TALENTED: An Advanced Guarantee Public Order Tool for Urban Inspectors

Mingchu Li(✉), Gang Tian, and Kun Lu

School of Software Technology, Dalian University of Technology, Dalian, China
mingchul@dlut.edu.cn

Abstract. In the streets of Chinese cities, we often see that illegal pedlars sell some fake and inferior products such as outdated food and inferior household goods to people who do not know about this, which may cause serious health problem. Besides, pedlars often cause people to gather and so may lead to traffic accidents. Thus, there are great requirements how to control illegal pedlars, and how to analyze, model and predict illegal pedlars activities. Such research will help urban inspectors decide better strategies to guarantee public order. Thus, in this paper, we explore this problem, and propose a model called TALENTED (Target Attributes LEarNing model with TEmporal Dependence) to deal with the problem. TALENTED provides three main contributions. First, a new learning model is proposed to predict the probability of each target being attacked, and our model consists of three aspects: (i) This model considers a richer set of domain features; (ii) Adversaries' previous behaviors affect their new actions; (iii) Each target has different attributes and the adversaries weight them differently. Second, we adopt a game-theoretic algorithm to compute the defender's optimal strategy. Finally, simulation results illustrate the reasonability and validity of our new model.

Keywords: Learning model · Public order
Stackelberg Security Game

1 Introduction

In the cities of China, illegal pedlars often sell fake and inferior products (household goods and outdated food) to people who do not know about these products in relatively prosperous places with the large flow of people, which may cause serious health problem. Besides, pedlars often cause people to gather and so may contribute to traffic accidents (as shown in Fig. 1(a)). We call the pedlars illegal sale problem as public order problem. To address the problem, the governments have to send urban inspectors to patrol the street and to catch the illegal pedlars (as shown in Fig. 1(b)) to maintain public order. Thus how to assign limited resources of urban inspectors to monitor these illegal actions is a very important issue problem for Chinese governments. In this paper, we model

L. Wang et al. (Eds.): QShine 2017, LNICST 234, pp. 228–237, 2018.
https://doi.org/10.1007/978-3-319-78078-8_23

this problem using Stackelberg game, and propose a model called TALENTED (Target Attributes LEarNing model with TEmporal Dependence). Our TALENTED is to help urban inspectors improve patrol efficiency such that illegal pedlars are deterred from selling in the streets of the city. In addition, different patrol strategies are generated for urban inspectors according to the distribution of illegal pedlars. However, illegal pedlars, in turn, can continually conduct surveillance on the urban inspectors' patrol strategy and then change the places of selling accordingly. Thus urban inspectors and pedlars form a game. As the urban inspectors, their primary objectives are to stop illegal sale, and their main method of doing so is to patrol the streets of city. During a patrol, urban inspectors will catch illegal pedlars who sell in the streets, confiscate any fake and inferior products, and a corresponding fine is imposed on the illegal pedlars. Therefore, it is important to help the urban inspectors to identify and predict the most likely spots/locations of illegal pedlars and to generate patrolling strategies so that public order problem is solved.

Defender-attacker Stackelberg Security Game (SSG) has been successfully applied to infrastructure security problems and wildlife protection [1–4]. In SSGs, In SSGs, the defender attempts to allocate her limited resources to protect a set of targets against the adversary who plans to attack one of the targets. Several models have been proposed to protect against perfectly rational and bounded rational adversaries [5–7]. In fact, previous work which (such as SSG-based anti-poaching tool called PAWS [3]) has been successfully applied in the wildlife protection domain. However, PAWS still has some limitations. First, PAWS is based on an existing adversary behavior model named as Subjective Utility Quantal Response (SUQR) [3], which has several limitations: (i) This model just relies on three domain attributes which can not provide a detailed description of the impact of environmental and topographic features on the poachers' behaviors; (ii) Poachers' activities are independent between time periods; (iii) The parameter which measures the weight of each factor in the decision making process for adversary is a single parameter vector. Second, in PAWS, the utility of players at each target is fixed. Actually, the utilities of players at each target vary with the migration of animals in real world.

