On False Data Injection Attacks Against Railway Traction Power Systems

Subhash Lakshminarayana[1] Zhan-Teng Teo[1] Rui Tan[2*] David K. Y. Yau[1,3] Pablo Arboleya[4]
[1]*Advanced Digital Sciences Center, Illinois at Singapore* [2]*Nanyang Technological University, Singapore*
[3]*Singapore University of Technology and Design* [4]*University of Oviedo, Spain*

***Abstract*—Modern urban railways extensively use computerized sensing and control technologies to achieve safe, reliable, and well-timed operations. However, the use of these technologies may provide a convenient leverage to cyber-attackers who have bypassed the air gaps and aim at causing safety incidents and service disruptions. In this paper, we study false data injection (FDI) attacks against railways' traction power systems (TPSes). Specifically, we analyze two types of FDI attacks on the train-borne voltage, current, and position sensor measurements – which we call *efficiency attack* and *safety attack* – that (i) maximize the system's total power consumption and (ii) mislead trains' local voltages to exceed given safety-critical thresholds, respectively. To counteract, we develop a global attack detection system that serializes a *bad data detector* and a novel *secondary attack detector* designed based on unique TPS characteristics. With intact position data of trains, our detection system can effectively detect the FDI attacks on trains' voltage and current measurements even if the attacker has full and accurate knowledge of the TPS, attack detection, and real-time system state. Extensive simulations driven by realistic running profiles of trains verify that a TPS setup is vulnerable to the FDI attacks, but these attacks can be detected effectively by the proposed global monitoring.**

I. Introduction

In modern cities, safe, reliable, and well-timed operations of urban railways are critical. A modern railway is a highly complex cyber-physical system (CPS) consisting of diverse subsystems including train motion control, traction powering, signaling, and etc, where deeply embedded information and communication technologies (ICTs) are used to operate each train and connect trains to an operation center. The extensive use of ICT may provide a convenient leverage to attackers, however, who may aim to hurt passengers' safety or cause widespread service disruptions. To date, the cybersecurity of modern railways has relied on air gaps that isolate their ICT systems from public networks. However, recent high-profile intrusions such as Stuxnet [1] and Dragonfly [2] have successfully breached the air gaps of critical CPS infrastructures and resulted in physical damage. For instance, the Stuxnet worm damaged nuclear centrifuges by injecting false control commands and forging normal system states. Its design and architecture are not domain-specific – they can be readily customized against other CPSes including transportation [1]. Insider attacks represent another major threat to air-gapped systems; their severe consequences have likewise been well documented [3]. It is thus critical to understand the cybersecurity risks of modern railways as a mission-critical CPS, and develop effective security countermeasures in their ICT design.

In this paper, we study the cybersecurity of direct current (dc) traction power systems (TPSes) that are widely deployed in urban electrical railways. The criticality of TPS is evidenced by prior severe incidents caused by TPS malfunctions. The 2014 Moscow metro derailment that led to 24 dead and 160 injured was caused by sudden braking of the train in question, when its traction voltage dropped abruptly [4]. In Singapore, a system-wide metro service disruption, triggered by TPS faults, affected almost half a million commuters during rush hours on July 7, 2015 [5]. Moreover, the computerized sensing and control in an automated TPS could be prime targets for cyber-attackers, who can sabotage the control and steer the system into inefficient and unsafe states.

Motivated by Stuxnet-type attacks that forge physical system states, in this paper, we study a general class of integrity attacks called *false data injection* (FDI), which tampers with train-borne sensor measurements required by TPS control. In a TPS, the electricity power supplied by substations is delivered by a network of overhead lines and third rails to the trains. According to its operation mode, a train's power consumption can be highly dynamic. In traction mode, it draws power from the TPS, causing a drop in the train's local voltage; in braking mode, it regenerates electricity from kinetic energy and injects this electricity back to the TPS,[1] causing a rise in the voltage. To prevent the voltage from exceeding safety-critical thresholds, trains apply *overcurrent control* and *squeeze control* [7] to throttle their power draw and injection, respectively. As these controls are performed based on train-borne voltage and current sensor measurements, FDI attacks on the measurements may mislead the train into erroneous power control decisions, which may in turn produce damaging and even catastrophic physical impacts on the train and the TPS. Recent results show that the measurements can be compromised in practice by precisely controlled electromagnetic interference in analog sensors [8], hardware trojans in chips [9], and malware infections in sensor firmwares [10]. Hence, FDI attack is

*Part of this work was completed while Rui Tan was with Advanced Digital Sciences Center, Illinois at Singapore.

[1]In electrical railways, trains are often equipped with regenerative brakes that generate electricity in deceleration [6].

a clear and present threat that requires immediate attention.

In this paper, we aim to answer the following two fundamental research questions:

(1) How to characterize the impact of FDI attacks on TPS system's efficiency and safety? Analysis of the impact based on an essential TPS model will provide basic understanding for developing countermeasures. However, the analysis is difficult, due to complex system dynamics arising from the trains' motion. In particular, a moving train does not only act as "load" and "generation" alternately over time, but also alters the power network's topology and electrical parameters continually. Moreover, because different TPS components (trains, substations, and etc) become physically interconnected through a common underlying power network, effects of an erroneous power control on a train during attack may propagate to the neighboring TPS components. The analysis must address these intricate and unique characteristics of TPSes.

(2) How to develop effective approaches for detecting the FDI attacks? Our thesis is that, because measurements from different trains are inherently correlated through interconnection over the same power network, for attack resilience we can apply a global detection that cross-checks the measurements collected from all trains based on an *a priori* global TPS model. However, in contrast to alternating current (ac) power grids that have well-established centralized monitoring and sensor data cross-check safeguards for reliable holistic control [11], [12], TPS is mainly concerned with individual trains' local operation (i.e., the overcurrent and squeeze controls), and therefore it is not traditionally subject to any global sensor data checks across trains. Thus, an existing dc TPS operation center seldom scrutinizes the sensor measurements, beyond their display and presentation for human operators. In this paper, we demonstrate the importance of these global, but hitherto ignored, sensor data cross checks in the TPS domain against FDI attacks.

In answering the above two research questions, our main contributions in this paper are as follows:

First, based on essential models of power substations, power flows, and train overcurrent and squeeze controls in a TPS, we formulate two types of FDI attacks that we call *efficiency attack* and *safety attack*. These attacks (i) maximize the total instantaneous power consumption of the TPS and (ii) mislead victim trains' local voltages to exceed given safety-critical thresholds, respectively. The efficiency attack formulation models an aggressive attacker who aims at maximizing the attack impact and provides insights into understanding the performance degradation limit caused by FDI attacks. Numerical results for a TPS section with two substations and two trains show that the efficiency attack can result in an instantaneous efficiency loss of about 20%, whereas the safety attack on a single train can indeed lead to significant safety breaches. These results substantiate the potency of FDI attacks on train-borne sensor measurements.

Second, we propose to apply a global *bad data detection* (BDD) method, similar to that widely used in ac power grids [12], to detect FDI attacks in a dc TPS. Despite a known vulnerability of the BDD – it can be bypassed by an attacker who knows enough details of its design – our numerical results show that, in order for an FDI attack to be stealthy against the BDD, it will have to settle for a significantly reduced damage on the system efficiency. Moreover, we observe that, given intact position data of trains, solutions of the BDD bypass condition will become discrete. Based on this observation, we develop a novel *secondary attack detection* (SAD) algorithm that can effectively detect the onset of an FDI attack on trains' voltage and current measurements after it has bypassed the BDD. Hence, the BDD and the SAD form in tandem a global attack detector under the Kerckhoffs's assumption (i.e., the attacker has full and accurate knowledge of the system model, attack detection, and real-time system state), provided that the integrity of trains' position information can be verified.

Third, we report extensive simulations, driven by realistic profiles of trains in operation, to evaluate our solutions. For a TPS with four trains each running over a distance of ten kilometers for 800 seconds, our results show that, without the global BDD, FDI attacks can increase the total system energy consumption by 28.3% and breach the system's safety condition. After applying the BDD, the system's total energy consumption increases by no more than 6.2% under the efficiency attack, and safety attacks become no longer successful. Moreover, the proposed SAD algorithm achieves a probability of 96% in detecting the onsets of the FDI attacks that have successfully bypassed the BDD.

