



Contents lists available at ScienceDirect

## Computers &amp; Security

journal homepage: [www.elsevier.com/locate/cose](http://www.elsevier.com/locate/cose)

# Designing effective masking strategies for cyberdefense through human experimentation and cognitive models<sup>☆</sup>

Palvi Aggarwal<sup>a,\*</sup>, Omkar Thakoor<sup>b</sup>, Shahin Jabbari<sup>c</sup>, Edward A. Cranford<sup>a</sup>,  
Christian Lebiere<sup>a</sup>, Milind Tambe<sup>c</sup>, Cleotilde Gonzalez<sup>a</sup>

<sup>a</sup> Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>b</sup> University of Southern California, Los Angeles, CA 90007, USA

<sup>c</sup> Harvard University, Cambridge, MA 02138, USA

## ARTICLE INFO

## Article history:

Received 8 April 2021

Revised 10 February 2022

Accepted 23 February 2022

Available online 10 March 2022

## Keywords:

Cybersecurity

Masking

Human experiments

Cognitive models

Game-theory

Decision making

## ABSTRACT

Masking strategies for cyberdefense (i.e., disguising network attributes to hide the real state of the network) are predicted to be effective in simulated experiments. However, it is unclear how effective they are against human attackers. We address three factors that challenge the effectiveness of the masking strategies in practice: (1) we relax the assumption of rationality of the attackers made by Game Theory/Machine Learning defense algorithms; (2) we provide a cognitive model of human attackers that can inform these defense algorithms; and (3) we provide a way to generate data on attacker's decisions through simulation with a cognitive model. Two masking strategies of defense were generated using Game Theory and Machine Learning (ML) algorithms. The effectiveness of these two masking strategies of defense, *risk averse* and *rational*, are compared in an experiment with human attackers. We collected attacker's decisions against the two masking strategies. With the limited human participant's data, the results indicate that the risk averse strategy can reduce the defense losses compared to the rational masking strategy. We also propose a cognitive model based on *Instance-Based Learning Theory* that accurately represents and predicts the attacker's decisions in this task. We demonstrate the model's process by generating simulated data and comparing it to the attacker's actual actions in the experiment. The model is able to capture the data at the aggregate and at the individual levels of attackers making decisions in both rational and risk averse defense algorithms. We propose that this model can be used to inform game theoretic defense algorithms and to produce synthetic data that can be used by ML algorithms to generate new defense strategies.

© 2022 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

The growth of cybercrime has increased the interest in designing effective cyberdefense strategies using game-theory and Machine Learning (ML) approaches (Goel and Perlroth, 2016; Gutzmer, 2017). One cyberdefense strategy is deception (i.e., planned actions

taken to mislead attackers for taking, or not taking certain actions Cohen, 1998). *Masking* is a cyberdeception strategy used to camouflage the network attributes to conceal information that can be confiscated by the attackers during the reconnaissance phase (De Gaspari et al., 2016; Ferguson-Walter et al., 2017; Heckman et al., 2013; Thinkst, 2015). To date most research on masking strategies has been either theoretical or tested only in simulations. Thus, it is unclear whether such defense strategies would be effective in practice, against *human* attackers. In fact, in a recent study, we found that a masking strategy that appeared successful in theory, was ineffective against human attackers: it was not better than a random camouflage strategy (Aggarwal et al., 2020b).

One possible explanation for the current results in masking strategies is the assumption of "rationality" of human attackers made by these algorithms. Generally, humans are limited cogni-

<sup>☆</sup> This research was sponsored by the Army Research Office and accomplished under MURI Grant Number W911NF-17-1-0370 and by the Army Research Laboratory under Cooperative Agreement Number W911NF-13-2-0045 (ARL Cyber Security CRA). The authors thank Jeffrey Flagg from the DDMLab at Carnegie Mellon University and Perspecta Labs for their help in conducting this experiment.

\* Corresponding author.

E-mail addresses: [pagggarwal@utep.edu](mailto:pagggarwal@utep.edu) (P. Aggarwal), [othakoor@usc.edu](mailto:othakoor@usc.edu) (O. Thakoor), [jabbari@seas.harvard.edu](mailto:jabbari@seas.harvard.edu) (S. Jabbari), [cranford@cmu.edu](mailto:cranford@cmu.edu) (E.A. Cranford), [cl@cmu.edu](mailto:cl@cmu.edu) (C. Lebiere), [milind\\_tambe@harvard.edu](mailto:milind_tambe@harvard.edu) (M. Tambe), [coty@cmu.edu](mailto:coty@cmu.edu) (C. Gonzalez).

tively in various ways, and they can only be *boundedly rational* (Simon, 1956; Tversky and Kahneman, 1979). Humans are limited in their memory and engage in sequential processing of information, which often results in decision making biases (Lemay and Leblanc, 2018; Sawyer and Hancock, 2018). Attackers maybe vulnerable to such biases and thus make cybersecurity-relevant mistakes. To illustrate, Gutzwiller et al. (2018) used oppositional human factors to exploit biases and deficiencies, related to limited attention, to disrupt cyberattacks. Gutzwiller et al. (2018) observed various kinds of biases including illusion of control, sunk cost fallacy, irrational escalation and attentional tunneling. Similarly, among cybersecurity experts, Gutzwiller et al. (2019) observed decision making biases such as anchoring bias, confirmation bias, and take-the-best heuristic bias. Unfortunately, current defense algorithms ignore the biases generated by human memory, instead of *exploiting* them to the benefit of cyberdefense. In addition, cyberdefense algorithms also ignore defender's biases which may become a bottleneck in their defense actions. In a network defense scenario, Bos et al. (2016) demonstrated the effect of gain and loss framing biases on defenders decisions. Defenders that began with gain framing (i.e. with a network already in quarantine) used a quarantine system more comparable to those that started in loss framing. To date, There has not been much work in how to mitigate such biases in defenders. On the attacker's side, Cranford et al. (2020) have demonstrated how defense algorithms can take advantage of biases (e.g., confirmation bias) in human attackers with the use of cognitive models that emulate the attacker's decision process computationally. Using a simple task, Cranford et al. (2020) have shown that it is possible to provide information about the attacker's behavior to the defense algorithms and improve the game theory/ML algorithms by making them more adaptive to the individual attacker's actions.

In this paper, we advance prior work in cyberdeception (i.e., masking techniques) by addressing two factors that limit the progress on the design of effective masking strategies against human attackers. First, we relax the assumption of attacker rationality that most ML and game theory approaches of defense make (Alpcan and Başar, 2010; Laszka et al., 2015; Schlenker et al., 2017; Serra et al., 2015). Assuming that humans will choose the best option available, in terms of expected values, is problematic, as psychologists have known for decades that humans can only be *boundedly rational* (Kahneman, 2003; Simon, 1956) and act according to simple heuristics (Gigerenzer and Todd, 1999). This was demonstrated recently in a human-subject experiment that evaluated an optimal defense strategy (proposed by Schlenker et al. (2018)) compared to a random strategy of masking (Aggarwal et al., 2020b). Their findings showed that the optimal strategy, which was theoretically most effective, only slightly reduced attacker's outcome compared to a random masking strategy (reduced by 10% whereas simulations predicted 20% reduction). The analysis by Aggarwal et al. (2020b) suggests that attackers acted in agreement with *risk aversion*, a form of boundedly-rational behavior, where humans appeared to attack machines with low rewards and high probability of success. This human attacker data was used to develop a new "risk averse" algorithm Thakoor et al. (2020). In this paper, we examine human attacker behavior in a new human-in-the-loop experiment, comparing the new risk averse masking strategy Thakoor et al. (2020) to a rational masking strategy proposed by Schlenker et al. (2018).

In addition, we demonstrate a strategy to improve game theory and ML defense algorithms by providing large amounts of data from a well-calibrated cognitive model of attackers' behavior. To make an accurate estimation of the parameters required by Thakoor et al. (2020)'s algorithm, large amounts of human data are required. Unfortunately, ML models may learn inaccurate estimates

of the parameters of the attacker model without sufficient human attacker data. To address this challenge, we developed an Instance-Based Learning (IBL) model (Gonzalez et al., 2003), that represents the process by which attackers make decisions and predicts the attacker's actions in a cyber attack situation. In this paper, we present the process by which human data collection can be used to calibrate the parameters of game theory and ML algorithms; we develop an IBL model of the human attacker to demonstrate the capability of the cognitive model to emulate the attacker's actions collected in a human experiment. The results suggest that this model can inform the adaptive cyber defense algorithms and may be used to generate large amounts of synthetic data regarding the attacker's actions to improve ML-based boundedly rational masking algorithms.

## 2. Background

Gonzalez et al. (2020) proposed a research framework for generating dynamic, adaptive, and personalized defense strategies using cognitive models. In this framework, game-theory defense algorithms are developed and deployed in experimental testbeds. An experimental testbed is used with human participants (e.g., attackers) for evaluating the performance (i.e., defender's utility) of defense algorithms. Importantly, cognitive models are used for emulating human decisions to inform the game-theory algorithms for adaptive defense. This general idea of adaptive cyberdefense based on cognitive models has been used by Cranford et al. (2020) to demonstrate the generation of adaptive and personalized signals in a simple insider attack game. In this paper, we leverage this work by first deploying defense strategies developed by game-theory/ML algorithms on an experimental testbed, CyberVAN (Chadha et al., 2016), conducting human experiments to evaluate the performance of the algorithms, and developing a cognitive model to simulate the attackers' decisions in a complex cyberdeception scenario.

In cyber camouflage games (Schlenker et al., 2018; Thakoor et al., 2019a), game theoretic models determines how the defender can mask the configurations of the machines to create uncertainty in an attacker's potential rewards. Almost all such models assume that the attacked machine is guaranteed to provide utility to the attacker. Furthermore, most of these models assume a rational attacker. However, these assumptions do not hold in practice (Chicoisne and Ordóñez, 2016; Cooney et al., 2019a). To address the issue of rationality assumptions in Stackelberg security games, Yang et al. (2011) developed optimal strategies against Prospect Theory models (Kahneman and Tversky, 1979). However, their model relies on using parameters from previous literature, ignoring the fact that model parameters could be population dependent.

ML models, such as decision trees and neural networks have also been deployed to learn human behavior (Cooney et al., 2019b). In addition to Yang et al. (2011), two particular defense algorithms MATCH (Pita et al., 2012a) and COBRA (Pita et al., 2012b) also provide defense mechanisms against deviations from rational behavior. However, they are only applicable to strictly competitive games. Thakoor et al. (2020) developed a game-theoretic/ML solution to strategically obfuscate the features of machines to reduce a defender's expected losses against boundedly rational attackers. In our work, we follow the defense strategies proposed in Thakoor et al. (2020) which we discuss in more detail in Section 2.1.

One of the challenges with game-theoretic/ML algorithms is that the predictive power of such models typically relies on large amounts of data to fit the model parameters. To understand how different defense algorithms would work in real scenarios usually requires human intervention and collecting large volumes of

human decisions in domains such as cybersecurity. Unfortunately, such interventions are very challenging. Generally, cognitive models are starting to play a direct role in applications where predictive models of human decision-making take the role of people in the task. For example, Sycara et al. (2015) developed a cognitive model based on the Adaptive Control of Thought-Rational (ACT-R) (Anderson, 1996) architecture that simulates human cognition for training a ML model in the control of a robotic swarm simulation. Similarly, Trafton et al. (2020) developed an ACT-R cognitive model for generating synthetic data that was used for a complex task. In cybersecurity, collecting data from actual human attackers and defenders has been a key challenge. In this research, we address the challenge of limited attackers' decisions, by generating large amounts of data on simulated human decisions using cognitive models. Specifically, we rely on Instance-Based Learning (IBL) Theory (Gonzalez et al., 2003) to construct this cognitive model. IBL Theory proposes basic principles and a process of how humans make decisions from experience: Decision makers recognize the similarity between a current decision situation and decisions made in the past to evaluate the expected benefits of available decision alternatives, and they learn from feedback on the decisions actually made (Gonzalez et al., 2003).

