SAND2019-0230C

Power Grid Bad Data Injection Attack Modeling in PRESTIGE

Yu-Cheng Chen, Dustin Campbell, Vincent Mooney and Santiago Grijalva
Georgia Institute of Technology, Atlanta, Georgia 30318
ychen414@gatech.edu, djrox316@gatech.edu,
mooney@ece.gatech.edu, sgrijalva@ece.gatech.edu

Brandon Eames, Alexander Outkin, and Eric Vugrin
Sandia National Laboratories
Albuquerque, New Mexico 87185
bkeames@sandia.gov, avoutki@sandia.gov, edvugri@sandia.gov

***Abstract*— The coupling between information, communication and computing elements with the physical components of critical infrastructure such as the electric power grid introduces new cyber and cyber-physical security concerns. Addressing these concerns requires novel methods that complement legacy and existing security solutions. Attacks such as bad data injection can cause disruptions that transcend the cyber realm and affect the physical world. This paper introduces a novel way to analyze security risk in critical infrastructure. We use the PRESTIGE Tool Chain developed by Sandia National Laboratories to model a bad data injection attack scenario in the power grid infrastructure. Analysis performed by this tool can guide the operator at the control center to take appropriate action to minimize disruption of the physical power system operation due to a bad data injection attack and to assess risk in order to optimize the use of resources to minimize the probability of a successful attack.**

Keywords—Bad Data Injection, Attack Graph, Cyber-Physical System, Cyber-Physical Security

I. Introduction

Recent cyber-physical attacks on electric power grids reinvigorate the drive by the research community, government, and industry to assess power grid vulnerabilities and the need to develop and deploy mechanisms to minimize the risk of cyber-attack. The cyber-attack on Ukraine's power grid, the cyber-attack on Burlington Electric, and the Stuxnet malware attack on electrical equipment powering Iran's Nuclear Technology Center [1] have brought more attention to vulnerabilities, threats and impacts of cyber-attack on critical power delivery systems.

To develop mechanisms that can detect and mitigate the effects of these attacks, modelling and simulation of how these attacks propagate through a cyber-physical power system need to be undertaken. We demonstrate a novel framework that leverages game theory to analyze a model of an attempt to inject malicious data into the power grid network. The goal of the attack is to cause the system operator to issue incorrect but legitimate commands which result in loss of electrical power and/or serious damage to power grid infrastructure.

II. Background

A. Bad Data Injection

Bad data injection (BDI) attacks have drawn widespread concerns in cyber-physical power grids and were first proposed by Liu and Ning [2]. Liu showed that attackers can manipulate field measurements to introduce bad data into certain state variables and bypass the existing techniques for bad measurement detection in power systems by exploiting their knowledge of the power system topology. Liu et al. proposed a bad data detection method based on adaptive partitioning state estimation which can raise the detection sensitivity by dividing the global power system into several subsystems [3]. Bad data then can then be identified even in a small area by multiple rounds of partitioning. In prior work by Chukwuka et al. [4], a graph-based attack propagation model was introduced that combines a Markov attack graph model together with state estimation to simulate attack propagation in an electric power grid scenario, shown in Fig 1.

B. Attack Graph Modeling

There are three prior modeling approaches relevant to this work as previous models were used to assess the impact, risk, and resilience of the power system. First [5] presents a model based on the topology of the electrical power grid network and utilizes a deterministic approach to calculate the loss of the power system given the number of compromised substations. Second, a statistical model uses multiple probabilistic simulations to evaluate which disruptions and mitigations are most impactful and thereby provides approximate risk assessment measurements [6]. The third modeling approach is a hybrid utilizing both topological information and statistical network descriptions to evaluate the performance and resilience of a specific power system [7].

C. Probabilistic Learning Attacker, Dynamic Defender (PLADD) and PRESTIGE

Jones et al. developed a game theoretic model, PLADD, to analyze the impact of moving target defenses [8]. In PLADD, a single attacker and defender contend for a resource. Attack dynamics is modeled as a stochastic process. The time required for the attacker to gain control is a random variable. The attacker "learns" from a successful attack, shortening the average time to complete subsequent attacks. The defender can regain the resource with (i) a lower cost "take" move that does not lessen the attacker's knowledge level, or (ii) a higher cost "morph" move that in addition undoes attacker learning. The defender has no information about attacker progress and must decide (i) when to act and (ii) which move to use to maximize defender utility and minimize attacker utility. "Utility" for both players is defined as the cumulative duration of controlling the resource minus costs.

