Game-theoretic Model of Trust to Infer Human's Observation Strategy of Robot Behavior

Sailik Sengupta[§†]
Amazon AI
sailiks@amazon.com

Zahra Zahedi[§]
Arizona State University
zzahedi@asu.edu

Subbarao Kambhampati
Arizona State University
rao@asu.edu

***Abstract*—We consider scenarios where a worker robot has incentive to deviate from a preferred plan when a human supervisor is not monitoring it. We show that in such scenarios, via human subject evaluation, human supervisors choose sub-optimal observation strategies. To address this, we first consider a game-theoretic framework of trust to formally model such interaction. Then, we leverage this model to infer an optimal supervision strategy that does not need the human to place their trust on the robot. Using a task-planning domain example, we showcase the efficacy of our inferred policies.**

I. Introduction

We consider a multi-agent scenario where a robot R makes and executes a plan and a human supervisor H is held accountable for the robot's behavior. In settings where R can deviate from the supervisor's expectation, a notion of trust becomes a key factor. While it is possible to develop trust in longitudinal setting [1, 15], in one-off interactions (where no trust exists) conventional wisdom often guides the supervisor to spend all their time in monitoring the robot's behavior to ensure it adheres to their expectations. In this work, we challenge the latter belief and by modeling the interaction in a game theoretic framework, show that H can consider resource-efficient monitoring strategies.

There are cases when a robot's expectation may deviate from its supervisor's expectations? First, a robot may have side-goals that do not align with a supervisor's expectation. For example, an autonomous car ride-sharing (or, in general, robot-as-a-)service may have certain expectations from its supervisor (eg. travel on shortest routes) but may need to adhere to passenger's expectation (eg. avoid hilly roads) that are in conflict with one another. Second, the robot may not be fully aware of the human's expectation of itself. In such scenarios, we formally model the inference problem related to the finding a monitoring strategy for the human supervisor.

Specifically, we present a notion of trust that a human supervisor H places on a worker robot R when H chooses to *not observe* R's plan (or its execution) by modeling the interaction in a game-theoretic framework of trust motivated by [10]. To capture the aforementioned scenarios, we assume the robot is unaware of the human's model of itself M_H^R, but has knowledge about all the possible models $\mathcal{M}_H^R$ (or constraints) the human may have; hence, $M_H^R \in \mathcal{M}_H^R$, M_H^R is not known to R, $\mathcal{M}_H^R$ is. We leverage the game theoretic framework to devise a probabilistic observation strategy for H that ensures (1) R does not deviate away from executing a plan that respects all constrains and in turn, (2) H's saves valuable resources such an monitoring time, effort, etc.

While we propose a novel type of assistance that can assist H on when to supervise R to ensure expected behavior, we explore if such assistance is required by performing human studies. We show that without assistance, humans are either too risk-averse (monitors R most of the time to ensure that R adheres to their expectation) or too risk-taking (minimizes their observation time even if the R deviates from expectation). In either case, assumptions made in earlier works [8, 2], where humans are expected to monitor the robot all the time fails to hold. Thus, it makes sense to analyse the supervision scenario and propose methods to suggest optimal monitoring strategies. Furthermore, answers to subjective questions show that participants prefer such automated assistance.

II. Related Work

Our supervision scenario is situated in a specturm of fully-cooperative settings to fully-adversarial ones. In fully-cooperative settings, researchers argue that the robot should only consider plans that adhere to the human's expectation; then these plans are said to be explicable [16], legible [2], adhering to social norms [7]. The assumption that robots sole-objective is to cater to a single human's expectation (the supervisor) may not be true in our case, and the supervisor's monitoring time may be costly. While some works suggest introducing impreciseness in specification on the human's expectation [3] as a solution, other consider robot producing explanations [14] to soothe the human; neither can guarantee behavior produced by R adheres to human's expectation. Other methods where the supervisor communicates implicit constraints [5], or their preferences [6] may not work in our scenario, as a two-way channel is necessary for the robot to identify conflicting constraints, communicate back to the supervisor and convince H the rational behind their behavior. In fully-adversarial settings, related work seek to find monitoring strategies to catch a perpetrator in physical and cyber defense scenarios [13, 11, 12] by framing the interaction in a game-theoretic manner. While our modeling shares similarities, existing works do not consider a cooperative aspect between the players of the game. This makes

[§]Equal contribution
[†]Work done while at Arizona State University.
Presented at the RSS Workshop on Robotics for People– Perspectives on Interaction, Learning, and Safety, 2021.

