Morshed: Guiding Behavioral Decision-Makers towards better Security Investment in Interdepndent Systems

Mustafa Abdallah
School of Electrical and Computer Engineering
Purdue University

Based on joint work with
Daniel Woods[3], Parinaz Naghizadeh[2], Issa Khalil[4], Timothy Cason[3], Shreyas Sundaram[1], and Saurabh Bagchi[1]
[1]School of Electrical and Computer Engineering, Purdue University
[2]School of Electrical and Computer Engineering, Ohio State University
[3]Krannert School of Management, Purdue University.
[4]Qatar Computing Research Institute (QCRI).

Funding sources:

Agenda

- Motivation
- Main Contributions
- Related Work
- System Overview
- Multi-round Analysis
- Evaluation
- Human Subject Experiment
- Conclusion

Motivation

- Security of large-scale systems (such as the power grid, industrial plants, and computer networks) depends critically on **human decisions**.

- Many papers on optimal decision making for protecting interconnected systems (e.g., [Laszka et. al., CSUR 2015, La, TON 2016, Alpcan et. al., CUS 2010]).
    - Rely on classical economic models of **perfectly rational** and optimal behavior for human decision-makers.

- However, behavioral economics shows humans are only **partly rational** and consistently deviate from the above-mentioned classical models.
    - Prospect theory (Kahneman and Tversky 2002 **Nobel Prize** in economics).

Behavioral Weighting Function

- Prospect theory showed that human perceptions of rewards and losses can differ substantially from their true values.

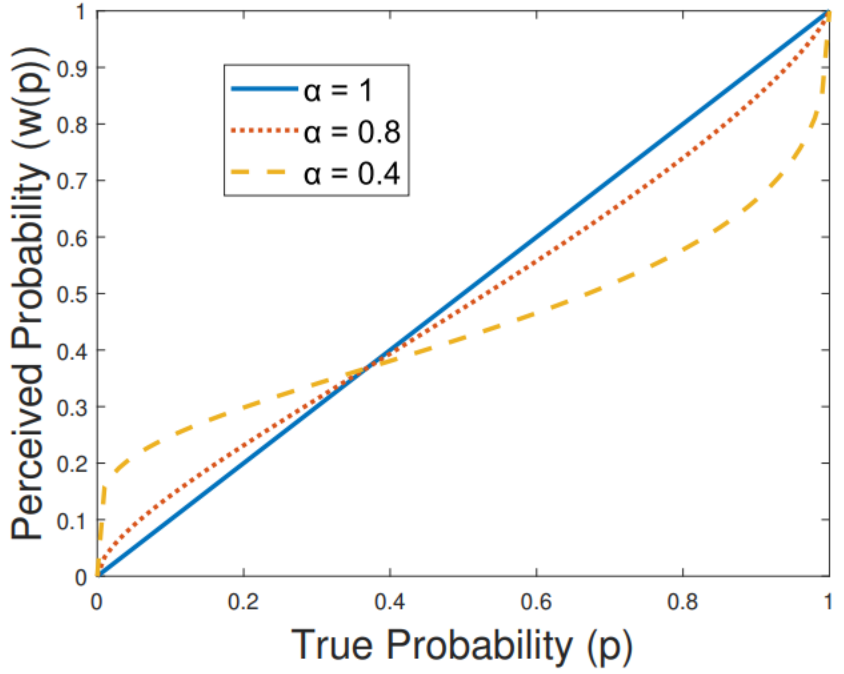
- These perceptions can have a significant impact on the investments made to protect the systems that the individuals are managing.

- Humans overweight low attack probabilities and underweight large attack probabilities.

- Example: Prelec [1998] weighting function:

  $w(\mathrm{p}) = \exp(-(-\ln(\mathrm{p}))^{\alpha})$ where $\alpha \in (0,1]$.

- The smaller is α, the greater is the degree of bias.



