We denote the number of stations in the system as N . The stations are labeled $1, \dots, N$. The train from station i to its neighboring station j is denoted as $i \rightarrow j$. We denote the number of locations as $N_l > N$. We divide time equally into time steps so that trains arrive at stations at the beginning of each time step. There are two phases in any time step. The first phase, the *decision phase*, is the beginning period when trains are at stations. In this phase, each passenger at each location can decide where to move next. There are two choices available. $Go\ i \rightarrow j$ means that (i) if a passenger is at station i , he/she gets on the

train $i \rightarrow j$; (ii) if he/she is on a train arriving at station i , he/she gets off the current train and gets on the train $i \rightarrow j$ unless the current train is $i \rightarrow j$. *Stay* means that the passenger stays at the station, so that if the passenger was on a train, then he/she gets off. We denote by $p_l^{i \rightarrow j}$ and s_l the probabilities of *Go* $i \rightarrow j$ and *Stay* for location l , respectively.

At the end of each decision phase, during the *action phase*, trains depart in all possible directions from all stations. In this phase, trains are in motion from one station to another. For simplicity we assume that the time it takes to travel between any two adjacent stations is identical; this assumption can be relaxed by including dummy stations. This model matches metro systems in Los Angeles, where trains leave stations at regular intervals in all directions. Next, we focus on two types of passengers: defenders and criminals.

3.2 Modeling Defenders

A defender is a team of police officer using trains for patrolling to mitigate crime. We start with a single defender and deal with multiple defenders later. The defender conducts randomized patrols using a *Markov Strategy* π , which specifies for each location a probability distribution over all available actions. Using a Markov strategy, rather than explicitly representing each pure strategy, essentially implies that we use a compact representation for defender strategies.

Example 1: Markov Strategy *Figure 2 shows a simple scenario with 3 stations (1, 2, 3) and 6 trains (1 \rightarrow 2, 1 \rightarrow 3, 2 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 1, 3 \rightarrow 2), which is a fully connected topology. A possible Markov strategy π for the defender is,*

$$\begin{aligned} s_1 &= 0.2, p_1^{1 \rightarrow 2} = 0.5, p_1^{1 \rightarrow 3} = 0.3; \\ &\dots \\ s_{3 \rightarrow 2} &= 0.1, p_{3 \rightarrow 2}^{2 \rightarrow 1} = 0.8, p_{3 \rightarrow 2}^{2 \rightarrow 3} = 0.1; \end{aligned}$$

$s_{3 \rightarrow 2} = 0.1, p_{3 \rightarrow 2}^{2 \rightarrow 1} = 0.8, p_{3 \rightarrow 2}^{2 \rightarrow 3} = 0.1$ means that if the defender is on the train from station 3 to station 2, then: she will have probability 0.1 to choose action *Stay*, probability 0.8 to choose action *Go* 2 \rightarrow 1, and probability 0.1 to choose action *Go* 2 \rightarrow 3 at the next decision phase.

Given π , the defender's movement is a Markov chain over the locations with *defender transition matrix* T_d , whose entry at column k , row l specifies the probability of the defender going from location k to location l in one time step. We choose an index such that i ($i \in 1, \dots, N$) represents station i ; indexes larger than N represent trains.

Example 2: For Example 1, T_d is as follows:

$$\begin{array}{l}
 1 \\
 2 \\
 3 \\
 1 \rightarrow 2 \\
 1 \rightarrow 3 \\
 2 \rightarrow 1 \\
 2 \rightarrow 3 \\
 3 \rightarrow 1 \\
 3 \rightarrow 2
 \end{array}
 \begin{pmatrix}
 1 & 2 & \dots & 2 \rightarrow 3 & 3 \rightarrow 1 & 3 \rightarrow 2 \\
 s_1 & 0 & 0 & s_{3 \rightarrow 1} & 0 & \\
 0 & s_2 & 0 & 0 & s_{3 \rightarrow 2} & \\
 0 & 0 & s_{2 \rightarrow 3} & 0 & 0 & \\
 p_1^{1 \rightarrow 2} & 0 & 0 & p_{3 \rightarrow 1}^{1 \rightarrow 2} & 0 & \\
 p_1^{1 \rightarrow 3} & 0 & 0 & p_{3 \rightarrow 1}^{1 \rightarrow 3} & 0 & \\
 0 & p_2^{2 \rightarrow 1} & 0 & 0 & p_{3 \rightarrow 2}^{2 \rightarrow 1} & \\
 0 & p_2^{2 \rightarrow 3} & 0 & 0 & p_{3 \rightarrow 2}^{2 \rightarrow 3} & \\
 0 & 0 & p_{2 \rightarrow 3}^{3 \rightarrow 1} & 0 & 0 & \\
 0 & 0 & p_{2 \rightarrow 3}^{3 \rightarrow 2} & 0 & 0 &
 \end{pmatrix}$$

