Cyber-Air-Gapped Detection of Controller Attacks through Physical Interdependencies

Sriharsha Etigowni, Mehmet Cintuglu†, Maryam Kazerooni*, Shamina Hossain*
Pengfei Sun, Katherine Davis*, Osama Mohammed†, Saman Zonouz
Electrical and Computer Engineering Department
Rutgers University, †Florida International University, *University of Illinois
{*se260, ps807, saman.zonouz*}*@rutgers.edu*, {*mcint015, mohammed*}*@fiu.edu*, {*shossai2, kazeron2, krogers6*}*@illinois.edu*

***Abstract*—Trustworthy operation of the power grid critical infrastructures requires real-time intrusion detection systems to identify compromised and malfunctioning controller devices. The past three decades of direct application of the traditional purely-cyber security solutions against these infrastructures has proved insufficient in practice due to emerging sophisticated malicious attacks against power grid control systems. In this paper, we propose PhiDS, a physics-aware intrusion detection system to identify compromised controllers through continuous observation of remote power system sensor measurements. Real-time remote sensor data analysis enables PhiDS to determine the power system state trajectory and infer the control commands issued by the distributed controllers on the plant. Given the power system safety requirements, PhiDS monitors the data stream and identifies the controllers that issue control commands that violates the safety of the power system. PhiDS does not require any cyber communication with the (potentially compromised) controller devices, and hence provides an *air-gap* between the the security monitor and the target device. Consequently, if the controller is infected, the adversary cannot compromise and corrupt the monitor's reports. The will ensure that the monitor will always remain away from the adversaries' access and hence provide trustworthy reports. We implemented and evaluated PhiDS on a real-world power system test-bed, where the programmable logic controllers are targets for and attacked by the remote network adversaries. PhiDS was able to identify all the infected controllers efficiently without any cyber link to the controllers. PhiDS's outcomes were instead purely based on the power system measurements from sensors that are deployed adjacent to the controllers.**

I. Introduction

The December 23, 2015 attack on the Ukrainian Kyivoblenergo energy service provider showed that attack sophistication on power grid elements has reached catastrophic potential. Due to its complexity A lesser known lesson was about detection, as the malware (BlackEnergy 3) first killed the intrusion detection systems and anti-virus processes on the computers it infected so that the operators were not notified about the ongoing attack while it would maliciously open the underlying circuit breakers.

The eternal war between the adversaries and the security monitoring software on computing devices boils down to who owns the higher level of privileges on the system. That is the entity with the higher privileges can neutralize the capabilities of the other running entity with the same or relatively lower privilege. For instance, an anti-virus software process running with a root/admin privilege on a power grid controller device can be killed or corrupted by a malicious rootkit that has compromised a root-level vulnerability. Consequently, such host-based security monitors are always vulnerable to malware on the same system that can take the control over the monitoring process(es).

The alternative solution to host-based security monitors in purely cyber systems is network-based security intrusion detection systems. The network-based monitors sit on a network device, such as a router, and monitor the network traffic to identify the compromised controller devices. The main advantage of network-based monitors is they do not run on the same controller as the potential malware. Hence, if the controller gets infected, the malware on the computer has a smaller attack surface to target in order to corrupt router-based monitor. Even in these cases, there have been attacks against network intrusion detectors and firewalls reported [25]. However, there are two major problems in practice with network-based monitors. First, they often produce relatively inaccurate outcomes compared to host-based monitors that have access to the target controllers internal information, such as the running process list, their corresponding executables and the filesystem details. Second, network-based monitors cannot detect intrusions effectively in practical settings where the traffic payload is encrypted end-to-end, e.g., using secure socket layer (SSL) connections.

Nowadays, both host-based (e.g., NortonAV or Samhain) and network-based intrusion detectors (e.g., Bro or Snort), are used in practical power grid control networks despite their above-mentioned limitations in accurate attack detection. Another major limitations of existing intrusion detectors is that they purely consider the cyber-based system behavior to identify adversaries. In cyber-physical power grid infrastructures, there exists an additional source of runtime information that can be leveraged to ensure the safe behavior of the controller devices. Almost all the existing intrusion detectors totally miss the physical dynamics of the underlying power system to distinguish anomalous and malicious controller behaviors more accurately.

In this paper, we present PhiDS, a physics-aware intrusion detection solution for cyber-physical power grid infrastructures that performs identifies malicious compromised controllers without executing any software module on those controllers. Additionally, there is no cyber network communication channel (and hence a potential attack vector) between the (compromised) controller and the intrusion detector. Therefore, there exists (by design) no way for the malware on the compromised controller to connect and infect PhiDS intrusion detector. Consequently, PhiDS's reports will (can) never be corrupted. PhiDS does this by leveraging the intrinsic exposure of the controllers' behavior on the physical plant and the physics interdependencies between the plant's (distributed) components. The controllers' issued control inputs to its local actuators often cause global physical system state change that can be detected remotely by PhiDS. PhiDS deploys power sensors to monitor different (possibly remote) power system parameters that cab be directly or indirectly affected by the controllers' outputs.