The balance of the paper is organized as follows. Section II reviews related work. Section III describes our TPS model. Section IV formulates the efficiency and safety attacks. Section V analyzes the effectiveness of the BDD and presents the proposed SAD algorithm. Section VI presents simulation results. Section VII concludes.

II. Related Work

Power flow analysis and optimization for TPS have received increasing research interest. Power flow analysis is a basic tool for TPS planning and operation. Prior work has analyzed dc power flows [13]–[15] and addressed the interactions between the dc TPS and a supporting ac power grid [16], [17]. We adopt existing electrical models for different TPS components [13]–[15] in this work. These models provide sufficient accuracy generally [14], and they are tractable for analysis. Based on power flow analysis, recent research has tried to improve the energy efficiency of railways by leveraging trains' power regeneration [18]. Techniques such as synchronizing the trains' speed profiles [19] and real-time substation voltage control [20] have been shown to provide efficient reuse of the regenerated power. To the best of our knowledge, none of the existing studies

on TPS control have addressed it from a cybersecurity perspective. The security problem is imperative, since TPS is a form of critical infrastructure that renders it an attractive target for attacks.

Different types of CPS can have vastly different properties and characteristics. Thus, their security concerns and admissible detection strategies can be totally different. Typically, their cybersecurity analysis must be carried out in a domain specific manner, with customized considerations given to main details and semantics of specific systems. Cybersecurity of various CPSes has been studied. Cárdenas et al. [21] investigate the impacts of integrity and denial-of-service attacks on a process control system of a chemical reactor. Amin et al. [22] perform security threat assessment of supervisory control and data acquisition systems for water supply. Other efforts [12], [23] have analyzed FDI attacks against ac utility power grids. They show that an attacker capable of tampering with grid sensor measurements or topology information can carefully construct attacks to bypass the detection by certain existing fault detectors. Recent studies have investigated the impact of such stealthy attacks on grid power flows [11], [24]. They show that maliciously biased estimates of the system state can cause grid operators to make erroneous decisions that will lead to degraded performance or even safety breaches. This paper is the first to analyze the efficiency and safety of TPS under FDI attacks. We provide new and nontrivial domain specific modeling and analysis to capture the targeted application's unique features and key properties. In particular, TPS involves real-time and complex interactions between two highly dynamic physical systems, namely a mechanical system of the trains' motion and an electrical system that governs the trains' power consumption and regeneration during this motion. Attackers could exploit the interactions to strengthen their attacks.

III. Traction Power System Model

In this section, we present a model of a dc TPS at a certain time instant. The TPS is modeled as a power network consisting of N nodes. Denote by $\mathcal{N} = \{1, 2, \ldots, N\}$ the set of nodes and $\mathcal{L}$ the set of resistive branches connecting the nodes. The substations and the trains are connected to different nodes. The sets of nodes for the substations, the tractioning trains, and the regenerating trains are denoted by $\mathcal{N}_{\text{sub}}$, $\mathcal{N}_{\text{tra}}$, and $\mathcal{N}_{\text{reg}}$, respectively. We define $\mathcal{N}_{\text{trains}} = \mathcal{N}_{\text{tra}} \cup \mathcal{N}_{\text{reg}}$. The positions of the nodes $1, \ldots, N$ are denoted by a set $\mathbf{s} = \{s_1, s_2, \ldots, s_N\}$, where s_1 is zero and s_i is the distance from node i to node 1. Fig. 1 illustrates a TPS section with two substations at nodes 1 and 4, as well as two trains at nodes 2 and 3. The train at node 2 is tractioning and the train at node 3 is braking and regenerating. Thus, $\mathcal{N} = \{1, 2, 3, 4\}, \mathcal{L} = \{(1, 2), (2, 3), (3, 4)\}, \mathcal{N}_{\text{sub}} = \{1, 4\}, \mathcal{N}_{\text{tra}} = \{2\}, \mathcal{N}_{\text{reg}} = \{3\}$. The electrical models for

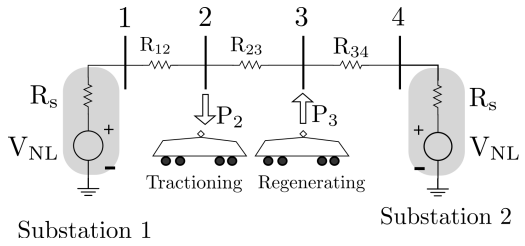


Fig. 1: Illustration of a TPS section.

the power network, substations, and trains are described as follows.

Power network: Let V_i and I_i denote the voltage and current at node i, respectively, and $\mathbf{v}$ and $\mathbf{i}$ the vectors of the nodal voltages and currents. For safe operations, all nodal voltages must be within a safety limit, i.e.,

$$V_i^{\min} \le V_i \le V_i^{\max}, \qquad \forall i \in \mathcal{N}, \tag{1}$$

where $V_i^{\min}$ and $V_i^{\max}$ are the safety thresholds for node i. By convention, we assume that the current injected into the TPS is positive. The resistance of the branch connecting the nodes i and j is denoted by $R_{i,j}(\mathbf{s})$ and its conductance by $G_{i,j}(\mathbf{s})$, where $G_{i,j}(\mathbf{s}) = 1/R_{i,j}(\mathbf{s})$. Note that branch resistance (and conductance) depends on the positions of the trains, i.e., $\mathbf{s}$. For instance, in Fig. 1, $R_{i,i+1} = \gamma(s_{i+1} - s_i)$, where γ is a constant depending on the electrical wire characteristics. From Kirchhoff's circuit laws, we have

$$\mathbf{Y}(\mathbf{s})\mathbf{v} = \mathbf{i}, \tag{2}$$

where $\mathbf{Y}(\mathbf{s}) \in \mathbb{R}^{N \times N}$ is the nodal conductance matrix and the (i, j)th element of $\mathbf{Y}(\mathbf{s})$, denoted by $Y_{i,j}(\mathbf{s})$, is given by

$$Y_{i,i}(\mathbf{s}) = \sum_{j:(i,j)\in\mathcal{L}} G_{i,j}(\mathbf{s}),$$

$$Y_{i,j}(\mathbf{s}) = \begin{cases} -G_{i,j}(\mathbf{s}), & \text{if } j \ne i \text{ and } (i,j) \in \mathcal{L}, \\ 0, & \text{if } j \ne i \text{ and } (i,j) \notin \mathcal{L}. \end{cases}$$

Substations: We consider inverting substations capable of both supplying and absorbing power. They are modeled as dc voltage sources governed by

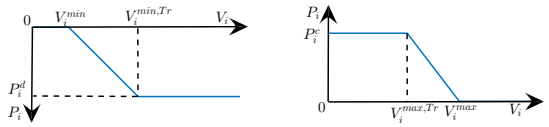
$$V_i = V_{\text{NL}} - R_s I_i, \qquad i \in \mathcal{N}_{\text{sub}}, \tag{3}$$

where V_{NL} and R_s are the no-load voltage and the internal resistance of the substation. When a substation supplies power, $I_i > 0$; when it absorbs power, $I_i < 0$ and the absorbed power is injected back to the supporting ac power grid. This dc substation model has been widely adopted in TPS analysis [14], [20].

Trains: Let P_i denote the power absorbed or injected by a tractioning train or a regenerating train at node i. We have

$$P_i = V_i I_i. \tag{4}$$

For safety, the trains adopt the following two local power controls [7].



(a) (b)

Fig. 2: (a) Overcurrent control for accelerating trains. (b) Squeeze control for regenerating trains.