We denote the defender's location coverage vector at time t as \mathbf{c}^t . Using T_d and \mathbf{c}^t , we can calculate the coverage vector at time step $t_1 > t$ through the formula

$$\mathbf{c}^{t_1} = (T_d)^{t_1-t} \cdot \mathbf{c}^t. \quad (1)$$

We restrict each element in π to be strictly positive so that T_d is ergodic, meaning it is possible to eventually get from every location to every other location in finite time. For an ergodic T_d , there is a unique *stationary coverage* \mathbf{c}^s , such that $T_d \cdot \mathbf{c}^s = \mathbf{c}^s$ [16]. The dependence of \mathbf{c}^s on T_d and hence on π is shown in Fig. 1. The defender's initial coverage, \mathbf{c}^1 , is set to \mathbf{c}^s so that the criminal will face an invariant distribution whenever he enters the system.

3.3 Modeling Opportunistic Criminals

Our model consists of three components: **The opportunistic criminal's probability to commit a crime at the current time step:** We ignore the possibility of crimes during the decision phases because they are instantaneous. In action phases, crimes may occur only at stations, as discussed in the Introduction, and occur with a probability determined by two factors. The first is the *attractiveness* of each target station [6], which measures the availability of crime opportunities at a station. Attractiveness measures precisely how likely a criminal located at that station during an action phase will commit a crime *in the absence of defenders*. \mathbf{Att} is the N vector of station attractiveness. The second factor is the presence of the defender. Specifically, if a criminal is at the same station as a defender, he will not commit a crime. Thus, his probability of committing a crime at station i will be influenced by $c^t(i)$. Using this strategy, the criminal will never be caught red handed by the defender, but may be forced toward a less attractive target. Considering both factors, the probability of the criminal committing a crime if located at station i during the action phase of time step t , denoted as $q_c(i, t)$, is $q_c(i, t) = (1 - c^t(i))Att(i)$.

Criminal's belief state of the defender: During the decision phase, the opportunistic criminal decides the next target station and moves on a path directly to that station. Hence, the criminal's motion within the metro system can be distilled down to a sequence of stations where he is located during action phases; we refer to these instances of attempted crime as *Strikes*. As shown in Fig. 3, only the time steps when the criminal is at stations are counted as strikes.

When making these target decisions, the criminal tends to choose stations with high *expected utilities*. He uses his knowledge of π and his real-time observations to make such decisions. Let T_{db} , \mathbf{c}_b^t and \mathbf{c}_b^s be his belief of T_d , \mathbf{c}^t and \mathbf{c}^s respectively. As the criminals are with limited surveillance capability, these beliefs may not be the same as T_d , \mathbf{c}^t , and \mathbf{c}^s . Therefore, we introduce the criminal's surveillance imperfection via *anchoring bias*, a cognitive bias that describes the human tendency to rely on a familiar, but not necessarily relevant, reference point when making decisions [17, 8]. We denote the level of the criminal's anchoring bias with the parameter b , where $b = 0$ indicates no anchoring bias, and $b = 1$ indicates complete reliance on such bias. We get $T_{db} = (1 - b) \cdot T_d + b \cdot T_u$, with corresponding stationary coverage \mathbf{c}_b^s , where T_u corresponds to a strategy in which the defender picks each available action with uniform probability.

At any given time step t , the criminal may use his beliefs and observations to calculate \mathbf{c}_b^t . It is reasonable to assume that the opportunistic criminal does not perform a perfect belief update using all observations; rather he only uses his current observation and belief to estimate \mathbf{c}_b^t . Specifically, if the criminal is at station i and the defender is also there, then \mathbf{c}_b^t is $(0, 0, \dots, 1, 0, \dots, 0)^T$, where row i is 1 and all others are 0. Otherwise the defender is not at i , and

$$\mathbf{c}_b^t = \frac{(c_b^s(1), c_b^s(2), \dots, 0, c_b^s(i+1), \dots, c_b^s(N_I))^T}{[1 - c_b^s(i)]}, \quad (2)$$

where row i is 0 and other rows are proportional to the corresponding rows in \mathbf{c}_b^s . Given \mathbf{c}_b^t and T_{db} , the belief coverage vector at time step t_1 ($t_1 > t$), $\mathbf{c}_b^{t_1}$, is calculated via Eq. 1.

