An Initial Study of Targeted Personality Models in the FlipIt Game

Anjon Basak[1]✉[0000−0001−8427−038X], Jakub Černý[2][0000−0002−9407−3782], Marcus Gutierrez[1][0000−0003−0746−5137], Shelby Curtis[1][0000−0001−5599−9462], Charles Kamhoua[3][0000−0003−2169−5975], Daniel Jones[1][0000−0003−4766−4827], Branislav Bošanský[2][0000−0002−3841−9515], and Christopher Kiekintveld[1][0000−0003−0615−9584]

[1] The University of Texas at El Paso, 500 W University Ave, El Paso, TX 79968, USA
abasak@miners.utep.edu, mgutierrez22@miners.utep.edu, srcurtis@miners.utep.edu, jonesdn@gmail.com, cdkiekintveld@utep.edu

[2] Czech Technical University in Prague, Technicka 2, 166 27 Prague 6, Czech Republic
jakub.cerny@agents.fel.cvut.cz, branislav.bosansky@agents.fel.cvut.cz

[3] Army Research Laboratory, 2800 Powder Mill Rd, Adelphi, MD 20783, USA
charles.a.kamhoua.civ@mail.mil

Abstract. Game theory typically assumes rational behavior for solution concepts such as Nash equilibrium. However, this assumption is often violated when human agents are interacting in real-world scenarios, such as cybersecurity. There are different human factors that drive human decision making, and these also vary significantly across individuals leading to substantial individual differences in behavior. Predicting these differences in behavior can help a defender to predict actions of different attacker types to provide better defender strategy tailored towards different attacker types. We conducted an initial study of this idea using a behavioral version of the FlipIt game. We show that there are identifiable differences in behavior among different groups (e.g., individuals with different Dark Triad personality scores), but our initial attempts at capturing these differences using simple known behavioral models does not lead to significantly improved defender strategies. This suggests that richer behavioral models are needed to effectively predict and target strategies in these more complex cybersecurity game.

Keywords: Game theory · Cybersecurity · Extensive-form game · Agent Quantal Response Equilibrium · Dark Triad personality.

1 Introduction

Game theory has a growing number of uses in cybersecurity, such as the strategic allocation of honeypots [20, 12] to learn more about the attacker or to slow the progress of the attacker. There are other examples [22, 5] where the game is dynamic (stochastic or in the extensive form). A common assumption in standard game models is that players are rational and the goal is to seek an optimal strategy in a form of a Nash Equilibrium [16]. However, in many cases, we deploy game strategies against humans or other types of opponents with limited rationality. The literature on behavioral game theory

has started to address the question of developing more predictive models of human behavior, but much of the work to date focuses on very simple games, and it typically ignores the substantial individual differences among humans. In addition, most of this work is not in the context of cybersecurity.

We take a first step towards developing targeted behavioral models that make specific predictions for different groups of human players. This is motivated in part by a long history of work in personality psychology that identifies different dimensions of personality in humans that lead to different behavioral predictions, such as the "Dark Triad" [19] that focuses on malicious behavior types. However, the general idea of developing targeted behavior models can extend beyond personality factors to many other aspects that might influence behavior. We investigate this idea in the context of cybersecurity, conducting a behavioral study for a variant of the FlipIt game.

There are some previous works that consider behavioral modeling of attackers. An initial study on the FlipIt game was done by Nochenson and Grossklags [18] to analyze the impact of participants' age and gender on performance. The authors built a regression model to predict the behavior of the users. Another work by Reitter et al. [21] did analysis on risk propensity and performance of the participants in the FlipIt game and found that high risk affects decision making. The authors also built a cognitive model based on ACT-R [2] which models an individual's risk propensity and decision making strategy.

Another approach is to use Instance Based Learning (IBL) [1] to model the attacker or to support the network administrator by modeling the defender. Another notable line of work is in Stackelberg Security Games (SSGs) where different behavioral models including Prospect Theory (PT) [10], Quantal Response (QR) [15] and Subjective Utility Quantal Response (SUQR) model [17], are used to model the attacker [25, 17, 24]. However, in these works, the game model is a repeated Stackelberg game where the dynamic nature of the real world interaction between an attacker and a defender is not represented in full generality. The attacker model parameters are estimated by considering repeated games where the defender is committed to a mixed strategy in a round rather than committing to a pure strategy. A Bayesian SUQR model has also been proposed where the parameters were fitted for every attack [24]. Kar et al. considered the attacker's attack history to improve the attacker model [11]. The main difference in our work is that we consider consider *targeted* models for different groups of attackers with distinctive behavior patterns.

