Empirical Game-Theoretic Methods for Adaptive Cyber-Defense

Michael P. Wellman(✉), Thanh H. Nguyen, and Mason Wright

University of Michigan, Ann Arbor, USA
wellman@umich.edu

Abstract. Game-theoretic applications in cyber-security are often restricted by the need to simplify complex domains to render them amenable to analysis. In the empirical game-theoretic analysis approach, games are modeled by simulation, thus significantly increasing the level of complexity that can be addressed. We survey applications of this approach to scenarios of adaptive cyber-defense, illustrating how the method operates, and assessing its strengths and limitations.

1 Introduction

Strategic analysis of a cyber-security situation starts with the understanding that attacker and defender are engaged in an adversarial interaction, driven by (largely) opposing objectives, and armed with distinct tools for assessing and shaping the cyber environment. Formalizing these elements almost inevitably leads the analyst to describe the situation in *game-theoretic* terms: available actions and observations of the respective actors (players), and utility functions representing objectives. Thus, it is not surprising to observe a large expansion of the literature on game theory applied to cyber-security,[1] and an associated increase in development of tools and applications (Manshaei et al. 2013; Roy et al. 2010; Sinha et al. 2018).

Many game-theoretic treatments of cyber-security domains start with major simplifications, due to the analytic complexity of high-fidelity representations of realistic environments. Analysis of such stylized models can often shed valuable light on a strategic situation. For example, Edwards et al. (2017) employ a coarse-grained "blame game" model to identify qualitative considerations for deciding how and when to attribute responsibility for suspected state-sponsored cyber-attacks. Simplicity in modeling facilitates reasoning and allows a given model to cover a broad class of relevant scenarios. Choosing the right abstractions to isolate exactly the strategic issues of interest is central to the game theorist's art, and when done well, it can provide deep insight for decision makers.

There are two significant drawbacks to the stylized approach, however. First, the models analyzed tend to be generic, and so do not necessarily help for

[1] Including dedicated annual conferences, such as GameSec (Bushnell et al. 2018; Rass et al. 2017a).

S. Jajodia et al. (Eds.): Adaptive Cyber Defense, LNCS 11830, pp. 112–128, 2019.
https://doi.org/10.1007/978-3-030-30719-6_6

determining particular solutions to specific situations. Work in the framework of Stackelberg security games (Tambe 2011) has effectively addressed this issue, by supporting decisions for specified problem instances rather than generic scenarios. Second, for complex scenarios there is danger that the abstractions applied may discard essential detail, and thus the resulting guidance is incomplete, or worse—potentially misleading. Cyber-systems are inherently complex environments, typically involving numerous computationally interacting entities, with considerable state and complicated patterns of communication and observation. Experts familiar with the intricacies of such systems are likely to view stylized game models as toy versions of reality, and thus take a skeptical stance to conclusions from such models.

Since any modeling approach will entail some abstraction of the real world, there is no way for an analysis method to completely avoid this second drawback. Simplification is a matter of degree, so extending game-theoretic reasoning to accommodate greater complexity will enable the models to capture more of the richness of realistic cyber-security situations. This is particularly important for treatments of *adaptive* cyber-defense, since the dynamic evolution of configuration and information is the essence of adaptation. To be considered adaptive, a defense policy must take into account the attack state of the system, in consideration of how successful attacks require a succession of actions to gain knowledge about and eventually compromise targeted resources (Evans et al. 2011). Incorporating dynamics in the game model is therefore an absolute requirement for this domain. Dynamic information in turn poses significant technical challenges for game-theoretic methods (Tavafoghi et al. 2019).

One interesting effort to capture complex security dynamics in an abstract game model is the *FlipIt* framework introduced by Dijk et al. (2013). In FlipIt, two players vie for control of a single resource. Each has a single action, which takes control of the resource at some cost. Neither player can observe when the other has acted, and so is uncertain about the state of control except at the instant it performs its own action. Though the FlipIt model is quite abstract, it captures key elements of system security not well-supported by previous models (Bowers et al. 2012). Analysis of FlipIt has led to interesting insights about the interplay of various strategy classes, the value of aggressive play, and the significance of information advantages. As a stylized model, however, the generic version of FlipIt misses many relevant features of adaptive cyber-defense and is not suitable for decision making in a particular situation. Extensions of FlipIt have covered additional relevant scenario features (Farhang and Grossklags 2016; Jones et al. 2015; Laszka et al. 2013, 2014; Pham and Cid 2012). These add to practical realism, but seriously complicate analysis of the FlipIt game, which to date has eluded complete analytic solution, even in its basic version.

Which brings us finally to the approach described in this chapter: *empirical game-theoretic analysis* (EGTA) (Wellman 2016). Rather than build an analytic model that may be amenable to direct game-theoretic solution, EGTA starts with a detailed environment model described in procedural form, that is, by a simulation. We then introduce a set of specific dynamic strategies, and systematically run simulations over combinations of these strategies. The simulation data form the basis for estimating a game model, which can be solved using standard techniques.

The advantage of simulation is its ability to handle complex, stochastic, and temporally extended scenarios. This allows us to include realistic features of adaptive cyber-defense domains, going beyond generic and toy models. In its iterative form, EGTA also supports exploration, allowing us to address a rich set of strategic questions without premature simplification, such as isolating all the key strategic variables in advance. There are also limitations, particularly regarding the difficulty of generalizing game-theoretic conclusions beyond the specific environments and strategy instances studied. Overall, we regard EGTA as a complement to traditional game-theoretic treatments, which sacrifice complexity for generality (within the simplified model).

2 Empirical Game-Theoretic Analysis

The general idea of EGTA is to apply game-theoretic reasoning to models derived from agent-based simulation. The approach is designed to combine the advantage of simulation models in accommodating complexity with principles of strategic analysis expressible in the framework of game theory.