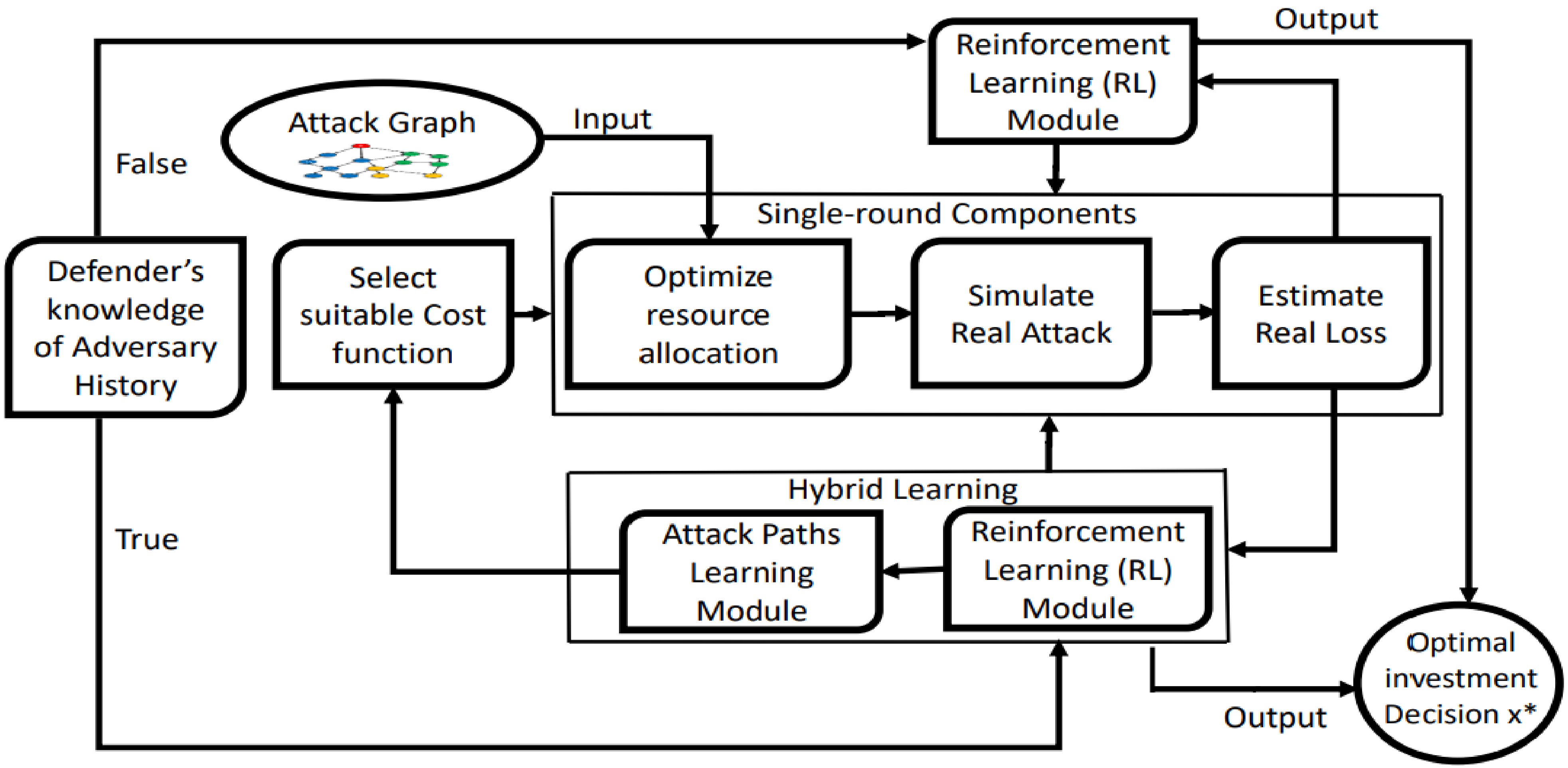
Main Contributions

- 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 under **different attack types**.
- We propose different learning techniques for a **multi-round setup** to enhance behavioral decision-making in our **game-theoretic framework** involving attack graph models of large-scale interdependent systems.
- We evaluate our algorithms via **five interdependent systems** with real attack scenarios and validate our findings by controlled **human subject experiment**.