IBL models have been used for decades in a wide range of domains including repeated binary choice decisions (Lejarraga et al., 2012), multi-choice sequential decisions (Gonzalez and Ben-Asher, 2014), prediction of human reliance on automation (Lebiere et al., 2021), prediction of human Theory of Mind in gridworlds (Nguyen and Gonzalez, 2021), and prediction of cognitive biases in human decision making (including confirmation bias, anchoring and adjustment, probability matching, and base rate neglect) (Lebiere et al. (2013)). In the domain of cybersecurity, IBL models have been widely used to replicate human decision processes in a variety of tasks involving deception in insider attack games (Cranford et al. (2018, 2021)), intrusion detection systems (Aggarwal et al., 2017; 2020a) and susceptibility to phishing emails (Cranford et al., 2019; 2021). Yet, despite this success, existing IBL models of human attackers often involve relatively simplistic tasks abstracting the complexity of cyber scenarios. Moreover, such tasks involve repeated attacker-defender interactions, since that helps IBL models capture the experiential learning process and develop more accurate predictions. In this paper, we demonstrate that IBL models can replicate human decisions in complex and more realistic cyber scenarios that rely on a large number of features and limited repeated interaction of attackers and the task.

### 2.1. Thakoor et al. (2020) Masking strategy

Thakoor et al. (2020) proposed a Risk-Based Cyber Camouflage Game (i.e., masking algorithm) to modify the responses to attackers' queries during the network reconnaissance. Their algorithm is based on a general sum Stackelberg game model, in which the defender configures the network with a deception strategy (i.e., how the system should respond to scan queries from an attacker) and the attacker scans the network and chooses a system to attack based on the system's responses. In this scenario, the rewards for attackers and losses for defenders could be different. The masking algorithm assumes the worst-case scenario against a risk-averse attacker (i.e., considers the minimum utility that a particular deception strategy would yield, and consequently, aims to compute the strategy that maximizes such utility). The authors show that this problem is NP hard and provides a mixed-integer linear program to compute the optimal solution.

A network comprises of a set of machines and each machine has a certain *True Configuration* (TC) reflecting its various attributes and vulnerabilities. The defender tries to obfuscate the attributes so that the *Observed Configuration* (OC) from the attacker's perspec-

tive can significantly differ from the TC of a machine. Any machine having TC  $i$  has associated values  $v_i^a$  and  $v_i^d$  that the attacker gains and the defender loses respectively, if the machine is successfully attacked. The interaction between the attacker and the defender is modeled as a Stackelberg Security Game (SSG) owing to the sequential nature of the decisions. The defender is the leader who knows the true state of the network (i.e., the number of machines of TC  $i$ ). Given this information, the defender masks the TCs with OCs, and this assignment strategy is represented as an integer matrix  $\Phi$  where each entry  $\Phi_{ij}$  denoting how many machines having TC  $i$  are masked with OC  $j$ . Deploying these strategies have several domain constraints: 1) feasibility constraint (i.e., some OCs can't feasibly mask with some TCs) and 2) masking any TC with an OC are capped by a budget for the defender. Under these constraints, a defender strategy  $\Phi$  is generated.

Given the defender strategy  $\Phi$ , the attacker chooses a pair  $(i, j)$  indicating that an exploit for TC  $i$  is launched on a machine masked with OC  $j$ . The attack is successful if the attacked machine is among the  $\Phi_{ij}$  machines of TC  $i$  masked by OC  $j$ . Since OC  $j$  masks  $\sum_i \Phi_{ij}$  machines in total, the success probability is  $\Phi_{ij} / \sum_i \Phi_{ij}$  and consequently, the expected attacker ( $U^a$ ) and defender ( $U^d$ ) utilities are:

$$U^a(\Phi, i, j) = \frac{\Phi_{ij}}{\sum_i \Phi_{ij}} v_i^a, \quad U^d(\Phi, i, j) = \frac{\Phi_{ij}}{\sum_i \Phi_{ij}} v_i^d.$$

A rational attacker attacks a pair  $(i, j)$  that maximizes the expected utility. In case of indifference, the defender must consider the worst-case tie-breaking for the attacker due to the restriction to a pure strategy, which leads to *Weak Stackelberg Equilibria* (WSE; Breton et al., 1988). Hence, the defender tries to choose a strategy to maximize utility, assuming a utility-maximizing rational attacker. We refer to this strategy as **WSE Model**, which assumes rational attackers.

For risk-averse attackers, prospect theory (Tversky and Kahneman, 1979) asserts that their decisions are governed by a value transformation function  $R$  that is monotone increasing, and concave. Any reward  $v$  (namely, attacked machine's value), gets perceived as  $R(v)$ . A typical parametric form proposed in literature is  $R_\lambda(v) = c(v/c)^\lambda$ , with  $\lambda \leq 1$  capturing the risk-aversion of the attacker, and  $c$ , a suitable constant. Learning the parameter  $\lambda$  is a challenging task. This can be done by obtaining attacker responses on randomly generated strategies and computing a maximum likelihood estimate of  $\lambda$  given the observed instances. Once  $\lambda$  is estimated, the defender computes an optimal strategy for the risk-averse attacker by simply modifying the WSE algorithm and replacing the valuations  $v_i^a$  with the transformed values  $R_\lambda(v_i^a)$ . We refer to this strategy as **Prospect Theory (PT) Model**.

In this paper, we test the WSE and PT masking algorithms developed in Thakoor et al. (2020) using human-subject experiments. We generated WSE and PT strategies using the algorithm briefly discussed above (refer to Thakoor et al., 2019b for more details). The generation of PT strategy requires a risk-aversion parameter  $\lambda$ . To learn the risk-aversion (parameter  $\lambda$ ), we collected human data where attackers play against a random strategy. Different subjects may have a different degree of risk aversion (parameter  $\lambda$ ). However, since defenders cannot estimate the level of risk-aversion of an individual attacker in advance when deploying the strategy, we aim to estimate a  $\lambda$  that is representative of the whole population (of attackers) and compute the optimal strategy against an attacker with this  $\lambda$ . We do so by obtaining the maximum likelihood estimate given the data collected.

We recruited 35 subjects in the random condition playing 10 rounds each, we have  $|\mathcal{N}| = 350$  observations. To create a diverse dataset, for each participant, matrices for 10 rounds were randomly chosen from a pool for 50 matrices. Each observation  $n \in \mathcal{N}$  corresponds to a particular round played by a particular human participant — suppose the subject plays against a defense strategy

**Table 1**  
Attacker's Rewards and Defender's Losses per True Configuration.

TC	Attacker's Rewards	Defender's Losses
slackware	15	9
xbox	11	10
ubuntu8	2	6
winxpemb	13	4
avayagw	14	3
freebsd	11	10
winxp	2	14
win2008	11	2
win2k	7	8
win7pro	10	5
win7ent	9	8
openwrt	7	12
openbsd	15	15
linux	6	15
cisco2500	13	12

$\Phi_n$ , and decides to attack an  $(i_n, j_n)$  that maximizes its prospect. Based on the collected data, we computed the  $\lambda$  parameter using the Maximum Likelihood Estimation approach described in Thakoor et al. (2019b) and obtained  $\lambda = 0.75$ . The data collected in the random condition was only used to develop the  $\Phi$  matrices in the PT masking algorithm. Thus, we do not analyze the random condition data otherwise.

As the PT model has been adapted to the risk aversion of the participants, we expect that defenders' losses using the PT model would be smaller compared to the WSE model. Similar to the study in Aggarwal et al. (2020b), we also expect to observe the risk aversion bias in both models (WSE and PT), i.e., participants would prefer a surer option even when the payoffs are lower. In what follows, we will evaluate the effectiveness of two masking strategies, e.g., WSE and PT, against human attackers, before we present an IBL model.

### 3. Human experiment

In this experiment, we tested two masking strategies against human attackers: (1) WSE (i.e., "rational") masking, and (2) PT (i.e., "boundedly rational") masking. The WSE strategy generates the  $\Phi$  matrices according to Thakoor et al. (2020)'s algorithm. Thakoor et al. (2020)'s algorithm minimizes the utility of the perfectly rational attacker and reduces the expected losses for defenders against a rational attacker. The PT strategy generates the  $\Phi$  matrices similar to the WSE algorithm, but it minimizes the expected losses for defenders against a risk-averse attacker. The utilities for each TC are defined in Table 1. The WSE strategy does not perform any transformations and assumes that attackers would perceive the utilities as defined in Table 1. We compute an optimal strategy for the attacker playing according to PT transformation by simply modifying the WSE algorithm, replacing the valuation  $v_i^a$  with the transformed values  $R_\lambda(v_i^a)$ . Given the number of systems in each matrix, the number of matrices to produce, and feasibility constraints (i.e., the list of TCs that cannot be masked with a particular OC), the WSE and PT algorithms produce the strategy matrices ( $\Phi$ ) with the mapping of TCs to OCs.

To test the effectiveness of these strategies, we develop a task in CyberVAN, a realistic cybersecurity testbed for conducting human-in-the-loop experiments (Chadha et al., 2016). The CyberVAN testbed provides capabilities such as virtual networks, synthetic traffic, substantial tools for scanning and attack, and a specific set of vulnerabilities to conduct sophisticated cybersecurity research (Chadha et al., 2016). For this experiment, we use virtual machines, scanning tools and Honeyd service for deploying decep-

### 3.1. Experimental setup in cyberVAN

In the CyberVAN testbed, we assigned 5 honeyd servers where we configure honeyd files to mask the TCs of virtual machines to OCs using the strategy matrices. The honeyd configuration file masks the operating systems and ports of TCs with OCs to trick the network scanning tools (Provos, 2003). Each of the Honeyd servers could communicate to a range of IP addresses via a router that associates various virtual machines to these Honeyd machines.

Participants were provided a link and login credentials to a virtual machine running Kali operating system. As shown in Fig. 1, step 1, participants login to the virtual machine using the credentials provided to them. These virtual machines were configured with a scanning tool (i.e., zenmap) and attack scripts. Using these machines, participants scanned and attacked various machines. The task consists of 10 rounds (preceded by 1 practice round). Participants were provided a different pre-generated  $\Phi$  matrix in each round that provides TC to OC mapping of 15 virtual machines.

After logging in to the virtual machine, participants were asked to start the task via the start script as shown in Fig. 1. The start script provides the IP address range and  $\Phi$  matrix for the practice round. Similar information is provided for the main rounds as well. The  $\Phi$  matrix describes the type and number of machines present in the network (TC) and their corresponding masked configuration (OC). The  $\Phi$  matrices were randomly selected for each participant and the configuration of virtual machines was different in each round. Specific details of the  $\Phi$  matrices used in each round are provided in the Appendix A. Fig. 2 presents an example of a  $\Phi$  matrix used in one of the conditions. To help interpret the matrix, participants were given information regarding the way the TCs were mapped into OCs. For example, in the sample matrix, there are 6 TCs (avayagw, Ubuntu8, Win7pro, Win7ent, WinXP, Slackware) which are mapped to 3 OCs (freeBSD, Win7pro, and Ubuntu8). In the given matrix, for example, 5 machines are shown as freebsd, out of which 3 are actually avayagw and 2 are Ubuntu8. In addition to the mapping information, we provide the utility of each TC along with the matrix. Participants were allowed to use this information to calculate their probability of success and expected utility of attacking a particular machine.