Motivated by the success of defender-attacker (SSG) applications, we model the pedlars' illegal deal problem as a SSG, in which the urban inspectors play as the defenders and the illegal pedlars are the attackers. The regions where illegal pedlars often appear represents a target. In essence, TALENTED (Target Attributes LEarNing model with TEmporal Dependence) attempts to address all aforementioned limitations in PAWS while providing the following two key contributions. First, TALENTED attempts to address SUQR's limitations in modeling adversary behavior. More specifically, TALENTED introduces a new behavioral model based on softmax [8] to predict illegal pedlars' actions, and consists of three aspects: (i) This model considers a richer set of domain features in addition to the three features used in SUQR in analyzing the probability of each target being attacked; (ii) We incorporate the dependence of the illegal pedlars' behavior on their activities in the past into the component for predicting the probability of each target being attacked; (iii) In our new learning model, each target has different attributes and the adversaries weight them differently. Second, TALENTED presents the dynamic rewards and penalties of defenders functions according to the number of illegal pedlars. In detail, the number of illegal pedlars corresponds to the number of rewards and penalties of the urban inspectors, and defenders generate patrol strategies according to the distribution of adversaries. At the same time, the patrol strategies of defenders also affect adversaries' decisions. Therefore, the rewards or penalties of the defenders vary at each target in different period.

(a) Illegal pedlars sell on the street (b) Urban inspectors catch illegal pedlar

Fig. 1. Illegal pedlars and urban inspectors

The rest of the paper is structured as follows. In Sect. 2, we give the domain description. In Sect. 3, we introduce our new learning model. In Sect. 4, we give the game-theoretic algorithm of computing the defender's optimal strategy. In Sect. 5, the results of simulation and performance analysis are presented. In Sect. 6, our conclusions are presented.

2 Domain

In China, we often see that illegal pedlars sell some fake and inferior products (household goods and outdated food) in the streets of the city, and often cause people to gather and so may lead to traffic accidents. To deal with the problem, the urban inspectors have to patrol the streets and catch the illegal pedlars. In addition to their normal patrol duties, urban inspectors also collect and analyze data on illegal pedlars' activities. These data will be used to obtain best patrol strategies for urban inspectors.

In the public order domain, the urban inspectors plays as the leaders and the illegal pedlars are the followers. City area is divided into grids, where each cell represents an attack target and contains potential customers for illegal pedlars. Note that each cell represents 1 km^2 (maybe less or more according to different city's requirement). An attack is that illegal pedlars sell items in the cell. If illegal pedlars attack a target which is uncovered by urban inspector, they receive a reward which is related to the number of potential customers in the target. Otherwise, they receive a penalty which corresponds to the fine being caught. At the same time, if urban inspectors patrol a target, they receive a dynamic reward which corresponds to the total fine received from the captured illegal pedlars. Otherwise, they receive a dynamic penalty. The dynamic rewards and penalties of defenders would vary with the number of adversaries in each target distribution. The purpose of the illegal pedlars is to sell as many goods as possible and not to be caught. Therefore, the flow of people, the flourishing degree of target and the urban inspectors' patrol strategies have an impact on illegal pedlars' decisions. Moreover, for a long-term benefit, illegal pedlars may tend to come back to the areas where they have attacked before. Our work will focus on incorporating all these factors into our model. Because of limited resources, urban inspectors can not patrol all potential targets. Thus, we propose a model called TALENTED to aid patrol managers and determine an optimal strategy so that urban inspectors can effectively cover these numerous places with their limited resources.

3 Behavioral Learning

As we know, in order to guarantee public order, we need to study a model to analyze and predict the probability of each target being attacked so that urban inspectors can effectively decide their strategies to solve public order problem. This paper introduces a new behavioral model to predict the probability of each target being attacked for urban inspectors in the public order domain.

In the public order domain, there are many illegal pedlars in the urban streets of China that affect urban transportation, city sanitation and public health. The public order domain is different from wildlife protection or illegal fishing [3,5,9,10]. We choose to represent the regions as targets, where illegal pedlars often appear. Therefore, we assume that all targets must be attacked. Specially, the new learning model is proposed to predict the probability of each target being attacked, and learn the different weights for each factor different target. This new model helps urban inspectors to find the attacked targets which is most likely attacked, and to generate their patrol strategies.

