Modeling of Multi-Defender Collaboration in a Cyber-Security Scenario

Yinuo Du (yinuod@andrew.cmu.edu)

Carnegie Mellon University
Palvi Aggarwal (palvia@utep.edu)
The University of Texas at El Paso
Kuldeep Singh (palvia@utep.edu)
The University of Texas at El Paso
Cleotilde Gonzalez (palvia@utep.edu)
Carnegie Mellon University

Abstract

While evidence shows that cyber attackers are good at coordinating and collaborating in their attacks, network defenders are notoriously poor at sharing information and collaborating among themselves. To help promote cooperation among defenders, one requires models that can explain and make predictions of emergent cooperation decisions of each defender in a cyber security scenario. We propose a Multi-Agent Instance-Based Learning (MDIBL-PD) cognitive model based on Instance-based Learning (IBL) theory, and founded on the Prisoner's Dilemma (PD) of cooperation. MDIBL-PD aims at explaining how collaborations emerge to share information with other defenders in a group. MDIBL-PD was created to interact in a Multi-Defender-Game (MDG) that was used in an experimental study with human participants, intended to determine the effect of different levels of information sharing on collaboration. MDIBL-PD uses an extension of the utility function in IBL theory to capture the emergence of cooperation with higher levels of social information. Through simulations with MDIBL-PD we collect synthetic data to compare to the data set collected in human studies. Our results help explain the emergence of cooperation at increasing levels of information regarding others' actions. We demonstrate the ability of MDIBL-PD to predict human cooperation decisions in the MDG in situations in which players have only their own information and in situations in which they have information about the sharing behavior of the other players.

Keywords: Cognitive Modeling; Multi-agent; Cooperation; Prisoners' dilemma; Cyber-Security

Introduction

In cybersecurity a major problem is the collaboration and coordination among defenders to share information on their vulnerabilities and experienced attacks. Sharing this information brings a major concern for companies and organizations: their privacy and competitive advantage can be damaged if other ill-intentioned people can take advantage of such information for their own benefit. In other words, organizations experience a social dilemma, in which there is a benefit to sharing information, but also put privacy at risk.

Singh, Aggarwal, and Gonzalez (2021) studied this social dilemma in cybersecurity using a Multi-Defender-Game (MDG) in human experiments, to learn about the conditions under which humans share information. MDG is a dynamic game in which sharing information may influence their future security and attack probability. Their experimental results demonstrated a decreasing trend of the average proportion of group-level sharing. Human participants also tended to share less after being attacked, suggesting that instead of making sharing decisions solely based on reciprocity to their groupmates, participants may also base their decisions on the breach status, and might erroneously attribute the breach loss to groupmates.

As suggested by the Hierarchy of Social Information (HSI) in Gonzalez and Martin (2011), an increase in cooperation can be promoted by additional levels of information regarding the other players' actions and outcomes. Thus, knowledge about others' actions and outcomes might make the associations of reciprocity more clear and direct. The similarity of other's predicament to one's own can help strengthen a sense of reciprocity and thus lead to greater cooperation. The HSI proposed an increased level of social information from having no information about the others to an increased level of descriptive social information, where increased information about the complete interaction structure may result in more effective promotion of cooperation. Gonzalez and Martin (2011) argued that ongoing visibility of the payoff matrix can assist in clarifying the trade-off between short-term and long-term rewards. The cognitive modeling work in (Gonzalez, Ben-Asher, Martin, & Dutt, 2015) also suggests that humans tend to consider the outcome of their opponent, dynamically weighted by their interaction experience.

In cognitive science, most models focus on the individual behaviors. Many models aim at studying the cognitive processes of the attacker in order to inform the defense strategies (e.g., masking Aggarwal, Thakoor, et al., 2022; signaling Cranford et al., 2021; anti-phishing Singh, Aggarwal, Rajivan, & Gonzalez, 2020). Other models describe the recognition and comprehension processes of an individual defender (Dutt, Ahn, & Gonzalez, 2011) or the interaction between attacker and defender (Aggarwal, Moisan, Gonzalez, & Dutt, 2022). However, there's a lack of cognitive modeling for groups of defenders in the context of cybersecurity.