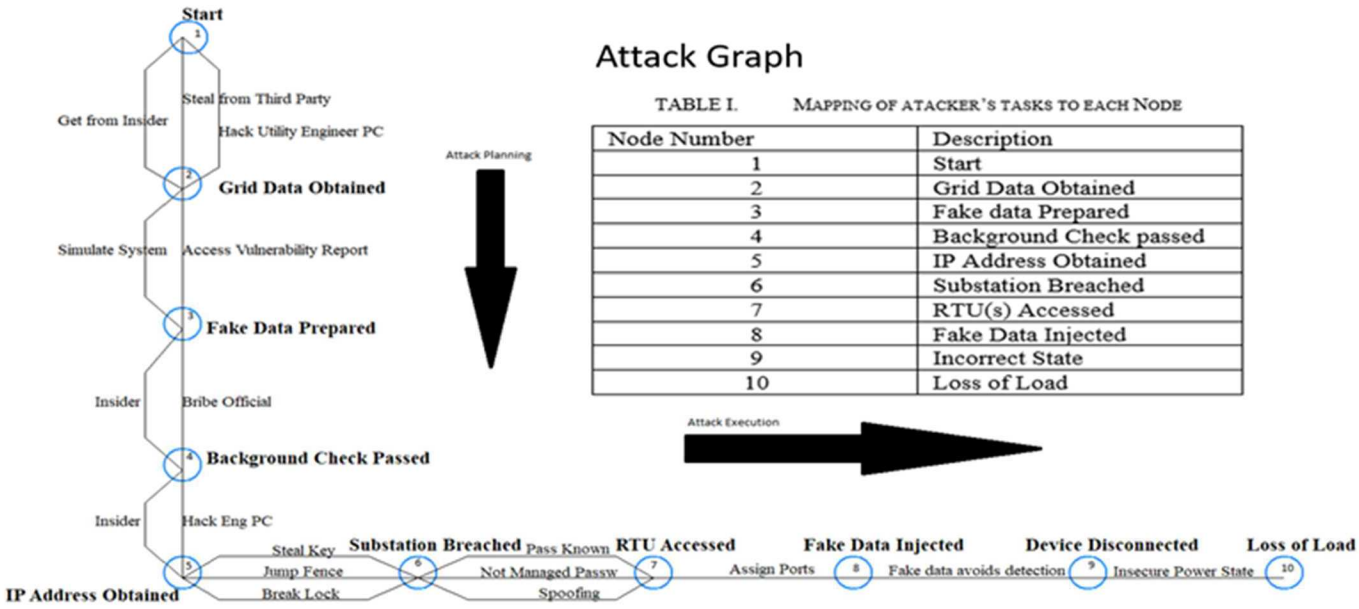

TABLE I. MAPPING OF ATACKER'S TASKS TO EACH NODE

Node Number	Description
1	Start
2	Grid Data Obtained
3	Fake data Prepared
4	Background Check passed
5	IP Address Obtained
6	Substation Breached
7	RTU(s) Accessed
8	Fake Data Injected
9	Incorrect State
10	Loss of Load

Figure 1. Attack graph capturing attacker's strategy

PRESTIGE (PRactical Evaluation and Synthesis of Trust In Government systEms) incorporates a system of PLADD games to model complex attack scenarios [9]. In PRESTIGE, system processes are captured as a set of producer-consumer processes (referred to as Actors), coupled to entities, which are the objects of production (referred to as Artifacts). Each Artifact is held under one or more access controls. PRESTIGE allows the specification of attack graphs, which represent sequences of specific steps to undermine system processes to achieve a nefarious outcome. Attackers must undermine access controls in order to perform certain attack actions. PRESTIGE incorporates PLADD to model ongoing contention between defender and attacker for "ownership" of an access control, where attacker ownership implies attacker access and defender ownership implies denied attacker access. Since attacks involve multiple artifacts, each of which potentially maps to different access controls, PRESTIGE evaluation of an attack scenario potentially incorporates multiple PLADD games. The current version of PRESTIGE does not incorporate "morph" moves. PRESTIGE also captures attacker actions performed after one or more access controls have been circumvented. Such actions could include file access, system modification, data injection, log manipulation, etc.