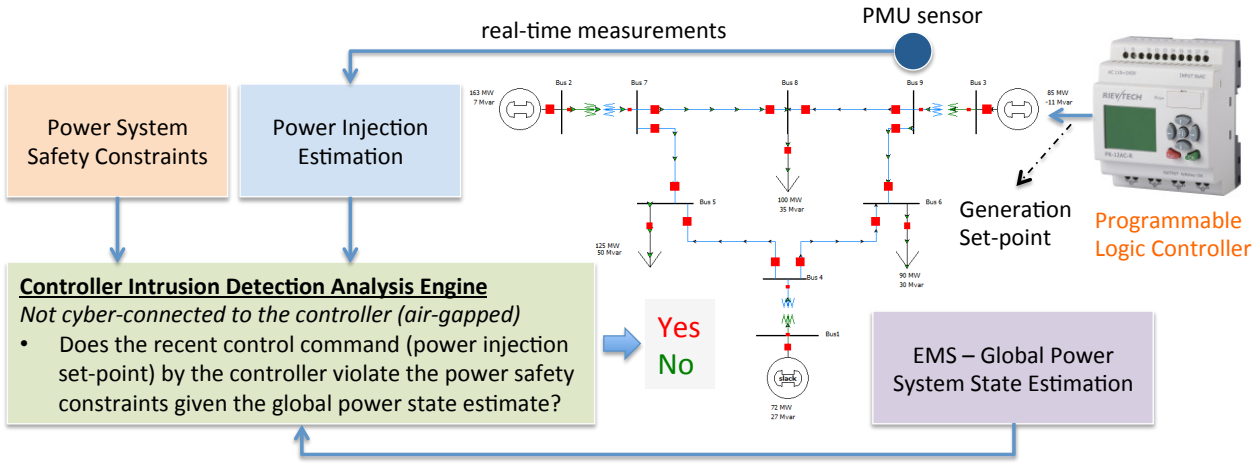

Fig. 1: PhiDS's High-Level Architecture - Cyber-Air-Gapped Controller Security Monitoring

To develop PhiDS, we leverage a power system-based sensitivity analysis that can be used as a communication channel between what the compromised controller and the PhiDS's detection engine. It is noteworthy that our solution is completely different from the traditional power line communication solutions, because we do not send any digital signal through the transmission lines. Instead, we leverage the physical inter-dependencies among the components and measure only the power system parameters such as bus voltages and line currents. We propose a complete analysis of how the malicious controller outputs to the power system can be sensed (received) by PhiDS's detection sensor remotely. Additionally, we present a formal logic-based representation of the power grid safety constraints and requirements. These requirements are used by PhiDS's detection sensors to identify and trigger alerts whenever the compromised controller drives the power system towards unsafe states through its malicious control inputs. PhiDS uses linear temporal logic formalism for accurate representation of the safety constraints with potentially temporal interdependencies such as in inter-locking requirements within a substation. Other simpler requirements like maximum generation set-points, bus voltage and line current capacities can be represented easily using the first-order propositional logic expressions as a part of linear temporal logic formula.

The contributions of this paper are as follows:

- We propose a power system-based remote security-oriented event monitoring solution that leverages the inter-component physical dependencies using linearized sensitivity matrices.
- We provide a novel formal logic-based description of the power system safety constraints that PhiDS uses to detect unsafe violations and controller compromises.
- We implemented and evaluate a real-world working prototype of PhiDS on a realistic power system environment. Our results are very promising and confirm the feasibility of PhiDS for practical deployments.

This paper is organized as follows: Section IV and Section III presents PhiDS s formal power system safety description and power system analyses in details. Section V provides the details of how PhiDS performs on a realistic power grid setup. We review the most related past work and how they fall short to achieve PhiDS s objectives in Section VI. Finally, Section VII concludes the paper.

II. PhiDS Design

We present PhiDS's high-level architecture, and how its achieves its objectives. As shown on the figure, PhiDS can detect infected malicious controllers without the need for a cyber connection to the controller. This creates a air gap between the intrusion detector (PhiDS) and the (possibly compromised) controller such that the attacker from the controller cannot infect PhiDS analysis engine under any circumstance.

