MORSHED: Guiding Behavioral Decision-Makers towards Better Security Investment in Interdependent Systems

Mustafa Abdallah
abdalla0@purdue.edu
Purdue University
West Lafayette, IN, USA

Daniel Woods
woods104@purdue.edu
Purdue University
West Lafayette, IN, USA

Parinaz Naghizadeh
naghizadeh.1@osu.edu
Ohio State University
Columbus, OH, USA

Issa Khalil
ikhalil@hbku.edu.qa
Qatar Computing Research Institute
Doha, Qatar

Timothy Cason
cason@purdue.edu
Purdue University
West Lafayette, IN, USA

Shreyas Sundaram
sundara2@purdue.edu
Purdue University
West Lafayette, IN, USA

Saurabh Bagchi
sbagchi@purdue.edu
Purdue University
West Lafayette, IN, USA

ABSTRACT

We model the *behavioral* biases of human decision-making in securing interdependent systems and show that such behavioral decision-making leads to a suboptimal pattern of resource allocation compared to non-behavioral (rational) decision-making. We provide empirical evidence for the existence of such behavioral bias model through a controlled subject study with 145 participants. We then propose three learning techniques for enhancing decision-making in multi-round setups. We illustrate the benefits of our decision-making model through multiple interdependent real-world systems and quantify the level of gain compared to the case in which the defenders are behavioral. We also show the benefit of our learning techniques against different attack models. We identify the effects of different system parameters (e.g., the defenders' security budget availability and distribution, the degree of interdependency among defenders, and collaborative defense strategies) on the degree of suboptimality of security outcomes due to behavioral decision-making.

CCS CONCEPTS

• **Security and privacy → Network security**; **Economics of security and privacy**.

KEYWORDS

Behavioral decision-making, Guiding security decision-makers, Security games, Learning attacks, Reinforcement Learning.

ASIA CCS '21, June 7–11, 2021, Hong Kong, Hong Kong

ACM ISBN 978-1-4503-8287-8/21/06…$15.00
https://doi.org/10.1145/3433210.3437534

ACM Reference Format:
Mustafa Abdallah, Daniel Woods, Parinaz Naghizadeh, Issa Khalil, Timothy Cason, Shreyas Sundaram, and Saurabh Bagchi. 2021. MORSHED: Guiding Behavioral Decision-Makers towards Better Security Investment in Interdependent Systems. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security (ASIA CCS '21), June 7–11, 2021, Hong Kong, Hong Kong*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3433210.3437534

1 INTRODUCTION

Most of the current IT-based systems are becoming more complex, however they are facing sophisticated attacks from external adversaries where the attacker's goal is to breach specific (critical) assets within the system. For each critical asset, the attacker typically utilizes different vulnerabilities to compromise such asset. In this context, the system operators, Chief Information Security Officer or security executives have to judiciously allocate their (often limited) security budgets to reduce security risks of the systems they manage. This resource allocation problem is further complicated by the fact that a large-scale system consists of multiple interdependent subsystems managed by different operators, with each operator in charge of securing her own subsystem.

Prior work has considered such security decision-making problems in both decision-theoretic and game-theoretic settings [22, 37] in which the security risk faced by an operator (defender) depends on her security investments. However, most of the existing work relied on *classical models* of decision-making, where all defenders and attackers are assumed to make fully rational risk evaluations and security decisions [17, 22, 27].

On the contrary, behavioral economics has shown that humans consistently deviate from these classical models of decision-making. Most notably, research in *behavioral economics*, has shown that humans perceive gains, losses and probabilities in a skewed, nonlinear manner [20]. In particular, humans typically overweight low probabilities and underweight high probabilities, where this weighting function has an inverse S-shape, as shown in Figure 3. Many empirical studies (e.g., [12, 20]) have provided evidence

for this class of behavioral models. These effects are relevant for evaluating security of such systems in which decisions on implementing security controls are not made purely by automated algorithms, but rather through human decision-making, albeit with help from threat assessment tools [19, 33].

There are many articles discussing the prevalence of human factors in security decision-making, both in popular press and in academic journals [9], none of which however shed light on the impact of cognitive biases on the overall system security and how we can mitigate such biases. Our work bridges this gap by showing how behavioral research can lead to better security decision-making for interdependent systems. Specifically, we study the effect of the aforementioned human behavioral decision-making bias on security allocations and propose multiple techniques to overcome such bias in both single-round and multi-round setups.

There are recent works [18, 32] that have started to leverage mathematical analysis to model and predict the effect of behavioral decision-making on the players' investments. However, these works have the following limitations. First, they have considered the impact of probability weighting in certain specific classes of interdependent security games. Second, these works did not consider multiple-round setups in which defenders can learn. In contrast to those, we consider general defense allocation techniques that can be applied to any system where its failure scenarios are modeled by an attack graph, and we propose multi-round learning algorithms to guide behavioral decision-makers in different setups and consider different types of attackers. The difference between Morshed[1] and previous related work is shown in Table 1.

Our contributions:

In this paper, we first study the effects of human behavioral decision-making on the security of interdependent systems with multiple defenders where each defender is responsible for defending a set of assets (i.e., a subnetwork of the whole system network). In interdependent systems, stepping-stone attacks are often used by external attackers to exploit vulnerabilities within the network in order to reach and compromise critical targets. These stepping-stone attacks can be captured via *attack graphs*, representing all possible paths an attacker may take to reach targets within the system [16, 27].

We design a reasoning and security investment decision-making technique that we call Morshed pronounced as *M-or-Sh-ed*. We first describe the model consisting of multiple behavioral defenders and an attacker, in which the interdependencies between the defenders' assets are captured via an attack graph by proposing a *behavioral security game* model. We show that behavioral decision-making leads to suboptimal resource allocation compared to non-behavioral decision-making. We then propose different learning-based techniques for guiding behavioral decision-makers towards optimal investment decisions for two different scenarios where each scenario represents whether the defender has knowledge of the adversary's history (i.e., chosen attack paths in previous rounds) or not. Our proposed techniques enhance the implemented security policy (in terms of reducing the total system loss when compromised by allocating limited security resources optimally). Morshed has components for both single-round and multi-round setups as shown in Figure 1. We consider two classes of defenders.

[1]Morshed is an Arabic word with the meaning of guiding people to the right place.

Table 1: Comparison between the prior related work and* Morshed *in terms of the available features.

System	Multiple Defenders	Interdependent subnetworks	Analytical Framework	Behavioral Biases	Various Attack Types	Multiple Rounds
RAID08 [27], MILCOM06 [23]	✗	✓	✗	✗	✗	✗
S&P02 [33], CCS12 [37]	✗	✗	✓	✗	✗	✗
S&P09 [4], EC18 [30], ACSAC12 [6]	✗	✗	✗	✓	✗	✗
ICC17 [32]	✗	✓	✓	✓	✗	✗
TCNS20 [2], TCNS18 [18]	✓	✓	✓	✓	✗	✗
Morshed	✓	✓	✓	✓	✓	✓

Behavioral defenders: These defenders make security investment decisions under two types of cognitive biases. First, following prospect-theoretic, non-linear probability weighting models, they misperceive the probabilities of a successful attack on each edge of the attack graph. Second, they have a bias toward spreading their budget so that a minimum, non-zero investment is allocated to each edge of the attack graph. This second kind of bias is motivated by behavior that we observe in our human subject experiments (see Section 3).

Non-behavioral or rational defenders: These defenders make security investment decisions based on the classical models of fully rational decision-making. Specifically, they correctly perceive the risk on each edge within the attack graph of the system network.

On the other hand, almost all research that have considered behavioral economics in security and privacy has the common theme of considering individual choices regarding privacy and how people treat their own personal data [4] or entirely based on psychological studies [6]. To the best of our knowledge, none of these research considered the defense choices made by people in organizational contexts with interdependent system under control. On the contrary, our work considers scenarios that can be applied to critical infrastructure systems (e.g., cyber-physical systems).

We perform a human subject study with N = 145 participants where they choose defense allocations in two simple attack graphs. We then evaluate Morshed using five synthesized attack graphs that represent realistic interdependent systems and attack paths through them. These systems are DER.1 [19], (modelled by NESCOR), SCADA industrial control system, modeled using NIST guidelines for ICS [17], IEEE 300-bus smart grid [21], E-commerce [27], and VOIP [27]. We do a benchmark comparison with two prior solutions for optimal security controls with attack graphs [23, 33], and quantify the level of the underestimation of loss compared to the Morshed evaluation where defenders are behavioral. In conducting our analysis and obtaining these results based on a behavioral model, we address several domain-specific challenges in the context of security of interdependent systems. These include augmenting the attack graph with certain parameters such as sensitivity of edges to security investments (Equation 2), the estimation of baseline attack probabilities (Table 4) and the types of defense mechanisms (Section 6.5) in our formulations.

In summary, this paper makes the following contributions:

(1) We propose a *security investment guiding* technique for the defenders of interdependent systems where defenders' assets have mutual interdependencies. We show the effect of *behavioral* biases of human decision-making on system security and we quantify the level of gain due to our decision-making technique where defenders are behavioral.

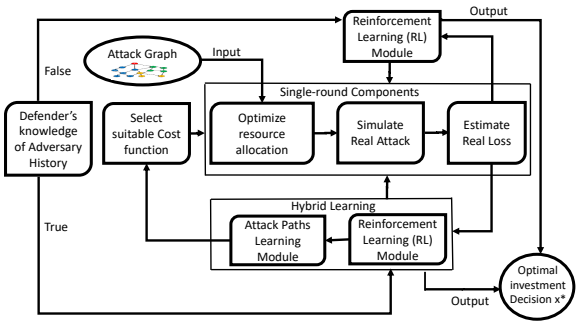


Figure 1: A high level overview of* MORSHED*'s flow, available features and main components (e.g., single-round and Hybrid Learning).

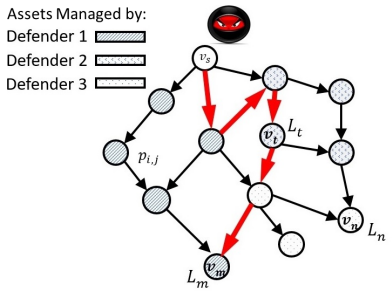


Figure 2: Overview of the interdependent security framework. The interdependencies between assets are represented by edges. An attacker tries to compromise critical assets using stepping stone attacks starting from node v_s. The bold (red) edges show one such attack path.

(2) We validate the existence of bias via a controlled subject study and illustrate the benefits of our decision-making through multiple real-world interdependent systems. We also analyze the different system parameters that affect the security of interdependent systems under our behavioral model.
(3) We propose three learning techniques to improve defense decisions in multi-round scenarios against different attack models that affect the security of interdependent systems. We incorporate such effects with behavioral decision-making.