3.1 Proposed Model

We use K to denote the set of locations that can be targeted by the illegal pedlars, where $i \in K$ represents the i^{th} target. We denote by M the number of resources, N the number of adversaries, L the number of domain features, and T the number of time periods. Overall, each target has a set of feature values $t_i = \{t_i^l\}$, where $l = 1, \dots, L$ and t_i^l is the value of the l^{th} feature at target i. In our model, we adopt five domain features: road number, number of residential areas, visitors flow rate, market distance, and station distance which impact illegal pedlars' decisions. In addition, $x_{t,i}$ is defined as the coverage probability of the resources in time period t on target i. Moreover, at each time step t, N^t is defined as the total number of illegal pedlars at all targets, N_i^t is defined as the number of illegal pedlars at target i, at time period t. In other words, we have $N^t = \sum_{i \in K} N_i^t$, and $N_i^t > 0$ for all time step t.

Our new model considers poachers' behavior to be dependent between different time steps, we incorporate the dependence of the illegal pedlars behavior into their activities in the past, as illegal pedlars may tend to come back to the areas where they have attacked before. Therefore, we define the exponential update function to describe the degree of target i being attacked at the past time period ($p'_{t-1,i}$). We evaluate the impact of illegal pedlars' activities in the previous period and prior behavior in the past period. Furthermore, the exponential update function of target i before time period t is shown as Eq. (1):

$$p'_{t-1,i} = \begin{cases} \alpha p'_{t-2,i} + (1-\alpha)\frac{N_i^{t-1}}{N^{t-1}} & \text{if} \quad t > 2 \\ \frac{N_i^{t-1}}{N^{t-1}} & \text{if} \quad t = 2 \end{cases}, \tag{1}$$

where α is the weight factor, and $0 \leq \alpha \leq 1$. Moreover, $\frac{N_i^{t-1}}{N^{t-1}}$ indicates illegal pedlars' activities in the previous period at each target. In other words, $\frac{N_i^{t-1}}{N^{t-1}}$ is the ratio of target i being attacked at time period $t-1$; $p'_{t-2,i}$ indicates the exponential update function of target i at time step $t-2$. N_i^{t-1} and N^{t-1} are the data which is collected in the past time period.

To predict the probability of each target being attacked, we adopt the softmax regression model which takes int account the several factors above. Thus, given the urban inspectors' coverage probability of target i at time period t: $x_{t,i}$, the exponential update function of target i in the past time step: $p'_{t-1,i}$, and the domain features: $t_i = \{t_i^l\}$, we aim at predicting the probability of target i being attacked at time period t as Eq. (2):

$$p(k = i | 1, x_{t,i}, p'_{t-1,i}, t_i) = \frac{e^{\theta_i^{\mathrm{T}}[1, x_{t,i}, p'_{t-1,i}, t_i]}}{\sum_j e^{\theta_j^{\mathrm{T}}[1, x_{t,j}, p'_{t-1,j}, t_j]}}, \tag{2}$$

where $k \in K$, and $\theta_i = \{\theta_{ij}\}$ is the $(L+3) \times 1$ parameter vector of target i which measures the importance of all factors with the target i being attacked and L are the number of domain features. θ_{ij} is the j^{th} parameter in θ_i. θ_{i1} is the free parameter and θ_i^{T} is the transpose vector of θ_i. $\theta = \{\theta_1^{\mathrm{T}}, \ldots, \theta_K^{\mathrm{T}}\}^{\mathrm{T}}$ is the parameter matrix of all targets. In essence, our new model learns all targets' weights for target attributes in predicting the probability of each target being attacked at the same time.

3.2 Parameter Estimation

We employ Maximum Likelihood Estimation (MLE) to learn the parameters matrix $\theta = \{\theta_1^{\mathrm{T}}, \ldots, \theta_K^{\mathrm{T}}\}^{\mathrm{T}}$ for each target [11]. First we formulate the log-likelihood of softmax, given the defender strategy $\mathbf{x} = \{x_{t,i}\}$, and a set of samples of the adversaries choices as Eq. (3):

$$\log L(\theta) = \sum_n \log \prod (p(y^{(n)} = i|\theta))^{1 \cdot \{y^{(n)} = i\}}, \tag{3}$$

where $y^{(n)}$ is the n^{th} sample, i is the chosen target in that sample, and $p(y^{(n)} = i|\theta)$ presents the probability that the target i is chosen in Eq. (2), $1 \cdot \{y^{(n)} = i\}$ is indicative function, if $y^{(n)} = i$ is true, $1 \cdot \{y^{(n)} = i\} = 1$, otherwise, $1 \cdot \{y^{(n)} = i\} = 0$. Then we have:

$$\log L(\theta) = \sum_n \sum_i 1 \cdot \{y^{(n)} = i\} \log(\frac{e^{\theta_i^{\mathrm{T}}[1, x_{t,i}, p'_{t-1,i}, t_i]}}{\sum_j e^{\theta_j^{\mathrm{T}}[1, x_{t,j}, p'_{t-1,j}, t_j]}}) \tag{4}$$

