Camera Placement Based on Vehicle Traffic for Better City Security Surveillance

Xiaobo Ma
Xi'an Jiaotong University

Yihui He
Xi'an Jiaotong University

Xiapu Luo
The Hong Kong Polytechnic University

Jianfeng Li
Xi'an Jiaotong University

Mengchen Zhao
Nanyang Technological University

Bo An
Nanyang Technological University

Xiaohong Guan
Xi'an Jiaotong University

Security surveillance is important in smart cities. Deploying numerous cameras is a common approach. Given the importance of vehicles in a metropolis, using vehicle traffic patterns to strategically place cameras could potentially facilitate security surveillance. This article constitutes the first effort toward building the link between vehicle traffic and camera placement for better security surveillance.

Carefully conducted surveillance has become a widely-used tactic in city-wide security efforts. Due to the continuous growth of cities in size and complexity, maintaining safety becomes critical to attracting skilled people and economic investment.

To counter the adversary, deploying a number of video cameras is a common approach.[1] At first glance, camera placement for adequate coverage could be achieved by purchasing numerous cameras. However, subject to any realistic budget, cameras cannot be placed unlimitedly. Additionally, as the camera resolution increases, the overhead of transmitting and processing videos would be considerably large. For example, a latest IP camera may require up to 10 Mbps bandwidth.[2]

Using the information gleaned from vehicle traffic patterns to strategically place cameras could potentially facilitate security surveillance. For example, criminals may drive vehicles before and after security breaches occur, or even use vehicles as weapons like in the 2017 Westminster attack.[3] Thus, monitoring vehicles on the roads has great potential in collecting evidence in criminal investigations.

Published by the IEEE Computer Society
1541-1672/18/$33.00 ©2018 IEEE

A straightforward strategy is to place cameras in busy regions that would naturally have more vehicle traffic, where people tend to commit crime or where terrorist attacks may be more likely to occur. However, this strategy may fail to cover vehicles that never enter busy regions. A good strategy would be to have as many individual vehicles recorded by surveillance cameras as possible and to record each vehicle as frequently as possible.

After setting up the camera infrastructure, the next requirement is that the cameras can provide high-resolution images for high-quality evidence. However, because of the overhead barrier, the number of cameras that could be simultaneously configured as high-resolution is limited. Determining which cameras should be configured as high-resolution is non-trivial. If it is a fixed set of cameras, criminals may use that knowledge to evade surveillance.

This article investigates security camera placement informed by vehicle traffic, and formulates it as a group of submodular function optimization problems solvable by the greedy algorithm with theoretical lower bounds. We propose five placement strategies with different goals. Furthermore, considering the video transmitting and processing overhead, we design a game-theoretic framework for randomized camera resolution upgrading with maximal utility to deter a sophisticated adversary.

To the best of our knowledge, we are the first to build the connection between vehicle traffic and camera placement for principled ways of security surveillance.

PLACEMENT STRATEGIES FOR BUILDING SURVEILLANCE INFRASTRUCTURE

Problem Description

Let V be the set of vehicles and $v \in V$ denote a vehicle. A record from a vehicle's motion sensor comprises: *vehicle ID, time, latitude, longitude*, and the *vehicle ID* uniquely identifies a vehicle.

We consider the metropolis as a rectangle characterized by maximum latitude, minimum latitude, maximum longitude, and minimum longitude. The rectangle is divided into small-sized $l \times l$ (e.g., $50m \times 50m$) blocks, resulting in a block set denoted by Ω. The advantages of dividing the map into blocks include map-independence, so it would be applicable for all metropolises, and exact coverage of the complete map.

Our goal is to find a subset of blocks $C \subseteq \Omega$ to place cameras maximizing security surveillance utility. Let N denote the maximum number of blocks where we can afford to place cameras. Obviously, we have $|C| \leq N$. Note that only blocks with GPS records can be selected for camera placement, hence covering at least one road.

Placement Strategies

We propose five placement strategies facilitating security surveillance in different aspects. The motivation is as follows. First, in real-life security surveillance, it is best to observe as many distinct vehicles as possible and as much vehicle traffic as possible. Thus, maximum unique vehicles (i.e., *S*1) and maximum vehicle traffic (i.e., *S*2) are two immediate functions to optimize concerning overall statistics of all vehicles. Second, for any given vehicle, we expect it can spend more time under surveillance, and hit unique cameras more frequently along its trajectory. Accordingly, we propose metrics *S*3, *S*4 and *S*5 as detailed below.

For all strategies, we denote the objective function by $F(C)$.

S1—Maximum Unique Vehicles. It selects a subset $C \subseteq \Omega$ to maximize the number of unique vehicles crossing one or more blocks in *C*, formally expressed as

$$\arg\max_{C} F(C) = \sum_{v \in V} I_v(C)$$

where $I_v(C)$ equals one if v crossed at least one block in C; otherwise zero. *S1* maximizes the total number of unique vehicles.