2 BACKGROUND AND PROBLEM SETUP

We begin by presenting a background on behavioral security games, establishing a theoretical basis that can be used to model any multi-defender interdependent system. A simple example of our setup is shown in Figure 2, which represents a system consisting of 3 interdependent defenders. An external attacker aims to exploit vulnerabilities within the network in order to reach and compromise critical targets [17, 19]. We formalize the attacker and defenders' goals and actions in this section.

2.1 Threat Model

We study security games consisting of one attacker and multiple defenders interacting through an attack graph $G = (V, \mathcal{E})$. The nodes V of the attack graph represent the assets in the system, while the edges $\mathcal{E}$ capture the attack progression between the assets. In particular, an edge from v_i to v_j, $(v_i, v_j) \in \mathcal{E}$, indicates that if asset v_i is compromised by the attacker, it can be used as a stepping stone to launch an attack on asset v_j (e.g., if an attacker gains the password required to access a power plant's control software (v_i), it can use it to attempt to alter the operation of a generator (v_j)). The baseline probability that the attacker can successfully compromise v_j given that it has compromised v_i, is denoted by the edge weight $p^0_{i,j} \in [0, 1]$. By "baseline probability" we mean the probability of successful compromise without any security investment in protecting the assets. The attacker initiates attacks on the network from a source node v_s (or multiple possible source nodes), and aims to reach a target node $v_m \in V_k$, i.e., a critical node for defender D_k.

2.2 Defense Model

Each defender $D_k \in D$ is in control of a subset of assets $V_k \subseteq V$. This is motivated by the fact that a large system comprises a number of smaller subnetworks, each owned by an independent stakeholder. Among all the assets in the network, a subset $V_m \subseteq V$ are *critical* assets, the compromise of which entails a financial loss for the corresponding defender. Specifically, if asset $v_m \in V_m$ is compromised by the attacker, any defender D_k for whom $v_m \in V_k$ suffers a financial loss $L_m \in \mathbb{R}_{>0}$.

To protect the critical assets from being reached through stepping stone attacks, the defenders can choose to invest their resources in strengthening the security of the edges in the network. Specifically, let $x^k_{i,j}$ denote the investment of a defender D_k on edge $(v_i, v_j) \in \mathcal{E}_k$, and let $x_{i,j} = \sum_{D_k \in D} x^k_{i,j}$ be the total investment on that edge by all eligible defenders. Then, the probability of successfully compromising v_j starting from v_i is given by $p_{i,j}(x_{i,j})$. In addition, let $s_{i,j} \in [1, \infty)$ denote the sensitivity of edge (v_i, v_j) to the total investment $x_{i,j}$. For larger sensitivity values, the probability of successful attack on the edge decreases faster with each additional unit of security investment on that edge; in other words, edges that are easier to defend will have larger sensitivity.

Let P_m be the set of all attack paths from v_s to v_m. The defender assumes the worst-case scenario, i.e., the attacker[2] exploits the most vulnerable path to each target.[3] Note that previous works considered such adversary model that chooses the most vulnerable path to target assets (e.g., [17, 22]). Mathematically, this can be captured via the following total loss function for D_k:

$$\hat{C}_k(\mathbf{x}) = \sum_{v_m \in V_k} L_m \left(\max_{P \in P_m} \prod_{(v_i, v_j) \in P} p_{i,j}(x_{i,j}) \right). \tag{1}$$

We let the probability of successfully compromising v_j starting from v_i be given by,

$$p_{i,j}(x_{i,j}) = p^0_{i,j} \exp\left(- s_{i,j} \sum_{D_k \in D \text{ s.t. } (v_i, v_j) \in \mathcal{E}_k} x^k_{i,j} \right). \tag{2}$$

That is, the probability of successful attack on an edge (v_i, v_j) decreases exponentially with the sum of the investments on that edge by all defenders. This probability function falls within a class commonly considered in security economics (e.g., [13, 17]).

[2] Our formulation also captures the case where each defender faces a different attacker who exploits the most vulnerable path from the source to that defender's assets.
[3] We will consider different types for the attacker (with partial knowledge) in Section 5.

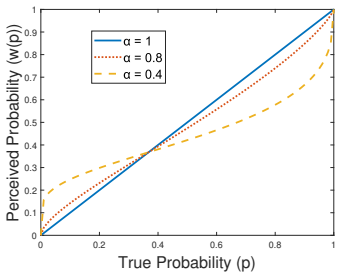


Figure 3: Prelec Probability weighting function which transforms true probabilities p into perceived probabilities $w(p)$. The parameter α controls the extent of overweighting and underweighting, with $\alpha = 1$ indicating non-behavioral or rational decision-making.

2.3 Behavioral Probability Weighting

As mentioned in the Introduction, the behavioral economics literature has shown that humans consistently misperceive probabilities by overweighting low probabilities, and underweighting high probabilities [20, 29]. More specifically, many humans perceive a "true" probability p as probability $w(p)$, where $w(\cdot)$ is known as a *probability weighting function*. A commonly studied functional form for this weighting function was formulated by Prelec in [29], shown in Figure 3, and is given by

$$w(p) = \exp\Big[-(-\log(p))^{\alpha}\Big], \quad p \in [0,1], \tag{3}$$

where $\alpha \in (0,1]$ is a parameter that controls the extent of misperception. When $\alpha = 1$, we have $w(p) = p$ for all $p \in [0,1]$, which corresponds to the situation where probabilities are perceived correctly, i.e., a non-behavioral defender.

2.4 Perceived Costs of a Behavioral Defender

We now incorporate this probability weighting function into the security game of Section 2.2. In a *behavioral security game*, each defender misperceives the attack success probability on each edge according to the probability weighting function in (3). She then chooses her investments $x_k := \{x^k_{i,j}\}_{(v_i,v_j)\in\mathcal{E}_k}$ to minimize her *perceived* loss

$$C_k(x_k, \mathbf{x}_{-k}) = \sum_{v_m \in V_k} L_m\Big(\max_{P\in P_m} \prod_{(v_i,v_j)\in P} w\left(p_{i,j}(x_{i,j})\right)\Big), \tag{4}$$

subject to her total security investment budget B_k, i.e., $\sum_{(v_i,v_j)\in\mathcal{E}_k} x^k_{i,j} \le B_k$, and non-negativity of the investments, i.e., $x^k_{i,j} \ge 0$. It is easy to show that the total loss (4) is convex.

2.5 Spreading Nature of Security Investments

We augment our model with another aspect of behavioral decision-making, which we call *spreading*. A defender with this characteristic spreads some of her investments on all edges of the attack graph, even when some edges are unlikely to be exploited for attacks. Spreading here is inspired by *Naïve Diversification* [7] from behavioral economics, where humans have a tendency to split investments evenly over the available options. This phenomenon has not been reported earlier for security decision-making, to the best of our knowledge, and we infer this behavior from our human subject study (detailed in Section 3). We capture this effect by adding another constraint to our model in (4): for each defender D_k, we set $x^k_{i,j} \ge \eta_k$, where η_k is the minimum investment D_k makes on any edge. The value $\eta_k = 0$ gives us the behavioral decision with no spreading, i.e., with only behavioral probability weighting.

3 HUMAN SUBJECT STUDY

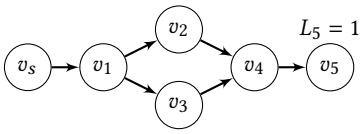
To validate the existence of behavioral bias in security allocations (captured by our model in Section 2), incentivized experiments were conducted on 145 students in an Experimental Economics Laboratory at a large public university. Subject demographics are presented in Appendix B. Subjects participated in the role of a defender, and allocated 24 discrete defense units over edges in each network. Subjects made their decisions on a computerized interface, and faced 10 rounds for each network, receiving feedback after each round indicating whether the attack was successful or not (i.e., the asset was compromised). Subjects received comprehensive written instructions on the decision environment that explained how their investment allocation mapped into the probability of edge defense, and what was considered a successful defense. Subjects received a base payment of $5.00 for their participation. In addition, we randomly selected one round from each network and if the subject successfully defended the critical node in that round, she received an additional payment of $7.50.

3.1 Network (A) with Critical Edge

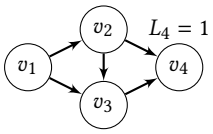
This human experiment is on a network similar to Figure 4a, except that there is only one critical edge (v_4, v_5) i.e., $v_s = v_1$. Figure 5 shows the average investment allocation to the critical edge, based on 1450 investment decisions (i.e., 10 decisions from each of the 145 subjects). It shows the proportion of subjects who are non-behavioral (those at the vertical red line of $\alpha = 1$, 27%), as well as heterogeneity in α, with observations further to the left being more behavioral. Subjects to the left of the $\alpha = 0.4$ line (approximately 10 units allocated to the critical edge) are not necessarily exhibiting $\alpha < 0.4$. Those who allocate between 5 and 10 units to the critical edge could have a strong preference for spreading. We observe that after round 4, the average investment on the critical edge in each round is higher than the initial investment in round 1 (Figure 18 in Appendix E). The average increase summed across the 10 rounds is one defense unit. This means that subjects become less behavioral on average through learning.

3.2 Network (B) with Cross-over Edge

This experiment used the attack graph from Figure 4b. This attack graph is suitable to separate the spreading behavioral bias from the behavioral probability weighting, since for any $0 < \alpha \le 1$, the optimal decision is to put zero defense units on the cross-over edge (v_2, v_3). Figure 6 shows the average investment allocation on the cross-over edge based on 1450 investment decisions. We see that the proportion of subjects that are non-behavioral, i.e., invest nothing on the cross-over edge, is 29%. We observe that the average of subjects' investments on the cross-over edge in each round, shows a weak downward trend (Figure 19 in Appendix E). Taken together, these human experiments provide support for our behavioral model with probability weighting and spreading factors.



(a) An attack graph with a min-cut edge.



(b) An attack graph with a cross-over edge.

Figure 4: The attack graph in (a) is used to illustrate the sub-optimal investment decisions of behavioral defenders. The attack graph in (b) is used in the human subject experiment to isolate the spreading effect.

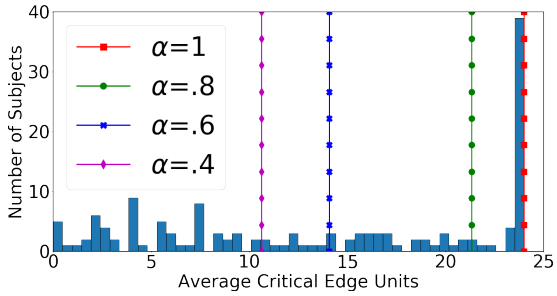


Figure 5: Subjects' investments on the critical edge. Vertical lines with dots show optimal allocations at specific behavioral levels (α).