Essentially, we can see that $\log L(\theta)$ in Eq. (4) is a concave function, this function has an unique local maximum point, since the Hessian matrix is negative semi-definite. Thus, we can compute the optimal weights matrix $\theta = \{\theta_1^{\mathrm{T}}, \ldots, \theta_K^{\mathrm{T}}\}^{\mathrm{T}}$ as follow.

$$\theta = arg \max_\theta \log L(\theta) \tag{5}$$

4 Patrol Planning

Once the model parameter vector of each target $\theta_i = \{\theta_{ij}\}$ is learned, we can compute the optimal strategies for the urban inspectors in the next time period with the new learning model, given the urban inspectors patrol strategies and domain features.

In our model, the number of illegal pedlars corresponds to the number of rewards and penalties of the urban inspectors. Defenders generate patrol strategies according to the distribution of adversaries. At the same time, the patrol strategies of defenders also affect adversaries' decisions. Therefore, the rewards or penalties of the defenders vary at each target in different period. We call the rewards or penalties as dynamic rewards or penalties for urban inspectors.

At each time period, if urban inspectors patrol target i, they receive a dynamic reward $R^d_{t,i}$, otherwise they receive a dynamic penalty $P^d_{t,i}$. At the same time, if illegal pedlars attack target i which is covered by urban inspector, they receive a penalty P^a_i, otherwise, they receive a reward R^a_i. The dynamic rewards and penalties of defenders will vary with the number of adversaries in each target distribution at different time steps. Therefore, given the probability of each target being attack, if urban inspectors patrol target i at time period t, they receive a dynamic reward $R^d_{t,i}$. Their dynamic rewards are computed as Eq. (6):

$$R^d_{t,i} = p_{t,i} N^t_i, \tag{6}$$

where $p_{t,i}$ is the abbreviations of the probability that target i is attacked at time period t in Eq. (2). Similarly, urban inspectors' dynamic penalty: $P^d_{t,i} = -R^d_{t,i}$.

In this paper, we consider a long-term benefit for the players. Suppose that N^t and N^t_i are known for the players, where $t = 1, \ldots, T$ and $i = 1, \ldots, K$. Similar to standard SSGs. We assume that if the urban inspectors patrol, they obtain a dynamic reward $R^d_{t,i}$, otherwise, they receive a dynamic penalty $P^d_{t,i}$. Therefore, at each time period, the urban inspectors' expected utility at target i is computed as Eq. (7):

$$U^d_{t,i} = x_{t,i} R^d_{t,i} + (1 - x_{t,i}) P^d_{t,i} \tag{7}$$

The purpose of the urban inspectors is to obtain the maximum expected utility. Thus, given the urban inspectors' patrol history data N^t and N^t_i, and the model parameters matrix θ, the problem of computing the optimal strategies $x_{t+1,i}$ for urban inspectors at the next time period $t+1$ can be formulated as follows:

$$\max_{x_{t+1,i}} \sum_i U^d_{t+1,i} \tag{8}$$

$$s.t. \qquad 0 \leq x_{t+1,i} \leq 1; \quad i \in K \tag{9}$$

$$\sum_i x_{t+1,i} \leq M; \quad i \in K \tag{10}$$

where M is the total number of resources and $U^d_{t+1,i}$ is the urban inspectors' expected utility in Eq. (7), and K is the set of locations that can be targeted by the illegal pedlars.

