Behavioral Cybersecurity: Investigating the influence of Patching Vulnerabilities in Markov Security Games via Cognitive Modeling

Zahid Maqbool[1], V. S. Chandrasekhar Pammi[2] and **Varun Dutt[1]**

[1]Applied Cognitive Science Laboratory, Indian Institute of Technology Mandi, India – 175005
[2]Centre of Behavioral and Cognitive Sciences, University of Allahabad, India – 211002

ABSTRACT

Current research in cyber-security is not focused on human decision-making. The primary objective of this study is to address this gap and investigate how cognitive processes proposed by Instance-based Learning Theory (IBLT) like reliance on recency and frequency, attention to opponent's actions, and cognitive noise are influenced by the effectiveness of vulnerability patching. Data involving participants performing as hackers and analysts was collected in a lab-based experiment in two patching conditions: effective (N = 50) and less-effective (N = 50). In effective (less-effective) patching, computer systems were in a non-vulnerable state (i.e., immune to cyber-attacks) 90% (50%) of the time after patching. An IBL model accounted for human decisions and revealed low (high) reliance on recency and frequency, attention to opponent's actions, and cognitive noise for hacker (analyst) in effective patching. Whereas, it revealed opposite results for less-effective patching. We highlight the implications of our findings for cyber decision-making.

Published online: 2nd December 2019

Keywords: *Analyst, Attack, Defend, Patching, Cyber Security, Hacker, Markov security games, Nash equilibrium, Instance-Based Learning Theory (IBLT), Cognitive Model*

1 INTRODUCTION

There has been a spurt in Internet growth recently and Internet is now being pervasively used across different socio-economic sectors (Humayed et al., 2017; Keller & Schaninger, 2019). With Internet's growth, protecting online data from illegal entry has become difficult (Economic Times, 2017). Hackers, people who attack computer networks, are finding newer ways of exploiting vulnerabilities present on computer systems (TechTarget, 2017). To remove vulnerabilities and safeguard against cyber-attacks, security-analysts, people who protect computer systems, may implement software fixes (patches) against vulnerabilities (Florida Tech, 2019). These software patches may be effective, and these may help remove vulnerabilities present in computer systems (Kissel, 2013). However, these software patches may also be less effective as they may partially remove vulnerability or introduce newer vulnerabilities in computer systems (Dunagan et al., 2004). Thus, a less effective software patch may remove vulnerabilities from only a small part of the computer system, or it may create newer vulnerabilities in a larger part of the system (Grimes, 2016). As these existing or newly created vulnerabilities may impact our subsequent patching decisions, it is important to study the influence of the effectiveness of the patching processes on decision-making of human analysts and hackers (Florida Tech, 2019).

Prior research has proposed game theory to be a promising tool to study human decision-making in cyber-attack situations (Roy et al., 2010). As per prior research, the influence of the effectiveness of the patching processes on cyber decision-making may be studied using Markov security games (Alpcan & Başar, 2006; 2010). In the Markov security game, human players perform as hackers and analysts, where hackers may take attack/not-attack actions and analysts may take defend (patch)/not-defend (not-patch) actions. As a result of both players' actions, both players may obtain payoffs (outcomes) and the interaction between hackers and analysts is recurrent over rounds. As per the Markov assumption, analyst's action in the last round influences the vulnerability of the computer system to an attack in the current round. In most cases, patching of vulnerabilities may improve the security of a computer system (i.e., patching may make the computer system non-vulnerable to cyber-attacks); however, in some cases patching may also lead to unresolved

vulnerabilities (i.e., patching may be less-effective making the computer system vulnerable to attacks).

Preliminary research has analyzed optimal decision-making in Markov security games (Xiaolin et al., 2008). According to reference (Xiaolin et al., 2008), in the absence of patches, cyber-attacks could produce damages that become problematic as the attacks spread among computer systems. In contrast, damages to computer systems become smaller when analysts are able to timely patch vulnerabilities present in computer systems (Xiaolin et al., 2008). These findings are in agreement the Markov security games dynamics. Using Markov security games, Xiaolin et al. (2008) have derived predictions about the Nash equilibria using mathematical simulation techniques; however, these authors did not attempt an empirical investigation of human actions against Nash predictions.

Building upon Xiaolin et al. (2008)'s limitations, Maqbool et al. (2018) investigated the influence of the patching process on the attack-and-defend decisions of human hackers and analysts. These authors found that the human attack-and-defend proportions deviated significantly from their Nash proportions across cases when the patching process was effective and less-effective. Maqbool et al. (2018) explained their results based upon Instance-based Learning Theory (IBLT; Gonzalez & Dutt, 2011; 2012; Gonzalez, Lerch, & Lebiere, 2003; Dutt & Gonzalez, 2012), a theory of decisions from experience. In cognitive literature, models built upon IBLT (referred to as "IBL models" hereafter) have successfully accounted for human decisions in cyber scenarios (Aggarwal et al., 2018; Arora & Dutt, 2013; Dutt, Ahn, & Gonzalez, 2013; Kaur & Dutt, 2013). These IBL models possess cognitive limitations about memory and recall (Gonzalez & Dutt, 2010). For example, these models assume that while making decisions, people may rely upon blending of the most recent and frequent experiences (instances) retrieved from memory. Also, these models may assume decision-makers to pay attention to opponent's actions and cognitive noise (Arora & Dutt, 2013). Thus, mechanisms like reliance on recency and frequency, attention to opponent's actions, and cognitive noise may influence hacker's decisions against varying effectiveness of patching processes.

Although Maqbool et al. (2018) provided an explanation of their results using cognitive assumptions of IBLT, these authors did not explicitly build IBL models for hackers and analysts using their experimental data. In this research, we overcome this limitation by building cognitive models using IBLT, which could provide a cognitive explanation to results obtained by Maqbool et al. (2018). Overall, the primary objective of this study is to

investigate how cognitive processes proposed by IBLT like reliance on recency and frequency, attention to opponent's actions, and cognitive noise are influenced by the effectiveness of patching processes. The IBL models use 2-players Markov security games and human data collected in a lab-based experiment by Maqbool et al. (2018) to investigate the differences in cognitive processes between effective and less-effective patching processes. Overall, the application of IBLT to the analyst's and hacker's experiential decisions in Markov security games is novel and it will enable us to explain how these decisions are influenced by the effectiveness of patching process. Also, it will enable us to explain how, on account of limitations of memory and recall, human decisions deviate from their Nash proportions.