2.1 Basic Steps

The basic steps of EGTA are as follows, illustrated in Fig. 1.

1. Define a space of strategies for each player.
2. Simulate various combinations, or *profiles*, of agent strategies.
3. Induce or estimate an *empirical game model* from the accumulated simulated payoff data.
4. Analyze the resulting empirical game model, for example to identify Nash equilibria or otherwise characterize solutions of the game.

We elaborate on each step in turn.

Define Strategy Space. In EGTA, we typically constrain attention to a strict subset of the strategies that could be implemented in principle, for example imposing a parameterized form for strategies or adopting a particular agent

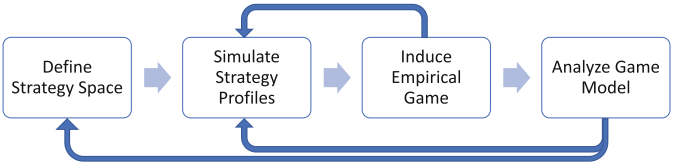


Fig. 1. Basic steps of empirical game-theoretic analysis. Feedback arrows show common patterns for iteration.

architecture. For this reason the available options are sometimes referred to as *heuristic strategies*. Though in general, a game may have any number of players, cyber-security games commonly focus on two: *attacker* (A) and *defender* (D). Let S_A and S_D denote their respective strategy sets.

Simulate Strategy Profiles. In a cyber-security game, we would simulate profiles (s_A, s_D) for various choices of $s_A \in S_A$ and $s_D \in S_D$. Each simulation yields a sample *payoff vector*, giving a numeric representation of the value of the outcome received by each player from one play of the given profile. Given stochastic factors in the simulation, we would typically require many samples of a profile to produce a reliable estimate of the expected payoffs to A and D.

Induce Empirical Game. In the most straightforward implementation of this step, we estimate a normal-form game model by sampling every profile $s \in S_A \times S_D$ a sufficient number of times. The payoff to player A in s, $u_A(s)$ is simply the sample average of A's payoffs in these simulations (and similarly $u_D(s)$ for player D). If the strategy spaces are very large, machine learning methods may be employed to generalize over the data to estimate payoffs for profiles not explicitly simulated (Vorobeychik et al. 2007).[2]

Analyze Game Model. The goal of analysis is to calculate Nash equilibria or another chosen solution concept, typically using off-the-shelf techniques. In the cyber-security context, let us define a *mixed profile* (σ_A, σ_D), with $\sigma_A \in \Delta(S_A)$ a probability distribution over A's strategy set (and similarly for σ_D) to be a joint strategy where each player independently chooses a strategy according to these distributions. Then (σ_A, σ_D) is a *Nash equilibrium* iff $\mathbb{E}[u_A(\sigma_A, \sigma_D)] \geq \mathbb{E}[u_A(s_A, \sigma_D)]$ for all $s_A \in S_A$, and similarly $\mathbb{E}[u_D(\sigma_A, \sigma_D)] \geq \mathbb{E}[u_D(\sigma_A, s_D)]$ for all $s_D \in S_D$.

Game analysis may also include reasoning about strategic relationships, such as dominance or ranking responses to particular opponents. Sensitivity analysis or statistical reasoning about candidate solutions would also be included in the game analysis step.

2.2 Iterative EGTA

It would be unusual for an EGTA study to proceed linearly according to steps 1-2-3-4 and complete. In practice, preliminary results at one step may inform reconsideration or elaboration of work at previous steps, and so the procedure would be iterative in nature. The key feedback links are shown in Fig. 1.

The simulation of strategy profiles (step 2) generates a collection of payoff samples. The number of samples required may not be straightforward to determine in advance. Feedback arrows from the game induction and analysis suggest that the results of these computations may be relevant in determining whether the collection is adequate, and if not, where additional simulation-based sampling is required. Such determination can be made on a principled basis through

[2] Such generalization is also often needed for the more general case of games where there are many players (Sokota et al. 2019; Wiedenbeck et al. 2018).

statistical analysis (Wiedenbeck et al. 2014), considering properties of the data collected and goals of the game analysis.

The longer feedback arrow from game analysis to strategy space represents *strategy exploration* (Jordan et al. 2010). Analysis of an accurate game model in step 4 gives us solutions to the game defined by the strategy space (S_A, S_D) defined in step 1. Since S_A and S_D are strict subsets of the true strategy sets available to players A and D, it is quite likely that the solutions found are not actually equilibria of the true game. We can bolster our confidence by considering additional strategies, thus defining augmented strategy sets $S'_A \supset S_A$ and $S'_D \supset S_D$. Solutions to the game over strategy space (S'_A, S'_D) are not actually guaranteed to be better approximations with respect to the full game (except in the limit when all strategies are included), but all else equal we expect improvement as more strategies are considered.

Of course, the interesting question in strategy exploration is which strategy or strategies to add at each iteration. A natural approach is to try to improve on the current equilibrium, by computing a *best response* to the other-player strategy. It turns out that the best response is generally not the optimal strategy to add in an iterative EGTA procedure (Jordan et al. 2010), as it does not consider opponent strategies outside the equilibrium, and it may not diversify the strategy set enough. Nevertheless, it is often a good heuristic, particularly if some stochastic exploration is conducted as well.

