Optimal Decoy Resource Allocation for Proactive Defense in Probabilistic Attack Graphs

Extended Abstract

Haoxiang Ma
University of Florida
Gainesville, United State
hma2@ufl.edu

Shuo Han
University of Illinois Chicago
Chicago, United State
hanshuo@uic.edu

Nandi Leslie
Raytheon Technologies
Arlington County, United State
nandi.o.leslie@raytheon.com

Charles Kamhoua
U.S. Army Research Laboratory
Rome, United State
charles.a.kamhoua.civ@mail.mil

Jie Fu
University of Florida
Gainesville, United State
fujie@ufl.edu

ABSTRACT

This paper investigates the problem of synthesizing proactive defense systems in which the defender can allocate deceptive targets and modify the cost of actions for the attacker who aims to compromise security assets in this system. We model the interaction of the attacker and the system using a formal security model– a probabilistic attack graph. By allocating fake targets/decoys, the defender aims to distract the attacker from compromising true targets. By increasing the cost of some attack actions, the defender aims to discourage the attacker from committing to certain policies and thereby improve the defense. To optimize the defense given limited decoy resources and operational constraints, we formulate the synthesis problem as a bi-level optimization problem, while the defender designs the system, in anticipation of the attacker's best response given that the attacker has disinformation about the system due to the use of deception. Though the general formulation with bi-level optimization is NP-hard, we show that under certain assumptions, the problem can be transformed into a constrained optimization problem. We proposed an algorithm to approximately solve this constrained optimization problem using a novel, incentive-design method for projected gradient ascent. We demonstrate the effectiveness of the proposed method using numerical experiments.

KEYWORDS

Deception; Attack Graph; Bi-Level Optimization; Markov Decision Process

ACM Reference Format:
Haoxiang Ma, Shuo Han, Nandi Leslie, Charles Kamhoua, and Jie Fu. 2023. Optimal Decoy Resource Allocation for Proactive Defense in Probabilistic Attack Graphs: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Proactive defense refers to a class of defense mechanisms for the defender to detect any ongoing attacks, distract the attacker with deception, or use randomization to increase the difficulty of an attack to the system. This paper proposes a mathematical framework and solution approach for synthesizing a proactive defense system with deception.

We start by formulating the attack planning problem using a probabilistic attack graph, which can be viewed as a Markov decision process with a set of attack target states. Attack graphs(AGs)[5] can be used in modeling computer networks. They are widely used in network security to identify the minimal subset of vulnerability/sensors to be used in order to prevent all known attacks[6, 7]. Probabilistic attack graphs introduce uncertain outcomes of attack actions that account for action failures in a stochastic environment. For example, in [3, 4], probabilistic transitions in attack graphs capture uncertainties originated from network-based randomization. Under the probabilistic attack graph modeling framework, we investigate how to allocate decoy resources as fake targets to distract the attacker into attacking the fake targets and how to modify the attack action costs to discourage the attacker from reaching the true targets.

The joint design of decoy resource allocation and action cost modification can be cast as a bi-level optimization problem, where the defender (at the upper level) designs the system, in anticipation of the attacker's (at the lower level) best response, given that the attacker has disinformation about the system due to allocated decoys. However, bi-level optimization problems are generally NP-hard [1]. Under the assumption that potential decoy states are predefined, and the defender only needs to allocate resources/rewards to decoys, we prove the bi-level optimization can be equivalently expressed as a constrained optimization problem. To solve the constrained optimization problem using a projected gradient ascent efficiently, we build two important relations: First, we show that the projection step of a defender's desired attack policy to the set of realizable attack policy space can be performed using Inverse Reinforcement Learning (IRL) [8]. Essentially, IRL shapes the attacker's perceived reward so that the attacker will mimic a strategy chosen by the defender. Second, the gradient ascent step can be performed using policy improvement, which is a subroutine in policy iteration with respect to maximizing the defender's total reward. The projected gradient ascent is ensured to converge to a (local) optimal solution to this nonconvex-constrained optimization problem.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom. © 2023 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