This interdisciplinary research presents theoretical perspectives from cognitive science, which currently are relatively rare in cybersecurity research. The majority of current research in cybersecurity is not informed by theories from cognitive science, nor focused on human decision-making. Furthermore, the current research relies on experimental methods that are capable of being reproduced and extended by others; whereas, the majority of current research in cybersecurity uses methods that are not suitable for reproduction or extension. Overall, this research proposes cognitive models for hackers and analysts that may be reused and refined as part of future research.

In what follows, we first introduce the Markov security game and the Nash equilibria for attack and defend (or patch) proportions. Next, we report an experiment by Maqbool et al. (2018), where these authors varied the effectiveness of the patching process in repeated Markov-security games. Furthermore, we report the results of our analyses of the experimental data as well as how IBLT could help explain human decisions against varying effectiveness of patching processes. More specifically, we create IBL models for hackers and analysts on the collected dataset and test the ability of these models to account for decisions of human hackers and analysts. We present results from our models and discuss the cognitive mechanisms used by human participants while performing against varying effectiveness of the patching process.

2 THE MARKOV SECURITY GAME

The Markov security game (see Figure 1) (Alpcan & Başar, 2010; Xiaolin et al., 2008) is a repeated 2 x 2 zero-sum game. Two opponents, hacker and analyst, play against each other in this game. The objective for both opponents is to maximize individual payoffs by repeatedly making decisions over

several rounds (the end-point is unknown to both opponents). The hacker can take an attack (a) and a not-attack (na) action; whereas, the analyst can take a defend (d) and a not-defend (nd) action. Attack actions correspond to attacking a computer system; whereas, defend actions correspond to patching vulnerabilities on a computer system. When the game is played between human players, one human player is randomly asked to perform as the hacker and the other human player is asked to perform as the analyst.

As shown in Figure 1(a), there are two possible states for a given set of actions available to hackers and analysts, vulnerable (v) and not-vulnerable (nv). In the v state, the probability of hacker to penetrate the computer system is very high. On the contrary, in non-vulnerable state, the probability of hacker to penetrate the computer system is low. The transition between the v and nv states is determined by the analyst's last decision (d or nd). If the analyst chooses to patch the computer system (i.e., initiate a d action) in a round t, then this patching action likely increases the computer system's probability of being in the nv state in round t+1. In contrast, if the analyst does not patch the computer (i.e., the analyst performs a nd action) in a round t, then the nd action increases the computer system's probability of being in the v state in round t+1.

The movement of state v to state nv or from state nv to state v between two successive rounds depend on the patching process' effectiveness. The probability of transiting from the state nv to the state v is small (= 0.1) and the probability of transiting from state v to state nv is large (= 0.8), if the patching process is effective. However, the probability of transiting from state nv to state v and from state v to state nv are equal (= 0.5), if the patching is less-effective. The following Markov process determines the probability of each state in a round t:

$$\text{Prob}(t) = M(.) * \text{Prob}(t-1) \qquad (1)$$

Where, Prob (t) and Prob (t - 1) refer to probabilities of being in states v and nv in round t and t – 1, respectively. Similarly, M (.) refers to state-transition matrix corresponding to different analyst actions (see Figure 1(a)). The probability of being in states v or nv at the start of the game is made equally likely (= 0.5):

$$\text{Prob}(1) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \qquad (2)$$

Where, the values in the first and second rows correspond to the v state and the nv state probabilities, respectively.

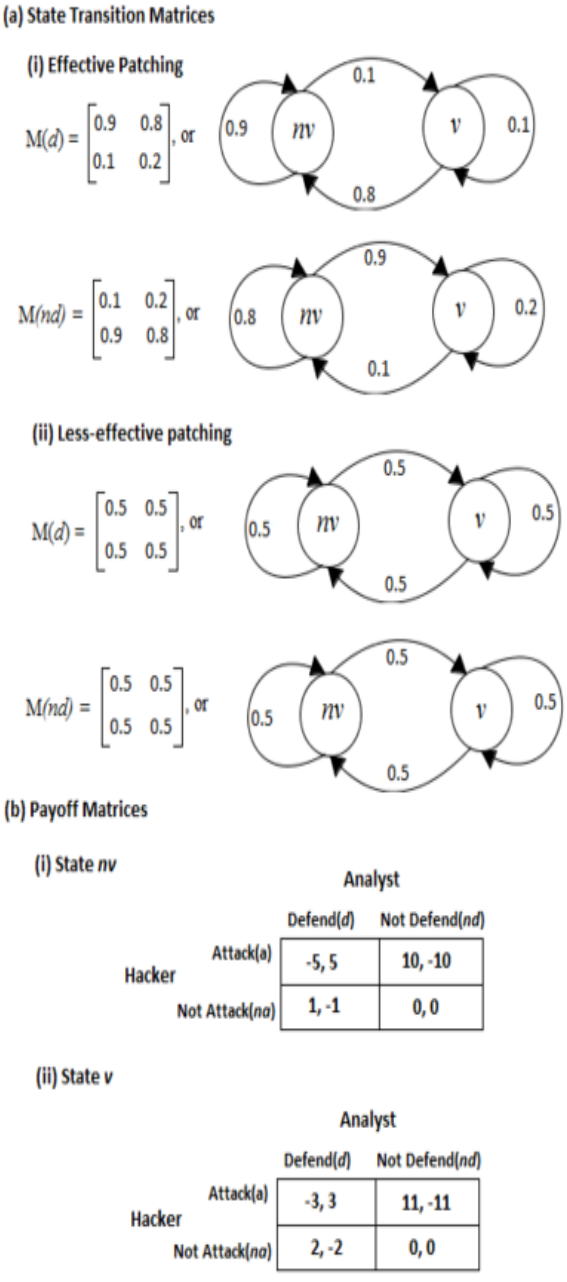

Figure 1. The Markov security game (Maqbool et al., 2018). (a) The M(.) matrices showing transitions between non-vulnerable (nv) and vulnerable

(v) states for different patching conditions. (b) The payoffs corresponding to nv and v states. In each cell, the first payoff is for the hacker and the second payoff is for the analyst.

As shown in Figure 1(b), there are separate sets of zero-sum payoffs associated with each state v and nv due to opponent's individual actions. For example, in state v, an a - d action results in +5 points for the analyst and a - 5 points for the hacker (the hacker is caught while attacking the computer system due to patching). For a - nd action, analysts get -10 points and hackers get +10 points. Similarly, Figure 1(b) shows payoffs associated with other action combinations. Upon comparing the payoffs in states v and nv, one would find larger (smaller) penalties and smaller (larger) benefits for the hackers (analysts) in the state v (nv).