PhiDS continuously receives inputs from three sources in order to determine whether the target controller has been infected by malware: *i)* PhiDS receives the power system safety constraints that need to be satisfied for the power system to maintain a safe global state. Please note that PhiDS does not require extra information in addition to what is already available within energy management systems (EMS). These safety constraints, e.g., the transmission line flow capacities, are currently used by the contingency analyses tools within the EMS. PhiDS creates a formal logic-based conjunctive linear temporal logic [7] expression ϕ that is composed of all EMS safety requirements; *ii)* PhiDS receives real-time power system state estimate $\hat{s}$ from the EMS state estimation server. PhiDS continuously determines whether the current state of the power system satisfies the above-mentioned safety requirements. To that end, PhiDS feeds the conjunctive temporal expression $\phi \wedge \hat{s}$ to a SMT satisfiability checker [4]. If the outcome is positive, it indicates the power system's current state satisfies the safety constraints. Otherwise, the constraints are violated at some point; *iii)* PhiDS receives from its local engine a continuous estimate of the controller's output (control command) to the power system. This could be done through either direct measurement of the controller's output, or indirect controller output estimation via measuring other remote system parameters that are affected by the controller's output. For instance, assume the controller's output determines the generation set-point on one of the generators. The real power measurement on the neighboring power bus by PhiDS can disclose the controller's output. To generalize such indirect inferences, we propose a power system-based security-oriented remote event monitoring analysis solution that enables PhiDS to infer the controller's output remotely based on the underlying global system-wide sensitivity matrices.

III. Remote Intrusion Detection

To recognize the physical plant manipulation details by the target controller, PhiDS makes use of the underlying power system reconfiguration and power flow model. PhiDS observes

its local power system parameters and tries to infer the control commands that the target controller sends to its local actuators. Needless to mention, if PhiDS has access to monitor the exact power system parameters that are controlled by the target controller, PhiDS does not need to perform any power system analysis. However, it is a cyber security practice to keep the intrusion detector as far as possible from the victim device (i.e., the target controller here) to minimize the possibility that the intrusion detector also gets compromised by the same attacker that targets the victim device.

To address the cases, where the power system parameters monitored by PhiDS's intrusion detection engine are different from the the system parameters controlled by the target controller, PhiDS leverages power system sensitivity analysis. Given a power system topology, PhiDS has to perform detailed offline analysis of the power system mathematical models to determine how various changes to the system by the target controller may affect the system parameters observed by PhiDS's intrusion detector. PhiDS obtains the current power system topology and the power flow models continuously from the energy management system within the control centers. Any perturbation of a particular system parameter by the target controller causes updates across other parameters such that all values will comply with the power flow equations.

An n-bus power system's dynamic behavior can be represented by parameterized differential equations [11].

$$\dot{x} = f(x, u, \lambda), \tag{1}$$

where f is a continuously differentiable function representing the physical plant's dynamic behavior; $x \in \mathcal{R}^{2n-1}$ represents the system state vector that includes the voltage magnitude and phase angles for each bus[1]; $u \in \mathcal{R}^m$ represents the plant's control input vector that can be manipulated by the target controller (e.g., generator's real power set point) monitored remotely by PhiDS's intrusion detection engine; λ represents a vector discrete events that change the plant's topology, and hence its continuous differential equations. Because of such intervention of discrete and continuous dynamics, the power plants are considered as hybrid systems [13].

$$\lambda^{k+1} = \Lambda(x, u, \lambda^k), \tag{2}$$

where, Λ accounts for the topological evolution of the plant due to operation of controllers (including the target controller), protective devices (e.g., relays), and transformer tap changes. Given the power system's state vector, all other system parameters, such as transmission line impedances and loads, can be obtained using the power plant's equations. The sensor measurements are correlated with the plant state and the operator's control inputs through

$$w = h(x, u), \tag{3}$$

where w is the sensor measurement vector, and h is called the measurement function. For a given plant topology and system parameters, we have

$$f(x_0, u_0, \lambda_0) = 0 \tag{4}$$

at the plant's exponentially stable[2] equilibrium state x_0 [13].

To perform security-oriented remote event detection and monitoring, PhiDS implements power system sensitivity analysis to determine how much each possible control command by the target controller affects the power system global state and specifically the power system parameters that are constantly monitored by PhiDS's intrusion detection engine. All these parameters are correlated through the power flow models that are used by PhiDS's sensitivity analysis engine. PhiDS investigates the plant's power flow equations given any stable point x_0 and calculates the margin by which each system parameter changes due to physical dependencies if the target controller applies a particular control input to the system.

To start its analysis, PhiDS obtains the underlying power system state estimate from the energy management system's state estimation server that maintains the most recent state estimate based on the past sensor measurements such as phasor measurement units. Given the power system state value, PhiDS alters the system parameters around the equilibrium point. To increase its intrusion detection perform for real-time event monitoring, PhiDS linearizes the power system model before performing the sensitivity analysis. PhiDS computes the plant's Taylor-expanded model around the plant's current state. While higher order derivatives could be computed at the cost of increased overhead, low order approximations proves to be accurate in our real-world power system test-bed experiments.