3 Example: A Moving Target Defense Game

We illustrate the EGTA approach to cyber-security by sketching the study of Prakash and Wellman (2015), which addressed an abstract scenario in moving-target defense (MTD). MTD covers a broad class of adaptive defenses where the main object is to defeat the attacker's ability to gain sufficient knowledge to compromise or take over a system (Jajodia et al. 2011). There are many MTD techniques, which accomplish this objective in various ways, generally involving some adaptation of the system to confuse the attacker or render its existing knowledge obsolete. We sought an abstract model that could fit the MTD approach broadly, without committing to a particular technology or system context. We thus adopted an extended version of the FlipIt model (van Dijk et al. 2013) discussed above in Sect. 1. The extension adds some complicating features present in prior work, such as multiple servers (Laszka et al. 2014) and asymmetric stealth (Laszka et al. 2013). It also incorporates a progressive concept of attack, in that unsuccessful attempts to compromise a server yield information that make subsequent attempts more likely to succeed (absent defender adaptation). This last feature is essential for capturing the primary dynamic of MTD (Albanese et al. 2019; Evans et al. 2011).

3.1 Game Description

In the specific MTD game studied, an attacker and defender compete for the control of 10 servers. (We could scale to many more servers with linear growth in the simulation time). Servers start out in control of the defender. The key actions are *probe* for the attacker, and *reimage* for the defender. A probe is essentially an attempt to compromise a server. If it succeeds, the attacker gains control, and if not, the attacker gains some information (not modeled explicitly) that increases its chance of succeeding on the next attempt. A reimage action by the defender takes a server down and resets its state. That is, any progress the attacker may have made on that server through probing is erased, such that probe success probability is reduced to its initial value.

The simulation proceeds for $T = 1000$ time steps. At each time step, the attacker may decide to probe any subset of the servers, and similarly the defender may choose some servers to reimage. Each faces a tradeoff, in that their actions help them achieve their goal of gaining or maintaining control of servers—but at a cost. For attackers, the probe actions bear an explicit cost, and for defenders the cost of reimaging is implicit in the downtime (7 time units in our setting) incurred for performing that action.

The state of the system at any point can be described by which player controls each server, and if the defender controls: whether it is down or up, and how many probes the attacker has attempted since the last reimage.

3.1.1 Observation Model

As argued above, cyber-security games are generally characterized by complex dynamics of state and observations, and this game is no exception. Technically, when agents cannot reliably observe each other's actions, the game is said to exhibit *imperfect information*. In this game, neither agent can perfectly observe the other. Precisely characterizing the model of what is and is not observed by each player is crucial for capturing the strategic interaction in an imperfect information game.

In the example MTD game, the defender has a partial ability to detect probes executed on any server, Specifically, if the server is up, the defender detects the probe with a specified probability, which varies across environment settings. However, the defender cannot tell whether a detected probe succeeded in compromising its target. The defender does of course know when it performs a reimage, and it is only at that point (and for the following downtime) that it can be sure it controls the server.

The attacker, on the other hand, does become aware when a probe succeeds. It also finds out when a server it controls is retaken by the defender through reimaging. Therefore, the attacker always knows the state of control of every server. However, it can only imperfectly track its progress in increasing success probability through probes, because it cannot tell when a defender reimages a server not in its control.

3.1.2 Utility

The primary objective of each player is to control servers. This is reflected in their utility functions, which quantify the value they attribute to any trajectory of states and actions. In the MTD game, players accrue utility each time period, based on the fraction of servers up and in their control, and also the fraction of servers *not* in the *other* player's control (i.e., either up and in the player's control, or not up).

This functional form of the utility function is designed to accommodate a variety of preference patterns, including objectives from the classic "CIA" (confidentiality, availability, integrity) triad (Pfleeger and Pfleeger 2012). For example, the *confidentiality* objective can be expressed through parameters encoding the defender's strong aversion to allowing the attacker to control servers. *Availability* from the defender's perspective can be expressed as requiring that a sufficient fraction of servers are in the defender's control and not down. We can categorize attacker utility in an analogous way. An attacker that accrues utility only by having servers in its control is termed a *control* attacker, whereas an attacker that accrues utility by having servers in its control or down is termed a *disrupting* attacker.

The utility function also includes threshold parameters governing the level of contention for servers in the associated environment. For example, by setting the threshold to 1/2 we impose the constraint that significant utility is accrued only if at least a majority are in control.

Finally, the utility model accounts for the cost of actions. The attacker pays a specified cost in utility per probe. The cost of the defender action is expressed implicitly in the utility function as the difference in utility accrued by servers being down as opposed to in the defender's control.

In the best case, a player accrues one utility unit per time period for keeping servers in their desired state, at no cost. The maximum overall utility for a game run is therefore T. The minimum is unbounded, as players may take unlimited costly actions without achieving their objective.

3.1.3 Strategies

In the EGTA approach, we focus on parameterized families of heuristic strategies, characterized by regular structures and patterns of behavior over time. Defining this strategy space is the first key step of EGTA (Fig. 1). The heuristic strategies defined for the MTD game generate actions based on the passage of time, or observed events in the system. If the actions are triggered by passage of time (in either a deterministic or probabilistic manner), we call the strategy *periodic*. The remaining strategies are triggered by observed events. They may apply actions to servers based on observations of that server, or based on combinations of observations across servers.

Specific families of heuristic strategies are defined for both attacker and defender. Within each family, there may be parametric options, so a large or even infinite number of possible instances. Overall, we considered 12 distinct attacker strategies and 20 defender strategies (i.e., $|S_A| = 12, |S_D| = 20$). These include for each player the *No-Op* strategy, in which the agent never takes any action.

Attacker Strategies. We consider two forms of periodic attacker strategy:

- *Uniform-Uncompromised.* Selects uniformly among those servers under the defender's control.
- *MaxProbe-Uncompromised.* Selects the server that has been probed the most since last reimage (that the attacker knows about), among those servers under the defender's control, breaking ties uniformly.

We also include one non-periodic attacker strategy that generates probe actions based on the number of servers that an attacker controls.