Overcurrent control: A tractioning train absorbs power from the power network, resulting in a drop in the train's nodal voltage. To prevent the nodal voltage from falling below the safety threshold $V_i^{\min}$, the overcurrent control is activated whenever the nodal voltage at the train node i drops below a triggering voltage $V_i^{\min,\mathrm{Tr}}$. Let P_i^d denote the power demand of a tractioning train at node i. The overcurrent control will command the train to absorb P_i units of power, where P_i is given by

$$P_i = \begin{cases} 0, & \text{if } V_i \le V_i^{\min}; \\ P_i^d \left(\frac{V_i - V_i^{\min}}{V_i^{\min,\mathrm{Tr}} - V_i^{\min}} \right), & \text{if } V_i^{\min} \le V_i \le V_i^{\min,\mathrm{Tr}}; \\ P_i^d, & \text{if } V_i \ge V_i^{\min,\mathrm{Tr}}. \end{cases}$$

This control law is illustrated in Fig. 2a. Specifically, if the nodal voltage at the train is greater than the triggering voltage, the train draws a power equal to its demand. Otherwise, the train curtails its power consumption according to the V_i. If the voltage drops below $V_i^{\min}$, the train draws no power to prevent safety incidents.

Squeeze control: A regenerating train injects power into the power network, resulting in an increase in the train's nodal voltage. To prevent the voltage from exceeding the safety limit $V_i^{\max}$, the squeeze control is activated whenever the nodal voltage at the train node i exceeds a certain triggering voltage level $V_{i,\max}^{\mathrm{Tr}}$. Let P_i^c denote the regeneration capacity of the train. The squeeze control will command the train to inject P_i units of power into the TPS, where P_i is given by

$$P_i = \begin{cases} P_i^c, & \text{if } V_i^{\max} \le V_i^{\max,\mathrm{Tr}}; \\ P_i^c \left(\frac{V_i^{\max} - V_i}{V_i^{\max} - V_i^{\max,\mathrm{Tr}}} \right), & \text{if } V_i^{\max,\mathrm{Tr}} \le V_i \le V_i^{\max}; \\ 0, & \text{if } V_i \ge V_i^{\max}. \end{cases}$$

This control law is illustrated in Fig. 2b. Specifically, if the nodal voltage at the train is lower than the triggering voltage, the train injects all the regenerated power. Otherwise, the train curtails the power injection according to the V_i by burning the remaining power in a rheostatic braking system [7]. If the voltage drops below $V_i^{\max}$, the train does not inject power into the TPS to prevent safety incidents.

The train's power demand P_i^d and regeneration capacity P_i^c depend on the train's running profile and real-time state. They can be provided by the train's motion control system. We note that the electrical models described in this section address the steady-state voltages and currents. They ignore the power transients of the trains due to their internal feedback control systems that implement the overcurrent/squeeze control decisions. However, it is safe to ignore these transients, because they can settle quickly before the next overcurrent/squeeze control action [25]. This steady-state analysis approach has been widely adopted in TPS power flow analysis [13]–[15].

IV. False Data Injection Attacks Against TPS

In this section, we study how an attacker can mislead the TPS into an inefficient or unsafe operating state. We focus on FDI attacks that tamper with the measurements of train-borne voltage and current sensors. Such an attack will cause the TPS to make wrong decisions of power absorption/injection, since a train's overcurrent and squeeze controls depend on the sensor measurements. We further consider attacks of two different objectives: (i) to increase the system's total instantaneous power consumption, and (ii) to cause breaches of the safety conditions in Eq. (1). We call them *efficiency attack* and *safety attack*, respectively. In this section, we first describe our threat model. Then, we analyze the attacker's approach to computing efficiency and safety attacks. Lastly, we present numerical results to illustrate the two kinds of attacks.

A. Threat Model

Real-world attackers against critical CPSes are often smart, resourceful, and highly strategic. Their strategies can be guided by detailed knowledge of their targets, which can be obtained in practice by malicious insiders, long-term data exfiltration [2], or social engineering against employees, contractors, or vendors of the operators in question [1]. In this paper, we follow Kerckhoffs's principle to consider an attacker who has accurate knowledge of the targeted system and read access to the system state. Knowledge of the system includes the electrical models and parameters given in Section III, as well as the system's method of attack detection. The system state includes present power demands, regeneration capacities, as well as voltage, current, and position measurements of all the trains. This information can be leaked through a compromised operation center, as in recent high-profile attacks [1], [2]. We assume that the attacker can corrupt the voltage, current, and position measurements of nodes in the set $\mathcal{N}_{\mathrm{a}}$, where $\mathcal{N}_{\mathrm{a}} \subseteq \mathcal{N}$. Recent studies have demonstrated that such unauthorized write access can be obtained for analog sensors, traditional electromechanical meters, and modern smart meters [8], [26], [27]. Analog sensors are vulnerable to controlled electromagnetic interference [8]; measurement devices can be affected by hardware trojans [9] and infected with malwares [10], [26].

Under the Kerckhoffs's assumption on the attacker's knowledge, we will analyze his strategies of achieving successful efficiency and safety attacks. Conversely, we will also develop countermeasures by a defender to detect these attacks. Our threat model is strong, but the conservative analysis is necessary because any underestimation of the

attacker's capability may have catastrophic consequences, including extremely costly infrastructure damage and loss of human lives.

We note that, alternatively, the attacker can launch FDI attacks against the *decisions* of the local controls (i.e., the P_i values for the trains). To detect such attacks, each train can compare the P_i value in question with that computed based on the train's voltage and current measurements and the *a priori* overcurrent and squeeze control laws. In the rest of this paper, we focus on the analysis and detection of FDI attacks on the voltage and current measurements only. This problem is comparatively much more challenging since information compromised right at the sources will preclude its use for any subsequent sanity checks.

B. FDI Attack Construction

In the rest of the paper, we will use x' to denote the compromised version of a sensor measurement x. In this section, we analyze how to compute an effective *attack vector*, as a vector of false voltage and current measurements to be injected into the sensing systems of the trains in $\mathcal{N}_a$. Note that, in this section, we ignore position measurements in the attack vector, because they will not affect the trains' overcurrent and squeeze controls. In the following, we first derive conditions for the attack vector to mislead the train into absorbing or injecting a certain amount of power. With the calculated power absorptions/injections of the trains, we can determine the system's total power consumption and hence its safety status. Thus, we can formulate the attacker's problem of finding an attack vector to achieve his goal (e.g., to maximize the total power consumption), under conditions that we will state presently for enforcing a certain amount of power absorption/injection.

The following conditions are sufficient to enforce that a train at node $i \in \mathcal{N}_a$ will absorb or inject P_i units of power:

$$V_i' I_i' = P_i, \qquad i \in \mathcal{N}_{\text{a}}; \tag{5}$$

$$P_i \geq P_i^d, \qquad i \in \mathcal{N}_{\text{tra}}; \tag{6}$$

$$P_i \leq P_i^c, \qquad i \in \mathcal{N}_{\text{reg}}; \tag{7}$$

$$V_i' \begin{cases} \geq V_i^{\min,\text{Tr}}, & \text{if } P_i = P_i^d, \\ = V_i^{\min} + \frac{P_i(V_i^{\min,\text{Tr}} - V_i^{\min})}{P_i^{\min}}, & \text{if } P_i^d \leq P_i \leq 0, \\ \leq V_i^{\min}, & \text{if } P_i = 0, \end{cases} \tag{8}$$
$$\forall i \in \mathcal{N}_{\text{a}} \cap \mathcal{N}_{\text{tra}};$$

$$V_i' \begin{cases} \geq V_i^{\max}, & \text{if } P_i = 0, \\ = V_i^{\max} - \frac{P_i(V_i^{\max} - V_i^{\max,\text{Tr}})}{P_i^c}, & \text{if } 0 \leq P_i \leq P_i^c, \\ \leq V_i^{\max,\text{Tr}}, & \text{if } P_i = P_i^c, \end{cases} \tag{9}$$
$$\forall i \in \mathcal{N}_{\text{a}} \cap \mathcal{N}_{\text{reg}}.$$

The conditions in Eqs. (8) and (9) are obtained by inverting the overcurrent and squeeze control laws given in Section III, and replacing the true voltage V_i by the compromised measurement V_i'. As a result, based on V_i', the train will follow the overcurrent/squeeze control law to regulate its power absorption/injection at the attacker's desired value P_i. This control process is often achieved in a closed loop, with the measurements V_i' and I_i' acting as the feedback and the desired value P_i as the setpoint. Under the condition Eq. (5), the actual power absorption/injection under the aforementioned closed-loop control will converge to P_i. Moreover, the condition Eq. (5) can hide the attack for trains that can directly measure the power consumption. The conditions Eqs. (6) and (7) ensure the feasibility of inducing the train to absorb/inject P_i units of power. Specifically, the attacker's desired P_i should not exceed a regenerating train's regeneration capacity. The condition Eq. (6), where both P_i and P_i^d are negative, prevents the mechanism from violating the overcurrent control. In summary, if the compromised measurements V_i' and I_i' satisfy the conditions in Eq. (5) to Eq. (9), the train will control its power absorption/injection to P_i. With this understanding, the attacker can carefully plan the attack vector to achieve his goal. Without the conditions in Eq. (5) to Eq. (9), the attacker cannot predict the impact of his attack, and therefore cannot implement his desired strategy.