PhiDS implements its cyber-air-gapped intrusion detection analysis of the target victim controller through dynamic behavior inspection of the plant around Equation 1's equilibrium state using the plant's Taylor approximate equivalent.

$$\begin{aligned} f(x_0 + \Delta x, u_0 + \Delta u, \lambda_0 + \Delta\lambda) \approx f(x_0, u_0, \lambda_0) + \\ + f_x \Delta x + f_u \Delta u + f_\lambda \Delta\lambda, \end{aligned} \tag{5}$$

which uses the first-order partial derivatives (Jacobian matrix) of the power plant's vector-valued function $f_x = \frac{\partial f}{\partial x}(x_0, u_0, \lambda_0)$, $f_u = \frac{\partial f}{\partial u}(x_0, u_0, \lambda_0)$, and $f_\lambda = \frac{\partial f}{\partial \lambda}(x_0, u_0, \lambda_0)$. Assuming that f_x is non-singular, we can reorder Equation 5 as follows

$$\Delta x = -f_x^{-1} f_u \Delta u - f_x^{-1} f_\lambda \Delta\lambda, \tag{6}$$

which formulates how the power plant's state changes every time the target controller modifies an actuation point. Equation 6 shows the parameteric value correlation between the target controller's actuation points and the state variables that are partially monitored by PhiDS's intrusion detection engine. PhiDS determine how each PhiDS's sensor measurement is affected as the result of a control input application by the target controllers anywhere in the system. Following Equation 3's first order Taylor expansion at (x_0, u_0), gives us

$$\Delta w = [w_u - w_x f_x^{-1} f_u] \Delta u, \tag{7}$$

where the changes in measurements Δw are calculated as the result of any change in the system Δu. For higher intrusion detection accuracy via considering higher order dynamics of the plant, PhiDS makes use of second order approximation

$$\begin{aligned} \Delta w = (w_u - w_x f_x^{-1} f_u) \Delta u + \\ + \Delta u^T (\frac{1}{2} (f_x^{-1} f_u)^T w_{xx} f_x^{-1} f_u - w_{ux} f_x^{-1} f_u + \frac{1}{2} w_{uu}) \Delta u, \end{aligned} \tag{8}$$

where the Jacobian matrices used for PhiDS's sensitivity analysis (Equation 8) are shown in Figure 2. PhiDS performs security-oriented remote power system sensitivity analysis for the plant's current state online during the system operation, since it requires the real-time state information. The analysis is sufficiently fast for practical control system settings as demonstrated in our evaluation. To further improve the framework's overall performance, PhiDS can be expanded to leverage the existing power plant's characteristics, i.e., their fairly well-defined behavior in terms of daily load patterns

[1] Except the reference (slack) bus angle that is assigned zero.

[2] Exponentially stable systems' convergence, after a perturbation, is bounded by exponential decay. A continuous linear time-invariant system is exponentially stable iff the system has eigenvalues with strictly negative real parts [13].

Jacobian matrix $w_u = \begin{bmatrix} w_{u_c} \end{bmatrix}$

$$\frac{\partial w_l}{\partial Z_{cl'}} = \begin{cases} \gamma_l (V_i^2 - V_i V_j \cos\theta_{ij}) - \beta_l V_i V_j sin\theta_{ij} & l = l' \\ 0 & l \neq l' \end{cases}$$

Jacobian matrix $F_u = \begin{bmatrix} P_{z_c} \\ Q_{z_c} \end{bmatrix}$

$$\frac{\partial P_i}{\partial Z_{Cl}} = \begin{cases} \gamma_l (V_i^2 - V_i V_j \cos\theta_{ij}) - \beta_l V_i V_j sin\theta_{ij} & i \neq SLK, l \in S^{L(i)} \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial Q_i}{\partial Z_{Cl}} = \begin{cases} \beta_l (-V_i^2 + V_i V_j \cos\theta_{ij}) - \gamma_l V_i V_j sin\theta_{ij} & i \in S^{PQ}, l \in S^{L(i)} \\ 0 & \text{otherwise} \end{cases}$$

Jacobian matrix $w_{ux} = \begin{bmatrix} w_{Z_c\theta} & w_{Z_cV} \end{bmatrix}$

$$\frac{\partial w_{Z_c}}{\partial \theta_i} = \begin{cases} \gamma V_i V_j sin\theta_{ij} - \beta_l V_i V_j \cos\theta_{ij} & i \neq SLK, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial w_{Z_c}}{\partial \theta_j} = \begin{cases} -\gamma_l V_i V_j \sin\theta_{ij} + \beta_l V_i V_j \cos\theta_{ij} & i \neq SLK, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial w_{Z_c}}{\partial V_i} = \begin{cases} 2\gamma_l V_i - \gamma_l V_j \cos\theta_{ij} - \beta_l V_j \sin\theta_{ij} & i \in S^{PQ}, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial w_{Z_c}}{\partial V_j} = \begin{cases} -\gamma_l V_i \cos\theta_{ij} - \beta_l V_i \sin\theta_{ij} & i \in S^{PQ}, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