Mermoud, Keupp, Huguenin, Palmié, and Percia David

(2019) proposed a behavioral framework that theorizes the association between human behavior and their frequency and intensity to participate in security information sharing. However, their analysis focused on the individuals rather than the interaction among them. A recent review by (Ask, Lugo, Knox, & Sütterlin, 2021) suggests that research on cyber threat communication are mostly interview-based exploratory studies and focused more on individual-organization interaction and internal collaboration (Ahrend, Jirotka, & Jones, 2016; Hámornik & Krasznay, 2017).

In what follows, we first describe the Multi-Defender-Game (MDG) paradigm that reveals the dynamics of defenders' sharing tendency in groups of three over the course of 50 trials. We then formalize a cognitive model of a defender, built in SpeedyIBL (Gonzalez, Lerch, & Lebiere, 2003; Nguyen, Phan, & Gonzalez, 2021). Using the data set collected in a human experiment, we demonstrate that cognitive models of defenders can be useful for understanding the factors affecting the continuation and break down of collaboration and how humans account for the outcome of others.

Multi-Defender-Game (MDG)

We have developed a Multi-Defender-Game (MDG) for data collection through human experiments. The MDG is designed for group experiments. In the MDG there is a group of defenders (human participants) that play an information sharing game in a cyber-security scenario. The participants are assigned in groups of three players, in which they will be identified as defenders Defender 1, Defender 2, and Defender 3, each of them defending their own network. Initially, each defender receive 1000 points as an endowment, which can be used to invest in security to defend their network. Each defender's network is independent, some defenders may be attacked when the others are not and each may have a different chance of being attacked. Then defenders start the game and play 50 trials of decision making on sharing/not sharing information with other defender in the network. The goal of each defender is to maximize their points in the game.

In each trial t, the defender's network may or may not get attacked determined by his *Probability of Breach* Pb^t. If the defenders' network gets attacked, then it costs them -30 points (*attack status* $C_a^t = 1$). They need to choose to share or not to share information with other defenders in the network about the attack/not attack. They will then receive feedback information after the other two group members make their decisions.

The cost of information sharing (-15 points) is deducted from the available points if defenders choose to share information with others. The defender (receiver) gets rewarded (35 points) for receiving information from each other defender. Collectively, the sharing interaction between two defenders forms a prisoner's dilemma (table.1). For example, the payoff in the share-share cell is $20 = 35 - 15$ for both the column player and the row player. *Sharing points* Z_i^t of defender i at trial t is the sum of receiving reward and sharing cost with the other two defenders in their group. The accumulated reward of player i at trial t of defender i is given by Eq.1. We assume the information shared is valuable and it helps the receiver to strengthen their security, thus information sharing also affects future probability of breach by Eq.2.

$$R_i^t = R_i^{t-1} + Z_i^t + (-30) \cdot C_a^t \tag{1}$$

$$Pb^{t+1} = Pb^t - \frac{0.95 \cdot Z_i^t}{2000} \tag{2}$$

Table 1: Payoff matrix

Defender 1 or Defender 2			
		Share	Not-Share
Defender 3	Share	20,20	-15,35
	Not-Share	35,-15	0,0

Human Dataset

As a baseline to compare the predictions of our IBL model, we used a data set collected from human participants who played together in groups using the MDG. This study recruited a total of 210 participants (about 46% female) from Amazon Mechanical Turk, to play a game in groups of 3 participants). On successful completion of the experiment, all participants received a base payment of $3 and they could earn up to $1.75 as additional bonuses based on the points available at the end. The average time taken to complete the experiment was 15 minutes.

The data set consists of two experimental conditions defined based on the information given to the participants regarding the sharing information of the other defenders in their group. The information levels were: *Own* and *Others*, where the *Own* condition provided only information on the actions of the other defenders in the group; while the *Others* condition also provided the outcomes of others and their breach status. Participants received this information in table 2 where the sharing decisions of each defender in the group, including the protagonist defender, were displayed in a separate column. The table also included their breach status, when this information was shared by the other defenders in the group. A total of 102 participants (34 groups) were in the *Own* condition, and 108 participants (36 groups) in the *Others* condition.