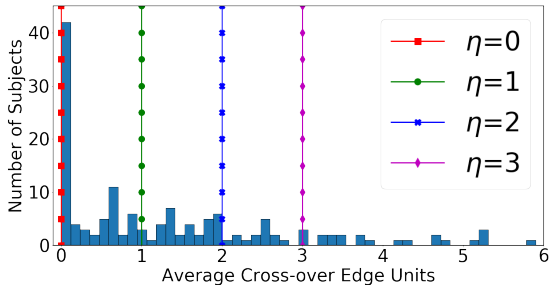


Figure 6: Subjects' investments on the cross-over edge. Vertical lines with dots show optimal allocations at specific spreading levels (η).

Generalizability of the study: The applicability of this subject study to security experts is motivated by the fact that numerous academic studies of even the most highly-trained specialists have shown that experts too have susceptibility to systematic failures of human cognition (e.g., [11, 14]). In the meta-review article [11], 9 of 13 studies that make a direct comparison between student and professional subject pools find no evidence of differing behavior, and only 1 out of 13 studies finds that professionals behave more consistently with theory. Moreover, recent research has shown that cybersecurity professionals' probability perceptions are as susceptible to systematic biases as those of the general population [25, 26]. Finally, even if security experts exhibit weaker biases, this can result in sub-optimal security investments and their effects may be magnified due to the magnitude of losses associated with compromised 'real-world' assets.

4 EFFECT OF BIAS ON INVESTMENTS

In this section, we provide a simple example to illustrate the investment decisions by behavioral and non-behavioral defenders, and provide some intuition on why the optimal defense strategies under the two decision-making models differ. In this example, we use the notion of a *min-cut* of the graph. Specifically, given two assets s and t in the graph, an edge-cut is a set of edges $\mathcal{E}_c \subset \mathcal{E}$ such that removing $\mathcal{E}_c$ from the graph also removes all paths from s to t. A min-cut is an edge-cut of smallest cardinality over all possible edge-cuts. As the example will show, the optimal investments by a non-behavioral defender (i.e., $\alpha = 1$) will generally concentrate the security investments on certain critical (i.e., min-cut) edges in the network. In contrast, behavioral defenders tend to spread their budgets throughout the network.

Consider the attack graph shown in Figure 4a, with a single defender D and a single target asset v_5 (with a loss of $L_5 = 1$ if successfully attacked). Let the defender's budget be B, and let the probability of successful attack on each edge (v_i, v_j) be given by $p_{i,j}(x_{i,j}) = e^{-x_{i,j}}$ (assuming $p^0_{i,j} = 1$). This graph has two possible min-cuts, both of size 1: the edge (v_s, v_1), and the edge (v_4, v_5). The total loss function (1) for the defender is given by

$$C(x) = \max\left(e^{-(x_{s,1}+x_{1,2}+x_{2,4}+x_{4,5})}, e^{-(x_{s,1}+x_{1,3}+x_{3,4}+x_{4,5})}\right),$$

which reflects the two paths from the source v_s to the target v_t. We note that the optimal solution of this constrained convex optimization problem satisfies the KKT conditions [15]. One can then verify (using KKT conditions [15]) that it is optimal for a non-behavioral defender to put all of her budget only on the min-cut edges, i.e., any solution satisfying $x_{s,1} + x_{4,5} = B$ and $x_{1,2} = x_{2,4} = x_{1,3} = x_{3,4} = 0$ is optimal. The intuition of the above result is that from a non-behavioral defender's viewpoint, the probability of successful attack on any given path is a function of the sum of the security investments on the edges in that path. Thus, any set of investments on min-cut edges would be optimal since the sum of investments would be the whole security budget on each path of the graph.

Now, consider a behavioral defender, i.e., a defender with $\alpha < 1$. With the above expression for $p_{i,j}(x_{i,j})$ and using the Prelec function (3), we have $w(p_{i,j}(x_{i,j})) = e^{-x^{\alpha}_{i,j}}$. Thus, the total (perceived) loss function (4) for a behavioral defender is

$$C(x) = \max\left(e^{-x^{\alpha}_{s,1}-x^{\alpha}_{1,2}-x^{\alpha}_{2,4}-x^{\alpha}_{4,5}}, e^{-x^{\alpha}_{s,1}-x^{\alpha}_{1,3}-x^{\alpha}_{3,4}-x^{\alpha}_{4,5}}\right),$$

which includes the two paths from the source v_s to the target v_5. Again, one can verify (using the KKT conditions [15]) that the optimal investments are

$$x_{1,2} = x_{2,4} = x_{1,3} = x_{3,4} = 2^{\frac{1}{\alpha-1}} x_{s,1}.$$
$$x_{s,1} = x_{4,5} = \frac{B-4x_{1,2}}{2} = \frac{B}{2+4(2^{\frac{1}{\alpha-1}})}.$$

Comparing these two cases, the optimal investments of the non-behavioral defender yield a total loss of e^{-B}, whereas the investments of the behavioral defender yield a total loss of $e^{-2^{\frac{\alpha}{\alpha-1}}} e^{-\frac{B}{1+2^{\frac{\alpha}{\alpha-1}}}}$, which is larger than that of the non-behavioral defender.

Interpretation: The reason for this discrepancy can be seen by examining the Prelec probability weighting function in Figure 3. Specifically, when considering an undefended edge (i.e., whose probability of successful attack is 1), the marginal reduction of the attack probability on that edge as *perceived* by a behavioral defender is much larger than the marginal reduction of true attack probability on that edge. Thus the behavioral defender is incentivized to invest some non-zero amount on that edge. Therefore, a behavioral defender splits her investments among the two non-critical sub-paths in the attack path. Note that the same insight holds for different baseline probabilities, but this shifting effect is greater when the slope of the behavioral probability weighting curve is higher (i.e., close to values of 1, 0, or where the cross-over happens between the behavioral curve and the diagonal). A rational defender, on the other hand, correctly perceives the drop in probability, and thus prefers not to invest on the non-critical sub-paths, instead placing her investment only on the critical edges (v_s, v_1) or (v_4, v_5) or both.

In the above example, we assumed all edges have the same sensitivity to investments. We provide the analysis for different edges' sensitivities in Appendix A.

5 LEARNING OVER ROUNDS

Here we consider a defender who plays multiple rounds of the game, learning from observing the attack in each round. In each round, each defender plays the single-shot game with the attacker, allocating all her security budget. She then uses information collected during this interaction to inform her future decisions. In particular, we consider two different forms of learning: (1) what can the defender learn about an attacker over time, and (2) how can repeated interactions lead to decrease in the defenders' extent of behavioral decision-making (i.e., increase in α)? We answer these questions through casting them as repeated resource allocation and reinforcement learning problems, respectively.

5.1 Learning about the Attacker

Now, we assume that the defender can observe the attacker's past actions, e.g., via an intrusion detection system [27] or user metrics [36].

We propose an algorithm through which the defender learns the attack paths over time, and distributes her investments optimally accordingly over the edges. In particular, the steps of this algorithm for this defense technique, as outlined in Algorithm 1, are as follows. First, for each round, we compute the empirical frequency of the attacker's actions over the past N moves (i.e., the probability of choosing every attack path based on the most recent N choices). Then, we compute the best response of the defender to a modified

Algorithm 1: Learning Attack Paths

Input: Set of attack paths $\mathcal{P}_m$, number of rounds N_R and history of attack paths $(P^{t-N}, \cdots, P^{t-1})$
Output: Vector of investments over rounds, O
Round Number = t = 0
while $t < N_R$ **do**
  for $v_m \in V_k$ **do**
    for *Path* $P \in \mathcal{P}_m$ **do**
      $\beta_P^t = \frac{1}{N} \sum_{\tau=t-N}^{t-1} [P^\tau = P]_1$
  $C_k^t(x_k) = \sum_{v_m \in V_k} L_m \Big(\sum_{P \in P_m} \beta_P^t \prod_{(v_i,v_j) \in P} w(p_{i,j}(x_{i,j})) \Big)$
  $x_k^t \in argmin_{x_k \in X_k} C_k^t(x_k)$
  Append (O, x_k^t)
Return O

version of the cost $C_k(x_k)$: this is a weighted version of the cost where each path P has a weight β_P (computed from the previous step). The complexity of the algorithm therefore depends on the number of attack paths.

In Section 6, we compare the investment decisions prescribed by Algorithm 1 with those from our earlier single-shot setup where the defender exhibits no learning. In these comparisons, we consider three types of attackers: replay attackers, randomizing attackers, and adaptive attacker. Specifically, a *replay attacker* chooses the same attack path for every critical asset in every round. Such behavior may be due to limited observations [5], or when the attack process is automated. A *randomizing attacker*, on the other hand, chooses an attack path (for every critical asset v_m) randomly each round, i.e., with probability following a uniform distribution over the possible attack paths in $\mathcal{P}_m$. Such attackers have also been studied in other work using attack graph models [34]. We consider a third attacker type, the *adaptive attacker*, who chooses the least chosen attack path in the past N moves (for every critical asset).

In contrast to replay attacker and randomizing attacker, we assume that the adaptive attacker is aware that the defender's strategy considers the most recent N attacks, and thus the attacker engineers its attack history over a period of time so as to make additional gains on the future attack by choosing the least chosen attack path in the past N moves. Note that the attacker does not have a budget, he just chooses an attack path to each critical asset.

5.2 Reinforcement Learning for Reducing Behavioral Decision-Making

As shown in Sections 3 and 4, the one-round investment decisions made by a behavioral defender D_k based on the decision model in Equation (4) are sub-optimal. It is therefore of interest to understand whether such defender can reduce her behavioral biases in a multi-round defense game by using her experience from previous rounds. In this section, we propose a learning technique through which the defender can make such progress towards a more rational model, i.e., leads to $\alpha^j > \alpha^i$, for some $j > i$, where α^i denotes the behavioral level in round i. Our proposed algorithm, outlined in Algorithm 2, uses a reinforcement learning approach. Our

algorithm is based on that of [10], adapted to our problem of security investment decision-making.

