CS 5390, Spring ’23

Project Report

Student Name: Saeefa Rubaiyet Nowmi

Graduate Student, Computer Science, University of Texas at El Paso

Student ID: 80782256

Project Title: FlipIT Game Model with IBL and Comparison with Human Data

1.Introduction

There is a lot of work going on in the cyber security domain, but most of them use game theory models. The main motivation of my project was to see if it is possible to create cognitive models to replicate such cyber security scenarios and to compare the results with human data to see how closely the cognitive models can capture human behavior.

Typically, Nash equilibrium and other solution notions in game theory require rational behavior. Nevertheless, this presumption is frequently broken when human agents engage in real-world situations, including cybersecurity. Game theory is finding more and more applications in cybersecurity, such as the strategic placement of honeypots (Kiekintveld et al., 2015; Píbil et al., n.d.)to gather information about the attacker or impede their progress. Other instances (Durkota Karel and Lisý, 2017; Shiva et al., 2010) of dynamic (stochastic or in the extensive form) games exist. Standard game models frequently assume that players are rational and that the objective is to find an optimal strategy in the form of a Nash equilibrium.

However, researchers frequently use limitedly rational game methods against humans or other types of opponents. There are several earlier efforts that take into account the attackers' behavioral modeling. Nochenson and Grossklags (Nochenson & Grossklags, n.d.) conducted an initial study on the FlipIt game to examine the effects of participant age and gender on performance. To forecast user behavior, the authors used a regression model. Another study by Reitter et al., n.d. examined how players in the FlipIt game reacted to risk and found that it had an impact on their performance. Additionally, the authors developed a cognitive model based on ACT-R that simulates a person's risk propensity and decision-making process. Another strategy is to model the attacker using Instance Based Learning (IBL) (Dehghani Abbasi et al., 2015)or the defender to support the network administrator.

Another strategy is to model the attacker using Instance Based Learning (IBL) (Dehghani Abbasi et al., 2015) or the defender to support the network administrator. A noteworthy area of research is Stackelberg Security Games (SSGs), in which the attacker is modeled using a variety of behavioral models, such as Prospect Theory (PT), Quantal Response (QR)

(Mckelvey & Palfrey, 1998), and the Subjective Utility Quantal Response (SUQR) model(Nguyen et al., n.d.; Yang et al., n.d., 2014).

For this project, I used Basak et al., 2018, paper as the primary reference. I created two cognitive IBL models, one with random defender and another with strategic defender, following there online Strata FlipIT Game and used there human data set to compare with my cognitive models.

2. Task Description

The tasks for this project were divided into four parts. Which are- 1) Developing IBL Model for the FliptIT game with an IBL attacker against a random defender with full information, 2) Developing IBL Model for the FlipIT game with an IBL attacker against an IBL strategic defender with full information, 3) Developing IBL Model with for the FlipIT game with an IBL attacker against an IBL strategic defender with partial information and 4) Analyze human data and the data from both the models and compare both the models with human attackers data.

2.1. Game Description for Human Players

According to Basak et al., 2018, in their Strata FlipIT game, two players compete over a network with multiple nodes. A node can be any machine in the network. In their experiment they used six nodes: Node A(10/8), Node B(10/2), Node C(4/2), Node D(4/8), Node E(10/5), Node F(0/0). Each node has a reward (ρ)and a cost(γ). For example, Node A has a reward of 10 and cost 8. The defender/attacker has to pay the cost each time he wants to defend/capture a node. Each node can be either captured by the attacker or not.

The game had five rounds. Initially, the defender has control over all the nodes. The purpose of the attacker/participant is to take over control from the defender by attacking nodes. In each round, the defender defends a node and the attacker attacks a node. If the attacker or defender chooses to attack/defend a node he/she has to pay the cost associated with that node.

If the defender and attacker do not make the same move, then the attacker takes control of the node and receives the reward associated with that node and pays the cost. If both players make the same move, then the previous controller of the resource retains the control and attacker and defender both pay the cost.

Each of the participants played total 6 rounds containing 5 trials in each round. In each round, the user interface shows the following information: total points, current round, time, action history (log), and who currently controls each node. After each round the attacker receives points for all the nodes he controls. Red and blue mean the attacker or the defender controls a node, respectively. The attacker is able to observe the effect of the defender action in the next round. Each player tries to maximize their utility by controlling the nodes.

