

TOWARDS ADVERSARY AWARE SURVEILLANCE SYSTEMS

Vivek K. Singh and Mohan S. Kankanhalli

Department of Computer Science, National University of Singapore, Singapore.

ABSTRACT

We consider surveillance problems to be a set of system- adversary interaction problems in which an adversary can be modeled as a rational (selfish) agent trying to maximize his utility. We feel that appropriate adversary modeling can provide deep insights into the system performance and also clues for optimizing the system's performance against the adversary. Further, we propose that system designers should exploit the fact that they can impose certain restrictions on the intruders and the way they interact with the system. The system designers can find the assumptions under which the surveillance system shall out-perform the intruder and then enforce those assumptions over the system-intruder interaction as part of a 'scenario engineering' approach. We study both these aspects using a game theoretic framework and undertake practical experiments to verify the proposed enhancements.

1. INTRODUCTION

Recently a significant amount of research has been undertaken by the visual surveillance community in areas like advanced detection, tracking and recognition etc [1]. However, sensor-adversary interaction which is pivotal to surveillance systems is very rarely studied. Very few works model the adversary and the typical models have been poisson, equiprobability distributions etc. over the entire possible action sets. This is in stark contrast with the multimedia security/cryptography studies where deep insights have been gained by the modeling the performance of a 'smart adversary' against the system [2].

Thus we propose to introduce the notion of a 'smart adversary' into surveillance research. Such modeling can be used for measuring the system performance and analysing the best and the worst case performances. It can also be used to study the effects of changing different surveillance attributes (sensor positioning etc.) on the system performance.

We assert that another important concept missing from surveillance research is that of 'scenario engineering'. Typically the surveillance setup is assumed to be 'given' and 'fixed'. This may however not always be the case. For example, the positioning of an ATM machine within a lobby can and should be changed if it provides significant improvements in the surveillance effectiveness. We propose using the appropriate adversary and scenario modeling to find the assumptions and/or

enforcements which will improve the surveillance effectiveness and then enforcing them over the scenario to help the surveillance system outperform the adversary. The idea is akin to exploiting the 'Home Ground Advantage' to benefit your team wherever possible. This idea is also loosely related to context based awareness works wherein certain system parameters can be adjusted based on user behavior [3], though indeed our focus is very different.

In our proposed work, both adversary and scenario modeling have been studied using a game theoretic framework. We assume our adversary to be a rational selfish agent who has a clear goal and utility gains/costs associated with each of his actions. We model the interaction problem as that of two selfish rational agents (system and the adversary), both trying to maximize their utilities while being acutely aware that the other agent is also trying to do the same. This fits in very well under a game theoretic framework which have extensively been used to study similar interaction problems between multiple nations, competing firms and online bidders over the last few decades [4, 5].

To summarize the key contributions of this paper are:

1. Modeling of rational selfish adversaries in surveillance scenarios through a game theoretic framework.
2. Adoption of 'scenario engineering' concept to allow modifications to the surveillance environment in order to benefit the surveillance system rather than the adversary.

To check the applicability of the proposed approaches in practical surveillance scenarios we have considered a surveillance scenario of an enclosed rectangular area such as that of a museum subsection or an ATM lobby. We assume that the adversary goal is to reach the important artifact (ATM machine or the expensive art-piece), while the system goal is to capture the intruder's facial images. We relax the often (tacitly) used assumption that the adversary does not know where the camera is focusing. Rather, we assume that the adversary knows where the camera is focusing and actively tries to avoid getting his images captured.

To the best of our knowledge there have been no attempts at using the game theory to model surveillance scenarios as yet. Similarly there have been no attempts at *explicit* adversary modeling or scenario engineering in Surveillance.

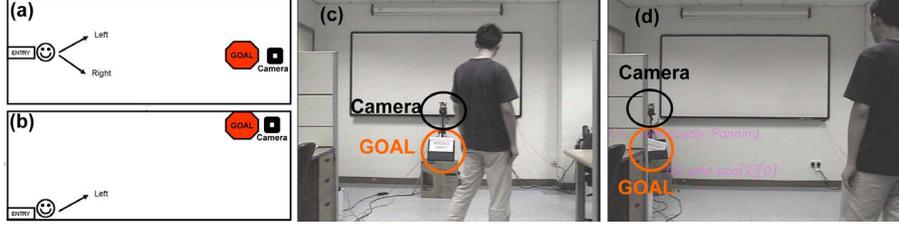


Fig. 1. (a) Schematic- Original scenario (b) After enforcement 1 (c) Physical- Original scenario (d) After enforcement 1

2. PROPOSED WORK

While our proposed ideas of *adversary modeling* and *scenario engineering* are generic, to understand their implications let us consider a base case surveillance scenario of an enclosed environment like a museum subsection or an ATM lobby. We assume that the adversary goal is to reach the important artifact (ATM machine or the expensive exhibit), while the system goal is to capture the adversary’s facial images.

