Towards Personalized Deceptive Signaling for Cyber Defense Using Cognitive Models

Edward A. Cranford (cranford@cmu.edu)[1], Cleotilde Gonzalez (coty@cmu.edu)[2], Palvi Aggarwal (palvia@andrew.cmu.edu)[2], Sarah Cooney (cooneys@usc.edu)[3], Milind Tambe (tambe@usc.edu)[3], and Christian Lebiere (cl@cmu.edu)[1]

[1]Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue
Pittsburgh, PA 15213 USA
[2]Dynamic Decision Making Laboratory, Social and Decision Sciences Department
Carnegie Mellon University, 5000 Forbes Avenue
Pittsburgh, PA 15213 USA
[3]USC Center for AI in Society, University of Southern California, 941 Bloom Walk
Los Angeles, CA 90089 USA

Abstract

Recent research in cybersecurity has begun to develop active defense strategies using game-theoretic optimization of the allocation of limited defenses combined with deceptive signaling. While effective, the algorithms are optimized against perfectly rational adversaries. In a laboratory experiment, we pit humans against the defense algorithm in an online game designed to simulate an insider attack scenario. Humans attack far more often than predicted under perfect rationality. Optimizing against human bounded rationality is vitally important. We propose a cognitive model based on instance-based learning theory and built in ACT-R that accurately predicts human performance and biases in the game. We show that the algorithm does not defend well, largely due to its static nature and lack of adaptation to the particular individual's actions. Thus, we propose an adaptive method of signaling that uses the cognitive model to trace an individual's experience in real time, in order to optimize defenses. We discuss the results and implications of personalized defense.

Keywords: cyber deception; cognitive models; instance-based learning; knowledge-tracing; model-tracing

Introduction

Cybersecurity often involves passive defense strategies which fail to discover a threat before major damage is done to a network. However, recent work within the domain of cybersecurity has focused on developing active defense strategies based on cognitive principles of deception (Al-Shaer et al., 2019; Cooney et al., 2019; Cranford et al., 2018). Deception is a form of persuasion where one intentionally misleads an agent into a false belief, in order to gain an advantage over the agent and achieve one's goals (Rowe & Rushi, 2016). In this line of research, the goal for security is to assist human administrators defend networks from cyber-attacks (Gonzalez et al., 2014). Limited defense resources cannot simultaneously protect all targets. In the event of an attack, truthful signals that divulge the protection status of a target can deter some attacks on protected targets. However, defenders can use a combination of truthful and deceptive signals to improve protection of the unprotected resources.

Game-theoretic principles have been employed to optimize the allocation of limited defense resources and determine how often to send a deceptive signal before it loses its effectiveness (Xu et al., 2015). While deception may reduce attacks on uncovered targets compared to no deception, the algorithms are static and tailored to an entire population. They fail to take into account the individual and their particular set of knowledge, experiences, and biases. The goal of this paper is to develop a personalized signaling strategy that can outperform traditional static methods.

Cranford et al. (2018) developed an instance-based learning (IBL) cognitive model (Gonzalez, Lerch & Lebiere, 2003) of attackers that accurately predicts human decision making from experience. We propose that such a model can be used to trace an individual's knowledge and experiences, and exploit their biases, to determine on-the-fly the best signal given the situation, to further reduce attacks.

The following section presents a line of research on game-theoretic models that have proven to optimize deceptive signaling for perfectly rational adversaries, and initial efforts toward optimizing for boundedly rational adversaries. We then describe an online game developed to investigate attacker behavior against deceptive signaling algorithms and a cognitive model that accurately predicts human behavior. Next, we describe a method for deceptive signaling that uses the cognitive model to drive adaptive signaling, personalized to the individual attacker. We highlight its applicability for optimizing defense by tracking human knowledge, experience, and biases. Finally, we discuss the implications of this line of research and avenues for future research.