We set the actual payoff for a crime to 1, but this can be generalized. The expected payoff for the criminal to choose station j as the next strike, given that the current strike is at station i at time step t , is $q_c(j, t + dt)$, where $dt \geq 1$ is the time needed to arrive at j from i . But, criminals are known to discount more distant locations when choosing targets [7]. Therefore, the *utility* that the criminal places on a given payoff is discounted over time. We implement this by dividing the payoff by the time taken. Finally, the criminal must rely on his belief of the defender's coverage when evaluating $q_c(j, t + dt)$. Altogether, station j has the expected utility $E(j|i, \mathbf{c}_b^t)$, which is:

$$E(j|i, \mathbf{c}_b^t) = \frac{(1 - [T_{db}^{dt} \mathbf{c}_b^t](j)) Att(j)}{dt}. \quad (3)$$

The criminal's probability distribution for each station being chosen as the next target: Finally, we use Quantal Biased Random Movement (QBRM) to model the criminal's bounded rationality. Instead of always picking the station with highest expected utility, his movement is modeled as a random process biased toward stations of high expected utility. Given the expected value for each station $\mathbf{E}(\cdot|i, \mathbf{c}_b^t)$, the probability distribution for each being chosen as the next strike, $\mathbf{p}(\cdot|i, \mathbf{c}_b^t)$ is determined by the equation

$$p(j|i, \mathbf{c}_b^t) = \frac{E(j|i, \mathbf{c}_b^t)^\lambda}{\sum_{h=1}^N E(h|i, \mathbf{c}_b^t)^\lambda}, \quad (4)$$

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 [18]. The

criminal may, as an alternative to choosing a further strike, leave the system at *exit rate* α .

To summarize, as shown in Fig. 1, the opportunistic criminal is modeled as follows. First, he commits a crime or not based on defender presence and attractiveness at current station. Next, he uses T_{db} and current observation to update c_b^t (Eq. 2). Finally, we use QBRM to model his next attack (Eq. 4). Algorithm 1, in the online appendix (<http://osgops.webs.com/>), is a full mathematical description of the criminal's movement.

4 Exact OSG (EOSG) algorithm

Given the defender and criminal models, the EOSG algorithm computes the optimal defender strategy by modeling the game as a finite state Markov chain. As all the criminals behave identically, we can focus on the interaction between the defender and one criminal without loss of generality.

Each state of the EOSG Markov chain is a combination of the criminal's station and the defender's location. Here we only consider situations where the criminal is at a station as states because he only makes decisions at stations. Since there are N stations and N_l locations, the number of states is $N \cdot N_l$. State transitions 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 defender and criminal model. For further analysis, we pick the element $p_{S1 \rightarrow S2}$ in T_s that represents the transition probability from state $S1$ to $S2$. Suppose in $S1$ the criminal is at station i while the defender is at location m at time step t , and in $S2$, the criminal is at station j while the defender is at location n at time step $t + dt$. We need two steps to calculate this transition probability. First, we find the transition probability of the criminal from i to j , $p(j|i, \mathbf{c}_b^t)$. Then, we find the defender's transition probability from m to n as described above, which is $c^{t+dt}(n) = (T_d^{dt} \cdot \mathbf{e}_m)(n)$, where \mathbf{e}_m is a basis vector for the current location m . The transition probability $p_{S1 \rightarrow S2}$ is therefore given by

$$p_{S1 \rightarrow S2} = p(j|i, \mathbf{c}_b^t) \cdot c^{t+dt}(n). \quad (5)$$

Since $p(j|i, \mathbf{c}_b^t)$ and $c^{t+dt}(n)$ are determined by π , $p_{S1 \rightarrow S2}$ is also in terms of π (see Figure 1), as is T_s .

Given this EOSG model, we can calculate the defender's utility at each strike. For each successful crime, the defender receives utility $u_d < 0$. If there is no crime, she receives utility 0. This is a non-zero-sum game because we do not consider the time discount factor in the defender's expected utility. Formally, we define a vector $\mathbf{r}_d \in \mathbb{R}^{N \cdot N_l}$ such that entries representing states with both criminal and defender at the same station are 0 while those representing states with criminal at station i and defender not present are $Att(i) \cdot u_d$. Then, the defender's expected utility $V_d(t)$ during strike number t is $V_d(t) = \mathbf{r}_d \cdot \mathbf{x}_t$, where \mathbf{x}_t is the state distribution at strike number t , which we can calculate from the initial state distribution \mathbf{x}_1 as follows: $\mathbf{x}_t = ((1 - \alpha) \cdot T_s)^{t-1} \mathbf{x}_1$. The

defender's total expected utility over all strikes is thus:

$$\begin{aligned}
Obj &= \lim_{\ell \rightarrow \infty} \sum_{t=1}^{\ell} V_d(t) \\
&= \lim_{\ell \rightarrow \infty} \sum_{t=1}^{\ell} \mathbf{r}_d \cdot ((1 - \alpha) \cdot T_s)^{t-1} \mathbf{x}_1 \\
&= \mathbf{r}_d \cdot (I - (1 - \alpha)T_s)^{-1} \mathbf{x}_1,
\end{aligned} \tag{6}$$

where I is an identity matrix. In this equation we use the geometric sum formula and the fact that the largest eigenvalue of T_s is 1, so that $I - (1 - \alpha)T_s$ is nonsingular for $0 < \alpha < 1$.

The objective is a function of the transition matrix T_s , which can be expressed in terms of π via Eqs. (1), (3), (4), and (5). We have thus formulated the defender'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).