	$O_{P,\neg E}$	$O_{\neg P,E}$	NO-OB
π_{pr}	$\color{blue}{-C_P^H(\pi_{pr})-I_P^H(\pi_{pr})}$, $-C_P^R(\pi_{pr})-C_{\tilde{E}}^R(\pi_{pr})-C_{\tilde{G}}^R$	$\color{blue}{-C_E^H(\tilde{\pi}_{pr})-I_E^H(\hat{\pi}_{pr})}$, $-C_P^R(\pi_{pr})-C_E^R(\tilde{\pi}_{pr})-C_{\tilde{G}}^R$	$\color{blue}{-V_I^H(\pi_{pr})}$, $-C_P^R(\pi_{pr})-C_E^R(\pi_{pr})$
π_s	$\color{blue}{-C_P^H(\pi_s)\overbrace{-I_P^H(\pi_s)}^{0}}$, $-C_P^R(\pi_s)-C_E^R(\pi_s)$	$\color{blue}{-C_E^H(\pi_s)\overbrace{-I_E^H(\hat{\pi}_H)}^{0}}$, $-C_P^R(\pi_s)-C_E^R(\pi_s)$	$\color{blue}{\overbrace{-V_I^H(\pi_s)}^{0}}$, $-C_P^R(\pi_s)-C_E^R(\pi_s)$

TABLE I
NORMAL-FORM GAME MATRIX FOR MODELING THE ROBOT-MONITORING SCENARIO. R (H) IS THE ROW (COLUMN) PLAYER.

it difficult to use their framework for longitudinal interaction where repeated interaction can build a sense of trust between H and R. While we do not explore this aspect explicitly, our framework keeps this at its core for future extension.

III. GAME THEORETIC FORMULATION

We formulate a two-player general-sum game between the human supervisor H and the robot R. In this section, we explain the components of this game shown in Table I.

A. Player Actions

R, the row-player, has two pure strategies– plans π_{pr} and π_s– plans that are probably risky (does not adhere to all constraints) and safe (one that does). H, the column player, has three strategies– (1) to only observe the plan made by the robot $O_{P,\neg E}$ and decide whether to let it execute (or not), (2) to only observe the execution $O_{\neg P,E}$ and stop R from executing at any point, and (3) not to monitor (or observe) the robot at all (NO-OB). We make two inherent assumptions in this formulation– (1) the robot cannot switch from a plan (or a policy) it commits to in the planning phase during execution phase and (2) the human only stops the robot from executing the plan if they believe that the robot's plan does not achieve the goal G while satisfying their constraints.

B. Utilities

The values on the top, highlighted in blue, indicates the H's utilities and the ones at the bottom represent R's utilities.

R's Utility Values: The utilities are defined in terms of:

$C_P^R(\pi)$	Cost of making a plan π.
$C_E^R(\pi)$	Cost to robot for executing plan π.
$C_{\tilde{G}}^R$	Penalty of not achieving the goal G.

We denote partial plans as $\hat{\pi}_{pr}$; partial plans arise H aborts execution of a probably risky plan. Note that depending on where the human stops the robot, the cost for the partial plans can be different. $C_{\tilde{G}}^R$ that represents the cost of not achieving the goal. The robustness $r(\in(0,1])$ of a plan represents the fraction of models in $\mathcal{M}_H^R$ where the plan π_{pr} is executable and can be obtained via model counting approaches [9]. As a particular M_H^R is sampled from $\mathcal{M}_H^R$, we model $C_{\tilde{G}}^R$ as a random variable drawn from the Bernoulli distribution s.t. it represents a non-zero penalty if the plan is not robust in a sampled human model ($Pr = 1-r$) or zero otherwise.

The cost incurred by the R when H chooses to observe the plan π_{pr} (before execution) is $C_{\tilde{E}}^R(\pi_{pr}) = \begin{cases} C_E^R(\pi_{pr}) & \text{if } C_{\tilde{G}}^R = 0 \\ 0 & o.w. \end{cases}$. If H chooses to monitor the execution directly, R's utility is $C_E^i(\tilde{\pi}_{pr}) = \begin{cases} C_E^i(\pi_{pr}) & \text{if } C_{\tilde{G}}^i = 0 \quad i \in \{R,H\} \\ C_E^i(\hat{\pi}_{pr}) & o.w. \end{cases}$.For execution costs, it is natural to assume $C_E^R(\pi_{pr}) \leq C_E^R(\pi_s)$, i.e cost of executing the plan that satisfies all constraints is greater than executing a plan that satisfies a sub-set of constraints. Similarly, for planning, coming up with π_{pr} will be easy if the value of r is small while coming up with the plan π_s will take considerably longer. Hence, we also assume $C_P^R(\pi_{pr}) \leq C_P^R(\pi_s)$.