A typical surveillance setup in such a scenario looks similar to as shown in figures 1 (a) and 1(c). The adversary can move towards the goal by going towards left or right direction. Clearly, a surveillance camera can only focus in one of the two directions, left or right. Hence, by purely random selection(s) by the Adversary Alice (A) and the Surveillance System Bob (B), B has a 50% probability of capturing A ’s image. Our claim is that through adversary modeling and scenario engineering B can almost always outperform A .

We assume both Alice and Bob to be rational agents trying to maximize their utilities. Their utilities for undertaking various actions are shown in table 1. The rows represent the direction of adversary’s motion while the columns represent the direction being focused by the camera. ν represents the gain for A in nearing the goal while σ is the cost incurred upon image capture. The matrix only shows values for A as the gains for her are considered the losses for B and vice-versa (Zero Sum Game).

The matrix shown in table 1 has 2 best cases (for A), no pure Nash Equilibrium and one impure Nash Equilibrium. Please recall that an impure equilibrium exists when both the players have no motivation in changing their adopted strategies *unilaterally*. Consequently it signifies that each player should become unbiased between his/her strategies. Thus equating the utilities achievable through A ’s two strategies:

$$(p_2)(\nu - \sigma) + (1 - p_2)(\nu) = (p_2)(\nu) + (1 - p_2)(\nu - \sigma) \quad (1)$$

where p_2 is the probability of B choosing strategy 1 i.e. Focusing Left.

Solving equation 1 we get $p_2 = 0.5$. Similarly we obtain p_1 i.e. probability of A choosing Left as 0.5 and Expected Utility for A (EU_A) is:

$$EU_A = \nu - 0.5\sigma \quad (2)$$

$Adversary \backslash System$	Left	Right
Left	$\nu - \sigma$	ν
Right	ν	$\nu - \sigma$

Table 1. Surveillance game between Adversary and System

Clearly, EU_A is high as Bob cannot monitor both the feasible options for Alice (Left and Right) at the same time. Thus, using the ‘Home Ground Advantage’, we modify this situation by our first enforcement.

Enforcement 1- Goal Placement: We enforce that the entry point and the goal be placed as further apart as possible in opposite directions as shown in figures 1(b) and 1(d). A ’s movement now (if at all possible) shall only take him further away from the goal and thus the game matrix changes to table 2. Clearly, A does not gain much now by going towards Right. Consequently there exists a pure Nash equilibrium at [Left, Left], provided $\nu - \sigma \geq 0$. Thus the utility now changes to:

$$EU_A = \nu - \sigma \quad (3)$$

This new value is clearly worse for Alice than earlier value (in Eq 2). However we as system designers hope to push it down further. We notice that this EU_A value holds only when $\nu \geq \sigma$ i.e. when cost of detection/image capture is less than the gain of nearing. This can only occur when images being captured are of poor quality (cannot be used for identification) or when the adversary is benign i.e. does not mind her images being captured.

Enforcement 2- Quality of Sensing: We enforce that the Quality of Sensing should be enough to act as a deterrence to the prospective adversaries. Thus the images ought to be captured at a minimum resolution required for identification. With this enforcement, there is no more pure Nash Equilibrium in the game. The impure Nash Equilibrium occurs when:

$$(p_2)(\nu - \sigma) + (1 - p_2)(\nu) = (p_2)(0) + (1 - p_2)(-\sigma) \quad (4)$$

and

$$(p_1)(-\nu + \sigma) + (1 - p_1)(0) = (p_1)(-\nu) + (1 - p_1)(\sigma) \quad (5)$$

which give p_1 as 0.5 and p_2 as $\frac{\nu + \sigma}{2\sigma}$ and net utility for A as:

$$EU_A = \left(1 - \frac{\nu + \sigma}{2\sigma}\right) \times (\nu - \sigma) \quad (6)$$

Adversary \ System	Left	Right
Left	$\nu - \sigma$	ν
Right	0	$-\sigma$

Table 2. Modified Surveillance game between Adversary and the System

Thus EU_A depends on the ratio between ν and σ but clearly shall be negative and worse than the earlier value (in Eq 3).

