Minmax Restless Bandits for Efficient Moving Target Defense

Pierre Charreaux, Alexandre Reiffers-Masson, Francoise Sailhan, Sandrine Vaton

Minmax Restless Bandits for Efficient Moving Target Defense

Pierre Charreaux, Alexande Reiffers-Masson, Francoise Sailhan and Sandrine Vaton
IMT-Atlantique, Lab-STICC Laboratory, UMR CNRS 6285
Email:{firstname.surname}@imt-atlantique.fr

***Abstract*—The static configuration of networks (including security-related tools) gives attackers a significant advantage. To address this issue, we propose a Moving Target Defense (MTD) that introduces changes in the virtual network configuration and disrupts the attacker's exploration phase. In particular, we formalise MTD as a restless bandit problem, considering IP and port shuffling.**

I. Introduction

Spurred by the need to provision network services faster, Network Service Providers (NSPs) have undertaken a major transformation of the network infrastructure by adopting Network Functions Virtualization (NFV) and Software-Defined Networks (SDN). NFV[1] entails implementing network functions - that are traditionally available on hardware-based middleboxes and proprietary network equipment - as software appliances. The potential of NFV is realized when paired with SDN that enables a flexible centralized management by decoupling the control and data planes. Overall, the resulting virtual network is made up of (virtualised) network elements (nodes and links) that are managed to form a virtual topology running on top of a Substrate Network (SN). The abstraction introduced by resource virtualization mechanisms allows network operators to manage and modify virtual networks in a flexible and dynamic way: multiple virtual network topologies can easily be hosted on the same physical hardware.

Nonetheless, this shift to virtual networks implies the sharing of hardware resources and the network softwarisation[2], which together expose virtual networks to new vulnerabilities that cybercriminals may exploit.

Furthermore, traditional defense tools use static network configurations and fail to leverage the flexibility of virtual network, thus leaving the attacker with the same advantage as on traditional networks.

To improve upon the previous issue, our approach relies on Moving Target Defense (MTD) [1], which makes changes across various system aspects by e.g. constantly moving software, changing open network ports. In particular, the proposed MTD, in our paper, is intended to introduce some changes to the network configuration that disrupt the exploration phase of the attacker, increase probing costs and render investigations inaccurate. The goal is also to prevent and/or delay attacks by increasing the uncertainty of the attacker. When designing an optimal MTD strategy, our aim is to answer the three following questions:

- "What to move?" - This involves modeling the virtual network and determining which attributes of the VN configuration should be modified to alter the exploration, attack, detection or prevention surfaces [2]. For this purpose, NFV MANO [3], NFV Infras, and NFVs software layers can be used to access a wider set of configuration attributes. In this paper, we focus on moving IP addresses and ports.
- "How to move?" - This defines ways of modifying attributes by devising strategies (algorithms, protocols) that shift the virtual network in a way that maximizes the security of the network. The challenge here is to improve the network security without compromising the quality of service of users. Changes in the attributes of network elements are made in a controlled manner by defenders, making the targeted network unpredictable, dynamic, heterogeneous and more reliable. For this purpose, three key methods are traditionally implemented [3]:
  - Shuffling: changes are made to the system configuration. They may consist in randomizing e.g., IP addresses and ports.
  - Increasing the diversity: systems are deployed with different implementations, e.g. different softwares, which reduces the risk of exploiting implementation-dependent vulnerabilities.
  - Redundancy: multiple replicas of system (e.g. NFV) are provided to increase reliability and availability.
- "When to move?" - This attempts to determine the best time to make the moves. This aspect, which is poorly investigated in the literature, is usually addressed by periodically[4] applying configuration changes.

In this paper, we present an MTD that shuffles, in an unpredictable way, the IP and port of the nodes composing the virtual network. We hereafter formalize the attacker-defender interactions as a restless bandits problem and find an optimal movement strategy.

II. Moving Target Defense

We introduce a MTD that adds movement to a virtualized network to make the information collected by the attacker

[1]https://www.etsi.org/technologies/nfv
[2]a.k.a introduction of a new software stack
[3]network functions virtualization management and orchestration
[4]With a constant interval time between two successive configurations.

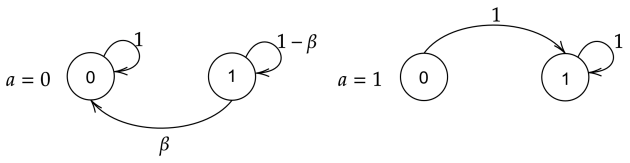

Fig. 1. Markov Decision Process of the arms with a shuffling probability β.

obsolete during a recognition phase. We exemplify our MTD with a use case, in which *the attacker* scans the virtualized network and *the defender* shuffles IP addresses and ports. The advantage of shuffling strategy is threefold: the shuffling (may) render obsolete information previously gleaned by the attacker ; an attacker ignores whether/which reconfiguration has been applied, unless the attacker performs another scan.