Jacobian matrix $w_{xx} = \begin{bmatrix} w_{\theta\theta} & w_{vv} \end{bmatrix}$

$$\frac{\partial^2 w_l}{\partial \theta_i^2} = \begin{cases} g_l V_i V_j \cos\theta_{ij} + b_l V_i V_j \sin\theta_{ij} & i \neq SLK, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2 w_l}{\partial \theta_j^2} = \begin{cases} g_l V_i V_j \cos\theta_{ij} + b_l V_i V_j \sin\theta_{ij} & i \neq SLK, i, j \in l \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2 w_l}{\partial V_i^2} = \begin{cases} 2g_l \\ 0 \quad \text{otherwise} \end{cases}$$

$$\frac{\partial^2 w_l}{\partial V_j^2} = \begin{cases} 0 \quad i \in S^{PQ}, i, j \in l \\ 0 \quad \text{otherwise} \end{cases}$$

Jacobian matrix $w_{uu} = \begin{bmatrix} w'_{z_c} \end{bmatrix}$

$$\frac{\partial w_l}{\partial Z_{Cl'}} = \begin{cases} \gamma'_l (V_i^2 - V_i V_j \cos\theta_{ij}) - \beta'_l V_i V_j sin\theta_{ij} & l = l' \\ 0 & l \neq l' \end{cases}$$

Jacobian matrix $w_x = \begin{bmatrix} w_\theta & w_v \end{bmatrix}$

$$\frac{\partial w_l}{\partial \theta_i} = \begin{cases} g_l V_i V_j \sin\theta_{ij} - b_l V_i V_j \cos\theta_{ij} & i \neq SLK, i, j \in l \\ o & \text{otherwise} \end{cases}$$

$$\frac{\partial w_l}{\partial \theta_j} = \begin{cases} -g_l V_i V_j \sin\theta_{ij} + b_l V_i V_j \cos\theta_{ij} & i \neq SLK, i, j \in l \\ o & \text{otherwise} \end{cases}$$

$$\frac{\partial w_l}{\partial V_i} = \begin{cases} 2g_l V_i - g_l V_j \cos\theta_{ij} - b_l V_j \sin\theta_{ij} & i \in S^{PQ}, i, j \in l \\ o & \text{otherwise} \end{cases}$$

$$\frac{\partial w_l}{\partial V_j} = \begin{cases} -g_l V_i \cos\theta_{ij} - b_l V_i \sin\theta_{ij} & i \in S^{PQ}, i, j \in l \\ o & \text{otherwise} \end{cases}$$

Notations:

n_l	number of transmission lines
l, l'	indices for transmission lines
z_l	series impedance of line l
y_l	series admittance of line l
Z_c	series capacitive resistances
P	active power injections at all nodes except slack node
Q	reactive power injections at PQ-nodes
w	active power line flows
$S^{L(i)}$	set of lines connected to bus i
S^{PV}	set of PV-nodes
S^{PQ}	set of PQ-nodes
SLK	slack node

Definitions:

z_l	$r_l + j(x_l - x_{cl})$
y_l	$g_l + jb_l$
Z_c	${x_{cl}}^{nl}$
θ_{ij}	$\theta_i - \theta_j$
γ_l	$\frac{\partial g_l}{\partial x_{cl}} = \frac{2r_l(x_l - x_{cl})}{(r_l^2 + (x_l - x_{cl})^2)^2}$
β_l	$\frac{\partial b_l}{\partial x_{cl}} = \frac{-(x_l - x_{cl})^2 + r_l^2}{(r_l^2 + (x_l - x_{cl})^2)^2}$
γ'_l	$\frac{-2r_l(r_l^2 + (x_l - x_{cl})^2) + 6r_l(x_l - x_{cl})^2}{(r_l^2 + (x_l - x_{cl})^2)^3}$
β'_l	$\frac{2(x_l - x_{cl})(r_l^2 + (x_l - x_{cl})^2)^2 + 4(r_l^4 - (x_l - x_{cl})^4)(x_l - x_{cl})}{(r_l^2 + (x_l - x_{cl})^2)^4}$

Fig. 2: Physical-Side Sensitivity-Based Information Flow Analysis

and hence next possible system states. PhiDS can complete its intrusion detection analysis for the plant's next potential states proactively, and use the results instantaneously in case those states occur next.