However, if we employ the ‘Theory of Moves’ [5] framework to study the matrix there shall be continuous cycles no matter what is the initial state. This means that B shall always try to focus on the side where A is, but she will change her side as soon as B does so and such a process gets continuously repeated. Recall that the ‘Theory of Moves’ promulgates that the a player in her 3rd best state should not try to go to her 2nd best state if her opponent can undertake a counteraction to push her to her 4th best state. Intuitively, this theory states that in a game of chess you should not take your opponents ‘pawn’, if doing so makes your ‘queen’ vulnerable. Thus A and B can *before-hand* calculate their expected pay-offs for various sequence of moves and then choose the best possible one. Further it states that cycles of moves shall result in some iteration loss for both players. Thus if we consider the iteration losses, the considered game becomes Non Zero-Sum as shown in table 3.

Enforcement 3- Power Play: We enforce that B ’s iteration loss (β) is negligible compared to A ’s iteration loss α . This makes sense as B can afford to be there for as long as required but A cannot stay inside the premises for a long duration. Thus instead of making continuous cycles of [Left] [Right] [Left]... movements A shall have to accept the ‘best-she-can-get’ situation of [Left, Left]. The [Left, Left] is the ‘best-she-can-get’ situation as the states [Left, Right] and [Right, Left] are unstable (B can unilaterally change its strategy to reduce A ’s utility) and state [Right, Right] provides an even lower utility value. If she does not undertake forward analysis to choose one stable state she shall suffer from significant iteration losses which will make her net utility even worse than the [Left, Left] state. Such a process of iteration between various possible utilities and eventual stabilization at [Left, Left] state is illustrated in figure 2. Please note that the Expected Utility for the A now becomes:

$$EU_A = \nu - \sigma \quad (7)$$

which shall have a negative value and shall be even worse than with enforcement 2 (Eq 6). Thus using a series of practically enforceable assumptions we have been able to push the Alice from having a 50% chance at best states to always getting her third best state outcome.

Adversary \ System	Left	Right
Left	$\nu - \sigma - n\alpha,$ $-\nu + \sigma - n\beta$	$\nu - n\alpha,$ $-\nu - n\beta$
Right	$-n\alpha, -n\beta$	$-\sigma - n\alpha, \sigma - n\beta$

Table 3. Further modified surveillance game between Adversary and the System

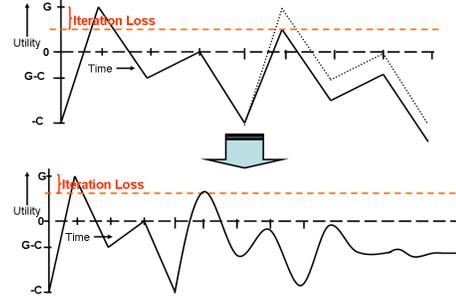


Fig. 2. Effect of powerplay enforcement

3. EXPERIMENTAL RESULTS

To check the veracity of our proposed approaches we conducted multiple rounds of experiments simulating a scenario where the adversary’s aim is to pick up a ‘precious object’ kept inside a room. The adversary aims to get as near as possible to the precious object without getting her facial images captured while the system tries to get as many high resolution facial images of her as possible.

We conducted 4 sets of experiments each consisting of 20 rounds. The experiments were conducted in an enclosed environment of 20ft by 15ft dimension using a Canon VC-C50i Pan Tilt Zoom camera. The volunteer adversaries were Multimedia Lab graduate students who were explained clearly the purpose of the experiments and the ‘gains’ and ‘losses’ that can be incurred by them. They were explicitly asked to beat the system by changing their trajectory as and when required but were asked to maintain a steady walking pace and not to hide their faces. In the base case setup we kept the precious object at the center of the room and allowed the adversaries to enter from a ‘virtual’ door in the center as shown in figure 1(a) (and 1(c)). The camera was employed at 768 by 576 pixel resolution with active Pan and Tilt (but no Zoom) to capture adversary face. After this, we iteratively enforced the three scenario engineering enforcements and checked their impact

Base case	Enfo. 1	Enfo. 2	Enfo. 3
13 (65%)	8 (40%)	5 (25%)	2 (10%)