2 METHOD

2.1 A Bi-level Optimization Formulation

We consider the adversarial interaction between a defender and an attacker in a system equipped with proactive defense. We use a probabilistic attack graph to capture how the attacker plans to achieve the attack objective. The probabilistic attack graph, also known as a Markov decision process (MDP) with a set of attack target states, can be modeled $M = (S, A, P, \nu, \gamma, F, R_2)$, where S is the set of statespace(nodes in the attack graph); A is the set of actions; $P : S \times A \to \text{Dist}(S)$ is a probabilistic transition function such that $P(s'|s, a)$ is the probability of reaching s' given action a taken at state s; $\nu \in \text{Dist}(S)$ is the initial state distribution; $\gamma \in (0, 1]$ is a discount factor. The attacker's objective is described by a final state set F and a reward function $R_2 : F \times A \to \mathbf{R}_{\geq 0}$ which assigns each state-action pair (s, a) where $s \in F$ and $a \in A$ to a nonnegative value of reaching that target for the attacker.

The attacker aims to find a policy π to maximize her value function

$$V_2^{\pi}(\nu) = \mathbf{E}_{\pi}[\sum_{k=0}^{\infty} \gamma^k R_2(S_k, A_k)|S_0 \sim \nu],$$

where $\mathbf{E}_{\pi}$ is the expectation given the probability measure $\mathbf{Pr}^{\pi}$.

We assume the defender knows the attacker's objective given by the tuple $\langle F, R_2 \rangle$. The defender has the following proactive defense mechanisms: decoy resource allocation and state-action reward modification. Let $\vec{x} \in \mathbf{R}_{\leq 0}^{|S \times A|}$ represent the state-action reward modification policy. The attacker's perceptual reward function is

$$R_2^{\vec{x}}(s, a) = \begin{cases} \vec{x}(s, a) & \text{if } \vec{x}(s, a) < 0, \\ R_2(s, a) & \text{if } \vec{x}(s, a) = 0. \end{cases}$$

Let $\vec{y} \in \mathbf{R}_{\geq 0}^{D}$ represent the decoy resource allocation policy, where D is the set of decoys. The attacker's perceptual reward function is defined by

$$R_2^{\vec{y}}(s, a) = \begin{cases} \vec{y}(s) & \text{if } \vec{y}(s) > 0, \\ R_2(s, a) & \text{if } \vec{y}(s) = 0. \end{cases}$$

The defender's reward function $R_1 : S \to \mathbf{R}$ defined by $R_1(s) = 1$ if $s \in D$, otherwise $R_1(s) = 0$. The defender aims to find a defending strategy $(\vec{x}, \vec{y})$ that the attacker's policy π maximizes his value function $V_1^{\pi}(\nu, \vec{y}) = \mathbf{E}_{\pi}[\sum_{k=0}^{\infty} \gamma^k R_1(S_k)|S_0 \sim \nu]$.

Then our problem can be formulated as follows:

Problem 1.

$$\begin{aligned} \max_{\vec{x} \in X, \vec{y} \in Y} \quad & V_1^{\pi^*}(\nu; \vec{y}) \\ \text{s.t.} \quad & \pi^* \in \operatorname*{argmax}_{\pi} V_2^{\pi}(\nu; \vec{x}, \vec{y}). \end{aligned}$$

2.2 Transforming into a Constrained Optimization Problem

The bi-level optimization problem is known to be strongly NP-hard [2]. However, under certain conditions, the bi-level optimization problem can be shown to be equivalent to a constrained optimization problem. Let $\Pi(\vec{x}, \vec{y})$ be the set of response policies in the attacker's perceived planning problem with respect to a choice of variables $\vec{x}$ and $\vec{y}$. The bi-level optimization problem is then equivalently written as the following constrained optimization problem:

$$\begin{aligned} \max_{\pi^*} \quad & V_1^{\pi^*}(\nu, \vec{y}) \\ \text{s.t.} \quad & \pi^* \in \overline{\Pi} \triangleq \bigcup_{\vec{y} \in Y, \vec{x} \in X} \Pi(\vec{x}, \vec{y}). \end{aligned}$$