Related Work

System	Multiple Defenders	Interdependent Subnetworks	Analytical Framework	Behavioral Biases	Various Attack Types	Multiple Rounds
RAID08 [Howard et. al.] MILCOM06 [Lipman et. al.]	✘	✓	✘	✘	✘	✘
S&P02 [Sheyner et. al.] CCS12 [Yan et. al.]	✘	✘	✓	✘	✘	✘
S&P09 [Acquisti] EC18 [Redmiles et. al.] ACSAC12 [Anderson]	✘	✘	✘	✓	✘	✘
TCNS20 [Abdallah et. al.] TCNS18 [Hota et. al.]	✓	✓	✓	✓	✘	✘
MORSHED	✓	✓	✓	✓	✓	✓

High Level System Overview



Single Round Gain for Different Systems

- We evaluate Morshed using **five synthesized attack graphs** that represent **realistic interdependent systems** and attack paths through them.
- The **Avg Gain** is 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 and 50% are behavioral defenders**.
- The **Max Gain** is the ratio of the total system loss of the **highest behavioral defenders** to that with rational defenders.

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

Analysis of Multi-Rounds

- We consider a defender who plays multiple rounds of the game.
- The defender learns from observing the attack in each round.
- In each round, each defender plays **single-shot** game with the attacker, allocating all her security budget.
- **Research Questions:** we explore two different forms of learning:
  - **Q1:** What can the **defender learn about an attacker over time**?
  - **Q2:** How can repeated interactions lead to **decrease in the defenders' extent of behavioral decision-making** (**i.e., increase in α**)?

Learning Attack Paths over Time

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** // Estimate Paths' weights for each critical asset

    for *Path* $P \in \mathcal{P}_m$ **do**

      $\beta_P^t = \frac{1}{N} \sum_{\tau=t-N}^{t-1} [P^\tau = P]_1$ // Compute frequency of opponent actions over past N moves

  $C_k^t(x_k) = \sum_{v_m \in V_k} L_m \left(\sum_{P \in P_m} \beta_P^t \prod_{(v_i, v_j) \in P} w(p_{i,j}(x_{i,j})) \right)$ // Modify the perceived cost based on estimated weights

  $x_k^t \in argmin_{x_k \in X_k} C_k^t(x_k)$

  Append (O, x_k^t)

Return O

Attack Types

- **Replay attacker:** chooses the **same attack path** for every critical asset in **every round** (limited observations or automated attack process).

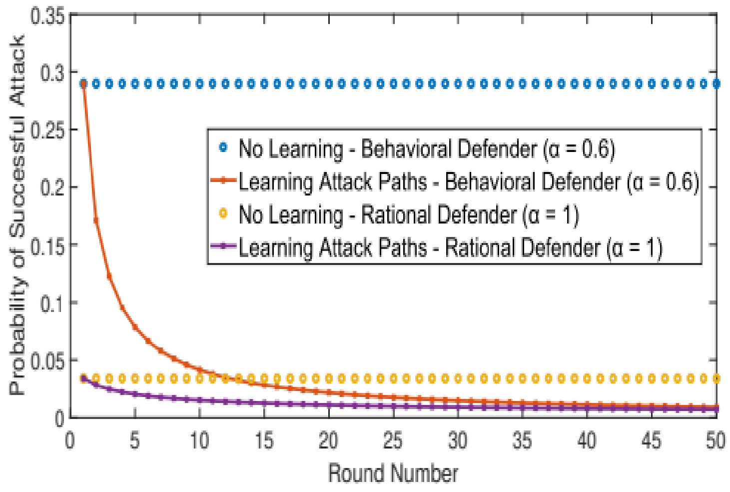
- **Randomizing attacker:** chooses an attack path (for every critical asset) randomly each round with a probability following a **uniform distribution over the possible attack paths** to that asset.