In each round, participants perform two phases: exploration and attack. In the exploration phase, we provided Zenmap utility for using nmap commands in the exploration phase as shown in Fig. 1. Participants probe the machines using the nmap command to obtain information of the open ports, operating systems, and running services (according to the OC). Participants are free to probe any machines in any order. The participants received observable features on scanning the machines as a response from the nmap command. After the exploration phase, participants go through the attack phase, where they decide which machine to attack and what type of exploit to use to conduct the attack. To decide which machine to attack after exploration, rational attackers are expected to consider the potential utility and probability of success of using the correct exploit during an attack. Note that the utility for the attacker is different than the losses of the defender for each TC. Participants were provided the rewards that they would obtain if they were successful in their attack. In real scenarios, attacker's usually gather information about the machines during the reconnaissance phase and estimate the utility of the machines. For the simplicity of our experiment, we provide the rewards for the successful attack of each system upfront. These rewards are presented in Table 1. Note, the participants were only aware of their rewards, not the defender's loss. The attacker's utilities are randomly allocated between a range of 2 and 15 to represent the low, medium, and high valued machines in the network. The corresponding defender's losses were assigned with an assumption that the value of a TC may or may not be the same for

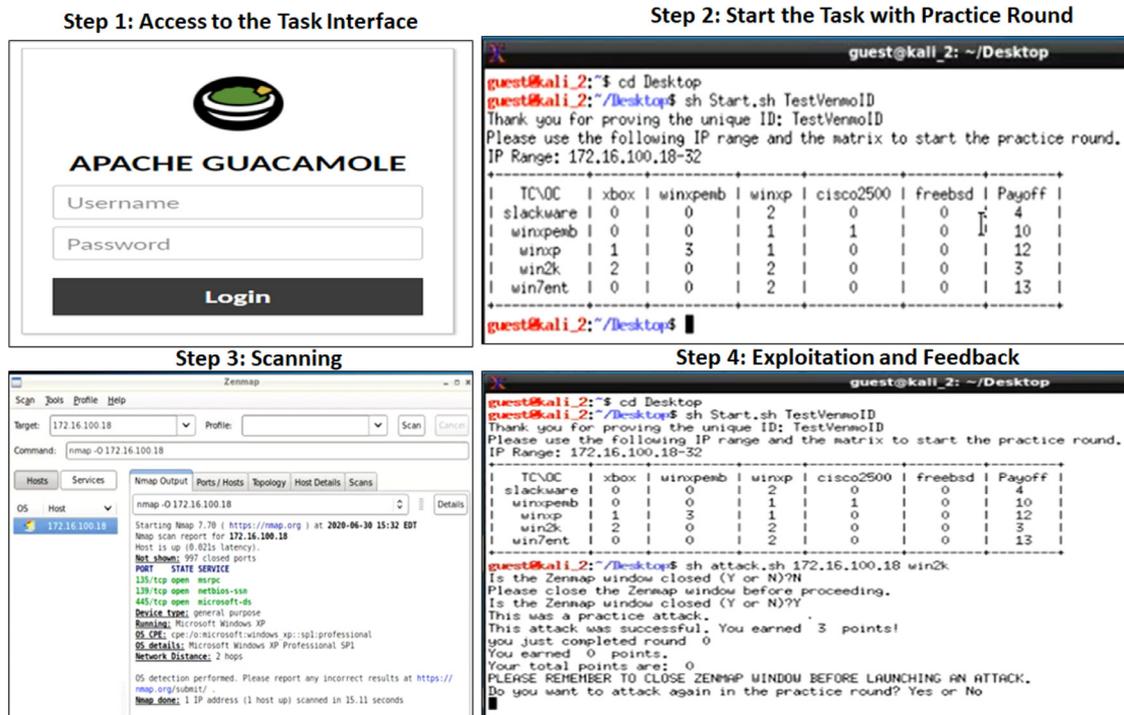


Fig. 1. Steps involved in the CyberVAN Task for Human participants.

$TC \backslash OC$	<i>freeBSD</i>	<i>win7pro</i>	<i>Ubuntu8</i>
<i>avayagw</i>	3	0	0
<i>Ubuntu8</i>	2	0	0
<i>win7pro</i>	0	2	0
<i>win7ent</i>	0	2	0
<i>winXP</i>	0	2	0
<i>Slackware</i>	0	0	1

Fig. 2. Sample  $\Phi$  Matrix: columns represent the observable configuration and rows represent the true configuration.

the attacker and defender. Thus, some TCs have equal defender's losses and others are either lower or higher than the attacker's gain. The attacker's rewards and defender's losses for the TCs remained the same across all 10 rounds. Participants earned the sum of the points accumulated across the 10 rounds, which were directly translated into a bonus monetary earning to the participant.

### 3.2. Participants

Participants were recruited through advertisements via various university email groups, social media, and cybersecurity targeted groups. To be qualified to participate, participants were required to pass an online test of basic cybersecurity knowledge, which included questions on various attacks, network protocols, scanning tools for networks, etc. The prescreening questions are included in Appendix C of the paper. The questions were adopted from previously published research by Ben-Asher and Gonzalez (2015). Only qualified participants were scheduled for an online study of 90 min.

An a priori power analysis was conducted using pwr library in R to test the difference between two independent group means using an ANOVA test, with a medium effect size ( $d = 0.40$ ), and an alpha of .05. Results showed that a total sample of participants with two equal sized groups of  $n = 25$  was required to achieve a power

of .80. Due to the highly specialized testbeds that require participants with good knowledge of cybersecurity, we could only recruit 25 participants in the WSE and 20 in the PT condition. In WSE condition, 84% participants reported themselves as male, (Age:  $Mean = 24.4$ ,  $SD = 4.2$ ) and in the PT algorithm, 70% participants reported themselves as male (Age:  $Mean = 28.7$ ,  $SD = 5.8$ ). Approximately 43% reported having or pursuing a bachelor's degree, 44% reported having or pursuing master's degree, 7% reported Ph.D degrees, and the rest reported to have another form of education. A majority of the participants reported having a type of hands-on experience (87%), 5% of participants reported themselves as experts, and only 5% of participants had no practical cybersecurity experience.

After the successful completion of the experiment, all participants were paid a base payment of \$18. In addition, for each successful exploit, participants received 1 point, which accumulated and were converted to a monetary bonus (\$1 per 10 points). Participants could earn up to \$15 in bonus based on their performance. The maximum time taken to complete the experiment was 90 min.

### 3.3. Procedure

First, participants provided informed consent and completed a demographic questionnaire. Next, they were asked to watch a video with instructions regarding the goal of the task and the general procedure. Participants were also provided with text instructions to which they could refer to during the experiment. Instructions were followed with a brief comprehension test. They received feedback if they incorrectly answered a question in the test. Participants were provided the contact details of the research assistant and they could ask any clarification questions before proceeding with the experiment.

During the instructions, participants were informed that the experiment would take up to 90 min and would consist of 11 rounds. After finishing the instructions, participants were provided with login and password information for the virtual machine. Once logged into their machine, participants could see a cheat sheet to help them throughout the task. In the terminal window, participants

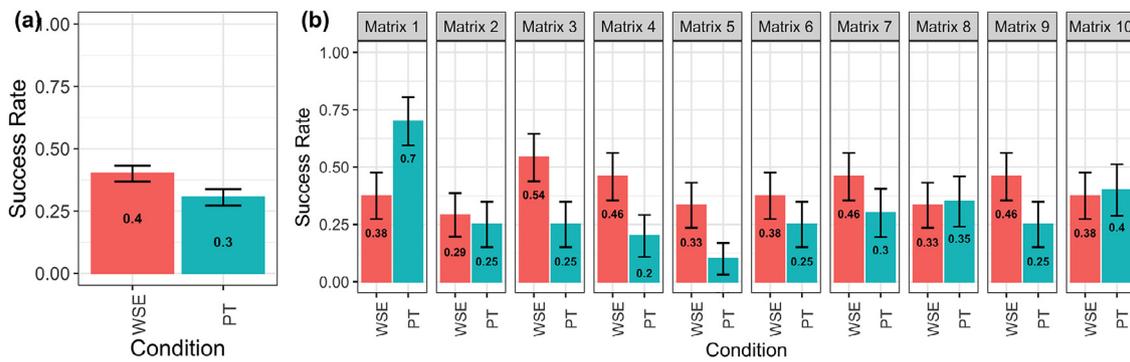


Fig. 3. Attacker's Success Rate: a) Average success and b) Matrix-wise success in WSE and PT conditions. The error bars represent the standard error.

Table 2  
ANOVA Table for all the dependent measures.

Measure	Effect	WSE		PT		df	F	MSE	p	$\eta^2$
		M	SD	M	SD					
Success	Condition	0.40	0.039	0.30	0.036	1, 42	3.02	0.33	0.08	0.07
	Matrix					9, 378	1.49	0.22	0.14	0.03
Attacker's Utility	Condition:Matrix					9, 378	1.76	0.21	0.07	0.04
	Condition	2.54	0.27	3.17	0.39	1, 42	1.83	23.90	0.18	0.04
Defender's Loss	Matrix					9, 378	1.56	21.92	0.12	0.03
	Condition:Matrix					9, 378	0.70	21.92	0.70	0.02
Defender's Loss	Condition	4.04	0.47	2.16	0.28	1, 42	10.40	36.97	0.002*	0.20
	Matrix					9, 378	2.22	22.11	0.01*	0.05
	Condition:Matrix					9, 378	3.19	22.12	0.00*	0.07

Note: MSE: Mean Square Error,  $\eta^2$ : partial eta-square.

started the task and received information such as IP addresses, the  $\phi$  matrix, and payoffs during each round. In each round, participants were asked to probe the machines using an nmap command like "nmap -O 172.16.31.31" to learn about open ports and operating systems on this IP address. They were also allowed to scan a specified range of IP addresses in each round together using a command like "nmap -O 172.16.31.31-61". After probing the machines, participants were allowed to calculate the likelihood of the true configuration of the machines by looking at the utilities of each of the configurations and the  $\Phi$  matrices. Next, using the attack script, they decided what IP addresses to attack by selecting an appropriate exploit. Participants received points if the exploit matched with the true configuration; otherwise, they received zero points. Once they finished all rounds, we asked for their feedback regarding the experiment.

### 3.4. Experiment results

Participants generally scanned all machines before launching an attack. In the practice round, each participant exploited between a minimum of 1 and a maximum of 7 machines.

We analyzed the data collected in the WSE and PT conditions during the 10 actual rounds. We randomly allocated 10 matrices to the participants during the 10 rounds. To measure the effectiveness of each matrix, we measured the average attacker's utility, their success rate, and their attacker's preferred option for the attack. We also analyzed the algorithm (i.e., defender) loss.

#### 3.4.1. Attacker's success rate

We calculated the rate with which participants used the correct exploit. Fig. 3(a) shows the average success rate of participants in the WSE and PT conditions. The difference in participants' average success rate when paired with the WSE than the PT algorithm was not statistically significant and an extremely low effect size was found (0.40 ~ 0.30;  $F(1, 42) = 3.02, p = 0.09, \eta^2 = 0.07$ ). The detailed statistics are presented in Table 2. The success rates in each

matrix of 10 rounds are shown in Fig. 3(b). Although it appears that human attackers exploited the machines more successfully in the WSE than the PT condition (except matrix 1), in most of the matrices the difference between WSE and PT was not significant.

