# Methods and Algorithms for Infinite Bayesian Stackelberg Security Games (Extended Abstract)

Christopher Kiekintveld[1], Janusz Marecki[2], and Milind Tambe[3]

[1] University of Texas at El Paso, El Paso, TX
Department of Computer Science
cdkiekintveld@utep.edu
[2] IBM T.J. Watson Research Center, Yorktown Heights, NY
marecki@us.ibm.com
[3] University of Southern California, Los Angeles, CA
Department of Computer Science
tambe@usc.edu

**Abstract.** Recently there has been significant interest in applications of game-theoretic analysis to analyze security resource allocation decisions. Two examples of deployed systems based on this line of research are the ARMOR system in use at the Los Angeles International Airport [20], and the IRIS system used by the Federal Air Marshals Service [25]. Game analysis always begins by developing a model of the domain, often based on inputs from domain experts or historical data. These models inevitably contain significant uncertainty—especially in security domains where intelligence about adversary capabilities and preferences is very difficult to gather. In this work we focus on developing new models and algorithms that capture this uncertainty using continuous payoff distributions. These models are richer and more powerful than previous approaches that are limited to small finite Bayesian game models. We present the first algorithms for approximating equilibrium solutions in these games, and study these algorithms empirically. Our results show dramatic improvements over existing techniques, even in cases where there is very limited uncertainty about an adversaries' payoffs.

**Keywords:** Bayesian Stackelberg games, security games, approximate algorithms, sampling techniques

## 1 Introduction

Game theory offers a powerful framework for modeling security decisions, both for protecting critical infrastructure [22, 6] and computer networks [3, 16, 18]. In two recent real-world applications game theory is used to make security resource allocation decisions at the Los Angeles International Airport (LAX) [20] and for the Federal Air Marshals Service (FAMS) [25]. The game models at the heart of these systems capture the capabilities of both the police and the adversaries, as well as information about the potential consequences of different outcomes. An important consideration in these domains is that the security policy should be unpredictable, which comes naturally from the game analysis assuming a rational and adaptive adversary.

Specifying an accurate game model to represent the domain is a crucial first step in any application of game theory. These models are typically based on input from domain experts, including the ones developed for the LAX and FAMS applications. Even though the models are based on the best available information they are *inherently uncertain*. In security games in particular it is very problematic to provide precise and accurate information about the preferences and capabilities of possible attackers. Our goal in this work is to develop new models and algorithms that explicitly reason about this uncertainty in the context of security games. Most game-theoretic models and solution algorithms make strong assumptions about perfect information and common knowledge. Bayesian games [11] are commonly used to represent uncertainty in games, but unfortunately the available algorithms for solving Bayesian games are limited. The best available algorithm for Bayesian security games is DOBSS [19] which applies to games with a finite number of attacker types. Unfortunately, this method does not scale well with the number of types and in practice can only solve relatively small games with few types.

We define a model for Bayesian Stackelberg Security Games with continuous payoff distributions for the attacker, leading to a infinite Bayesian game. For example, in this model we can represent the uncertainty that security forces have about the attacker's payoffs using normal distributions, or uniform distributions over an interval. Solving these game to find equilibrium solutions presents significant challenges; even for the finite case the problem is NP-hard [9], no exact method is known for the infinite case. We explore methods for approximating equilibrium solutions for these problems, and test these approaches experimentally. Specifically, we describe two different methods for computing the defender's optimal strategy. The first applies DOBSS to a find optimal solutions for games with a finite number of sample attacker types. The second uses replicator dynamics to approximate an optimal strategy using the Monte-Carlo sampling for approximating the attacker response. In this shortened version of the paper we present only an abbreviated summary of the main results; in a longer version we will present a full detailed experimental evaluation.