2.2 Description of the Human Data Set

Basak et al., 2018, designed the experiment with two types of games and two types of defenders. Game type 1 is the full information game and game type 0 is the minimal/partial information game. Here, full information refers to the fact that participants were fully aware of the payoff structure and game environment before playing the game. And the minimal structure refers to the fact that the participants did not know about the payoff structure before starting the game.

The random defender type 0 indicates that the defender is of strategic type and the random defender type 1 indicates that the defender is of random type.

But in the data set, I did not find any data with random defender type 1. Rather found some data with random defender type 2 and random defender type 3. To avoid any confusion, I cleaned the data set by eliminating the data with random defender type anything but 0. After cleaning the data set, there remained 152 participants of game type 1 and random defender type 0 , and 151 participants of game type 2 and random defender type 0.

So, for my project I used two types of human attackers' data set. Both types played against the strategic defender, but one type (game type 1) played knowing all the information before starting the game and another type (game type 0) played not knowing any information before starting the game.

2.3 Game Description for Cognitive Models

I followed the Strata FlipIT Game by Basak et al., 2018, for all the models. In my models, two players competed over a network with six nodes. Node A(10/8), Node B(10/2), Node C(4/2), Node D(4/8), Node E(10/5), Node F(0/0). Each node has a reward and a cost. For example, Node A has a reward of 10 and costs 8. The defender/attacker must pay the cost each time he wants to defend/capture a node. If the attacker and the defender choose the same nodes, the defender wins. Otherwise, the attacker wins. Whoever wins gets the reward associated with their chosen node.

3. Model Description

3.1 Model 1: IBL Model for the FliptIT game with an IBL attacker against a random defender with full information

The model was created with PyIBL. One attacker agent was created with default noise and decay parameters (0.25 and 0.5 respectively). Reward and cost were incorporated into the attacker agent as attributes. One function was created for feeding the attacker agent with the outcomes or responses according to the payoff structure described in section 2.3. Linear similarity function was used for both the attributes with a mismatch penalty value of 2.5. The agent was prepopulated with all six options with respective positive reward, negative cost and net reward outcome. For example, to populate the attacker agent with node A, I used the following:

```
>>>attacker_agent.populate([{"reward":10,"cost":-8}],2)
```

As the model was a full information model, I prepopulated the model with all six options. To match with the human data set, the model was run for 152 participants and with 30 trial for each participant.

3.2 Model 2: IBL Model for the FlipIT game with an IBL attacker against an IBL strategic defender with full information

This model was also created with PyIBL. One attacker agent and one defender agent both were created with default noise and decay parameters (0.25 and 0.5 respectively). Reward and cost were incorporated into both the agents as attributes. Two different functions were created for feeding the attacker agent and the defender agent with respective outcomes or responses according to the payoff structure described in section 2.3. The same linear similarity function was used for both the attributes with a mismatch penalty value of 2.5 for both the agents. Both the agents were prepopulated with all six options with respective positive reward, negative cost and net reward outcome. For example, to populate the attacker agent and the defender agent respectively with node A, I used the following:

```
>>>attacker_agent.populate([{"reward":10,"cost":-8}],2)
```

```
>>>defender_agent.populate([{"reward":10,"cost":-8}],2)
```

As the model was a full information model, I prepopulated both the agents with all six options. To match with the human data set, the model was run for 152 participants and with 30 trial for each participant.

3.3 Model 3: IBL Model for the FlipIT game with an IBL attacker against an IBL strategic defender with partial information

This model was also created with PyIBL. One attacker agent and one defender agent both were created with default noise and decay parameters (0.25 and 0.5 respectively). Reward and cost were incorporated into both the agents as attributes. Two different functions were created for feeding the attacker agent and the defender agent with respective outcomes or responses according to the payoff structure described in section 2.3. The same linear similarity function was used for both the attributes with a mismatch penalty value of 2.5 for both the agents.

As the model was a partial information model, none of the agents were prepopulated with any of the options with actual outcome. Rather, default_utility was used for both the agents with a higher value than the actual outcomes of all the nodes. This was done so that the agents could explore the options in initial few trials and learn from the exploration to depict the scenario with the human participants with partial information game.

This model was also simulated to run for 152 participants and with 30 trial for each participant.

4. Results

I analysed the data from three major aspects: 1) proportion of choosing different nodes over 30 trials and comparing the trend of the models and human data, 2) proportion of total attacks on different nodes and comparison among the models and human data 3) winning ratio of the models and the humans and comparison among them.