PRESTIGE implements a discrete event simulation of attacker-defender interactions. Each attacker and defender step is modeled as a cost-bearing, potentially time-consuming move. All time parameters are modeled as probability distributions, to capture the variance in time required for both attack and defense operations. If the defender is able to successfully complete all steps in her graph, while simultaneously preventing the attacker from completing his objectives, the defender wins the game. Whereas, if the attacker is successfully able to achieve his objective before the defender completes her game-specific objectives, the attacker wins. PRESTIGE uses Monte Carlo simulation to compute the likelihood of attack success and costs incurred by defender and attacker. PRESTIGE analysis facilitates computation of risk for individual attacks, followed by a relative ranking of risk associated across multiple attacks. In this paper, we focus on a single, multistep attack.

The authors are not aware of prior work attempting to model BDI power grid attacks using multiple PLADD games.

III. ATTACK SCENARIO

We focus our efforts on a power grid BDI attack. In our scenario, we assume that the attacker has sufficient knowledge to understand the power grid such as (i) electrical conditions when the grid may be stressed or specific electrical actions which may cause the grid to become unstable and (ii) computing technology used to control and protect the grid. However, in order to carry out a malicious action, the attacker must obtain detailed knowledge of the power system topology as well as knowledge of the TCP/IP addresses of the monitoring system and the remote terminal units (RTUs) at the substations [4]. Once the relevant background information is obtained, a specific day is chosen to carry out the attack. Fig. 2 shows in the upper right-hand corner a conceptual view of this attack. On the specific day, the attacker breaks into a room containing the RTU and communication radio. The attacker connects a computing device to the Human Machine Interface (HMI) and accesses the RTU(s). The attacker compromises data sent back to the monitoring system by constructing packets containing fake field readings and transmits these fake readings to the monitoring system on the TCP listening port for the monitoring system. The step-by-step attack graph is shown in Fig. 1. Attack planning is shown on the left-hand side of Fig. 1 (nodes 1 through 5) and may occur over the course of several months or longer. Attack execution steps are shown in nodes 6 through 10 and typically would occur on a single day. The power grid infrastructure relationships are shown in Fig. 2. The goal of the attack is malicious: the result in node 10, if reached, indicates that the power grid operator was fooled by the fake data and issued command(s) resulting in loss of power load, e.g., a partial blackout.