## 2 Related Work

Recent interest in applying game theory to security decision includes fielded applications at the Los Angeles International Airport [20] and the Federal Air Marshals Service [25], work on patrolling strategies for robots and unmanned vehicles [10, 2, 5], policy recommendations for protecting critical infrastructure [22, 6], and applications in computer networks [3, 16, 18]. Bayesian games [11] are the dominant paradigm for modeling uncertainty in game theory, and there are many examples of specific games that have been solved analytically, including many types of auctions [14]. Unfortunately, algorithms for finding equilibria of Bayesian games are quite limited, and no general algorithms exist for infinite Bayesian games. Recent research efforts have focused primarily on developing approximation techniques [21, 4, 8]. Monte-Carlo sampling approaches similar to those we consider in our work have been applied to some kinds of auctions [7]. In addition, the literature on stochastic choice [15, 17] studies problems that are simplified versions of the choices that attackers face in our model.

Finally, the literature on robust optimization has also inspired distribution-free alternatives to Bayes-Nash equilibrium [1].

## 3 Bayesian Security Games

We define a new class of Bayesian Security Games, extending the model in Kiekintveld et. al. [12] to include uncertainty about the attacker's payoffs. The key difference between our model and existing approaches (such as in Paruchuri et. al [19]) is that we allow the defender to have a continuous distribution over the possible payoffs of the attacker. Previous models have restricted this uncertainty to a small, finite number of possible attacker types, limiting the kinds of uncertainty that can be modeled.

A security game has two players, a *defender*, $\Theta$, and an *attacker*, $\Psi$, a set of *targets* $T = \{t_1, \ldots, t_n\}$ that the defender wants to protect (the attacker wants to attack) and a set of *resources* $R = \{r_1, \ldots, r_m\}$ (e.g., police officers) that the defender may deploy to protect the targets. Resources are identical in that any resource can be deployed to protect any target, and any resource provides equivalent protection. A defender's pure strategy, denoted $\sigma_\Theta$, is a subset of targets from $T$ with size less than or equal to $m$ An attacker's pure strategy, $\sigma_\Psi$, is exactly one target from $T$. $\Sigma_\Theta$ denotes the set of all defender's pure strategies and $\Sigma_\Psi$ is the set of all attacker's pure strategies.

We model the game as a Stackelberg game [23] which unfolds as follows: (1) the defender commits to a mixed strategy $\delta_\Theta$ that is a probability distribution over the pure strategies from $\Sigma_\Theta$, (2) nature chooses a random attacker type $\omega \in \Omega$ with probability $Pb(\omega)$, (3) the attacker observes the defender's mixed strategy $\delta_\Theta$, and (4) the attacker responds to $\delta_\Theta$ with a best-response strategy from $\Sigma_\Psi$ that provides the attacker (of type $\omega$) with the highest *expected* payoff given $\delta_\Theta$.

The payoffs for the defender depend on which target is attacked and whether the target is protected (covered) or not. Specifically, for an attack on target $t$, the defender receives a payoff $U^u_\Theta(t)$ if the target is uncovered, and $U^c_\Theta(t)$ if the target is covered. The payoffs for an attacker of type $\omega \in \Omega$ is $U^u_\Psi(t, \omega)$ for an attack on an uncovered target, and $U^c_\Psi(t, \omega)$ for an attack on a covered target.We assume that both the defender and the attacker know the above payoff structure exactly. However, the defender is uncertain about the attacker's type, and can only estimate the expected payoffs for the attacker. We choose not to model uncertainty that the attacker may have over the defender's payoffs because the attacker already observes the defender's strategy perfectly.

### 3.1 Bayesian Stackelberg Equilibrium

A Bayesian Stackelberg Equilibrium (BSE) for a security game consists of a strategy profile where every attacker type is playing a best-response to the defender strategy, and the defender is playing a best-response to the distribution of actions chosen by the attacker types. We first define the equilibrium condition for the attacker and then the equilibrium condition for the defender. We conveniently represent the defender's mixed strategy $\delta_\Theta$ by the compact *coverage vector* $C = (c_t)_{t \in T}$ that gives the probabilities $c_t$ that each target $t \in T$ is covered by at least one resource. Note that $\sum_{t \in T} c_t \leq m$

because there are $m$ resources at the defender's disposal. In equilibrium each attacker type $\omega$ must best respond to the coverage $C$ with a pure strategy $\sigma^*_\Psi(C,\omega)$ given by:

$$\sigma^*_\Psi(C,\omega) = \arg\max_{t\in T}(c_t \cdot U^c_\Psi(t,\omega) + (1-c_t)\cdot U^u_\Psi(t,\omega)) \tag{1}$$