4.1. proportion of choosing different nodes over 30 trials and comparing the trend of the models and human data

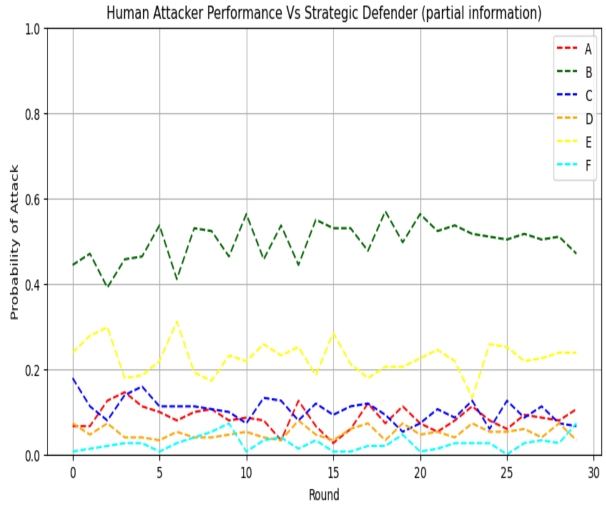

Figure 1.proportion of attack on different nodes (Humans Vs Strategic Defender (partial information))

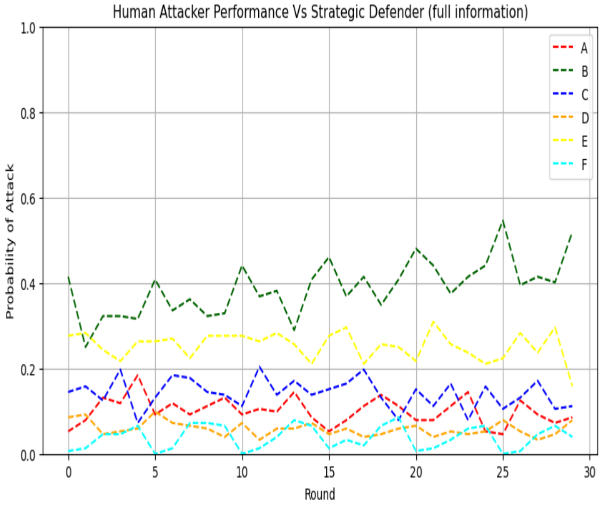

Figure 2.proportion of attack on different nodes (Human Vs Strategic Defender (full information))

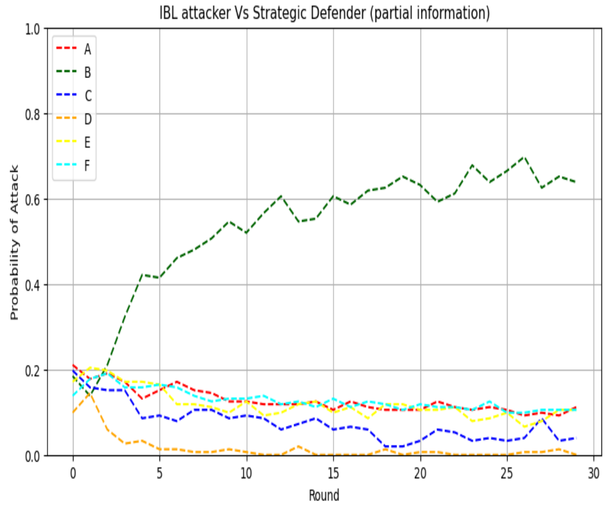

Figure 3.proportion of attack on different nodes (IBL Vs Strategic Defender (partial information))

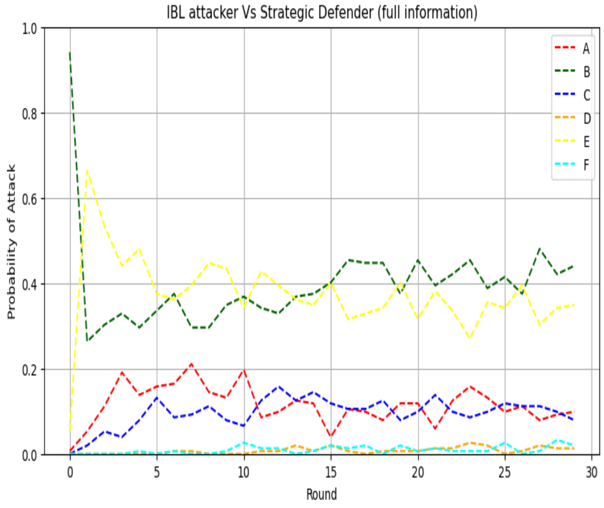

Figure 4.proportion of attack on different nodes (IBL Vs Strategic Defender (full information))