S2—Maximum Vehicle Traffic. *S2* maximizes the amount of vehicle traffic, rather than unique vehicles, crossing blocks in C. It equals

$$\arg\max_{C} F(C) = \sum_{v \in V} T_v(C)$$

where $T_v(C)$ denotes the amount of traffic of blocks in C contributed by v, and could be measured by the total time when v stays inside C. $T_v(C)$ can be calculated by $\sum_{c \in C} T_v(c)$, where $T_v(c)$ denotes the total time when v stays in c, depending on the times when v enters and leaves c.

S3—Minimum Mean OITR (Out-Camera to In-Camera Time Ratio). *S3* minimizes the mean OITR across all vehicles. The OITR of a vehicle represents the proportion of time when a vehicle is out of surveillance. Intuitively, smaller mean OITR across all vehicles indicates better security surveillance. We express S3 as

$$\arg\max_{C} F(C) = \Phi - \sum_{v \in V} \left(\frac{S}{T_v(C)+1} - 1\right) / |V|$$

where S is the measurement time. *S3* does not consider the uniqueness of cameras. Therefore, it encourages more unique vehicles visible to (whichever) cameras, and meanwhile each vehicle visible to cameras as long as possible. *Note that $T_v(C)$ is increased by one to avoid zero denominator, and Φ is a constant derived from $F(\varnothing) = 0$ so that $F(C) \geq 0$. The same operation exists for S4 and S5.*

S4—Minimum Mean ACIs (Average Camera-Hit Intervals). We define the ACIs to represent the average time interval to hit a camera for a vehicle. Formally, we minimize mean ACIs across all vehicles, and *S4* is

$$\arg\max_{C} F(C) = \Phi - \sum_{v \in V} \left(\frac{S}{H_v(C)+1}\right) / |V|$$

where $H_v(C)$ denotes the number of times that v hits (whichever) cameras along its trajectory during measurement time S. *S4* encourages more unique vehicles visible to (whichever) cameras, and meanwhile each vehicle to hit cameras more frequently.

S5—Minimum Mean AUIs (Average Unique-Camera Hit Intervals). We further adapt *S4* by considering the uniqueness of cameras that a vehicle hits. Accordingly, we define the AUIs to represent the average time interval to hit a new camera for a vehicle. We minimize mean AUIs across all vehicles, and *S5* is

$$\arg\max_{C} F(C) = \Phi - \sum_{v \in V} \left(\frac{S}{U_v(C)+1}\right) / |V|$$

where $U_v(C)$ denotes the number of unique cameras that v hits during measurement time S. $U_v(C)$ is the number of unique blocks in C that v crosses. *S5* encourages more unique vehicles visible to cameras, and meanwhile each vehicle hits more new cameras.

Solving Optimal Placement Strategies

All strategies are formulated as maximization problems. To find an optimal subset C, we need to solve these problems, which are NP-hard. All these maximization problems could be solved with a greedy algorithm (GA) because of their non-increasing monotony and submodularity.[4-5] The non-increasing monotony means that, for any two sets C_1, $C_2 \subseteq \Omega$ and $C_1 \subseteq C_2$, we have $F(C_1) \leq F(C_2)$. Apparently, the functions defined in *S1~S5* are non-decreasing.

The submodularity means that a non-decreasing set function has the property of diminishing returns when a new element c is added to an input set C, as compared to c is added to an input set that is a subset of C. Specifically, for a non-decreasing function, such as $F(C)$ in *S1~S5*, given any two sets C_1, $C_2 \subseteq \Omega$ and $C_1 \subseteq C_2$, and a new block $c \in \Omega \backslash C_2$, the submodularity refers to $F(C_1 \cup \{c\}) - F(C_1) \geq F(C_2 \cup \{c\}) - F(C_2)$, i.e., smaller sets have more function value increment when added with a new block. Interested readers are referred to Krause and Golovin[5] for a comprehensive survey of submodularity. It is easy to prove the submodularity for *S1* and *S2*. The proof of the submodularity for *S3~S5* is given below.

For $F(C)$ in S3, given any two sets $C_1, C_2 \subseteq \Omega$ and $C_1 \subseteq C_2$, we have

$$F(C_*) = \Phi - \sum_{v \in V} \left(\frac{S}{T_v(C_*)+1} - 1\right) / |V|.$$

Note C_* means the equation holds for both C_1 and C_2. We add a new block $c \in \Omega \backslash C_2$ to C_1 and C_2, and obtain

$$F(C_* \cup \{c\}) = \Phi - \sum_{v \in V} \left(\frac{S}{T_v(C_* \cup \{c\})+1} - 1\right) / |V|.$$

We derive function value increment as

$$F(C_* \cup \{c\}) - F(C_*) = \sum_{v \in V} \left(\frac{S}{T_v(C_*)+1} - \frac{S}{T_v(C_* \cup \{c\})+1}\right) / |V|.$$