We consider Extensive Form Games (EFG), a richer representation of the multi-agent interaction than repeated Stackelberg games. The defender and attacker mixed strategy can change in every round depending on previous actions of the players and, as a result, the dynamic nature of the attacker and defender interactions are fully represented. We introduce a variant of an EFG we call a Type-revealing game and a standard Bayesian Game to model dynamic interactions with bounded rational players in cybersecurity settings. Using this model, we evaluated the quality of defender strategies depending on different attacker groups based on different grouping techniques.

We created an online game called *StrataFlip* [3]. We recruited users from Amazon Mechanical Turk (AMT) to play the game as the attacker against the defender, including an equlibrium "strategic" defender. We used the QR behavioral model to fit parameters

to the attackers' behavior using Maximum Likelihood Estimation (MLE), and then we used this model in a variant of an EFG and in a Bayesian Game. To differentiate between different attacker types we used the gameplay of the AMT participants in the StrataFlip game. We also considered the personality of the AMT participants. We studied three personality traits that have been linked to deception and interpersonal harm (Machiavellianism, and subclinical versions of psychopathy and narcissism), which is called the "Dark Triad" [19]. Our initial results show that while there are behavioral differences between the groups, the simple one-parameter QR model appears not to capture them well enough for the defender to effectively target the different groups, so we will need to consider richer behavioral models (such as SUQR) in future work.

2 Background

First, we describe a model of EFGs, a representation of sequential interactions between players. EFGs is a rich representation that is able to represent partial knowledge of the players. Formally, a two-player EFG is defined as a tuple $G = (\mathcal{N}, \mathcal{H}, \mathcal{Z}, \mathcal{A}, u, \mathcal{I})$: $\mathcal{N} = \{d, a\}$ is a set of players, the defender and the attacker. We use i to refer to one of the players, and $-i$ to refer to his opponent. $\mathcal{H}$ denotes a finite set of *nodes* in the game tree. Each node corresponds to a unique *history* of actions taken by all players and chance from the root of the game; hence, we use the term history and node interchangeably. $\mathcal{A}$ denotes the set of all actions. $\mathcal{Z} \subseteq \mathcal{H}$ is the set of all *terminal nodes* of the game. For each $z \in \mathcal{Z}$ we define a *utility function* for each player i ($u_i : \mathcal{Z} \rightarrow \mathbb{R}$). In this work we consider only zero-sum EFGs, for which $u_d = -u_a$.

The imperfect observation of player i is modeled via *information sets* $\mathcal{I}_i$ that form a partition over $h \in \mathcal{H}$ where i chooses an action. We use $A(I_i)$ to denote possible actions available in each node from an information set $I_i \in \mathcal{I}_i$. We assume *perfect recall*, which means that players remember the history of their own. As a consequence, all nodes in any information set I_i have the same history of actions for player i.

Pure strategies Π_i assign one action for each $I \in \mathcal{I}_i$. A *mixed strategy* $\delta_i \in \Delta_i$ is a probability distribution over Π_i. A *best response* of player i to the opponent's strategy δ_{-i} is a strategy $\delta_i^{BR} \in BR_i(\delta_{-i})$, such that δ_i^{BR} is optimal against δ_{-i} according to a given criterion (see Section 3). Strategies in EFGs with perfect recall can be compactly represented by using the sequence form [13]. A *sequence* $\sigma_i \in \Sigma_i$ is an ordered list of actions taken by a single player i in history h. We use $seq_i(I_i)$ and $seq_i(h)$ to denote the sequence of i leading to I_i and h, respectively. A mixed strategy of a player can now be represented as a *realization plan* ($r_i : \Sigma_i \rightarrow \mathbb{R}$). A realization plan for a sequence σ_i is the probability that player i will play σ_i under the assumption that the opponent plays to allow the actions specified in σ_i to be played.