H's Utility Values: The utilities are defined in terms of:

$C_P^H(\pi)$	Cost of observing a plan π.
$C_E^H(\pi)$	Cost of observing the robot's execution of π.
$V_I^H(\pi)$	Cost incurred by the supervisor when R violates a constrain due to lapse in H's monitoring
$I_P^H(\pi)$	Inconvenience to H if R presents a plan π that H cannot allow R to execute (note $I_P^H(\pi_s)=0$).
$I_E^H(\pi)$	Inconvenience to H if R is stopped from executing π(note $I_E^H(\pi_s)=0$).

When R proposed π_{pr}, it is only executable in a sub-set of models in $\mathcal{M}_H^R$. As this sub-set may not contain the human's actual model M_H^R, we need to factor in this uncertainty into $V_I^H(\pi)$, $I_P^H(\pi)$ and $I_E^H(\pi)$. We leverage the robustness value r and the Bernoulli distribution for this purpose. We assume R violating a constraint due to lapse in H's monitoring has the highest penalty for (the supervisor) H; thus,

$$V_I^H(\pi_{pr}) > C_P^H(\pi_{pr}) + I_P^H(\pi_{pr}) \quad (1)$$
$$V_I^H(\pi_{pr}) > C_E^H(\tilde{\pi}_{pr}) + I_E^H(\hat{\pi}_{pr}) \quad (2)$$

We also assume that (1) $C_E^H(\pi) > C_P^H(\pi)$ (the cost of observing and the a plan is less than observing the execution of a plan) and (2) $I_E^H(\hat{\pi}_{pr}) > I_P^H(\pi_{pr})$ (same assumption for the inconvenience caused).

IV. GAME-THEORETIC NOTION OF TRUST

In our game, the amount of trust placed in R increases as the H selects $O_{P,\neg E} < O_{\neg P,E} <$NO-OB. When H selects NO-OB, it exposes itself to a vulnerability– R executes π_{pr} resulting in the high negative reward, V_I^H, for H. On the

other hand, if H chooses $(O_{P,\neg E})$, H has the least amount of risk– even before R can execute, the plan is verified by H. There exists a trade-off due to this notion of trust– monitoring depletes H's resources (time, concentration etc.), but if R cannot be fully trusted, H needs to monitor costs to ensure R adheres to constraints.

The No-Trust Scenario: In this setting, H should never play an action that exposes them to a risk of a high negative utility. If a pure-strategy Nash Equilibrium exists, the players should consider it as neither can deviate to get a better utility [10]. Given we consider a Bayesian game where the rewards represent random variable, the expected utility values need to satisfy the following inequalities for a pure-strategy Nash Equilibrium to exists,

$$(1-r)V_I^H(\pi_{pr}) < C_P^H(\pi_{pr}) + (1-r)I_P^H(\pi_{pr})$$
$$C_P^R(\pi_{pr}) + (1-r)C_{\tilde{G}}^R + rC_E^R(\pi_{pr}) < C_P^R(\pi_s) + C_E^R(\pi_s) \quad (3)$$

If $r = 1$, we can guarantee that $(\pi_{pr}, NO - OB)$ is the Nash equilibrium. But, to reduce costs, $r << 1$ (otherwise, $\pi_{pr} = \pi_s$), leading to the following proposition.

Proposition 1. *The game defined in Table I has no pure strategy Nash Equilibrium where π_{pr} is not executable in some of the models.*

Absence of Pure Strategy Nash Equilibrium: The absence of a pure-strategy Nash eq. makes it difficult to define a human's best course of action in this no-trust setting [10]. Thus, we devise the notion of a trust boundary.