The attacker's utility for each condition is shown in Fig. 4(a & b). For each successful exploit, the attacker gained points in accordance to Table 1. We observe that although the attackers gained slightly more points in the PT masking algorithm compared to WSE algorithm, however, the statistical test revealed no significant difference between the masking conditions, (2.54 ~ 3.17;  $F(1, 42) = 1.83, p=0.18, \eta^2 = 0.04$ ) with a low effect size and power. None of the differences within each matrix was significant ( $p > 0.05$ ) (see Table 2). The dotted line represents the utility when the best option based on expected values is selected. We observe in Fig. 4(a) that overall human attackers earned fewer points compared to the best option utility. The matrix-wise analysis in Fig. 4(b) shows that human attackers consistently earned fewer points when working against the WSE algorithm. We also observe lower attacker's utility compared to the optimal utility in the majority of the matrices for WSE and PT algorithms.

#### 3.4.2. Defender's losses

The losses for each of the two defense algorithms against humans are shown in Fig. 5(a & b). For each successful exploit, the attacker gained points and the defender lost points in accordance with Table 1. According to Fig. 5(a), overall, the defender's losses were higher in WSE condition compared to the PT condition. We also observed that defenders' losses are higher than the expected losses in both the conditions. To support these observations, we conduct a mixed-ANOVA to evaluate the effect of conditions and matrices.

The statistical test revealed a significant difference between the masking conditions (4.04 > 2.16;  $F(1, 42) = 10.40, p<0.002, \eta^2 = 0.20$ ). We also found that there is a significant differences between matrices ( $F(9, 378) = 2.23, p=0.02, \eta^2 = 0.05$ ) and interaction between conditions and matrices, ( $F(9, 378) = 3.19, p<0.001$ ,

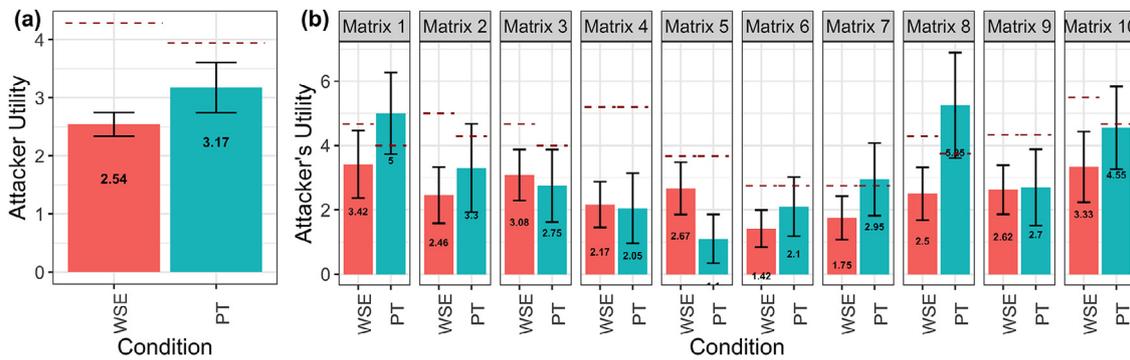


Fig. 4. a) Average Attacker's Utility and b) Matrix-wise Attacker's Utility in WSE and PT conditions. The dotted lines represent the average expected utility for the attacker. The error bars represent the standard error.

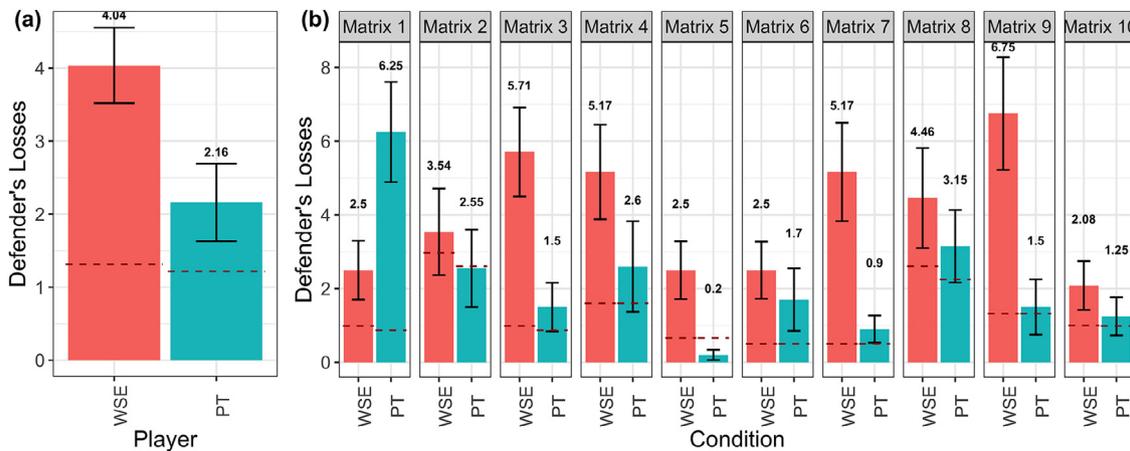


Fig. 5. (a) Average Defender's Losses in WSE and PT conditions. (b) Matrix-wise Defender's Losses in WSE and PT conditions. The dotted lines represent the average expected losses.

$\eta^2 = 0.07$ ). The detailed statistics are presented in Table 2. The average defender's losses per defense strategy ( $\Phi$  matrix) are shown in Fig. 5(b). The defender's losses were higher for WSE algorithm compared to the PT algorithm for all matrices except matrix 1. We also performed the post-hoc analysis to evaluate the effect of the condition in each matrix. The post-hoc analysis shows that Matrix 1 ( $2.5 < 6.25$ ;  $F(1, 42) = 6.10, p < 0.05, \eta^2 = 0.13$ ), Matrix 3 ( $5.71 > 1.5$ ;  $F(1, 42) = 8.38, p < 0.01, \eta^2 = 0.17$ ), Matrix 5 ( $2.5 > 0.20$ ;  $F(1, 42) = 7.02, p < 0.01, \eta^2 = 0.14$ ), Matrix 7 ( $5.17 > 0.90$ ;  $F(1, 42) = 8.06, p < 0.01, \eta^2 = 0.16$ ) and Matrix 9 ( $6.75 > 1.5$ ;  $F(1, 42) = 8.36, p < 0.01, \eta^2 = 0.17$ ) shows a significant difference in defender's losses in the two conditions.

To explore the contrast between attacker utility and defender loss, we plotted the frequency of attacks on each machine in each of the matrices. Fig. 8 presents these results for WSE and PT in human experiments (and IBL model discussed later). The machines are sorted based on their expected value: in each matrix, the left-most bar is the machine with the lowest expected value, and the right-most bar is the machine with the highest expected value. The corresponding attacker's payoff and probability of success for each of the TCs are provided at the top of each bar.

We observe that participants in the WSE condition prefer a more certain option compared to the one with more uncertainty. For example, in Matrices 3, 4, 6, and 7 in Fig. 8, a significant number of participants attacked the TC with utility 2 and probability of success 1.0. This observation is in agreement with past findings regarding risk aversion in Aggarwal et al. (2020b) where humans

tended to attack targets that were more likely to result in success, regardless of reward.

In the PT condition of Fig. 8, only Matrix 1 has a TC with a success probability of 1.0. We observe that even though the potential reward for this machine is low (i.e., only worth 2 points for a successful attack), a significant number of participants chose that machine compared to the other machines in that matrix. The sure option resulted in the loss of 14 points to the PT defender, which contributed to the heavy losses observed under phi Matrix 1. This result suggests that the PT algorithm helped avoid the certainty effect by reducing the number of matrices that would have machines with a certain outcome. However, even with the PT algorithm, one of the matrices had one machine with a safe option, and we observe how human participants once again fell into the trap of the risk aversion and certainty bias found in our previous experiment with CyberVAN (Aggarwal et al., 2020b). Humans preferred a sure option regardless of the low benefits.

#### 4. Instance-based learning (IBL) model

To gain a better understanding of human decision making in the task, we developed a cognitive model of attack decisions in the CyberVAN scenario using IBLT (Gonzalez et al., 2003). According to IBLT, a human makes decisions by generalizing across past experiences that are similar to the present decision situation. Each experience (i.e. an instance) is represented as a triplet, including the contextual features of the selected target, the decision, and

outcome. Instances are accumulated in memory when options are evaluated and decisions are made in the environment.

#### 4.1. IBL Theory

Generally, the IBL procedure is as follows. When a new decision is to be made, the similarity is computed between the current situation and the existing instances in memory. For each possible decision, the model computes an expected utility, using a *blending* mechanism involving the average across past outcomes weighted by their probability of memory retrieval. The memory retrieval probability is calculated by weighing the memory *activation* of an instance against all the instances in memory. The activation of an instance is a concept formalized in the ACT-R cognitive architecture (Anderson et al., 2004). The activation depends on the contextual similarity to past instances, on the frequency of experiencing similar instances, and on the recency with which an instance has been experienced in the past. According to IBLT, after evaluating the alternatives and determining their blended value, a decision is made for the option that has the highest blended value (i.e., the highest expected utility). Finally, the expected utility is updated with an experienced utility once the outcome of a decision made is known. These instances are reused for making future decisions.

The blended value  $V_{k,t}$  of option  $k$  at trial  $t$  is computed as follows:

$$V_{k,t} = \sum_{i=1}^n P_{i,k,t} * X_{i,k,t} \quad (1)$$

where  $X_{i,k,t}$  represents the outcome of an instance  $i$  for option  $k$  at trial  $t$  and  $P_{i,k,t}$  represents the probability of retrieval of an instance  $i$  for option  $k$  at any trial  $t$  (value of  $k$  is the options in each round). The retrieval probability of an instance  $i$  is the ratio of activation of the  $i$ th instance corresponding to the activation of all instances (1, 2, ...,  $n$ ; where  $n$  is total instances) created within the option  $k$  at trial  $t$ . The retrieval probability is defined as:

$$P_{i,k,t} = \frac{e^{A_{i,k,t}/\tau}}{\sum_{i=1}^n e^{A_{i,k,t}/\tau}} \quad (2)$$

Here,  $\tau = \sigma * \sqrt{2}$  and  $\tau$  is a free noise parameter. Noise captures the inaccuracy of remembering past experiences from memory.

At each trial,  $t$ , activation of an instance  $i$  on option  $k$  represents the linear aggregation of three cognitive elements: frequency and recency, the similarity of the instance to past experiences, and the noise that introduces stochasticity in the activation value (Anderson et al., 2004):

$$A_{i,k,t} = \ln \sum_{t_i=1,t-1} (t - t_i)^{-d} + MP \sum_k Sim(v_k, c_k) + \sigma * \ln \left( \frac{1 - \gamma_{i,k,t}}{\gamma_{i,k,t}} \right) \quad (3)$$

The first term reflects the power law of experience and forgetting.  $t_i$  represents all the previous trials where the instance  $i$  was either created or its activation was reinforced due to its recurrence.  $t_j$  is the time since the  $j$ th occurrence of instance  $i$  and  $d$  is the decay rate of each occurrence which is set to the default ACT-R value of 0.5. The activation of an instance can increase with the frequency of observing that outcome, as well as with the recency (i.e., by small differences in  $t - t_i$ ). This term represents the frequency and recency of events in the memory. The decay parameter accounts for the rate of forgetting the experienced events: the higher the decay, the faster will be the rate of forgetting the past events and the reliance on recent events will increase.

The second term is a partial matching process reflecting the similarity between the current situation ( $c_k$ ) and the instances that are stored in memory ( $V_k$ ), scaled by a mismatch penalty (set to

2.5). Similarity between numerical 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.

The third term represents the Gaussian noise mechanism for capturing the variability in individual choices and  $\gamma_{i,k,t}$  is a random number drawn uniformly between 0 and 1. The  $\sigma$  (i.e., the variance in the noise term) is set to the default ACT-R value of 0.5.