Recall the definition of $T_v(C)$. Because of $c \notin C_1$ and $c \notin C_2$, we have $T_v(C_1 \cup \{c\}) = T_v(C_1) + T_v(\{c\})$, and $T_v(C_2 \cup \{c\}) = T_v(C_2) + T_v(\{c\})$. Therefore, we derive

$$F(C_* \cup \{c\}) - F(C_*) = \sum_{v \in V} \left(\frac{S}{T_v(C_*)+1} - \frac{S}{T_v(C_*)+1+T_v(\{c\})}\right) / |V|.$$

Given $C_1 \subseteq C_2$, we have $T_v(C_1) \leq T_v(C_2)$. Thus,

$$\frac{S}{T_v(C_1)+1} - \frac{S}{T_v(C_1)+1+T_v(\{c\})} \geq \frac{S}{T_v(C_2)+1} - \frac{S}{T_v(C_2)+1+T_v(\{c\})}.$$

Finally, we have

$$F(C_1 \cup \{c\}) - F(C_1) \geq F(C_2 \cup \{c\}) - F(C_2).$$

The submodularity holds for $F(C)$ in *S3*. Similarly, we can deduce the submodularity for $F(C)$ in *S4* and *S5*.

Submodularity allows us to derive a solution lower bounded by $1-1/e \approx 63$ percent of the optimal solution.[4] GA runs for at most N rounds to obtain a set C of size $|C| \leq N$. In each round, it selects a new block $c \in \Omega \backslash C$ maximizing the reward gain, $\delta_c(C) = F(C \cup c) - F(C)$, and inserts c into C. This process repeats until $|C| = N$ or $\delta_c(C) = 0$.

RANDOMIZED CAMERA RESOLUTION UPGRADING FOR HIGH-QUALITY SURVEILLANCE

We have designed camera deployment strategies, given the budget of the maximum number of cameras. The aim of these strategies is to deploy camera infrastructure, offering basic surveillance of vehicle movement (e.g., plate number, trajectories). However, in real-life, high-quality crime evidence may be needed to identify fine-grained information (e.g., criminals' faces), entailing the deployment of high-resolution cameras.

Modern cameras could be remotely configured as different resolutions. Nevertheless, if the number of deployed cameras is pretty large, it is unlikely to configure all cameras as high-resolution due to the intractable overhead of handling videos (e.g., transmission, and processing). That is, only a limited number of cameras could be of high-resolution simultaneously.

Next, we study which cameras should be configured as high-resolution among all the deployed cameras. If a fixed set of cameras are configured as high-resolution, the adversary may evade them deliberately in the long run. To avoid being evaded, we propose to randomly choose a set of placed cameras and upgrade their resolution in a principled way, for deterring the adversary with the fact that any camera might be of high-resolution.

Note that the high-resolution of cameras are not a parameter in the functions to optimize in the previous section. The reason is that those objective functions are designed for deploying camera infrastructure, offering basic surveillance of vehicle movement without the need to consider camera resolutions. We assume that all deployed cameras could be remotely configured as high-resolution. In this section, the high-resolution of cameras would be taken into account by solving the probabilities to set cameras as high-resolution under a game-theoretic model.

A Game-theoretic Formulation

Since blocks differ in their surveillance priorities, we consider the block importance when designing the randomized strategy. Meanwhile, the adversary aims to evade high-resolution cameras, while the defender tries to observe the adversary.

To describe such confrontation, we leverage the Stackelberg security game (SSG).[6] A standard SSG has two players, a leader and a follower. Each player has their own set of pure strategies to select. The players act sequentially as follows.

Step 1. The leader (i.e., defender) commits to a mixed strategy, i.e., playing a probability distribution over pure strategies, maximizing her utility.

Step 2. After learning the mixed strategy chosen by the leader, the follower (i.e., adversary), as a response, selects a pure strategy, optimizing his utility.

We consider a threat model wherein the adversary is *sophisticated*. Specifically, the adversary can learn the defender's mixed strategy (i.e., the probability distribution), and select the pure strategy maximizing his utility (i.e., a best response adversary).

Meanwhile, the defender is forward-looking. That is, she takes into account the adversary's threat model when designing strategies, thereby making her strategy robust against the sophisticated adversary.

Although the adversary can learn the defender's mixed strategy, he cannot predict which specific pure strategy the defender would adopt at the time of scheduled criminal activities.

Player Strategies

A pure strategy of the defender is a set of cameras whose resolution can be upgraded simultaneously. Our aim is to deter the adversary by randomly upgrading the resolution of the placed cameras. Thus, the adversary committing crimes in blocks without cameras is not considered, and a pure strategy of the adversary is a set of blocks with placed cameras.

Utility Functions

Consider that the adversary commits crimes in the ith block c_i. If c_i is covered by the defender's pure strategy, the defender receives reward R_i^d; otherwise penalty P_i^d. Similarly, the adversary receives penalty P_i^a in the former case, and reward R_i^a in the latter case. The reward and the penalty can be assigned according to the block importance, such as the amount of vehicle traffic, the number of unique vehicles, and historical crime activity severity, and so forth.