The algorithm proceeds as follows. Let $q^t(\alpha_i)$ denotes the defender's propensity to invest according to the behavioral level α_i at round t. We first initialize these propensities to the defender's initial behavioral level (i.e., $\alpha^0 = \alpha_i$, $q^0(\alpha_i) = A$, and $q^0(\alpha_j) = B, \forall j \neq i$).[4] Then, for every round t, the defender does not know her behavioral level but she draws her defense budget decision in accordance to her reinforcement level. After the defender distributes her defense budget, she receives corresponding reinforcement R^t (which is the difference between the true loss $\hat{C}^t(x_k^t)$ calculated with the investments (budget allocation on edges) in round t, denoted by x_k^t, and the maximum possible true loss $\hat{C}_{max}$ (which is the initial loss). Thus, if the defender invests according to a more rational behavior (i.e., higher α) in round t, she receives higher reinforcement and thus the propensity to choose this investment again in next rounds ($q^{t+1}(\alpha_i)$) increases. For all other investments that are not observed in this round, the propensities of the corresponding behavioral levels do not change. Then, we update the probability distribution for the investments (resp. behavioral levels) for the next round. We repeat the process until we reach convergence (where the reinforcement learning model chooses $\alpha_i = 1$ with a probability sufficiently close to 1) or we reach the maximum number of rounds N_R. The output of our algorithm is a time-series of behavioral level values. We emphasize that the learning comes from the reinforcements received each round which controls the propensity of the defender to choose particular budget distributions in next rounds and that the defender does not know the optimal investments apriori.

Convergence of Algorithm 2 to rational behaviour: The convergence of Algorithm 2 depends on the relation between the total loss (true cost) under rational behavior $\alpha = 1$ and the total loss (true cost) under bias $\alpha < 1$. In the interest of space, we state this result in Lemma C.1 and provide its proof in Appendix C.

5.3 Hybrid-Learning Algorithm

In Algorithm 2 the defender learns through observing her payoffs in the last recent rounds. In Algorithm 1, the defender learns the attacker's chosen paths. Here, we combine these two forms of learning to create a hybrid learning algorithm. This algorithm is a modified version of Algorithm 2 where the cost $C_k^t(x_k, \alpha_i)$ is the cost proposed in Algorithm 1, which changes each round as the defender updates the weights of each path according to the history of attack paths. We will evaluate this hybrid-learning algorithm in Section 6 and will compare it with both of Algorithm 1 and Algorithm 2, described earlier in this section.

6 EVALUATION

Our evaluation of Morshed aims to answer the following questions:

- What is the gain of using Morshed for guiding behavioral decision-makers towards rational decision-making?
- How can we decrease level of behavioral bias over rounds?

[4]In our evaluation, we show the convergence of Algorithm 2 under different possible values of the initial propensities of different behavioral levels (i.e., A and B in Algorithm 2). We also show in Appendix C that the convergence of Algorithm 2 depends on the true total loss of the investment, not the initial propensities.

Algorithm 2: Reinforcement Learning to Reduce Behavioral Biases

Input: Set of behavioral levels α and number of rounds N_R
Output: Vector of behavioral level over rounds O
Round Number = t = 0
$q^0(\alpha_i) = A$ and $q^0(\alpha_j) = B \forall j \neq i$
while $t < N_R$ *or not Convergence to* $\alpha_i = 1$ **do**
  for $\alpha_i \in \alpha$ **do**
    if α_i *was observed in round* t **then**
      $x_k^t \in argmin_{x_k \in X_k} C_k^t(x_k, \alpha_i)$
      $R^t = \hat{C}_{max} - \hat{C}_k^t(x_k^t)$
      $q^{t+1}(\alpha_i) = q^t(\alpha_i) + R^t$
    else
      $q^{t+1}(\alpha_i) = q^t(\alpha_i)$
    $p^{t+1}(\alpha_i) = \frac{q^{t+1}(\alpha_i)}{\sum_{\alpha_i \in \alpha} q^{t+1}(\alpha_i)}$
  Sample random α_i with probability $p^{t+1}(\alpha_i)$ to get α^{t+1}
  Append (O, α^{t+1})
Return O

Table 2: The one-round gain of* Morshed *compared to behavioral investment decisions for the five studied interdependent systems.

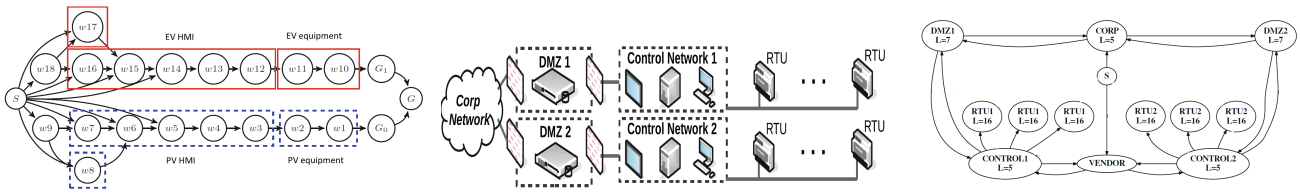
System	# Nodes	# Edges	# Min-cut Edges	# Critical Assets	Avg Gain	Max Gain
SCADA-external	13	20	2	6	1.43	2.63
SCADA-internal [17]	13	26	8	6	4.43	9.42
DER.1 [19]	22	32	2	2	1.29	2.38
E-Commerce [27]	18	26	1	4	3.70	18.28
VOIP [27]	20	28	2	4	4.46	18.66
IEEE 300-bus [21]	300	822	98	69	5.85	11.25

- How does each system parameter affect the overall security level of the system with behavioral decision-making?

6.1 Experimental Setup

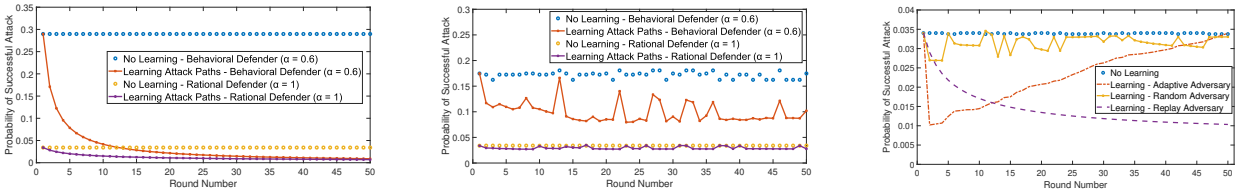
Dataset Description: We use five synthesized attack graphs that represent real-world interdependent systems with different sizes to evaluate our setups, i.e., different attacks, defense, and learning (See Table 2). Specifically, we consider 5 popular interdependent systems from the literature which are: DER.1 [19], SCADA (with internal attacks) [17], SCADA (with only external attacks), IEEE 300-bus smart grid [21], E-commerce [27], and VOIP [27]. In all of these systems, nodes represent attack steps (e.g., taking privilege of control unit software in SCADA, accessing customer confidential data such as credit card information in E-commerce). Now, we give a detailed explanation of one of these systems; the SCADA system (see Appendix D and [19, 21, 27] for detailed description of the rest of the systems). We generate the attack graphs using the CyberSage tool [19] which maps the failure scenarios of the system automatically into an attack graph given the workflow of that system, the security goals, and the attacker model.

SCADA system description: The SCADA system (shown in Figure 7b) is composed of two control subsystems, where each incorporates a number of cyber components, such as control subnetworks and remote terminal units (RTUs), and physical components, such as, valves controlled by the RTUs. This system is architected following the NIST guidelines for industrial control systems. For example, each subsystem is separated from external networks through a demilitarized zone (DMZ). The purpose



(a) Attack Graph of DER System *(b) A high level network overview of the SCADA system* *(c) Attack Graph of SCADA System*

Figure 7: Attack graphs of DER.1 and SCADA case studies. The attack graphs of the remaining systems are given in Appendix D.



(a) Attacker chooses same attack paths *(b) Attacker chooses attack paths randomly* *(c) Different attack types comparison*

Figure 8: The effect of learning attack paths over the rounds. The learning is useful for both behavioral and rational defenders. Moreover, behavioral defender with learning attack paths can eventually reach same security level as rational defender (specifically if the attacker chooses same attack path for each critical asset over rounds). The adaptive attacker is the most challenging attack type.

of a DMZ is to add an additional layer of security between the local area networks of each control subsystem and the external/corporate networks, from where external attackers may attempt to compromise the system. The system implements firewalls both between the DMZ and the external networks, as well as between the DMZ and its control subnetwork. Therefore, an adversary must bypass two different levels of security to gain access to the control subnetworks.

Mapping this system to our proposed security game model, each control subnetwork is owned by a different defender. These two subsystems are interdependent via the shared corporate network, as well as due to having a common vendor for their control equipment. The resulting interdependencies map to the attack graph shown in Figure 7c. The "Corp" and the "Vendor" nodes connect the two subnetworks belonging to the two different defenders and can be used as jump points to spread an attack from one control subsystem to the other. This system has six critical assets (i.e., 3 RTUs, Control Unit, CORP, and DMZ). The compromise of a control network "CONTROL i" will lead to loss of control of all 3 connected RTUs. Now, we present the various system parameters.

Baseline Probability of successful attack: Each edge in the attack graphs represents a real vulnerability. To create the baseline probability of attack on each edge (i.e., without any security investment), we first create a table of CVE-IDs (based on real vulnerabilities reported in the CVE database for 2000-2019). We then followed [16] to convert the attack's metrics (i.e., attack vector (AV), attack complexity (AC)) to a baseline probability of successful attack (e.g., Table 4 in Appendix illustrates such process for SCADA and DER.1). Interestingly, we show that the gain of rational vs. behavioral investments exists for any combination of baseline probabilities (as will be shown in Section 6.5).

Security Budget: We assume that the total budget available at the defenders' organization is B, and that an amount BT of this budget is set aside for security investments. We refer to $BT < 0.3B$, $0.3B < BT < 0.6B$, and $BT > 0.6B$, as low, medium, and high security budgets. For instance, BT = \$10 and \$20 reflect low and moderate budgets, respectively and $BT \geq$ \$30 reflects high budgets in SCADA system given that B = \$50. We emphasize that the gain of our proposed techniques exists for any choice of budget (as will be shown in Section 6.5).

Convergence to Optimal Solution: In our experiments, to find the optimal investments, we use the notion of *best response dynamics*, where the investments of each defender D_k are iteratively updated based on the investments of the other defenders. In each iteration, the optimal investments for defender D_k can be calculated by solving the convex optimization problem in (4).[5] Note that the best response dynamics converge to a Nash equilibrium [17] and we study the security outcomes at that equilibrium.

6.2 Gain from Using Morshed in One Round

Here, we show the gain that behavioral security decision-maker would have using Morshed.

Reduction in Defender's Total Loss: To show the gain of our proposed algorithm, we quantitatively compare the total system loss of the aforementioned five systems in two scenarios which are assuming behavioral decision-maker without the help of Morshed and with the help of Morshed investments, respectively. We then calculate the gain as the ratio of the total system loss by behavioral decision-maker to the total system loss by Morshed to quantify the benefit of using our proposed algorithm.