Instance-Based Learning Model of Defender's Collaborations

We propose an Instance-Based Learning (IBL) cognitive model to make predictions about human sharing behavior in the MDG, at different levels of information. The model, Multi-Defender IBL - Prisoner's dilemma (MDIBL-PD),

Table 2: An example output table provided as feedback in the *Others* condition of the (Du et al., n.d.) experiment

Defender 3 Decision (Me)	Defender 1 Decision	Defender 2 Decision
Information not shared with Defender 1, Information shared with Defender 2	Defender 1 shared information with me, He was attacked	Defender 2 didn't share any information
	My Payoff with Defender 1 : 35 Defender 1's Payoff with me : -15	My Payoff with Defender 2 : 0 Defender 2's Payoff with me : 0

is based on a model of individual learning and decisions from experience in repeated two-player prisoner's dilemma (Gonzalez et al., 2015), and expands that concept to a multiplayer situation beyond a dyad. Like all IBL models, the MDIBL-PD model relies on the IBL Theory (i.e., IBLT) (Gonzalez et al., 2003), a well-known cognitive theory of experiential decision making. The key idea of this theory is that decisions are made by recognition of similar past experiences, their integration into the generation of expected utility of decision alternatives, and the selection of the alternative with the maximal expected utility. An IBL model can accurately represent the content of human memory, recognition, learning, and recall of experiences in decision making.

The IBLT process and mechanisms are general to every IBL model. These have been published in the past, but we repeat the mathematical formulations of the theory here for completeness. In IBLT, an "instance" is a memory unit that results from the potential alternatives evaluated. These memory representations consist of three elements that are constructed over time: a situation state s that is composed of a set of characteristics f; a decision or action a taken corresponding to an alternative in state s; and an expected utility or experienced outcome x of the action taken in a state. Concretely, for an IBL agent, an option $k = (s, a)$ is defined by the action a in the state s. At time t, assume that there are n_{kt} different instances (k_i, x_{ik_it}) for $i = 1, ..., n_{kt}$, associated with k. Each instance i in memory has an *activation* value, which represents how readily available this information is in memory (Anderson & Lebiere, 1998). Here, the equation captures recency, frequency, similarity, and noise in memory.

$$\Lambda_{ik_it} = \ln\left(\sum_{t'\in T_{ik_it}} (t-t')^{-d}\right) + \alpha \sum_j Sim_j(f_j^k, f_j^{k_i}) + \sigma \ln \frac{1-\xi_{ik_it}}{\xi_{ik_it}}, \tag{3}$$

where d, α and σ are the decay, mismatch penalty, and noise parameters, respectively, and $T_{ik_it} \subset \{0, ..., t-1\}$ is the set of the previous timestamps in which the instance i was observed, f_j^k is the j-th attribute of the state s, and Sim_j is a similarity function associated with the j-th attribute. The rightmost term represents noise to capture individual variation in activation, and ξ_{ik_it} is a random number drawn from a uniform distribution $U(0,1)$ at each step and for each instance and option.

The activation of an instance i is used to determine the probability of retrieving an instance from memory. The probability of an instance i is defined by a soft-max function:

$$P_{ik_it} = \frac{e^{\Lambda_{ik_it}/\tau}}{\sum_{j=1}^{n_{kt}} e^{\Lambda_{jk_jt}/\tau}}, \tag{4}$$

where τ is the Boltzmann constant (i.e., the "temperature") in the Boltzmann distribution. For simplicity, τ is often defined as a function of the same σ used in the activation equation $\tau = \sigma\sqrt{2}$.

The expected utility of option k is calculated based on *Blending* as specified in the choice tasks:

$$V_{kt} = \sum_{i=1}^{n_{kt}} P_{ik_it} x_{ik_it}. \tag{5}$$

The choice rule is to select the option with the maximum blended value.

MDIBL-PD model of Information Sharing

The IBL model of the individual defender is primarily concerned with the learning processes determined by the various levels of information available to the model. We denote the within-group defender index as $x \in \{1, 2, 3\}$ and their sharing decisions as $D_x \in \{C(\text{Share}), D(\text{Not-Share})\}$.

The new MDIBL-PD model was developed for both the own and others information conditions described above. Each IBL agent in the MDIBL-PD model makes decisions using the same procedure defined in the previous section. The human participants in the condition *Others* receive information on the outcome and the breach status of other players (Table 2). To capture this interdependence, we modified the blending equation (Eq.5) to account for the outcome of the other player, as suggested in (Gonzalez et al., 2015).