Solution Concepts in EFGs. We provide a formal definition of Nash Equilibrium (NE) and its extension into games with subrational players. We say that a strategy profile $\delta_{NE} = (\delta_1, ..., \delta_n) \in \Delta$ is a *Nash equilibrium* if and only if for each player i it holds that δ_i is a best response to δ_{-i}. NE assumes that the structure of the game is always common knowledge among the players. Now we explain a concept introducing uncertainty in the game being played. The uncertainty is expressed as a probability distribution over the set of possible opponents the player can face. We con-

sider a simplified scenario in which the defender is of a given type, but there might be several types of the attackers: a *Type-revealing game* based on EFG G is a tuple $B_G = (G, n, \overline{p}, \overline{BR}, \overline{u},)$, such that $\overline{p_k}$ is the probability that a defender plays a modified game G with the attacker's utility function $\overline{u_k}$ and best-response function $\overline{BR_k} : \Delta_d \to \Delta_a$, $k \in \{1, \ldots, n\}$.

In the Type-revealing game, the type of the attacker is revealed to the defender after the game begins. In contrast, in a standard *Bayesian game* [7] the type of the attacker is not revealed. An example of a Bayesian game is depicted in Figure 1. The attacker is one of the two possible types: the defender faces the first type with probability 0.7 and the second type with probability 0.3. Since he cannot distinguish the individual types, the former singleton game states are now grouped into information sets. Both types of attackers are purely rational utility maximizing players, but with different utilities. For example, in case the defender decides to play action b_3, the best response of the first type is to play b_8, while the best-response of the second types is b_7.

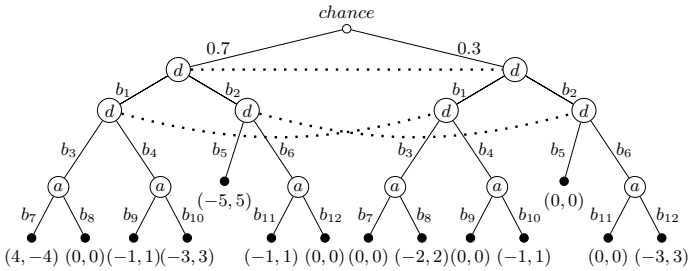

Fig. 1: A Bayesian extension game. Each internal node is labeled by a player who acts in this node. Under every terminal node is a tuple of utilities obtained by the defender and the attacker, respectively. Nodes in the same information set are connected by a dashed line.

Computing Solution Concepts. The baseline approach for computing an exact NE in zero-sum games with imperfect information is via mathematical programming [13, 4]. For computing strategies in large EFGs we use a state-of-the-art algorithm for approximating NE called *Counterfactual Regret Minimization (CFR)* [26]. CFR is an iterative algorithm, which in every iteration updates the strategies of the players in order to minimize a weighted sum of regret at each decision. The average strategies then approach NE. Because the individual attackers choose their strategies according to a specific best-response method, we use a variant of CFR called CFR-BR [8]. In this algorithm, one player updates his strategy using CFR (the defender), while the second player (the attacker) computes his best response against this strategy. The algorithm can be modified for a Type-revealing or Bayesian game by considering a game tree as in Figure 1. It can consider several methods for computing response functions, so that each attacker type can behave according to a different behavioral model.

3 Game Model

The domain we use in this work is a variant of the "FlipIt" game [23]. This game is motivated by a cybersecurity scenario where an attacker can perform a stealthy attack

to gain control of a resource (e.g., install malware on a host or steal a password) that may not be immediately detected by the defender. However, the defender can take actions to restore control to the defender (e.g., performing a virus scan or resetting a password).

A two-player FlipIt game is defined as a tuple $F = (V, t, \rho, \gamma)$. The game is played by a defender and an attacker on an empty graph with nodes V for a finite number of simultaneous rounds t. There is a positive reward $\rho : V \rightarrow \mathbb{R}^+$ and a positive cost $\gamma : V \rightarrow \mathbb{R}^+$associated with each node $v \in V$. At the beginning of the game, the defender controls all nodes. In each round, each player selects one node to flip, i.e to attempt to gain control of. The flipping action is successful when the current owner of the node does not also flip it. For every flipping action, the players pay the cost assigned to the node. At the end of every round the players collect the total rewards from all nodes they control. After t rounds the game ends and the final utilities are the sum of the rewards collected in the individual rounds. We consider the version of the game when after every round the players are provided with the action of the other player. To