- *Control-Threshold.* If the attacker controls less than a threshold fraction of the servers, it chooses to probe the server that has been probed the most since last reimage (as far as it is aware) among those it does not currently control. Ties are broken uniformly among all eligible servers. A minimum waiting period of one time unit separates any two consecutive actions.

Defender Strategies. We consider periodic defender strategies employing two different criteria for server selection:

- *Uniform.* Selects uniformly among all up servers.
- *MaxProbe.* Selects the server that has been probed most since its last reimage, breaking ties uniformly.

The non-periodic defender strategies trigger a reimage operation based on probe activity or inactivity.

- *ProbeCount-or-Period* (PCP). Reimages a server whenever it detects that a threshold number of probes since the last reimage, *or* if it has been probed at least once but not within the specified period. The rationale for reimaging a server that is not being probed is that this could be an indication that the attacker has already compromised it and thus ceased attack.
- *Control-Threshold.* Analogous to the attacker's strategy by the same name, we include a defender strategy that performs a reimage action when the fraction of servers controlled falls below a threshold. Unlike the attacker, however, the defender cannot directly observe control state. Instead, the defender estimates the number of servers compromised based on the probes it has observed since reimaging on each server.
- *Control-Target.* Like Control-Threshold, except based on a target rather than a threshold.

3.2 Simulation and Analysis

We performed EGTA over a variety of environment and agent utility settings. The experiments covered a variety of environment settings, with systematic analysis over the possible combinations. Specifically, we varied games over the following features:

- Defender objective: confidentiality or availability.
- Attacker objective: disruption or control.

- Threshold on server control: low, majority, or high.
- Attacker probe cost: low, medium, or high.
- Defender probe detection: perfect, or imperfect at levels low, medium, or high.

Altogether, these settings define 144 distinct game instances. For 43 of these instances we conducted a full empirical game analysis: steps 2 through 4 of Fig. 1. These include all 18 instances with availability defender objective and perfect probe detection (Fig. 2), and another 18 with availability defender objective, imperfect probe detection, and medium p;robe cost (Fig. 3). We also ran seven with confidentiality defenders; as discussed below the confidentiality objective generally leads to an obvious equilibrium. For steps 2 and 3, we estimated through repeated simulation the joint payoffs for all $12 \times 20 = 240$ strategy profiles (s_A, s_D). Finally (step 4) we computed Nash equilibria for each game instance, using the Gambit software package (McKelvey et al. 2014). Most of the games had multiple equilibria—often similar, but sometimes quite diverse.

Our goal for this analysis was to gain strategic insight into a generic MTD scenario. As such, we were interested not so much in specifics of individual equilibria, but rather understanding at a qualitative level the kinds of equilibria observed. We found that equilibria could be classified into four qualitatively distinct groups.

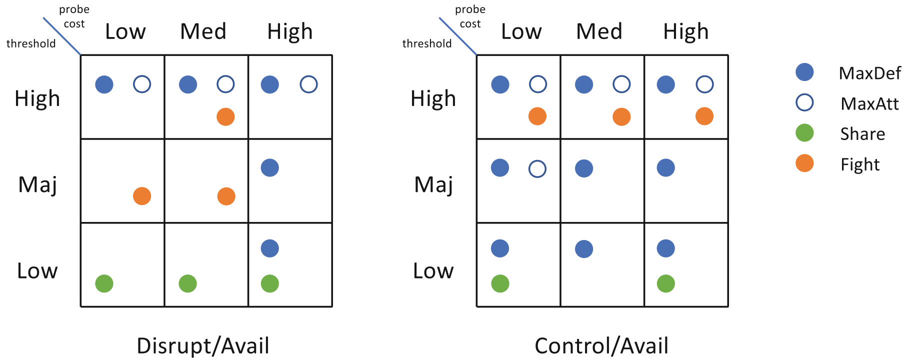
1. *MaxDef.* In a maximal defense equilibrium, the defender reimages so aggressively that it is futile for the attacker to even try to compromise the servers. Aggressive defense means a frequent periodic reimaging strategy or one that reimages based on a low-threshold probe trigger. Faced with such an aggressive defense, the attacker plays No-Op. As a result, the attacker gets no utility and the defender may get near maximum.
2. *MaxAtt.* We classify a profile as maximal attack if the attacker probes aggressively and in response the defender either plays No-Op or reimages only infrequently or ineffectively. This category is the dual of MaxDef, and corresponds to outcomes that are poor for the defender.
3. *Share.* We classify an equilibrium profile as a sharing if attack and defense levels are moderate, and both players are able to achieve their objectives.
4. *Fight.* Fight equilibria are characterized by robust attack and defense activity, resulting in persistent contention such that neither player achieves its objective to a satisfactory degree.

First, we observe that games with confidentiality defenders always have MaxDef equilibria. Such defenders care only that the attackers do not control their servers, and they can trivially achieve this objective by frequently reimaging—essentially keeping the servers down and unavailable. This result actually shows that a focus purely on confidentiality is not very realistic, so we devote the main part of our attention to games where defenders have an availability objective.

Figure 2 presents results for the 18 games where the defender has the availability objective and can perfectly detect attacker probes. The games cover all combinations of settings for attacker objective, probe cost, and utility thresholds for control of servers. As we see, the various game instances lead to different qualitative categories of equilibria. For the disrupting attacker (table on left), we have MaxDef equilibria for cases of high threshold or high probe cost. Those settings are particularly challenging for the attacker, enabling the defender to effectively deter attack through aggressive reimaging. Since the threshold setting applies to both players, the high threshold games also have MaxAtt equilibria, where an aggressive attack can cause the defender to give up. With low thresholds, both players need only achieve their objective with a minority of servers, so sharing equilibria are possible. Some of the intermediate settings support fight equilibria, where both players accrue some utility, but neither can keep the majority of servers in their preferred state on a consistent basis.