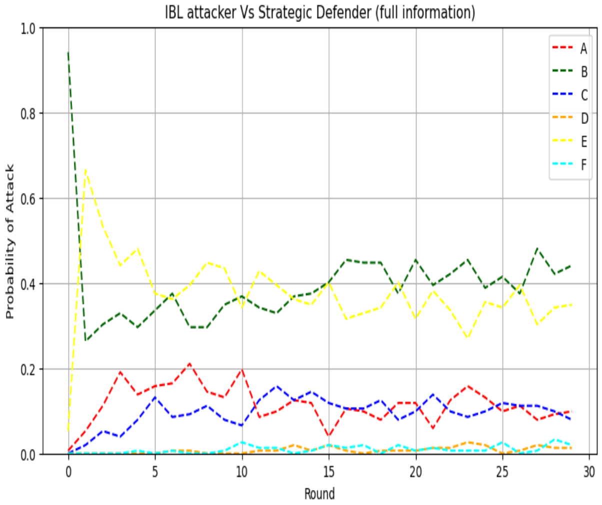

Figure 5.proportion of attack on different nodes (IBL Vs Random Defender (full information))

From figure 1 and figure 2, we can see the proportion of attack on different nodes by human attackers against strategic defender in partial information game and full information game respectively. For the full information game, from the trend of choosing node B,it can be seen that humans are choosing option B in an increasing manner. It can be because, once they have captured node B, in later trials they prefer to protect node B rather than capturing another node. Because, as the game is full information one, they know the payoff structure and know that they need to protect the captured node till the end of the rounds. And as B is the most beneficial node for them, they try to protect the node B rather than capturing other nodes. But for the partial information game, no such trend is noticeable. It can be for the reason that the participants are not aware of this information about the game.

Figure 3 and figure 4 show the proportion of attack on different nodes by IBL attackers against strategic defender in partial information game and full information game respectively. For the partial information game, in the initial few rounds, the IBL attacker explores the options and learns from the outcomes and it's learning is gradually reflected in later rounds. There is an increasing trend for option B which shows that the IBL agent is learning that node B is the most beneficial option for it. As attacks on node B increase drastically, attacks on other nodes decrease consequently. In figure 4, it can be seen that trend for choosing node B is less steep than figure 3. It is because the agent already new that node is the most beneficial option for it and thus from the begging, attacks on node was higher comparatively than the figure 3. From figure 4, it can also be seen that there is no exploration phase, as the agent already knew all the information. So, it shows more or less same proportion of attack on particular node over the trials. The initial steep decrease for node B and steep increase for node E are due to the prepopulation at trial 0.

Figure 5 shows the proportion of attacks on different nodes by IBL attackers against a random defender in a full information game. As the game was full information, the increasing attack curve on node B is noticeable but less steep than the partial information graphs. The initial steep decrease for node B and steep increase for node E are due to the prepopulation at trial 0.

From all five figures, it can clearly be seen that figure 2 and figure 4 have the most similarity between them. So, it can be said that IBL model against strategic defender for full information game captured the human behavior from the human attackers against strategic defender in full information game satisfactorily.

4.2 proportion of total attacks on different nodes and comparison among the models and human data