[5] Note that in the results of learning attack paths and Hybrid learning techniques, we use different cost function (shown in Algorithm 1).

1) Average Gain: We define the Average Gain as the ratio of the weighted sum of total system loss by behavioral decision-maker to the total system loss by Morshed assuming that 50% of the decision-makers are fully rational (with $\alpha = 1$) and 50% are behavioral defenders ($\alpha \in [0.4, 1)$); this is consistent with the range of behavioral parameters from prior experimental studies [12] and our subject study. Average Gain for all systems is shown in Table 2.

2) Maximum Gain: We define the Maximum Gain as the ratio of the total system loss by the highest behavioral defender ($\alpha = 0.4$) to the total system loss by rational ($\alpha = 1$) decision-maker (computed by Morshed). Table 2 shows maximum gains which are 2.38, 9.43, 2.63, 11.25, 18.28, and 18.66 for the DER.1, SCADA-internal, SCADA-external, IEEE 300-bus, E-commerce, and VOIP respectively.

6.3 Learning over Rounds Results

Now, we consider the different setups where the defender learn over rounds using our proposed algorithms in Section 5. For some results, we only show results on the SCADA attack graph as we observe similar patterns on the remaining studied attack graphs.

1) Learning of attack paths: We show the effect of learning attack paths over the rounds for all of the possible attack scenarios described in Section 5. We consider the five systems described earlier and simulate our learning algorithm over 50 rounds with considering medium budget. For each round, the attacker chooses one path for compromising each critical asset (for SCADA, we have six critical assets (i.e., 3 RTUs, Control Unit, CORP, and DMZ) and thus each round the attacker chooses six paths, one for each critical asset) and then the overall probability of successful attack is calculated. We show that the learning of attack paths is useful for both behavioral and rational defenders. Specifically, Figure 8a shows such effect of learning if the attacker chooses same attack paths for each critical asset. Also, Figure 8b shows that our proposed algorithm helps enhancing system security even if the attacker chooses attack paths randomly over rounds since it captures an approximate distribution of the attacker choice of the paths over the rounds. Interestingly, behavioral defender that learns attack paths can eventually reach comparable security level as rational defender (with same security level if the attacker chooses same attack path for each critical asset over rounds; here, after 40 rounds as shown in Figure 8a). Moreover, we compare the learning effect for all attack types, defined in Section 5, in Figure 8c which shows that adaptive attacker is the most challenging attack type.

2) Reinforcement learning of behavioral level: Now, we show the performance of our reinforcement learning algorithm to guide behavioral decision-makers to rational behavior. Here, we consider the attacker who chooses the most vulnerable path to each target asset introduced in Section 2. For each system of the five systems, we run our learning algorithm over 500 rounds with considering medium budget. For each round, the attacker chooses the most vulnerable path for compromising each critical asset. First, Figure 9a shows the convergence of our algorithm over the rounds to rational behavior (i.e., $\alpha = 1$) for all of the five systems where the probability of having rational behavior after learning over 100 rounds is more than 0.9 and approaches 1 by the end of 500 rounds (for four systems from the five systems). Note that here we show the convergence when the initial behavior was $\alpha = 0.2$ (i.e., $\alpha^0 = 0.2$). Such convergence would happen for any behavioral defender (with any $\alpha < 1$) given enough learning. This also shows that behavioral defender with our proposed Reinforcement learning algorithm can eventually reach optimal investment decisions (that leads to comparable security level as rational defender).

Figure 9a shows the rate of convergence of our Reinforcement-learning algorithm for the five case studies. It worth noting that learning is slower for VOIP compared to the other four systems. The reason is the higher criss-cross edges across the VOIP system (see Figure 14 in Appendix D). This also sheds the light that each system has its own characteristics and may need further parameter tuning for enhancing convergence.

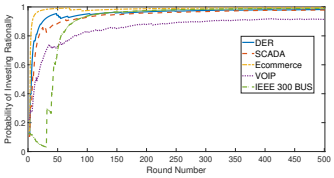
Initial values of propensities: Recall from Algorithm 2 that A and B represent the initial propensities for investing with the initial behavioral level and the propensities of other possible behavioral levels, respectively. To test convergence under different setups, we iterate A over the values $\{0.1, 0.2, 0.6, 1, 1.6, 2\}$ while keeping $B = 0.1$ to simulate different propensities for investing with initial behavior level α_0. In all of the experiments, the algorithm converges to rational behaviour with an average of 400 iterations (i.e., $P^t(1) > 0.95$ at $t \geq 400$).

3) Hybrid-learning Results: Here, we show the performance of our proposed Hybrid-learning Algorithm in Section 5. Figure 9b shows the enhancement of defense (represented by total system loss) over rounds under the Hybrid-learning for all proposed attack types. Note that we let the initial attack to be the same for all of the four attack types. We note that our proposed Hybrid-learning algorithm is effective in reducing total system loss for all attack types with emphasizing that it is more effective with the replay attacker (i.e., the attacker that chooses same attack path for each critical asset every round) and the minmax attacker (i.e., the attacker that chooses the attack path with the highest probability for each critical asset every round). The enhancement is also noticeable for the two other attack types (i.e., randomizing and adaptive) but with less magnitude since capturing the attack patterns by the defender is more challenging in these two attack types. Note also that the spikes in the figure corresponds to the rounds in which the defender invest sub-optimally. These spikes decrease with rounds since the probability of investing behaviorally decreases as defender learns and enhance her budget distribution over rounds.

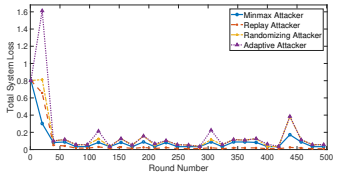
Benefit of Hybrid Learning: We show the benefit of the learning techniques by calculating the Average Gain of Learning (which is the ratio of total system loss after learning to the total system loss with no learning averaged over the four attack types we study in this paper). Figure 9c shows the Average Gain of learning for the three earning techniques: Learning attack paths only (Algorithm 1), Reinforcement Learning only (Algorithm 2), and Hybrid Learning. We observe the superiority of Hybrid Learning compared to using only one of the two learning techniques for all of the five systems. The intuition is that this Hybrid learning combines both learning behavioral level with learning attack paths.

6.4 Baseline Systems

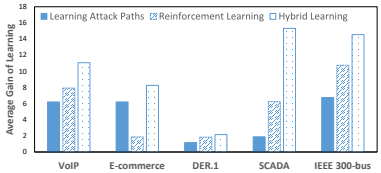
We compare Morshed with two baseline systems: the seminal work of [33] for security investment with attack graphs on attack



(a) Convergence of Reinforcement learning to rational behavior for the five studied interdependent systems.



(b) Defense Enhancement under Hybrid Learning for different attack types. The spikes (that represents investing suboptimally) decreases in later rounds.



(c) Average Gain in Total System Loss for the different Learning techniques. The Hybrid Learning is superior for all five systems.

Figure 9: (a) shows the convergence of Reinforcement learning for all systems. (b) shows the effect of Hybrid learning for each attack type. In (c), we show the average gain of learning for all systems.

graph generation and investment decision analysis[6] and [23] for placing security resources using defense in depth technique which traverses all edges that can be used to compromise each critical asset and distribute resources equally on them. In [33], the defense mechanism is to select the minimal set C of edges that, if removed from the attacker's arsenal, will prevent her from reaching the target asset (there can be multiple sets in case of non-uniqueness). This is equivalent to our min edge-cut. We compare [33] and [23] with Morshed under both single and multi-round setups. We compare the two methods in Table 3 by calculating the probability of successful attack (PSA) and show the superiority of Morshed in multi-round for all different attack types. Note that the defense investments given by Morshed for non-behavioral defenders is identical to that determined by [33] in single-round setup.

6.5 Evaluation of Multiple-defender Setups

Here, we evaluate our proposed algorithm in multiple-defender setups. There are six parameters that could affect the total loss of the defender. The six parameters are: defenders' security budget availability (Low, Moderate, and High), the defense mechanism (Individual, and Joint), the budget distribution among defenders (Symmetric, and Asymmetric), the degree of interdependency (number of edges between defenders' subnetworks), the sensitivity of edges to investments (the hyperparameter $s_{i,j}$), and the edges' baseline probabilities of successful attacks (the hyperparameter $p^0_{i,j}$). When studying the impact of a specific parameter, we fix the remaining parameters to their default values. Next, we study the impact of each system parameter with the behavioral decision-making and identify the effects of these system parameters on the degree of suboptimality of security outcomes due to behavioral decision-making in the two-defenders SCADA system.

1) Effect of defense mechanism: We observe the merits of cooperation (i.e., joint defense) in decreasing the total loss to the defenders as shown in Figure 10. The effect is more pronounced for a higher degree of behavioral bias of the defenders. For example, at moderate budget ($BT = 20$), the relative decrease in total system loss due to joint defense at $\alpha_1 = \alpha_2 = 0.4$ is 25% while $\alpha_1 = \alpha_2 = 0.8$, the decrease is lower (10%). Thus, as the defenders exhibit higher degree of cognitive bias, it is more advantageous to adopt joint defense mechanisms.

2) Interdependency among different defenders. Here, we observe effect of interdependency between defenders on the security of the SCADA system. In the SCADA system, the degree of interdependency increases if assets from one subnetwork can access assets in the other, without going through the Corporate or Vendor nodes. For example, if the attacker gets access to Control unit 1, this enables her to compromise RTU2 as well, in addition to RTU1.

[6]More recent approaches (e.g., [38]) follow the same strategy proposed in [33].

Table 3: Comparison of Morshed and baseline systems for different attacks scenarios. We consider here rational defender for Morshed. The column "System Setup" shows the specific scenario; the second, third, and forth columns show the respective probability of successful attack (PSA) under [33], [23], and Morshed for each scenario.

System Setup	[33]	[23]	Morshed
DER.1			
	PSA		
Single-round	**0.075**	0.208	**0.075**
Multi-round, Random Att.	0.095	0.205	**0.080**
Multi-round, Replay Att.	0.075	0.208	**0.037**
Multi-round, Adaptive Att.	0.091	0.209	**0.080**
SCADA			
Single-round	**0.035**	0.110	**0.035**
Multi-round, Random Att.	0.034	0.582	**0.029**
Multi-round, Replay Att.	0.033	0.110	**0.010**
Multi-round, Adaptive Att.	**0.035**	0.582	**0.035**
VOIP			
Single-round	**0.337**	0.556	**0.337**
Multi-round, Random Att.	0.348	0.559	**0.313**
Multi-round, Replay Att.	0.337	0.556	**0.084**
Multi-round, Adaptive Att.	0.354	0.559	**0.313**
E-commerce			
Single-round	**0.124**	0.276	**0.124**
Multi-round, Random Att.	0.139	0.572	**0.097**
Multi-round, Replay Att.	0.124	0.276	**0.007**
Multi-round, Adaptive Att.	0.139	0.569	**0.097**
IEEE 300-BUS			
Single-round	**0.431**	0.653	**0.431**
Multi-round, Random Att.	0.439	0.680	**0.168**
Multi-round, Replay Att.	0.431	0.653	**0.086**
Multi-round, Adaptive Att.	0.448	0.680	**0.186**

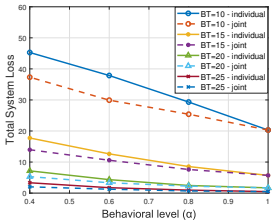


Figure 10: Comparison between individual and joint defense mechanisms. Joint defense is superior under asymmetric budget distribution.