Each sensor in the TPS may apply data quality checks on its measurements. For instance, the measurements at present time instant should not differ significantly from those predicted based on the measurements at previous time instant. Intuitively, if the compromised measurement is bounded around the true measurement, the data quality checks, designed to be insensitive to natural random noises of measurement, will not raise an alarm. Thus, we assume that the compromised measurements need to satisfy:

$$\mathbf{v} - \Delta\mathbf{v} \preceq \mathbf{v}' \preceq \mathbf{v} + \Delta\mathbf{v}, \tag{10}$$

$$\mathbf{i} - \Delta\mathbf{i} \preceq \mathbf{i}' \preceq \mathbf{i} + \Delta\mathbf{i}, \tag{11}$$

where $\Delta\mathbf{v} = [\Delta V_1, \ldots, \Delta V_N]^T$ and $\Delta\mathbf{i} = [\Delta I_1, \ldots, \Delta I_N]^T$ are the maximum errors allowed by the data quality checks; $\mathbf{x} \preceq \mathbf{y}$ means that each element of $\mathbf{x}$ is no greater than the corresponding element in $\mathbf{y}$. We note that, if $i \notin \mathcal{N}_{\text{a}}$, $\Delta V_i = 0$.

Based on the above conditions for the compromised measurements, we formulate the efficiency and safety attacks.

1) Efficiency Attack: An efficiency attack causes an increase or decrease in the total instantaneous power injected or absorbed by the substations. In particular, we consider an aggressive attacker who aims to maximize or minimize such injected or absorbed power. Formally, the attacker solves the following constrained optimization problem to compute the attack vector $\{V_i', I_i' | \forall i \in \mathcal{N}_a\}$:

$$\max_{\{V_i', I_i' |\ \forall i \in \mathcal{N}_a\}} \quad \sum_{i \in \mathcal{N}_{\text{sub}}} V_i I_i \tag{12}$$
$$s.t. \qquad \text{constraints in Eq. (2) to Eq. (11).}$$

The above formulation captures the physical laws governing the power network and the substations (i.e., Eq. (2) to Eq. (4)), as well as how the attack vector induces the trains to make erroneous power control decisions (i.e., Eq. (5) to Eq. (9)). Specifically, for any $\{V_i', I_i' \mid \forall i \in \mathcal{N}_a\}$ satisfying Eq. (5) to Eq. (9), the attacker can predict the trains' power absorptions/injections $\{P_i = V_i' I_i' \mid \forall i \in \mathcal{N}_{\text{trains}}\}$. He then uses the physical laws in Eqs. (2), (3), and (4) to determine the actual voltages and currents of the substations (i.e., $\{V_i, I_i \mid \forall i \in \mathcal{N}_{\text{sub}}\}$) and predict the system's total power consumption $\sum_{i \in \mathcal{N}_{\text{sub}}} V_i I_i$.

Solving the constrained optimization problem in Eq. (12) can be computationally intensive, mainly because the constraints in Eqs. (8) and (9) are non-smooth and non-differentiable. Existing constrained non-linear optimization solvers (e.g., the `fmincon` function of MATLAB) often require smooth objective and constraint functions. To use these solvers, the attacker can adopt a *divide-and-conquer* approach that divides the problem Eq. (12) into multiple subproblems, in which a piece of Eq. (8) or Eq. (9) is selected as a constraint for a train. By comparing the optimization results of all the subproblems, the attacker can obtain a global optimal solution to the problem in Eq. (12). Because each train has three choices in Eq. (8) or Eq. (9), this approach will generate a total of $3^{|\mathcal{N}_a|}$ subproblems, where $|\mathcal{N}_a|$ is the number of trains under FDI attacks. As the subproblems are mutually independent, the attacker can solve the subproblems in parallel, to reduce computation time. The ability to solve the problem in Eq. (12) in real time can be important to the attacker. This is because, to accumulate large energy loss, the attacker needs to keep at the FDI attacks by solving Eq. (12) continually, based on the latest system state given by $\mathbf{s}$, P_i^d, and P_i^c. The attacker will need to procure sufficient computing resources for achieving the real-time objective.

2) Safety Attack: For safety attacks, we model the space of attack vectors that can cause the voltages at a subset of the TPS nodes, denoted by $\mathcal{N}_{\text{unsafe}}$, to cross the safety limits in Eq. (1). The attack space is defined by all the constraint conditions in the optimization problem Eq. (12), and $V_i \notin [V_{i,\min}, V_{i,\max}], i \in \mathcal{N}_{\text{unsafe}}$. As long as the attacker can find an attack vector satisfying the above constraints, he will be able to achieve the safety breaches.

We now discuss a heuristic approach that the attacker can use to aggressively increase the extent of the safety breaches. Specifically, the attacker maximizes the total power injected into the TPS by the regenerating trains, i.e., $\sum_{i \in \mathcal{N}_{\text{reg}}} V_i I_i$, subject to all the constraints of the optimization problem in Eq. (12). The intuition is that injecting more power into the TPS will result in higher catenary voltages. This constrained optimization problem can also be solved by the aforementioned divide-and-conquer approach. Our numerical results in Section IV-C show that, under this heuristic approach, tampering with the sensor measurements of a single train can already lead to safety breaches.

Table I: TPS model parameters.

Parameters	V_{NL}	γ	R_s	$V_i^{\max,\text{Tr}}$	$V_i^{\max}$
Value	750V	30mΩ/km	29.56mΩ	850V	900V

C. Numerical Examples

We now present numerical examples to illustrate the efficiency and safety attacks. These examples are based on the TPS shown in Fig. 1, in which both trains are decelerating and regenerating. The system model parameters are given in Table I. We consider a time instant at which the system state in the absence of attack is given by the first part of Table II, where the total instantaneous power absorbed by the substations and injected back into the supporting ac power grid is 3.601 MW. In these examples, we assume that the attacker can only compromise the voltage and current measurements of the train at node 2.

1) Efficiency Attack: The attacker solves the constrained optimization problem in Eq. (12) and tampers with V_2 and I_2 accordingly. We set $\Delta V_i = 50$V and $\Delta I_i = 200$A, $\forall i \in \mathcal{N}_a$. The compromised measurements and the true state of the system under attack are given in the second part of Table II. We can see that the compromised voltage measurement at node 2 is greater than the true value. Consequently, the train injects less power into the TPS because of the squeeze control, resulting in less power absorption by the substations. Specifically, the total power absorption is 2.888 MW, a 20% reduction compared with the case of no attack. Thus, the power efficiency of the system is degraded.

2) Safety Attack: The attacker uses the heuristic approach in Section IV-B2 to compute the safety attack. The compromised measurements and the true system state are given in the third part of Table II. The compromised voltage measurement at node 2 is lower than its true value. Thus, the train at node 2 injects more power into the TPS because of the squeeze control, causing the actual voltage at node 2 to exceed the safety limit. We can see that it is possible for an attacker to tamper with the measurements of a single train and already achieve a safety attack. In this example, since both the trains are regenerating, the catenary voltages are closer to the safety limit. This makes it easier for the attacker to achieve the safety attack. Thus, for an attacker with limited write access to the trains' measurements (i.e., a small set $\mathcal{N}_a$), he can continuously monitor the system and wait for feasible moments for launching safety attacks.