Consider a human chooses the mixed strategy $\vec{q} = [(1-q_E-q_N), q_E, q_N)]^T$ over the actions $O_{P,\neg E}, O_{\neg P,E}$ and NO-OB respectively. In order to ensure that the robot cannot deviate away from making and executing π_s, we have to ensure that the expected utility (U) for the robot given $\vec{q}$ is greater for π_s than for π_{pr}.

$$\begin{aligned} \mathbb{E}_{\vec{q}}[U(\pi_s)] &> \mathbb{E}_{\vec{q}}[U(\pi_{pr})] \Rightarrow \quad (4)\\ r\ -C_P^R(\pi_s) - C_E^R(\pi_s) &> (-C_P^R(\pi_{pr}) - C_{\tilde{G}}^R - C_{\tilde{E}}^R(\pi_{pr})) \\ &\quad \times(1-q_E-q_N) \\ &\quad +(-C_P^R(\pi_{pr}) - C_E^R(\tilde{\pi}_{pr}) - C_{\tilde{G}}^R) \times q_E \\ &\quad +(-C_P^R(\pi_{pr}) - C_E^R(\pi_{pr})) \times q_N \end{aligned}$$

where $\mathbb{E}_{\vec{q}}[U(\pi)]$ denotes the expected utilities. This inequality is linear w.r.t. the variables q_N and q_E. Thus, in the region on one side of the linear boundary, the robot always executes π_s. Thus, we call this linear boundary the ***trust boundary***.

V. Experimental Setup and Evaluation

The aim of this section is to first describe a task-planning scenario in which we can compute the trust boundary and then, perform human subject studies in a simplified version of this supervision scenario. To do so, we initially describe the robot-delivery domain that we will use throughout the section.

A. Robot Delivery Domain

We used a robot delivery domain [8] in which the robot can collect and deliver parcels (that may not be waterproof) or coffee by picking it from the reception desk and taking it to a particular location. The robot in the *PDDL* domain has the following actions: {*pickup, putdown, stack, unstack, move*}.

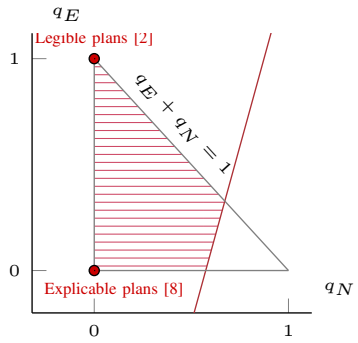

Fig. 1. An observation strategy in the trust region (shaded) ensures that the robot sticks to π_s. In contrast to observation strategies discussed in existing works, one can reduce monitoring costs while ensuring explicable/legible/safe behavior.

Problem Instance: The problem instance in our setting has the initial setting where (1) the robot is standing at a position equidistant to the reception and the kitchen, (2) there is a parcel located at the reception that is intended for the employee, (3) there is brewed coffee in the kitchen that needs to be delivered in a tray to the employee. The goal for the robot is to collect and deliver the coffee and the parcel to the employee.

Robot Plans: There are two plans in which the robot achieves the goal of collecting coffee from the kitchen and parcel from the reception desk and delivers them to an employees' desk. (a) π_s, the robot (1) collects coffee, (2) delivers it to the employee, (3) goes back along the long corridor to collect the parcel from the reception desk and finally (4) delivers it back to the same employee. (b) π_{pr}, the robot collects coffee from the kitchen, (2) collects parcel from the reception desk and puts them on the same tray and finally, (3) delivers both of them to the employee.[§]

B. Computing the Trust Boundary in a Task-Planning Scenario

In order to compute the trust boundary, we calculate the utility values for our game leveraging Table I and the cost incurred by R and H in this robot delivery domain.

In this example, if the robot makes π_{pr}, it will be executable (or safe) as per one of the two observers whose models make up the set $\mathcal{M}_H^R$, because the coffee and parcel taken in the same tray runs the risk of the coffee spilling, thereby ruining the package. Thus, the robustness for π_{pr} is $r = \frac{1}{2} = 0.5$. On the other hand, the plan π_s is executable (and thus, overall safe) in both the models in $\mathcal{M}_H^R$.

We used the Fast Downward planner [4] on the robot delivery domain [8] to find the costs for R and H and to compute them we used a machine with an Intel Xeon CPU (clock speed 3.4 Ghz) and 128GB RAMv. Given the calculated costs, we can define the utility matrix for the players (R, H). [§]

According to Proposition 1, this game does not have a pure Nash Eq. strategy with probability 0.5. Therefore, we now find the boundary in the space of mixed strategies for second type of H who can choose to adopt which will ensure that the robot

[§]Given the (actual and the human's) domain models and the problem instance, these plans can simply be computed using available open-source software like `Fast-Downward` or web-services like `planning.domains`.