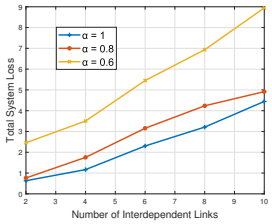
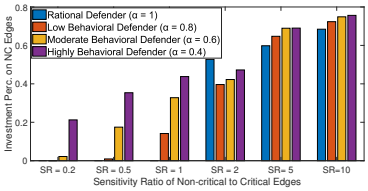
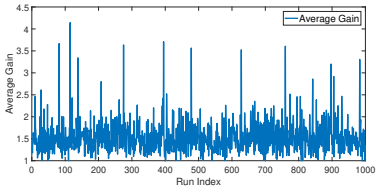


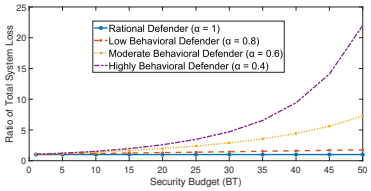
Figure 11: The effect of increasing the degree of interdependency on the total system loss. Such effect is more pronounced when the defender is more behavioral.



(a) Percentage of investments on non-critical edges for different edge sensitivities ratios (SR).



(b) Average gain for different random attack success probabilities over all of the edges.



(c) Effect of behavioral investments for different security budgets.

Figure 12: (a) The effect of edges' sensitivities on investments for different behavioral levels. (b) The average gain of rational decision-making for randomly chosen baseline probabilities of successful attacks. (c) The effect of sub-optimal investments for different choices of security budget.

Figure 11 illustrates this effect—as the number of interdependent edges between the two defenders increases, the total system loss increases in both non-behavioral and behavioral security games. The highest level of interdependency is when there are two edges between DMZ1 and DMZ2, between Control1 and Control2, and the controller to the 3 RTUs of the other defender. An example of this phenomenon is that if both defenders are non-behavioral and the level of interdependency is the highest, the total system loss is higher by 462% over the case of the lowest level of interdependency (2 interdependent links). We also see that as the interdependency between the different defenders increases, the suboptimal security decisions have greater adverse impact on the total system loss.

3) Sensitivity of edges to investments: We next consider the effects of different sensitivities of edges to security investments. Recall that higher sensitivity edges are those for which the probability of successful attack decreases faster with each unit of security investment. We show the result in Figure 12a by using as the independent variable the ratio of sensitivity of non-critical to critical edges. First, assume critical edges correspond to mature systems that are already highly secure and difficult to secure further. For our model, this translates to high (resp. low) $s_{i,j}$ for non-critical (resp. critical) edges. We observe that as the sensitivity ratio increases, all defenders put more investments on the non-critical edges, but the increase is slower in behavioral defenders. However, lower sensitivity ratio will result in investing almost all budget on these critical edges, even for behavioral defenders.

4) Baseline probabilities of successful attacks: We show that the gain of rational vs. behavioral investments exists for any combination of baseline probabilities by performing 1000 runs and in each run, for each edge, we draw the baseline probability of successful attack on that edge from a uniform distribution $p^0_{i,j} \sim \mathcal{U}(0,1]$. We consider a symmetric budget distribution and medium security budgets. Figure 12b shows that the gain for rational over behavioral decision-making (with mean 1.53X) exists for any randomly chosen baseline probability of successful attacks.

5) Amount of security budget: We next show that the total system loss of rational defenders is less than that of behavioral defenders for any choice of security budget (as shown in Figure 12c).

6) Security budget distribution among defenders: Finally, we analyze the effect of asymmetric budget distribution between the defenders facing the attacker. Figure 17 (in Appendix E) illustrates the total loss as a function of the fraction of defender 1's budget. For the individual-defense loss, we observe that the suboptimality of behavioral decision-making is more pronounced with higher budget asymmetries. For example, if defender 1 has 20% of the total budget, the relative increase in total loss from $\alpha = 1$ to $\alpha = 0.4$ is 25%. In contrast, the same change of α when the budget is symmetric results in only a 6% relative increase in the total loss. This observation can be explained by two facts. First, with suboptimal behavioral allocation, the poorer defender wastes even her constrained budget on non-critical edges. Second, the richer player also allocates her resources suboptimally. This leads to this magnified relative increase in losses under budget asymmetry.

In the interest of space, we present system parameters evaluation of DER.1, which has similar insights as SCADA, in Appendix E.

7 LIMITATIONS AND DISCUSSION

Guiding security decision-makers: We believe that our work opens up a new dimension of *intervention* in securing interdependent systems. Our framework allows a quantification of the improvements in security that can be obtained by training security professionals to reduce their behavioral biases. In this context, we can quantitatively show the decision-maker the improvement in system security when moving from her current (sub-optimal) investments to that given by a (rational) algorithm (e.g., Morshed with $\alpha = 1$). Furthermore, our framework can guide security audits by system operators of large-scale interdependent system, by allowing the operator to investigate subsystems within the system where sub-optimal security investments might have been made by subordinates operating those subsystems. While such an operator may not be able to check every single aspect of every subsystem, she may be able to "zoom in" to portions of the overall system where an audit may be warranted due to evidence of sub-optimality from our framework.

Behavioral level of the attacker: We assume that defenders perceive the attacker as non-behavioral; in reality the attacker can be behavioral as well. Our assumption of a non-behavioral attacker gives the worst case loss for the system; as a behavioral attacker may not choose the path of true highest vulnerability due to probability misperceptions. This can open the interesting question "how a rational defender, who uses the security investments recommended by Morshed, can deceive a behavioral attacker to choose harder attack paths?" This can help the defender to misguide the attacker and make the target system more secure.

Multi-hop dependence: In several cybersecurity scenarios, the ease of an attacker in achieving an attack goal depends not just on

the immediate prior attack step but on steps farther back. In such scenarios, the simpler formulation of using probabilities on each edge and assuming independence of the events of traversing the different edges can lead to inaccurate estimates. However, we follow several prior works (e.g., [27, 36]) that leveraged the property that in most cases, a node has the highest dependence on the previous node, in order to build computationally tractable analysis tools. Moreover, to handle this issue in our model, the notion of *k-hop dependence* [24] can be used, whereby the probability of reaching a particular node depends nodes up to k hops away.

8 RELATED WORK

Security in interdependent systems: The problem of securing systems with interdependent assets has been handled in several prior works [27, 36]. The common theme is that a successful attack to one asset may be used to compromise a dependent asset. The notion of attack graphs [16] is a popular abstraction for capturing the security interdependencies. The specific works differ in what the assets are (physical or virtual, resource-constrained nodes, networking assets, etc.), the level of observability into the states of the assets, and the probabilistic reasoning engine used. Our work here differs from these works in that the prior work creates algorithms to make the security control decisions, while we are considering humans with cognitive biases making these decisions.
Game-theoretic modeling of security: Game theory has been used to describe the interactions between attackers and defenders and their effects on system security. A commonly used model in this context is that of two-player games, where a single attacker attempts to compromise a system controlled by a single defender [5, 31]. Game theoretic models have been further used in [37] to study the interaction between one defender and (multiple) attackers attempting Distributed Denial of Service attacks. Game theoretic models have also been proposed for studying critical infrastructure security (See the survey [22]). The major difference of our work with all aforementioned literature is that existing work has focused on classical game-theoretic models of rational decision-making, while we analyze behavioral models of decision-making.
Human behavior in security and privacy: Notable departure from classical economic models within the security and privacy literature is [4], which identifies the effects of behavioral decision-making on individual's personal privacy choices. The importance of considering similar models in the study of system security has been recognized in the literature [8]. Prior works [6, 30] considered models from behavioral economics in the context of security applications. However, these works are based only on psychological studies [6] and human subject experiments [30] for end-user. Our work differs from these in that we explore a rigorous mathematical model of defenders' (decision-makers) behavior, model the interaction between multiple defenders (in contrast to the study of only one defender for all of these studies), and consider interdependent assets (in contrast to these studies which reason about binary decisions on isolated assets). To the best of our knowledge, the exceptions that provide a theoretical treatment of behavioral decision-making in certain specific classes of interdependent security games are [1, 2, 18, 32]. These works, however, are theoretical in scope and do not consider the more realistic attack scenarios and types that we consider, do not validate bias of decision-makers via subject experiments, and don't consider multi-round setups or learning algorithms that we consider here.
Multi-round in Security: Reinforcement-based learning models have been used in literature where players' strategies receive reinforcement related to the payoffs they earn and adjust their moves over time seeking higher payoffs. Specifically, [10] proposed reinforcement learning for an environment with only two possible actions. Such Reinforcement-based learning models have been used in different security applications such as the robustness of smart grid [28]. Our work differs from these works that we guide the behavioral decision-maker towards rational decision-making where the reinforcements are received from the true loss that the defender accrues when investing with behavioral bias. Likelihood method of discovering attack paths using Bayesian attack graphs has been proposed in [36]. However, to the best of our knowledge, no previous work has the idea of minimizing the adapted defender's cost and generate optimal allocations each round while weighting attack paths based on previous rounds that we consider in our Hybrid-learning algorithm.

