Turing-like Experiment in a Cyber Defense Game

Yinuo Du[1], **Baptiste Prebot**[1], **Cleotilde Gonzalez**[1]

[1]Carnegie Mellon University
yinuod@cmu.edu, baptiste.prebot@ensc.fr, coty@cmu.edu

Abstract

This work aims at advancing the development of cognitive agents that learn and make defense dynamic decisions during cyber attacks and to evaluate the capability of these agents in "Turing-like" experiments, comparing the decisions and performance of these agents against human cyber defenders. We present an initial demonstration of a cognitive model of a defender that relies on a cognitive theory of dynamic decision making, Instance-Based Learning Theory (IBLT) and we demonstrate the execution of the same defense task by human defenders. We rely on OpenAI Gym, CybORG, and adapt an existing scenario (i.e., CAGE) to generate a simulation experiment using an IBL defender. We also offer a new Interactive Defense Game (IDG), where *human* defenders can perform the same CAGE scenario simulated with the IBL model. Our results suggest that the IBL model makes decisions against intelligent attack agents in a way similar to that observed in a human experiment. We conclude with a description of the cognitive foundations required to build intelligent autonomous cyber defense agents that can collaborate with humans in autonomous cyber defense teams.

Introduction

The cyber battlefield of the future will certainly see autonomous systems fight other autonomous systems (Kott 2018). These autonomous systems, characterized by some degree of freedom in decision making and action, will need to operate in uncertain and complex environments (David and Nielsen 2016). To achieve this goal, research must address two major challenges: (1) develop intelligent defense systems that are able to learn and understand the dynamic strategies of attackers to efficiently anticipate and counter their decisions, and (2) evaluate the ability of these intelligent defense systems to produce defense behaviors that are comparable to those of expert cyber defenders (Vieane et al. 2016; Dhir et al. 2021; Kott et al. 2020).

Human cognition and our ability to computationally represent the dynamic decision-making process of a cyber analyst are key for the future of cyber security (Gonzalez et al. 2014; Kott et al. 2020). In the past decade, cognitive models have been developed in the context of cyber security to represent human defense decisions (Dutt, Ahn, and Gonzalez 2011), human attacker decisions that can inform cyber defense strategies (Cranford et al. 2020a,b; Gonzalez et al. 2020), and end-user phishing classification decisions that can help improve cyber defense (Cranford et al. 2019). All these models are based on the well-known cognitive theory of dynamic decision making, Instance-Based Learning Theory (IBLT) (Gonzalez, Lerch, and Lebiere 2003). IBLT is a comprehensive account of how humans make decisions based on experience during dynamic tasks, and has been used to represent the dynamic decision-making process in cyber security and many other domains (Gonzalez 2022). For example, an IBL model first recognizes cyber events (e.g., execution of a file on a server) in the network based on the attributes of the situation and the similarity of the attributes of the events to past experiences (instances) stored in the memory of the analyst (Dutt, Ahn, and Gonzalez 2011). Then, the model reasons about whether a sequence of observed events is a cyber attack or not, based upon instances retrieved from memory and the risk tolerance of a simulated analyst. Execution of the IBL model generates predictions of the analyst's decisions that are evaluated on the basis of their timeliness.

Initial work that created autonomous cyber defenders (Dutt, Ahn, and Gonzalez 2011) was not evaluated against human cyber defenders. Additionally, the cybersecurity scenarios were quite simple, where an attacker would attempt to access a company's server indirectly through a web server. In the current work, we advance this initial work with the following contributions: (1) present an IBL cognitive model of the dynamic decision process of cyber defense in a complex OpenAI gym, called CybORG (Baillie et al. 2020), and a challenging cyber attack scenario against two different attack strategies (Standen et al. 2021); (2) develop a new Interactive Defense Game (IDG) that integrates the attack scenario into an interactive tool where human defenders confront the same attackers in CybORG; (3) provide simulation results of the performance of the IBL model against the two attack strategies; and (4) evaluate the capabilities of the IBL models against human performance in the same scenarios using IDG. Our results suggest that the IBL model can make reasonable predictions regarding defense behavior against the two attacks strategies and that these predictions are similar to the behavior observed in human defenders.