Deceptive Signaling for Cybersecurity

Research on Stackelberg Security Games (SSGs) led to the development of algorithms that have greatly improved physical security systems (e.g., protecting ports, scheduling air marshals, and mitigating poachers) through the optimal allocation of limited defense resources (Pita et al., 2008; Shieh et al., 2012; Sinha et al., 2018; Tambe, 2011). Xu et al. (2015) extended these models by incorporating elements of *signaling*, in which a defender (sender) strategically reveals information about their strategy to the attacker (receiver) in order to influence the attacker's decision making (Battigalli, 2006; Cho & Kreps, 1987). Their solution, the Strong Stackelberg Equilibrium with Persuasion (peSSE), improves

defender utility against a perfectly rational attacker compared to strategies that do not use signaling. For a given target, the peSSE finds the optimal combination of bluffing (sending a deceptive message that the target is covered when it is not) and truth-telling (sending a truthful message that the target is covered) so the attacker continues to believe the bluff.

The goal of the peSSE is to reduce attacks on uncovered targets. Attackers earn a reward for successful attacks, suffer a loss for failed attacks, and earn zero for withdrawing. When a target is covered, the peSSE will always send a truthful signal. When uncovered, the peSSE will send a deceptive signal with a probability that brings the attacker's expected value of attacking, given a signal, to zero. This makes it equal to the utility of withdrawing the attack and, based on standard game-theoretic assumptions of perfect rationality, the attacker will break ties in favor of the defender and withdraw.

The peSSE is suitable for cyber defense where optimizing the probability of sending a deceptive signal can mitigate attacks on uncovered targets with little overhead. However, it is based on the assumption of perfect rationality while humans exhibit, at best, bounded rationality (Simon, 1956). To address this weakness of the peSSE, researchers have begun to develop signaling algorithms for security against boundedly rational attackers (Cooney et al., 2019). However, these algorithms do not offer substantial improvement over the peSSE in terms of reducing attacks and minimizing defender loss.

In what follows, we present an IBL cognitive model that accurately predicts human attacker behavior playing against the peSSE in a laboratory experiment. We propose that a personalized deceptive signaling scheme based on insights from the IBL model, in combination with model-tracing mechanisms, can be used to adapt defense signaling to the individual experiences of attackers at each point in time.

Cognitive Models of Human Attackers Playing Against Deceptive Signaling Algorithms

The Insider Attack Game (IAG) was designed to investigate the interaction between an attacker and defender in a cybersecurity scenario (Cranford et al., 2018). As shown in Figure 1, players take the role of the attacker (a company employee) and their goal is to score points by "hacking" computers to steal proprietary data. There are six potential computers to attack, but only two security analysts (defenders controlled by a computer algorithm) that can monitor one computer each. If the player attacks a computer that is monitored, they lose points, but if the computer is not monitored then they win points. Each computer shows its reward for winning, penalty for losing, and the probability that the computer is being monitored (reflecting the SSE for the game). On each turn, the player must select a computer to attack; after which, the signaling algorithm determines whether to send a truthful signal or a deceptive signal (with the signal, the player is presented the probability that the given signal is deceptive). The player must decide whether to continue their attack or withdraw and earn zero points. Players play four rounds of 25 trials each (after an initial 5 trials of practice). The payoff structures and monitoring probabilities of the targets are different in each round. Coverage and signaling of targets were precomputed for each trial. Therefore, each individual player experiences the same coverage and signaling schedule.

Figure1: Screenshot of the IAG. The attacker is in the center surrounded by six targets. The monitoring probability is displayed as a percentage in text and represented visually by red bars, the yellow stars represent the potential reward, and the red stars represent the potential penalty.

Attacker Cognitive Model

Cranford et al. (2018) developed an IBL cognitive model of the attacker using the ACT-R cognitive architecture (Anderson & Lebiere, 1998; Anderson et al, 2004). Following collection of human data in the peSSE condition, we modified this model to better represent human behavior playing the IAG. For brevity, details of the model described below, and its underlying equations, can be found in Cranford et al., while specific changes are footnoted.

In the current model, decisions are made by generalizing across past experiences, or instances, that are similar to the present situation. For the IAG, instances include slots to represent the context of the selected target, the decision, and the outcome. The context includes the monitoring probability [0.0, 1.0], reward [1, 10], penalty values [-1, -10], and warning signal [present, absent]. The possible decisions are attack or withdraw, and the outcome is the reward or penalty based on the decision. In a given situation, for each possible decision, an associated utility is computed through blended memory retrieval weighted by contextual similarity to past instances. The decision with the highest utility is made. In the present game there are two decisions: attack or withdraw. However, withdrawing always results in zero points. Therefore, the model only needs to determine the utility of attacking in order to make a choice.