5 OSG for multiple defenders

If multiple defenders all patrol the entire metro, applying the same Markov strategy, then they will often be at the same station simultaneously, and such duplication carries no benefit. Instead, we construct our multiple-defender strategy by dividing the metro up into K contiguous segments, and designating one defender for each segment, as is the case in typical real-world Metro patrolling. Each defender will have a different strategy that is particular to her segment.

Defenders: In k -th segment, the number of locations is n_i^k . Defender k patrols with the Markov strategy π_k . Her transition matrix is $T_{dk} \in \mathbb{R}^{n_i^k \times n_i^k}$. Her coverage vector at time t is \mathbf{c}_k^t , and \mathbf{c}_k^s is the stationary coverage. Hence, defender k 's behavior is the same as that in a single-defender OSG, while their collective behavior is described by the Markov strategy $\pi = (\pi_1, \pi_2, \dots, \pi_K)$. The transition matrix T_d is as follows, where we have dropped the trains between neighboring segments from the basis for T_d :

$$T_d = \begin{pmatrix} T_{d1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & T_{dK} \end{pmatrix}. \tag{7}$$

The coverage of all units at time step t is \mathbf{c}^t , and is defined as the concatenation of coverage vectors $(\mathbf{c}_1^t; \mathbf{c}_2^t; \dots; \mathbf{c}_K^t)$. \mathbf{c}^t sums to K since each \mathbf{c}_k^t sums to 1. The vector \mathbf{c}^t evolves to future time steps t_1 in the same way as before, via Equation 1. The overall stationary coverage is $\mathbf{c}^s = (\mathbf{c}_1^s; \mathbf{c}_2^s; \dots; \mathbf{c}_K^s)$.

Opportunistic criminals: The previous model for criminals still applies. However, any variables related to defenders (T_d , \mathbf{c}^t , \mathbf{c}^s) are replaced by their counterparts for the multiple defenders. Furthermore, the criminal in segment k at time t cannot observe defenders other than k . As a result, his belief of defender coverage is $\mathbf{c}_b^t = (\mathbf{c}_{b1}^s; \mathbf{c}_{b2}^s; \dots; \mathbf{c}_{b(k-1)}^s)$.

$\mathbf{c}_{bk}^t; \mathbf{c}_{b(k+1)}^s; \dots; \mathbf{c}_{bK}^s$). That is, his belief of coverage for segment k , \mathbf{c}_{bk}^t , is calculated as in a single defender scenario, while his belief of coverage for other segments is the stationary coverage. Algorithm 2 (online appendix) describes a criminal’s behavior in multiple defenders settings.

Markov chain: In optimizing defender strategies via a Markov chain, each state records the station of the criminal and the location of *each* defender. As a result, each state is denoted as $S = (i, m_1, \dots, m_K)$, where the criminal is at station i and defender k is at location m_k . Since defender k can be at n_k^k different locations, the total number of states is $N \cdot n_1^1 \cdot \dots \cdot n_K^K$. To apply EOSG for multiple defenders, T_s is still calculated using the defender and criminal models. The transition probability $p_{S_1 \rightarrow S_2}$ from $S_1 = (i, m_1, \dots, m_K)$ at time t to $S_2 = (j, n_1, \dots, n_K)$ at time $t + dt$ is

$$p_{S_1 \rightarrow S_2} = p(j|i, \mathbf{c}_b^t) \prod_k e^{t+dt}(n_k),$$

where $e^{t+dt}(n_k) = (T_d^{dt} \cdot \mathbf{e}_{m_1, m_2, \dots, m_K})(n_k)$ and $\mathbf{e}_{m_1, m_2, \dots, m_K}$ is an indicator vector with 1 at entries representing locations m_1, m_2, \dots, m_K and 0 at all other entries. The state distribution \mathbf{x} and revenue \mathbf{r}_d are both $N \cdot n_1^1 \cdot \dots \cdot n_K^K$ vectors. The defenders’ total expected utility over all stations and strikes is given by Equation 6. Similarly, we are faced with the problem of finding a π to maximize *Obj*.

6 The COPS Algorithm

Unfortunately, EOSG fails to scale-up due to the size of T_s in Equation (6), which is $N \cdot N_l$ by $N \cdot N_l$ for one defender, and is much greater for multiple defenders. We propose the Compact Opportunistic Security game state (COPS) algorithm to accelerate the computation by compactly representing the states. The size of the transition matrix in COPS is $2N \times 2N$, *regardless of the number of defenders*, which is dramatically smaller than in EOSG. COPS is inspired by the Boyen-Koller(BK) algorithm on Dynamic Bayesian Networks [19], but improves upon a direct application of this algorithm.

In OSG with one defender, there are two components in each state, the criminal’s station S_c^t and the defender’s location θ_d^t , which are correlated when evolving. We introduce an intermediate component, the criminal’s observation O_c^t , which is determined by S_c^t and θ_d^t . Given S_c^t and O_c^t , we can compute S_c^{t+1} . The evolution of θ_d^t is independent of S_c^t , as shown in Figure 4(a). This is an instance of a Dynamic Bayesian Network: S_c^t , O_c^t , and θ_d^t are the random variables, while edges represent probabilistic dependence.