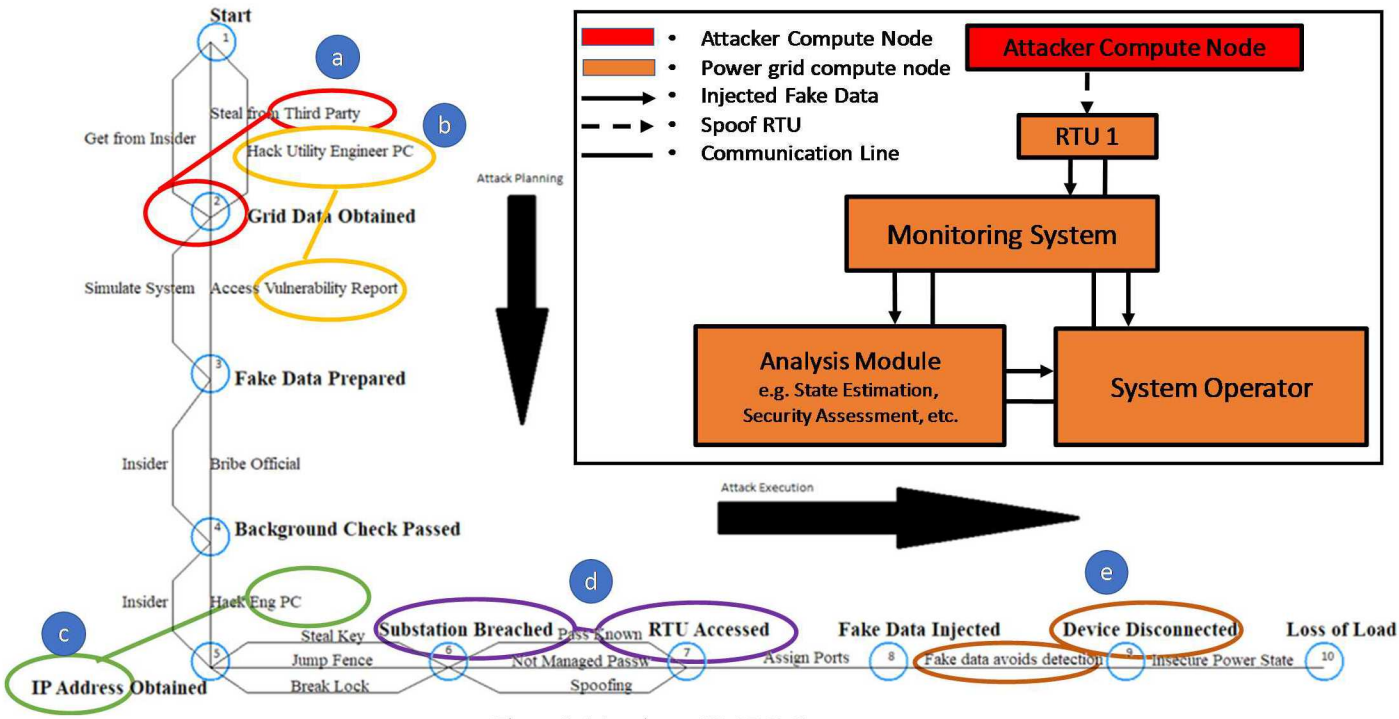

Figure 2. Mapping to PLADD Games

IV. Game Theoretic Model

We aim to utilize PRESTIGE to create a game theoretic model for the BDI power grid attack. To do so, we transition from a Markov model to a PLADD-inspired model. Our previous Markov-based attack graph is shown in Fig. 1 [4]. We analyze Fig. 1 and map the scenario to five distinct PLADD games as shown in Fig. 2 by the solid blue circles with letters 'a' through 'e.' Each game has a resource over which the attacker and the defender contend for control with corresponding actions to influence the desired control.

Game 1: Blue circle 'a' in Fig. 2 refers to a game circled in red where the attacker steals power grid data. The grid data is assumed to be stored at a third-party data repository. The attacker's goal is to gain access to the grid data by running password cracking software on the third party's computer. The defender may counteract attacker access to the grid data in several ways including by physically changing the grid topology, which causes the power flow in the grid to change rendering past grid data useless as the physical relationships (power flow relationships) from the past are now invalid.

Game 2: Blue circle 'b' in Fig. 2 refers to a game circled in yellow where a PC is hacked to obtain an electric vulnerability report. This report refers to a document that is developed using electric grid simulation software in real-time, which describes (i) the current topology and electrical conditions of the system such as margins to operational limits that operators must take into consideration, (ii) the events or actions that can take place in the system that would cause significant impact such as partial blackouts or equipment damage, and (iii) worst-case conditions of the grid such as times of peak demand or forecasted high demand. Electric grid engineers conduct these simulations to guide the day- to-day secure electric operation. We assume here that the electric vulnerability report is produced daily. The electric vulnerability report is stored on a computer in the control center. The attacker can gain access to the electric vulnerability report by running password cracking software to hack into the computer storing it. This computer is assumed to be behind a corporate firewall and to be connected to the power grid internal network. The defender can invalidate the information contained in a stolen vulnerability report by altering the grid state so that it has higher electric security with respect to known critical events, e.g., one specific action may be to bring additional generators online to provide higher reserve in case of a loss of generation.

Game 3: Blue circle 'c' in Fig. 2 refers to a game circled in green. IP address information is stored on another computer in the control center. The attacker can learn the required IP address information by running password cracking software to gain access to the computer storing the IP address. Our model assumes that the level of difficulty faced by the attacker to gain access to the computer storing the IP address information is commensurate with the difficulty of accessing the computer storing the electric vulnerability report. To invalidate stolen IP address information, the defender must change the IP addresses assigned to the RTUs and the passwords protecting the addresses.

Game 4: Blue circle 'd' in Fig. 2 refers to a game corresponding to the actions circled in purple. Physical access protection is provided by the substation in which the RTU is stored. The substation has fences and a locked door which restricts access to the room containing the RTU. The attacker can gain access to the area by jumping the fence and breaking the lock. The defender can regain control of the area by sending a utility engineer to check out the substation, causing the attacker to leave or risk discovery or arrest.