- **Adaptive attacker:** chooses the **least chosen attack path in the past N moves** (for every critical asset).

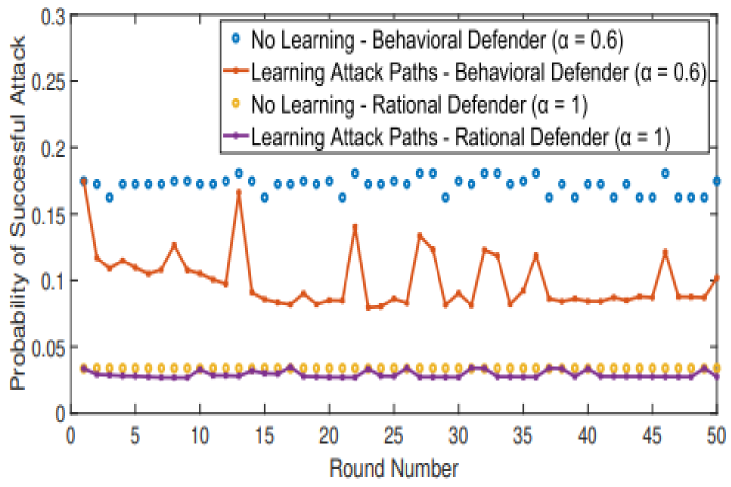
- **Minmax attacker:** chooses the attack path with the **highest probability of successful attack** (for every critical asset).

Results of Learning History Attack Paths

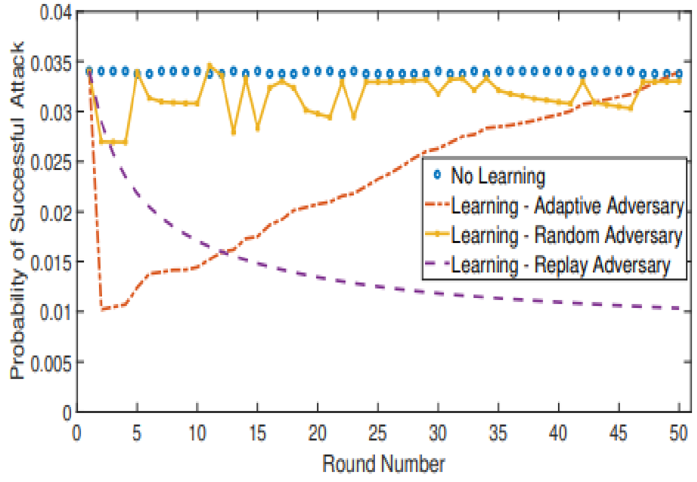
- **Replay attacker pattern** can be expected in less rounds and thus the defender can decreases its adverse effects.
- **Random attacker distribution** can be expected in some sense.
- **Adaptive attacker** is the most challenging attack type.



(a) Attacker chooses same attack paths



(b) Attacker chooses attack paths randomly



(c) Different attack types comparison

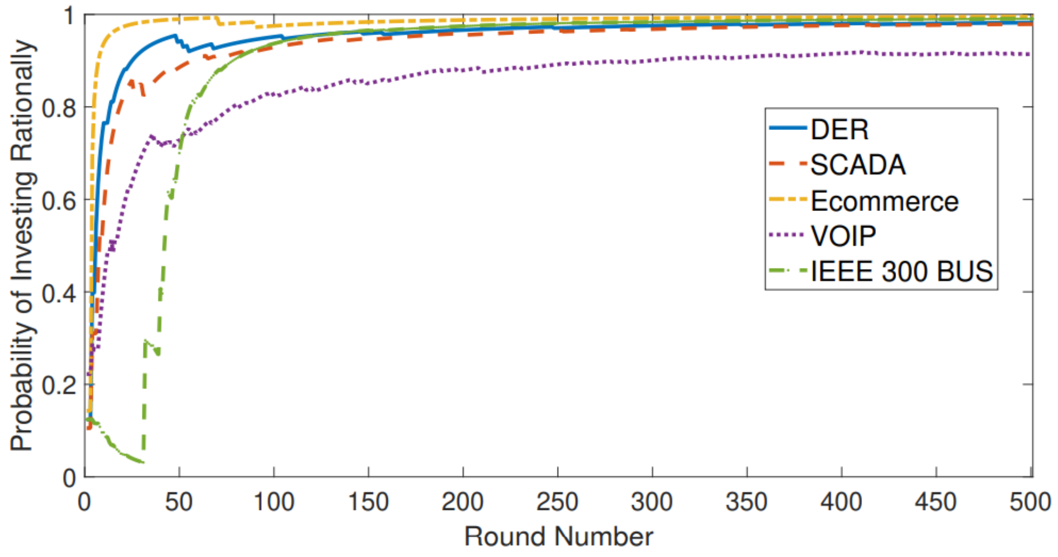
Reinforcement Learning of Behavioral Bias