CybORG: An Autonomous Cyber Operations Gym

The CybORG AI Gym (Baillie et al. 2020) was initially developed and released as an experimental simulation platform to train reinforcement learning agents for cyber defense. It uses the OpenAI Gym interface (Brockman et al. 2016) together with a wide range of scenarios cyber-operation adversarial scenarios in a realistic but simple training environment.

For the present work, we adapted the Cage Challenge scenario (Standen et al. 2021) (i.e., CAGE) that was implemented in CybORG. In CAGE, a defense agent (blue agent) is assigned to defend a network against an attacker (red agent). Green agents can also be included to simulate normal activity generated by regular users performing *Scans* on a system. These three types of agents interact with the scenario alternatively by performing high-level actions in a game-like episode with a fixed number of steps. Red agents can choose to do reconnaissance, exploitation, privilege escalation, and pivoting. Blue agents are enabled to conduct network monitoring, host analysis, malicious code removal, and system recovery from backups. From the CAGE scenario, we were able to manipulate different aspects of the simulation: the type of attacker, and the presence or absence of the 'normal' activity of the Green agent and the size of the network, as illustrated in Figure 1.

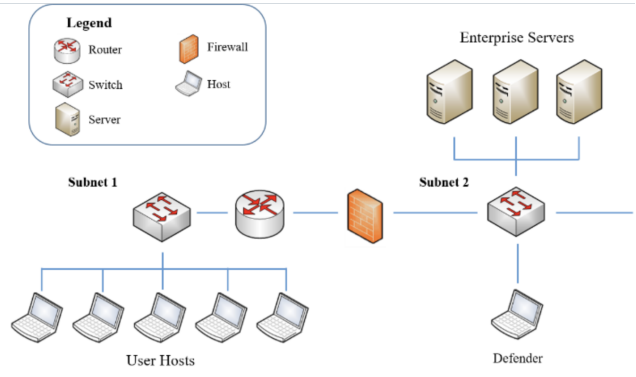

Figure 1: Example of a network created in the CybORG scenario

The type of attacker can represent different deterministic strategies. Two of them, *Beeline* and *Meander*, were provided in the initial Cage Challenge scenario, to resemble the diversity of attackers and to evaluate the capability of the defender agent more comprehensively. They differentiate by their prior knowledge of the network and the way they route through the network accordingly. Beeline can be assimilated to an agent performing a blitz, targeted attack, while Meander is seeking to gain privileged access on all hosts, stealthy establishing a long-term presence in the network. Although these are two types of attackers provided in the CybORG scenario, other AI models of attackers can also be plugged in, to simulate more dynamic, nondeterministic, realistic attackers and to study the adaptation capacities of defenders.

On the defense side, the simulation environment allows for testing of multiple IBL agents, built with different instances structures or based on different cognitive models, as well as other types of AI agents, for example, RL agents, used to compare their performance in predicting human behavior. However, in order to do so, an interactive interface dedicated to human testing had to be developed.

In terms of experimental settings, the number of episodes to be simulated and their duration (i.e., the number of steps of each episode) can also be manipulated. For initial experiments, the scenario runs on 25 steps-long episodes on a small network.

Interactive Defense Game

We created a new Interactive Defense Game (IDG) which is a Django-based web application. IDG offers a web-based graphical user interface to allow human participants to perform the task proposed in our adapted CAGE scenario (Prebot, Du, and Gonzalez 2023).

The IDG interface shown in Figure 2 consists of a central interactive table representation of the network and the related information on each host or server: IP Address, name, subnet, last detected activity, and compromised level. In this task, a human defender can select from a set of actions represented in the buttons on the bottom right of the screen: *Monitor, Analyze, Remove, Restore*.

Human defenders can select a host by clicking on its row in the table and then choose one of the four actions to perform on that particular host. Then by clicking on the "Next" button, the action selected takes effect, and the defender can see the result (i.e. points lost) from the execution of that action in the "Last round" value. A new and updated version of the environment is presented to the human defender, demonstrating the status (activity and compromised levels) of the network elements. The "Last round" outcome provides immediate feedback regarding the effectiveness of the past action, and the "Total loss" presents the human defender with a cumulative account of the loss during the episode.