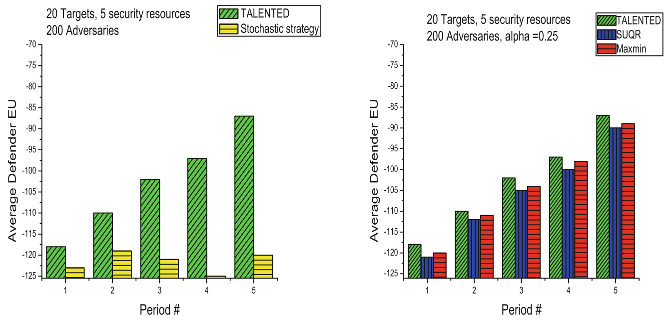
Therefore, we can piecewise linearly approximate $U^d_{t+1,i}$ and represent (8–10) as a Mixed Integer Program which can be solved by CPLEX. The details of piecewise linear approximation can be found in [12]. Essentially, the piecewise linear approximation method provides an $O(\frac{1}{P})$-optimal solution for (8–10) where P is the number of piecewise segments [12].

5 Experiments

In this section, we aim to evaluate the solution quality and runtime of the TALENTED planning for generating patrols. The results are obtained using CPLEX to solve the MILP for TALENTED. All experiments are conducted on a standard 2.00 GHz machine with 4 GB main memory. In the following, we provide a brief description of experiment settings to our new model.

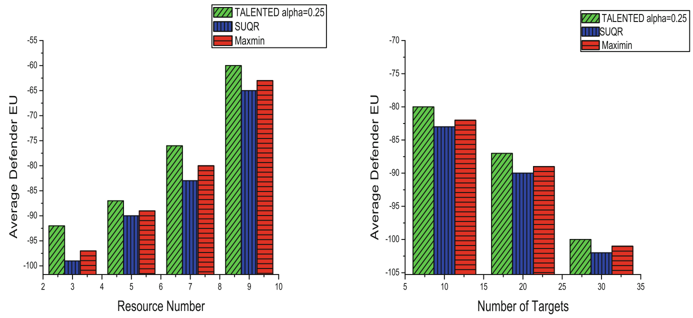
In the first time period, we randomly generate a defender strategy, then we simulate the target choices made by illegal pedlars according to the strategy. We call this process as a round of games. At next time period, we change the defender strategy according to the adversary's behavior. At each time period, we conduct 10 rounds game, after each period, we count up the number of illegal pedlars choosing each target i, N^t_i. We assume that these date are known by players. Then, we learn the parameter vector $\theta_i = \{\theta_{ij}\}$ for each target and compute the average expected utility of defenders and constantly update the parameter vector $\theta_i = \{\theta_{ij}\}$.

In both Fig. 2(a) and (b), the y-axis displays the average EU (Expected Utility) of the urban inspectors after each time period, and the x-axis displays the number of time period. In Fig. 2(a), we compare the average EU of TALENTED with stochastic strategy at different periods. In Fig. 2(b), we compare two different approaches: SUQR, and maximin strategy. In both Fig. 2(a) and (b), we set the number of adversaries to 200, the number of targets to 20, the number of resources to 5, and $\alpha = 0.25$ in Eq. (1). The parameter vector $\theta_i = \{\theta_{ij}\}$ are learned by our new model in Sect. 3.2. As shown in the figure, taking into account the dependence of the illegal pedlars behavior on their activities in the past is critical for urban inspectors to predict the probability of each target being attacked. TALENTED outperforms stochastic strategy, SUQR that just consider single parameter vector and independent adversaries activities between time periods and maximin strategy.



(a) Cumulative EU (b) Cumulative EU

Fig. 2. Simulation results over period

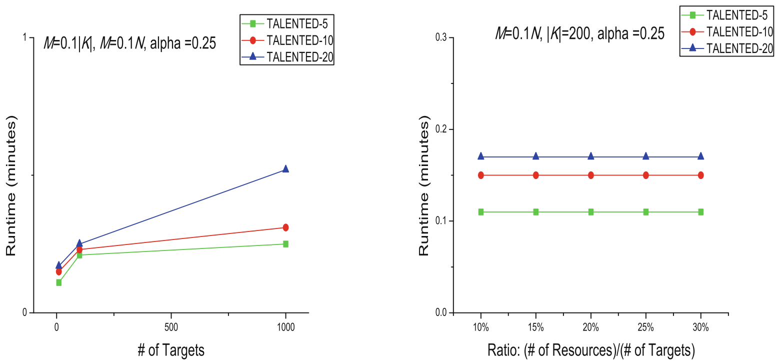


(a) 20 Targets, 200 Adversaries (b) 5 security resources, 200 Adversaries

Fig. 3. Comparing cumulative EU at period 5

Then, we compare the average EU achieved by the three different methods under different number of targets and different amount of resources. In both Fig. 3(a) and (b), the y-axis displays the average EU of the urban inspectors after 5 time periods. In both figures, we also simulate 200 illegal pedlars. In Fig. 3(a), we vary the number of resources on the x-axis while fixing the number of targets to 20. It shows that the average EU increases as more resources are added. In addition, TALENTED outperforms the other two approaches regardless of resource quantity. Similarly, we vary the number of targets on the x-axis in Fig. 3(b) while fixing the amount of resources to 5. The better performance of TALENTED over the other two methods can be seen from the figure regardless of the number of targets.