A direct application of the BK compactly represents the states by the marginal distribution of S_c^t and θ_d^t , denoted as $\Pr(S_c^t)$ and $\Pr(\theta_d^t)$ respectively, and then restores the Markov Chain states by multiplying these marginal distributions. We set $\Pr(\theta_d^t) = \mathbf{c}^s$ for all strikes; thus, we do not need θ_d^t in the new state and the number of the new states is just N . However, such an approximation throws away the strong correlation between the criminal’s station and defender unit’s location — in our preliminary experiments, this approximation led to low defender expected utility.

In contrast, our COPS algorithm compactly represents the original states with less information lost. Instead of just considering the marginal distributions of each component $\Pr(\theta_d^t)$ and $\Pr(S_c^t)$, we also include O_c^t in the approximate states. O_c^t is binary:

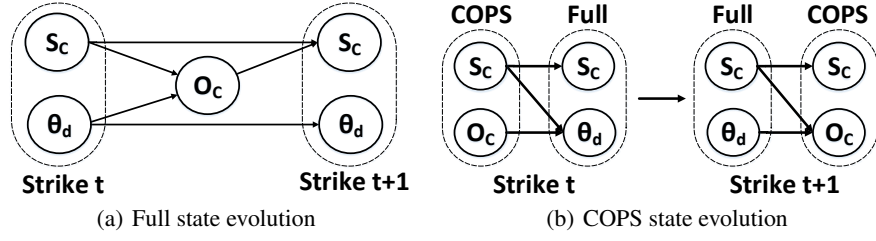


Fig. 4. COPS algorithm

1 if the defender is at the same station with the criminal, 0 otherwise. The new approximate states, named *COPS states*, only keep the marginal probability distribution of $\Pr(S_c^t, O_c^t)$. So, the new state space is the Cartesian product of the sets of S_c^t and O_c^t , which has size $2N$.

One subtask of COPS is that, given our state representation $\Pr(S_c^t, O_c^t)$, recover the distributions over the full state space (S_c^t, θ_d^t) . We cannot restore such distribution by multiplying $\Pr(\theta_d^t)$ and $\Pr(S_c^t)$ in COPS. This is because S_c^t , O_c^t , and θ_d^t are not independent. For example, in COPS state $S_c^t = 2$, $O_c^t = 1$, θ_d^t can only be 2. In other words, the defender's location distribution $\Pr(\theta_d^t | S_c^t, O_c^t)$ is no longer c^s . Instead, we approximate $\Pr(\theta_d^t | S_c^t, O_c^t)$ as follows. In each COPS state (S_c^t, O_c^t) , the defender's *estimated marginal distribution*, $\widehat{\Pr}(\theta_d^t | S_c^t, O_c^t)$, is found in a manner similar to that for the criminal's belief distribution c_b^t . Specifically, if $O_c^t = 1$, $\widehat{\Pr}(\theta_d^t | S_c^t, O_c^t) = (0, 0, \dots, 1, 0, \dots, 0)^T$, where the row representing station S_c^t is 1 and all others are 0; if $O_c^t = 0$, then $\widehat{\Pr}(\theta_d^t | S_c^t, O_c^t)$ is found through Eq. 2, but with the $c_b^s(j)$ replaced by $c^s(j)$. Next we recover the estimated distribution over (S_c^t, θ_d^t) , using $\widehat{\Pr}(S_c^t, \theta_d^t | S_c^t, O_c^t) = \widehat{\Pr}(\theta_d^t | S_c^t, O_c^t)$. Estimated distributions evolve the same way as exact distributions do, as described above. At the future strike, we can project the evolved estimated distribution to distributions over COPS states. Figure 4(b) shows the evolution of COPS states. However, such a process appear to involve T_s , negating the benefit of the factored representation; we avoid that by using a transition matrix $T_{COPS} \in \mathbb{R}^{2N \times 2N}$. Each element of T_{COPS} , i.e., transition probability $\Pr(S_c^{t'}, O_c^{t'} | S_c^t, O_c^t)$, can be calculated as follows:

$$\Pr(S_c^{t'}, O_c^{t'} | S_c^t, O_c^t) = \Pr(S_c^{t'} | S_c^t, O_c^t) \sum_{\theta_d^{t'}} \Pr(O_c^{t'} | S_c^{t'}, \theta_d^{t'}) \cdot \sum_{\theta_d^t} \Pr(\theta_d^{t'} | S_c^{t'}, S_c^t, \theta_d^t) \cdot \widehat{\Pr}(\theta_d^t | S_c^t, O_c^t), \quad (8)$$

where $\Pr(S_c^{t'} | S_c^t, O_c^t)$ and $\Pr(\theta_d^{t'} | S_c^{t'}, S_c^t, \theta_d^t)$ correspond to $p(j|i, c_b^{t_0})$ and $c^{t_0+dt}(n)$, respectively. Detailed derivation of Eq. 8 can be found in the online appendix.