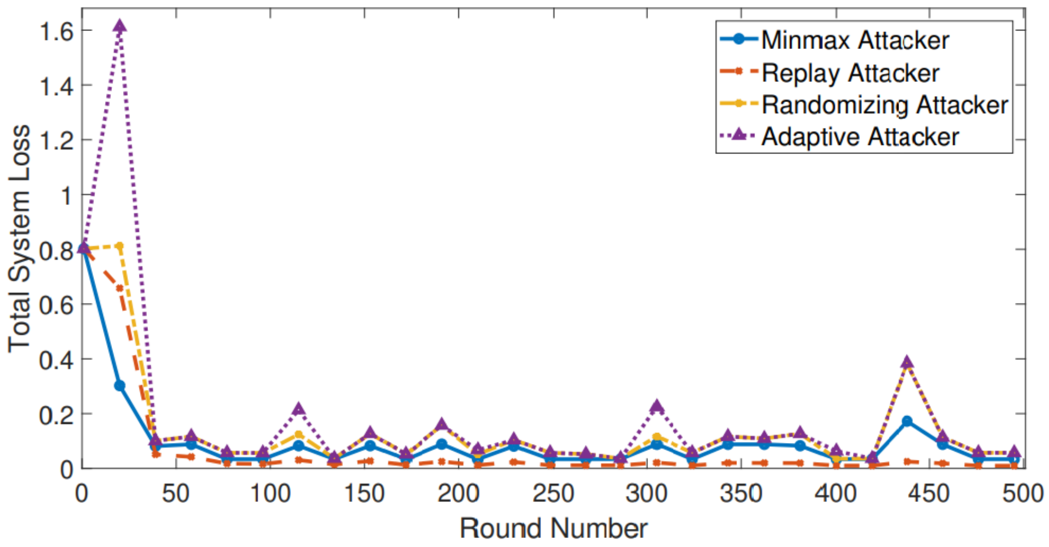
Algorithm 2: Reinforcement Learning to Reduce Behavioral Biases

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**
      $x_k^t \in argmin_{x_k \in X_k} C_k^t(x_k, \alpha_i)$
      $R^t = \hat{C}_{max} - \hat{C}_k^t(x_k^t)$ // Receive reward (punishment) of that round
      $q^{t+1}(\alpha_i) = q^t(\alpha_i) + R^t$
    else
      $q^{t+1}(\alpha_i) = q^t(\alpha_i)$
    $p^{t+1}(\alpha_i) = \frac{q^{t+1}(\alpha_i)}{\sum_{\alpha_i \in \alpha} q^{t+1}(\alpha_i)}$ // Update probability of playing such action in next round
  Sample random α_i with probability $p^{t+1}(\alpha_i)$ to get α^{t+1}
  Append (O, α^{t+1})
Return O

Results of Reinforcement Learning

- Our Reinforcement learning algorithm **converges to rational behavior** for the five studied interdependent systems.
- The defense is enhanced under learning **(in terms of Total System Loss)**. The spikes (that represents investing suboptimally) decrease in later rounds.





Comparison with Baselines

- <u>We compare our system with two baselines:</u>
  O. Sheyner, S&P 2002 [31] (allocates security investments using classical decision-making models).