Game 5: Blue circle 'e' in Fig. 2 refers to a game corresponding to the two adjacent circled actions. State estimation can be used to mitigate bad data obtained during grid operations due to communication errors or faulty sensors. The

adversary gains control of the state (estimated by the state estimator) by sending well-crafted bad data that avoids detection from the bad data detection algorithm (chi-squared test) [2,3,5]. The defender regains control of the state by resetting the state (estimated by the state estimator) to a known good state by inserting known good data (as a replacement for the bad data detected by the state estimator).

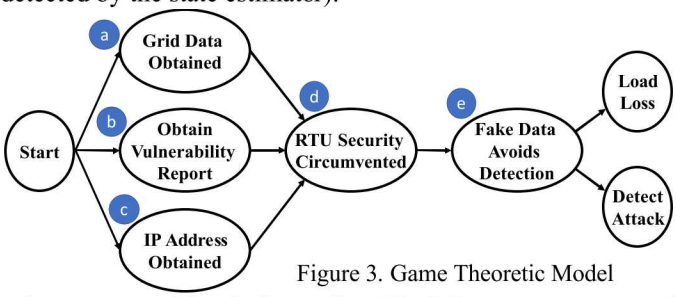

Figure 3. Game Theoretic Model

In summary, Fig. 2 shows five PLADD games as mapped from the prior Markov-based attack model. Resources from Fig. 1 are paired with actions that can be used by an attacker to obtain the resources. Games 1, 2 and 3 identified in Fig. 2 are within the planning phase of the attack and can execute in any order. Games 4 and 5 are in the execution phase and must be in sequence.

Fig. 3 shows our game theoretic model of the bad data injection attack. The five games identified in Fig. 2 (blue circles 'a' through 'e') are shown in Fig. 3 as PLADD games. The left-most node in Fig. 3 is simply the beginning or start node. The last two nodes on the right of Fig. 3 are the two possible endings; they do not denote games.

V. Implementation of Bad Data Injection Attack Scenario in Prestige Tool Chain

The PRESTIGE tool chain [9] enables the defined attack scenario to be implemented and tested as a series of PLADD games. These tests generate results for multiple defender strategies for each game, and from this data it is possible to determine the most cost-efficient strategies, which games are the most vulnerable to attack, and the influence a particular game has on the overall attack.

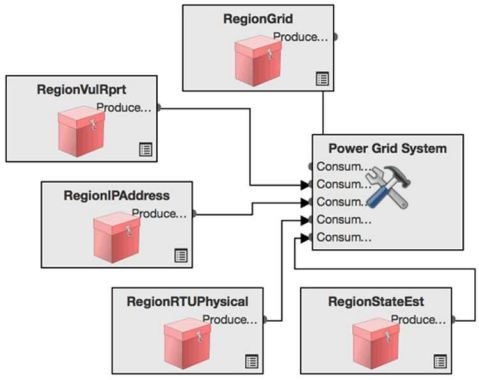

Figure 4. PRESTIGE Model of Power Grid System

Fig. 4 depicts the power grid system as modeled in the PRESTIGE tool. Each pink box (e.g., RegionVulnRprt) models a specific access control region, where the artifacts contained in a region (not shown) are subject to a set of access controls protecting that specific region.

Fig. 5 shows a model of the processes used to generate and consume the electric vulnerability report by the power grid provider. These structures are internal to the RegionVulnRprt model depicted in Fig 4. In Fig. 5, the box labeled "Vulnerability Report" is of type Artifact and is subject to the set of access controls protecting the RegionVulnRprt region. The access controls are modeled separately and are each characterized with cost and delay parameters modeling the attacker-defender game. For this specific model, each region depicted in Fig. 4 is associated with exactly one distinct access control. Each region in Fig. 4 contains a model similar to that of Fig. 5, with a single artifact modeling a resource of interest to the attacker.

Figure 5. Contents of the RegionVulnRprt block, modeling electric vulnerability report generation and consumption

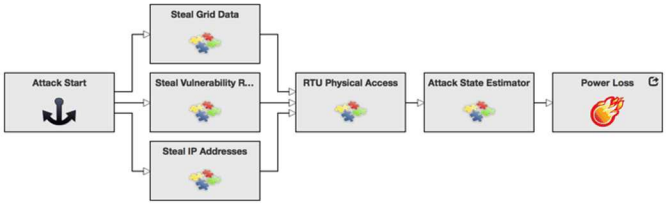

Figure 6. PRESTIGE model of the bad data injection attack

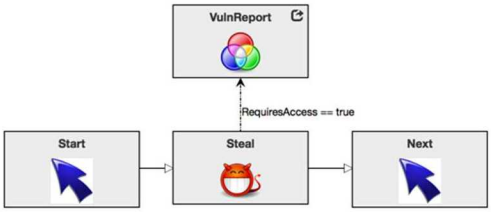

Figure 7. PRESTIGE model of the Steal Electric Vulnerability Report phase of the bad data injection attack