In ACT-R, the retrieval of past instances is based on the activation strength of the relevant instance in memory and its

similarity to the current context. The activation of an instance reflects the power law of practice and forgetting, and includes a partial matching process[1] reflecting the similarity between the current context elements and the corresponding context elements for the instance in memory. A variance parameter *s* introduces stochasticity in retrieval. Similarities between numeric slot values are computed on a linear scale from 0.0, an exact match, to -1.0. Symbolic values are either an exact match or maximally different, -2.5, to prevent bleeding between memories for different actions and signal types.

A Boltzmann softmax equation[2] determines the probability of retrieving an instance based on its activation strength. The IBL model uses ACT-R's blending mechanism (Lebiere, 1999; Gonzalez et al., 2003) to calculate an expected outcome of attacking a target based on similarity to past instances. The expected outcome is the value that best satisfies the constraints of all matching instances weighted by their probability of retrieval.

In summary, the outcomes of past instances are weighted by their recency, frequency, and similarity to the current instance to produce an expected outcome. If the value is greater than zero then the model attacks, else it withdraws.

IBL Model Procedure To begin the IAG, the model is initialized with seven instances[3]: five represent a simulated practice round, and two represent knowledge gained from instructions (one instance had a signal value of *absent* and an outcome of 10, representing that attacking when a signal is absent will result in a reward; another instance had signal value of *present* and an outcome of 5, representing that attacking when a signal is present could result in either a penalty *or* a reward). On a given trial, the model first selects a target to attack. The model cycles through each target and generates an expected outcome of attacking via blending. The model selects the target with the highest expected outcome. Target selection is a passive process; therefore, no instances are saved in memory that could influence future decisions.

After selecting a target, the context is augmented with the value of the signal[4] (i.e., present or absent). The model then decides whether to attack or withdraw by generating an expected outcome via blended retrieval. The similarity of the selected target's context to past instances is based solely on the value of the signal[5] (monitoring probability, reward, and penalty values are ignored). In the IAG, the pop-up warning message covers all information about the selected target. Therefore, we inferred that humans base their decisions only on the value of the signal and ignore, or forget, the occluded target information.

After determining the expected outcome, an instance is saved in memory that represents the model's expectations[6]. Humans tend to remember not only the actual experience, but also their expectations prior to the experience (Gonzalez et al., 2003). This serves as an implementation of confirmation bias, in which one's preconception of winning/losing can increase the likelihood of attacking/withdrawing on future trials (i.e., generating positive/negative expected outcomes).

After generating an expected outcome, a decision is made, and the action and outcome slots of the current instance are updated to reflect the action taken by the model and the ground-truth outcome. This final instance is saved in memory and thereby influences future decisions.

The model continues for four rounds of 25 trials each. The model behavior reflects its experiences. If an action results in a positive/negative outcome, then its future expectations will be increased/decreased, and the model will be more/less likely to select and attack that target in the future. Also, the impact of a particular past experience on future decisions strengthens with frequency and weakens with time.

IBL Model Evaluation Against Human Players

The attacker IBL model was compared to human behavior in the IAG. In a laboratory experiment, human participants (i.e., "attackers") played against the peSSE signaling scheme. Participants were recruited via Amazon Mechanical Turk. All participants resided in the United States. For completing the experiment and submitting a completion code, participants were paid $1 plus $0.02 per point earned in the game, up to a maximum of $5.50. Four participants were removed from analysis because they had incomplete data (e.g., data recording errors) or restarted the experiment after gaining experience, resulting in a final sample size of 100.

The data was analyzed for the probability of attack and the number of points earned by attackers across rounds. The probability of attack was calculated as the proportion of players that continued the attack on a given trial. Points were separated into mean losses and gains per round. Losses/gains were calculated as the total number of points lost/gained per round by attacking targets that were/weren't monitored.

The model played the IAG 1000 times to generate stable predictions of the probability of attack and total number of points obtained per round. At the end of each run, the model was reset to its initial state and its memory cleared. Due to the stochastic nature of the model, and the influence its experiences have on its future decisions, the model behaves differently on each run and can therefore represent a diverse population of human attackers without the need to parameterize for individual differences.