In practice, the attacker explores the network to determine the networked services (along with related vulnerabilities) that are accessible via the hosts ports. With the nmap tool, the attacker may follow different scanning strategies ranging from scanning the network either quickly (at the risk of almost certain detection) or slowly (with a smaller probability of detection). After scanning the network, the attacker will be able to exploit the discovered vulnerabilities.

A. Problem Formulation

We formalize the behavior of the attacker and defender as a N-arms Restless Bandit problem [4], with each of the N arms corresponding to a host port. The state of each arm evolves following a Markov process with an attacker that decides at each time step which of the available arms to pull (i.e., which hosts ports to scan). In this setting, a controlled Markov chain governs each arm. We assume that the transition matrices of each arm are statistically equivalent. As depicted in Figure 1, an arm of the bandit is a host port that can be either scanned (state 1) or not in which case the state of the host port is unknown (state 0), which means that the port has never been scanned or was shuffled by the defender. The actions are either to scan the port (i.e. the arm is pulled, noted $a = 1$) or do nothing (i.e. the arm is not pulled, noted $a = 0$).

To prevent the attacker from being detected, we assume that as long as the attacker is not scanning more than a given number of ports per time unit, the attacker is not detected. Each pulled arm produces a positive reward in a stochastic manner and the attacker goal is to maximize the reward accumulated over time. We hereafter detail the model.

Rewards: Each pulled arm generates a reward. The attacker decides at each step which of the N arms to pull over a sequence of trials, with the aim of maximizing the long-term reward. Note that it is beneficial for the attacker to search for the most popular services (e.g. as HTTPS, DNS) and to a lesser extent those that are rare. The reward depends on the service popularity, which is categorised into 3 service classes corresponding to well known services (ports 0 to 1023), common services (registered port 1024 to 49151), and remaining services (49152 to 65535).

Budget constraint: the attacker activates a fraction of the available arms ; the scanning remains undetected as long as the fraction of ports scanned at time t is less than α (with $0 \leq \alpha \leq 1$).

Shuffling: the shuffling is modeled as the probability of going from the scanned state to the unknown state, as depicted in the transition matrix of the arms. The shuffling strategy varies across time, according to the port class and refers to:

- IP shuffling, which complicates the attacker's task that has to re-scan IP addresses,
- Port Shuffling, which makes the attacker loose information about ports.

Given that a shuffled service (resp. host) is still running on host (resp. network), the attacker needs to find these latter again. In practice, IP and port shuffling is managed by a SDN controller that handles virtual ports/IPs and attempts to maintain the service availability.

B. Conclusion

We formalized the attacker-defender interactions using min-max weakly coupled Markov decision processes. The attacker is modeled as a restless bandit, which pulls arms to realize attacks. On the other hand, the defender shuffles ports and IPs, by changing the transition probabilities of the arms of the attacker.

Our future work involves:

- Solving the bandits problem, using the tools developed in [4].
- Find the optimal shuffling probability (using gradient descent) that varies across time and depending on the class. This optimization should reduce (i) the rewards accumulated by the attacker and (ii) the shuffling cost, expressed in term of e.g., resources consumption and QoS/QoE metrics.
- So far, we considered IP and port scanning/shuffling, we plan to adapt and extend our work to deal with other types of attack threatening virtual networks [5], [6] .

REFERENCES

[1] R. Ross, V. Pillitteri, G. Guissanie, R. Wagner, R. Graubart, and D. Bodeau, "Enhanced security requirements for protecting controlled unclassified information: A supplement to nist special publication 800-171 (final public draft)," National Institute of Standards and Technology, Tech. Rep., 2020.

[2] S. Sengupta, A. Chowdhary, A. Sabur, A. Alshamrani, D. Huang, and S. Kambhampati, "A survey of moving target defenses for network security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, 2020.

[3] J.-H. Cho, D. P. Sharma, H. Alavizadeh, S. Yoon, N. Ben-Asher, T. J. Moore, D. S. Kim, H. Lim, and F. F. Nelson, "Toward proactive, adaptive defense: A survey on moving target defense," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, 2020.

[4] N. Gast, B. Gaujal, and C. Yan, "Linear program-based policies for restless bandits: Necessary and sufficient conditions for (exponentially fast) asymptotic optimality," *Mathematics of Operations Research*, 2023.

[5] L. R. Bays, R. R. Oliveira, M. P. Barcellos, L. P. Gaspary, and E. R. Mauro Madeira, "Virtual network security: threats, countermeasures, and challenges," *Journal of Internet Services and Applications*, vol. 6, no. 1, 2015.

[6] T. Penner and M. Guirguis, "Combating the bandits in the cloud: A moving target defense approach," in *17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*, 2017.