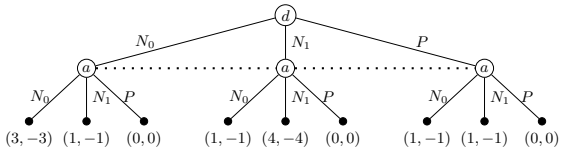

Fig. 2: An example of EFG representation of a FlipIt game with 2 nodes (N_0 and N_1) played for 1 round. The game also contains a pass action (P). The figure follows a standard denotation of an extensive-form game.

represent the FlipIt game, we use the EFG formalism. For example, a representation of a FlipIt game with 2 nodes: N_0 and N_1, played for 1 round is depicted in Figure 2. We assume $\rho(N_0) = 2, \gamma(N_0) = 3, \rho(N_1) = 3$ and $\gamma(N_1) = 4$. Then we define both a Type-revealing and a Bayesian game based on this game with multiple attacker types. Each types is described by its behavioral model (see the next section). A NE in these games are computed using CFR-BR.

4 Models For Attacker Behavior Prediction

The existing attacker models (QR, SUQR, PT) for SSGs consider only one-shot or repeated SSGs with only one generic type of attacker or as many attackers as the number of attacks. However, these assumptions may cause the defender to adopt a suboptimal strategy if the correct model is not used for the attacker. In this work, we model the interaction between a defender and an attacker using a Type-revealing game and a Bayesian game models. We solve the game with CFR-BR where the BR is provided by the attacker. Next, we present a behavioral game theoretic model for modeling attacker behavior to provide best responses (BR) against defender strategy using CFR in our behavioral game based on the EFG model where the defender knows which type of attacker he is facing.

Agent Quantal Response Model. In this work we consider the Agent Quantal Response Equilibrium (AQRE) [15] which is compatible with the behavioral strategy representation to support the sequential interaction between the attacker and the defender. Agent

Quantal Response (AQR) model in an EFG assumes that different information sets of a player are played by different agents. Each agent of each player has an additive payoff disturbance that is added to the continuation payoff for each possible action at that agent's information set. All the agents share the same payoff function. In the model each agent, i simply chooses the maximum of $\hat{u}_{a,I}$ at information set $I \in \mathcal{I}_a$ and acts independently of the other agents of the same player.

For the AQR model there is only one parameter, λ, which has a value from 0 to ∞. When $\lambda = 0$ the model behaves like a pure random agent, and when $\lambda = \infty$ the AQR model converges to the rational best response model. A logit-AQRE is any solution to the set of k equations: one equation for each action in each information set of each agent. In Equation 1, Agent Quantal Best Response $AQBR(r_d, a)$ gives the probability for playing action a for the attacker in information set $I \in \mathcal{I}_a$. Equations 2, 3, and 4 define how to compute attacker's expected utility u_a using defender's realization plan r_d in case h is an inner node. Otherwise, the leaf utility is used.

$$AQBR(\lambda|a, I, r_d) = \frac{e^{\lambda \hat{u}(r_d,a)}}{\sum_{a' \in A(I)} e^{\lambda \hat{u}(r_d,a')}} \ \ \forall I \in \mathcal{I}_a, \forall a \in A(I) \tag{1}$$

$$\hat{u}(r_d, a) = \frac{\sum_{h \in I} r_d(seq_d(h)) u_a(h, a)}{\sum_{h \in I} r_d(seq_d(h))} \ \ \forall I \in \mathcal{I}_a, \forall a \in A(I) \tag{2}$$

$$u_a(h, a) = \sum_{\substack{I' \in \mathcal{I}_a, h' \in I' \\ seq_a(h') = seq_a(h)a}} \frac{u_a(h') r_d(seq_d(h'))}{r_d(seq_d(h))}$$
$$\forall I \in \mathcal{I}_a, \forall h \in I, \forall a \in A(I) \tag{3}$$

$$u_a(h) = \sum_{a \in A(h)} u_i(h, a) AQBR(r_d, a) \ \ \forall I \in \mathcal{I}_a, \forall h \in I \tag{4}$$

5 Parameter Estimation

We now describe how we collected data from the AMT participants using the StrataFlip game so that we can analyze different attacker behavior and fit the parameters of the QR model according to different attacker groups.