Using payoffs in Figure 1(b), one could compute the mixed strategy Nash equilibria for the v and nv states, respectively. Let p represent the attack proportions and 1 - p represent the not-attack proportions. Similarly, let q represent the defend (patch) proportions and 1- q represent the not-defend proportions. In the absence of a pure Nash strategy equilibrium, both opponents would be indifferent between the payoffs from their actions. Thus, we get the following Nash proportions:
For the state *v*:

$$3*p - 2*(1 - p) = -11*p + 0 \text{ and } -3*q + 11*(1 - q) = 2*q + 0 \qquad (3)$$

$$\Rightarrow p = 1/8\ (= 0.125) \text{ and } q = 11/16\ (= 0.687)$$

For the state *nv*:

$$5*p - 1*(1 - p) = -10*p + 0 \text{ and } -5*q + 10*(1 - q) = 1*q + 0 \qquad (4)$$

$$\Rightarrow p = 1/16\ (= 0.062) \text{ and } q = 5/8\ (= 0.625)$$

These Nash equilibria proportions were compared against human action proportions in a lab-based experiment performed by Maqbool et al. (2018).

3 EXPECTATIONS IN THE MARKOV SECURITY GAME

According to IBLT, people tend to maximize their perceived payoff across actions (Gonzalez & Dutt, 2011; 2012; Gonzalez, Lerch, & Lebiere, 2003; Lejarraga, Dutt, & Gonzalez, 2012). In IBLT, the perceived payoffs are determined by the blended values computed for different actions (Lejarraga, Dutt, & Gonzalez, 2012). As opponents performing as hackers and analysts would experience different payoffs across the effective and less-effective patching conditions, these players would likely possess dissimilar perceived

payoffs in both conditions. Thus, based upon IBLT, we expect differences in cognitive parameters concerning reliance upon recency and frequency of outcomes, attention to opponent's actions, and cognitive noise across different effective and less-effective patching conditions. Furthermore, according to IBLT, we expect human decisions to deviate significantly from their Nash proportions across different patching conditions. That is because, human opponents would possess cognitive limitations on memory and recall processes and human beings would tend to rely upon recency and frequency of outcomes, attention to opponent's actions, and cognitive noise to make their repeated decisions. The reliance upon recency and frequency processes would likely not allow opponents to form optimal Nash expectations for their actions. In the next section, we detail an experiment performed by Maqbool et al. (2018), which allowed us to test different expectations from IBLT.

4 EXPERIMENT

In this section, we report a lab-based experiment performed by Maqbool et al. (2018) involving people performing as hackers and analysts in the Markov security game (Figure 1). Using data collected in the experiment, we develop cognitive models and investigate how the effectiveness of the patching process influences the decisions of human players.

Experimental Design

Maqbool et al. (2018) randomly assigned 100 participants to one of two between-subjects patching conditions: effective (N = 50) and less-effective (N = 50). In each condition, 25 participants performed as hackers; whereas, 25 participants performed as analysts. The game was 50-rounds in length in each condition and it involved an interaction between participants performing as hackers and analysts in real time.

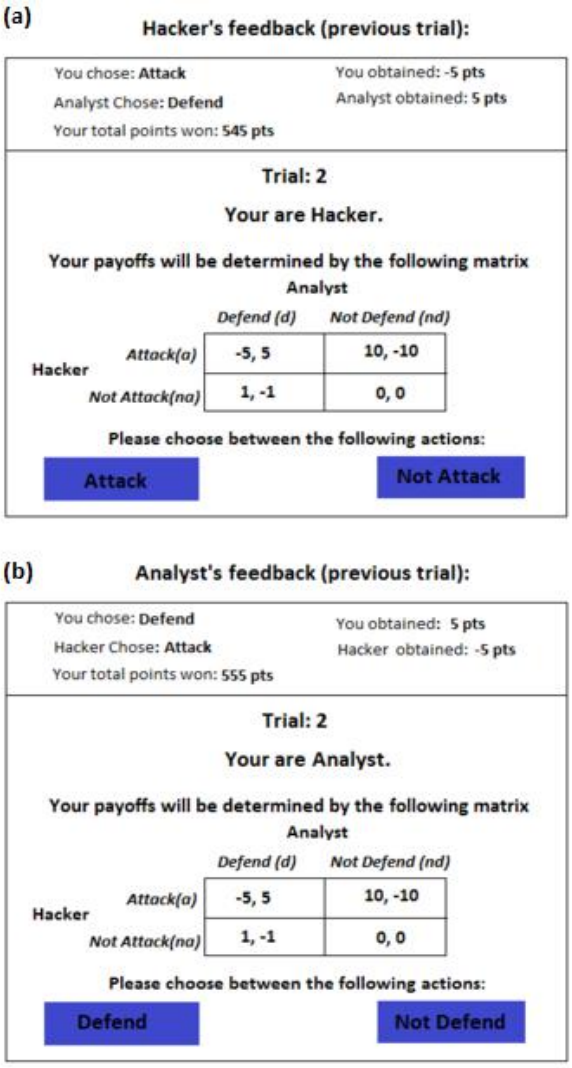

Figure 2. The Graphical User Interface shown to participants acting in the roles of hackers(a) and analysts (b) across different patching conditions.

Figure 2 shows the graphical user interface shown to participants performing in hacker (A) and analyst (B) roles across both conditions. In a round, participants performing as hackers and analysts were shown the actions chosen by them and their opponents, current payoffs obtained by them and their opponents, and total payoffs obtained by them since the start of the game (see Figure 2). Both hacker and analyst roles were also presented with the

payoff matrices resulting in different states (v or nv) in each round, and they were asked to choose between attack/not-attack and defend/not-defend actions. The payoff changed for both players across rounds depending upon whether the computer system was in the v state or nv state (the two payoff matrices are shown in Figure 1(b)).

For testing our expectations, we reanalyse the data of Maqbool et al. (2018) and we compare the proportion of attack and defend actions from human players across different conditions and states. Furthermore, we compare human action proportions with the corresponding Nash action proportions (computed in equations 3 and 4). We use mixed-factorial ANOVAs for testing our expectations. Also, we perform t-tests to compare human and Nash proportions in different states and conditions. For our analyses, we use an alpha level of 0.05 and power level of 0.8 across all statistical comparisons.