[§]The details of each cost and the hasn't been given due to page limit

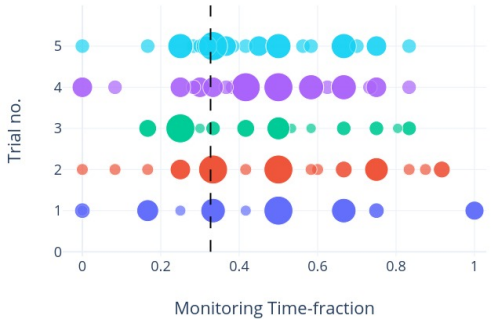

Fig. 2. Participant's monitoring strategies across multiple trials. Trust boundary indicated using the black vertical line.

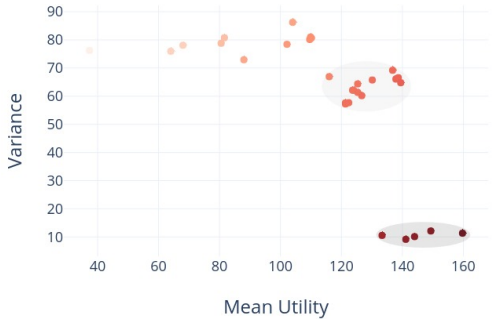

Fig. 3. Average utility and its variance for each of the participants across the five trials.

always executes π_s. To do so, we use the calculated values plug them into equation 4 and obtain,

$$10 \times q_N - 3 \times q_E - 5.74 \quad < \quad 0 \tag{5}$$

In Figure 1, we plot the trust boundary represented by the lines in Eqn. 5. The three black lines (sides of the larger triangle) represent the feasible region for the human's mixed strategy $\vec{q}$. Monitoring strategy in the shaded region guarantees the robot, being a rational agent, executes π_s. The strategy that optimizes H's monitoring cost and yet ensures the robot adheres to π_s lies on the trust boundary indicated using the red line.

C. Human Studies

The human-subjects study was designed to evaluate whether (1) the human can find a good strategy to cut-down the monitoring time while ensuring the constraints structured manner from the robot and (2) the humans tend to deviate to more split-time strategies where some of the time, originally meant for monitoring, can be used for other tasks. We designed a user-interface to represent the robot-delivery scenario. The participants in the study play the role of a student in a robotics department who are asked to monitor the robot for an hour. In order to make the monitoring action be associated with a cost, we added a second task in which participants could choose to grade exam papers (and get paid for it) instead of just monitoring the robot and this represents the action to not-monitor the robot. For simplicity, we combine the actions to monitor the plan and monitor the execution as a single 'monitor the robot' action. We ask them to give us a time slice for which they would choose a particular action (eg. 30 minutes to monitor the robot and 30 minutes to grade exam papers). We let each participant do five trials and after each trial, the overall utility based on the participant's monitoring strategy and the robot's strategy is reported to them. The robot does not adapt itself to the human's strategy in the previous trial (which intents to preserve the non-repeated nature of our game). We collected data from 32 participants who were all graduate students across various engineering departments at our university.

Aggregate Results – Changes in Monitoring Strategy across Trials: Note that a participant, given the information on the interface, can formulate a simplified version of the game-theoretic model proposed in this paper and find the optimal strategy for monitoring (which is to monitor the robot for 0.327 or 19.62 minutes of an hour and use the remaining time to grade papers). The participants' time slice allocated for monitoring, across the five trials, are shown in Fig. 2. Given that there are only two actions for the participant, the strategy can be represented using a single variable (fraction to monitor the robot) and thus, is plotted along the x-axis. The size of each bubble is proportional to the number of participants who selected a particular strategy. The optimal strategy is shown using a black vertical line (i.e. $x = 0.327$). In the first trial, most users ($n = 18$) choose a risk-averse strategy, i.e. monitored the robot to ensure it performs a safe plan even if it meant losing out on money that could be earned from grading. As the trials progressed, participants started discarding extreme strategies (i.e. only monitor or only grade papers) and started considering strategies closer to the optimal. In Fig 2, note that for the first two trials, the strategies are well spread out in the range $[0, 1]$ where as in the last two trials, the strategies are clustered around the optimal decision boundary, with very few data points below 0.25 and very few above 0.7. This shows humans hardly can find an optimal monitoring strategy when there is no prior interaction with the robot and finding an near optimal monitoring strategy after many trial and error can cause a lot of loss. So, a strategy suggestion is needed to provide an assistant to the human to deal with unsafe robots.