StrataFlip Game. The StrataFlip game is based on the description provided in section 3. In this game, two players compete over a network with multiple nodes. A node can be any machine in the network. In our experiment we used six nodes: Node A($10/8$), Node B($10/2$), Node C($4/2$), Node D($4/8$), Node E($10/5$), Node F($0/0$). Each node has a reward (ρ) and a cost (γ). For example, Node A has reward 10 and cost 8. The defender/attacker has to pay the cost each time he wants to defend/capture a node. Each node can be either captured by the attacker (red) or not (blue) as shown in Figure 3.

The game has five rounds. Initially, the defender has control over all of the nodes. The purpose of the attacker/participant is to take over control from the defender by attacking nodes. In each round, the defender defends a node and the attacker attacks a node. If the attacker or defender chooses to attack/defend a node he/she has to pay the

Fig. 3: Game interface of the StrataFlip game

cost associated with that node. If the defender and attacker do not make the same move then the attacker takes control of the node and receives the reward associated with that node and pays the cost. If both players make the same move then the previous controller of the resource retains the control and attacker and defender both pay the cost. In each round, the user interface shows the following information: total points, current round, time, action history (log), and who currently controls each node. After each round the attacker receives points for all the nodes he controls. Red and blue mean the attacker or the defender controls a node, respectively. The attacker is able to observe the effect of the defender action in the next round. Each player tries to maximize their utility by controlling the nodes. The defender uses a pre-computed Nash Equilibrium strategy by formulating the StrataFlip game as a zero-sum EFG (strategic defender) and a random strategy where the actions played by the defender were completely random (random defender). The attackers are the AMT participants. In the next section, we describe how we collected data from AMT using the StrataFlip game to analyze attacker behavior.

AMT Experiment. We recruited 155 participants using AMT. After agreeing to participate the participants filled out the Short Dark Triad (SD3) [9, 14] scale. The participants were given instructions on how to play the StrataFlip game. Next, they answered some comprehension check questions. The participants were not allowed to go ahead until they answered correctly. If they answered incorrectly, proper instructions were given on why the answer is wrong. Then they played a practice game where the participants could make themselves familiar with the StrataFlip game.

Each participant played six StrataFlip games: three against the strategic defender and three against the random defender. Fifty percent of all the participants played against the strategic defender first and later against the random defender and the other fifty percent played the random defender first and then the strategic defender. We collected all the participant's gameplay in each round including temporal data.

Parameter Estimation for AQR Model. We now describe how we estimate the λ parameter of the AQR model which represents the degree of rationality of a player in a certain sense. As λ reaches ∞ the response of the player approaches perfectly rational behavior. We use Maximum Likelihood Estimation (MLE) [6] to estimate the parameter of the AQR model. The idea behind MLE is to find parameter estimates that maximize the probability of seeing what is observed in the data.

Given the defender's realization plan r_d and M samples of the players' choices, the following equation defines the log-likelihood function for a given λ:

$$\log(L(\lambda|r_d)) = \sum_{I \in \mathcal{I}_a} \sum_{a \in A(I_k)} M_a(I) \log(AQBR(\lambda|a, I, r_d)), \tag{5}$$

where $AQBR(\tau_j, I, r_d)$ is the probability of playing action τ_j in sample j in information set $I \in \mathcal{I}_a$ for attacker, and $M_a(I)$ is the number of samples taking action a in information set I. An optimal $\hat{\lambda}$ is then selected as λ maximizing the loglikelihood function. For the computation we used a modified binary search.

5.1 Attacker Grouping

The AQR model described above is used to model different attacker types depending on the parameter values that can be fitted to different attacker types. For example the parameter λ_P in the AQR model can either be fitted to one generic group with the full attacker population P, or we can have $\{\lambda_{p_1}, \lambda_{p_2}, ..., \lambda_{p_q}\}$ fitted to different subgroups where $P = \{p_1, p_2, ...p_q\}$. Detecting the attacker groups can be done in many ways. Next we present how we detect different attacker groups $\{p_1, p_2, ...p_q\}$ among the attacker population P.

Clustering using attacker behaviors (c_{bhv}). Clustering using attacker behaviors (c_{bhv}) is based on the actions played by the AMT participants in the online StrataFlip experiment. For features we used the node value, node cost, and points received for the node attacked in each round. We used k-means and density based clustering to detect the best groupings of attackers $\{p_1, p_2, ...p_q\}$. The best number of possible groups we could find is three $(k = 3)$.