V. Global Attack Detection

As discussed in Section I, dc TPSes mainly rely on trains' local controls (i.e., overcurrent and squeeze controls) to avoid unsafe states. The TPS does not otherwise cross-check sensor data from different trains based on an *a priori* TPS model to ensure the data's global consistency. However, such global monitoring is clearly advantageous,

Table II: System state and compromised measurements under efficiency and safety attacks. Distance is measured in kilometers, voltage in volts, current in amperes, and power in megawatts.

Node		1	2	3	4
TPS State (Without Attack)	s_i	0	0.9	1.2	2
	P_i^c	-	5.5	1.8	-
	V_i	815.6	875.5	867.7	815
	I_i	-2218.8	3079.2	1338	-2198.4
	P_i	-1.81	2.696	1.161	-1.792
Efficiency Attack	V_i'	-	888.6	-	-
	I_i'	-	1409.6	-	-
	V_i	801.1	847.7	850	805.2
	I_i	-1728.2	1477.6	2117.6	-1867.1
	P_i	-1.384	1.253	1.8	-1.503
Safety Attack	V_i'	-	862.9	-	-
	I_i'	-	4731.6	-	-
	V_i	828.9	**901**	884.2	824.1
	I_i	-2669.1	4531.6	643.2	-2505.7
	P_i	-2.212	4.083	0.569	-2.065

because anomalies in the data relationships, can help flag the occurence of an FDI attack.

An attacker that wishes to remain stealthy under global monitoring thus becomes more constrained, and his actions may become less effective. In this section, we present the design of a global monitor for detecting FDI attacks under the Kerckhoffs's assumption.

Fig. 3 overviews our global attack detection approach, in which the trains' voltage, current, and position measurements are sent to a central *TPS monitor* periodically. The TPS monitor applies state estimation (SE), bad data detection (BDD), position integrity verification (PIV), and secondary attack detection (SAD) in sequence to detect attacks. In ac utility power grids, similar SE and related BDD are widely used for detecting faulty data or reducing the impact of noisy sensor measurements [28]. In Section V-A, we propose a new BDD design specific to the application domain of dc TPS. By checking the consistency among measurements based on a TPS model, the BDD can detect a range of FDI attacks. However, the detection is not complete – an attacker under the Kerckhoffs's assumption can bypass it using his knowledge of the system. In Section V-B, we provide numerical results to illustrate the impacts of these stealthy attacks. To counter them, in Section V-C, we further propose a novel SAD algorithm to supplement the BDD, under an additional assumption that the trains' position data is intact, which is ensured by the PIV.

A. TPS Bad Data Detection and Its Vulnerability

Recall that in Section IV-B, the trains apply local controls based on their own voltage and current measurements only. Hence, the trains' position information does not matter. Under global detection, however, compromise of the trains' position information becomes relevant, since it may enable the

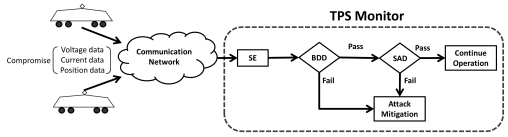


Fig. 3: Global attack detection. SE: State estimation; BDD: Bad data detection; PIV: Position integrity verification; SAD: Secondary attack detection.

attacker to mislead the TPS monitor into deriving a wrong TPS model that is consistent with the compromised voltage and current measurements. Tampering with the position data can thus help the attacker evade detection. Although in practice it is extremely difficult for the attacker to hide the compromise of train position data because multiple sources of this data are often available (see Section V-C for the details), in this section, for generality of analysis, we account for possible compromise of the position data.

We use $\tilde{x}$ to represent a possibly compromised measurement x, i.e., $\tilde{x} = x$ in the absence of attack and $\tilde{x} = x'$ in the presence of attack. The state of the TPS is a vector of the nodal voltages, i.e., $\mathbf{v}$. The set of measurements includes nodal positions $\tilde{\mathbf{s}} = [\tilde{s}_1, \ldots, \tilde{s}_N]^T \in \mathbb{R}^{N\times 1}$, and nodal voltage and current readings $\tilde{\mathbf{z}} = [\tilde{\mathbf{v}}, \tilde{\mathbf{i}}]^T \in \mathbb{R}^{2N\times 1}$. In the absence of attack, the measurement vector $\mathbf{z}$ is related to the system state $\mathbf{v}$ as $\mathbf{z} = \mathbf{H}(\mathbf{s})\mathbf{v} + \mathbf{n}$, where $\mathbf{H}(\mathbf{s}) = [\mathbb{I}_N; \mathbf{Y}(\mathbf{s})]$ is a *measurement matrix* depending on the positions $\mathbf{s}$, $\mathbb{I}_N$ is an N-dimensional identity matrix, and $\mathbf{n} \in \mathbb{R}^{2N\times 1}$ is a random measurement noise vector. We assume that $\mathbf{n}$ follows a multivariate Gaussian distribution. The maximum likelihood (ML) estimate of $\mathbf{v}$, denoted by $\hat{\mathbf{v}}$, is given by $\hat{\mathbf{v}} = (\mathbf{H}(\tilde{\mathbf{s}})^T\mathbf{\Sigma}^{-1}\mathbf{H}(\tilde{\mathbf{s}}))^{-1}\mathbf{H}(\tilde{\mathbf{s}})^T\mathbf{\Sigma}^{-1}\tilde{\mathbf{z}}$, where $\mathbf{\Sigma}$ is the covariance matrix of $\mathbf{n}$ [28, Chap. 12]. The SE's BDD raises an alarm if $(\tilde{\mathbf{z}} - \mathbf{H}(\tilde{\mathbf{s}})\hat{\mathbf{v}})^T\mathbf{\Sigma}^{-1}(\tilde{\mathbf{z}} - \mathbf{H}(\tilde{\mathbf{s}})\hat{\mathbf{v}}) > \tau$, where τ is a constant threshold that can be determined to meet a given false alarm rate under random measurement noise. The BDD is originally designed to detect faulty sensor data caused by natural malfunction of sensors. Thus, it is effective in detecting a range of FDI attacks that are not specifically designed to bypass it. However, the attacker that we consider in this paper, following the Kerckhoffs's principle, can design FDI attacks to bypass the BDD. In the following, we formulate these stealthy attacks.

From an existing result [12], if the compromised measurement vector $\mathbf{z}'$ is in the column space of the compromised measurement matrix $\mathbf{H}(\mathbf{s}')$, $\mathbf{z}'$ can bypass the BDD. Applying this result to the TPS, we have the following lemma.

Lemma 1: Any compromised measurements that satisfy

$$\mathbf{Y}(\mathbf{s}')\mathbf{v}' = \mathbf{i}' \tag{13}$$

can bypass the BDD.

Proof: Lemma 1 holds since any $\mathbf{z}'$ that satisfies Eq. (13) is in the column space of $\mathbf{H}(s')$, i.e., $\mathbf{z}' = [\mathbf{v}', \mathbf{i}']^T = [\mathbb{I}_N; \mathbf{Y}(\mathbf{s}')]\mathbf{v}' = \mathbf{H}(\mathbf{s}')\mathbf{v}'$. ■

In addition to Eq. (13), the TPS monitor may use two

other sensor data checks. First, to meet the constraint in Eq. (13), the attacker may need to compromise the voltage and current measurements at the substations. The TPS monitor may check the substation measurements, i.e., V_i and I_i, $\forall i \in \mathcal{N}_{\text{sub}}$, against the substation model in Eq. (3). To be stealthy to this check, the attacker can impose an additional constraint of

$$V_i' = V_{\text{NL}} - R_s I_i', \quad \forall i \in \mathcal{N}_{\text{sub}}. \tag{14}$$

Second, the TPS monitor can also apply data quality checks similar to those in Eqs. (10) and (11) to check the trains' position measurements. Thus, if the attacker can compromise the position measurements, he needs to satisfy

$$\mathbf{s} - \Delta\mathbf{s} \preceq \mathbf{s}' \preceq \mathbf{s} + \Delta\mathbf{s}, \tag{15}$$

where $\Delta\mathbf{s} = [\Delta s_1, \ldots, \Delta s_N]^T$ are the maximum allowed errors for position measurements and $\Delta s_i = 0$ if $i \notin \mathcal{N}_{\text{a}}$.