  Lippmann, MILCOMM 2006 [21] (uses **defense in depth** technique by traversing all edges that can be used to compromise each critical asset and distribute resources equally on them).

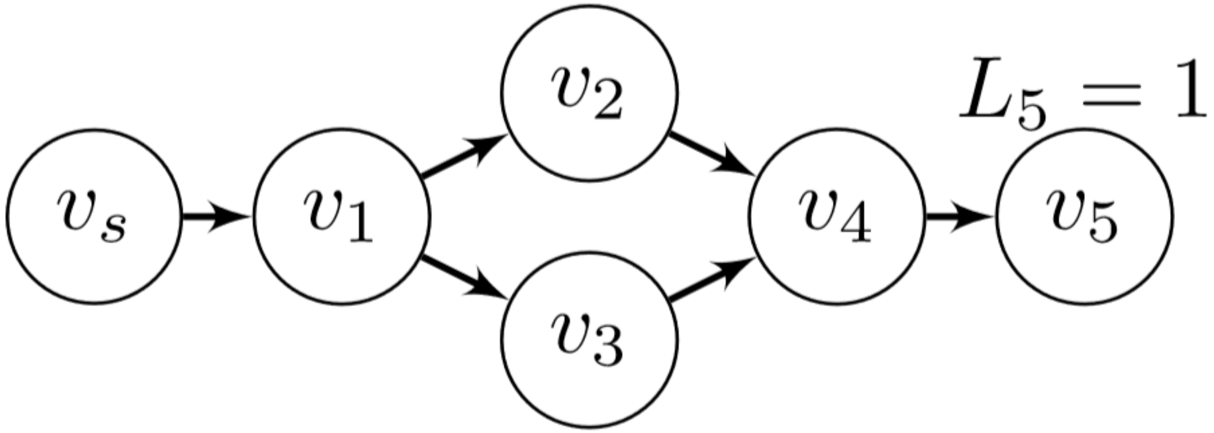
- Same performance (**probability of successful attack (PSA)**) in single-round.

- In multi-round, learning in Morshed is dynamic in contrast to the baselines which results in better performance (i.e., **lower PSA**).

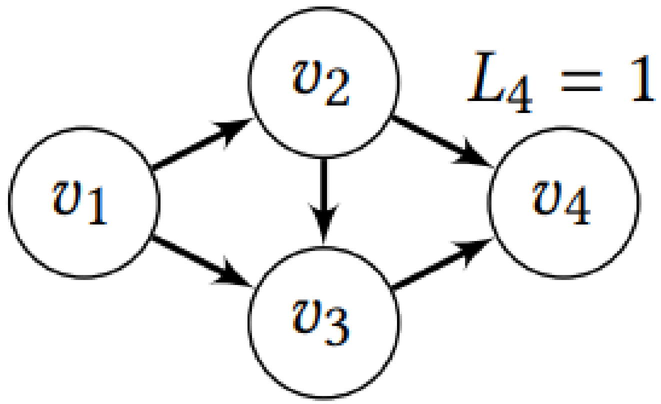
System Setup	[31]	[21]	MORSHED
DER.1			
		PSA	
Single-round	**0.075**	0.208	**0.075**
Multi-round, Random Att.	0.095	0.205	**0.080**
Multi-round, Replay Att.	0.075	0.208	**0.037**
Multi-round, Adaptive Att.	0.091	0.209	**0.080**
SCADA			
Single-round	**0.035**	0.110	**0.035**
Multi-round, Random Att.	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**

Human Subject Experiments

- All experiments have been performed by **Daniel Woods**.
- 145 Students from different departments and different levels.
- Each subject took **10 rounds** of investments for four different networks.
- Instructions about experiments were written and provided to subjects.
- Monetary awards were given to the subject who defends correctly (**by choosing one random round**).