For the control attacker (table on right), the objective is more challenging than disruption. As a result, the defender always has the possibility of deterring attacks through sufficient aggression in a MaxDef equilibrium. MaxAtt can be sustained under the high threshold, or with majority threshold and low probe cost. Sharing equilibria appear for a couple of the low threshold environments, and fight equilibria in all the high threshold environments.

Results for 18 environments with imperfect probe detection are presented in Fig. 3. By comparing the two figures, it is obvious that maintaining a MaxDef equilibrium is much harder when the defender may miss some probes. On the other hand, degraded detection opens the door for aggressive attack, as MaxAtt and Fight are the only equilibria found in the majority or high threshold environments. With low threshold, sharing remains possible, and indeed this equilibrium is most prevalent.



threshold \ probe cost	Low	Med	High
High	🔵 ○	🔵 ○ 🟠	🔵 ○
Maj	🟠	🟠	🔵
Low	🟢	🟢	🔵 🟢

Disrupt/Avail

threshold \ probe cost	Low	Med	High
High	🔵 ○ 🟠	🔵 ○ 🟠	🔵 ○ 🟠
Maj	🔵 ○	🔵	🔵
Low	🔵 🟢	🔵	🔵 🟢

Control/Avail

🔵 MaxDef
○ MaxAtt
🟢 Share
🟠 Fight

Fig. 2. Qualitative categorization of equilibria across 18 game settings, with availability defender and perfect probe detection. The left table is for a disruptive attacker, and the right for a control attacker. In each cell, colored circles indicate which categories of equilibria are found. (Color figure online)

3.3 Discussion

This example is meant to illustrate general several features of EGTA for cybersecurity domains. First, that the method can address a strategically complex scenario, and evaluate a variety of heuristic strategies. Second, that through systematic exploration, we can uncover regular qualitative patterns of strategic behavior. Once identified, these patterns can deepen our understanding of the strategic tradeoffs in the domain. In this case, the findings can all be rationalized straightforwardly. Cases where multiple behaviors are possible (e.g., instances with both MaxDef and MaxAtt equilibria) are natural candidates for further study, toward characterizing refinements that would support one or the other.

4 Survey of Literature

The first application of EGTA to a security domain was the study of privacy attacks by Duong et al. (2010). This work started from the well-understood fact that an attacker's ability to compromise the privacy of a target depends on the background knowledge it already has about the target. In a scenario with multiple attackers, a coalition can increase their collective prospects of privacy breach by sharing background knowledge. There is a tradeoff, however, in that the value of a successful attack may decrease if it is non-exclusive. The study employed EGTA to characterize rational sharing in a variety of settings. The ability to predict sharing is relevant in particular to a database publisher, who must decide how much to degrade the published information (at a cost) in order to protect privacy.

A second security domain studied using EGTA by some of the same authors addressed incentives for compliance with a network security protocol (Wellman et al. 2013). Compliance is an important strategic problem for security,

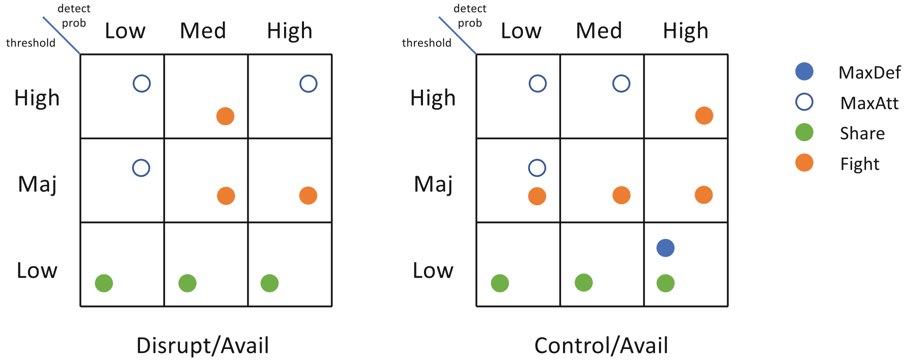
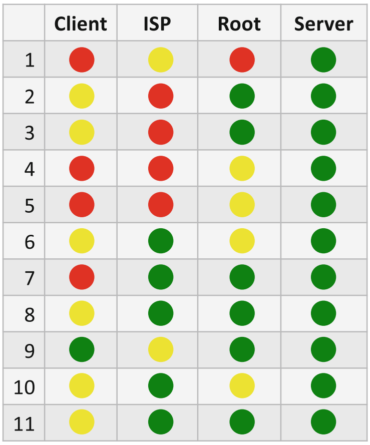


Fig. 3. Qualitative categorization of equilibria for imperfect probe detection. Columns represent three levels where the probability the defender detects a given probe action is 0.2 (low), 0.5 (med), or 0.7 (high).

as often participants will have an incentive to free-ride on the security contributions of others (Čagalj et al. 2005; Naghizadeh and Liu 2016). This study included several methodological innovations, including a systematic procedure to extend the strategy space through local search, and scaling the number of agents by exploiting symmetry across multiple roles. Specifically, the work modeled the introduction-based routing protocol (Frazier et al. 2011) on a network with four kinds of nodes: clients, ISPs, roots, and servers. The game is *role-symmetric*, meaning that players corresponding to a given role (in this context, node type) had the same strategy sets and utility functions, but these generally varied between roles. This enabled use of an aggregation technique called *player reduction* (Wellman et al. 2005), in which a many-player game is approximated by an empirical game with much fewer players. For example, one reported analysis simulated a 4956-node network to estimate a game with six players. Results for that instance are shown in Fig. 4. As we can see, tendency toward compliance varies by role, and there are qualitatively distinct equilibria. Overall, we found over several game settings that compliance was not universal, but typically at a sufficient level to deter attacks.