Participants

In data collected by Maqbool et al. (2018), seventy-nine percent of participants were males. Ages ranged from 18 years to 30 years (Mean = 21.2 years and standard deviation = 1.92 years). Participants were from different education levels: 74% undergraduates and 26% graduates. All participants were from Science, Technology, Engineering, and Mathematics (STEM) backgrounds. Discipline-wise the demographics were the following: 42% pursing degrees in computer-science and engineering, 18% pursing degrees in mechanical engineering, 38% pursing degrees in electrical engineering, and 2% pursing degrees in basic sciences. Participants were asked to maximize their payoffs and were compensated a flat participation fee of INR 30 (~ USD 0.5). In addition, participants could get up to INR 20 based on their performance. For calculating the performance incentive, a participant's final score was converted to real money in the following ratio: 55 points = INR 1.0. No participant took more than 20 minutes to finish the study.

Procedure

Maqbool et al. (2018) recruited their participants through an email advertisement, where participation was completely voluntary. Participants gave their written consent before starting their study and the study was approved by the ethics committee at the Indian Institute of Technology Mandi. Participants were given instructions about the goal in the task (to maximize their total payoff) and they were instructed about the game's working. As part of the instructions, payoff matrices as well as actions possible were explained to participants. Questions in the instructions, if any, were answered before participants could begin their study. Participants

possessed complete information about their own and their opponent's actions and payoffs in all conditions (the payoff matrices were given to both players). In a round, both participants decided their actions simultaneously and then received feedback about each other's actions and payoffs. After feedback, participants were asked to make the next round's decision. Once the study ended, participants were thanked and given their participation fee.

5 RESULTS

Proportion of attack and defend actions across conditions

In our reanalysis of Maqbool et al. (2018) data, we first calculated the attack-and-defend proportions in each patching condition (see Figure 3). As shown in Figure 3, for the hacker, there was no significant differences in the attack proportions between less-effective condition and the effective condition (0.31 ~ 0.36; $F(1, 49) = 0.32$, $p = .57$, $\eta p^2 = .007$). Furthermore, for the analyst, there was again no significant difference in the defend (patching) proportions between the effective condition and the less-effective condition (0.67 ~ 0.69; $F(1, 49) = 0.13$, $p = .71$, $\eta p^2 = .003$). Thus, overall, the proportion of action was similar in effective and less-effective conditions.

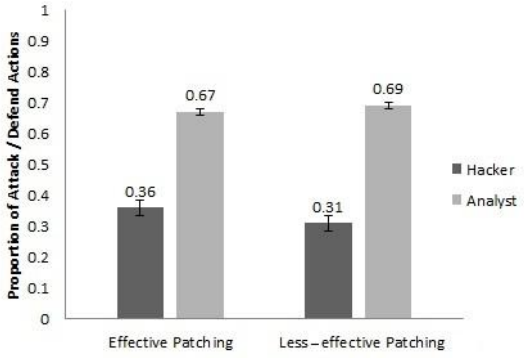

Figure 3. Proportion of attack and defend actions across the two conditions.

Proportion of attack and defend actions across states

By reanalyzing Maqbool et al. (2018)'s data, we compared the human proportion of attack-and-defend actions with the respective Nash proportions in the two states, *v* and *nv* (see Figure 4). For hackers, the attack proportions were significantly different from their Nash proportions across both the *v* and

nv states (state *v*: $t(49) = 10.34$, $p < .05$, $r = .82$; state *nv*: $t(49) = 7.563$, $p < .05$, $r = .73$). For analysts, the defend proportions were not significantly different from their Nash proportions in the *v* state ($t(49) = -0.481$, $p = .63$, $r = .068$); however, the defend proportions were significantly different from their Nash proportion in the *nv* state ($t(49) = 3.040$, $p < .05$, $r = .40$). Thus, overall, these results agree with our expectations from IBLT.

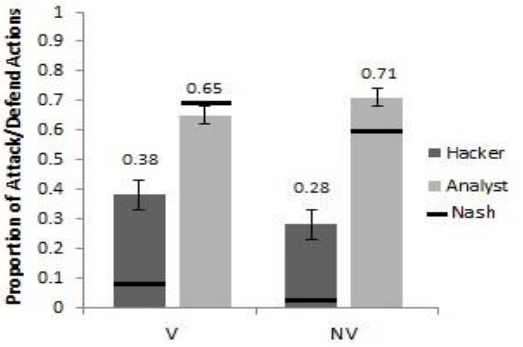

Figure 4. Proportion of attack and defend actions across the states.

Proportion of attack and defend actions across patching conditions and states

In our reanalysis of Maqbool et al. (2018), we next evaluated how the proportion of attack-and-defend actions differed from their Nash proportions across the two patching conditions and the two network states. Figure 5 shows the proportion of attack-and-defend actions across the patching conditions and states with respect to the Nash proportions. For hackers, the proportion of attack actions were significantly different from their Nash proportions across all conditions and states (effective and state *v*: $t(49) = 12.43$, $p < .05$, $r = .87$; effective and state *nv*: $t(49) = 14.56$, $p < .05$, $r = .90$); less-effective and state *v*: $t(49) = 12.40$, $p < .05$, $r = .87$; and, less-effective and state *nv*: $t(49) = 12.12$, $p < .05$, $r = .86$). Thus, these results for hackers agree with our expectations from IBLT. For analysts, the defend proportions were not significantly different from their Nash proportions in the *v* state across both effective and less-effective conditions (effective: $t(49) = -1.95$, $p = .06$, $r = .26$; less-effective: $t(49) = -1.34$, $p = .18$, $r = .18$). However, the defend proportions were significantly different from their Nash proportions in the n*v* state across both effective and less-effective conditions (effective: $t(49) = 3.76$, $p < .05$, $r = .28$; less-effective: $t(49) = 5.53$, $p < .05$, $r = .61$). Overall, these results partially agree with our expectations in the *nv* state.

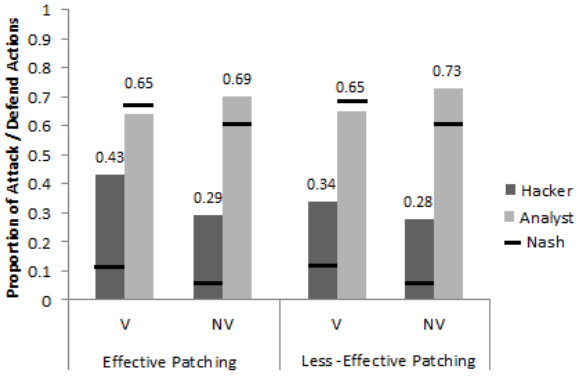

Figure 5. Proportion of attack/defend actions across the patching conditions and network states.