Let Γ_j denote the jth defender pure strategy, and A_{ij} denote the coverage indicator of Γ_j on c_i, where $A_{ij} = 1$ for $c_i \in \Gamma_j$, and $A_{ij} = 0$ for $c_i \notin \Gamma_j$. Let J be the number of defender pure strategies.

The number of adversary pure strategies is N, the (maximum) number of blocks where cameras are placed. We denote the probability of the defender choosing Γ_j by a_j, and

$$\sum_{j=1}^{J} a_j = 1$$
.

The marginal probability x_i for the defender to upgrade c_i (i.e., upgrade the resolution of the camera in c_i) is

$$x_i = \sum_{j=1}^{J} \mathrm{a}_j A_{ij}, \quad i = 1, 2, ..., N$$
.

We denote $(a_1,a_2,...,a_J)$ by **a**, and $(x_1,x_2,...,x_N)$ by **x**, where **x** is determined by **a**.

We express the defender's expected utility on upgrading c_i as

$$U_i^d(x_i) = x_i R_i^d + (1-x_i) P_i^d,$$

and the adversary's expected utility on committing crimes in c_i is

$$U_i^a(x_i) = x_i P_i^a + (1-x_i) R_i^a .$$

Mixed Strategy Against a Best Response Adversary

A sophisticated adversary takes the best response to commit crimes to maximize his utility. Formally, the probability that he selects c_i equals

$$B_i = \begin{cases} 1 & U_i^a(x_i) \geq U_j^a(x_j), \ \forall j = 1,...,J \\ 0 & \text{otherwize} \end{cases}$$
.

This means that the adversary *knows* the marginal probability x_i for the defender to upgrade c_i, and he selects the target with maximal expected utility. The utility functions of the adversary and the defender are

$$U^a = \sum_{i=1}^{N} B_i U_i^a(x_i) ,$$

$$U^d = \sum_{i=1}^{N} B_i U_i^d(x_i) .$$

Simultaneously, the defender selects an optimal mixed (i.e., randomized) strategy in consideration of the sophisticated adversary's best response. The defender maximizes her utility U^d as

$$\max_{\mathrm{a}} \sum_{i=1}^{N} B_i U_i^d(x_i) .$$

Substituting $U_i^d(x_i) = x_i R_i^d + (1-x_i) P_i^d$, we rewrite the above equation as

$$\max_{\mathrm{a}} \sum_{i=1}^{N} B_i (x_i R_i^d + (1-x_i) P_i^d) .$$

To calculate the defender's optimal strategy, problem **P1** needs to be solved.

$$\text{P1:} \begin{cases} \max_{\mathrm{a}} \sum_{i=1}^{N} B_i (x_i R_i^d + (1-x_i) P_i^d) \\ s.t. \qquad x_i = \sum_{j=1}^{J} \mathrm{a}_j A_{ij}, \forall i \\ \qquad\qquad \sum_{j=1}^{J} a_j = 1 \\ \qquad\qquad 0 \leq a_j \leq 1, \forall j \end{cases}$$
.

P1 can be solved by the branch-and-cut algorithm.[7] The defender finally adopts the mixed strategy below.

Mixed Strategy: play Γ_j with probability a_j
where $a_j \in \mathbf{a}$ is the solution of problem **P1**, $j = 1, 2, \ldots, J$

EXPERIMENTAL EVALUATION

Dataset

The data contains one-week GPS trajectories of 10,357 taxis in Beijing.[8] We remove the outlier GPS points, including those indicating an impossible speed and significantly deviating the moving average.

We then divide Beijing into $50m \times 50m$ blocks Ω, yielding the total number of such blocks $|\Omega| = 14{,}473{,}089$. We can place cameras in 438,674 blocks where vehicles arrive, denoted as $R \subset \Omega$.