The defenders' total expected utility in COPS is calculated in a similar way as the exact algorithm, which is

$$Obj = \mathbf{r}_{d,COPS} \cdot (I - (1 - \alpha)T_{COPS})^{-1} \mathbf{x}_{1,COPS}, \quad (9)$$

where $\mathbf{r}_{d,COPS}$, $\mathbf{x}_{1,COPS}$ are the expected utility vector and the initial distribution for COPS states. Similar to \mathbf{r}_d , $\mathbf{r}_{d,COPS}(S)$ is 0 if in S the defender is at the same station

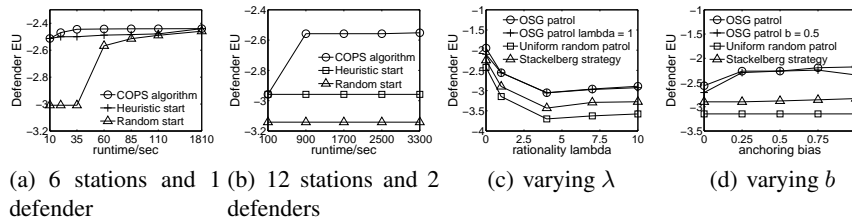


Fig. 5. (a) and (b) shows COPS vs. EOSG; (c) and (d) shows performance of different patrol strategies

with the criminal, else it is u_d . COPS is faster than EOSG because the number of states is reduced dramatically.

7 Heuristic Warm Start

Both the EOSG (9) and COPS are based on nonlinear optimization. We thus apply a local optimization solver. The initial configuration for the local optimization solver, which is the initial defenders' Markov strategy, affects the quality of the final solution obtained. Therefore, we propose a heuristic to generate good initial configurations. We divide the N station train line into some sublines – say two – without overlapping stations. The sub-line on the left has N_1 stations while the one on the right has N_2 , where $N_1 + N_2 = N$ and $|N_1 - N_2| \leq 1$. For the left sub-line, we add an imaginary station on the right end. The attractiveness of this imaginary station is the average attractiveness over the N_2 stations in the right sub-line. This imaginary station represents the effect of the right sub-line on this left sub-problem. The strategy generated by solving this $N_1 + 1$ station sub-problem is used as the initial configuration for the left N_1 stations; a similar procedure is repeated for N_2 . By generating small enough partitions, the warm start procedure can be run quickly.

8 Experiments

Settings: Motivated by metro system in Los Angeles and other cities, we deal with scenarios where stations lie along a straight line in our experiments. We solve the nonlinear optimization in OSG using the `FindMaximum` function in Mathematica. This function automatically chooses the best available local-optimization solver for the nonlinear problem instance. Possible solvers include Levenberg Marquardt, Newton, Quasi Newton, and Interior Point. Each data point we report is *an average of 30 different problem instances*, each based on a different attractiveness setting; these instances were generated through a uniform random distribution from 0 to 1 for each station. For multiple patrol unit scenarios, we use segment patrolling except for Figure 6, in which we show segment patrolling is better than full scale patrolling. The defender's utility of a successful crime was $u_d = -1$. In our experiments, the criminal's initial distribution was set to a uniform distribution over stations, while the defenders' initial distribution was the stationary distribution c^s , which can be computed from the transition matrix

T_d . We set the exit rate of the criminal $\alpha = 0.1$. All key results involving COPS are *statistically significant* ($\rho < 0.01$)

Results: First, we compare the performance for COPS against EOSG. Figures 5(a) and 5(b) show the performance of three different algorithms – the COPS algorithm with heuristic initial configuration, the exact algorithm with heuristic initial configuration, and the exact algorithm with random initial configuration – in two metro settings. In both, we set $\lambda = 1$. As stated previously, given our nonlinear problem, Mathematica uses locally optimal solvers, and there is always a current best feasible solution available although the quality of the solution keeps improving as the algorithm iterates over time. Therefore, one way to compare solutions is to provide a fixed run-time and check the level of solution quality reached in that time. In Figure 5(a), we show results of 6 stations with 1 defender. The x-axis in this figure shows the fixed runtime in seconds, which includes the runtime necessary to generate the warm start. The y-axis maps the defenders’ average expected utility against a single criminal. Unsurprisingly, the heuristic method generates significantly better initial configurations than using random initial configurations. The figure also shows that COPS outperforms the EOSG for any runtime within 100 s, even though COPS operates under an approximated version of the exact algorithm. This is because COPS reaches the local optima much faster than the exact algorithm. Further, even for runtime long enough to allow EOSG to reach its local optimum, where it outperforms COPS, the difference in utility between EOSG and COPS is less than 1%. Hence, by using COPS, we gain computational efficiency without a significant loss in solution quality. In Figure 5(b), we show results of 12 stations with 2 defenders. The conclusions are essentially the same as in Figure 5(a), but the advantage of COPS is even more obvious in this larger scale problem. As shown in Figure 5(b), in most instances of the 12-station problem, COPS reaches a local optimal strategy in 800 s. Meanwhile, the output strategies of EOSG are the same as their initial values even after 3200 s.