Network with critical edge (Probability Weighting Bias)

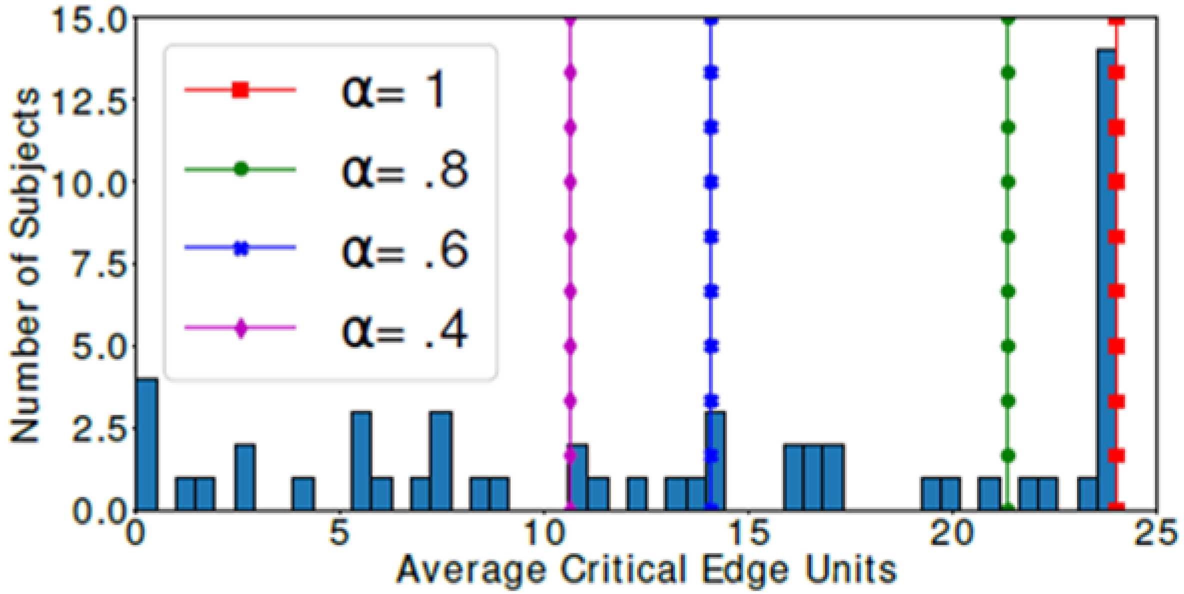


Network with cross-over edge (Spreading Bias)

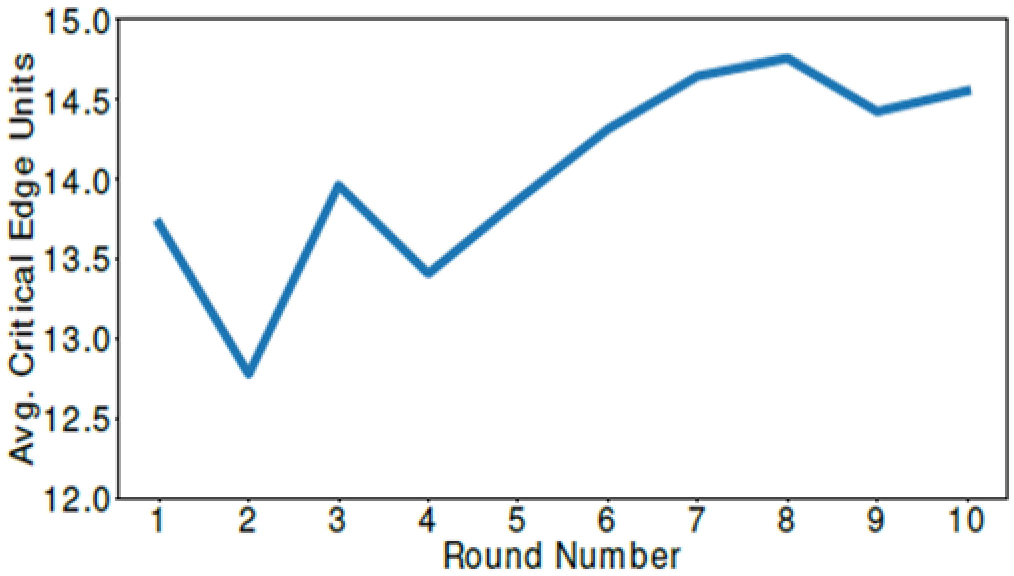
Human Subject Experiments

A) Probability Weighting Bias

- **24%** of the subjects make rational decisions
- **76%** of the subjects are behavioral