[1] The mismatch penalty parameter for the activation equation was originally set high at 2.5, but was reduced to the ACT-R default 1.0.

[2] The temperature parameter was changed from the ACT-R default of $\sqrt{2} * s$ to a neutral value of 1.0 which results in retrieval probability reflecting the original presentation probability.

[3] The model was originally initialized with 8 instances representing edges of the decision space, but we believe the current method is a more accurate representation of participants' experience.

[4] Representing the deception probability as an additional context slot in the instance resulted in a poorer model fit. It appears that humans do not consider, or know how to utilize, the information. Therefore, the deception probability was excluded from the context.

[5] In the original model, the full context was used, but this resulted in an over-selection and reduced attack rate of high-reward targets.

[6] The model did not originally save this instance in memory and attacked far less often than humans. Saving this instance increased the mean probability of attack. This insight was key to understanding the biases humans have in the game and why they attacked so often.

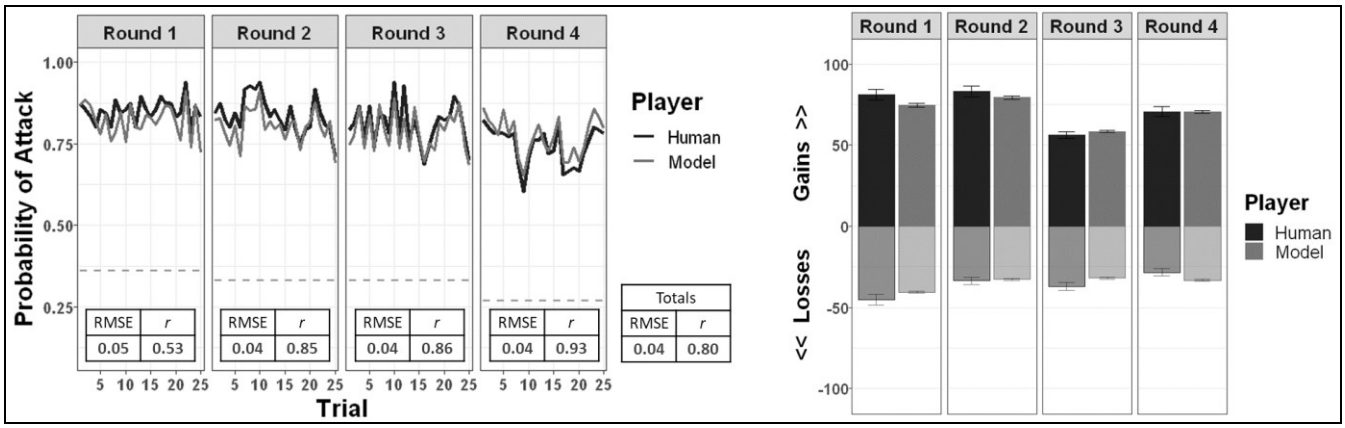

Figure 2: Probability of attack across trials and rounds (left side) and mean gains/losses per round (right side) for the humans compared to the IBL model. For probability of attack, RMSE and correlations (*r*) between human and model data are displayed under each round, and the aggregate values across the entire game are on the right under the legend.

Figure 2, left side, shows the mean probability of attack across trials and rounds for humans compared to the model. The dashed, gray line represents the peSSE predictions under assumptions of perfect rationality. Humans attack far more than perfectly rational attackers. Meanwhile, compared to the human data, the model is an excellent predictor of performance. RMSE and correlations, comparing the model to human data, are included at the bottom of the graph. The model is sensitive to the schedule of coverage, just as humans are, which produces the spiking pattern across trials.

Figure 2, right side, shows the average gains/losses for the humans compared to the model. Humans attack at a high rate, earning many points from attacks on uncovered targets, while incurring fewer losses. Moreover, the model accurately predicts this behavior. The peSSE suffers because human biases (e.g., recency, frequency, and confirmation) lead them to attack at higher rate, resulting in more experiences of wins than losses. The IBL model captures these biases, and therefore, can feasibly be used as a predictive tool for personalizing deceptive signals for an individual attacker. Notably, the model has accurately predicted human performance against other signaling algorithms (not reported here) prior to collection of human data.