More recently, we have conducted several EGTA studies within a broader project on adaptive cyber-defense. The first was the MTD study illustrated in Sect. 3 (Prakash and Wellman 2015; Wellman and Prakash 2014). The second employed EGTA to evaluate a moving-target defense against distributed denial of service (DDoS) attacks (Wright et al. 2016). The defense, called MOTAG, had originally been designed and modeled in non-game-theoretic terms (Jia et al. 2013; Venkatesan et al. 2016). Like the MTD game study, the MOTAG



	Client	ISP	Root	Server
1	🔴	🟡	🔴	🟢
2	🟡	🔴	🟢	🟢
3	🟡	🔴	🟢	🟢
4	🔴	🔴	🟡	🟢
5	🔴	🔴	🟡	🟢
6	🟡	🟢	🟡	🟢
7	🔴	🟢	🟢	🟢
8	🟡	🟢	🟢	🟢
9	🟢	🟡	🟢	🟢
10	🟡	🟢	🟡	🟢
11	🟡	🟢	🟢	🟢

Fig. 4. Top 11 approximate symmetric mixed equilibria for a 4956-node instance of the introduction-based routing compliance game. Strategies are classified as compliant or non-compliant. Each row represents a mixed profile, indicating whether the role plays strategies that are compliant (green), non-compliant (red), or a mixture of these (yellow). (Color figure online)

investigation covered a two-player game with 10–20 strategies per player, and systematically evaluated a set of parametric variations on the game environment (41 game instances overall). We found that strategy ideas proposed in prior literature for this setting can be effective under certain conditions, but the ideal strategies varied considerably across these conditions. The study was helpful for making these conditions precise, and generally illuminating the strategic landscape for DDoS mitigation in the MOTAG framework.

The third EGTA study in this broader project addressed strategic behavior in domains that can be modeled by attack graphs (Nguyen et al. 2017). The basic idea of an attack graph model is to represent the progress of an attack in terms of following paths in a graph of security conditions (Kordy et al. 2014; Phillips and Swiler 1998). The work in this project specifically builds on a Bayesian framework for attack graphs developed by Miehling et al. (2015). The EGTA study extended the framework to a game, where at each time the attacker chooses edges representing available exploits, and a defender chooses nodes to defend. The strategy sets for both attacker and defender were populated by sophisticated heuristics developed as approximate solution of corresponding optimization problems. The study found that these heuristics successfully beat several baselines, and were robust to variation in the environment settings.

In work outside of this project, Chapman (2016) developed an abstract cybersecurity game based on an extension of hide-and-seek game models. The extensions were motivated by adaptive attack behavior in network security, and render the model infeasible for analytic solution. Chapman therefore adopted a simulation-based approach, and appealed to EGTA methods for game-theoretic treatment. Rass et al. (2017b) likewise appeal to EGTA for a game involving mitigation of advanced persistent threats, citing uncertainty as a complicating factor requiring this approach. Qi et al. (2018) model a scenario similar to the MTD game of Sect. 3 on a switching network using simulation to estimate game payoffs.

5 Conclusion and Extensions

As established by the MTD example and review of related literature, EGTA has by now been employed in a wide variety of adaptive cyber-defense applications. These works demonstrate the value of combining agent-based simulation and game-theoretic analysis in support of principled strategic reasoning for complex security domains. In each case, game-theoretic concepts were applied to scenarios of a complexity far exceeding the capacity of purely analytic methods to tackle.

Results of these analyses in many cases are compelling, though not necessarily definitive. Since by definition an EGTA study restricts attention to chosen strategies, conclusions are always subject to refutation based on refined analysis. Moreover, as for any modeling approach, assumptions incorporated in simulation or approximation methods are open for debate, or relaxation in subsequent studies. Indeed, there remain many areas where improvement in technique could significantly increase the power and scope of EGTA methodology. Here we briefly catalog some of the open issues and opportunities for extensions of EGTA in service of cyber-security analysis.

Covering Large Strategy Spaces. For a two-player game, profile space grows quadratically with strategy sets. This often allows consideration of a rich variety of attack and defense strategy candidates, albeit far from the full space of strategies available (typically highly dimensional or even infinite). Moreover, it is often possible to identify equilibria without evaluating all strategy combinations (Fearnley et al. 2013), which can sometimes dampen even quadratic growth. Limitations on strategy space become more acute when there are greater than two players. Though the standard setup in cyber-security domains is attacker versus defender, some scenarios naturally feature a broader set of strategic actors.

Automating Strategy Search. An effective approach to dealing with limitations on strategy space is to incrementally extend coverage, based on an iterative exploration using feedback from analysis of games of progressively increased size (Jordan et al. 2010). Given some formal description of the strategy space, strategy exploration can be automated in terms of a search in that space. Previous work has employed automated strategy generation for EGTA using local search (Wellman et al. 2013) or reinforcement learning (Lanctot et al. 2017; Schvartzman and Wellman 2009; Wright and Wellman 2018). Recent advances in deep learning have demonstrated breakthrough performance on two-player board games (Silver et al. 2017), and are demonstrating promise in cyber-security games as well (Wang et al. 2019; Wright et al. 2019).

Statistical Reasoning About Results. In the EGTA approach, the game model is estimated or induced from simulation data. The simulations are generally samples of a stochastic system, which means that results are subject to sampling error. This error may be mitigated by devoting more resources to sampling, though naturally that presents tradeoffs regarding alternative uses of that computation (e.g., to exploring more strategies or profiles). There has been some progress on developing principled methods for statistical reasoning in EGTA (Vorobeychik 2010; Wiedenbeck et al. 2014), but further work in this area is needed.