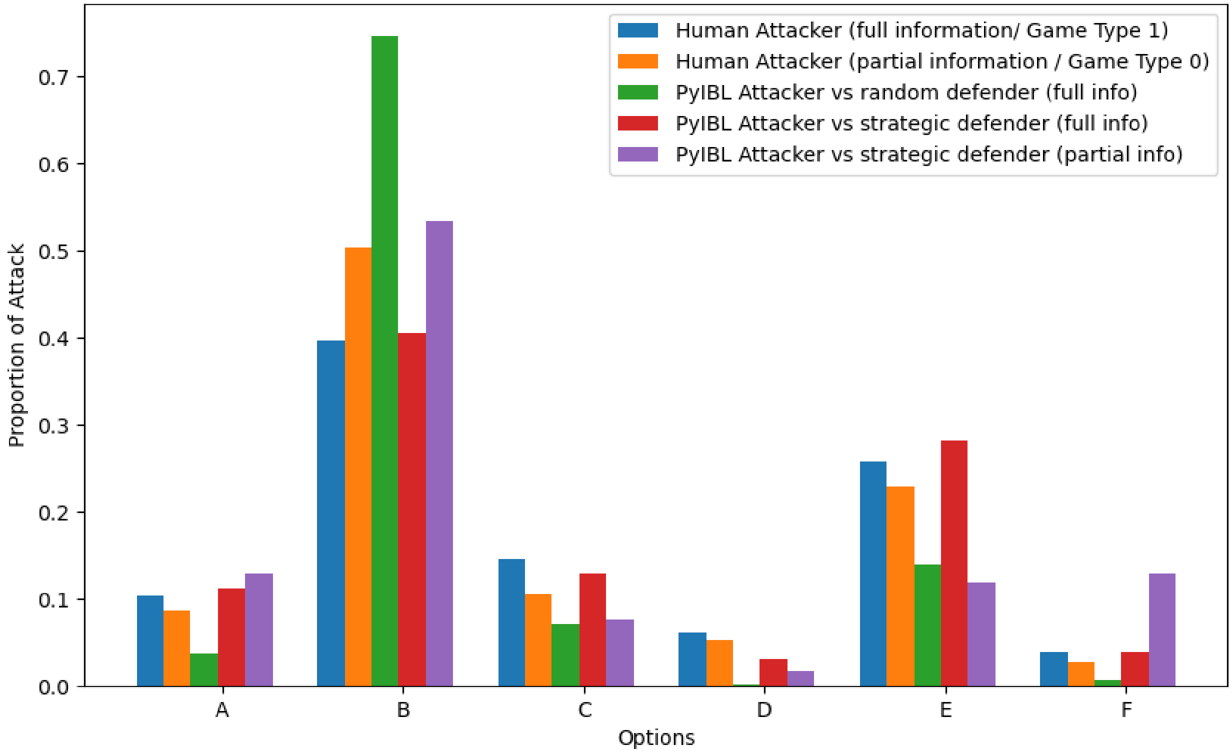

Figure 6. proportion of total attack on different nodes

Figure 6 shows the proportion of total attacks on different nodes for all the three models and two sets of human data. From the graph, it can be seen that the red and the blue bars for all the nodes have the most close values. From this graph, it can again be established that the IBL model against the strategic attacker for full information game could capture the human behavior against the strategic attacker for the full information game.

4.2 Total winning ratio comparison among the models and human data

Figure 7 shows the overall winning proportion of the three models and human attackers. IBL (m-1) indicates the first IBL model with random defender and full information, IBL (m-2) indicates the second IBL model against strategic defender and full information and IBL (m-3) indicates the third IBL model against strategic defender and partial information. Both the human data set were against strategic defenders.