To define the equilibrium condition for the defender we first define the *attacker response function* $A(C) = (a_t(C))_{t\in T}$ that returns the probabilities $a_t(C)$ that each target $t \in T$ will be attacked, given the distribution of attacker types and a coverage vector $C$. Specifically:

$$a_t(C) = \int_{\omega\in\Omega} Pb(\omega)\mathbf{1}_t(\sigma^*_\Psi(C,\omega))d\omega \tag{2}$$

where $\mathbf{1}_t(\sigma^*_\Psi(C,\omega))$ is the indicator function that returns 0 if $t = \sigma^*_\Psi(C,\omega)$ and 0 otherwise. Given the attacker response function $A(\cdot)$ and a set of all possible defender coverage vectors $\mathcal{C}$, the equilibrium condition for the defender is to execute its best-response mixed strategy $\delta^*_\Theta \equiv C^*$ given by:

$$\delta^*_\Theta = \arg\max_{C}\sum_{t\in T} a_t(C)(c_t \cdot U^c_\Theta(t) + (1-c_t)\cdot U^u_\Theta(t)). \tag{3}$$

When the set of attacker types is infinite, calculating the attacker response function from Equation (2) is impractical. For this case we instead replace each payoff in the original model with a continuous distribution over possible payoffs. Formally, for each target $t \in T$ we replace values $U^c_\Psi(t,\omega)$, $U^c_\Psi(t,\omega)$ over all $\omega \in \Omega$ with two continuous probability density functions:

$$f^c_\Psi(t,r) = \int_{\omega\in\Omega} Pb(\omega)U^c_\Psi(t,\omega)d\omega \tag{4}$$

$$f^u_\Psi(t,r) = \int_{\omega\in\Omega} Pb(\omega)U^u_\Psi(t,\omega)d\omega \tag{5}$$

that represent the defender's *beliefs* about the attacker payoffs. For example, the defender expects with probability $f^c_\Psi(t,r)$ that the attacker receives payoff $r$ for attacking target $t$ when it is covered. This provides a convenient and general way for domain experts to express uncertainty about payoffs in the game model, whether due to their own beliefs or based on uncertain evidence from intelligence reports.

## 4 Solution Methods

To solve the model described in the previous section we need to find a Bayesian Stackelberg equilibrium, describing an optimal coverage strategy for the defender and the optimal response for every attacker type. If the space of possible attacker types is finite, an optimal defender strategy can be found using DOBSS [19]. Unfortunately, there are

no known methods for finding exact equilibrium solutions for infinite Bayesian Stackelberg games, and DOBSS only scales to small numbers of types. Here we focus on methods for approximating solutions for infinite Bayesian Stackelberg games. The problem can be broken down into two parts:

1. Computing or estimating the attacker response function (Equation 2)
2. Optimizing over the space of defender strategies, given the attacker response function

Similarly to the previous work [13] we compute the attacker response function using Monte-Carlo sampling from the space of possible attacker types. In addition, we consider both optimal and approximate methods for optimizing the defender's strategy given the attacker response calculations. We now briefly describe these two methods for solving infinite Bayesian Stackelberg security games.

### 4.1 Sampled Bayesian ERASER

The first method we describe combines Monte-Carlo sampling from the space of attacker types with an exact optimization over the space of defender strategies. This approach is based on the DOBSS solver [19] for finite Bayesian Stackelberg games. However, we also incorporate several improvements from the ERASER solver [12] that offer faster solutions for the restricted class of security games. The resulting method can be encoded as a mixed-integer linear program (MIP), which we call *Bayesian ERASER* (not presented here due to space constraints).