9 CONCLUSION

We presented *behavioral security games* to study the effects of human behavioral decision-making on the security of interdependent systems with multiple defenders where we model stepping-stone attacks by the notion of *attack graphs*. While behavioral decision-makers tend to allocate their budget across the network, Morshed helps decision-makers concentrate their budget on critical edges to make the system more secure. We performed a controlled subject experiment to validate our behavioral model. In multi-round setups, we proposed different learning algorithms to guide behavioral decision-makers towards optimal decisions. We evaluated Morshed on five real case studies of interdependent systems where we studied the effects of several system parameters. The insights gained from our analysis would be useful for configuring real-world systems with optimal parameter choices and guiding behavioral decision-makers toward rational decision-making that can ultimately lead to improvements in interdependent systems' security.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their valuable comments to improve the quality of this paper. This material is based in part upon work supported by the National Science Foundation under Grant Number CNS-1718637, Wabash Heartland Innovation Network (WHIN) project from Lilly Endowment Inc. NSF CCF-1919197, and Army Research Lab under Contract number W911NF-2020-221. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

[1] M. Abdallah, P. Naghizadeh, T. Cason, S. Bagchi, and S. Sundaram. 2019. Protecting Assets with Heterogeneous Valuations under Behavioral Probability Weighting. In *2019 IEEE 58th Conference on Decision and Control (CDC)*. 5374–5379.
[2] M. Abdallah, P. Naghizadeh, A. R. Hota, T. Cason, S. Bagchi, and S. Sundaram. 2020. Behavioral and Game-Theoretic Security Investments in Interdependent

Systems Modeled by Attack Graphs. *IEEE Transactions on Control of Network Systems* (2020).
[3] M. Abdallah, D. Woods, P. Naghizadeh, I. Khalil, T. Cason, S. Sundaram, and S. Bagchi. 2020. BASCPS: How does behavioral decision making impact the security of cyber-physical systems? *arXiv preprint arXiv:2004.01958* (2020).
[4] A. Acquisti. 2009. Nudging privacy: The behavioral economics of personal information. *IEEE security & privacy* 7, 6 (2009).
[5] T. Alpcan and T. Basar. 2006. An intrusion detection game with limited observations. In *12th Int. Symp. on Dynamic Games and Applications*, Vol. 26.
[6] R. Anderson. 2012. Security economics: a personal perspective. In *Proceedings of the 28th Annual Computer Security Applications Conference*. ACM, 139–144.
[7] S. Benartzi and R. H. Thaler. 2001. Naive diversification strategies in defined contribution saving plans. *American economic review* 91, 1 (2001), 79–98.
[8] L. F. Cranor. 2008. A framework for reasoning about the human in the loop. *Proc. 1st Conference on Usability, Psychology, and Security, Usenix Assoc.* (2008).
[9] D. Dor and Y. Elovici. 2016. A model of the information security investment decision-making process. *Computers & security* 63 (2016), 1–13.
[10] N. Feltovich. 2000. Reinforcement-based vs. belief-based learning models in experimental asymmetric-information games. *Econometrica* 68, 3 (2000), 605–641.
[11] G. R. Fréchette and A. Schotter. 2015. *Handbook of experimental economic methodology*. Oxford University Press, USA.
[12] R. Gonzalez and G. Wu. 1999. On the shape of the probability weighting function. *Cognitive psychology* 38, 1 (1999), 129–166.
[13] L. A. Gordon and M. P. Loeb. 2002. The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC)* 5, 4 (2002), 438–457.
[14] L. Haynes, B. Goldacre, D. Torgerson, et al. 2012. Test, learn, adapt: developing public policy with randomised controlled trials. *Cabinet Office-Behavioural Insights Team* (2012).
[15] F. S. Hillier. 2012. *Introduction to operations research*.
[16] J. Homer, S. Zhang, X. Ou, D. Schmidt, Y. Du, S. R. Rajagopalan, and A. Singhal. 2013. Aggregating vulnerability metrics in enterprise networks using attack graphs. *Journal of Computer Security* 21, 4 (2013), 561–597.
[17] A. R. Hota, A. Clements, S. Sundaram, and S. Bagchi. 2016. Optimal and game-theoretic deployment of security investments in interdependent assets. In *International Conference on Decision and Game Theory for Security*. 101–113.
[18] A. R. Hota and S. Sundaram. 2018. Interdependent Security Games on Networks Under Behavioral Probability Weighting. *IEEE Transactions on Control of Network Systems* 5, 1 (March 2018), 262–273. https://doi.org/10.1109/TCNS.2016.2600484
[19] S. Jauhar, B. Chen, W. G. Temple, X. Dong, Z. Kalbarczyk, W. H. Sanders, and D. M. Nicol. 2015. Model-based cybersecurity assessment with nescor smart grid failure scenarios. In *Dependable Computing (PRDC), 2015 IEEE 21st Pacific Rim International Symposium on*. IEEE, 319–324.
[20] D. Kahneman and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society* 47 (1979), 263–291.
[21] M. Khanabadi, H. Ghasemi, and M. Doostizadeh. 2012. Optimal transmission switching considering voltage security and N-1 contingency analysis. *IEEE Transactions on Power Systems* 28, 1 (2012), 542–550.
[22] A. Laszka, M. Felegyhazi, and L. Buttyan. 2015. A survey of interdependent information security games. *ACM Computing Surveys (CSUR)* 47, 2 (2015), 23.
[23] R. Lippmann, K. Ingols, C. Scott, K. Piwowarski, K. Kratkiewicz, M. Artz, and R. Cunningham. 2006. Validating and restoring defense in depth using attack graphs. In *IEEE Military Communications Conference*. IEEE, 1–10.
[24] R. Maheshwari, J. Gao, and S. R. Das. 2007. Detecting wormhole attacks in wireless networks using connectivity information. In *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications*. IEEE, 107–115.
[25] K. Mersinas, B. Hartig, K. M Martin, and A. Seltzer. 2015. Experimental Elicitation of Risk Behaviour amongst Information Security Professionals.. In *14th Workshop on the Economics of Information Security (WEIS)*.
[26] K. Mersinas, B. Hartig, K. M. Martin, and A. Seltzer. 2016. Are information security professionals expected value maximizers?: An experiment and survey based test. *Journal of Cybersecurity* 2, 1 (12 2016), 57–70.
[27] G. Modelo-Howard, S. Bagchi, and G. Lebanon. 2008. Determining placement of intrusion detectors for a distributed application through bayesian network modeling. In *International Workshop on Recent Advances in Intrusion Detection*. Springer, 271–290.
[28] Z. Ni and S. Paul. 2019. A multistage game in smart grid security: A reinforcement learning solution. *IEEE transactions on neural networks and learning systems* 30, 9 (2019), 2684–2695.
[29] D. Prelec. 1998. The probability weighting function. *Econometrica* 66, 3 (1998), 497–527.
[30] E. M. Redmiles, M. L. Mazurek, and J. P. Dickerson. 2018. Dancing pigs or externalities?: Measuring the rationality of security decisions. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. ACM, 215–232.
[31] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu. 2010. A survey of game theory as applied to network security. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*. IEEE, 1–10.
[32] A. Sanjab, W. Saad, and T. Başar. 2017. Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game. In *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 1–6.
[33] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing. 2002. Automated generation and analysis of attack graphs. In *Proceedings 2002 IEEE Symposium on Security and Privacy*. IEEE, 273–284.
[34] B. Wang and N. Z. Gong. 2019. Attacking graph-based classification via manipulating the graph structure. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS)*. 2023–2040.
[35] D. Woods, M. Abdallah, S. Bagchi, S. Sundaram, and T. Cason. 2020. Network Defense and Behavioral Biases: An Experimental Study. (2020).
[36] P. Xie, J. H. Li, X. Ou, P. Liu, and R. Levy. 2010. Using Bayesian networks for cyber security analysis. In *Dependable Systems and Networks (DSN), 2010 IEEE/IFIP international conference on*. IEEE, 211–220.
[37] G. Yan, R. Lee, A. Kent, and D. Wolpert. 2012. Towards a bayesian network game framework for evaluating DDoS attacks and defense. In *Proceedings of the 2012 ACM conference on Computer and communications security (CCS)*. 553–566.
[38] M. Zhang, L. Wang, S. Jajodia, A. Singhal, and M. Albanese. 2016. Network diversity: a security metric for evaluating the resilience of networks against zero-day attacks. *IEEE Transactions on Information Forensics and Security* 11, 5 (2016), 1071–1086.

A MOTIVATIONAL EXAMPLE WITH DIFFERENT SENSITIVITIES

In the above example, we assumed all edges have the same sensitivity to investments. In cases where critical edges have equal or higher sensitivity than non-critical edges, the same insight as above holds. Specifically, when edge (v_i, v_j) has sensitivity $s_{i,j}$, one can verify (using KKT conditions) that the optimal investments by a behavioral defender are given by

$$x_{1,2} = x_{2,4} = x_{1,3} = x_{3,4} = 2^{\frac{1}{\alpha-1}} \left(\frac{s_{i,j}}{s_{s,1}}\right)^{\frac{\alpha}{1-\alpha}} x_{s,1}.$$

$$x_{s,1} = \left(\frac{s_{s,1}}{s_{4,5}}\right)^{\frac{\alpha}{1-\alpha}} x_{4,5};\ x_{4,5} = B - \sum_{\forall (i,j) \neq (v_4, v_5)} x_{i,j}.$$

The insight here is that the investment decision has two dimensions: behavioral level and sensitivity ratio of non-critical edges to critical edges. Specifically, as the defender becomes more behavioral, she puts less investments on edges with higher sensitivity.

B HUMAN SUBJECT DEMOGRAPHICS

The 145 human subjects in our experiment are comprised of 78 males (53.79%) and 67 females (46.21%). They belong to various majors on campus, with the three largest being Management/Business (24.8%), Engineering (24.2%), and Science (23.5%). Regarding year in college, 6.9% are 1st year, 13.1% are 2nd year, 21.38% are 3rd year, 35.86% are 4th year, and 22.76% are graduate students. Regarding the GPA distribution, 44.83% have GPAs between 3.5 and 4, 35.17% between 3 and 3.5, and 17.93% between 2.5 and 3.

C CONVERGENCE OF REINFORCEMENT LEARNING OF BEHAVIORAL LEVEL

Lemma C.1. *Let N_1 and N_{α_i} represent the number of rounds in which the defender chose to invest rationally and with behavioral level α_i, respectively. Let $\hat{C}_{opt}$ and $\hat{C}_i$ be the total real loss incurred by the defender when investing rationally and with behavioral level α_i, respectively. Then, we have*

(1) If $q^0(\alpha_i) = q^0(1)$, then Algorithm 2 converges to rational behavior in the round N_R if $\frac{\hat{C}_{max} - \hat{C}_{opt}}{\hat{C}_{max} - \hat{C}_i} >> \frac{N_{\alpha_i}}{N_1}$.

Table 4: Baseline probability of successful attack for vulnerabilities in SCADA and DER.1 failure scenarios.