The BDI attack is captured in PRESTIGE as a sequence of attacker operations, each of which attempts to gain the ability to use an access-controlled resource. Fig. 6 depicts the PRESTIGE model of the BDI attack. Each interior node of the graph in Fig. 3 maps to an attack stage in Fig. 6. Fig. 7 depicts the contents of the Steal Vulnerability Report phase of the attack. The "Start" and "Next" nodes represent, respectively, the commencement and completion of this attack phase. The box labeled "Steal" represents the step in which the attacker attempts to gain access to and exfiltrate the electric vulnerability report. The box labeled "VulnReport" is a pointer to the "Vulnerability Report" artifact from Fig. 5, and represents the intended target or "context" of the Steal operation. The dashed line connecting the Steal to its context depicts an attribute, labeled "RequiresAccess=true". This attribute allows the modeler to specify whether the attacker must circumvent any access controls under which the target artifact is held. If RequiresAccess were set to false, this would imply that the attacker has insider access to the access controls protecting the RegionVulnReport region and therefore does not need to explicitly circumvent them to access the vulnerability report. However, in this model, we represent an outsider attacker without implicit access. Consequently, the attacker must first circumvent access controls prior to access and exfiltration of the target artifact. Each Steal node is assigned two attributes, a delay and a cost. The delay models the time required to execute the

attack operation, once access controls have been circumvented. In the case of the vulnerability report, this represents the time required to discover the location of the report on the filesystem and to then exfiltrate the file into attacker control. The cost parameter on the Steal node represents the cost the attacker incurs when performing the steal operation. Note that the cost and delay associated with the Steal operation is separate from costs and delays incurred by the attacker while circumventing access controls. PRESTIGE models access control contention separate from the attack steps.

In PRESTIGE, contention for ownership of access controls is modeled as a set of PLADD games. Each PLADD game is treated as an independent, concurrent game, with a state variable that tracks which party is currently in control of the resource. An attack step (e.g., stealing the vulnerability report) can only proceed when the state of the PLADD game modeling the contention for ownership of the access control protecting the target artifact indicates attacker ownership. The attacker must maintain ownership of the access control for the time required to perform the individual attack operations that require access to the entity in question. In the case of the electric vulnerability report, the attacker must maintain control of the vulnerability report's access control for enough time to acquire and exfiltrate the vulnerability report from the defender network. If the defender happens to execute a move on the access control protecting the electric vulnerability report after the attacker gains access to the report, but before the attacker is able to completely exfiltrate it, the attack operation fails to complete and must restart, forcing the attacker to delay until the attacker regains control of the corresponding access control.

The PRESTIGE model of the BDI attack consists of five distinct attack phases as shown in Fig. 6. Each phase is structured similar to the model depicted in Fig. 7, consisting of a Steal node targeting an artifact defined in one of the regions in Fig 4. The PRESTIGE modeling approach potentially allows the capture and simulation of additional detail associated with each attack phase. However, for the purposes of this paper, we abstract the attack into a set of "steal" operations coupled with PLADD games modeling contention for access controls.

Table 1. PLADD Parameters in PRESTIGE

Games	α	β	τ	μ	C
Grid Data	0.2	0.02	60	.25	200
Electric Vulnerability Report	0.4	0.04	60	.25	200
IP Address	0.4	0.04	1	.25	1
RTU Security	0.2	0.02	6	.00139	2
Fake Data Avoids Detect	0.1	0.01	.00984	.0171	2

Table 1 shows the assigned parameters for the five PLADD games, each corresponding to an access control protecting a different region defined in Fig. 6. The parameters of the PLADD game determine cost and delay incurred by each party as they contend for ownership of the access control (or more specifically, as they contend for access to the artifacts protected by the access control). The columns are defined as follows: α represents a fixed cost incurred by the attacker at the beginning of an attack, β represents an ongoing attack cost per unit time, τ represents the time between defender-executed take moves, μ represents the mean time to success for an attacker, and C represents the cost of a take move (recall that defenders carry out take moves, thus C is a per-move cost incurred by the defender). We estimated these numbers based on expert knowledge of the power grid. The time-related parameters are expressed in months and the costs in USD.