Toward Personalized Deception

To personalize deception, we can run the IBL model alongside the human to predict an individual's behavior and optimize the rate of deceptive signals. To make accurate predictions of an individual, two methods have proven useful to align the model behavior with the human's decisions: model-tracing and knowledge-tracing. Model-tracing aligns the model's *actual* actions and outcomes to those *observed* of the human. Knowledge-tracing aligns the *expected* actions and outcomes to match those *inferred* of the human.

Model-tracing

Model-tracing is a method used to align a model's behavior with that of the human and is commonly used to adjust feedback provided to the student in intelligent tutoring systems (see Anderson et al., 1995). The alignment helps in a way that future model predictions are adapted and optimized to the interaction with the human. For example, geometry tutors use model-tracing to keep track of where errors are made so that the learning experience can be tailored to the individual (Anderson, Boyle, & Yost, 1986).

We use model-tracing to synchronize the IBL model with the human's *observed* actions and experience in the IAG task. After each trial, the instance saved in memory that represents the model's decision and outcome is changed to reflect the human's action and outcome (i.e., the action and outcome slots are changed to match the human's). Therefore, on the next trial the model makes predictions based on the exact experience of the human and not on what it would have done based on its own past instances. With more trials, the model is expected to make more accurate predictions of a particular human's actions, as the model's memory aligns better with that of the human. Model-tracing changes the instances representing the *observed* ground truth decision and outcome. However, in order to generate accurate predictions, we must also align the model's expectations to those of the human.

Knowledge-tracing

The model produces instances that represent the expected outcome of attacking, which contributes to confirmation bias, and these must also be changed. *Knowledge-tracing* can be used to *infer* the expectations humans had prior to making a decision that would contribute to confirmation bias. For example, if the model and human both decided to attack (or both withdraw), then nothing need change and the expected outcome generated by the model can be used to infer the human's expectation. However, if the model expects a positive outcome for attacking, but the human withdrew the attack, then we can *infer* that the human expected to lose (or vice versa). For these instances, we can modify the expected outcome slot to match the expectations of the player. We cannot infer this expectation precisely, so we set the expected outcome to either the reward or penalty of the selected target.

Model Predictions with Model & Knowledge-tracing

To test the effectiveness of model- and knowledge-tracing for predicting human decision making, the model was run alongside human data in the peSSE condition. On each trial, the model simply makes a decision, which is recorded, and is then updated via model-tracing and knowledge-tracing. The model decision was then compared to the decision the human made to generate a probability of agreement between the model and human. The mean probability of agreement for rounds 1-4 are 86.4% (*SD* = 12.3%), 90.8% (*SD* = 11.4%), 89.6% (*SD* = 12.4 %), and 86.8% (*SD* = 15.5%), respectively.

The trial-to-trial agreement is highly accurate, just short of accounting for the entirety of human stochasticity. In fact, even at the 1st trial the model is accurate to 83.3%. Moreover, the model adapts well to the individual's probability of attack. Figure 3 shows the overall probability of attack of individual model runs compared to the human it traced. The model is exceptionally accurate in adapting to the human, $r^2 = 0.95$. Using techniques of model-tracing and knowledge-tracing, the model makes very accurate predictions and could feasibly be used in designing a personalized signaling scheme.

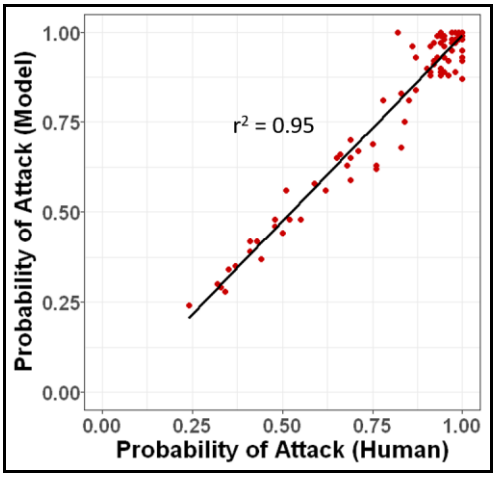

Figure 3: Overall mean probability of attack comparing individual humans to the model run that traced him/her, in the peSSE condition using personalized signaling.