6 IBLT MODEL

We developed a model based upon IBLT that was built to explain human behavior the Markov security game. An instance, i.e., smallest unit of experience, in the IBL model consists of three parts: a situation in a task (a set of attributes that define the decision situation), a decision in a task, and an outcome resulting from making that decision in that situation (Gonzalez & Dutt, 2011; 2012). Different parts of an instance are built through a general decision process: creating a situation from attributes in the task, a decision and expectation of an outcome when making a judgment and updating the outcome in the feedback stage when the actual outcome is known. In the IBL model, instances collect over time in memory, are retrieved from memory, and are used repeatedly to make decisions. This availability is measured by a statistical mechanism called activation, originally implemented in the ACT-R cognitive architecture (Anderson & Lebiere, 1998). We develop our model for two-player security games by simply allowing two single-person models to interact with each other in the game, repeatedly.

In the IBL model, each instance consists of a label that identifies a decision option presented to each player (i.e., to Attack or Not Attack for the hacker and Defend or Not Defend for the analyst) and the outcome obtained (e.g., 10 points). As the situation remains the same for each binary decision, the structure of an instance is simply (alternative, outcome) (e.g., Defend, 10) for both players. In each round t of the game, the process of selection of decision options in the model starts with calculation of the blended value of the options. Next, the decision option with the highest blended value is selected. The blended value of an option depends on outcomes occurring in the option and the probability of retrieval of instances from memory corresponding to

those outcomes (Equation 5 below). Furthermore, the probability of retrieval of instances from memory is a function of their activation in memory, governed by the recency and frequency of instance retrievals from memory (Equations 7 and 8 below).

In the IBL model, the selected option is one with the highest blended value V (Gonzalez & Dutt, 2011). The blended value of option j is defined as:

$$V_j = \sum_{i=1}^{n} p_i x_i \qquad [5]$$

where x_i is the value of the observed relative outcome in the outcome slot of an instance i corresponding to the option j and p_i is the probability of that instance's retrieval from memory (for the security game, the value of j is either to Defend (patch) or Not-defend (not-patch) for analysts and Attack or Not-Attack for the hacker. Similarly, sx_i are the relative outcomes for the decision-maker, hacker or analyst, depending upon the decision choices in Fig. 1). Thus, x_i is defined as per the following equation:

$$x_i = O_{pH} + w * O_{pA} \qquad [6]$$

Where, O_{pH} and O_{pA} are outcomes from the payoff matrix of the hacker and analyst players (see Figure 1(b)) and w is the cognitive attention parameter that measures the attention to opponent's actions. The blended value of an option is the sum of all observed outcomes x_i in the corresponding instances in memory, weighted by their probability of retrieval. In any trial t, the probability of retrieval of instance i from memory is a function of that instance's activation relative to the activation of all other instances corresponding to that option, given by:

$$P_{i,t} = \frac{e^{A_{i,t}/\tau}}{\sum_j e^{A_{j,t}/\tau}} \qquad [7]$$

where τ is random noise defined as $s * \sqrt{2}$, and s is a free noise parameter (see below for its description). Noise in equation 7 captures the imprecision of recalling instances from memory.

The activation of each instance in memory depends upon the activation mechanism originally proposed in the ACT-R architecture (Anderson & Lebiere, 1998). A simplified version of the activation mechanism that relied on recency and frequency of use of instances in memory was sufficient to capture human choice behavior in several binary-choice tasks (Lejarraga, Dutt, & Gonzalez, 2012) and has been used in the IBL model reported in this paper. For each trial t, activation $A_{i,t}$ of instance i is:

$$A_{i,t} = \ln\left(\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right) + s \cdot \ln\left(\frac{1-\gamma_{i,t}}{\gamma_{i,t}}\right) \quad [8]$$

where, d is a free decay parameter, t is the current round, and t_i is the previous round where the instance i was created or its activation was reinforced due to an outcome in the task. The summation will include a number of terms that coincides with the number of times that an outcome has been observed in previous rounds and that the corresponding instance i's activation has been reinforced in memory. Therefore, the activation of an instance corresponding to an observed outcome increases with the frequency of observation (i.e., by increasing the number of terms in the summation) and with the recency of those observations (i.e., by small differences in $t_i \in \{1,\ldots, t\text{-}1\}$ of outcomes that correspond to that instance in memory). The decay parameter d has a default value of 0.5 in ACT-R and it affects the activation of the instance directly, as it captures the rate of forgetting. The higher the value of the d parameter, the more is the reliance on recency, and the faster is the decay of memory.

The $\gamma_{i,t}$ term is a random draw from a uniform distribution bounded between 0 and 1, and the $s \cdot \ln\left(\frac{1-\gamma_{i,t}}{\gamma_{i,t}}\right)$ term represents Gaussian noise important for capturing the variability of human behavior. The higher the s value, the more variability there will be in the retrieval of information from memory.

7 IMPLEMENTATION AND EXECUTION OF THE IBL MODEL IN THE SECURITY GAME

We implemented the model in two settings: (1) calibrated model, where the values of the free parameters were obtained by calibration using a genetic algorithm; (2) ACT-R model, where the ACT-R default values for the free parameters were used in the model. Two identical IBL model agents performed as a pair of participants for 50 rounds in the Markov security game in different conditions, just as human participants performed in the two conditions. Agents in both the model settings used blending and activation mechanisms independently with a separate set of parameters, where decisions made by both agents in a trial determined each version's outcomes across the two settings.

Each agent had three free parameters: w attention to opponent's actions, noise s, and decay d. We calibrated the model's parameters using human data in both less-effective and effective conditions. In these calibrations, we minimized the sum of mean square distances (MSDs) on attack and defend actions between model and human data. A genetic algorithm (GA) program

was used to optimize values of the w, d and s parameters for both the model participants. The w, d, and s parameters were varied between 0 and 1, between 0.0 and 10.0, and between 0.0 and 10.0, respectively, in the GA program. These ranges ensured that the optimization was able to capture the optimal parameter values with high confidence. The GA had a crossover rate of 80% and a mutation rate of 1%. To check the performance of the calibrated model, we compared its performance with the ACT-R model. In the ACT-R model, the parameters were kept at their default values: w = 0.0, d = 0.5, and s = 0.25.