For Game 1, Grid Data Obtained, the attacker cost to steal power grid data is $\alpha + \beta t$ (where t = time in minutes) = 0.2 + 0.02t. We measure cost in dollars, so, for example, if it takes the attacker a month (30 days) to obtain the Grid Data, the cost is $864 (recall as explained in Section IV that the attacker gains access to the grid data by running password cracking software). The cost to "take" back the Grid Data is C = 200 dollars. The time τ between take moves is 60 months. The average time μ to success for an attacker is 0.25 months or one week.

For Game 2, Obtain Electric Vulnerability Report, the attacker cost to steal the report is 0.4 + 0.04t. Note that this attacker cost is greater than Game 1 because we assume that Game 2's password is harder to crack. The cost to "take" back the Electric Vulnerability Report is C = 200 dollars (e.g., if an additional generator is brought online). The time τ between take moves is 60 months. The average time μ to success is 0.25 months.

For Game 3, IP Address Obtained, the attacker cost to obtain the IP address is 0.4 + 0.04t. This cost is the same as the Vulnerability Report because we assume that the IP address information is stored on a computer with a similar configuration. The cost to "take" back the IP address is only C = 1 dollar. The time τ between take moves is one month. The average time μ to success for an attacker is 0.25 months or one week.

For Game 4, RTU Security Circumvented, we are no longer in the planning stage. The attacker cost is 0.2 + 0.02t. The cost to "take" back the RTU is small – only $2 (C = 2 in Table 1) – but the time τ between moves is 6 months because the power grid substation containing the RTU is not frequently monitored. Given the prior attack planning, the average time μ to success for an attacker is 0.00139 months or one hour. The attacker jumps the substation fence, accesses the room containing the RTU and then uses the IP address and a fake RTU computational component to spoof the data measurements.

Finally, for Game 5, Fake Data Avoids Detection, the attacker cost is 0.1 + 0.01t which is very small because this attack is passive. The time between take moves is 0.00984 months (425 minutes). This value was calculated empirically using a power grid state estimator written in Matlab and sample RTU data including modifications that an attacker may be able to make to the data given a reasonable level of education (e.g., a Master's degree in Electrical Engineering and five years of working experience in the power grid industry). The average time μ to success for an attacker is 0.0171 months or 740 minutes which again is a rough estimate of how long on average it would take to fool the system operator (a human being looking at the power grid data in a control center) into taking an action based on the fake data resulting in loss of load.

The costs and delays for Games 1-5 represent the parameters associated with circumventing access controls protecting the artifacts in the PRESTIGE process model. A fully specified PRESTIGE model must also account for the cost and delay associated with the attack steps contained in the attack graph.

Specifically, the Steal operations, such as depicted in Fig. 7, each incur delay and cost for the adversary, in addition to the costs and delays associated with circumventing access controls. The bad data injection analysis focuses on access control

circumvention rather than the attack steps. Hence, for experimentation purposes, we assigned unity delay to each attack steal operation, and a cost of 100. A more detailed model could potentially alter these parameters to account for nuances associated with each attack step. However, we did not incorporate this level of detail in this paper.

VI. Simulation Results

The PRESTIGE model outlined in the previous section characterizing the power grid process and BDI attack was simulated by the PRESTIGE simulation tool. The simulation tool evaluates the attack graph and attempts to ascertain the likelihood of success for the attacker. When executing the attack for 20 repetitions, and then computing average results over each repetition, the PRESTIGE simulator computed attack success rate to be 45%. The simulator computed the average time for the defender to complete his goals to be 1105 time units. In those runs where the attacker was successful, the attacker achieved his goals after only 11.67 time units. The defender execution time is dominated by the delay time associated with the "consumption" actors. The PRESTIGE simulator imposes the requirement that every attack operation be associated with a delay of finite duration, and every actor in a process model be associated with a delay of finite duration. PRESTIGE currently does not allow analysis on a graph with "always executing" elements. Consequently, the power grid process model had to be structured to mimic an "always executing" process using long-running consumption actors. Each consuming actor (e.g., Apply Electric Vulnerability Report in Fig. 5) was set to have a delay parameter of 1000 time units, modeling a duration much longer than other defender and attacker activities.