Furthermore, we give the runtime of our model. We present the runtime results in Fig. 4(a) and (b). In all two figures, the y-axis display the runtime, the x-axis displays the variables which we vary to measure their impact on the runtime of the algorithms. In both Fig. 4(a) and (b), M is the number of resources, N is the number of adversaries, and K is the set of targets. α is the weight factor in Eq. (1). We compare the runtime when $P = 5$, $P = 10$, and $P = 20$, where P is the number of piecewise segments in [12]. According to the two figures, we find that TALENTED can deal with large-scale problems and get better results.



(a) Runtime v.s. $|K|$ (b) Runtime v.s. M

Fig. 4. Evaluate the runtime

6 Conclusions

As we know, pedlars' illegal deal problems seriously affect urban transportation and public health in China and may lead heavily transportation accidents. In this paper, we propose a new method called TALENTED to deal with the problem in public order domain, and will be applied in the Kaifa district of Dalian, China. TALENTED adopts a novel learning model which considers a richer set of domain features and incorporates the dependence of the illegal pedlars behavior into their activities in the past. Our new learning model can effectively predict the probability of each target being attacked in the public order domains. Moreover, we adopt a game-theoretic algorithm to compute the defender's optimal strategy. Finally, we have a large number of simulation experiments to testify the reasonability and validity of our new model. The experimental results demonstrate the superiority of our model compared to other existing models.

Acknowledgments. This paper is supported by Nature Science Foundation of China under grant No. 61572095.

References

1. Basilico, N., Gatti, N., Amigoni, F.: Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In: AAMAS (2009)
2. Ford, B., Brown, M., Yadav, A., Singh, A., Sinha, A., Srivastava, B., Kiekintveld, C., Tambe, M.: Protecting the NECTAR of the Ganga River through game-theoretic factory inspections. In: Demazeau, Y., Ito, T., Bajo, J., Escalona, M.J. (eds.) PAAMS 2016. LNCS (LNAI), vol. 9662, pp. 97–108. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39324-7_9
3. Yang, R., Ford, B., Tambe, M.: Adaptive resource allocation for wildlife protection against illegal poachers. In: Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, pp. 453–460. AAMAS (2014)
4. Nguyen, T.H., Sinha, A., Gholami, S.: CAPTURE: a new predictive anti-poaching tool for wildlife protection. In: Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems, pp. 767–775. AAMAS (2016)
5. Kar, D., Fang, F., Delle, F.F.: A game of thrones: when human behavior models compete in repeated Stackelberg security games. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, pp. 1381–1390. AAMAS (2015)
6. Kiekintveld, C., Jain, M., Tsai, J.: Computing optimal randomized resource allocations for massive security games. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, pp. 689–696. AAMAS (2009)
7. Carthy, S., Marie, M., Tambe, M.: Preventing illegal logging: simultaneous optimization of resource teams and tactics for security. In: AAAI Conference on Artificial Intelligence (2016)
8. Duan, K., Keerthi, S.S., Chu, W., Shevade, S.K., Poo, A.N.: Multi-category classification by Soft-Max combination of binary classifiers. In: Windeatt, T., Roli, F. (eds.) MCS 2003. LNCS, vol. 2709, pp. 125–134. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-44938-8_13
9. Gholami, S., Wilder, B., Brown, M., Sinha, A.: A game theoretic approach on addressing collusion among human adversaries. In: Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems (2016)
10. Haskell, W.B., Kar, D., Fang, F.: Robust protection of fisheries with compass. In: AAAI, pp. 2978–2983 (2014)
11. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics. Springer, Berlin (2001)
12. Yang, R., Ordonez, F., Tambe, M.: Computing optimal strategy against quantal response in security games. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, pp. 847–854. AAMAS (2012)