Actions a: In the MDG, the choice options are defined by the actions that each defender can take. The defender D_x can choose not to share information, to share information with one or both of the other defenders, denoted as None, $D_{(x+1) \mod 3}, D_{(x+2) \mod 3}, Both.$

State S_i^t: The situation state of the defender consists of four attributes: the breach status $A_x \in \{1(\text{attacked}), 0(\text{safe})\}, probabilityofbreach(\text{Pb}_i^t),$ and the expectation of receiving information from each player ($E_{D_x}^t$). Thus, the situation state s of participant i (Defender x) at trial t is $s_i^t = (A_i^t, Pb_i^t, E_{D_{(x+1) \mod 3}}^t, E_{D_{(x+2) \mod 3}}^t)$.

Breach status A_i^t and probability of breach Pb_i^t have direct and indirect affect on the outcome of a trial, thus are included as the context information whose pure appearance might affect human's information sharing tendency. As suggested by (Zhang, Lin, Jing, Feng, & Gu, 2019), beliefs and behavior correlate within rounds in repeated prisoners' dilemma game, and beliefs in one round vary with behavior in the previous round. Thus, we include $E^t_{D_x}$ to capture the association between the expectation of receiving information from peers and the decision of whether to share information with them. It is approximated with the accumulated proportion of receiving information from D_x (Eq.6). Here, we assume that participants can keep track of the interaction experience with their peers. This assumption can be relaxed by manipulating the window of proportion calculation. After receiving the actual sharing decisions at the trial t, the $E^t_{D_x}$ slots will be updated to $T^t_{D_x}$ to store the real interaction experience in memory. When the expectation $E^t_{D_x}$ is closer to 1, memory instances of receiving information from peer x ($T^{t'}_{D_x} = 1, t' \in [0,t)$) have greater similarities to the current situation, resulting in higher activation values (3), and higher likelihood to be recovered (4). Similarly, when the expectation $E^t_{D_x}$ is closer to 0, memories of defected by peer x ($T^{t'}_{D_x} = 0, t' \in [0,t)$) are more likely to be retrieved. The similarity of these numeric attributes is calculated linearly and normalized to $[0,1]$.

$$E^t_{D_x} = \frac{\sum_{i=0}^{t-1} T^i_{D_x}}{t-1} \tag{6}$$

Utility U^t_x: Depending on the experimental condition, the players in the MDG received only information on their own actions (Own) or about the sharing decisions of other defenders and the effect on their outcome of themselves (Others). Therefore, the utility of the defender x in the trial t is the points gained or lost exclusively at that trial, constituted with the benefit of receiving information (35 points), the cost of sharing information (-15 points) and the cost of being attacked (Eq.7). The cost-benefit of information sharing forms the dyadic prisoner's dilemma as shown in Table 1. The cost of the breach is included as part of the utility, since the status of the breach has an effect on the sharing decisions of human defenders.

$$U^t_x = \Delta^t_x = Z^t_x + (-30) \cdot A^t_x \tag{7}$$

$$U^t_x = \Delta^t_x + w^t_1 \cdot \Delta^t_{(x+1) \mod 3} + w^t_2 \cdot \Delta^t_{(x+2) \mod 3} \tag{8}$$

$$w^t_1 = \frac{1 - Surprise^t_1}{2} \tag{9}$$

$$w^t_2 = \frac{1 - Surprise^t_2}{2} \tag{10}$$

To simulate how humans account for the outcome of others, the utility for the blended value calculations is set as the weighted sum of the point update of the defender D_x and his peers (Eq.8). Inspired by the notion of *Social value orientation (SVO)* (Balliet, Parks, & Joireman, 2009), w represents the degree to which a player is willing to consider the outcome of the other player for each option when making a decision that maximizes the gains in each trial.

Research in (Gonzalez et al., 2015) finds that the dynamic w dependent on individual experiences can best explain human cooperation behavior. Under this hypothesis, a player will account for the outcome of the opponent as a function of a normalized gap between expected and actual outcomes (surprise). The value of w^t_i (with respect to the opponent's outcome in the trial t) will be reduced by surprise (Eq.9 and Eq.10). We assume that the players evaluate the benefit of sharing information with each other independently with different weights, updated according to separate *surprises* and *gaps*.