To use Bayesian ERASER to approximate a solution for an infinite game we can draw a finite number of sample attacker types from the type distribution, assuming that each occurs with equal probability. The payoffs for each type are determined by drawing from the payoff distributions specified in Equations 4 and 5. This results in a constrained, finite version of the infinite game that can be solved using the Bayesian ERASER MIP. We refer to this method as *Sampled Bayesian ERASER* (SBE) and use SBE-$x$ to denote this methods with $x$ sample attacker types. Armantier et al. [4] develop an approach for approximated general infinite Bayesian games that relies on solving constrained versions of the original game. Given certain technical conditions, a sequence of equilibria of constrained games will converge to the equilibrium of the original game. Here, increasing the number of sample types corresponds to such a sequence of constrained games, so in the limit as the number of samples goes to infinity the equilibrium of SBE-$\infty$ will converge to the true Bayes-Nash equilibrium.

### 4.2 Sampled Replicator Dynamics

The second approach we consider uses a local search method (replicator dynamics) to approximate the defender's optimal strategy, given the attacker response function. Given that we are already using sampling techniques to estimate the attacker response, it makes a great deal of sense to explore approximation methods for the defender optimization as well. This allows us to trade off whether additional computational resources

should be devoted to improving the attacker response estimate, or improving the defender strategy. In our experimental results we show that this is key to scaling to large problem instances.

We implemented an approximation algorithm based on replicator dynamics [24], which we call *Sampled Replicator Dynamics* (SRD). Since this method is a form of local search, all we require is a black-box method to estimate the attacker response function, such as Monte-Carlo sampling. As above, we use SRD-$x$ to denote the Monte-Carlo version of SRD with $x$ sample attacker types. SRD proceeds in a sequence of iterations. At each step the current coverage strategy $C^n = (c_t^n)_{t \in T}$ is used to estimate the attacker response function, which in turn is used to estimate the expected payoffs for both players. A new coverage strategy $C^{n+1} = (c_t^{n+1})_{t \in T}$ is computed according to the replicator equation:

$$c_t^{n+1} \propto c_t^n \cdot (E_t(C) - U_\Theta^{min}), \tag{6}$$

where $U_\Theta^{min}$ represents the minimum possible payoff for the defender, and $E_t(C)$ is the expected payoff the defender gets for covering target $t$ with probability 1 and all other targets with probability 0, given the estimated attacker response to $C^n$. The search runs for a fixed number of iterations, and returns the coverage vector with the highest expected payoff. We introduce a learning rate parameter $\alpha$ that interpolates between $C^n$ and $C^{n+1}$, with $C^{n+1}$ receiving weight $\alpha$ in the next population and $C^n$ having weight $1 - \alpha$. Finally, we introduce random restarts to avoid becoming stuck in local optima. After initial experiments, we settled on a learning rate of $\alpha = 0.8$ and random restarts every 15 iterations, which generally yielded good results (though the solution quality was not highly sensitive to these settings).

## 5 Evaluation

We omit the majority of our evaluation due to space constraints, but present one result demonstrating the importance of modeling uncertainty rather than using a perfect-information approximation. We generate 500 random game instances with 5 targets and 1 defender resource. The defender's payoffs for a covered target are drawn from $U[0, 100]$, and the uncovered payoffs from $U[-100, 0]$. The attacker's payoffs are represented by Gaussian distributions, with mean values drawn from $U[-100, 0]$ for covered targets and $U[0, 100]$ for uncovered targets; we vary the standard deviation. A sample attacker type is defined by drawing one value from each of these distributions (two values for each target).

The baseline algorithm uses a single point to estimate each payoff, rather than a distribution. This is motivated by the standard methodology for eliciting game models from domain experts, where no information about the uncertainty of the parameters is included in the model. We model this with a perfect-information model where the attacker has only one type, corresponding to the mean value for each payoff distribution. This can be solved exactly using the SBE algorithm with a single attacker type, which we refer to as "SBE-Mean."