Vulnerability (CVE-ID)	Edge(s)	Attack Vector	Score
SCADA application			
Control Unit (CVE-2018-5313)	(Vendor,Control1),(Vendor,Control2)	Local	0.78
Remote authentication (CVE-2010-4732)	(S, Vendor)	Network	0.9
Remote cmd injection (CVE-2011-1566)	(Control,RTU1),(Control,RTU2)	Network	1.0
Authentication bypassing (CVE-2019-6519)	(Corp,DMZ1),(Corp,DMZ2)	Network	0.75
DER.1 application			
Physical access (CVE-2017-10125)	$(w_9, w_7),(w_{18}, w_{16})$	Physical	0.71
Network access (CVE-2019-2413)	$(w_9, w_8),(w_{18}, w_{17})$	Network	0.61
Software access (CVE-2018-2791)	$(w_7, w_6),(w_8, w_6)$	Network	0.82
Sending cmd (CVE-2018-1000093)	$(w_6, w_5),(w_{15}, w_{14})$	Network	0.88

(2) If $q^0(\alpha_i) = A_i$ and $q^0(1) = B$, then Algorithm 2 converges to rational behavior in the round N_R if $(N_1 - N_{\alpha_i})\hat{C}_{max} + N_{\alpha_i}(\hat{C}_i) - N_1(\hat{C}_{opt}) >> A_i - B$.

Proof. We first calculate the propensities for each behavioral level. From Algorithm 2, we have

$$q^{N_R}(1) = q^0(1) + N_1(\hat{C}_{max} - \hat{C}_{opt})$$
$$q^{N_R}(\alpha_i) = q^0(\alpha_i) + N_{\alpha_i}(\hat{C}_{max} - \hat{C}_i)$$

(i) To reach convergence, $\forall \alpha_i \neq 1$ and $q^0(\alpha_i) = q^0(1)$, we have

$$q^{N_R}(1) >> q^{N_R}(\alpha_i) \iff N_1(\hat{C}_{max} - \hat{C}_{opt}) >> N_{\alpha_i}(\hat{C}_{max} - \hat{C}_i)$$
$$\iff \frac{\hat{C}_{max} - \hat{C}_{opt}}{\hat{C}_{max} - \hat{C}_i} >> \frac{N_{\alpha_i}}{N_1}$$

(ii) With a similar argument to (i), $\forall \alpha_i \neq 1$ and $q^0(\alpha_i) \neq q^0(1)$ where $q^0(\alpha_i) = A_i$ and $q^0(1) = B$, we have

$$q^{N_R}(1) >> q^{N_R}(\alpha_i)$$
$$\iff (N_1 - N_i)\hat{C}_{max} + N_i(\hat{C}_i) - N_1(\hat{C}_{opt}) >> A_i - B$$

Note that in all of the possible cases, $\hat{C}_i - \hat{C}_{opt} >> 0$ ensures convergence under any choice of A and $B \neq 0$ which is realistic where the real loss associated with suboptimal investments decisions is much higher compared to the real loss associated with optimal (i.e., rational) investments decisions. □

D ATTACK SCENARIOS OF CASE STUDIES

In this section, we provide explanations of the system overview and equivalent attack graph for the DER.1, E-commerce, VOIP, and IEEE 300-BUS, respectively.

D.1 DER.1

System Description: The US National Electric Sector Cybersecurity Organization Resource (NESCOR) Technical Working Group has proposed a framework for evaluating the risks of cyber attacks on the electric grid. A distributed energy resource (DER) is described as a cyber-physical system consisting of entities such as generators, storage devices, and electric vehicles, that are part of the energy distribution system. The DER.1 failure scenario has been identified as the riskiest failure scenario affecting distributed energy resources according to the NESCOR ranking. As shown in Figure 15, there are two critical equipment assets: a PhotoVoltaic (PV) generator and an electric vehicle (EV) charging station. Each piece of equipment is accompanied by a Human Machine Interface (HMI), the only gateway through which the equipment can be controlled. The DER.1 failure scenario is triggered when the attacker gets access to the HMI. The vulnerability of the system may arise due to various reasons, such as hacking of the HMI, or an insider attack. Once the attacker gets access to the system, she changes the DER settings and gets physical access to the DER equipment so that they continue to provide power even during a power system fault. Through this manipulation, the attacker can cause physical damage to the system.

D.2 E-commerce

System Description: In that E-commerce system (shown in Figure 13), all servers are running a Unix-based operating system. The web server sits in a demilitarized zone (DMZ) separated by a firewall from the other two servers, which are connected to a network not accessible from the Internet. All connections from the Internet and through servers are controlled by the firewall. Rules state that the web and application servers can communicate, as well as the web server can be reached from the Internet. Here, the attacker is assumed to be an external one and thus her starting point is the Internet which uses stepping-stone attacks with the goal of having access to the MySQL database (specifically access customer confidential data such as credit card information), represented by node 19 in the attack graph. For this system, we follow the attack graph generated by [27] (shown in Figure 13), which is based on the vulnerabilities associated with specific versions of the particular software, and are taken from popular databases.

D.3 VoIP

System Description: As shown in Figure 14, the VoIP system is composed of three zones; a DMZ for the servers accessible from the internet cloud, an internal network for local resources (e.g., computers, mail server and DNS server), and an internal network that is consisted of only VoIP components. This architecture follows the security NIST guidelines for deploying a secure VoIP system. In this context, the VoIP network consists of a Proxy, voicemail server and software-based and hardware-based phones. The firewall has the rules to control the traffic between the three zones. Note that the DNS and mail servers in the DMZ are the only accessible hosts from the Internet. The PBX server can route calls to the Internet or to a public-switched telephone network (PSTN). The ultimate goal of this multi-stage attack is to eavesdrop on VoIP communication.

D.4 IEEE 300 BUS

System Description: Finally, we consider the widely used benchmark IEEE 300 bus power grid network [21]. We define the network itself as the interdependency graph where each node represents a bus (i.e., the network has 300 nodes), and the physical interconnection between the buses represent the edges. Each bus has generators and/or load centers associated with it. As shown in Figure 16, the 300 bus network data divides the buses or nodes into 3 different regions containing 159, 78 and 63 nodes respectively. We assume that each region is managed by an independent entity.

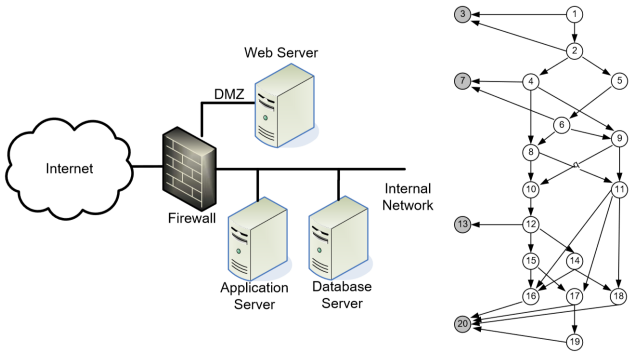


Figure 13: A high level network overview of the E-commerce (on the left) adapted from [27]. The resultant attack graph (on the right).

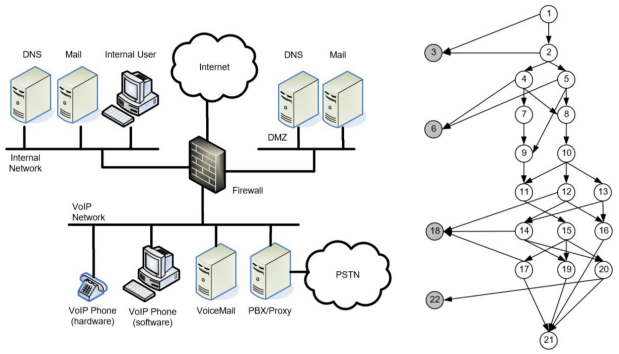
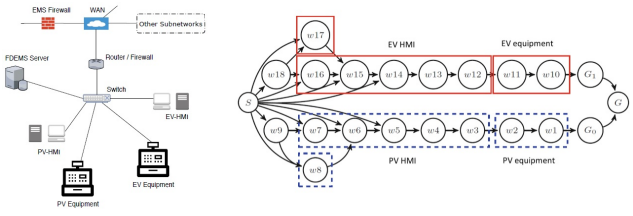


Figure 14: A high level network overview of the VoIP (on the left) adapted from [27] and its resultant attack graph (on the right).



(a) (b)

Figure 15: A network level overview in (a) of a Distributed Energy Resource (DER) system (adapted from [19]). In (b), the equivalent attack graph of that failure scenario is shown.

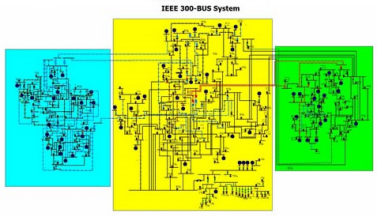


Figure 16: A high level overview of the IEEE 300- BUS (adapted from [21]). Each area has a different color.

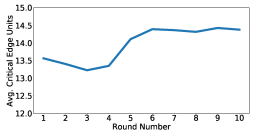


Figure 18: Average of all subjects' investments on the critical edge vs experiment rounds. The upward trend indicates that on average, subjects are learning.

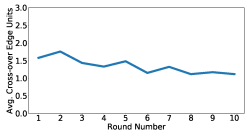


Figure 19: Average of all subjects' investments on the cross-over edge vs experiment rounds. There is only a weak downward trend in spreading behavior.

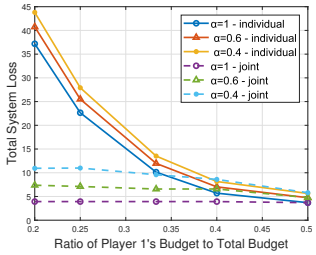
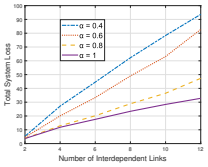
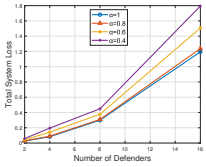


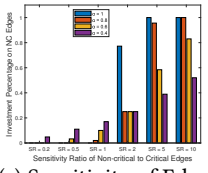
Figure 17: The total system loss as a function of the fraction of defender 1's budget. We observe that joint defense outperforms individual defense at higher budget asymmetry.



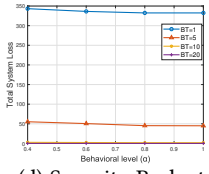
(a) Interdependency Effect



(b) Number of Defenders



(c) Sensitivity of Edges



(d) Security Budget

Figure 20: Results of Multi-defenders for DER.1 system.

E EVALUATION-EXTENDED

E.1 Muti-Defenders: DER.1

We present the system parameters results (shown in Figure 20) for the DER.1 interdependent system. We observe similar insights to SCADA's results (Section 6) and the remaining systems. The detailed explanations of such evaluations are on the arXiv version [3].

E.2 Average Investments of Multi-rounds

We show the average investments for each round for both of the attack graphs tested in our human subject study. Note that we emulated the reinforcement learning environment where in each round as simulated attack is run and result shown to the subject [35].