Performance of Camera Placement Strategies

Metrics

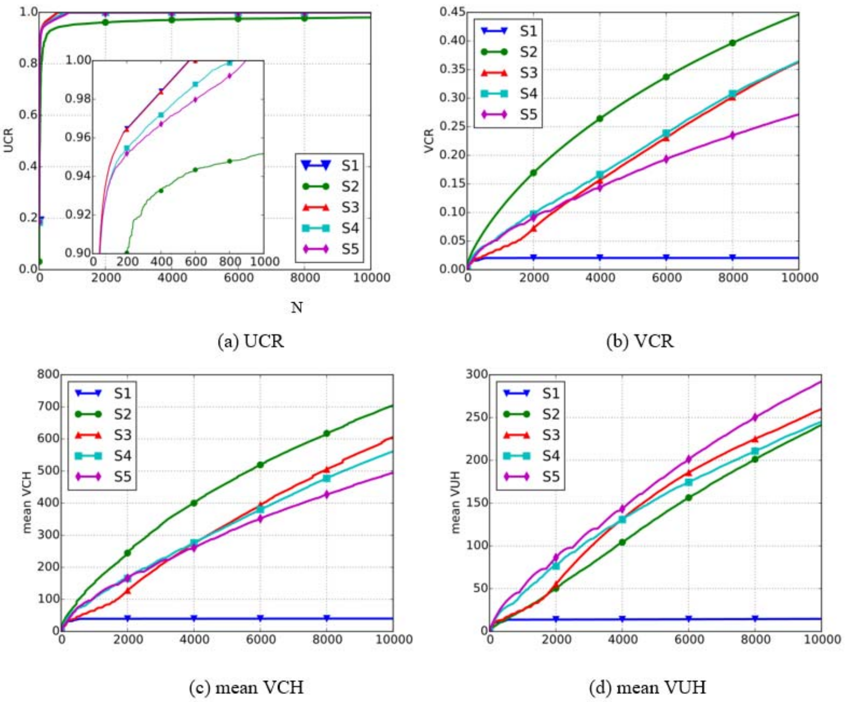
Larger values of these metrics indicate better security surveillance.

- UCR (Unique Vehicle Coverage Ratio). The ratio of observed unique vehicles to all vehicles, calculated by $\sum_{v\in V} I_v(C) / \sum_{v\in V} I_v(R)$.
- VCR (Vehicle Traffic Coverage Ratio). The ratio of traffic observed by all cameras to the total amount of traffic, which equals $\sum_{v\in V} T_v(C) / \sum_{v\in V} T_v(R)$.
- VIT (Vehicle In-Camera Time). The total amount of time when a vehicle v is under surveillance, i.e., $T_v(C)$
- VCH (Vehicle Camera-Hits). The number of times that v hits cameras along its trajectory, calculated as $\sum_{c\in C} I_v^{'}(\{c\})$, where $I_v^{'}(\{c\})$ is the number of times v hits $c \in C$.
- VUH (Vehicle Unique-Camera-Hits). The number of unique cameras v hits, calculated as $\sum_{c\in C} I_v(\{c\})$, where $I_v(\{c\})$ equals one if v hits c; otherwise zero.

Results

For each strategy, we calculate UCR, VCR, and VIT by varying the number of cameras N from 1 to 10,000. UCR and VCR reveal the proportion of unique vehicles and vehicle traffic covered under each strategy. Additionally, we calculate VIT, VCH, and VUH for each vehicle to have a closer look at the total amount time when a vehicle is under surveillance, the total number of cameras a vehicle hits, and the total number of unique cameras a vehicle hits, respectively.

Figure 1 depicts the performance of all strategies. Generally, the metrics increase slower as N becomes larger, meaning all strategies accomplish diminishing returns. Particularly, all metrics rise rapidly before N reaches 200, accounting for no more than 0.05 percent of all possible blocks where we can place cameras (i.e., R).



(a) UCR

(b) VCR

(c) mean VCH

(d) mean VUH

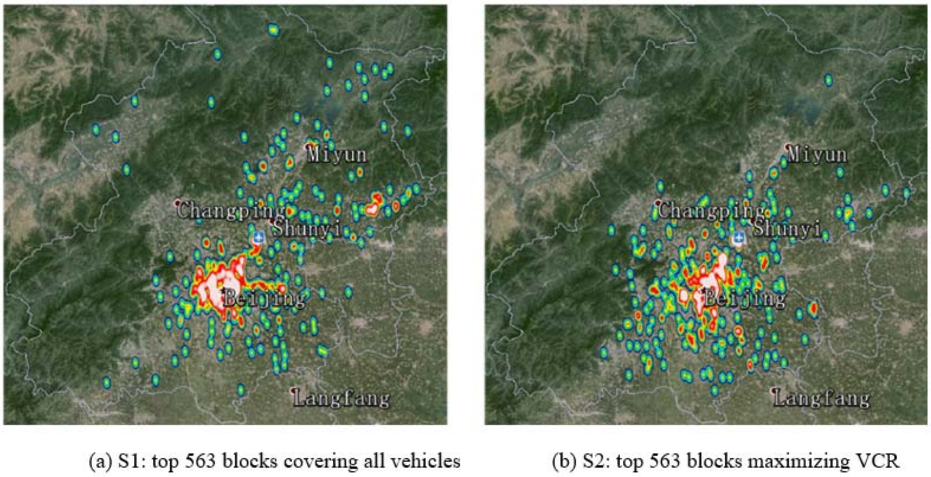
Figure 1. The performance metrics over the number of cameras *N* of different placement strategies.

As Figure 1(a) shows, a very small proportion of cameras need to be placed to cover the vast majority of vehicles. Specifically, all strategies exhibit a rapid rise in UCR to at least 90 percent before *N* reaches 200, indicating that we need to place cameras in no more than 0.05 percent of blocks to cover at least 90 percent of vehicles. Moreover, all strategies except *S2* achieve a 100 percent UCR before *N* reaches 900 (i.e., 0.2 percent of *R*), while *S2* almost has no increasing returns after *N* = 4,000.