Therefore, the efficiency attacks that are stealthy to the BDD can be computed by solving the constrained optimization problem Eq. (12) with the additional constraints Eq. (13) to Eq. (15). Similarly, the attack space for BDD-stealthy safety attacks is characterized by the constraints of the optimization problem Eq. (12), $V_i \notin [V_{i,\min}, V_{i,\max}], i \in \mathcal{N}_{\text{unsafe}}$, and the additional constraints Eq. (13) to Eq. (15). Naturally, BDD reduces the attack space since the attacker now needs to satisfy additional constraints to remain undetected. In the simulation results presented in Section VI, we show that, under a realistic TPS setting, the BDD significantly reduces the impact of attacks.

B. Numerical Examples

We now present numerical examples to illustrate the efficiency and safety attacks that can bypass the BDD as analyzed in Section V-A. The TPS model and parameters are identical to those in Section IV-C. The true system state and the compromised measurements are given in Table III. We set $\Delta s_i = 0.6\,\text{km}$, $\forall i \in \mathcal{N}_a$. To illustrate a powerful attacker, we assume that the attacker can corrupt the voltage and current measurements of all the four nodes in Fig. 1, as well as the positions of both the trains.

1) Efficiency Attack: Under the efficiency attack, the total power injected back to the supporting power grid by the substations is 3.431 MW, which is a reduction of about 4.7% compared with no attacks. This reduction is much less than the 20% caused by the efficiency attack in Section IV-C, which was achieved by compromising the voltage and current measurements of node 2 only in the absence of BDD. This result illustrates the ability of the BDD in limiting the impact of efficiency attacks.

2) Safety Attack: We observe that by compromising the nodal measurements and the trains' position information, the attacker can increase the voltage at node 2 to $901.4\,\text{V}$ while bypassing the BDD. Furthermore, if the attacker can gain write access to any one train (i.e., $|\mathcal{N}_a| = 1$), he cannot launch a successful safety attack. This is in contrast to the example in Section IV-C, where the attacker could launch a successful safety attack by compromising the measurements of a single train only.

Table III: System state and compromised measurements under efficiency and safety attacks that have bypassed the BDD. Distance is measured in kilometers, voltage in volts, current in amperes, power in megawatts.

Node		1	2	3	4
Efficiency Attack	s_i'	0	1	1	2
	V_i'	812	874.9	874.9	812
	I_i'	-2096.7	3159	1034.5	-2096.8
	V_i	813.2	871	861.7	811.6
	I_i	-2138.8	3173.2	1050.4	-2084.8
	P_i	-1.739	2.764	0.905	-1.692
Safety Attack	s_i'	0	0.43	1.8	2
	V_i'	835	872.3	847.3	830.9
	I_i'	-2876.8	3487.8	2124.5	-2735.4
	V_i	829.1	**901.4**	895.1	830.1
	I_i	-2676.8	3375.3	2010.9	-2709.4
	P_i	-2.219	3.043	1.8	-2.249

In summary, the above examples suggest that the global monitoring and BDD can significantly limit the impact of stealthy FDI attacks on the TPS even if the attacker can compromise the measurements of multiple trains. To accomplish a safety attack, the attacker needs to compromise more trains compared with no BDD.

C. Secondary Attack Detection

In this section, we propose a novel secondary attack detection (SAD) algorithm that can effectively detect the *onset* of an FDI attack that has bypassed the BDD. SAD requires that the trains' position data communicated to the TPS monitor is intact. It is feasible for the TPS monitor to verify the integrity of the position data. For example, real-world railway systems invariably provide multiple sources of train position information including train-borne wheel sensors, GPS, track-side Balise [29], etc. By cross-checking position measurements from the multiple sources, we can readily identify FDI attacks on the position data unless the attacker succeeds in compromising all the data sources, which is highly challenging. Such cross checks constitute the PIV illustrated in Fig. 3. If FDI attacks on the position data are identified, the TPS should immediately apply attack mitigation approaches. Due to space limitation, the details of attack mitigation are omitted in this paper and can be found in [30].

The analysis in the previous sections is for a particular time instant, and the attacker can use the techniques in Sections IV and V-A to launch attacks continually over time. Once the SAD detects an attack's onset, the system can activate the attack mitigation to render subsequent FDIs ineffective. Thus, in this section, we focus on analyzing the property of the system and designing the SAD accordingly

for the onset time instant only of an attack.

1) A Discrete Solution Property: The requirement of intact position data and the design of the SAD algorithm are based on a key observation as follows. If the attacker can compromise the trains' position data, the three equality conditions Eqs. (5), (13), and (14) that the attacker must obey form an underdetermined problem with $3N$ variables and $2N$ equations. Since the other conditions that the attacker needs to follow (i.e., Eqs. (6) to (11), and (15)) are inequalities, the attacker's problem of finding stealthy FDI attack vectors most likely has infinitely many solutions that are continuous. However, if the trains' position data is intact, the three equality conditions Eqs. (5), (13), and (14) with $\mathbf{s}'$ replaced by the known $\mathbf{s}$, will form a determined problem with $2N$ variables and $2N$ equations. As a result, the attacker's problem most likely has a finite number of discrete solutions[2] and the attacker must choose one of them that is different from the true measurement vector. The SAD algorithm uses this discrete solution property to detect the onset of a BDD-stealthy attack.

2) SAD Algorithm: Based on the discrete solution property, we design the SAD algorithm as follows.

Algorithm 1 (Secondary Attack Detection Algorithm):
Inputs: Trains' true positions $\mathbf{s}$, possibly compromised measurement vector $\tilde{\mathbf{z}}$, intact nodal voltage vector $\mathbf{v}_{\text{pr}}$ at the previous time instant
Output: Attack onset detection result

1. Using $\tilde{\mathbf{z}}$, compute $P_i = \tilde{V}_i \tilde{I}_i,\ i \in \mathcal{N}_{\text{trains}}$.
2. Solve the following constrained optimization problem

$$J^* = \min_{\mathbf{v}_a \neq \mathbf{v}_b} \quad ||\mathbf{v}_a - \mathbf{v}_b||_p \tag{16a}$$
$$s.t. \quad \mathbf{Y}(\mathbf{s})\mathbf{v}_a = \mathbf{c}_a, \tag{16b}$$
$$\mathbf{Y}(\mathbf{s})\mathbf{v}_b = \mathbf{c}_b, \tag{16c}$$

   where $||\mathbf{x}||_p$ represents the p-norm of a vector $\mathbf{x}$ and the kth element of $\mathbf{c}_x \in \mathbb{R}^{N\times 1}$ is given by

$$\mathbf{c}_x[k] = \begin{cases} \frac{V_{\text{NL}} - \mathbf{v}_x[k]}{R_s}, & \text{if } k \in \mathcal{N}_{\text{sub}}; \\ \frac{P_k}{\mathbf{v}_x[k]}, & \text{if } k \in \mathcal{N}_{\text{trains}}, \end{cases} \tag{17}$$

   where the label x is a or b.
3. Extract $\widetilde{\mathbf{v}}$ from $\widetilde{\mathbf{z}}$. If $||\widetilde{\mathbf{v}} - \mathbf{v}_{\text{pr}}||_p \leq J^*$, report no attack; Otherwise, report onset of attack.

In Step 1 of the algorithm, given the possibly compromised measurement vector $\tilde{\mathbf{z}}$, the TPS monitor computes the actual power absorption or injection of each train. This follows from Eq. (4). Based on the trains' true positions $\mathbf{s}$ and powers, in Step 2, the TPS monitor solves the constrained optimization problem Eq. (16). The constraints in Eqs. (16b) and (16c) are compact representations of the BDD bypass condition given by Eqs. (13) and (14), for two distinct solutions $\mathbf{v}_a$ and $\mathbf{v}_b$. By the observation that the BDD bypass condition given the trains' true positions has discrete solutions, the J^* given by Eq. (16) is the minimum distance defined in p-norm between any two distinct solutions.