Figure 5(c) compares the performance of four different strategies against criminals with various levels of rationality. The first strategy is the uniform random strategy, which is a Markov strategy with equal probability for all available actions at each location; the second is a Strong Stackelberg equilibrium strategy, which is the optimal strategy to commit to against a strategic attacker that picks a single target to attack; the third is a COPS OSG strategy; the last is also an OSG strategy, but unlike the third strategy, the defenders do not know the real rationality of the criminals, and set a fixed $\lambda = 1$ during computation. $\lambda = 1$ for this last strategy was picked from a set of sampled λ so that the OSG strategy with this fixed λ performs best against criminals with various levels of rationality. The max runtime for each strategy is 1800 s, which is enough time for most algorithms to reach a local optimum. In this set of experiments, criminals have no anchoring bias ($b = 0$). Results with other b are similar and shown in the online appendix. The system consists of 12 stations and 2 defenders. The x-axis shows λ , the rationality level of the criminal in QBRM; $\lambda = 0$ means the criminal randomly picks the next target, and as λ increases, the criminal is more biased toward the station with higher expected utility.

As shown in Figure 5(c), the COPS OSG strategy outperforms the random and Stackelberg strategies significantly for any λ . Even though the OSG strategy assuming

$\lambda = 1$ performs slightly worse than that using the correct λ value, it is still better than other strategies. We conclude that OSG is a better model against opportunistic criminals even with an inaccurate estimation of λ .

The OSG strategy, the Stackelberg strategy, and the uniform random strategy are compared in Figure 5(d) against criminals with different levels of anchoring bias. The performance of the OSG strategy using an accurate value for anchoring bias and that using a fixed assumption of anchoring bias $b = 0.5$ are both shown. $b = 0.5$ is also picked from a set of sampled b so that the OSG strategy with this fixed b performs best against criminals with various levels of anchoring bias. In this set of experiments, λ is fixed to 1, but experiments with other λ generate similar results, and are shown in the online appendix. The x-axis maps the anchoring bias b and the y-axis maps the defender’s expected utility against a single criminal. Again, our COPS OSG strategy outperforms the uniform random and Stackelberg strategies even with a fixed b . Thus, OSG generates a better strategy even with an imprecise assumption of the criminal’s anchoring bias.

To show the scalability of COPS, we compare the performance of COPS with different numbers of defenders in metro systems with varying numbers of stations. 16 stations is very comparable to stations patrolled at the LA Metro. Five different configurations are compared in Figure 6: one defender, two defenders with full length patrolling, four defenders with full length patrolling, two defenders with segment patrolling, and four defenders with segment patrolling. The x-axis shows the number of stations and y-axis shows the average defender expected utility. The max runtime for each strategy is 1800 s. As expected, within the same patrol techniques, more defenders provide higher expected utility. But, with the same amount of resources, segment patrolling outperforms full length patrolling, again with significance $\rho < 0.01$ as with other results.

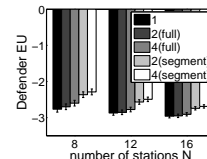


Fig. 6. Different number of defenders (varying N)

9 Summary and Related work

This paper introduces OSG, a new computational framework for opportunistic crime, opening the door to a new research area at the intersection of game theory and criminology. In OSG, this paper contributes (i) a new model of an opportunistic criminal; (ii) a new algorithm EOSG to compute defender patrol strategies; and (iii) an approximate algorithm, COPS, to speed up defender allocation to real-world scale scenarios. Given our experimental results, COPS is also deployed in the Los Angeles Metro systems [20].

Before reviewing SSG and criminology, we note that pursuit evasion games (PEGs) [21, 22] deal with similar problems in a graph setting. However, the pursuer’s goal is only to capture the evader and not to minimize evader’s influence in PEGs, which makes PEG not directly applicable in our scenario. Besides, the evader usually knows only the current position of the pursuer but not its strategy in most PEGs. In recent research

in PEGs, evaders are also equipped with perfect knowledge and rationality [23, 24]. However, neither of these assumptions can describe our opportunistic criminals.

Previously, the SSG framework has been successfully applied in security domains to generate randomized patrol strategies, e.g., for counter-terrorism and fare evasion checks on trains [25–27]. Recent work in SSG has begun to consider bounded rationality of adversaries [17] and incorporate some limited flexibility in adversary execution [28]. However, as discussed earlier, even though SSG can be extended to arbitrary topologies [2], SSG model remains mostly focused on highly strategic adversaries and fails to account for opportunistic criminals.