IV. Formal Power System Safety Monitoring

A. Physical Safety Specification

To formulate power system safety requirements, PhiDS makes use of the linear temporal logic formalism [18]. Let us define A to be a finite set of atomic logical propositions about the system $\{b_1, b_2, \cdots, b_{|A|}\}$, e.g., `relay` R_1 `is open.` and $\Sigma = 2^A$ a finite alphabet composed of the above-mentioned propositions. Every element of the alphabet is a possibly empty set of propositions from A, and is denoted by a_i, e.g., $a_i = b_1, b_4, b_9$. As PhiDS deals with runtime verification of the past and current traces of the power system states, we define Σ^* to be all of the possible finite traces over Σ, e.g., $(a_0 a_1 a_2)$, where two subsequent events a_i and a_j are represented by symbolic concatenation $a_i a_j$. Similarly, Σ^ω is defined to be the set of infinite system traces.

The set of linear temporal logic-based security requirements is inductively defined by the grammar

$$\varphi ::= true \mid b \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \, \mathbf{U} \, \varphi \mid \mathbf{X} \, \varphi, \tag{9}$$

where φ is a logical predicate; $\mathbf{U}$ and $\mathbf{X}$ represent the temporal *until* and *next* operators, respectively. For safety description simplicity, PhiDS also makes use of the following three redundant notations: $\varphi \wedge \psi$ instead of $\neg(\neg\varphi \vee \neg\psi)$, $\varphi \rightarrow \psi$ instead of $\neg\varphi \vee \psi$, $\mathbf{F}\varphi$ (eventually) instead of $true\mathbf{U}\varphi$, and $\mathbf{G}\varphi$ (globally) instead of $\neg(true\mathbf{U}\neg\varphi)$.

For example, consider a power system substation system with Boolean variables r_1 and r_2 that activate (closes) the cir-

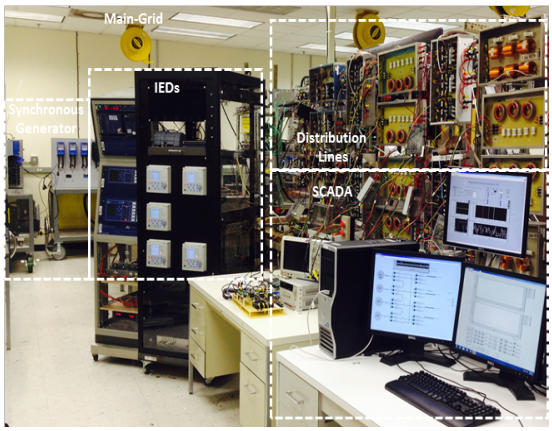

Fig. 3: Real-World Evaluation Test-Bed

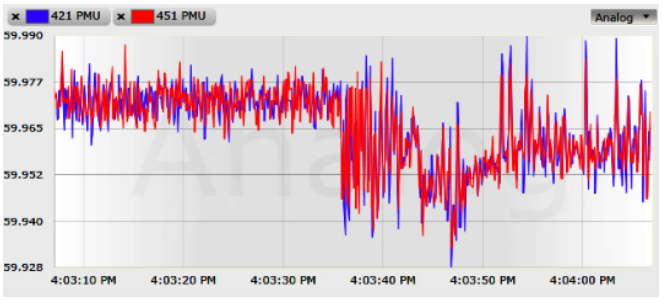

Fig. 4: Real-time System Frequency Measurements

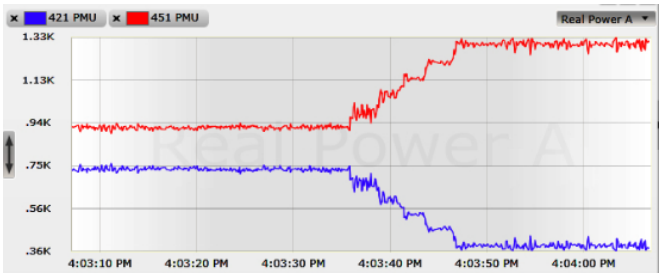

Fig. 5: Real-time Real-Power Measurements

cuit breakers for interlocking transmission lines when true. The property that both relays are never open at the same time has two atomic propositions: $a \equiv r_1 = false$, and $b \equiv r_2 = false$. The global LTL property is then stated $G\ \neg a\ \vee\ \neg b$.

It is important that we do not require definition of power system safety requirements from scratch. Instead, PhiDS makes use of the system safety constraints that are already available within energy management systems and typically used for "what-if" $N-1$ contingency analyses. Most of the existing safety requirements are individual atomic propositions that each can take on either true or false values at any time instant. Consequently, PhiDS creates a conjunctive linear temporal logic formula by combining all the existing contingency analysis safety requirements and composes a single formula for its runtime malicious controller detection as discussed below.

B. Runtime Malicious Controller Detection

Given the formally specified safety description for the power system, PhiDS's final objective is to determine if the power system is currently unsafe, and if so, whether the controller's output is responsible for such an unsafe state. The responsible controller is then identified as malicious that requires more indepth security/forensics analysis and clean up.