8 MODEL RESULTS

Proportion of attack and defend actions across conditions

We first calculated the proportion of attack-and-defend actions in each patching condition for the calibrated and ACT-R models and compared these proportions with the corresponding human proportions (see Figure 8). The MSDs from the calibrated model and ACT-R models were 0.0039 and 0.0089, respectively, across both hacker and analyst players. As shown in Figure 8, overall, the calibrated model fitted the human data better compared to the ACT-R model for both the hacker and analyst participants

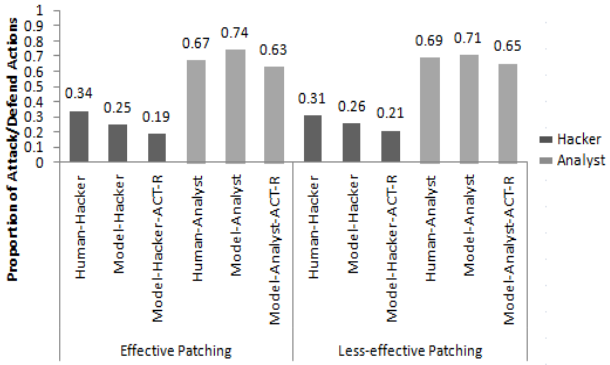

Figure 6. Proportion of attack and defend (Model and Human) actions across the two patching conditions. The Model-Hacker and Model-Analyst refer to the calibrated model for both players.

Proportion of attack and defend actions across states

Next, we calculated the proportion of attack-and-defend actions across different *v* and *nv* states from the calibrated and ACT-R models and compared these proportions with the corresponding human action proportions (see

Figure 9). The MSDs from the calibrated model and ACT-R models were 0.0092 and 0.0149, respectively, across both hacker and analyst players. Overall, the calibrated model fitted the human data better compared to the ACT-R model for both the hacker and analyst in both *v* and *nv* states, respectively.

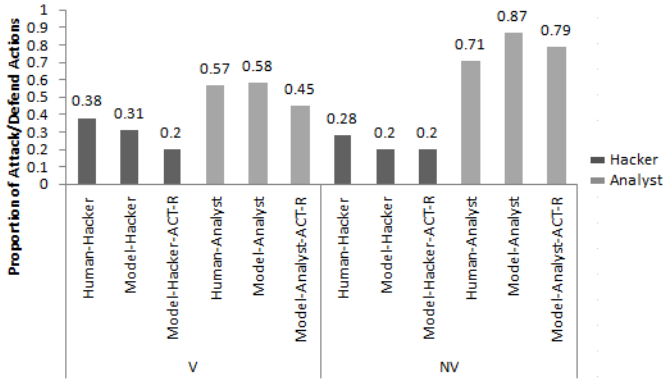

Figure 7. Proportion of attack/defend actions across the states. The Model-Hacker and Model-Analyst refer to the calibrated model for both players.

Proportion of attack and defend actions across patching conditions and states

Furthermore, we also analyzed the proportion of attack-and-defend actions from the calibrated and ACT-R models across the two patching conditions and the two network states and compared these proportions with the corresponding human proportions. Figure 10 shows the proportion of attack-and-defend actions across the patching conditions and states for both models and human participants. The MSDs from the calibrated model and ACT-R models were 0.0535 and 0.0741, respectively, across both hacker and analyst players for the effective condition. Furthermore, the MSDs from the calibrated model and ACT-R models were 0.0051 and 0.0058, respectively, across both hacker and analyst players for the less-effective condition. Overall, the calibrated model fitted human data better compared to the ACT-R model across both patching conditions and states.

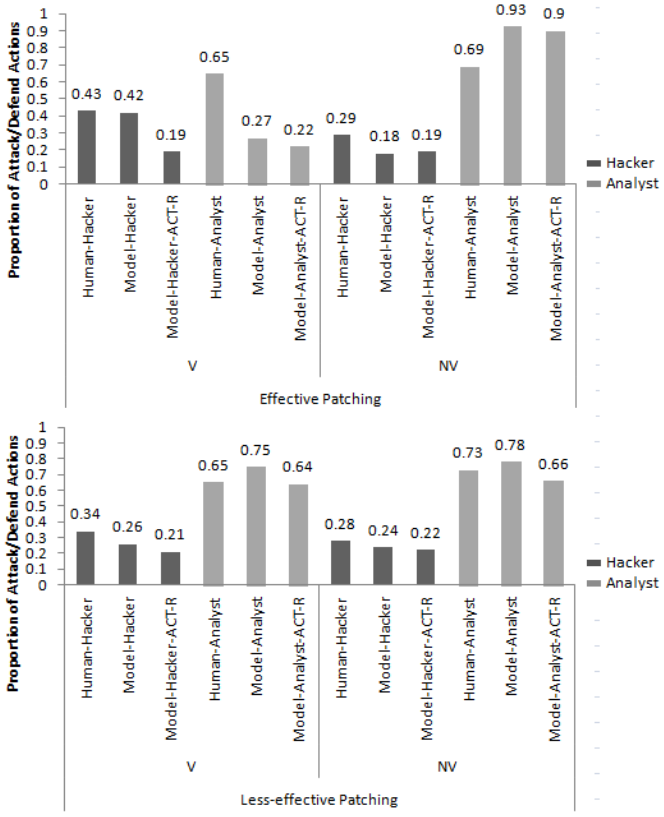

Figure 8. Proportion of attack/defend actions across the patching conditions and network states. The Model-Hacker and Model-Analyst refer to the calibrated model for both players.

Calibrated parameters of the model

Table 1 shows the values of the calibrated parameters and overall MSDs from the calibrated model in the two patching conditions. As shown in Table 1, the MSDs obtained for both model players across the two conditions were very low. As per our expectation from IBLT, the parameters showed contrasting values across the effective and less-effective conditions. In fact, the model revealed low (high) reliance on recency and frequency, attention to opponent's actions, and cognitive noise for hacker (analyst) in effective condition. Whereas, it revealed opposite results for less-effective condition: high (low) reliance on recency and frequency, attention to opponent's actions, and cognitive noise for hacker (analyst).

Table 1. IBL model with calibrated parameters and MSD values across the two patching conditions.

Condition	d_H	d_A	s_H	s_A	w_H	w_A	MSD_H	MSD_A
Effective Patching	1.25	7.68	0.55	0.97	0.41	0.94	0.022	0.024
Less-effective Patching	6.44	0.07	1.3	0.05	0.97	0.87	0.017	0.014

Note. The subscript 'A' is for analyst and subscript 'H' is for hacker.

9 DISCUSSION AND CONCLUSIONS

Due to the rapid increase in cyber-attacks, there is an urgent need to patch vulnerabilities present in computer systems (Humayed et al., 2017). However, the vulnerability patching process may not be foolproof (Grimes, 2016). In some cases, the patching may be effective, and it may make the computer systems less-vulnerable to cyber-attacks; however, in other cases, patching may be less-effective and it may leave computer systems vulnerable to cyber-attacks. In this research, using data from a lab-based experiment conducted by Maqbool et al. (2018), we investigated the influence of effectiveness of the patching processes on cyber decision-making. Our results revealed that the proportion of attack and defend actions were similar when patching processes were effective and less-effective. Furthermore, a majority of time, both players deviated significantly from their optimal Nash proportions in different conditions and states. We explain these results based upon expectations from models built using Instance-based Learning Theory (IBLT; Gonzalez & Dutt, 2011; 2012; Gonzalez, Lerch, & Lebiere, 2003; Lejarraga, Dutt, & Gonzalez, 2012).