Participant Types: In Figure 3, we plot the average utility of each participant across five trials on the x-axis. The y-axis represents the variance. Highlighted in dark, at the bottom right, are five participants that chose observation probabilities in the trust region but not exactly at the trust boundary, i.e. sub-optimal w.r.t. the optimal trust boundary strategy (at 0.327) that yields a reward of 173.77. After that, they did behave in a greedy fashion to reduce the observation time in the hope to make more money by grading papers and stuck to the good policies they initially discovered. Towards the top-right corner, the set of points circled in light gray, we saw a dense cluster of participants ($= 15$) who obtained a high average utility but tried to tweak their strategies significantly, sometimes observing less and therefore, allowing the robot to choose the riskier plan. which eventually lead to a large loss in reward. This implies that the human often takes risk and deviates to more split-time strategies since the time meant to monitoring can be used for other tasks.

VI. Conclusions and Future Work

We model the notion of trust that a human supervisor places on a worker robot by modeling this interaction as a Bayesian Game. We show that existing notions of game-theoretic trust break down in our setting when the worker robot cannot be trusted due to the absence of pure strategy Nash Equilibrium. Thus, we introduce a notion of trust boundary that optimizes the supervisor's monitoring cost while ensuring that the robot workers stick to safe plans.

References

[1] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. Planning with trust for human-robot collaboration. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 307–315, 2018.

[2] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and predictability of robot motion. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pages 301–308. IEEE Press, 2013.

[3] Dylan Hadfield-Menell, Anca Dragan, Pieter Abbeel, and Stuart Russell. The off-switch game. *arXiv preprint arXiv:1611.08219*, 2016.

[4] Malte Helmert. The fast downward planning system. *Journal of Artificial Intelligence Research*, 26:191–246, 2006.

[5] Emmanuel Johnson and Jonathan Gratch. The impact of implicit information exchange in human-agent negotiations. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, pages 1–8, 2020.

[6] Joseph Kim, Christopher Banks, and Julie Shah. Collaborative planning with encoding of users' high-level strategies. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[7] Uwe Köckemann, Federico Pecora, and Lars Karlsson. Grandpa hates robots-interaction constraints for planning in inhabited environments. In *AAAI*, pages 2293–2299, 2014.

[11] Sailik Sengupta, Satya Gautam Vadlamudi, Subbarao Kambhampati, Adam Doupé, Ziming Zhao, Marthony Taguinod, and Gail-Joon Ahn. A game theoretic approach to strategy generation for moving target defense in web

[8] Anagha Kulkarni, Tathagata Chakraborti, Yantian Zha, Satya Gautam Vadlamudi, Yu Zhang, and Subbarao Kambhampati. Explicable robot planning as minimizing distance from expected behavior. *CoRR, abs/1611.05497*, 2016.

[9] Tuan Nguyen, Sarath Sreedharan, and Subbarao Kambhampati. Robust planning with incomplete domain models. *Artificial Intelligence*, 245:134–161, 2017.

[10] Vidyaraman Sankaranarayanan, Madhusudhanan Chandrasekaran, and Shambhu Upadhyaya. Towards modeling trust based decisions: a game theoretic approach. In *European Symposium on Research in Computer Security*, pages 485–500. Springer, 2007.
applications. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 178–186. International Foundation for Autonomous Agents and Multiagent Systems, 2017.

[12] Sailik Sengupta, Ankur Chowdhary, Abdulhakim Sabur, Adel Alshamrani, Dijiang Huang, and Subbarao Kambhampati. A survey of moving target defenses for network security. *IEEE Communications Surveys & Tutorials*, 22 (3):1909–1941, 2020.

[13] Arunesh Sinha, Thanh H Nguyen, Debarun Kar, Matthew Brown, Milind Tambe, and Albert Xin Jiang. From physical security to cybersecurity. *Journal of Cybersecurity*, 1(1):19–35, 2015.

[14] Sarath Sreedharan, Subbarao Kambhampati, et al. Explanations as model reconciliation—a multi-agent perspective. In *2017 AAAI Fall Symposium Series*, 2017.

[15] Anqi Xu and Gregory Dudek. Optimo: Online probabilistic trust inference model for asymmetric human-robot collaborations. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 221–228. IEEE, 2015.

[16] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz Hankui Zhuo, and Subbarao Kambhampati. Plan explicability and predictability for robot task planning. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 1313–1320. IEEE, 2017.