The normalization of *surprises* limited the value of $Surprise^t_i$ within the range of $[0,1]$, the value of w^t_i within $[0,0.5]$, and the sum of weights on the benefit of *others* within $[0,1]$. This formulation assumes that the way a player accounts for the opponent's outcomes will vary between *extreme selfish* when $w^t_1 = w^t_2 = 0$ and extreme fairness when $w^t_1 = w^t_2 = 0.5$.

$$Surprise^t_i = \frac{Gap^t_i}{[Mean(Gap^t_i) + Gap^t_i]} \tag{11}$$

$$Gap^t_i = Abs(V^{t-1}_j - (X_{ij} + O_{ij})) \tag{12}$$

$$Mean(Gap^t_i) = Mean(Gap^{t-1}_i)(1 - \frac{1}{50}) + Gap^t_i(\frac{1}{50}) \tag{13}$$

Pre-Population: From human data, we observed that more than 70% of the human participants chose to share information with both peers at the beginning. (Andreoni & Miller, 1993) show that some fraction of the population actually has altruistic motives. This ingrained tendency to share between human subjects can be the consequence of the experience of cooperation in recent years, or it could be an experimental effect of human participants who expected to cooperate in a *Multi-Defender Collaboration Game*. To capture this preference, and inspired by the conclusion in (Kelley & Stahelski, 1970) that there are two stable types of individuals that can be described as cooperative and competitive, we prepopulate the IBL agents with instances that represent these initial tendencies. 70% of IBL agents are prepopulated with *Share* instances with positive rewards (0, 20, 40 for zero, one, two *sharing - receiving* with peers), while 30% of IBL agents are prepopulated with *Not-share* instances with negative rewards (0, -15, -30 for zero, one, two *sharing - not receiving* with peers). Cooperatively biased agents and defectively biased agents are randomly formed groups of three. Each group contains random number (0 to 3) of cooperatively biased agents. The assumption is that the decrease in the proportion of information sharing is caused by the pairing of *cooperative* participants with *defective* participants.

Simulation Procedure: The MDIBL-PD model with default parameters was run for 100 simulated groups of players in each of the two information conditions. Each group plays the game for 50 trials. The utility assignment for *Own* condition follows Eq.7. The utility for *Others* condition follows Eq.8 with w_1, w_2 defined by Eq.9 and Eq.10.

Dependent Measures: We calculate the overall proportion of sharing in *Own* and *Others* conditions, the proportion of sharing with *Both*, *One*, or *None* of the other defenders, and the sequential dependencies that emerged from the interaction between IBL agents in a group (Martin, Gonzalez, Juvina, & Lebiere, 2014). Sequential dependencies measures include: *Mistrust*, the decision a player makes to defect at time t, after both players mutually defected at time $t-1$; *Forgiveness (Not Share - Share)*, the decision to continue cooperating at time t, although mutual cooperation was not achieved due to the defection of the other at time $t-1$; *Abuse (Share - Not Share)*, the decision to continue defecting at time t after a profitable defection at $t-1$; and *Trust*, the decision to continue cooperating at time t, after successful mutual cooperation at time $t-1$. To assess the precision of the predictions of the model with respect to human data, we calculated the mean squared deviation (MSD) using the average of the dependent measure (e.g., the average proportion of cooperation per trial) and using the Pearson correlation coefficient (r) to assess the similarity of time trends between the model and human data.

Results

Overall Information Sharing

Figure 1 illustrates the proportion of sharing for the MDIBL-PD model compared to human data in the conditions *Own* and *Others* conditions over the course of 50 trials. The proportion of sharing in human data is higher in the *Others* condition (Mean=0.74, SD=0.44) than in the *Own* condition(Mean=0.59, SD=0.49). As shown in Fig.1, the MDIBL-PD model captures these observed trends very accurately. The MSD between human data and model data in *Own* condition is 0.0029, with $r = .86, p < 0.001$. The MSD in Others condition is 0.0022, with $r = .76, p < 0.001$.

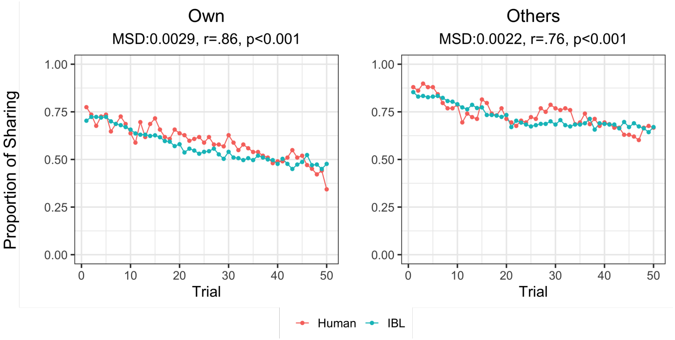

Figure 1: Overtime Sharing Proportion for the *Own* condition (left panel) and the *Others* condition (right panel)

Proportion of sharing with None, One or Both

Figure 2-Top panel, represents the proportion of information sharing with both one and none of the other players in the Own condition. More than 70% human participants choose to share with *Both* peers at the beginning. The proportion decreases over time, and some participants shift to sharing with *One* of the peers, and more participants choose to share with *None*. Most importantly, in the *Own* condition, where participants only receive feedback about their own actions and outcomes, the proportion of sharing with none of the other players increases over the 50 trials.