First, we found that the proportion of attack and defend actions were similar across the two patching conditions. A likely reason for this finding could be the similarity in payoff magnitudes and valances across the two patching conditions. As mentioned above, according to IBLT, people maximize their perceived payoff across actions (Lejarraga, Dutt, & Gonzalez, 2012). As participants performing as hackers and analysts faced similar payoffs across different patching conditions, they likely possessed similar perceived payoffs in both conditions.

Second, we found that the proportion of attack and defend actions deviated significantly from their Nash proportions. Again, this expectation can be

explained based upon IBLT. According to IBLT, human participants possess cognitive limitations on memory and recall processes and human beings tend to rely upon recency and frequency of outcomes to make their repeated decisions (Lejarraga, Dutt, & Gonzalez, 2012). It seemed that the reliance upon recency and frequency processes in our experiment did not allow participants to form optimal Nash expectations for their actions causing them to deviate significantly from the Nash proportions in several conditions and states.

Our results also revealed that analyst players did not deviate from their Nash proportions in the vulnerable state although we expected them to deviate from these Nash proportions. One likely reason for this result is that the Nash proportions were simply higher in the vulnerable state compared to those in the non-vulnerable state. As analysts continued to exhibit high patching proportions across both states, their action proportions seem to agree with the Nash proportions in the vulnerable state.

Also, we found that the calibrated IBL model better accounted for human decisions across both patching conditions compared to the ACT-R model. One likely reason for it is that the default ACT-R parameters seem to not allow the model to capture human data. However, a recalibration of these parameters drastically helped the model to improve its own results.

We observed a lower value of d parameter for the hacker in the effective patching condition compared to the less-effective patching condition. This d parameter's value indicates that when the patching mechanism is effective and the computer system is in nv state most of the time, then hackers tend to ignore recent experiences (due to a small d value). Whereas, when the patching mechanism is less-effective and computer system state transitions are equally likely between the v and nv state, hackers tend to rely heavily on recent experiences to attack the system. In contrast, we observed a significantly higher d value for the analyst in effective patching condition compared to the less-effective patching condition. This result indicates that when the patching mechanism is effective and the computer system is in nv state most of the time, the analysts tend to rely on recent experiences. Whereas, when the patching process is less-effective and the computer system state transitions are equally likely between the v and nv state, analysts tend to ignore recent experiences and continue to patch the system.

Similarly, we observed a higher value of the w parameter for the hacker in less-effective patching condition compared to the effective patching condition. This result means that participants in the hacker's role tended to

focus more on the opponent's (analyst's) last actions for an unreliable patching process. However, we also observed a higher value of the w parameter for participants in the analyst's role in effective patching condition compared to the less-effective patching condition. Thus, analysts tended to rely upon opponent's (hacker's) last actions to take advantage of the effective patching process.

Furthermore, we observed a higher value of the s parameter for the hacker in less-effective patching condition compared to the effective patching condition. Thus, an unreliable patching process caused hackers to show greater variability in their actions. However, we also observed a higher value of the s parameter for the analyst in effective patching condition compared to the less-effective patching condition. Thus, a reliable patching process also caused analysts to show exploration in their patching actions perhaps because they tend to trust the patching process.

In this research, we performed a lab-based experiment involving simple Markov security games. Although there are differences between lab-based environments and real-world environments, our results may have important implications for the real world. First, based upon our results, we expect that analysts would continue to excessively patch computer systems in the real-world irrespective of the optimality and the effectiveness of these patching decisions. Second, it seems that hackers, while attacking networks, do not seem to worry about whether computer systems are patched effectively or not. However, hackers do worry about the vulnerability of computer systems to their attacks. Thus, this perception of vulnerability is likely to influence hacker's cyber-attack decisions. In the real-world, it may be important to showcase computer networks as less vulnerable to cyber-attacks. One could do so via a number of methods including social networks, newspapers, reports, and multimedia. Furthermore, models based upon IBLT could be used to account for cyber decisions. For example, the hacker models could be used to simulate hacker decision against patches and vulnerabilities. Similarly, analyst models can be used against different kinds of cyber-attacks by automating the patching processes. This research may also help in assessing the cognitive factors associated with the patching processes.

There are likely to be a number of limitations of this research. First, data was collected using participants who possessed STEM backgrounds and computer science degrees. However, such participants may still be different from hackers in the real world. Second, this paper assumed a simple scenario where a hacker repeatedly chose to either attack or refrain from attacking a single system based on their perception of its potential vulnerability. In the

meantime, the analyst chose to patch the system or otherwise. Thus, the design of the experiment in this paper was simplistic, and it may not completely fit certain automated modus operandi of hackers. For example, in opportunistic/light touch attacks, a hacker may attempt to exploit a vulnerability against a population of systems. These attacks may be large scale and they may not involve human judgments exercised on a system-by-system basis. Similarly, in targeted attacks, there may be many vulnerabilities tested against a single system. For example, in targeted attacks, for a given system of interest, the hacker may sequentially test an extensive library of known exploits. Again, targeted attacks may be triggered using automated processes rather than one with human judgement at each step, and such attacks may not fit the modus operandi of hackers assumed in the design. Here, the proposed analyst's model may help understand their decision-making and cognitive processes against such automated attacks. Overall, this paper's experimental design may not completely capture the above mentioned automated situations and such situations may be tested as part of future research.

10 FUTURE RESEARCH DIRECTIONS

There are a number of research directions that one could undertake as part of future research. Our results revealed that the perception of vulnerability influenced hacker's decisions and there could be several ways in which this perception could be shaped. For example, one could involve deception in computer networks via honeypots, where these honeypots are easily attackable systems. Second, one could involve intrusion-detection systems (IDSs) and provide the knowledge of their existence and accuracy to hackers. For example, if hackers are told that IDSs are not present or they are told that IDSs are present but these are less accurate, then this information is likely to influence the hacker's perception of network's vulnerability to her attacks. Again, in this case, the IDSs may be effective in making hackers attack certain systems (e.g., honeypots) over others and causing them to get caught while waging such attacks. In the real-world, hackers and analysts may not possess information about opponent's actions. Thus, it would be interesting to investigate how the availability and unavailability of information about opponent's actions among hackers and analysts impacts their respective decisions. In addition to these ideas, we would also like to investigate the role of other cognitive mechanisms like similarity, spreading activation, and cognitive inertia in our models.