#### 4.2. IBL model of attacker

Fig. 6, represents the IBLT decision process for the human attacker model in CyberVAN. We learned through a post-survey questionnaire and interactions with human participants that participants combine the number of TCs and the number of OCs to calculate the probability of success. Participants then use the probability of success and payoffs to decide among the options. The contextual information in CyberVAN involves the TC, OC, the number of TCs, the number of OCs and payoffs, and in the instances we used, TC, OC, and the expected value of the option (i.e. the ratio of the number of TCs to the number of OCs multiplied by the payoff), assuming that participants are able to calculate such probabilities from the information provided and compute the expected values. The decision in the instance is the OC/TC combination to exploit, and the payoffs are the obtained utilities received after attacking one of the available options (i.e., the attacker's utility in 1).

To begin the task, the model is initialized with instances in the practice round corresponding to either successful attacks or failed attacks. These initial instances represent the payoff expectations that human participants are likely to acquire during the practice round, which are used in the task rounds after practice.

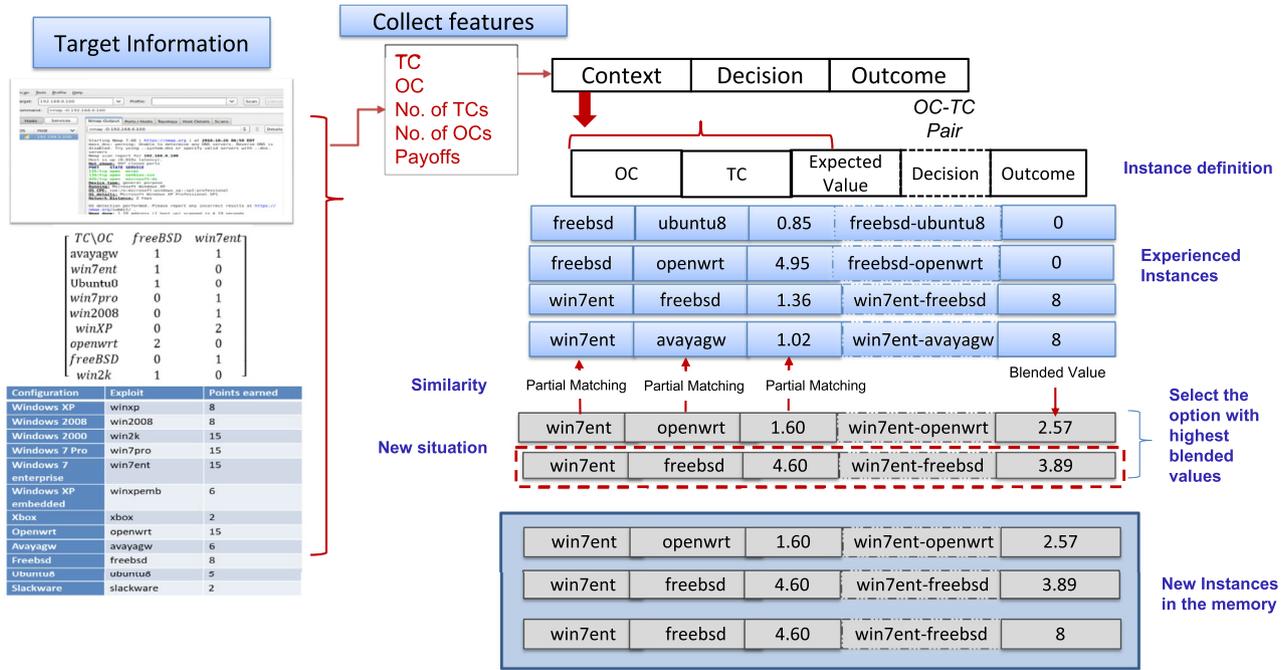
In each of the 10 rounds of the task, the model first processes all decision options available in the  $\Phi$  matrices. The attacker's model calculates the expected utility for each of the decision options using the blending mechanism. The model stores each of these instances of the options evaluated and their blended values. This process represents the way humans might scan different machines during the exploration phase and develop expectations by processing the information given in the form of the  $\Phi$  matrix.

Once the model has calculated the expected utility of all options, the model selects the option with the highest blended value to attack. The selected option with the experienced outcome is then stored in the memory.

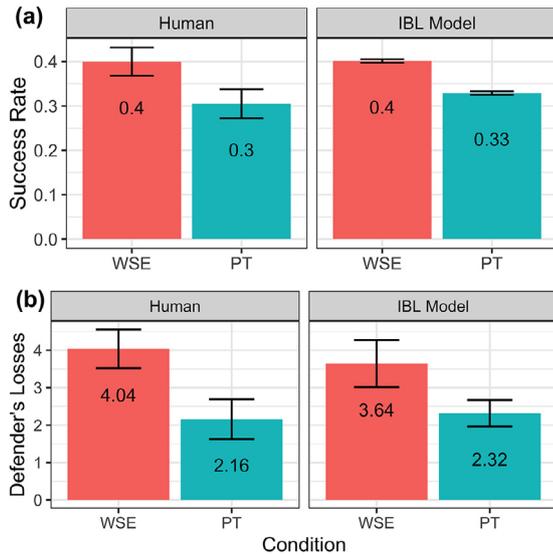
The exploration and exploitation process is repeated for each of the 10 rounds in the task. Using this IBL model, we ran individual simulations representing each individual attacker and evaluated the model's performance against the empirical results of the two experimental conditions WSE and PT, explained above. The IBL model described above was run 1500 times in each condition to generate stable estimates of the participants' performance during the CyberVAN experiment. Due to the stochasticity in the model, the data generated from the 1500 agents show variability just like in human data, but the large number of agents run in this simulation help in providing "stable" predictions of the model. Each run of the model (i.e., an agent) involved the same procedure that each human participant went through. Next, we present the results from the model's simulation against empirical data in the WSE and PT conditions.

#### 4.3. IBL Model results

We analysed the model's success rate, average points earned in each matrix, and the target selection preferences. We also compared the model's results from each of these measures against human data using the Root Mean Square Error (RMSE). The RMSE is calculated by subtracting the average human action from model



**Fig. 6.** IBL Model Process representing the 1) instance structure based on the task environment 2) process of comparing the new situation with existing instances using partial matching and calculating blended values and 3) making decisions based on the highest blended values and storing new instances in memory for future decisions.



**Fig. 7.** (a) Success Rate in PT and WSE algorithms from Human Data (left) and IBL model (right) and (b) Defender's Loss in PT and WSE algorithms from Human Data (left) and IBL model (right).

actions using the following formula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Human_i - Model_i)^2} \quad (4)$$

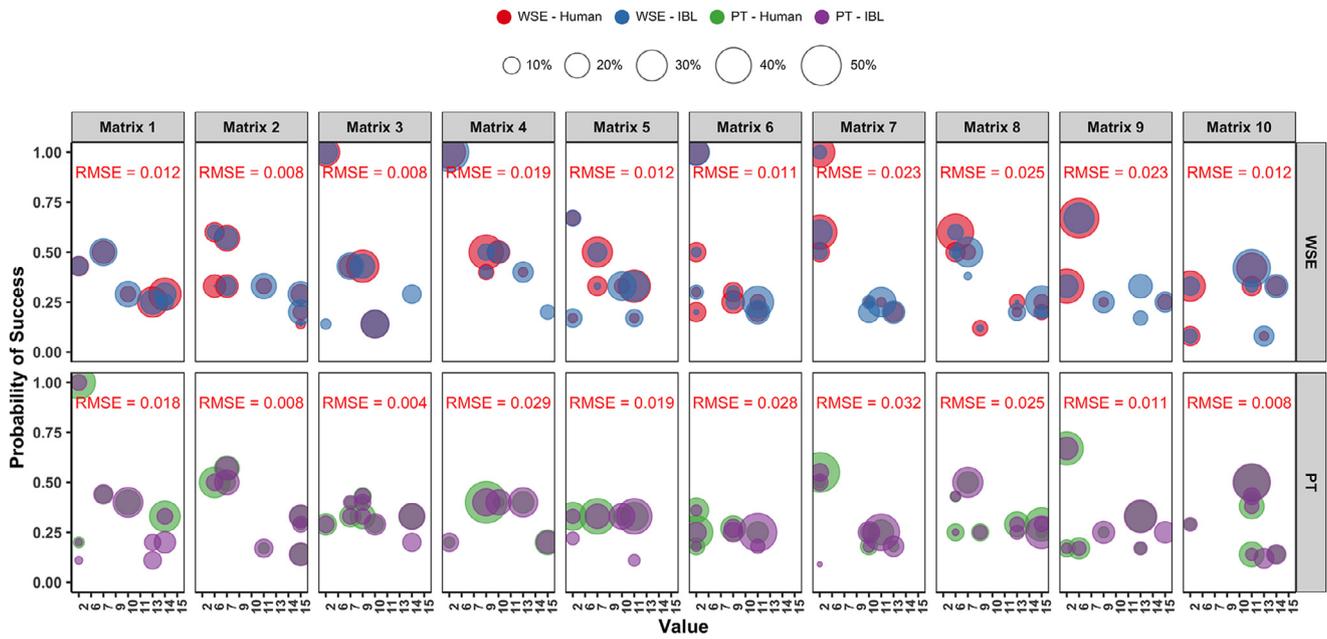
Fig. 7 (a) and (b) show the overall average success rate and the defender's utility in the WSE and PT conditions, resulting from the human data and the simulations of IBL model. As observed, the general patterns of the model results correspond well with those of human participants. The overall RMSE between the human data and the model data is 0.079 for the success rate in the WSE strategy and 0.106 for the PT strategy. Similar to human data, the model

predicts that the success rate is slightly higher for the WSE than the PT strategies in most of the matrices. The RMSEs for the success rates between human and model data for individual matrices are shown in Table 3. These values suggest that the model is able to predict the human data in most of the matrices quite accurately.

The defender algorithm's losses calculated from the IBL simulations are shown in Fig. 7(b). Again, the IBL model reproduces the general trends found in human data: the model predicts greater defender losses in the WSE than the PT defense strategy, overall and for each of the matrices. The RMSE values comparing the defender's losses between the human and model data for each matrix and overall are presented in Table 3. Note that we compute the RMSEs for the defender's loss after normalizing the defender's loss between 0 to 1. The corresponding values of RMSE for the defender's losses are 0.076 and 0.069 for the WSE and PT strategies, respectively. Generally, the RMSE values for defender losses in both WSE and PT algorithms suggest that the IBL model is able to predict the defender's losses accurately in a majority of matrices.

**Individual Selection Behaviour** The results of the IBL model showed that it is able to predict defender's loss and success rate reasonably well for both the PT and WSE conditions at the average level. However, a model that makes good predictions at the average level, might not be able to predict the individual decision variability (Dutt and Gonzalez, 2015). In this section, we analyze the distribution of the selection of individual machines as predicted by the IBL model against the human data.

Fig. 8 shows the distribution of individual machines organized according to their outcome (X-axis) and their probability of success (Y-axis). The figure overlays the human data with the model's predictions of selection preferences. The size of the circle represents the normalized frequency of participants in that particular option. The IBL model is generally able to capture the distributions of human preferences in both the WSE and PT strategies. The IBL model not only predicts human actions on the aggregate level but also captures the selection behavior in both the conditions. The figure also makes some differences apparent and the possible reason for those differences is the noise within the IBL model.



**Fig. 8.** Frequency of Selection of individual option from the human participants and the IBL model across 10 Matrices in WSE and PT algorithm. The RMSEs on each matrix represent the deviation of frequency among human and model participants for each option.

**Table 3**

Average success rate and defender's losses for each matrix and overall in WSE and PT algorithms for Human and IBL Model and their corresponding RMSE values. The visual representation for this table is included in the [Appendix B](#) of the paper.