A Personalized Deceptive Signaling Scheme

The peSSE signaling scheme uses deceptive signals on uncovered targets but not on covered targets. These schemes invite attacks with impunity when no signal is given. Therefore, a broader and more symmetrical approach may be warranted, as has been explored in recent game-theoretic research (Cooney et al., 2019). The following signaling scheme also uses deceptive signals when a target is covered.

If the goal is to minimize the probability of attack as a function of the warning signal then it can be shown that we must reach an equilibrium where the probability of attack given a warning, P(A|W), is equal to the probability of attack given no warning, P(A|NW). A signal must be generated at a rate that preserves this equality. We can examine the impact of the presence or absence of a signal in various situations. For example, given an attack, if a target is covered, the attacker will lose, and their future probability of attack will be lower. If a target is uncovered, the attacker will win, and their future probability of attack will be higher. Each outcome thus increases or decreases one of the attack probabilities. In particular, the change in attack probability (decrease or increase) is determined by whether the selected target is covered or not, respectively, while the probability impacted (signal or no signal) is determined by the presence or absence of a signal, respectively. This results in the following algorithm for signal generation: if the selected target is covered, if P(A|W) is greater than the P(A|NW) then generate a signal, otherwise do not generate a signal; but if the selected target is not covered, if P(A|W) is greater than the P(A|NW) then do not generate a signal, otherwise generate a signal.

The role of the cognitive model in this algorithm is to determine P(A|W) and P(A|NW). We know the model generates expected outcomes of attacking and decides to attack if the value is greater than zero. Therefore, we can simply generate the expected outcome of attacking given the presence or absence of a signal and compare them to compute the conditions used in the algorithm above. An essential point is that those expected values are not the true expected values, but the model's subjective expected value given its limited experience and its reflection of human cognitive biases.

Intuitively, if the selected target is covered, then we decide on whether to generate a signal or not depending on which condition is most likely to lead to an attack. This corresponds to trying to catch the attacker when the target is covered, lowering the future probability of attack. Conversely, if the selected target is not covered, select the condition (signal or not) least likely to lead to an attack. Again, the accuracy of the cognitive model is essential in this approach to capture the subject's intention to attack or not. We can use the current model to track an individual's decisions and generate predictions of their probability of attack given the situation.

Effectiveness of Personalized Signaling Scheme

To generate predictions of the effectiveness of this personalized signaling scheme, we ran the IBL model through the IAG while using the personalized signaling scheme described above to make predictions about the expected outcome of attacking, given a signal and given no signal. Based on those predictions and the underlying coverage of the selected target, the scheme determined whether or not to give a signal on each trial.

Compared to the human performance in peSSE, the personalized signaling method is expected to reduce the probability of attack by an average of 2.7% (RMSE = 6.6%). Meanwhile, Figure 4 shows that personalized signaling will result in fewer gains and more losses. Looking further into the data, Figure 5 plots the probability of attack across the various targets, based on their monitoring probability. Compared to human performance, the personalized signaling method seems to shift the distribution of attacking toward targets with a higher monitoring probability, and therefore the IBL model incurs more penalties.

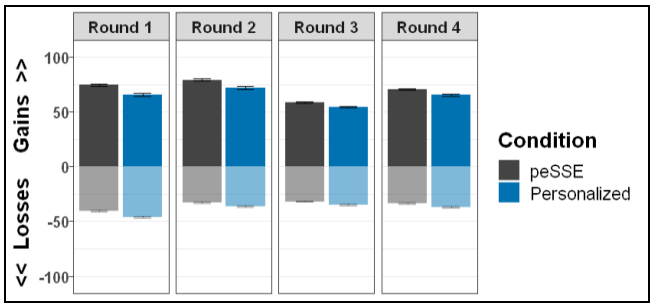

Figure 4: Comparing the mean gains/losses across rounds in the personalized signaling model to humans in the peSSE.

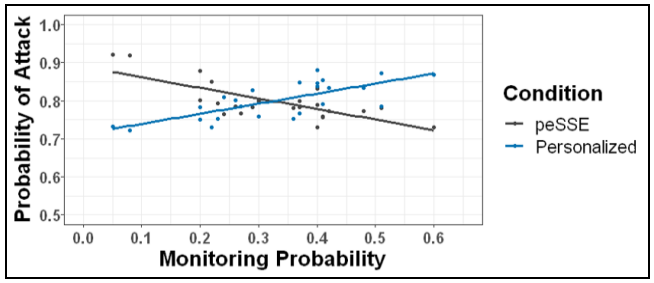

Figure 5: Mean probability of attack across targets, by their monitoring probability, comparing the personalized signaling model to humans in the peSSE condition.