In Step 3, the TPS monitor compares the J^* with the p-norm distance between the possibly compromised voltage measurement vector and the intact nodal voltage vector $\mathbf{v}_{\text{pr}}$ at the previous time instant, to determine the possible onset of an attack. This step is based on that, if the attacker launches a BDD-stealthy attack without tampering with the trains' position information, the p-norm distance between the compromised voltage vector and the voltage vector in the absence of attack must be no less than J^*. As the voltage vector in the absence of attack is unknown to the TPS monitor, a *practical approach* is to use the $\mathbf{v}_{\text{pr}}$ that is not compromised before the onset of the attack. Since the TPS monitor can run the SAD periodically and frequently (e.g., every second), the TPS state will not change significantly over one monitoring time interval. In Section VI, extensive simulations demonstrate the effectiveness of this practical approach by comparing it with an *oracle approach* that uses the voltage vector at the present time instant in the absence of attack in Step 3. If and when the onset of an attack is detected, the TPS switches to an attack mitigation mode [30] to prevent safety breaches.

[2]As the determined problem is non-linear, it could yield infinitely many solutions that are continuous. However, this is not the case for any of the numerical examples and extensive simulations we conducted. Confirmation of the discrete solution property by rigorous analysis is left for future work.

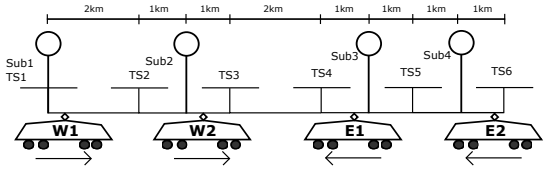


Fig. 4: Simulation setup. Sub - Substations, TS - Train stations, W - Trains departing from the west, E - Trains departing from the east.

VI. Simulations

Our analyses in the previous sections address a particular time instant only. In this section, we conduct time-domain simulations with realistic running profiles of trains to illustrate the impact of FDI attacks. We also show the effectiveness of the BDD in reducing the impact of the attacks, and that of the SAD in detecting those attacks that are BDD-stealthy.

A. Simulation Settings and Methodology

As Fig. 4 illustrates, we simulate a TPS consisting of four trains (labeled "W1", "W2", "E1", and "E2"), four substations (labeled "Sub1", ..., "Sub4"), and six train stations (labeled "TS1", ..., "TS6"). The parameters of the TPS are identical to those in Table I. The positions of the substations and the train stations are shown in Fig. 4. The trains "W1" and "W2" start their journeys from "TS1" and travel from west to east, whereas the trains "E1" and "E2" start their journeys from "TS6" and travel from east to west.

The trains "W1" and "E1" depart at time zero and the trains "W2" and "E2" depart at the 170th second. At each of the train stations, the trains stop for a duration of 20 seconds. Each train follows the same speed profile as shown in the top part of Fig. 5. The second plot of Fig. 5 shows the trains' positions over time. Each train switches between traction and braking modes during the simulation, and its power demand and regeneration capacity over time are shown in the bottom plot of Fig. 5. This plot is derived based on mechanical energy consumption of the train under the specified running profile, and with an efficiency ratio of 70% for the traction mode [19] and 40% for the braking mode [31] of converting kinetic energy into electrical energy. We simulate the TPS for 800 seconds at a time granularity of one second.

To simulate attacks, the attacker injects an attack vector computed using the methods given in Sections IV and V every second. In the absence of BDD, the attacker compromises the voltage and current measurements of all the train nodes. In the presence of BDD, the attacker tampers with the voltage and current measurements of all the train and substation nodes as well as the position information of the train nodes. The position information of substations cannot be compromised since their locations are fixed and known *a priori*. The maximum errors that the attacker can introduce to the voltage, current, and position measurements, as described in Eqs. (10), (11), and (15), are set as $\Delta V_i = 50\,\mathrm{V}$, $\Delta I_i = 200\,\mathrm{A}$, and $\Delta s_i = 500\,\mathrm{m}$ for $i \in \mathcal{N}_a$, unless otherwise specified.

The simulations are carried out in MATLAB. The constrained optimization problems are solved using the `fmincon` function of MATLAB with the `MultiStart` algorithm. In the absence of attack, to compute the system state, we use the `fmincon` with a constant objective function and the electrical models and trains' local control laws presented in Section III as the constraints. We also use the function to compute the safety attack vectors under the heuristic approach and the optimal efficiency attack vectors. If at any time instant, the `fmincon` function returns an attack vector that is the same as the true system state, the attacker does not launch an attack. Step 2 of the SAD algorithm is also implemented using the `fmincon` function.

Although our analysis in this paper is general and applicable to a TPS network of arbitrary size and topology, for simulations, we consider a small-scale TPS in Fig. 4. The rationale is two fold. First, the attacker may find it difficult to coordinate his attacks on a large number of geographically distributed trains. Computing resources may present another barrier for large-scale attacks. A more credible scenario is for the attacker to focus on one or a few trains in a TPS section. Second, since real-world TPS networks are mostly radial [16], the impact of a focused and localized attack will not propagate over long distances. In view of these factors, we use the small-scale TPS to represent well a TPS section in a large system.

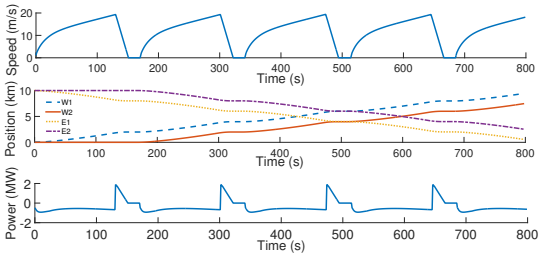


Fig. 5: Train speed (top plot), position (middle plot), power demand and regeneration capacity (bottom plot) over time. Power demand is negative and regeneration capacity is positive.

Moreover, to simplify our simulations, we do not consider overcurrent control. Specifically, we set the triggering threshold $V^{\mathrm{Tr}}_{i,\min}$ to a low value, so that overcurrent control will not be activated. As a result, the trains' speed profiles will not change because the trains need not curtail their power consumption. Therefore, at any time instant, a train's power consumption equals its power demand during acceleration. Because of this simplification, we do not simulate attacks on tractioning trains, which would alter the tractioning trains' power consumption and change their running profiles. Although we can simulate overcurrent control and attacks on tractioning trains by extending our simulator to admit changeable running profiles, the simulations reported in this paper already provide interesting understanding and insights into the impact of attacks and the effectiveness of countermeasures.

B. Simulation Results

1) Efficiency Attacks: The first set of simulations evaluates the impact of efficiency attacks on the TPS without BDD. Fig. 6 shows the power absorbed/injected by the train "E1" in the presence and absence of attacks. We can see that the efficiency attacks induce the regenerating trains to inject less power into the power network (e.g., from 302th to 315th second for the train "E1"). To calculate the loss in system efficiency, we ignore the time instants when all the trains are in traction mode, since we do not simulate attacks on the tractioning trains as discussed in Section VI-A. As a result, the efficiency attacks cause a reduction of 28.3% in the total energy adsorbed by the substations compared with the case of no attacks, during the time periods when there is at least one regenerating train under attack.

The second set of simulations evaluates the impact of efficiency attacks on the TPS with BDD. Fig. 7 shows the power absorbed/injected in the absence and presence of attacks. Although the efficiency attack can still induce the regenerating trains to inject less power to the power network, it causes a reduction of 6.2% only in the total energy adsorbed by the substations, during the time periods when there is at least one regenerating train under attack. This is in contrast to the 28.3% for the TPS without BDD. This result is consistent with our discussion in Section V-B that the BDD can reduce the impact of efficiency attacks.

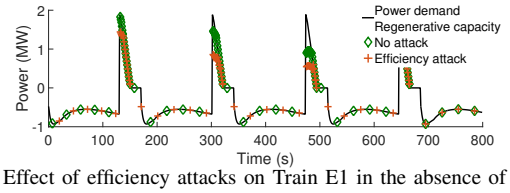


Fig. 6: Effect of efficiency attacks on Train E1 in the absence of BDD.

