A Survey of Attack Models for Cyber-Physical Security Assessment in Electricity Grid

Yu-Cheng Chen, Vincent Mooney and Santiago Grijalva
Georgia Institute of Technology
Atlanta, USA
ychen414@gatech.edu, mooney@ece.gatech.edu, sgrijalva@ece.gatech.edu

***Abstract*—This paper surveys some prior work regarding attack models in a cyber-physical system and discusses the potential benefits. For comparison, the full paper will model a bad data injection attack scenario in power grid using the surveyed prior work.**

Keywords— Bad Data Injection, Attack Graph, Attack Propagation, Markov Chain, PLADD

I. INTRODUCTION

The electric power grid is a cyber-physical system comprised of cyber and physical layers. The physical layer includes components such as power generators, transmission lines, and loads. The cyber layer consists of computer devices and communication lines. The purpose of the cyber layer is to provide control over the grid in order to improve the grid's performance and increase its overall reliability.

It is important to address the cybersecurity aspects of a cyber-physical system (CPS). However, security analysts are often constrained by time and money. Security analysts need to determine a reasonable way of spending resources (time and/or money) to protect a CPS. Securing a CPS goes beyond securing the individual system components. A motivated adversary often uses the inter-dependency of vulnerabilities to carry out multistage attacks. While each phase of the attack may not pose a serious threat to the corresponding component, the combined effect, however, may be catastrophic.

The rest of the paper is organized as follows. Section II describes the attack scenario of this research. Section III describes some of the prior works in attack modeling of CPS. Section IV discusses the possibility of using attack modeling for attack mitigation and conclusion.

II. ATTACK SCENARIO

A. Bad Data Injection Attack Surface

The attack scenario discussed in this paper is bad data injection attack. Figure 1 shows the steps of executing a bad data injection attack on a power grid system. Figure 1 shows that the adversary must gain access to a remote terminal unit (RTU) data, a vulnerability report, and IP address information before executing the bad data injection attack against a power grid. In reality, an adversary needs to prepare for an attack by spending time and money to gather information. During the preparation stage of the bad data injection attack, the information gathered may be invalidated due to actions done by the defender. For example, the adversary may spend time to steal login credentials to remote login to a power grid's communication network. However, if the defender is resetting the password every month, the stolen login information will only be valid for at most one month. In this paper's attack scenario, we assume the adversary may gain access and lose access to RTU data, vulnerability report, and IP information multiple times before actually executing a bad data injection attack on a power grid. In addition, RTU data, vulnerability report, and IP information may be attacked in any order. For the execution stage of the attack, the adversary must breach the substation, inject fake data, and avoid detection from state estimation.

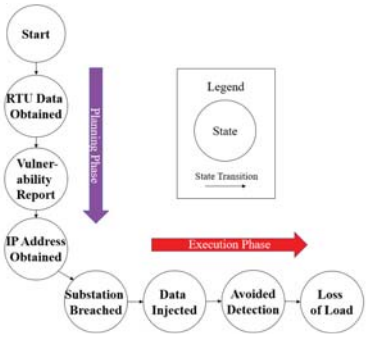

Figure 1. Attack graph capturing attacker's strategy

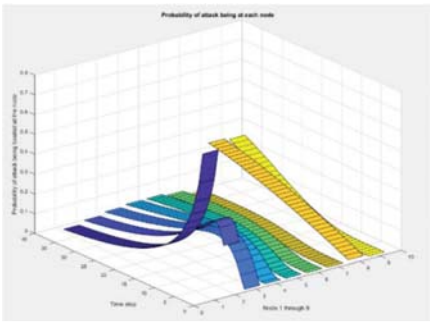

Figure 2. Probability of an attack being at each node with respect to time step for Markov Chain model

III. ATTACK MODELS

A. Markov chain model

In our previous work[1], we used a Markov approach to model the steps of our bad data injection attack, as shown in Figure 1. The probabilities of a successful attack are

978-1-7281-3915-9/19/$31.00 ©2019 IEEE

estimated but are intended to represent upper bounds on the probability of success of each attack step. As described in [1], the probability of the adversary's attack at each node is calculated and plotted in Figure 2.

B. PLADD model

Another interesting prior work is the PLADD (probabilistic learning attacker, dynamic defender) model [2]. The PLADD model well-suited for modeling persistent attacks over a long time. The PLADD model can model the interactions between attacker and defender. However, the PLADD model does not support modeling of attacker actions that cannot be revoked by a defender action. For example, attacker actions such as "jumping a fence" or "breaking into a substation" which is unsupervised in our scenario, cannot be modeled by the PLADD model because there is no defender action to revoke attacker's access. More formally, PLADD models as defined in [2] have the following two requirements for any game it is modeling:

- The defender does not know who owns the resource and is unable to use detection techniques;
- The defender has a fixed periodic action capable of retaking the resource.