The model is able to approximate the trends of three types of options accurately. As shown in Fig.2, the deviation between human and model in the proportion of sharing with *Both*, *One*, and *None* is trivial, especially for the *None* option with ($MSD = 0.0029, r = .86, p < 0.001$). We note that the model seems to show a stronger preference for sharing with *One*, while human participants share more with *Both*. A possible explanation is that a fraction of human participants are altruistic or are trying to build an altruistic reputation by indiscriminately sharing with *Both*. The model's decisions, driven by the utility exclusively, converge relatively quickly to the more rewarding options, i.e., sharing with the more reciprocal peer.

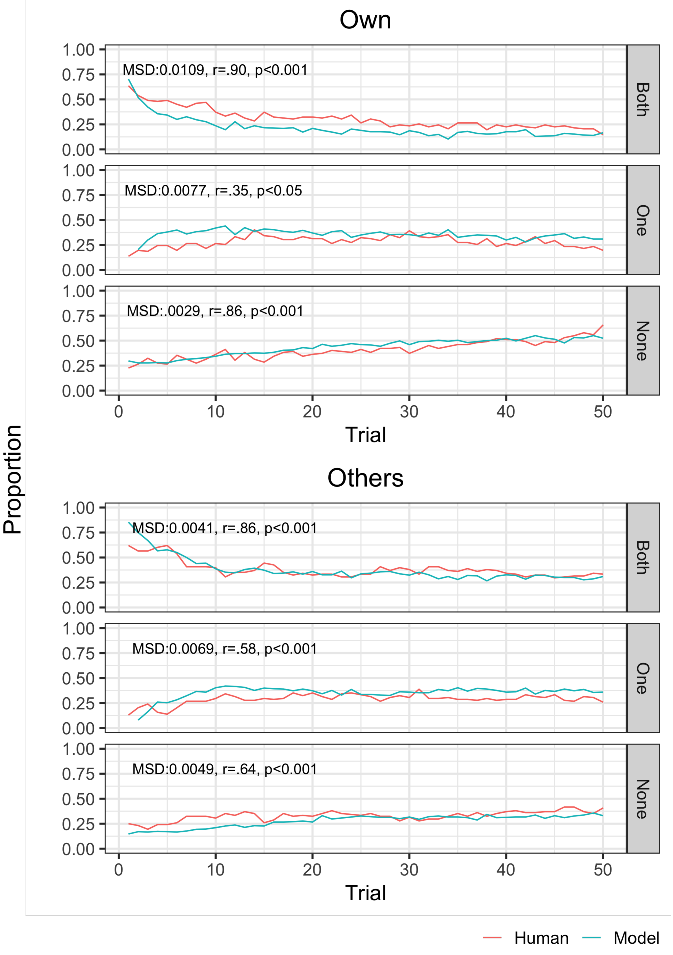

Figure 2: Sharing proportions with Both, One, or None of the other players for the *Own* condition (top panel) and the *Others* condition (bottom panel)

Figure 2-Bottom panel, represents the proportion of information sharing with both, one and none of the other players in the Others condition. The model can account for the dynamics of choosing three types of option (*Both*: $MSD = 0.0.0041, r = .86, P < 0.001$, *One*: $MSD = 0.0069, r = .58, p < 0.001$, *None*: $MSD = 0.0049, r = .64, p < 0.001$). Similar to *Own* condition, human participants demonstrate an initial preference to share with *Both* other players. Although still increasing, the upward trend of sharing with *None* is more flat, indicating that the information of the actions and results of others is effective in maintaining cooperation.

Sequential Dependencies

Fig.3-Left panels demonstrate the comparison between human and model in terms of sequential dependency metrics in *Own* condition. The model fits *Mistrust*, *Trust*, and *Forgiveness* reasonably well with a significant positive correlation with human data (*Trust*: $MSD = 0.0203, r = .55, P < 0.001$, *Mistrust*: $MSD = 0.0052, r = .92, p < 0.001$, *Forgiveness*: $MSD = 0.0207, r = .83, p < 0.001$), but exhibits approximately 25% more *Abuse* than human players ($MSD = 0.0708, r = .16, p > 0.05$).