A second thread of previous research has modeled opportunistic criminal behavior, and security forces deployment against such adversaries. In [6] burglars are modeled as biased random walkers seeking “attractive” targets, and [29] follows up on this work with a method for generating effective police allocations to combat such criminals. However, these works make the extreme assumption that 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. Also, in [29] police behave in a similarly reactionary way, allocating their resources in an instantaneously optimal way in response to the current crime rather than optimizing over an extended time horizon, and there is also no notion of the “movement” of police — rather, police officers distribute instantaneously. Our current approach is an attempt to generalize these two threads of research.

References

1. Tambe, M.: *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press (2011)
2. Basilico, N., Gatti, N., Amigoni, F.: Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In: *AAMAS*. (2009)
3. Brantingham, P.J., Tita, G.: Offender mobility and crime pattern formation from first principles. *Artificial crime analysis systems: using computer simulations and geographic information systems* (2008) 193–208
4. Felson, M., Clarke, R.V.G.: *Opportunity makes the thief: Practical theory for crime prevention*. Volume 98. Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate (1998)
5. Ratcliffe, J.H.: A temporal constraint theory to explain opportunity-based spatial offending patterns. *Journal of Research in Crime and Delinquency* (2006)
6. Short, M.B., D’Orsogna, M.R., Pasour, V.B., Tita, G.E., Brantingham, P.J., Bertozzi, A.L., Chayes, L.B.: A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences* **18**(supp01) (2008) 1249–1267
7. Bernasco, W., Luykx, F.: Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighborhoods. *Criminology* (2003)
8. See, K.E., Fox, C.R., Rottenstreich, Y.S.: Between ignorance and truth: Partition dependence and learning in judgment under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **32**(6) (2006) 1385
9. Block, R., Block, C.R.: *The bronx and chicago: Street robbery in the environs of rapid transit stations* (2000)

10. Matthews, S.A., Yang, T.C., Hayslett, K.L., Ruback, R.B.: Built environment and property crime in seattle, 1998–2000: A bayesian analysis. *Environment and planning. A* **42**(6) (2010) 1403
11. Brantingham, P., Brantingham, P.: Criminality of place. *European Journal on Criminal Policy and Research* **3**(3) (1995) 5–26
12. Chaiken, J.M., Lawless, M., Stevenson, K.A.: The impact of police activity on crime. (1974)
13. Cornish, D.B., Clarke, R.V.: Understanding crime displacement: An application of rational choice theory. *Criminology* (1987)
14. Loukaitou-Sideris, A., Liggett, R., Iseki, H.: The geography of transit crime documentation and evaluation of crime incidence on and around the green line stations in los angeles. *JPER* (2002)
15. Sherman, L.W., Weisburd, D.: General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly* (1995)
16. Perron, O.: Zur theorie der matrices. *Mathematische Annalen* (1907)
17. Pita, J., Jain, M., Ordóñez, F., Tambe, M., Kraus, S., Magori-Cohen, R.: Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In: *AAMAS*. (2009)
18. McKelvey, R.D., Palfrey, T.R.: Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* **10**(1) (1995) 6–38
19. Boyen, X., Koller, D.: Tractable inference for complex stochastic processes. In: *UAI*, Morgan Kaufmann Publishers Inc. (1998) 33–42
20. : Reference under review. In: *AAMAS*. (2014)
21. Isler, V., Kannan, S., Khanna, S.: Randomized pursuit-evasion in a polygonal environment. *Robotics, IEEE Transactions on* (2005)
22. Vidal, R., Shakernia, O., Kim, H.J., Shim, D.H., Sastry, S.: Probabilistic pursuit-evasion games: theory, implementation, and experimental evaluation. *Robotics and Automation, IEEE Transactions on* (2002)
23. LaValle, S.M., Hinrichsen, J.E.: Visibility-based pursuit-evasion: The case of curved environments. *Robotics and Automation, IEEE Transactions on* (2001)
24. Hespanha, J.P., Prandini, M., Sastry, S.: Probabilistic pursuit-evasion games: A one-step nash approach. (2000)
25. Yin, Z., Jiang, A., Johnson, M., Tambe, M., Kiekintveld, C., Leyton-Brown, K., Sandholm, T., Sullivan, J.: Trusts: Scheduling randomized patrols for fare inspection in transit systems. In: *IAAI*. (2012)
26. Jiang, A.X., Yin, Z., Zhang, C., Tambe, M., Kraus, S.: Game-theoretic randomization for security patrolling with dynamic execution uncertainty. In: *AAMAS*. (2013)
27. Varakantham, P., Lau, H.C., Yuan, Z.: Scalable randomized patrolling for securing rapid transit networks. In: *IAAI*. (2013)
28. Vorobeychik, Y., Singh, S.P.: Computing stackelberg equilibria in discounted stochastic games. In: *AAAI*. (2012)
29. Zipkin, J.R., Short, M.B., Bertozzi, A.L.: Cops on the dots in a mathematical model of urban crime and police response. (2013)