Insight 1. *To cover all vehicles, we need to place cameras only in 0.2 percent of blocks at most using S1, S3, S4, S5. Among these strategies, S1 could cover all vehicles with the smallest N, followed by S3, S4, and S5 in ascending order. However, S2 could hardly converge to cover all vehicles as N increases.*

This insight reveals the existence of a small set of "core" blocks covering all vehicles. Figure 2(a) shows the minimum set of such "core" blocks derived from *S1*.

Figure 1(b) demonstrates the vehicle traffic coverage ratio, where VCR rises roughly linearly as *N* increases for all strategies except *S1*. Among all strategies, *S2* exhibits the largest growth rate, followed by *S4*, *S3*, *S5*, and *S1*. Figure 2(b) shows top 563 blocks where we can place cameras to maximize VCR using *S2*. We observe that the blocks in Figure 2b are significantly less geographically dispersed than those in Figure 2(a). The growth rates of *S3* and *S4* are comparable. *S1* almost has no increasing returns after *N* = 200, implying that emphasizing vehicle coverage may fail to maximize traffic coverage.



(a) S1: top 563 blocks covering all vehicles

(b) S2: top 563 blocks maximizing VCR

Figure 2. Camera placement heatmaps derived using one-week taxi data in Beijing based on strategies S1 and S2.

Insight 2. *Although S1 achieves the best UCR, its performance regarding VCR degrades drastically to an extent where no increasing return is accomplished after N = 200. In contrast, S3~S5 achieve significant increasing returns regarding VCR as N increases. Their increasing returns, although not as fast as that of S2, can converge to cover all vehicles quickly but S2 cannot.*

This insight reveals that *S3~S5* can achieve well-balanced performance between UCR and VCR, because they encourage more vehicles under surveillance while considering more (unique) camera hits and traffic coverage per vehicle.

Figure 1(c) presents mean VCH across all vehicles, which has similar trend as VCR. This reflects that the number of camera hits is positively correlated to in-camera time. For the same reason, VIT exhibits the same trend as VCH. We therefore omit the plot of mean VIT. Among all strategies, *S2* achieves the best average surveillance performance (i.e., longer in-camera time, more camera hits) per vehicle, though it may result in slow or even incomplete vehicle coverage. Despite the performance degradation in mean VCH compared to *S2*, *S3~S5* achieve median values of VIT and VCH across all vehicles comparable to *S2*, while simultaneously exhibiting much smaller standard deviation of VIT and VCH. Furthermore, the Gini coefficient of VCH for *S2* is larger than that for *S3~S5*, confirming that *S3~S5* achieve relatively fairer surveillance across all vehicles regarding in-camera time and camera hits than *S2*.

Insight 3. *S3~S5 achieve more balanced surveillance across all vehicles regarding in-camera time and camera hits than S2, avoiding little surveillance of some vehicles and too much surveillance (possibly unnecessary) of others. Therefore, they provide relatively fairer surveillance services across all vehicles.*

Figure 2(d) presents the performance of mean VUH (i.e., the number of unique vehicles that a vehicle hits) across all vehicles. We observe that *S5* achieves the largest mean VUH, while *S1* has no increasing returns after *N* reaches 200. The remaining strategies exhibit moderate performance worse than that of *S5*. Actually, *S5* encourages deploying cameras in the blocks where more different blocks along a vehicle's trajectory can be observed, and meanwhile more vehicles. In other words, *S5* can capture more places where a vehicle has been.

Insight 4. *S5 is a strategy in favor of capturing more places where a vehicle has ever been. Other strategies perform worse than S5 regarding mean VUH.*

A Case Study of Randomized Resolution Upgrading

We consider the surveillance scenario where *N* = 10,000 cameras are placed while only 1,000 cameras could be configured as high-resolution. Suppose the surveillance focuses on criminal activities that tend to be conducted at places with more vehicle traffic. We thus measure the importance of a block c_i according to its amount of traffic, i.e., $T(c_i) = \sum_{v \in V} T_v(\{c_i\})$.

We consider that the defender and the adversary play a zero-sum game, i.e., each player's gain or loss of utility is exactly balanced by the losses or gains of the utility of the other player. Formally, we set $R_i^d + P_i^a = 0$, $R_i^a + P_i^d = 0$. Obviously, we have $U^d + U^a = 0$. Meanwhile, the importance values of blocks are normalized through dividing each value by the maximum value. Therefore, we have R_i^d, $R_i^a \in [0,1]$, and P_i^d, $P_i^a \in [-1,0]$. It is easy to derive $U^a, U^d \in [-1,1]$.

We then randomly generate 1,000 defender pure strategies in total. The union of these pure strategies covers all the placed cameras to ensure that the resolution of any camera has a chance to be upgraded. Moreover, there are 10,000 adversary pure strategies, each of which is a block where crimes are committed.