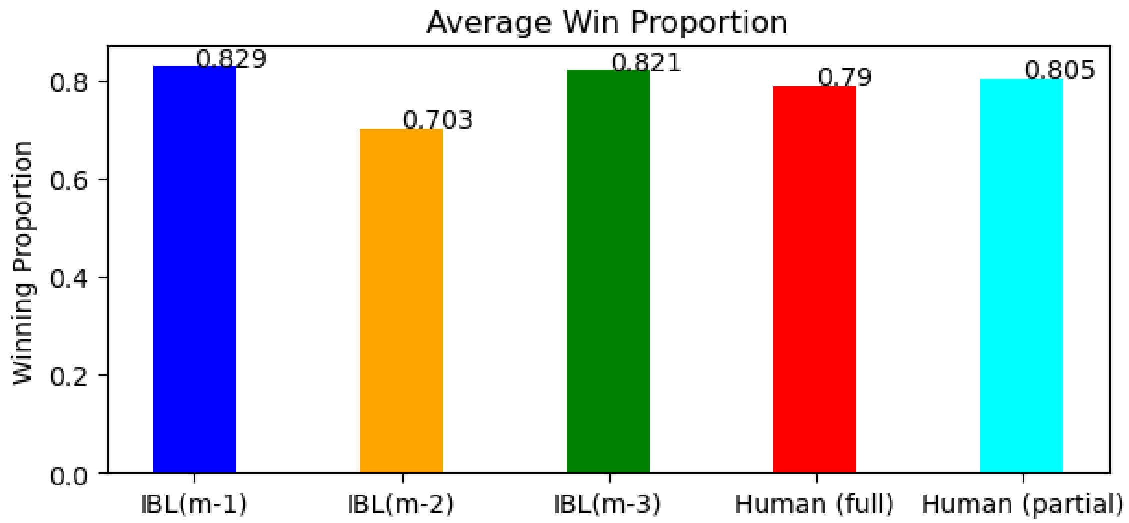

Figure 7. proportion of win

From figure 7, it is seen that the closest match with both human (full) and Human (partial) is the IBL (m-3). But IBL (m-2) and Human (full) are also in proximity.

5. Discussion

From the comparison between line graphs and proportion of attacks on each node, we can say the IBL model against strategic defender with full information has been able to capture the behavior of human attackers against the strategic defenders in full information game. The difference between the winning ratio of these two may be due to the slight difference in their payoff structure.

6.Conclusion

There are plenty of scopes to improve the models to replicate the actual scenario more closely. Model tracing can be introduced to feed the IBL attacker agents the history log of the defender's actions. Also, there is more scope for analyzing the data. Root Mean Square Error for all the nodes over all the trials can be a good measure to get the idea of a difference between IBL agents' actions and human attackers' actions. As time was short, I could not do all these detailed analyses.

There is also scope for comparing these cognitive models with game theory models. Data generated from these cognitive models can also be used to match with the behavior of different human personality types, such as -the dark triad.

7.Contribution statement

This project is fully done by Saeefa Rubaiyet Nowmi, UTEP ID: 8078256 under the supervision of Dr. Palvi Aggarwal, Assistant Professor, Computer Science , UTEP .

8. References

Basak, A., Černý, J., Gutierrez, M., Curtis, S., Kamhoua, C., Jones, D., Bošanský, B., & Kiekintveld, C. (2018). An initial study of targeted personality models in the flipit game. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11199 LNCS*, 623–636. https://doi.org/10.1007/978-3-030-01554-1_36

Dehghani Abbasi, Y., Short, M., Sinha, A., Sintov, N., Zhang, C., & Tambe, M. (2015). Human Adversaries in Opportunistic Crime Security Games: Evaluating Competing Bounded Rationality Models. In *Advances in Cognitive Systems X*.

Durkota Karel and Lisý, V. and K. C. and H. K. and B. B. and P. T. (2017). Optimal Strategies for Detecting Data Exfiltration by Internal and External Attackers. In B. and K. C. and F. F. and S. S. Rass Stefan and An (Ed.), *Decision and Game Theory for Security* (pp. 171–192). Springer International Publishing.

Kiekintveld, C., Lisý, V., & Píbil, R. (2015). Game-Theoretic Foundations for the Strategic Use of Honeypots in Network Security. *Advances in Information Security*, *56*, 81–101. https://doi.org/10.1007/978-3-319-14039-1_5

Mckelvey, R. D., & Palfrey, T. R. (1998). Quantal Response Equilibria for Extensive Form Games. *Experimental Economics*, *1*(1), 9–41. https://doi.org/10.1023/A:1009905800005

Nguyen, T. H., Yang, R., Azaria, A., Kraus, S., & Tambe, M. (n.d.). *Analyzing the Effectiveness of Adversary Modeling in Security Games*. www.aaai.org

Nochenson, A., & Grossklags, J. (n.d.). *A Behavioral Investigation of the FlipIt Game*.

Píbil, R., Lis´y, V. L., Kiekintveld, C., Bošansk´bošansk´y, B., & Pěchouček, M. (n.d.). *Game Theoretic Model of Strategic Honeypot Allocation in Computer Networks*. www.honeynet.org

Reitter, D., Grossklags, J., & Nochenson, A. (n.d.). *Risk-Seeking in a Continuous Game of Timing*.

Shiva, S., Roy, S., & Dasgupta, D. (2010). Game Theory for Cyber Security. *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*. https://doi.org/10.1145/1852666.1852704

Yang, R., Ford, B., Tambe, M., & Lemieux, A. (2014). Adaptive Resource Allocation for Wildlife Protection against Illegal Poachers. *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Yang, R., Kiekintveld, C., Ordonez, F., Tambe, M., & John, R. (n.d.). *Improving Resource Allocation Strategy Against Human Adversaries in Security Games*. http://www.ampl.com/