Matrix	Success Rate						Defender's Loss					
	WSE			PT			WSE			PT		
	Human	IBL	RMSE	Human	IBL	RMSE	Human	IBL	RMSE	Human	IBL	RMSE
1	0.375	0.306	0.069	0.7	0.351	0.349	2.5	2.087	0.028	6.25	2.63	0.241
2	0.292	0.362	0.070	0.250	0.359	0.109	3.542	4.196	0.044	2.550	4.238	0.113
3	0.542	0.446	0.096	0.25	0.321	0.071	5.708	4.409	0.087	1.501	2.052	0.037
4	0.458	0.691	0.232	0.2	0.345	0.145	5.167	8.063	0.193	2.601	2.493	0.007
5	0.333	0.350	0.016	0.101	0.326	0.226	2.501	1.805	0.046	0.201	1.179	0.065
6	0.375	0.412	0.037	0.25	0.252	0.002	2.5	2.695	0.013	1.701	1.260	0.029
7	0.458	0.351	0.108	0.301	0.299	0.002	5.167	2.627	0.169	0.901	1.444	0.036
8	0.333	0.374	0.041	0.350	0.318	0.032	4.458	4.455	0.004	3.150	3.517	0.024
9	0.458	0.396	0.062	0.250	0.346	0.096	6.750	4.801	0.130	1.501	3.365	0.128
10	0.375	0.325	0.049	0.400	0.372	0.027	2.083	1.291	0.053	1.250	0.999	0.017
Mean	0.399	0.401	0.0781	0.305	0.329	0.106	4.037	3.643	0.076	2.160	2.318	0.069

### 5. Discussion

In the cybersecurity domain, it is difficult to gain an understanding of the attacker's decision-making due to the lack of such decision data. Defense algorithms often rely on the assumption that attackers are rational decision makers and that they take the best course of action. Using human experiments, [Aggarwal et al. \(2020b\)](#) demonstrated that human attackers have a risk-aversion bias while making cyber-attack decisions. To exploit the risk-aversion bias of human attackers, [Thakoor et al. \(2020\)](#) developed a masking strategy using Prospect Theory (PT). Specifically, [Thakoor et al. \(2020\)](#) developed two masking algorithms: one strategy that assumes full rationality (WSE) while the other strategy exploits bounded rationality in the form of risk-aversion (PT). In this paper, we test PT and WSE masking strategies of defense against human attackers in an experiment.

The PT strategy developed by [Thakoor et al. \(2020\)](#) was calibrated using human attacker's data collected in an experiment where humans were pitted against random strategies. This data set helped in estimating the risk-averse parameter,  $\lambda = 0.75$ , for the PT strategy. The results from the comparison between WSE and PT strategies showed that the strategies were not different

with respect to the attackers success, but they were different with respect to the defender loss. The PT strategy resulted in lower defender losses compared to WSE. These results against human attackers are in agreement with the numerical findings in [Thakoor et al. \(2020\)](#) which evaluated these strategies against simulated risk-averse attacker populations. In other words, these results support the idea that game theoretic and ML methods that account for human bounded rationality can produce better defense strategies than methods that assume full rationality, both in theory and in practice, against human attackers.

Importantly, the PT algorithm would try to avoid generating  $\phi$  matrices in which there was a true configuration that would exactly correspond to an observable configuration, given the human bias towards a safe option ([Aggarwal et al., 2020b](#)). We observed that the only situation in which the PT algorithm produced a larger defender loss and larger attacker success than the WSE algorithm was the case in which the matrix (Matrix 1, [Fig. 5](#)) had one sure option. This again, suggests that such human bias towards certainty is inescapable for humans, and that the PT algorithm would need to be revised to ensure that such cases are prevented.

According to Instance-Based Learning Theory, human decisions in complex and uncertain environments, such as cyberdefense, are

made through exploration of the available options and the aggregation of past decisions using the similarity of situations, recency, and frequency of events (Gonzalez et al., 2003). In this paper, we used IBL models to replicate human attacker’s decisions in the presence of a cyberdeception strategy (i.e., masking). Humans generate expectations while exploring machines using the nmap command and a  $\phi$  matrix. Similar to humans, the IBL model also generates expectations about the utility of attacking various options using the information in the  $\phi$  matrix. The model makes a choice, selecting a machine with the maximum value. The IBL model of the attacker is able to predict human actions in both WSE and PT algorithms. The IBL model also predicts the decision making bias (i.e., risk-aversion as observed in human data). Past research has also demonstrated the applications of IBL models in predicting biases such as confirmation bias, probability matching, anchoring bias and representativeness (Lebiere et al., 2013). The IBL model in this paper reflects the probability matching behaviour and capture the risk averseness among human participants. Thus, the IBL model could be used to predict attacker decisions and cognitive biases, which could help build better defense algorithms.

Game theory/ML algorithms of defense are often data driven and usually do not consider insights about human behavior. The cognitive models in such situations could be used in multiple ways, including providing interpretation of human behaviour and acting as a data source for ML algorithms by generating accurate predictions about human data. Through human experimentation, Aggarwal et al. (2020b) provided insights about human’s risk-aversion bias and Thakoor et al. (2020) developed a masking algorithm to exploit such behavior in attacker’s decisions. To accurately represent the risk-aversion, we collected human data with a random masking strategy and adapted the PT model to the risk-aversion parameter. This research validates the numerical findings of Thakoor et al. (2020)’s masking algorithm in a human experiment. Furthermore, the IBL model of an attacker replicates human decisions and provides a methodology to generate accurate predictions of human decisions when data is limited. The IBL model with accurate representation of human data could help in validating new defense algorithms given that human data is challenging to collect. Moreover, the IBL model could be used to calibrate model parameters in real-time and game-theory models could be

adapted to new behavioral patterns. Similar to the adaptive deception strategies in Cranford et al. (2020), by leveraging the capability of an IBL model, the masking algorithms could predict human biases in real time and exploit them using defense algorithms. For example, we demonstrated that IBL model agents also produced certainty bias as observed in human data.

A potential limitation in our conclusions from the human experiment comes from the fact that our experiment is underpowered, given the low number of participants. Recruiting participants with specialized expertise such as cybersecurity knowledge is always challenging, and collecting 45 participants implied a significant effort in data collection. However, the close replication of the decisions with the IBL model allows us to simulate many more participants in each condition. Although the algorithm and the experiments in this paper have been conducted for a limited number of nodes and simple network structures, the masking algorithms are capable of including network constraints that apply in other realistic settings. Similarly, the IBL models could also adapt to different cybersecurity scenarios with a possible limitation of the run time, which could be a bottleneck for large applications. Through experiments and models, we developed an understanding of how human attackers make decisions. Attackers are not rational; instead they act according to decision biases including certainty and risk-aversion. Human attackers shift from the expected optimal actions that some defense algorithms assume; they make suboptimal decisions. When defense algorithms are designed to exploit such biases in attacker decision making, they could reduce the overall losses incurred from cyberattacks.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Appendix A.  $\Phi$  matrices for WSE and PT algorithms**

We include the  $\Phi$  matrices that were provided to human participants during the experiment. Fig. A.1 represents the WSE condition and Fig. A.2 represents the PT condition.

<b>Matrix 1</b>	TC\OC	winxp	win7pro	openwrt	freebsd						
	winxpemb	0	2	0	0						
	avayagw	2	2	0	0						
	winxp	3	0	0	0						
	win2k	0	4	0	0						
	win7pro	2	0	0	0						
<b>Matrix 2</b>	TC\OC	winxpemb	win2k	win7pro	freebsd						
	slackware	1	0	2	0						
	freebsd	0	1	0	0						
	openwrt	0	1	4	0						
	openbsd	1	0	1	0						
	linux2.4.7	3	1	0	0						
<b>Matrix 3</b>	TC\OC	win2k	openbsd	linux2.4.7	freebsd						
	avayagw	0	0	2	0						
	winxp	0	1	1	0						
	win2k	3	0	0	0						
	win7pro	1	0	1	0						
	win7ent	3	0	3	0						
<b>Matrix 4</b>	TC\OC	ubuntu8	winxpemb	freebsd	openwrt						
	winxpemb	0	0	0	2						
	winxp	1	1	0	0						
	win7pro	0	0	4	0						
	openbsd	0	0	0	1						
	win7ent	0	0	4	2						
<b>Matrix 5</b>	TC\OC	winxpemb	winxp	openbsd	freebsd						
	ubuntu8	1	2	1	0						
	freebsd	0	0	1	0						
	win2008	2	0	0	0						
	win2k	3	0	2	0						
	win7pro	0	1	2	0						
<b>Matrix 6</b>	TC\OC	slackware	linux2.4.7	cisco2500	freebsd						
	ubuntu8	2	1	0	0						
	freebsd	2	0	0	0						
	win2008	0	0	1	0						
	winxp	3	0	2	0						
	win7ent	3	0	1	0						
<b>Matrix 7</b>	TC\OC	slackware	ubuntu8	winxpemb	linux2.4.7						
	ubuntu8	0	0	3	3						
	winxpemb	0	0	1	1						
	win2008	0	1	0	0						
	winxp	1	2	0	0						
	win7pro	0	1	1	1						
<b>Matrix 8</b>	TC\OC	ubuntu8	windows2C	winxp	freebsd						
	slackware	1	0	2	0						
	openwrt	0	1	3	0						
	linux2.4.7	3	1	0	0						
	cisco2500	1	0	2	0						
	win7ent	0	0	1	0						
<b>Matrix 9</b>	TC\OC	winxp	openbsd	openwrt	freebsd						
	winxpemb	1	2	0	0						
	winxp	0	4	0	0						
	openbsd	0	3	0	0						
	linux2.4.7	2	0	0	0						
	win7ent	0	3	0	0						
<b>Matrix 10</b>	TC\OC	windows2C	openwrt	winxp	freebsd						
	xbox	5	1	0	0						
	ubuntu8	1	1	0	0						
	winxpemb	1	0	0	0						
	avayagw	0	1	0	0						
	win2008	5	0	0	0						

Fig. A.1. WSE Matrices.

<b>Matrix 1</b>	TC\OC	winxp	win7pro	openwrt	<b>Matrix 6</b>	TC\OC	slackware	cisco2500	win2k
	winxpemb	1	1	0		ubuntu8	2	1	0
	avayagw	1	3	0		freebsd	2	0	0
	winxp	1	1	1		windows2C	0	1	0
	win2k	0	4	0		winxp	4	1	0
	win7pro	2	0	0	<b>Matrix 7</b>	TC\OC	ubuntu8	winxpemb	win2k
<b>Matrix 2</b>	TC\OC	winxpemb	win2k	win7pro		ubuntu8	0	6	0
	slackware	2	0	1		winxpemb	0	2	0
	freebsd	1	0	0		windows2C	1	0	0
	openwrt	0	1	4		winxp	2	1	0
	openbsd	0	0	2		win7pro	1	2	0
	linux2.4.7	3	1	0	<b>Matrix 8</b>	TC\OC	ubuntu8	winxp	win7pro
<b>Matrix 3</b>	TC\OC	avayagw	win2k	linux2.4.7		slackware	2	1	0
	avayagw	1	1	0		openwrt	0	2	2
	winxp	0	0	2		linux2.4.7	3	0	1
	win2k	2	1	0		cisco2500	2	0	1
	win7pro	0	0	2		win7ent	0	1	0
	win7ent	2	1	3	<b>Matrix 9</b>	TC\OC	win7pro	openbsd	win2k
<b>Matrix 4</b>	TC\OC	freebsd	openwrt	win2k		winxpemb	1	2	0
	winxpemb	0	2	0		winxp	2	2	0
	winxp	2	0	0		openbsd	0	3	0
	win7pro	4	0	0		linux2.4.7	0	2	0
	openbsd	0	1	0		win7ent	0	3	0
	win7ent	4	2	0	<b>Matrix 10</b>	TC\OC	windows2C	openwrt	win2k
<b>Matrix 5</b>	TC\OC	winxpemb	openbsd	win2k		xbox	3	3	0
	ubuntu8	2	2	0		ubuntu8	0	2	0
	freebsd	0	1	0		winxpemb	1	0	0
	windows2C	2	0	0		avayagw	0	1	0
	win2k	2	3	0		windows2C	4	1	0
	win7pro	0	3	0					

Fig. A.2. PT Matrices.