PhiDS receives its local power system measurements (Figure 1) and leverages the power system-based security-oriented remote event monitoring analyses (Section III) to determine the controller's output change (if there exists any). If the controller's output change is detected, and it is in the direction that worsens the situation, i.e., drives the system farther from the safe states, PhiDS triggers an alert indicating that the controller has been infected.

V. Evaluations

The forensic analysis are implemented in a reconfigurable small scale power system available at the state-of-the-art test bed at Florida International University, Smart Grid Test Bed (Figure 3). The test bed offers a unique platform with actual synchronous generators, transmission line models, and controllable active and reactive loads. IEEE 9 bus topology including three synchronous generators was implemented on the reconfigurable test bed for study model. One of the synchronous generators is operated as slack, while the others are in constant torque mode of operation.

PMUs operate using the IEEE Std. C37.118 synchrophasor communication protocol for measurements and data reporting. PMUs continuously reports to a PDC on a central computer. PMUs are also capable of protection functions as well as measurement. A PLC using Modbus protocol is connected as central operator which is responsible from all operational functions.

OPC drivers are used to establish a common information exchange platform from industrial equipment, which use different communication protocols. The Common Format for transient Data Exchange (COMTRADE), IEEE C37.111, is a common example of a digital data format to acquire forensic data. Most PMUs that capture high-speed data can export their data in the COMTRADE format. However, investigation of the complete data can be cumbersome, therefore only compromised measurements and data should be examined with a detailed forensic analysis.

In our experiments, we placed PhiDS local sensor (PMU) on the power bus that was adjacent to the controller's bus. Any change in the controller's output was detected by PhiDS's power sensors remotely. We considered the system frequency's safe range as the formal power system safety constraints.

We uploaded a malicious controller logic on the controller (programmable logic controller) that was responsible to determine the generation set-point on the generator 2. In this attack scenario, the legitimate controller code on the programmable logic controller should try to increase the generator 2 generation to relieve generator 1 slack units power export and reduce the power flow in inter-tie. However, the malicious controller code on the controller device is compromised and instead of increasing the generation of generator 2, it issues the control command to gradually decrease the generation set-point on generator 2. This situation would result instability as the generator 1 would reach to generation capacity. PhiDS was able to correctly identify the controller's output change and the fact that its change causes the power system instability over time. Figure 4 shows the power system frequency measurements, while its safety requirements were violated by the malicious controller's output to the generator 2. Figure 5

demonstrates the real power flow measurements by the two installed phasor measurement units (PMUs). PhiDS triggers the alert and identifies the compromised controller device calling for more in-depth forensics analysis of the controller by the security administrators.

VI. Related Work

Since the past real-world critical infrastructure attacks, there has been an increasing number of security protection solutions proposed. We review the most related work.

Control system safety. Stouffer et al. [21] present a series of NIST guideline security architectures for the industrial control systems that cover supervisory control and data acquisition systems, distributed control systems, and PLCs. Such guidelines are also used in the energy industry [22], [16]. It has, however, been argued that compliance with these standards can lead to a false sense of security [26], [17]. There have also been efforts to build novel security mechanisms for control systems. Mohan et al. [14] introduced a monitor that dynamically checks plant behavior safety. A similar approach using model based intrusion detection was proposed in [3]. Goble [9] introduce mathematical analysis techniques to quantitatively evaluate aspects of a control system such as safety and reliability, including PLC devices. However, the proposed solution focuses mainly on accidental failures and does not investigate intentionally malicious actions.

As a fundamental power system monitoring tool, state estimation is the process of fitting power sensor data to a system model [10] and determining the current system state [1], e.g., using weighted least squares [19]. The estimated state is then used in stability analysis [8] through solving nonlinear AC [2] or linear DC [20] power flow equations for a series of "what if" scenarios, or contingency analysis [24], [8] that investigate the potential power system state in the case of an event, e.g., a generator outage. Almost all the current solutions, e.g., contingency analyses, do not consider the cyber-side controllers and/or take into account adversarial settings, and hence those solutions miss maliciously induced topological errors in modern cyber-physical infrastructures. Additionally, power system stability analysis concentrates on continuous dynamics only, and does not fully consider the possibility of subsequent discrete logic events in the system.

Recently, cyber security solutions have been proposed to harden critical infrastructures. These include practical best-effort techniques such as regulatory compliance such as attack tree analysis [], NIST guidelines [21], and perimeter protection recommendations [12]. These approaches have been confirmed to be insufficient by the past major security incidents [5], and recently discovered fundamental security flaws in power grid control devices [23] and popular human machine interfaces (HMIs) [15] from major vendors. From an adversarial viewpoint, the past cyber attacks are mostly not physics-aware, and do not complete the attack path by sending malicious control inputs to the underlying physical plant components. The very few real-world security incidents with physical impact [6], however, use manually crafted malicious control parameters such as setting them to an unsafe high value like in Stuxnet. Those trivial strategies are to be addressed by NERC-CIP regulations [16] that mandate local safety measure deployment to protect unsafe component operational points.