Table 4. Number of rounds with successful steals

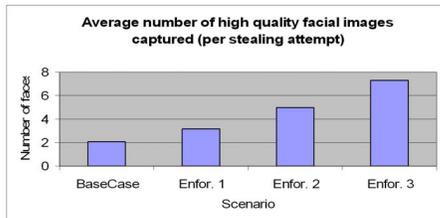


Fig. 3. Number of faces detected

on the number of ‘successful steals’ and the number of high resolution facial images found. We define ‘successful steals’ as those cases in which the system cannot obtain at least three high resolution facial images. We define high resolution images as those which capture facial information at minimum 100 by 100 px resolution (which can also be used for automatic face recognition at an accuracy of around 90% [6]). The underlying assumption here is that the adversary shall lose enough utility to counter her gain upon reaching the precious object, if her three high resolution images are captured.

As shown in table 4 and figure 3, we found that in the base case setup the adversaries were on average able to successfully steal the precious object 65% of times. Also, the average number of faces captured per stealing attempt was 2.1. To study enforcement 1: Goal Placement, we moved the precious object to the extreme left position and changed the virtual entry point to extreme right as shown in figures 1(b) and 1(d). The results show that successful steals decreased by 25% and the average number of facial images captured increased to 3.2 by this enforcement. However, we noticed that there were a significant number of frames with insufficient resolution facial images. Thus we used enforcement 2: Quality of Sensing and added a feature to zoom towards adversary faces upon detection. This reduced the number of successful steals by another 15% and the average number of images captured increased to 5. We noticed that some adversaries were able to counter the system by going away from the camera and approaching it in an obtuse trajectory.

Lastly, we employed the Power Play enforcement by allotting the intruders only 12 seconds to steal the precious object. This duration was assumed to be enough for picking the object as this was the average time taken by adversaries to steal the object in round 2 (using just Enforcement 1). We noticed that this enforcement further enhanced the system performance and only 10% of adversaries were able to get away without having 3 facial images captured. The average number of facial images captured also increased to 7.3. Those adversaries which still managed to perform well were those which constantly kept changing their trajectory and walked at a ‘brisk’ pace.

Based on these 4 rounds of experiments we observed that the 3 ‘scenario engineering’ based enforcements did indeed help in increasing the surveillance system performance. We

also noticed many similarities in the way game theory predicted and the way people behaved to the various setups. For example, multiple adversaries took a longer path to ensure not getting captured after enforcement 2 (Quality of Sensing to be good enough to act as deterrent) while they had taken relatively simpler trajectories after enforcement 1. Indeed not all behavior was consistent as what game theory predicted.

4. CONCLUSIONS

In this work, we have employed a game theoretic framework to model selfish adversaries in surveillance applications. We found that such a modeling of adversaries can provide us with performance bounds for surveillance systems and also provide prescriptive guidelines to improve the system performance. We further employed the ‘scenario engineering’ approach to modify the scenario itself so that the surveillance system can perform better against the adversary. Through experimental results we have verified the impact of various reasonable enforcements upon the system performance.

To demonstrate the key ideas we have modeled a simple but practical scenario of an enclosed environment. We hope to extend this work to more complex wide area scenarios in our future work. We have also used a *rational* model to study selfish adversaries. In future, we want to study how *non-rational* adversaries react in such scenarios and how we can handle different types of adversaries (benign and malicious) within the same system.

5. REFERENCES

- [1] W. Hu, T. Tan, L. Wang, and S. Maybank, “A survey on visual surveillance of object motion and behaviors,” in *Systems, Man and Cybernetics, Part C, IEEE Transactions on*, 2004.
- [2] M. Wu, W. Trappe, Z.J. Wang, and K. Liu, “Collusion-resistant fingerprinting for multimedia,” *Signal Processing Magazine, IEEE*, vol. 21 (2), pp. 15–27, 2004.
- [3] T. Kanter, “Event-driven, personalisable, mobile interactive spaces,” in *Symposium on Handheld and Ubiquitous Computing*, Bristol, UK, 2000.
- [4] E. Rasmusen, *Games and Information: An introduction to game theory*, Blackwell Publishers, USA, 2001.
- [5] S. J. Brams, *Theory of Moves*, Cambridge University Press, UK, 1993.
- [6] J. Wang, C. Zhang, and H. Shum, “Face image resolution versus face recognition performance based on two global methods,” in *Asian Conference on Computer Vision*, Jeju Island, Korea, January 2004.