Bandits for Cybersecurity: Adaptive Intrusion Detection Using Honeypots

Marcus Gutierrez and Christopher Kiekintveld
mgutierrez22@miners.utep.edu and cdkiekintveld@utep.edu
University of Texas at El Paso, USA

Abstract
Intrusion detection is a fundamental problem in network security, and honeypots are one method for actively detecting malicious activity by using deception to fool attackers into interacting with fake hosts or services. We consider the problem of how to strategically select which configurations of honeypots to use to maximize the detection capability in a network. This problem is complicated by the extremely large space of possible configurations, as well as the ability of attackers to evolve their strategies over time. We propose that this problem can be addressed by combining methods from computational game theory and learning strategies for multi-armed bandits. We describe a preliminary model for allocating honeypots using this approach.

Introduction
Intrusion detection and response is one of the fundamental challenges in cybersecurity. Many common approaches to intrusion detection are passive; they inspect patterns of activity on computer networks and individual hosts to detect the signatures of known attacks, or to identify suspicious anomalies. One exception to this is the use of honeypots to detect attackers and collect detailed intelligence about the goals and tactics of attackers. Honeypots are fake hosts or services that are designed to attract attention from attackers and gather detailed information about incidents (Spitzner 2003). Honeypots differ from other methods of intrusion detection in that they use deception to actively draw out attackers and reveal their presence and strategies.

The basic principle of honeypots has been used in network security for many years, and a great deal of research has considered the problem of making honeypots as believable as possible so they cannot be easily identified by attackers. For example, systems like Honeyd (Provos 2003) and Honeycomb (Kreibich and Crowcroft 2004) provide platforms and tools for deploying large numbers of honeypots on computer networks to aid in intrusion detection. However, very little work has been done to date on how to use honeypots strategically to maximize detection capability given the limited resources available for deploying and maintaining honeypots. We describe the several key challenges for optimizing the value of honeypots, and propose an approach for deploying honeypots that is both strategic and adaptive using methods from computational game theory and machine learning.

Challenges in Honeypot Allocation
At first glance, using a honeypot (or many honeypots) for intrusion detection may appear simple—just design a believable system, deploy it on the network, and wait for attackers to interact with it. However, this ignores important decisions about how many honeypots to deploy, and exactly what types of honeypots to deploy. A honeypot must be given a specific configuration, and there are an astronomical number of potential configurations to consider. Each different configuration exposes a different attack surface (set of vulnerabilities), and any specific configuration will detect only a small subset of potential attacks. In addition, attackers are intelligent and will avoid configurations that appear “too good to be true.” Attackers also evolve over time, as they learn to avoid common types of honeypots, and new vulnerabilities are discovered and integrated into hacking tools and worms. All of this means that network administrators using honeypots as part of an intrusion detection strategy need to act strategically and adapt over time to maximize the effectiveness of this approach.

Game Theory and Multi-Armed Bandits
We argue that two areas of artificial intelligence offer complementary tools to address this challenge: game theory and machine learning. Game theory provides a mathematical framework for modeling multi-agent decision problems, including adversarial situations with deception and information manipulation. Security games have become an important paradigm for optimally randomizing the allocation of limited security resources (e.g., police patrols) for infrastructure protection. Several recent papers have applied this basic approach to the problem of deciding how to allocate limited honeypots in a network (Garg and Grosu 2007; Carroll and Grosu 2011; Pibil et al. 2012). These models typically use one-shot extensive form games to model the problem, and often use abstract ways to specify what types of honeypots to use, such as the “value” of the host.

While the current game-theoretic models provide a useful way to formalize the problem of allocating honeypots and some useful strategic insights, they have several important
limitations. First, the solution algorithms have limited scalability, and cannot be applied beyond modest network sizes or when there are a very large number of host configurations to consider. Second, the analysis is static, so they propose a one-time selection of honeypot configurations that will be used forever, with no learning or adaptation. Finally, they assume knowledge of the attacker’s capabilities and preferences to accurately model their decision making.

To address these limitations, we turn to another area of artificial intelligence: machine learning, and particularly online learning strategies based on the multi-armed bandit model. Multi-armed bandits (MAB) capture the tradeoff between exploration and exploitation in an agent that needs to make repeated choices over time (Robbins 1985). The basic scenario considers a decision maker playing slot machines (bandits) where each machine has a different, initially unknown distribution of payouts. Over time, the player can learn about the payouts and must balance gathering more information with maximizing the current expected payout. While the basic MAB problem assumes fixed distributions of payouts, the adversarial MAB assumes that the payouts can be altered by an adversary (Auer et al. 1995).

Multi-armed bandits provide a useful way to think about several aspects of the honeypot allocation problem. First, they explicitly model repeated interactions with the potential for learning over time. Know policies for playing MAB provide bounds on the regret (maximum loss) as the number of interactions increases, including for the adversarial case. They also allow us to model situations where the defender does not have accurate information about the attacker’s payoffs, and instead must learn about the attacker by observing their behavior over time. However, the basic MAB problem does not capture all of the features of the honeypot problem. For example, it ignores the structure and topology of the network, and the potential for strategic interactions between the players.

Proposed Model

We propose using a combination of methods from computational game theory and multi-armed bandits to address the problem of strategically allocating honeypots in a dynamic way. Our model considers a repeated interaction between the defender and the attacker. In each round, the defender observes the network as a set of hosts where each host is represented by a set of vulnerabilities. The defender selects a limited number of honeypots to add to the network; each honeypot is also represented as a set of vulnerabilities, modeling different configurations. The attacker observes the modified network including the honeypots, but without knowledge of which hosts are real and which ones are honeypots. The attacker chooses a particular vulnerability to attempt to exploit. The rewards for both players depend on whether the attacker interacts with a honeypot or not when choosing this particular vulnerability.

Choosing which configurations of honeypots to add for the defender maps to the choice of an arm in the MAB problem. Each honeypot configuration consists of a collection of vulnerabilities, effectively forming a combinatorial bandit problem (Cesa-Bianchi and Lugosi 2012). However, there is also a strategic component to our model, since the network structure can be analyzed to provide additional information about which vulnerabilities are most important. We have performed preliminary experiments on this model with known policies for MAB such as EXP3 (Auer et al. 1995), which demonstrate that we can learn effective honeypot deployment strategies even against intelligent adversaries. We are currently working to develop more sophisticated learning policies for the defender that account for the combinatorial aspect of selecting honeypot configurations, as well as incorporating strategic analysis based on the network structure.

Conclusion

We have identified the problem of how to use deception strategies such as deploying honeypots strategically as an important problem for improving intrusion detection capabilities in cybersecurity. The problem of deciding how to use limited resources to deploy the most effective configurations of honeypots poses several interesting challenges. We propose using a combination of methods from computational game theory and multi-armed bandits to address these challenges, and have developed a preliminary model to evaluate this approach.

References