Clustering using DT scores (c_{DT}). Clustering using DT scores (c_{DT}) is based on the three personality scores, Machiavellian score (s_m), Narcissist score (s_n) and Psychopath score (s_p), computed from the Dark Triad survey given to the AMT participants at the beginning of the AMT experiment. There is no straightforward way to differentiate these three personalities since there is no pure Machiavellian or no pure Narcissist or no pure Psychopath. We tried density-based clustering and k-means clustering with $k = 2, 3$ using the three personalities scores as features. We found the best clustering when $k = 3$.

Grouping using DT Maximum score (c_{DTM}). Another way we can find different groups inside a generic group is to use the individual maximum DT score among the three scores: s_m, s_n, s_p. So, we can have three groupings where each group represent each personality. An individual is assigned to a group for which he has the maximum DT score.

6 Experiments

We now show that differences in opponent models and differences in the ways in which these models are integrated into the game model can significantly impact the quality of the strategy of the defender. We consider three cases: a generic subrational attacker population P in an EFG model, different attacker groups (p_1, p_2, p_3) in a Type-revealing

game model and also in a Bayesian game model. We also show the rationality and strategic differences between different attacker groups after separating them using different grouping techniques.

Game Values. In our first experiment, we show defender game values considering different variations of attacker groups using different game models. Game value is the expected utility of a player when following an optimal strategy. The left and the middle graphs in the Figure 4 show the results. In the first column the defender considers a rational attacker, but actually, he faces a subrational attacker (*r-sr* scenario). We assumed λ_P for population P for the actual attacker. In the second column, we show the game value when the defender knows the actual attacker with λ_P for population P (*sr-sr* scenario). The next three columns show game values when the defender considers different groups of attacker (p_1, p_2, p_3) in the variant of EFG model (where he knows the attacker types) and in the Bayesian Game model. We consider three different grouping techniques: c_{bhv}, c_{DT} and c_{DTM}. In Figure 4 (ignore the legends for the first two graphs) we see that the game value is noticeably higher when the defender models a subrational attacker than when he models a rational attacker. That means the defender was able to exploit the attacker strategies rather than being too conservative.

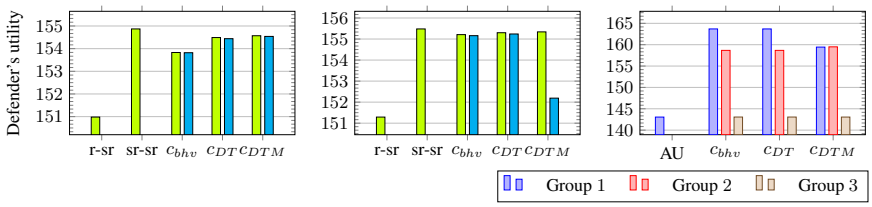

Fig. 4: Apply legends to the rightmost graph only. Defender's expected utility when attacker faced (Left) a strategic defender and (Middle) a random defender in the former EFG (first two columns), in the Type-revealing game (last three results in lime) and in the Bayesian game (last three results in cyan). (Right) Defender's expected utility against different attacker's empirical strategy including the generic attacker (AU) against a strategic defender.

To be sure that different attacker groups have significant strategical differences, we computed the defender's utility against the empirical strategies of different attacker groups. In Figure 4 (rightmost graph), where the defender considered the attackers separately in separate EFGs, we can see that there are significant differences between the defender's expected utilities. That means that the defender can potentially exploit the strategical differences of different attacker groups by using tailored strategies targeted towards specific attacker subgroups. Overall, the results confirm that the defender can explicitly exploit a single subrational opponent. However, such a strategy can be exploitable by the attacker if the model is not precise. When facing a group of several types of the attackers, the strategy of the defender is more conservative since there is a more rational attacker present in the group. The overall quality of the strategy (in terms the expected utility of the defender) in the settings with groups of models of the attackers, however, is comparable to the setting with just a single model of the attacker. This may present a way of exploiting weak attackers while still being able to defend against more rational ones.