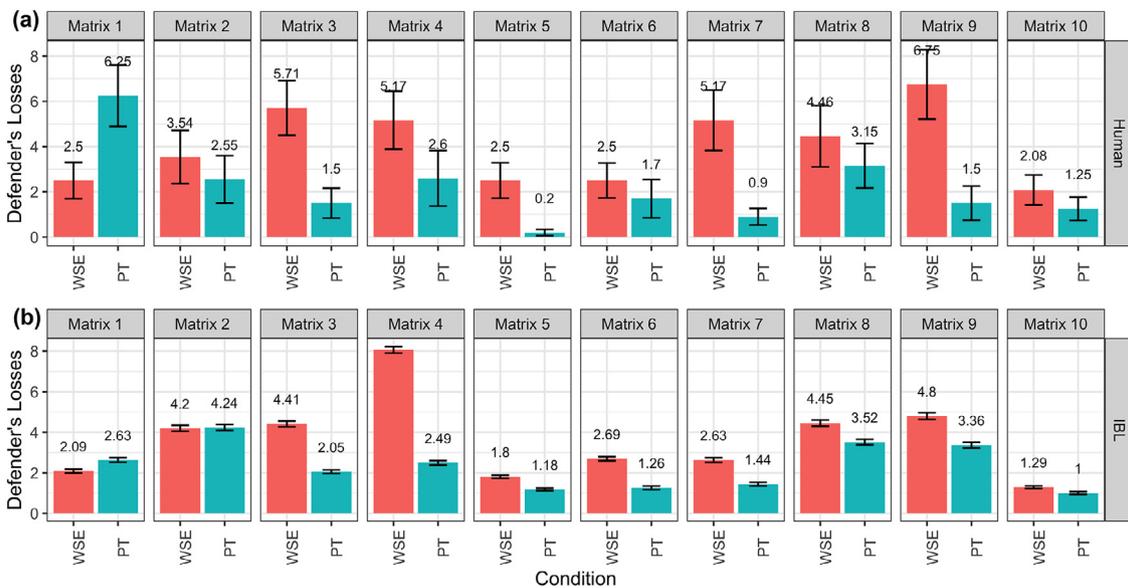


Fig. B.1. Defender's Losses in PT and WSE for IBL Model (matrix-wise).

### Appendix B. Matrix-wise Comparison for Human and IBL Model

We include a matrix-wise comparison of the average success rate and defender's losses obtained from IBL model and human participants in WSE and PT algorithms. Fig. B.1 represents the defender's loss and Fig. B.2 represents the success rate.

#### Defender's Loss for Human and IBL Model

Fig. B.1 presents the matrix-wise defender's losses from human participants (top panel) and IBL model (bottom panel). We observe that, similar to human participants, IBL model also produces higher defender's losses except in matrix 1 in WSE condition compared to the PT model. The difference of defender's loss between WSE

and PT is not well captured by the IBL model. We have earlier discussed the RMSE values in Table 2 which reflect the model performance across all matrices.

#### Attack Success Rate for Human and IBL Model

Fig. B.2 presents the matrix-wise attacker's success rate for human participants (top panel) and IBL model (bottom panel). We observe that similar to human participants, the IBL model also produces higher success in all matrices except in matrix 1 in WSE condition compared to PT model. The model better captures humans in WSE condition compared to the PT condition. The RMSE values were earlier presented in Table 2, which reflect the model performance across all matrices.

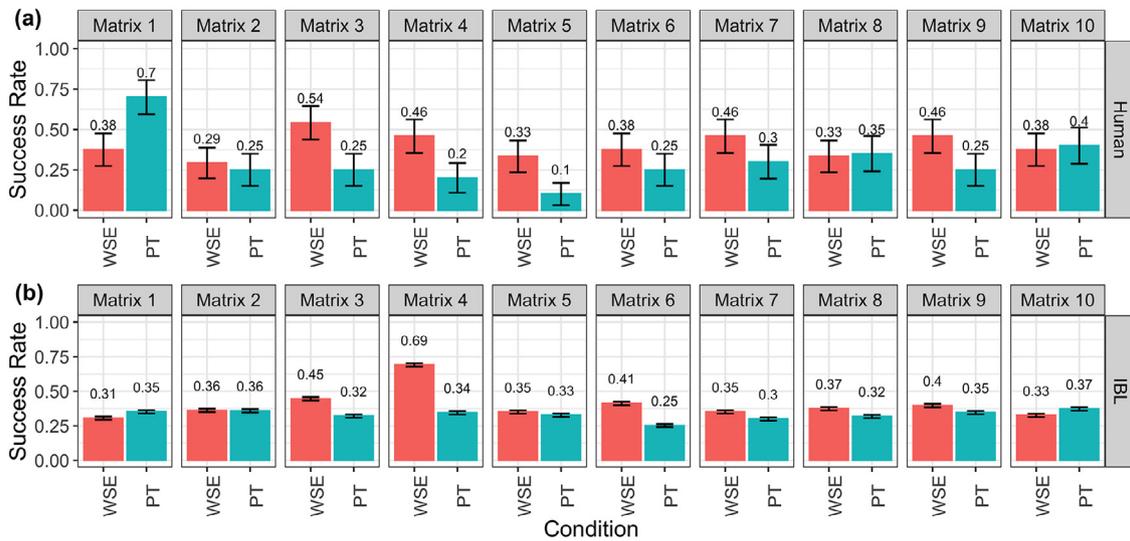


Fig. B.2. Success Rate in PT and WSE for IBL Model (matrix-wise).

### Appendix C. Survey Questions

#### Screening Test

The following questions relate to your practical knowledge in network and information security. We have marked correct answers in bold-italic font for the readers.

Note: In order to participate in the main study, you need to get 7 out 10 questions right. Please attempt carefully.

**Q1: What does the “https://” at the beginning of a URL denote, as opposed to “http://” (without the “s”)?**

- That the site has special high definition
- That information entered into the site is encrypted
- That the site is the newest version available
- That the site is not accessible to certain computers
- None of the above
- Not sure

**Q2: Criminals access someone’s computer and encrypt the user’s personal files and data. The user is unable to access this data unless they pay the criminals to decrypt the files. This practice is called:**

- Botnet
- Ransomware
- Driving
- Spam
- None of the above
- Not sure

**Q3: Which of the following statements are true?**

- Secure File Transfer Protocol (SFTP) runs by default on port 22
- Secure Shell (SSH) runs by default on port 22
- File Transfer Protocol over TLS/SSL (FTPS) runs by default on port 22
- Trivial File Transfer Protocol (TFTP) runs by default on port 22

**Q4: TCP port 80 is assigned to:**

- Hypertext Transfer Protocol (HTTP)
- Hypertext Transfer Protocol over TLS/SSL (HTTPS)
- Internet Message Access Protocol (IMAP)
- Lightweight Directory Access Protocol (LDAP)

**Q5: A command-line tool that can be used for banner grabbing is called:**

- tcpdump
- netcat
- Nmap
- Wireshark

**Q6: Which of the command-line utilities listed below can be used to perform a port scan?**

- Zenmap
- Nmap
- tcpdump
- nslookup

**Q7: Zero-day attack exploits are:**

- New accounts
- Patched software
- Vulnerability that is present in already released software but unknown to the software developer
- Well known vulnerability

**Q8: Which of the following is not checked by the Nmap command?**

- services different hosts are offering
- what OS they are running
- what kind of firewall is in use
- what type of antivirus is in use

**Q9: What are the port states determined by Nmap?**

- Active, inactive, standby
- Open, half-open, closed
- Open, filtered, unfiltered
- Active, closed, unused

**Q10: Which of the following is not an objective of scanning?**

- Detection of the live system running on network
- Discovering the IP address of the target system
- Discovering the services running on target system
- Detection of spyware in a system

## Post-Survey Questions

**Q1: What strategy did you use for attacking a machine?**

- Randomly chose among available options
- Calculated the probability of a OC being TC using phi matrix
- Carefully studied the nmap command output
- Used additional commands to identify the true configuration
- Attacked the option with highest payoff
- Attacked the option with sure success

**Q2: What strategy did you use for scanning the machines in each round before launching an attack?**

- Scanned all the systems and carefully studied output of nmap command
- Scanned one machine at time and carefully studied output of nmap command
- Did not scan the machines

**CRedit authorship contribution statement**

**Palvi Aggarwal:** Conceptualization, Methodology, Methodology, Data curation, Formal analysis, Writing – original draft. **Omkar Thakoor:** Conceptualization, Investigation, Writing – original draft. **Shahin Jabbari:** Conceptualization, Investigation, Writing – review & editing. **Edward A. Cranford:** Conceptualization, Formal analysis, Writing – original draft. **Christian Lebiere:** Conceptualization, Formal analysis, Writing – original draft. **Milind Tambe:** Conceptualization, Investigation, Writing – review & editing. **Cleotilde Gonzalez:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft.