Figure 1 presents results for the solution quality for SBE-Mean, SBE, and SRD. We vary payoff uncertainty along the x-axis, measured by the standard deviation of the Gaussian distributions for the attacker payoffs (in the same units as the payoffs). We

run each algorithm to generate a coverage strategy for the defender, and evaluate this coverage strategy against the true distribution of attacker types. Since we do not have a closed-from solution to compute this exactly, we rely on a very close approximation generated by sampling 10000 attacker types to evaluate the payoffs for each algorithm. The expected payoffs are shown on the vertical axis. We run SBE with up to 7 sample types and SRD with up to 1000 due to large differences in the computational scalability of the algorithms. With only 7 types, SBE takes roughly 2 seconds to run, while SRD runs in less than half a second with 1000 types and 5000 search iterations.

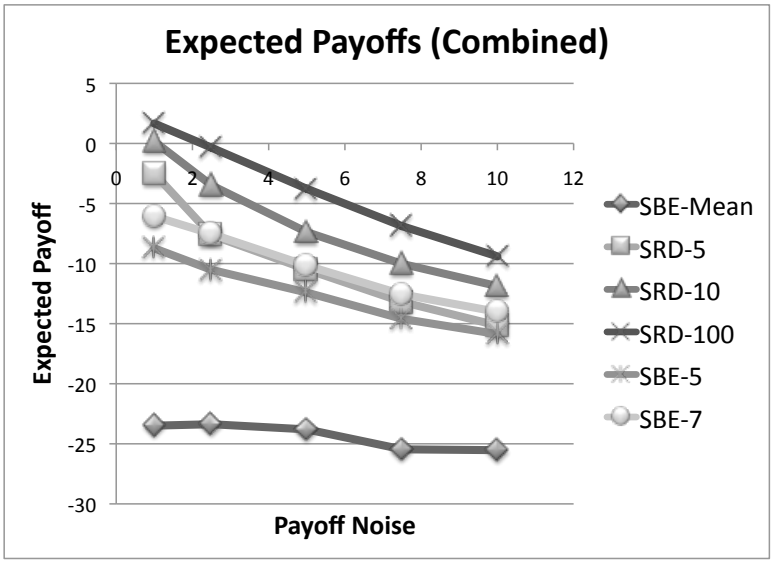


**Fig. 1.** Expected payoffs for SBE-Mean, SBE, and SRD with varying numbers of sample attacker types.

In Figure 1 we see that the solution quality for both SBE and SRD is dramatically higher than the SBE-Mean baseline when there is payoff uncertainty, even if the uncertainty is relatively small. SBE and SRD show improvements over the baseline even with very small numbers of sample attacker types, with diminishing returns as the number of types increases. This is a strong indication that the perfect-information approach is not a good approximation for security games with uncertainty about the attacker's payoffs. SBE and SRD represent the first steps towards more robust methods that give high-quality solutions even when there is payoff uncertainty.

## 6 Conclusion

Stackelberg games are increasingly important in the analysis of a broad range of security domains, including deployed applications. The existing method to model uncertainty in

these games are restricted to simple games with small, finite numbers of attacker types. We develop a class of infinite Bayesian Stackelberg security games in which attacker payoffs are provided as distributions, rather than point estimates of the payoffs. These games are able to more accurately capture the real payoff uncertainties in security domains, but present new computational challenges. We develop methods for approximating equilibrium solutions for these games using sampling and local search techniques. The SBE method exploits existing techniques for finite Bayesian Stackelberg games to solve constrained version of the infinite games. The SRD method combines replicator dynamics for searching the space of defender strategies with Monte-Carlo sampling techniques to estimate the attacker response function.

Our first important finding is that the baseline method that ignores uncertainty yields very poor results. *Modeling payoff uncertainty is critical in security games.* Both SBE and SRD give solutions with dramatically higher quality than the mean-approximation benchmark, even with just a few sample attacker types. The second major finding is that approximating the defender strategy enables scaling to much larger games and improves overall solution quality by enabling better approximations of the attacker response function. SRB is able to scale to very large problems (hundreds of targets) while using many more sample types than SBE. These results have immediate implications for the use of game theory in security domains, and open an exciting new research area in developing better approximation methods for Bayesian security games.

**Acknowledgments.** This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE) under grant number 2007- ST-061-000001. However, any opinions, conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the Department of Homeland Security. Janusz Marecki was supported in part by the DARPA GALE project, Contract No. HR0011-08-C-0110.