VII. Conclusions

In this paper, we presented PhiDS, a physics-aware intrusion detection systems to identify compromised controllers without having to communicating with the controller processes neither through running on the controllers nor monitoring the controller's network behavior. PhiDS is given with the safety constraints of the power system (that are currently used in power system contingency analysis routines). PhiDS converts those constraints to linear temporal logic expressions. PhiDS introduces and leverages power system-level information flow analysis to detect when the controller behaves in an unsafe manner that violates the power system safety constraints.

Acknowledgement

This material is based upon work supported by the Department of Energy under Award Number DE-OE0000780.

References

[1] A. Abur and A. Expósito. *Power System State Estimation: Theory and Implementation*. Marcel Dekker, 2004.

[2] J. Arrillaga and B. Smith. *AC-DC Power Systems Analysis*. The Institution of Electrical Engineers, 1998.

[3] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes. Using Model-based Intrusion Detection for SCADA Networks. In *Proc. SCADA Security Scientific Symposium*, 2007.

[4] L. De Moura and N. Bjørner. Z3: An efficient smt solver. In *Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer, 2008.

[5] C. S. Edge and C. Blast. Ics-cert advisory. 2011.

[6] N. Falliere, L. O. Murchu, and E. Chien. W32.Stuxnet Dossier. Technical report, Symantic Security Response, Oct. 2010.

[7] R. Gerth, D. Peled, M. Y. Vardi, and P. Wolper. Simple on-the-fly automatic verification of linear temporal logic. In *Protocol Specification, Testing and Verification XV*, pages 3–18. Springer, 1996.

[8] J. Glover, M. Sarma, and T. Overbye. *Power System Analysis and Design*. Cengage Learning, 2011.

[9] W. M. Goble. *Control Systems Safety Evaluation and Reliability*. International Society of Automation, 2010.

[10] J. J. Grainger and W. D. Stevenson. *Power System Analysis*. McGraw Hill, 1994.

[11] S. Greene. *Margin and sensitivity methods for security analysis of electric power systems*. PhD thesis, University of Wisconsin–Madison, 1998.

[12] T. G. Lewis. *Critical infrastructure protection in homeland security: defending a networked nation*. John Wiley & Sons, 2006.

[13] D. Liberzon. *Switching in systems and control*. Springer Science & Business Media, 2012.

[14] S. Mohan, S. Bak, E. Betti, H. Yun, L. Sha, and M. Caccamo. S3A: Secure System Simplex Architecture for Enhanced Security of Cyber-Physical Systems. http://arxiv.org, 2012.

[15] T. H. Morris, A. K. Srivastava, B. Reaves, K. Pavurapu, S. Abdelwahed, R. Vaughn, W. McGrew, and Y. Dandass. Engineering future cyber-physical energy systems: Challenges, research needs, and roadmap. In *North American Power Symposium (NAPS), 2009*, pages 1–6. IEEE, 2009.

[16] National Energy Regulatory Comission. NERC CIP 002 1 - Critical Cyber Asset Identification, 2006.

[17] L. Pièrre-Cambacédès, M. Trischler, and G. N. Ericsson. Cybersecurity Myths on Power Control Systems: 21 Misconceptions and False Beliefs. *IEEE Transactions on Power Delivery*, 2011.

[18] A. Pnueli. The Temporal Logic of Programs. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, pages 46–57. IEEE Computer Society, 1977.

[19] S. J. Sheather. Weighted least squares. In *A Modern Approach to Regression with R*, Springer Texts in Statistics, pages 115–123. Springer New York, 2009.

[20] B. Stott and O. Alsac. Fast decoupled load flow. *IEEE Transactions on Power Apparatus and Systems*, 93(3):859–869, May 1974.

[21] K. Stouffer, J. Falco, and K. Scarfone. Guide to Industrial Control Systems (ICS) Security. *NIST Special Publication*, 800:82, 2008.

[22] U.S. Department of Energy Office of Electricity Delivery and Energy Reliability. A Summary of Control System Security Standards Activities in the Energy Sector, October 2005.

[23] S. E. Valentine. *PLC code vulnerabilities through SCADA systems*. PhD thesis, University of South Carolina, 2013.

[24] M. Vutsinas. *Contingency Analysis Using Synchrophasor Measurements*. PhD thesis, Clemson University, 2008.

[25] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham. A taxonomy of computer worms. In *Proceedings of the 2003 ACM Workshop on Rapid Malcode*, WORM '03, pages 11–18, New York, NY, USA, 2003. ACM.

[26] J. Weiss. Are the NERC CIPS making the grid less reliable. *Control Global*, 2009.