Because the above problem is a standard-constrained optimization problem, one can obtain a locally optimal solution using the projected gradient method. A key step in performing Projected Gradient Ascent (PGA) is to evaluate the projection in the PGA. However, this is nontrivial because the set $\overline{\Pi}$ includes a set of attack policies, each of which corresponds to a choice of vectors $(\vec{x}, \vec{y})$. As a result, $\overline{\Pi}$ does not have a compact representation. However, it is noted that the projection step is equivalent to minimizing the distance between policy π and $\hat{\pi}$, which is equivalent to

$$\begin{aligned} \min_{\pi} \quad & \mathbf{D}(\pi, \hat{\pi}) \\ \text{s.t.} \quad & \pi \in \overline{\Pi}, \quad \vec{y}(s) > 0; \forall s \in D. \end{aligned}$$

where $\mathbf{D}(\pi, \hat{\pi})$ is the distance between the two policies $\pi, \hat{\pi}$.

Thus we can use IRL [8] to perform the projection step. Then use policy improvement to perform the gradient ascent step.

3 EXPERIMENT

We have evaluated our proposed method using a probabilistic attack graph.The attacker obtains a reward by reaching the true target or fake decoys. The defender gets rewards when the attacker enters the fake decoys.

When we do not allocate fake target rewards, the attacker has a probability 60.33% of reaching the target set F from the initial state. When we assign rewards to the fake target, the attacker has a probability 8.63% of reaching the target set F from the initial state. By assigning resources to decoys to attract the attacker, the defender significantly reduces the attacker's probability of reaching the target state (85% reduction) and improves the defender's value by 3.38 times.

4 CONCLUSION

We present a mathematical framework and algorithms for decoy allocation and reward modification in a proactive defense system. Our technical approach can be applied to many safety-critical systems where the probabilistic attack graphs are constructed from known vulnerabilities in a system. The formulation and solutions can be extended to a broad set of adversarial interactions in which proactive defense with deception can be deployed.

ACKNOWLEDGMENTS

Research was sponsored by the Army Research Office under Grant Number W911NF-22-1-0166 and W911NF-22-1-0034 and partially by NSF under Award #2144113. 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 Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

[1] Stephan Dempe and Alain Zemkoho. 2020. *Bilevel optimization.* Springer.

[2] Pierre Hansen, Brigitte Jaumard, and Gilles Savard. 1992. New branch-and-bound rules for linear bilevel programming. *SIAM Journal on scientific and Statistical Computing* 13, 5 (1992), 1194–1217.

[3] Jin Hong and Dong-Seong Kim. 2012. HARMs: Hierarchical Attack Representation Models for Network Security Analysis. In *Australian Information Security Management Conference*. SRI Security Research Institute, Edith Cowan University, Perth, Western Australia, 9.

[4] Jin B. Hong and Dong Seong Kim. 2016. Assessing the Effectiveness of Moving Target Defenses Using Security Models. *IEEE Transactions on Dependable and Secure Computing* 13, 2 (March 2016), 163–177.

[5] S. Jha, O. Sheyner, and J. Wing. 2002. Two Formal Analyses of Attack Graphs. In *Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15.* 49–63.

[6] Steven Noel and Sushil Jajodia. 2008. Optimal ids sensor placement and alert prioritization using attack graphs. *Journal of Network and Systems Management* 16, 3 (2008), 259–275.

[7] Oleg Sheyner, Joshua Haines, Somesh Jha, Richard Lippmann, and Jeannette M Wing. 2002. Automated generation and analysis of attack graphs. In *Proceedings 2002 IEEE Symposium on Security and Privacy*. IEEE, 273–284.

[8] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, Anind K Dey, et al. 2008. Maximum entropy inverse reinforcement learning.. In *Aaai*, Vol. 8. Chicago, IL, USA, 1433–1438.