Similarly, Fig.3-Right panels show that the model matches human behavior for the *Others* condition in terms of *Mistrust* ($MSD = 0.0158, r = .85, p < 0.001$) and *Forgiveness* ($MSD = 0.0404, r = .80, p < 0.001$), but deviates on *Trust* ($MSD = 0.0291, r = .12, p > 0.05$) and *Abuse* ($MSD = 0.0583, r = .36, p > 0.05$). The model is still more likely to *Abuse* and *Forgive* than humans. *Defect* is getting increasingly rewarding as the game progresses, and it becomes more affordable to lose a cooperator.

Discussion

In this paper, we propose a cognitive model that represents the dynamics of cooperation among defenders in a multi-defender game. The MDIBL-PD model builds on and advances the model proposed in (Gonzalez et al., 2015) for a dyad playing the PD game. The model proposes that direct information on the actions of others, whether they share or not with the own player, will influence the emergence of cooperation in the group. The outcomes of the other players in the group are used by each player to make their own decisions. However, the outcomes of the other players are only considered to a certain extent (i.e., "w"). The main insight from (Gonzalez et al., 2015) is that such "w" is dynamic and depends greatly on how the other players behave with the own player in each round of the game. That is, the regard that the self gives to others depends on the dynamic behaviors of others. This idea was used in the MDIBL-PD model and simulation results were produced to replicate the conditions of an experiment carried out with human data.

The results demonstrate that the model performed similarly to the actions taken by humans. First, with more information on Others, individuals share information more often in the MDG. Second, humans tend to decrease the proportion of sharing with both players and increase the proportion of their no-sharing behavior over time. This happens particularly in the Own condition. There are also some differences between the model's predictions and human data. For example, in the Own condition, the model initially tends to share more with one of the other players. The model also shows a higher proportion of "abuse" of the other players, defined as the proportion of defections (not sharing) the model makes after the other player has cooperated (shared). It seems that the model is more "selfish" than humans are regardless of the level of information, as clearly the level of abuse in the model is higher than that of human participants.

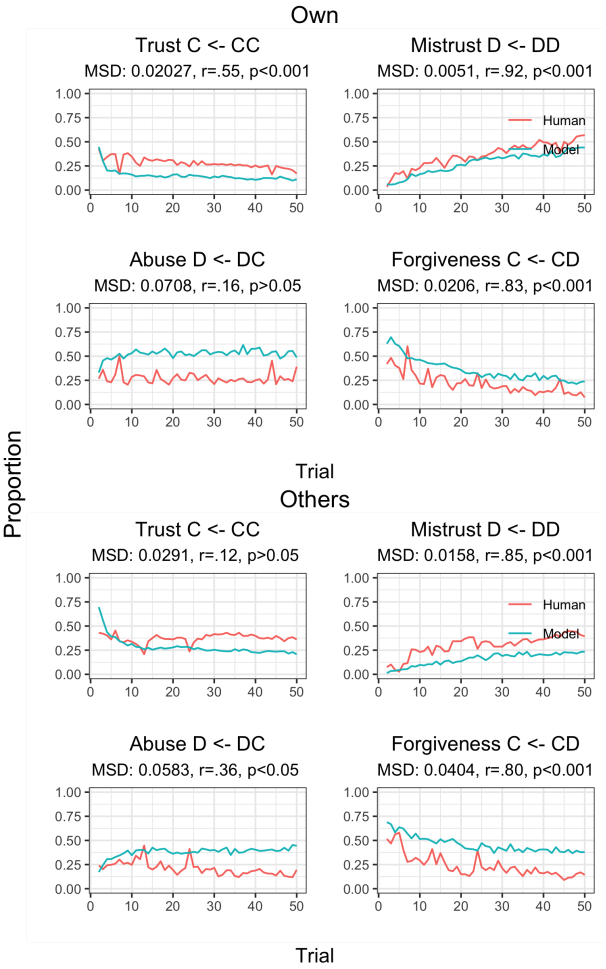

Figure 3: Sequential dependencies in the *Own* condition (top panel) and the *Others* condition (bottom panel), showing Trust, Mistrust, Abuse, and Forgiveness behaviors of the model and human participants

Sequential dependencies also indicate that humans have difficulty sharing information with other players, increasing the level of mistrust of other players over time. This pattern is particularly strong in the Own condition, and the model replicates such trends.