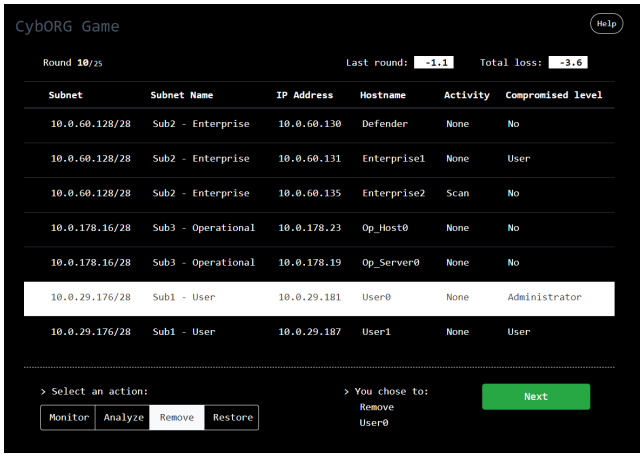

Figure 2: Interactive Defense Game user interface.

IBL Model of Blue Agents

In IBLT, an "instance" is a memory unit that results from the potential alternatives evaluated. These memory representations consist of three elements that are constructed over time: a situation state that is composed of a set of characteristics; a decision or action taken corresponding to an alternative in state; and an expected utility or experienced result of the action taken in a state. Concretely, for an IBL agent, an option is defined by the action in the state. At each time different instances are considered for each option. Each instance in memory has a value *Activation*, which represents the ease with which that information is available in memory (Anderson and Lebiere 2014).

The activation includes two parameters; a decay and noise parameters. The activation of an instance is used to determine the probability of retrieving an instance from memory and the expected utility of option is calculated based on *Blending*, which is a form of expected value, including the probability and the outcome of all instances for each option. The choice rule is to select the option that corresponds to the maximum blended value. When the agent receives delayed results, the agent updates expected utilities using a credit assignment mechanism (Nguyen, McDonald, and Gonzalez 2021). Since IBLT's process and mechanisms have been widely published, and they are common regardless of the particular task for which the models are developed, we do not repeat the mathematical formulations of the theory here given space restrictions. Please refer to (Nguyen, Phan, and Gonzalez 2023).

We developed an IBL cognitive model of cyber defense and demonstrate the predictions of the model in the scenario described above against Beeline and Meander (Du et al. 2022). The contextual features in the instances of the model, are constructed to resemble the information that would be presented to a human defender in the scenario. Specifically, there are two slots for each host or server, representing the observed activity and the known compromised status of that host at a certain step in an episode. The order of *(Activity, Compromised Status)* pairs for each host is fixed to encode the identity of each host, that is, the *Host name*, *IP address* and *Subnet*. The *Step Index* slot is included to resemble the step counter within each episode. The decision is for the IBL agent to choose a host to protect and the tool to protect it with. Each action consists of a host and a command in the format of *cmd host*. The model makes decisions by storing the instances in memory and following the general mechanisms of IBLT described above. The complete description of this model and the simulation results are presented in (Du et al. 2022).

Simulation and Human Experiments

In the simulation experiment, we designed four conditions involving different strategies of the red agent and the presence or absence of green agents. The Beeline Red agent without Green agents; the Beeline Red agent in the presence of Green agents; the Meander Red agent without Green agents; and the Meander Red agent in the presence of Green agents. We ran 40 IBL simulated defenders, each in 2000 episodes of 25 steps in each condition.