Table 1 shows the defender utility (i.e., U_d) when the defender adopts the mixed strategy, in comparison to two baseline defender strategies, namely, uniform strategy and best strategy. The uniform strategy means the defender adopts a pure strategy uniformly at random. The best strategy allows the defender to constantly adopt the pure strategy that contains the most important block. We see that the mixed strategy achieves significantly larger utility than baseline strategies, no matter which strategy (S1~S5) is employed to place the camera infrastructure. This indicates that, once the importance of each block is defined, one can derive the mixed strategy that outperforms baseline strategies and maximizes the defender utility against a sophisticated adversary. Thus, it is feasible and effective to strategically randomize camera resolution upgrading for high-quality evidence of criminal activity.

Table 1. The Defender's Expected Utility Against a Best-Response Adversary Under Different Strategies and Different Camera Placement Strategies

defender placement utility	mixed strategy	baseline strategies	
		uniform strategy	best strategy
S1	-0.07313	-0.74600	-0.19582
S2	-0.13279	-0.81400	-0.50654
S3	-0.10210	-0.77800	-0.47290
S4	-0.20533	-0.81000	-0.68853
S5	-0.12091	-0.82800	-0.49978

DISCUSSION

Some roads may belong to different blocks. However, this does not affect the correctness of the final result derived using our method. The reason is that our method could be aware of the overlapping vehicle traffic that cameras in different blocks observe. For example, if one road belongs to two blocks, to maximize vehicle coverage, only one block needs to be selected for camera placement; otherwise two blocks would observe the same vehicles, thereby failing to maximize vehicle coverage.

Different types of vehicles may differ in surveillance priorities. For example, compared with individual cars, taxis and school buses need higher priorities. Every vehicle v can be assigned a weight w_v indicating its surveillance priority. One can also assign such weights according to a vehicle's reputation from accident records, a vehicle owner's criminal records, etc. The weight w_v can be directly used as a multiplier of the parameters $I_v(C)$, $T_v(C)$, $H_v(C)$, and $U_v(C)$ in the proposed strategies, without altering submodularity of $F(C)$. Incorporating weights allows us to customize placement strategies with a bias towards vehicles with higher priorities.

Customizing security surveillance not only anticipates GPS data sharing from vehicles, but also relies on social data sources helpful to surveillance. The data sharing may introduce privacy issues, which could be resolved via data obfuscation. Whenever a security event occurs, the obfuscated data (e.g., vehicles IDs) could again be recovered from recorded videos using plate number recognition techniques, after being authorized. Tracking individual vehicles by camera may also have privacy issues. Thus, we suggest the public should be notified of placed cameras, and the collected videos must be strictly managed according to law.

Privacy issues may exist for people (i.e., non-criminals) who object to being under surveillance. If the video data were leaked or misused, a sophisticated adversary may track people's trajectories across multiple cameras using object identification/tracking techniques to infer home and work locations. To address these issues, comprehensive measures in management, law enforcement, and computer vision techniques should be combined and implemented in conjunction with the surveillance efforts. Interested readers are referred to Rajpoot and Jensen[9] for detailed measures. Despite the privacy risks, video surveillance is indispensable in combating security threats. It must be used in ways that protect people against crime, without compromising privacy.

RELATED WORK

Transportation monitoring includes tasks like road traffic monitoring, urban environment modeling.[10] It undoubtedly accelerates the operation efficiency of cities. However, none of the existing studies uses vehicle traffic to optimize camera placement for better security surveillance.

There is a rising trend towards game-theoretic security patrolling on roads.[11] A defender can schedule checkpoints on roads to detect adversaries. Different from these studies, we leverage game theory to control placed cameras. Moreover, camera placement complements road patrolling, because the adversaries missed by road checkpoints could be further inspected by cameras. Deploying cameras in each selected block is another problem particularly studied by the computer vision community. Typical purposes include performing accurate three-dimensional reconstruction, detecting vehicle violations,[12] and tracking and recognizing human motions and activities.[13] These techniques consider camera specifications (e.g., visual distance, orientation, Field of View) and road maps in a small area (e.g., office, road intersections).[14] All these techniques can be borrowed to place cameras in each selected block.

CONCLUSION

This article explored the connection between vehicle traffic and camera placement for principled ways of security surveillance. We proposed five camera placement strategies that are submodular and solvable with theoretical bounds. Using real-world data, we demonstrated that the proposed strategies could facilitate security surveillance in different aspects. We also studied the problem that high-resolution video is desired for high-quality evidence of criminal activity, while only a limited number of placed cameras can be configured as high-resolution. The results illustrated that, using the game-theoretic randomized strategy, we can deter a sophisticated adversary with maximal utility. In an era of terrorism, we expect our work could influence the decision making of surveillance camera placement in smart cities.