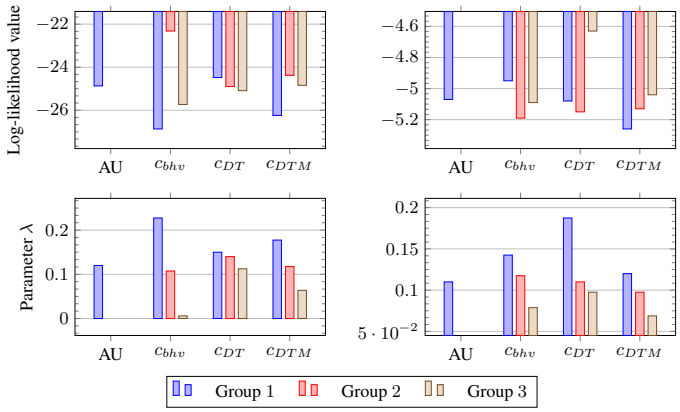

Fig. 5: (Up) Log-likelihood values and (Down) λ parameter values for different grouping techniques compared to one generic group (AU) when attacker faced (Left) a strategic defender and (Right) a random defender.

The MLE technique fits parameters when the Log-Likelihood Value (LLV) is maximized or the negative of the LLV is minimized. In our experiment, we fit the parameter by maximizing the LLV. As we can see in Figure 5, for each of the game instances Log-likelihood value is almost same when we used different grouping techniques instead of considering only one generic group P.

Targeted λ Value. In our next experiment, we show more targeted values of the the AQR model parameter considering different grouping techniques. Figure 5 shows the λ value we found for all the users against the strategic and random defender. The results show that dividing the users into different sub-groups highlight differences between their level of rationality (different values of their λ). We found that the error percentage of the MLE to estimate λ is about 6.7%. However, this may vary depending on the number of data points available.

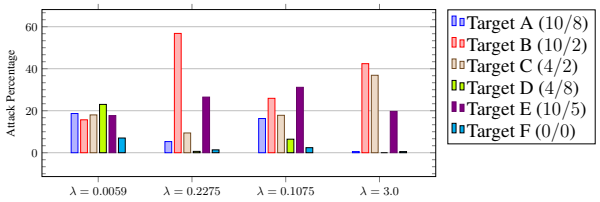

Fig. 6: Attack pattern by different groups of attackers on different targets for c_{bhv}.

Attack pattern. Next, we present Figure 6, which shows the percentage of attacks on different targets by different groups of attackers against a strategic defender when we used c_{bhv} clustering. To put this into perspective we also added the attack percentages of an attacker with $\lambda = 3.0$, which is very close to rational. We can clearly see that the more rational group ($\lambda = .2275$), not the simulated one, made more rational choices by attacking targets with higher rewards and lower costs and avoiding targets with higher costs.

7 Conclusion

Game theory assumes rational agents when computing Nash Equilibrium. However, in a real-world scenario like cybersecurity where humans are involved, this assumption can have consequences if we do not consider human factors into the game model. Our goal is to find different types of attacker based on their personality so that we can incorporate those human factors to identify different attack types and predict their actions to provide better defender strategy against different types of attackers. We show that there are strategic differences between different groups of attackers and the defender can benefit by modeling sub-rational attackers. In our initial analysis we have considered only the quantal-response model, which is well established in the behavioral game theory literature. While this model does capture at least some of the limited rationality of the human players, it appears to be too simple with a single parameter to allow the defender to effectively target different groups of attackers. In future work we plan to investigate other models that allow for more fine-grained predictions such as SUQR. In addition, we plan to investigate ways that a defender can safely exploit (uncertain) knowledge of the type of opponent he is facing when using such behavioral models.

Acknowledgment. This research was sponsored by the 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 Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes not with standing any copyright notation here on. The authors also acknowledge the support of the OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16˙019/000 0765 “Research Center for Informatics”. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme “Projects of Large Research, Development, and Innovations Infrastructures” (CESNET LM2015042), is greatly appreciated.

References

1. Abbasi, Y.D., Short, M., Sinha, A., Sintov, N., Zhang, C., Tambe, M.: Human adversaries in opportunistic crime security games: Evaluating competing bounded rationality models. In: Proceedings of the Third Annual Conference on Advances in Cognitive Systems ACS. p. 2 (2015)
2. Anderson, J.R.: Act: A simple theory of complex cognition. American Psychologist **51**(4), 355 (1996)
3. Basak, A., Shelby, C., Gutierrez, M., Černý, J.: StrataFlip Game. `http://iasrl1.cs.utep.edu/` (2017), [Online;]
4. Bosansky, B., Kiekintveld, C., Lisy, V., Pechoucek, M.: An exact double-oracle algorithm for zero-sum extensive-form games with imperfect information. Journal of Artificial Intelligence Research **51**, 829–866 (2014)
5. Durkota, K., Lisý, V., Kiekintveld, C., Horák, K., Bošanský, B., Pevný, T.: Optimal strategies for detecting data exfiltration by internal and external attackers. In: International Conference on Decision and Game Theory for Security. pp. 171–192. Springer (2017)