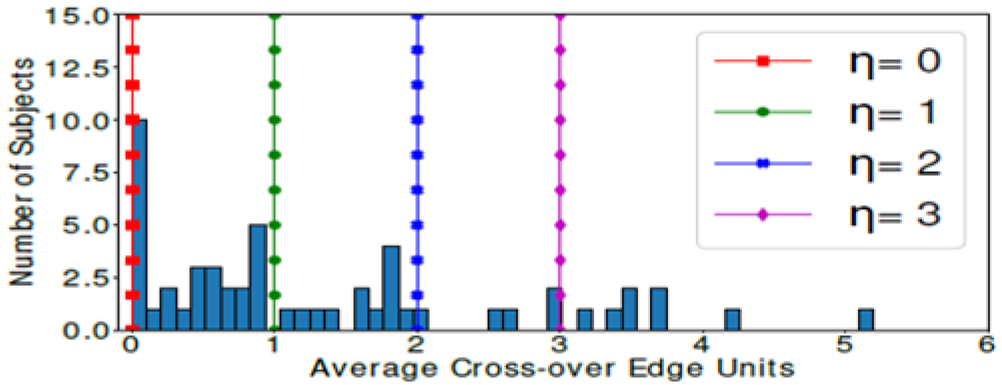
- **20.45%** make worse decisions in later rounds,
- **45.45%** exhibit no learning across rounds,
- **34.10%** improve their investments.

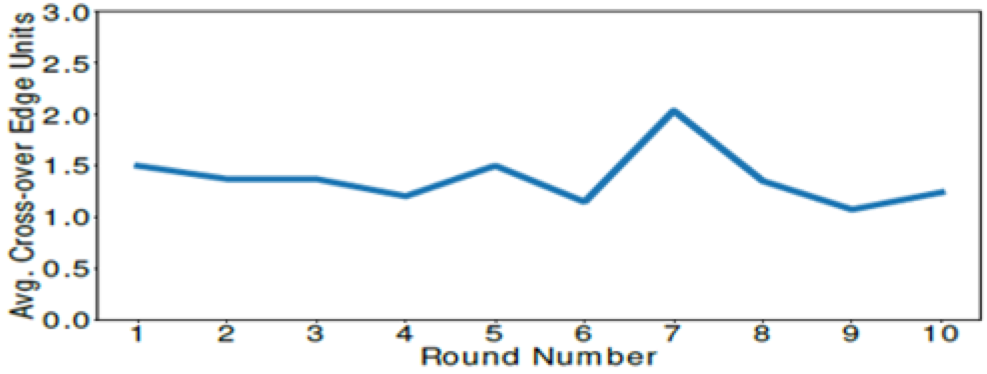


Human Subject Experiments

B) Spreading Heuristics Bias

18.5% of the subjects are non-spreaders
81.5% of the subjects are spreaders
Weak downward trend across rounds





- Experiments motivated a new bias parameter (**spreading level η**), which shows that **human tends to spread the budget even over the edges that does not affect the loss.**
- **In sum**, Human subject Experiments **validated our results about sub-optimal investments made by human security decision-makers.**

Conclusion

- Proposed a **game-theoretic framework** involving attack graph models of large-scale interdependent systems and multiple **behavioral** defenders.
- Proposed different **learning modules** for enhancing decision-making.
  - **Learning History**: Predict chosen attack paths over time.
  - **Reinforcement Learning:** Learn rational behavior over time.
- Evaluated our system via **five interdependent systems** with real attack paths.
- Human experiments **validated** our predictions.

Thank you

Questions!