Generalizing Over Environments. The results produced from EGTA studies apply directly to the game instance modeled by the given simulator. Often in security settings, guidance about action in a specific instance is exactly what we care about. However, deriving broad insights about strategic issues in cyber-securities entails lifting results from specific instances to broad categories of game scenarios. The current state of art in EGTA is to systematically explore a range of environments, and attempt to identify patterns in the mapping to solutions. This approach is illustrated well by the qualitative characterization of equilibrium patterns in Figs. 2 and 3. Further work should attempt to codify and automate this systematic search and generalization process.

Acknowledgment. This work was partially supported by the Army Research Office under grant W911NF-13-1-0421.

References

Albanese, M., Connell, W., Venkatesan, S., Cybenko, G.: Moving target defense quantification. In: Jajodia et al. (2019)

Bowers, K.D., van Dijk, M., Griffin, R., Juels, A., Oprea, A., Rivest, R.L., Triandopoulos, N.: Defending against the unknown enemy: applying FlipIt to system security. In: Grossklags, J., Walrand, J. (eds.) GameSec 2012. LNCS, vol. 7638, pp. 248–263. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34266-0_15

Bushnell, L., Poovendran, R., Başar, T. (eds.): GameSec 2018. LNCS, vol. 11199. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01554-1

Čagalj, M., Ganeriwal, S., Aad, I., Hubaux, J.-P.: On selfish behavior in CSMA/CA networks. In: 24th IEEE International Conference on Computer Communications, pp. 2513–2524 (2005)

Chapman, M.: Cyber Hide-and-Seek. Ph.D. thesis, King's College London (2016)

Duong, Q., LeFevre, K., Wellman, M.P.: Strategic modeling of information sharing among data privacy attackers. Informatica **34**, 151–158 (2010)

Edwards, B., Furnas, A., Forrest, S., Axelrod, R.: Strategic aspects of cyberattack, attribution, and blame. Proc. Natl. Acad. Sci. **114**, 2825–2830 (2017)

Evans, D., Nguyen-Tuong, A., Knight, J.: Effectiveness of moving target defenses. In: Jajodia et al. (2011)

Farhang, S., Grossklags, J.: FlipLeakage: a game-theoretic approach to protect against stealthy attackers in the presence of information leakage. In: Zhu, Q., Alpcan, T., Panaousis, E., Tambe, M., Casey, W. (eds.) GameSec 2016. LNCS, vol. 9996, pp. 195–214. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47413-7_12

Fearnley, J., Gairing, M., Goldberg, P., Savani, R.: Learning equilibria of games via payoff queries. In: 14th ACM Conference on Electronic Commerce (2013)

Frazier, G., Duong, Q., Wellman, M.P., Petersen, E.: Incentivizing responsible networking via introduction-based routing. In: McCune, J.M., Balacheff, B., Perrig, A., Sadeghi, A.-R., Sasse, A., Beres, Y. (eds.) Trust 2011. LNCS, vol. 6740, pp. 277–293. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21599-5_21

Jajodia, S., Ghosh, A.K., Swarup, V., Wang, C., Sean Wang, X. (eds.): Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats. Springer, New York (2011). https://doi.org/10.1007/978-1-4614-0977-9

Jajodia, S., Cybenko, G., Liu, P., Wang, C., Wellman, M.P. (eds.): Adversarial and Uncertain Reasoning for Adaptive Cyber Defense. Springer, Champ (2019). https://doi.org/10.1007/978-3-030-30719-6

Jia, Q., Sun, K., Stavrou, A.: MOTAG: moving target defense against internet denial of service attacks. In: 22nd International Conference on Computer Communications and Networks (2013)

Jones, S., et al.: Evaluating moving target defense with PLADD. Technical report 8432R, Sandia National Lab (2015)

Jordan, P.R., Schvartzman, L.J., Wellman, M.P.: Strategy exploration in empirical games. In: 9th International Conference on Autonomous Agents and Multi-Agent Systems, pp. 1131–1138 (2010)

Kordy, B., Piètre-Cambacédès, L., Schweitzer, P.: DAG-based attack and defense modeling: don't miss the forest for the attack trees. Comput. Sci. Rev. **13**, 1–38 (2014)

Lanctot, M., et al.: A unified game-theoretic approach to multiagent reinforcement learning. In: 31st Annual Conference on Neural Information Processing Systems (2017)

Laszka, A., Johnson, B., Grossklags, J.: Mitigating covert compromises. In: Chen, Y., Immorlica, N. (eds.) WINE 2013. LNCS, vol. 8289, pp. 319–332. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-45046-4_26

Laszka, A., Horvath, G., Felegyhazi, M., Buttyán, L.: FlipThem: modeling targeted attacks with FlipIt for multiple resources. In: Poovendran, R., Saad, W. (eds.) GameSec 2014. LNCS, vol. 8840, pp. 175–194. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12601-2_10

Manshaei, M.H., Zhu, Q., Alpcan, T., Başar, T., Hubaux, J.-P.: Game theory meets network security and privacy. ACM Comput. Surv. **45**(25), 1–39 (2013)

McKelvey, R.D., McLennan, A.M., Turocy, T.L.: Gambit: software tools for game theory, Version 13.1.2 (2014). www.gambit-project.org

Miehling, E., Rasouli, M., Teneketzis, D.: Optimal defense policies for partially observable spreading processes on Bayesian attack graphs. In: Second ACM Workshop on Moving Target Defense, pp. 67–76 (2015)

Naghizadeh, P., Liu, M.: Opting out of incentive mechanisms: a study of security as a non-excludable public good. IEEE Trans. Inf. Forensics Secur. **11**, 2790–2803 (2016)