Future research will explore more of how to account for others' decisions while making decisions, for example the surprise and *w* values to explain human behavior. We will also look at the triads in more detail and see the proportion of sharing with each of the two other players.

Acknowledgements

This research was sponsored by the U.S. Army Combat Capabilities Development Command Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-13-2-0045 (ARL Cyber Security CRA). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Combat Capabilities Development Command Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

Aggarwal, P., Moisan, F., Gonzalez, C., & Dutt, V. (2022). Learning about the effects of alert uncertainty in attack and defend decisions via cognitive modeling. *Human Factors*, *64*(2), 343–358.

Aggarwal, P., Thakoor, O., Jabbari, S., Cranford, E. A., Lebiere, C., Tambe, M., & Gonzalez, C. (2022). Designing effective masking strategies for cyberdefense through human experimentation and cognitive models. *Computers & Security*, *117*, 102671.

Ahrend, J. M., Jirotka, M., & Jones, K. (2016). On the collaborative practices of cyber threat intelligence analysts to develop and utilize tacit threat and defence knowledge. In *2016 international conference on cyber situational awareness, data analytics and assessment (cybersa)* (pp. 1–10).

Anderson, J. R., & Lebiere, C. J. (1998). *The atomic components of thought* (J. R. Anderson & C. J. Lebiere, Eds.). New York: Psychology Press. doi: https://doi.org/10.4324/9781315805696

Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The economic journal*, *103*(418), 570–585.

Ask, T. F., Lugo, R. G., Knox, B. J., & Sütterlin, S. (2021). Human-human communication in cyber threat situations: A systematic review. In *International conference on human-computer interaction* (pp. 21–43).

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations*, *12*(4), 533–547.

Cranford, E. A., Gonzalez, C., Aggarwal, P., Tambe, M., Cooney, S., & Lebiere, C. (2021). Towards a cognitive theory of cyber deception. *Cognitive Science*, *45*(7), e13013.

Du, Y., Aggarwal, P., Kuldeep, S., & Gonzalez, C. (n.d.). Multi-defender collaborations in a cyber-security scenario. *Human Factors*.

Dutt, V., Ahn, Y.-S., & Gonzalez, C. (2011). Cyber situation awareness: Modeling the security analyst in a cyber-attack scenario through instance-based learning. In *Ifip annual conference on data and applications security and privacy* (pp. 280–292).

Gonzalez, C., Ben-Asher, N., Martin, J. M., & Dutt, V. (2015). A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive science*, *39*(3), 457–495.

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

Gonzalez, C., & Martin, J. M. (2011). Scaling up instance-based learning theory to account for social interactions. *Negotiation and Conflict Management Research*, *4*(2), 110–128.

Hámornik, B. P., & Krasznay, C. (2017). A team-level perspective of human factors in cyber security: security operations centers. In *International conference on applied human factors and ergonomics* (pp. 224–236).

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of personality and social psychology*, *16*(1), 66.

Martin, J. M., Gonzalez, C., Juvina, I., & Lebiere, C. (2014). A description–experience gap in social interactions: Information about interdependence and its effects on cooperation. *Journal of Behavioral Decision Making*, *27*(4), 349–362.

Mermoud, A., Keupp, M. M., Huguenin, K., Palmié, M., & Percia David, D. (2019). To share or not to share: a behavioral perspective on human participation in security information sharing. *Journal of Cybersecurity*, *5*(1), tyz006.

Nguyen, T. N., Phan, D. N., & Gonzalez, C. (2021). Speedyibl: A solution to the curse of exponential growth in instance-based learning models of decisions from experience. *CoRR*, *abs/2111.10268*. Retrieved from `https://arxiv.org/abs/2111.10268`

Singh, K., Aggarwal, P., & Gonzalez, C. (2021, 06). A social dilemma for cybersecurity: Sharing information among defenders..

Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez, C. (2020). What makes phishing emails hard for humans to detect? In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 64, pp. 431–435).

Zhang, D., Lin, Y., Jing, Y., Feng, C., & Gu, R. (2019). The dynamics of belief updating in human cooperation: findings from inter-brain erp hyperscanning. *NeuroImage*, *198*, 1–12.