As an illustration, Figure 3 shows the performance and learning over the 2000 episodes for the Blue IBL agents, when confronted with the two Red agents, Beeline and Meander. These results are obtained in the absence of Green agents. The complete set of results of these simulations is shown in (Du et al. 2022). The results are shown in terms of the loss suffered by the IBL agents during the execution of the scenario in each episode. As observed, the IBL agents confronted with the Beeline Red agents show a larger loss initially but are able to learn over episodes. The IBL agents confronted against the Meander Red agents, suffer a less severe loss initially, and again are able to learn by reducing such loss over episodes. An experiment was conducted in-

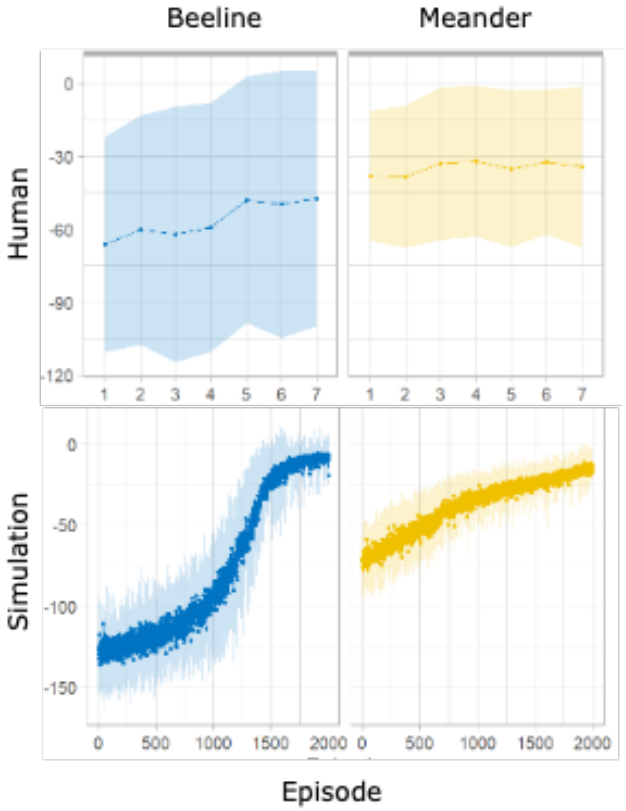

Figure 3: Top: Human defender average Loss against two different attacker strategies. Bottom: IBL agents average Loss against two different attackers strategies.

volving human defenders performing the same scenario as in the simulation experiment. The human experiment involved participants who defend the small network shown in Figure 1, against one of the two types of attackers, Beeline or Meander, using IDG. 120 subjects carried out the task of defending the network against one of the two attackers, with the goal of minimizing the loss of defense points. In the scenario, there were no Green agents, each participant performed 7 episodes of 25 steps in each episode. Such human data provide some initial evaluations of the IBL model of de-

fense. As observed in the figures above, human participants in the IDG perform worse against the Beeline than the Meander attacker. But also, just like IBL models, humans are able to learn and improve their performance against Beeline more than against Meander attackers.

Conclusions and Future Work

Cognitive models can help emulate the behavior of defenders, attackers, and users, allowing them to inform game-theoretic, Machine Learning, optimization algorithms, and other advanced technologies with predictions about human dynamic decision-making. This process has been demonstrated in recent work and provides great potential for the future of autonomous cyber defense (Aggarwal et al. 2022)

Simulation environments such as CybORG will need to be advanced and used in experimentation. We developed a research environment where IBL defense agents can perform a cyber defense task. Using the IBL model, we conducted a simulation experiment to test the performance of the IBL model against the two attacker agents of the adapted CAGE scenario. We also developed an Interactive Defense Game for human defenders, which shows similar learning trends with higher efficiency as the cognitive models on average. In future research, we will go beyond the aggregate level and investigate the model's capability to predict each individual's decision by model tracing.

References

Aggarwal, P.; Thakoor, O.; Jabbari, S.; Cranford, E. A.; Lebiere, C.; Tambe, M.; and Gonzalez, C. 2022. Designing Effective Masking Strategies for Cyberdefense through Human Experimentation and Cognitive Models. *Computers & Security*, 102671.

Anderson, J. R.; and Lebiere, C. J. 2014. *The atomic components of thought*. Psychology Press.

Baillie, C.; Standen, M.; Schwartz, J.; Docking, M.; Bowman, D.; and Kim, J. 2020. Cyborg: An autonomous cyber operations research gym. *arXiv preprint arXiv:2002.10667*.

Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; and Zaremba, W. 2016. OpenAI Gym.

Cranford, E.; Gonzalez, C.; Aggarwal, P.; Cooney, S.; Tambe, M.; and Lebiere, C. 2020a. Adaptive cyber deception: Cognitively informed signaling for cyber defense. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Cranford, E. A.; Gonzalez, C.; Aggarwal, P.; Tambe, M.; and Lebiere, C. 2020b. What Attackers Know and What They Have to Lose: Framing Effects on Cyber-attacker Decision Making. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, 456–460. SAGE Publications Sage CA: Los Angeles, CA.

Cranford, E. A.; Lebiere, C.; Rajivan, P.; Aggarwal, P.; and Gonzalez, C. 2019. Modeling cognitive dynamics in (End)-user response to phishing emails. *Proceedings of the 17th ICCM*.

David, R. A.; and Nielsen, P. 2016. Defense science board summer study on autonomy. Technical report, Defense Science Board Washington United States.

Dhir, N.; Hoeltgebaum, H.; Adams, N.; Briers, M.; Burke, A.; and Jones, P. 2021. Prospective artificial intelligence approaches for active cyber defence.

Du, Y.; Prebot, B.; Xi, X.; and Gonzalez, C. 2022. Towards Autonomous Cyber Defense: Predictions from a cognitive model. *Under review*.

Dutt, V.; Ahn, Y.-S.; and Gonzalez, C. 2011. Cyber Situation Awareness: Modeling the Security Analyst in a Cyber-Attack Scenario through Instance-Based Learning. In Li, Y., ed., *Data and Applications Security and Privacy XXV*, 280–292. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-642-22348-8.

Gonzalez, C. 2022. Learning and dynamic decision making. *Topics in Cognitive Science*.

Gonzalez, C.; Aggarwal, P.; Lebiere, C.; and Cranford, E. 2020. Design of Dynamic and Personalized Deception: A Research Framework and New Insights. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Gonzalez, C.; Ben-Asher, N.; Oltramari, A.; and Lebiere, C. 2014. Cognition and technology. In *Cyber defense and situational awareness*, 93–117. Springer.

Gonzalez, C.; Lerch, F. J.; and Lebiere, C. 2003. Instance-based learning in dynamic decision making. *Cogn. Sci.*, 27: 591–635.

Kott, A. 2018. Bonware to the Rescue: the Future Autonomous Cyber Defense Agents, a keynote at the Conference on Applied Machine Learning for Information Security. Washington DC.

Kott, A.; Théron, P.; Mancini, L. V.; Dushku, E.; Panico, A.; Drašar, M.; LeBlanc, B.; Losiewicz, P.; Guarino, A.; Pihelgas, M.; et al. 2020. An introductory preview of Autonomous Intelligent Cyber-defense Agent reference architecture, release 2.0. *The Journal of Defense Modeling and Simulation*, 17(1): 51–54.

Nguyen, T. N.; McDonald, C.; and Gonzalez, C. 2021. Credit Assignment: Challenges and Opportunities in Developing Human-like AI Agents. Technical report, Carnegie Mellon University.

Nguyen, T. N.; Phan, D. N.; and Gonzalez, C. 2023. SpeedyIBL: A comprehensive, precise, and fast implementation of instance-based learning theory. *Behavior Research Methods*, 55(4): 1734–1757.

Prebot, B.; Du, Y.; and Gonzalez, C. 2023. Learning About Simulated Adversaries from Human Defenders using Interactive Cyber-Defense Games. *arXiv preprint arXiv:2304.01142*.

Standen, M.; Lucas, M.; Bowman, D.; Richer, T. J.; Kim, J.; and Marriott, D. 2021. CAGE Challenge 1. In *IJCAI-21 1st International Workshop on Adaptive Cyber Defense.* arXiv.

Vieane, A.; Funke, G.; Gutzwiller, R.; Mancuso, V.; Sawyer, B.; and Wickens, C. 2016. Addressing Human Factors Gaps in Cyber Defense. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1): 770–773.