REFERENCES

1. A. Stutzer and M. Zehnder, "Is camera surveillance an effective measure of counterterrorism?," *Defence and Peace Economics*, 2013.
2. "Bandwidth and storage calculator,"; http://www.stardot.com/ bandwidth-and-storage-calculator.
3. "2017 Westminster attack,"; https://en.wikipedia.org/wiki/2017 Westminster attack.
4. G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Mathematical Programming*, 1978.

5. A. Krause and D. Golovin, “Submodular Function Maximization,” *Tractability: Practical Approaches to Hard Problems*, Cambridge University Press, 2014.
6. M.H. Manshaei et al., “Game theory meets network security and privacy,” *Game theory meets network security and privacy*, 2013.
7. D. Gavalas et al., “A survey on algorithmic approaches for solving tourist trip design problems,” *Journal of Heuristics*, 2014.
8. “T-Drive data,”; http://research.microsoft.com/apps/pubs/?id=152883.
9. Q.M. Rajpoot and C. D. Jensen, “Video surveillance: Privacy issues and legal compliance,” *Promoting Social Change and Democracy through Information Technology*, IGI Global, 2015.
10. Y. Zheng et al., “Urban computing: Concepts, methodologies, and applications,” *ACM Transaction on Intelligent Systems and Technology*, 2014.
11. B. An, M. Tambe, and A. Sinha, “Stackelberg security games (ssg) basics and application overview,” *Improving Homeland Security Decisions*, Cambridge University Press, 2016.
12. R. Marikhu et al., “Police eyes: Real world automated detection of traffic violations,” *10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2013.
13. P. Rahimian and J. K. Kearney, “Optimal camera placement for motion capture systems in the presence of dynamic occlusion,” *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, 2015.
14. A. Sforza, S. Starita, and C. Sterle, “Optimal location of security devices,” *Railway Infrastructure Security*, Springer, 2015.

ACKNOWLEDGMENTS

This work is supported in part by National Natural Science Foundation (61602371), Natural Science Basic Research Plan in Shaanxi Province (2016JQ6034), Postdoctoral Science Foundation (2015M582663), Shaanxi Postdoctoral Science Research Project, Fundamental Research Funds for the Central Universities, of China. Xiapu Luo is supported in part by the Hong Kong ITF (No. ITS/197/17FP), the National Natural Science Foundation of China (No. 61471228), the Key Project of Guangdong Province Science & Technology Plan (No. 2015B020233018).

ABOUT THE AUTHORS

Xiaobo Ma is an associate professor with MOE Key Lab for Intelligent Networks and Network Security, School of Electronic and Information Engineering, Xi'an Jiaotong University. His research interests include network security and privacy. He received a Ph.D. in Control Science and Engineering from Xi’an Jiaotong University, and was a Post-Doctoral Research Fellow with The Hong Kong Polytechnic University. He is a member of IEEE. Contact him at xma.cs@xjtu.edu.cn.

Yihui He is currently working toward the MS degree at the Carnegie Mellon University. His research interests include CNN acceleration, pruning networks, and deep learning. Contact him at yihuihe.yh@gmail.com.

Xiapu Luo served as the corresponding author for this paper. He is an assistant professor with the Department of Computing and an Associate Researcher with the Shenzhen Research Institute, The Hong Kong Polytechnic University. His research focuses on smartphone security and privacy, network security and privacy, and Internet measurement. He received the Ph.D. degree in Computer Science from The Hong Kong Polytechnic University, and was a Post-Doctoral Research Fellow with the Georgia Institute of Technology. Contact him at csxluo@comp.polyu.edu.hk.

Jianfeng Li is a Ph.D. student with School of Electronic and Information Engineering, Xi’an Jiaotong University. His research interests include network modeling and measurement. Jianfeng received the B.S. degree in automation engineering from Xi'an Jiaotong University. Contact him at jfli.xjtu@gmail.com.

Mengchen Zhao is a Ph.D. student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research lies in the areas of game theory, machine learning, reinforcement learning and their applications to real-world problems. Prior to enrolling at NTU, Mengchen received B.S in Applied Mathematics from Sun Yat-Sen University, China. Contact him at zhao0204@e.ntu.edu.sg.

Bo An is an associate professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include artificial intelligence, multiagent systems, game theory, and optimization. He received the Ph.D degree in Computer Science from the University of Massachusetts, Amherst. He is a member of the editorial board of JAIR and the Associate Editor of JAAMAS. He was elected to the board of directors of IFAAMAS. Contact him at boan@ntu.edu.sg.

Xiaohong Guan has been with the Systems Engineering Institute, Xi'an Jiaotong University, where he is currently a Cheung Kong Professor of Systems Engineering and the Dean of the School of Electronic and Information Engineering. His research interests include allocation and scheduling of complex networked resources, and network security. He received the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs, in 1993. Since 1995, he is also with the Department of Automation, Tsinghua National Laboratory for Information Science and Technology, and the Center for Intelligent and Networked Systems, TNLIST, Tsinghua University. He is an IEEE Fellow and an Academician of Chinese Academy of Sciences. Contact him at xghuan@xjtu.edu.cn.