Table 2. Average duration of attack on nodes in attack graph.

Attack Node	time spent
Grid Data	5.23
Electric Vulnerability Report	5.20
IP Address	610
RTU Security	0.09
Fake Data Avoids Detect	0.01

Table 2 depicts results obtained through simulation, showing where the attacker spent his or her time. An effective defense denies the attacker the opportunity to advance from one attack graph node to another, for long periods of time. The average total time of the attack ranges from 11.67 units for successful attacks to greater than 1100 for failed attacks (the attack never succeeds). The table shows that the attacker spends a disproportionate amount of time in the IP Address phase. This is due to the frequency with which the defender executes moves on that resource. Defense of other resources requires more extreme measures such as bringing extra power generation online, which cannot be performed as frequently as cyber defenses. Consequently, attackers, under this analysis, can more easily circumvent the corresponding access controls and hold valid resources for longer periods of time.

Each move executed by each agent (defender or attacker) incurs a cost. The average cost incurred by the attacker regardless of win or loss is 149,000. The average attacker cost when the attacker wins the game is 2830, and when the attacker loses is 269,000. For the defender, average overall cost is 234,000, and does not change when the defender wins or loses. This independence between defender incurred cost and the state of the game is a representation of the property of PLADD games that the defender is not able to observe the state of the PLADD game. Defender actions are not in reaction to attacker activity. Rather, we model a defender who must defend a power grid system without knowledge or ability to react to the attacker.

An observation based on cost is that when the attacker loses, the attacker incurs a significantly increased set of costs. When the attacker wins, the win comes quickly, and consequently, the attacker avoids incurring significant costs.

VII. Discussion and Conclusion

PRESTIGE has been used to represent and evaluate the difficulty faced by an adversary to attempt to use a bad data injection attack to undermine a power grid, with the goal of disrupting power distribution. The analysis demonstrates that simple cyber defenses applied to key phases of a network can frustrate would-be attackers. Specifically, the attacker was forced to spend a significant amount effort in acquiring valid IP addresses to power grid equipment. A defense involving frequent changes to IP addresses frustrates the would-be attacker.

PRESTIGE has been designed as a tool to perform relative risk assessment, allowing comparison of risk and prioritization of mitigations. In this example, we present a single attack graph and its associated analysis. However, this analysis does not yet answer whether mitigation of this specific bad data injection attack should be prioritized over other types of attacks. The PRESTIGE tools can be used to model and evaluate a variety of attacks, in similar fashion to what has been presented here. However, the identification of different attacks, and the application of PRESTIGE to compute and compare the corresponding risk, is left as future work.

References

[1] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Baeyens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Electr. Power Syst. Res.*, vol. 149, pp. 210–219, 2017.

[2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *Ccs*, vol. 14, no. 1, pp. 1–33, 2009.

[3] T. Liu, Y. Gu, D. Wang, Y. Gui, and X. Guan, "A novel method to detect bad data injection attack in smart grid," *2013 Proc. IEEE INFOCOM*, pp. 3423–3428, 2013.

[4] V. Chukwuka, Y.-C. Cheng, S. Grijalva and V. Mooney, "Bad Data Injection Attack Propagation in Cyber-Physical Power Delivery Systems," Power System Conference, Sep. 2018.

[5] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the North American power grid," *Eur. Phys. J. B*, vol. 46, no. 1, pp. 101–107, 2005.

[6] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, 2012.

[7] S. Arianos, E. Bompard, A. Carbone, and F. Xue, "Power grids vulnerability: a complex network approach," Chaos: An Interdisci-plinary Journal of Nonlinear Science, Vol. 19, Issue 1, March 2009.

[8] S. Jones, A. Outkin, J. Gearhart, J. Hobbs, J. Siirola, C. Phillips, S. Verzi, D. Tauritz, S. Mulder, and A. Naugl, "Evaluating Moving Target Defense with PLADD," Sandia National Laboratories, 2015.

[9] M. Galiardi, E. Vugrin, B. Eames, A. Outkin, G. Wyss, J. Hamlet, R. Helinski, B. Anthony, M. Napier, J. Eldridge, A. Bertels and M. Holmes, "On Modeling Detection for Quantitative Trust Analysis", GOMACTech 2018, Miami FL