**References**

- Aggarwal, P., Gonzalez, C., Dutt, V., 2017. Modeling the effects of amount and timing of deception in simulated network scenarios. In: 2017 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA). IEEE, pp. 1–7.
- Aggarwal, P., Moisan, F., Gonzalez, C., Dutt, V., 2020a. Learning about the effects of alert uncertainty in attack and defend decisions via cognitive modeling. *Hum. Factors*. 0018720820945425
- Aggarwal, P., Thakoor, O., Mate, A., Tambe, M., Cranford, E.A., Lebiere, C., Gonzalez, C., 2020b. An exploratory study of a masking strategy of cyberdeception using cyberVAN. *HFES*.
- Alpcan, T., Başar, T., 2010. Network security: a decision and game-theoretic approach.
- Anderson, J.R., 1996. ACT: a simple theory of complex cognition. *Am. Psychol.* 51 (4), 355.
- Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y., 2004. An integrated theory of the mind. *Psychol. Rev.* 111 (4), 1036.
- Ben-Asher, N., Gonzalez, C., 2015. Effects of cyber security knowledge on attack detection. *Comput. Hum. Behav.* 48, 51–61.
- Bos, N., Paul, C.L., Gersh, J.R., Greenberg, A., Piatko, C., Sperling, S., Spitaletta, J., Arendt, D.L., Burtner, R., 2016. Effects of gain/loss framing in cyber defense decision-making. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 60. SAGE Publications Sage CA: Los Angeles, CA, pp. 168–172.
- Breton, M., Alj, A., Haurie, A., 1988. Sequential stackelberg equilibria in two-person games. *J. Optim. Theory Appl.*
- Chadha, R., Bowen, T., Chiang, C.Y.J., Gottlieb, Y.M., Poylisher, A., Sapello, A., Serban, C., Sugrim, S., Walther, G., Marvel, L.M., et al., 2016. CyberVAN: a cyber security virtual assured network testbed. In: MILCOM 2016-2016 IEEE Military Communications Conference. IEEE, pp. 1125–1130.
- Chicoisne, R., Ordóñez, F., 2016. Risk averse stackelberg security games with quantal response. In: International Conference on Decision and Game Theory for Security. Springer, pp. 83–100.
- Cohen, F., 1998. A note on the role of deception in information protection. *Comput. Secur.* 17 (6), 483–506.
- Cooney, S., Vayanos, P., Nguyen, T.H., Gonzalez, C., Lebiere, C., Cranford, E.A., Tambe, M., 2019a. Warning time: optimizing strategic signaling for security against boundedly rational adversaries. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, pp. 1892–1894.
- Cooney, S., Wang, K., Bondi, E., Nguyen, T., Vayanos, P., et al., 2019b. Learning to signal in the goldilocks zone: improving adversary compliance in security games. *ECML/PKDD*.
- Cranford, E., Gonzalez, C., Aggarwal, P., Cooney, S., Tambe, M., Lebiere, C., 2020. Adaptive cyber deception: cognitively informed signaling for cyber defense. In: Proceedings of the 53rd Hawaii International Conference on System Sciences.
- 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. 2018.
- Cranford, E.A., Lebiere, C., Rajivan, P., Aggarwal, P., Gonzalez, C., 2019. Modeling cognitive dynamics in (end)-user response to phishing emails. In: Proceedings of the 17th Annual Meeting of the International Conference on Cognitive Modelling, Montreal, CA.
- Cranford, E.A., Singh, K., Aggarwal, P., Lebiere, C., Gonzalez, C., 2021. Modeling phishing susceptibility as decisions from experience. In: Proceedings of the 19th Annual Meeting of the International Conference on Cognitive Modelling, Montreal, CA.
- De Gaspari, F., Jajodia, S., Mancini, L.V., Panico, A., 2016. AHEAD: a new architecture for active defense. *SafeConfig*.
- Dutt, V., Gonzalez, C., 2015. Accounting for Outcome and Process Measures in Dynamic Decision-Making Tasks through Model Calibration. Technical Report. Carnegie Mellon University Pittsburgh United States.
- Ferguson-Walter, K., LaFon, D., Shade, T., 2017. Friend or faux: deception for cyber defense. *J. Inf. Warfare*.
- Gigerenzer, G., Todd, P.M., 1999. Simple Heuristics That Make Us Smart. Oxford University Press, USA.
- Goel V., Perliroth N., Yahoo says 1 billion user accounts were hacked; 2016. <https://www.nytimes.com/2016/12/14/technology/yahoo-hack.html>.
- Gonzalez, C., Aggarwal, P., Lebiere, C., 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., 2014. Learning to cooperate in the prisoner's dilemma: robustness of predictions of an instance-based learning model. In: Proceedings of the Annual Meeting of the Cognitive Science Society, vol. 36.
- Gonzalez, C., Lerch, J.F., Lebiere, C., 2003. Instance-based learning in dynamic decision making. *Cogn. Sci.* 27 (4), 591–635.
- Gutzmer I. Equifax announces cybersecurity incident involving consumer information; 2017. <https://investor.equifax.com/news-and-events/news/2017/09-07-2017-213000628>.
- Gutzwiller, R., Ferguson-Walter, K., Fugate, S., Rogers, A., 2018. "Oh, Look, A Butterfly!" a framework for distracting attackers to improve cyber defense. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 62. SAGE Publications Sage CA: Los Angeles, CA, pp. 272–276.
- Gutzwiller, R.S., Ferguson-Walter, K.J., Fugate, S.J., 2019. Are cyber attackers thinking fast and slow? Exploratory analysis reveals evidence of decision-making biases in red teamers. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 63. SAGE Publications Sage CA: Los Angeles, CA, pp. 427–431.
- Heckman, K.E., Walsh, M.J., Stech, F.J., O'boyle, T.A., DiCato, S.R., Herber, A.F., 2013. Active cyber defense with denial and deception: a cyber-wargame experiment. *Comput. Secur.* 37, 72–77.
- Kahneman, D., 2003. A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58 (9), 697.
- Kahneman, D., Tversky, A., 1979. On the Interpretation of Intuitive Probability: A Reply to Jonathan Cohen. Elsevier Science.
- Laszka, A., Vorobeychik, Y., Koutsoukos, X.D., 2015. Optimal personalized filtering against spear-phishing attacks. *AAAI*.
- Lebiere, C., Blaha, L.M., Fallon, C.K., Jefferson, B., 2021. Adaptive cognitive mechanisms to maintain calibrated trust and reliance in automation. *Front. Rob. AI* 8, 135.
- Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., Anderson, J.R., 2013. A functional model of sensemaking in a neurocognitive architecture. *Comput. Intell. Neurosci.* 2013.
- Lejarraga, T., Dutt, V., Gonzalez, C., 2012. Instance-based learning: a general model of repeated binary choice. *J. Behav. Decis. Mak.* 25 (2), 143–153.
- Lemay, A., Leblanc, S., 2018. Cognitive biases in cyber decision-making. In: Proceedings of the 13th International Conference on Cyber Warfare and Security, p. 395.
- Nguyen, T.N., Gonzalez, C., 2021. Theory of mind from observation in cognitive models and humans. *Top. Cogn. Sci.*
- Pita, J., John, R., Maheswaran, R., Tambe, M., Kraus, S., 2012a. A robust approach to addressing human adversaries in security games. In: *ECAL*, pp. 660–665.
- Pita, J., John, R., Maheswaran, R., Tambe, M., Yang, R., Kraus, S., 2012b. A robust approach to addressing human adversaries in security games. In: *AAMAS*, pp. 1297–1298.
- Provos, N., 2003. Honeyd-a virtual honeypot daemon. In: 10th DFN-CERT Workshop, Hamburg, Germany, vol. 2, p. 4.
- Sawyer, B.D., Hancock, P.A., 2018. Hacking the human: the prevalence paradox in cybersecurity. *Hum. Factors* 60 (5), 597–609.
- Schlenker, A., Thakoor, O., Xu, H., Fang, F., Tambe, M., Tran-Thanh, L., Vayanos, P., Vorobeychik, Y., 2018. Deceiving cyber adversaries: a game theoretic approach. *AAMAS*.
- Schlenker, A., Xu, H., Guirguis, M., Kiekintveld, C., Sinha, A., Tambe, M., Sonya, S., Balderas, D., Dunstatter, N., 2017. Don't bury your head in warnings: A game-theoretic approach for intelligent allocation of cyber-security alerts. *IJ-CAL*.
- Serra, E., Jajodia, S., Pugliese, A., Rullo, A., Subrahmanian, V.S., 2015. Pareto-optimal adversarial defense of enterprise systems. *ACM Trans. Inf. Syst. Secur. (TISSEC)* 17 (3), 11.
- Simon, H.A., 1956. Rational choice and the structure of the environment. *Psychol. Rev.* 63 (2), 129.

- Sycara, K., Lebiere, C., Pei, Y., Morrison, D., Tang, Y., Lewis, M., 2015. Abstraction of analytical models from cognitive models of human control of robotic swarms. In: *Proceedings of ICCM 2015-13th International Conference on Cognitive Modeling*. University of Pittsburgh, pp. 13–18.
- Thakoor, O., Jabbari, S., Aggarwal, P., Cleotilde, G., Tambe, M., Vayanos, P., 2020. Exploiting bounded rationality in risk-based cyber camouflage games. In: *International Conference on Decision and Game Theory for Security*.
- Thakoor, O., Tambe, M., Vayanos, P., Xu, H., Kiekintveld, C., Fang, F., 2019a. Cyber camouflage games for strategic deception. In: *International Conference on Decision and Game Theory for Security*. Springer, pp. 525–541.
- Thakoor, O., Tambe, M., Vayanos, P., Xu, H., Kiekintveld, C., Fang, F., 2019b. Cyber camouflage games for strategic deception. *GameSec*.
- Thinkst. Canary; 2015. <https://canary.tools/>.
- Trafton, J.G., Hiatt, L.M., Brumback, B., McCurry, J.M., 2020. Using cognitive models to train big data models with small data. In: *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1413–1421.
- Tversky, A., Kahneman, D., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47 (2), 263–291.
- Yang, R., Kiekintveld, C., Ordóñez, F., Tambe, M., John, R., 2011. Improving resource allocation strategy against human adversaries in security games. In: *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22. Cite-seer, p. 458.

**Palvi Aggarwal** is a Postdoctoral Research Fellow at the Dynamic Decision Making Lab, Carnegie Mellon University. She earned her Ph.D in cybersecurity from Indian Institute of Technology, Mandi, India. Prior to IIT Mandi, she did a master's in information security and a bachelor's degree in Computer Science. Her research interests broadly include the application of human factors and cognitive modeling to cybersecurity and human-machine teaming. Palvi studies how experiments and cognitive models could be used to understand the behavioral aspects of cybersecurity actors, i.e., attackers, defenders, and end-users to develop better defense algorithms and a safe cyber space.

**Omkar Thakoor** is a fourth year Ph.D student in the Computer Science department at the University of Southern California. He previously earned his master's degree in Computer Science at the University of Illinois at Urbana-Champaign, and bachelor's degree from Indian Institute of Technology, Bombay. His research aims at providing mathematical modelling and analysis for diverse real-world problems, notably cyber security, and primarily deploying techniques from Artificial intelligence and Game theory.

**Shahin Jabbari** is a CRCS postdoctoral fellow in the School of Engineering and Applied Sciences at Harvard hosted by Milind Tambe. He recently completed his Ph.D in the Computer and Information Science Department at University of Pennsylvania

where he was advised by Michael Kearns. Shahin study the interactions between machine learning and a variety of contexts, ranging from crowdsourcing to game theory, AI for social good and algorithmic fairness.

**Edward A. Cranford** earned his Ph.D in Cognitive Science from Mississippi State University in 2016 and is currently a postdoc in the Department of Psychology at Carnegie Mellon University, in the Functional Modeling Systems group, directed by Christian Lebiere. His research interests broadly include comprehension, prediction/anticipation, problem-solving, learning, and decision making, and the application of cognitive models to human-machine interactions. Drew's current research is focused on understanding and modeling the human decision making of adversaries in a cyber-security domain, modeling end-user responses to phishing attacks, and developing adaptive, personalized interventions. In other research, he investigates how experts generate and select appropriate courses of action in dynamic and time-pressured situations.

**Christian Lebiere** is a Research Faculty in the Psychology Department at Carnegie Mellon University, having received his Ph.D from the CMU School of Computer Science. During his graduate career, he studied connectionist models and was the coauthor with Scott Fahlman of the Cascade-Correlation neural network learning algorithm. Since 1991, he has worked on the development of the ACT-R cognitive architecture and was co-author with John R. Anderson of the 1998 book *The Atomic Components of Thought*. Most recently, he has been involved with John Laird and Paul Rosenbloom in defining the Common Model of Cognition, a community-wide effort to consolidate and formalize the scientific progress resulting from the 40-year research program in cognitive architectures.

**Milind Tambe** is Gordon McKay Professor of Computer Science and Director of Center for Research on Computation and Society at Harvard University; he is also Director "AI for Social Good" at Google Research India. He is a fellow of AAAI (Association for Advancement of Artificial Intelligence), ACM (Association for Computing Machinery) and has received the IJCAI John McCarthy Award, as well as ACM SIGART Autonomous Agents Research Award. Previous to his positions at Harvard and Google, he was Helen N. and Emmett H. Jones Professor in Engineering and a Professor of Computer Science and Industrial and Systems Engineering at the University of Southern California, Los Angeles.

**Cleotilde Gonzalez** is a Research Professor in the Department of Social and Decision Sciences at CMU. She earned a Ph.D in Management Information Systems from Texas Tech University in 1996. Her research lies at the intersection of Human Behavioral Decision Making and Technology. Her research program is motivated by real-world decision making and by the challenges involved in studying dynamic decision making in the laboratory. Her research is embedded within a theoretical framework that emphasizes the role and development of decisions from experience, the similarity of contexts, and the cognitive abilities of decision makers.