Furthermore, this paper assumed the analyst to have a free hand in attempting to optimize the application of patches, which may not be the case in practice.

For example, in the real-world, most patching may be governed by an operational policy (e.g., Common Vulnerability Scoring System) where analysts may apply patches addressing vulnerabilities of a given severity in a given time scale. Thus, critical patches may be applied within days, whereas non-critical patches may be applied in an extended time frame. Thus, future research may experiment with scenarios where analysts may decide between critical and less-critical patches and the effectiveness of the patching may work across both kinds of patches.

Some of these ideas form the immediate next steps for us to undertake as part of our ongoing research program in game theory and cyber-security.

11 ACKNOWLEDGEMENTS

This research was supported by the Department of Science and Technology, Government of India award ("A Game Theoretic Approach involving Experimentation and Computational Modeling using Deception in Cybersecurity," Award number: IITM/DST-ICPS/VD/251) to Varun Dutt. Also, we are grateful to the Indian Institute of Technology Mandi, for providing the necessary computational resources for this project.

12 REFERENCES

Aggarwal P., Moisan F., Gonzalez C., & Dutt, V. (2018). Understanding Cyber Situational Awareness in a Cyber Security Game involving Recommendations. *International Journal on Cyber Situational Awareness*, *3*(1), 11-38.

Alpcan, T., & Başar, T. (2006, July). An intrusion detection game with limited observations. In *12th Int. Symp. on Dynamic Games and Applications, Sophia Antipolis, France* (Vol. 26).

Alpcan, T., & Başar, T. (2010). Network security: A decision and game-theoretic approach. Cambridge University Press.

Anderson, J. R., & Lebiere, C. J. (1998). Hybrid modeling of cognition: Review of the atomic components of thought.

Arora, A., & Dutt, V. (2013). Cyber Security: Evaluating the Effects of Attack Strategy and Base Rate through Instance Based Learning. In *12th International Conference on Cognitive Modeling*. Ottawa, Canada.

Dunagan, J., Roussev, R., Daniels, B., Johnson, A., Verbowski, C., & Wang, Y. M. (2004, May). Towards a self-managing software patching process using black-box persistent-state manifests. In *International Conference on Autonomic Computing, 2004. Proceedings.* (pp. 106-113). IEEE.

Dutt, V., Ahn, Y. S., & Gonzalez, C. (2013). Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Human Factors*, *55*(3), 605-618.

Florida Tech. (2019). Cybersecurity Analyst Career Guide. Retrieved from https://www.floridatechonline.com/blog/information-technology/cybersecurity-analyst-career-guide/

Gonzalez, C., & Dutt, V. (2010). Instance-based learning

models of training. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 54, No. 27, pp. 2319-2323). Sage CA: Los Angeles, CA: SAGE Publications.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological review*, *118*(4), 523.

Gonzalez, C., & Dutt, V. (2012). Refuting data aggregation arguments and how the IBL model stands criticism: A reply to Hills and Hertwig (2012).

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

Grimes, R. A. (2016, January 26). Why patching is still a problem -- and how to fix it. Retrieved from CSO - India: https://www.csoonline.com/article/3025807/why-patching-is-still-a-problem-and-how-to-fix-it.html

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems security—A survey. *IEEE Internet of Things Journal*, *4*(6), 1802-1831.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, *25*(2), 143-153.

Kaur, A., & Dutt, V. (2013). Cyber situation awareness: modeling the effects of similarity and scenarios on cyber attack detection. In *12th International Conference on Cognitive Modeling.* Ottawa, Canada (Vol. 250).

Keller, S., & Schaninger, B. (2019, July). A better way to lead large-scale change. Retrieved from https://www.mckinsey.com/business-functions/organization/our-insights/a-better-way-to-lead-large-scale-change

Kissel, R. L. (2013, June 5). Glossary of Key Information Security Terms. Retrieved from https://www.nist.gov/publications/glossary-key-information-security-terms-1

Maqbool, Z., Pammi, V. C., & Dutt, V. (2018, June). Cyber security: Influence of patching vulnerabilities on the decision-making of hackers and analysts. In *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)* (pp. 1-8). IEEE.

Roy, S., Ellis, C., Shiva, S., Dasgupta, D., Shandilya, V., & Wu, Q. (2010, January). A survey of game theory as applied to network security. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1-10). IEEE.

TechTarget. (2017). Information security threats. Retrieved from http://searchsecurity.techtarget.com/definition/hacker

Economic Times (2017, October 30). Internet security 101: Six ways hackers can attack you and how to stay safe. New Delhi: The Economic Times.

Xiaolin, C., Xiaobin, T., Yong, Z., & Hongsheng, X. (2008, December). A markov game theory-based risk assessment model for network information system. In *2008 International Conference on Computer Science and Software Engineering* (Vol. 3, pp. 1057-1061). IEEE.

KEY TERMS

Security Analyst: A person who is in-charge of enforcing cybersecurity in a company or organization.
Hacker: A person who tries to attack computer systems by waging different kinds of cyber-attacks.
Patch: A software solution that fixes vulnerabilities present on computer systems.
Markov security game: A game between a hacker and an analyst, where the analyst's last action determines to vulnerability of the network.
Cognitive model: A computational algorithm that accounts for decisions of human participants using cognitive assumptions and parameters.

BIOGRAPHICAL NOTES

Zahid Maqbool is currently working as an assistant professor at the Government Degree College Dooru, India. He is currently pursuing his Ph.D. degree from the Indian Institute of Technology Mandi, India. His areas of expertise are in game theory, cybersecurity, and cognitive modeling.

V. S. Chandrasekhar Pammi worked as a professor at the Centre of Behavioral and Cognitive Sciences, University of Allahabad, India. His areas of expertise included in decision-making, game theory, cybersecurity, and cognitive science.

Varun Dutt works as an associate professor at the Indian Institute of Technology Mandi, India. His areas of expertise included in decision-making, human-computer interaction, AI, and cognitive science.

REFERENCE

Reference to this paper should be made as follows: Maqbool Z., Chandrasekhar Pammi V. S., and Dutt V. (2019). Behavioral Cybersecurity: Investigating the influence of Patching Vulnerabilities in Markov Security Games via Cognitive Modeling. *International Journal on Cyber Situational Awareness*, Vol. 4, No. 1, pp185-209.