As a result, we do not consider modeling our scenario completely in PLADD.

C. Hybrid attack model

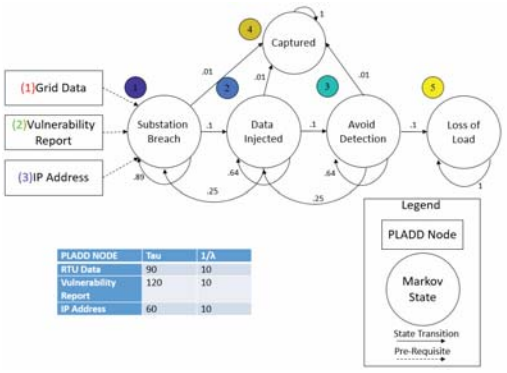

Figure 3. Hybrid model attack graph, where the table shows the parameters used for each PLADD node

The hybrid attack model (HAM) is another prior work that combines the benefits of the Markov chain and PLADD model. An advantage of the hybrid attack model in comparison to the Markov Chain and PLADD model is that the "time resolution" in the preparation stage and the execution stage can be significantly different. For example, it may be possible that the attacker needs to run keyboard logger for months before being able to determine the correct password to a cloud system. However, once the attacker has determined the correct password, the actual attack, which is to steal data from the cloud, may take less than a day to complete.

We implemented HAM for our scenario shown in Figure 1 using MATLAB. Figure 4a (left figure) is the state of the PLADD nodes with respect to time. Figure 4a shows that there are four instances during the simulation that the attacker's progress is propagated to the execution stage. The execution stage of our bad data injection attack scenario is modeled by Markov nodes, as shown in Figure 3. Figure 4b shows the probability of attack at each Markov node for one attack on the power grid, which happens to be one day long. The duration of the attack is one day long because the first attack in the execution stage happened on the 14th day. Then, the defender invalidated at least one information needed by the attacker to continue attacking in the execution stage on the 15th day. Figure 4(b) also shows that at the end of the 14th day, the probability that the attacker has reached node 5 in Figure 3 is 32.38%, and the probability that the attacker is captured by the defender is 5.36%.

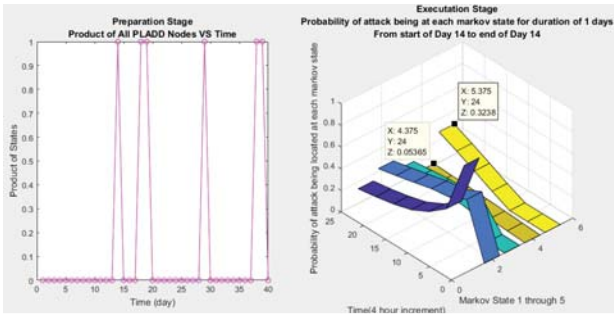

Figure 4. (a) The preparation stage for day 1 through 40 is shown on the left. (b) The first execution stage happens on the 14th day and the corresponding attack propagation is shown on the right.

IV. Discussion and Conclusion

Our experimental results show that the hybrid attack model is capable of modeling long periods of attacker and defender interactions from the attacker's point of view. From the attacker's point of view, whenever information in the preparation stage is revoked by the defender, the attacker cannot keep attacking the power grid or stay in the execution stage. The attacker must gather the necessary information in the preparation stage to start another attack in the execution stage. HAM is also able to simulate time frames where the attacker is capable of attacking the power grid. The time frame where the attacker is allowed to attack in the execution stage is dependent upon the defender's schedule to periodically revoke the attacker's access to the necessary information in the preparation stage. From the security analysts or defender's point of view, it may be possible to mitigate against the attacker by scheduling the defender actions such that there is always one PLADD node being taken back control by the defender. By making sure that the attacker does not have access to all the information needed to start the attack in the execution stage, the attacker will be forced to spend more time in the preparation stage to gather valid information.

References

[1] V. Chukwuka, Y. Chen, S. Grijalva and V. Mooney, "Bad Data Injection Attack Propagation in Cyber-Physical Power Delivery Systems," *2018 Clemson University Power Systems Conference (PSC)*, Charleston, SC, USA, 2018, pp. 1-8.

[2] S. Jones, A. Outkin, J. Gearhart, J. Hobbs, J. Siirola, C. Phillips, S. Verzi, D. Tauritz, S. Mulder, and A. Naugle. Evaluating moving target defense with pladd. Technical report, Sandia National Labs-NM, Albuquerque, 2015.

[3] Y. Chen, T. Gieseking, D. Campbell, V. Mooney and S. Grijalva, "A Hybrid Attack Model for Cyber-Physical Security Assessment in Electricity Grid," *2019 IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, USA, 2019, pp. 1-6.