6. Friedman, J., Hastie, T., Tibshirani, R.: The elements of statistical learning, vol. 1. Springer series in statistics New York (2001)
7. Harsanyi, J.C.: Games with incomplete information played by "bayesian" players, i–iii part i. the basic model. Management science **14**(3), 159–182 (1967)
8. Johanson, M., Bard, N., Burch, N., Bowling, M.: Finding optimal abstract strategies in extensive-form games. In: AAAI (2012)
9. Jones, D.N., Paulhus, D.L.: Introducing the short dark triad (sd3) a brief measure of dark personality traits. Assessment **21**(1), 28–41 (2014)
10. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. In: Handbook of the fundamentals of financial decision making: Part I, pp. 99–127. World Scientific (2013)
11. Kar, D., Fang, F., Delle Fave, F., Sintov, N., Tambe, M.: A game of thrones: when human behavior models compete in repeated stackelberg security games. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems. pp. 1381–1390. IFAAMAS (2015)
12. Kiekintveld, C., Lisý, V., Píbil, R.: Game-theoretic foundations for the strategic use of honeypots in network security. In: Cyber Warfare, pp. 81–101. Springer (2015)
13. Koller, D., Megiddo, N., von Stengel, B.: Efficient Computation of Equilibria for Extensive two-person Games. Games and Economic Behavior pp. 247–259 (1996)
14. Maples, J.L., Lamkin, J., Miller, J.D.: A test of two brief measures of the dark triad: The dirty dozen and short dark triad. Psychological assessment **26**(1), 326 (2014)
15. McKelvey, R.D., Palfrey, T.R.: Quantal response equilibria for extensive form games. Experimental economics **1**(1), 9–41 (1998)
16. Nash, J.: Non-cooperative games. Annals of mathematics pp. 286–295 (1951)
17. Nguyen, T.H., Yang, R., Azaria, A., Kraus, S., Tambe, M.: Analyzing the effectiveness of adversary modeling in security games. In: AAAI (2013)
18. Nochenson, A., Grossklags, J., et al.: A behavioral investigation of the flipit game. In: Proceedings of the 12th Workshop on the Economics of Information Security (WEIS) (2013)
19. Paulhus, D.L., Williams, K.M.: The dark triad of personality: Narcissism, machiavellianism, and psychopathy. Journal of research in personality **36**(6), 556–563 (2002)
20. Píbil, R., Lisý, V., Kiekintveld, C., Bošanský, B., Pěchouček, M.: Game theoretic model of strategic honeypot selection in computer networks. In: International Conference on Decision and Game Theory for Security. pp. 201–220. Springer (2012)
21. Reitter, D., Grossklags, J., Nochenson, A.: Risk-seeking in a continuous game of timing. In: Proceedings of the 13th International Conference on Cognitive Modeling (ICCM). pp. 397–403 (2013)
22. Shiva, S., Roy, S., Dasgupta, D.: Game theory for cyber security. In: Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research. p. 34. ACM (2010)
23. Van Dijk, M., Juels, A., Oprea, A., Rivest, R.L.: Flipit: The game of "stealthy takeover". Journal of Cryptology **26**(4), 655–713 (2013)
24. Yang, R., Ford, B., Tambe, M., Lemieux, A.: Adaptive resource allocation for wildlife protection against illegal poachers. In: Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. pp. 453–460. IFAAMAS (2014)
25. Yang, R., Kiekintveld, C., Ordonez, F., Tambe, M., John, R.: Improving resource allocation strategy against human adversaries in security games. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence. vol. 22, p. 458 (2011)
26. Zinkevich, M., Johanson, M., Bowling, M., Piccione, C.: Regret minimization in games with incomplete information. In: Platt, J., Koller, D., Singer, Y., Roweis, S. (eds.) Advances in Neural Information Processing Systems 20 (NIPS), pp. 1729–1736. MIT Press, Cambridge, MA (2008)