## References

1. Aghassi, M., Bertsimas, D.: Robust game theory. Mathematical Programming: Series A and B 107(1), 231–273 (2006)
2. Agmon, N., Kraus, S., Kaminka, G.A., Sadov, V.: Adversarial unvertainty in multi-robot patrol. In: IJCAI (2009)
3. Alpcan, T., Basar, T.: A game theoretic approach to decision and analysis in network intrusion detection. In: Proc. of the 42nd IEEE Conference on Decision and Control. pp. 2595–2600 (2003)
4. Armantier, O., Florens, J.P., Richard, J.F.: Approximation of bayesian nash equilibrium. Journal of Applied Econometrics 23(7), 965–981 (December 2008)
5. Basiloco, N., Gatti, N., Amigoni, F.: Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In: AAMAS (2009)
6. Bier, V.M.: Choosing what to protect. Risk Analysis 27(3), 607–620 (2007)
7. Cai, G., Wurman, P.R.: Monte Carlo approximation in incomplete information, sequential auction games. Decision Support Systems 39(2), 153–168 (2005)
8. Ceppi, S., Gatti, N., Basilico, N.: Computing bayes-nash equilibria through support enumeration methods in bayesian two-player strategic-form games. In: Proceedings of the

ACM/IEEE International Conference on Intelligent Agent Technology (IAT). pp. 541–548. Milan, Italy (September 15-18 2009)
9. Conitzer, V., Sandholm, T.: Computing the optimal strategy to commit to. In: ACM EC. pp. 82–90 (2006)
10. Gatti, N.: Game theoretical insights in strategic patrolling: Model and algorithm in normal-form. In: ECAI. pp. 403–407 (2008)
11. Harsanyi, J.C.: Games with incomplete information played by Bayesian players (parts i–iii). Management Science 14 (1967–8)
12. Kiekintveld, C., Jain, M., Tsai, J., Pita, J., Ordóñez, F., Tambe, M.: Computing optimal randomized resource allocations for massive security games. In: AAMAS (2009)
13. Kiekintveld, C., Marecki, J., Tambe, M.: Robust bayesian methods for stackelberg security games. In: Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-agent systems (2010)
14. Krishna, V.: Auction Theory. Academic Press (2002)
15. Luce, R.D., Raiffa, H.: Games and Decisions. John Wiley and Sons, New York (1957), dover republication 1989
16. wei Lye, K., Wing, J.M.: Game strategies in network security. International Journal of Information Security 4(1–2), 71–86 (2005)
17. McFadden, D.: Quantal choice analysis: A survey. Annals of Economic and Social Measurement 5(4), 363–390 (1976)
18. Nguyen, K.C., Basar, T.A.T.: Security games with incomplete information. In: Proc. of IEEE Intl. Conf. on Communications (ICC 2009) (2009)
19. Paruchuri, P., Pearce, J.P., Marecki, J., Tambe, M., Ordonez, F., Kraus, S.: Playing games with security: An efficient exact algorithm for Bayesian Stackelberg games. In: AAMAS. pp. 895–902 (2008)
20. Pita, J., Jain, M., Western, C., Portway, C., Tambe, M., Ordonez, F., Kraus, S., Paruchuri, P.: Depoloyed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. In: AAMAS (Industry Track) (2008)
21. Reeves, D.M., Wellman, M.P.: Computing best-response strategies in infinite games of incomplete information. In: UAI (2004)
22. Sandler, T., M., D.G.A.: Terrorism and game theory. Simulation and Gaming 34(3), 319–337 (2003)
23. von Stackelberg, H.: Marktform und Gleichgewicht. Springer, Vienna (1934)
24. Taylor, P., Jonker, L.: Evolutionary stable strategies and game dynamics. Mathematical Biosciences 16, 76–83 (1978)
25. Tsai, J., Rathi, S., Kiekintveld, C., Ordóñez, F., Tambe, M.: IRIS - A tools for strategic security allocation in transportation networks. In: AAMAS (Industry Track) (2009)