Nguyen, T.H., Wright, M., Wellman, M.P., Singh, S.: Multi-stage attack graph security games: heuristic strategies, with empirical game-theoretic analysis. In: Fourth ACM Workshop on Moving Target Defense, pp. 87–97 (2017)

Pfleeger, C.P., Pfleeger, S.L.: Analyzing Computer Security: A Threat/Vulnerability/Countermeasure Approach. Prentice Hall, Upper Saddle River (2012)

Pham, V., Cid, C.: Are we compromised? Modelling security assessment games. In: Grossklags, J., Walrand, J. (eds.) GameSec 2012. LNCS, vol. 7638, pp. 234–247. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34266-0_14

Phillips, C., Swiler, L.P.: A graph-based system for network-vulnerability analysis. In: Workshop on New Security Paradigms, pp. 71–79 (1998)

Prakash, A., Wellman, M.P.: Empirical game-theoretic analysis for moving target defense. In: Second ACM Workshop on Moving Target Defense, pp. 57–65 (2015)

Qi, C., Jiangxing, W., Cheng, G., Ai, J., Zhao, S.: Security analysis of dynamic SDN architectures based on game theory. Secur. Commun. Netw. **4123736**, 2018 (2018)

Rass, S., An, B., Kiekintveld, C., Fang, F., Schauer, S. (eds.): Decision and Game Theory for Security. LNCS, vol. 10575. Springer, Cham (2017a). https://doi.org/10.1007/978-3-319-68711-7

Rass, S., König, S., Schauer, S.: Defending against advanced persistent threats using game-theory. PLoS ONE **12**, e0168675 (2017b)

Roy, S., Ellis, C., Shiva, S.G., Dasgupta, D., Shandilya, V., Wu, Q.: A survey of game theory as applied to network security. In: 43rd Hawaii International Conference on System Sciences (2010)

Schvartzman, L.J., Wellman, M.P.: Stronger CDA strategies through empirical game-theoretic analysis and reinforcement learning. In: 8th International Conference on Autonomous Agents and Multi-Agent Systems, pp. 249–256, Budapest (2009)

Silver, D.: Mastering chess and shogi by self-play with a general reinforcement learning algorithm. Technical report, arXiv 1712.01815 (2017)

Sinha, A., Fang, F., An, B., Kiekintveld, C., Tambe, M.: Stackelberg security games: looking beyond a decade of success. In: 27th International Joint Conference on Artificial Intelligence, pp. 5494–5501 (2018)

Sokota, S., Ho, C., Wiedenbeck, B.: Learning deviation payoffs in simulation-based games. In: 33rd AAAI Conference on Artificial Intelligence, pp. 1266–1273 (2019)

Tambe, M.: Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned. Cambridge University Press, Cambridge (2011)

Tavafoghi, H., Yi, O., Teneketzis, D., Wellman, M.P.: Game theoretic approaches to cyber security: issues, results and challenges. In: Jajodia et al. (2019)

van Dijk, M., Juels, A., Oprea, A., Rivest, R.L.: FlipIt: the game of “stealthy takeover”. J. Cryptol. **26**, 655–713 (2013)

Venkatesan, S., Albanese, M., Amin, K., Jajodia, S., Wright, M.: A moving target defense approach to mitigate DDoS attacks against proxy-based architectures. In: IEEE Conference on Communications and Network Security (2016)

Vorobeychik, Y.: Probabilistic analysis of simulation-based games. ACM Trans. Model. Comput. Simul. **20**(3), 16:1–16:25 (2010)

Vorobeychik, Y., Wellman, M.P., Singh, S.: Learning payoff functions in infinite games. Mach. Learn. **67**, 145–168 (2007)

Wang, Y.: Deep reinforcement learning for green security games with real-time information. In: 33rd AAAI Conference on Artificial Intelligence (2019)

Wellman, M.P.: Putting the agent in agent-based modeling. Auton. Agents Multi-Agent Syst. **30**, 1175–1189 (2016)

Wellman, M.P., Prakash, A.: Empirical game-theoretic analysis of an adaptive cyber-defense scenario (preliminary report). In: Poovendran, R., Saad, W. (eds.) GameSec 2014. LNCS, vol. 8840, pp. 43–58. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12601-2_3

Wellman, M.P., Reeves, D.M., Lochner, K.M., Cheng, S.-F., Suri, R.: Approximate strategic reasoning through hierarchical reduction of large symmetric games. In: 20th National Conference on Artificial Intelligence, pp. 502–508 (2005)

Wellman, M.P., Kim, T.H., Duong, Q.: Analyzing incentives for protocol compliance in complex domains: a case study of introduction-based routing. In: Twelfth Workshop on the Economics of Information Security (2013)

Wiedenbeck, B., Cassell, B.-A., Wellman, M.P.: Bootstrap techniques for empirical games. In: 13th International Conference on Autonomous Agents and Multi-Agent Systems, pp. 597–604 (2014)

Wiedenbeck, B., Yang, F., Wellman, M.P.: A regression approach for modeling games with many symmetric players. In: 32nd AAAI Conference on Artificial Intelligence, pp. 1266–1273 (2018)

Wright, M., Wellman, M.P.: Evaluating the stability of non-adaptive trading in continuous double auctions. In: 17th International Conference on Autonomous Agents and Multi-Agent Systems, pp. 614–622 (2018)

Wright, M., Venkatesan, S., Albanese, M., Wellman, M.P.: Moving target defense against DDoS attacks: an empirical game-theoretic analysis. In: Third ACM Workshop on Moving Target Defense (2016)

Wright, M., Wang, Y., Wellman, M.P.: Iterated deep reinforcement learning in games: history-aware training for improved stability. In: 20th ACM Conference on Economics and Computation (2019)