Conclusions

The present research shows that we can leverage the predictive power of a generalizable IBL model to infer an individual's knowledge, trace their experience, and exploit their biases to design an adaptive signaling scheme that is personalized for an individual. The current method is an initial attempt toward developing a personalized deceptive signaling scheme for cyber defense. Although the current scheme did not greatly reduce the probability of attack, the cognitive model proved to be an accurate predictor of human behavior. Future research will test the personalized signaling scheme against human attackers. Insight gained from human experiments will provide information about how to modify the signaling logic to create a more effective scheme.

Acknowledgments

This research was sponsored by the Army Research Office and accomplished under Grant Number W911NF-17-1-0370.

References

Al-Shaer, E., Wei, J., Hamlen, K. W., & Wang, C. (2019). Dynamic Bayesian Games for Adversarial and Defensive Cyber Deception. In *Autonomous Cyber Deception*. Cham: Springer.

Anderson, J. R., Boyle, C. F., & Yost, G. (1986). The geometry tutor. *Journal of Mathematical Behavior*, 5-20.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences*, *4*, 167-207.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036-1060.

Battigalli, P. (2006). Rationalization in signaling games: Theory and applications. *International Game Theory Review 8,* 01, 67–93.

Cranford, E. A., Lebiere, C., Gonzalez, C., Cooney, S., Vayanos, P., & Tambe, M. (2018). Learning about cyber deception through simulations: Predictions of human decision making with deceptive signals in Stackelberg Security Games. *Proceedings of the 40th annual conference of the Cognitive Science Society* (pp.258-263). Madison, WI: Cognitive Science Society.

Cho, I.-K. & Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics, 102*(2), 179–221.

Cooney, S., Vayanos, P., Nguyen, T. H., Gonzalez, C., Lebiere, C., Cranford, E. A., & Tambe, M. (2019). Warning time: optimizing strategic signaling for security against boundedly rational adversaries. *Proceedings of the 18th AAMAS.* Montreal, Canada: IFAAMS.

Gonzalez, C., Ben-Asher, N., Oltramari, A., Lebiere, C. (2014). Cognition and Technology. In Kott, C., Wang, A. & R. Erbacher (eds.), *Cyber defense and situational awareness*. Switzerland: Springer International Publishing.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance based learning in dynamic decision making. *Cognitive Science, 27*(4), 591-635.

Lebiere, C. (1999). A blending process for aggregate retrievals. *Proceedings of the 6th ACT-R Workshop*. George Mason University, Fairfax, Va.

Pita, J., Jain, M., Ordónez, F., Portway, C., Tambe, M., Western, C., & Kraus, S. (2008). ARMOR Security for Los Angeles International Airport. *Proceeding of the 23rd AAAI conference on artificial intelligence* (pp. 1884-1885). Menlo Park, CA: AAAI Press.

Rowe, N. C., & Rrushi, J. (2016). *Introduction to Cyberdeception*. Switzerland: Springer.

Shieh, E., An, B., Yang, R., Tambe, M., Baldwin, C., & Meyer, G. (2012). Protect: A deployed game theoretic system to protect the ports of the United States. *Proceedings of the 11th AAMAS* (pp. 13-20). IFAAMS.

Simon, H. A. (1956). Rational choice and the structure of the environment. Psychological Review, 63(2), 129-138.

Sinha, A., Fang, F., An, B., Kiekintveld, C., & Tambe, M. (2018). Stackelberg Security games: Looking beyond a decade of success. *Proceedings of the 27th IJCAI* (pp. 5494–5501). IJCAI.

Tambe, M. (2011). *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned.* Cambridge University Press.

Xu, H., Rabinovich, Z., Dughmi, S., & Tambe, M. (2015). Exploring information asymmetry in two-stage security games. *Proceedings of the National Conference on Artificial Intelligence* (2, pp. 1057-1063). Elsevier B.V.