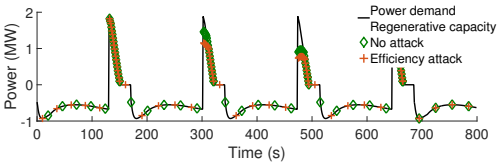
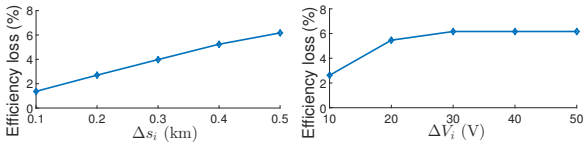


Fig. 7: Effect of efficiency attacks on Train E1 in the presence of BDD.



(a) Impact of Δs_i ($\Delta V_i = 50\,\text{V}$, $\Delta I_i = 200\,\text{A}$)

(b) Impact of ΔV_i ($\Delta s_i = 0.5\,\text{km}$, $\Delta I_i = 200\,\text{A}$)

Fig. 8: Effect of efficiency attacks on the TPS with BDD under different settings of Δs_i and ΔV_i.

We also examine the effect of efficiency attacks on the TPS with BDD under different settings of Δs_i and ΔV_i in Fig. 8a and Fig. 8b, respectively. From these figures, we can see that, under smaller settings of Δs_i and ΔV_i, the efficiency loss caused by the FDI attack diminishes. For instance, the efficiency loss is as low as 1.37% when $\Delta s_i = 0.1\,\text{km}$. In practice, the TPS monitor can estimate the present train position based on the train's speed and its position at the previous time instant when it was known that there were no attacks. The present position reading can be compared with the estimated position using Eq. (15). The setting of Δs_i should consider natural errors of train positioning systems and the estimation error. Existing train positioning systems such as GPS and Balise can achieve an accuracy of five to ten meters [32], [33]. Thus, it is reasonable to assume that the combined effect of the train positioning system error and the estimation error is less than $0.1\,\text{km}$. Our results show that by properly tuning the BDD's attack detection parameters (e.g., Δs_i and ΔV_i), the efficiency loss caused by FDI attacks can be significantly reduced.

2) Safety Attacks: The first set of simulations evaluates the impact of safety attacks on the TPS without BDD. Under safety attacks, the regenerating trains inject more power into the power network than that under no attacks, resulting in increased voltages. We say that the TPS experiences a safety breach when at least one node in the TPS experiences a safety breach. In the simulations, we observed that the TPS experiences safety breaches for a total of eight seconds. The prolonged overvoltage may cause safety incidents.

The second set of simulations evaluates the impact of safety attacks on the TPS with BDD. In these simulations, the TPS experiences safety breaches for a total of four seconds, compared with eight seconds in the absence of BDD. Table IV summarizes the time durations of safety breaches under different settings of Δs_i. When Δs_i is in the range of $0.1\,\text{km}$ to $0.3\,\text{km}$, the attack causes no safety breaches during the simulations. As discussed previously, the setting $\Delta s_i = 0.1\,\text{km}$ is appropriate in practice. Hence, this set of results shows that by appropriately setting the BDD parameters, safety breaches can be nearly eliminated.

Table IV: Time duration while the TPS experiences safety breaches under different settings of Δs_i in the presence of BDD.

Δs_i (km)	0.1	0.2	0.3	0.4	0.5
Time duration of safety breaches(s)	0	0	0	1	4

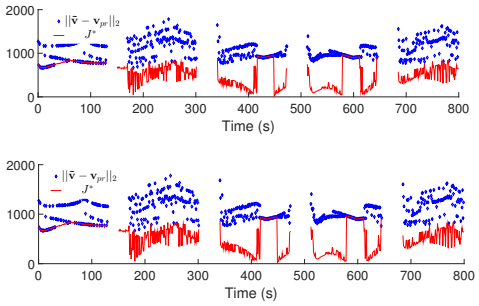


Fig. 9: The $||\tilde{\mathbf{v}} - \mathbf{v}_{\text{pr}}||_p$ and the adaptive threshold J^* computed by the Step 2 of the SAD algorithm with $p = 2$. Top figure: Oracle approach; Bottom figure: Practical approach.

3) SAD Algorithm: The last set of simulations evaluates the effectiveness of SAD in detecting attacks that have bypassed the BDD. For each time instant, the attacker tactically chooses the attack vector closest to the true system state in the sense of the p-norm distance, among the discrete solutions to the BDD bypass condition discussed in Section V-C. We compare our *practical approach* where the $\mathbf{v}_{\text{pr}}$ is the nodal voltage vector at the previous time instant (cf. Algorithm 1), with an *oracle approach* where the $\mathbf{v}_{\text{pr}}$ is the nodal voltage vector at the present time instant in the absence of attack. Fig. 9 shows the J^* computed by Step 2 of Algorithm 1 during the entire simulation, as well as the $||\tilde{\mathbf{v}} - \mathbf{v}_{\text{pr}}||_p$ under the oracle approach and the practical approach, where $p = 2$. In Fig. 9, we skip plotting a data point for a time instant when the attacker cannot bypass the BDD. For the oracle approach, from the top figure of Fig. 9, the $||\tilde{\mathbf{v}} - \mathbf{v}_{\text{pr}}||_p$ is consistently higher than the J^* for the entire simulation. This suggests that the oracle approach can detect the onset of a BDD-stealthy attack launched at any time instant. For the practical approach, from the bottom figure of Fig. 9, the $||\tilde{\mathbf{v}} - \mathbf{v}_{\text{pr}}||_p$ is higher than the J^* for 96% of the simulation time. For the remaining 4% of simulation time, the practical approach will miss the attack onset because of a significant change of $\mathbf{v}$ from the previous time instant to the present. This shows that the practical approach can detect the attack onset with a high probability.

VII. Conclusions

In this paper, we studied FDI attacks on train-borne sensor measurements used in railway TPSes. To the best

of our knowledge, ours is the first effort studying TPSes from a cybersecurity perspective. To account for the safety-criticality of TPS, we adopted the Kerckhoffs's principle and addressed two fundamental problems of importance, namely, characterization of the impact of FDI attacks on TPSes, and development of detection techniques for these attacks. We formulated and analyzed the efficiency and safety attacks that aim to minimize the system energy efficiency and breach system safety conditions, respectively. To detect these attacks, we proposed a global detection system that serializes the proposed BDD and SAD algorithms, both of which may be implemented at a central TPS monitor. Our simulation results verified the susceptibility of a TPS setup to the FDI attacks, but these attacks can be detected effectively by the proposed global detection system.

ACKNOWLEDGMENT

This work was supported in part by the National Research Foundation (NRF), Prime Minister's Office, Singapore, under its National Cybersecurity R&D Programme (Award No. NRF2014NCR-NCR001-31) and administered by the National Cybersecurity R&D Directorate, in part by the Human-Centered Cyber-physical Systems Programme at the Advanced Digital Sciences Center from Singapore's Agency for Science, Technology and Research (A*STAR), and in part by a Start-up Grant at NTU.

REFERENCES

[1] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *Conf. IEEE Industrial Electronics Society*, 2011.
[2] "The dragonfly attack," https://bit.ly/1RzGx2P.
[3] "U.S. DHS insider threat to utilities," https://bit.ly/1YPFoZH.
[4] "Deadly derailment in Moscow metro," https://bit.ly/1D4PF5F.
[5] SMRT, "Press release," Jul. 2015, https://bit.ly/1RxGBSk.
[6] R. Fletcher, "Regenerative equipment for railway rolling stock," *Power Engineering Journal*, vol. 5, no. 3, pp. 105–114, May 1991.
[7] Y. Okada, T. Koseki, and K. Hisatomi, "Power management control in DC-electrified railways for the regenerative braking systems of electric trains," *Advances in Transport*, vol. 15, pp. 919–929, 2004.
[8] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, "Ghost talk: Mitigating EMI signal injection attacks against analog sensors," in *IEEE Symp. Security and Privacy*, 2013.
[9] R. Karri, J. Rajendran, K. Rosenfeld, and M. Tehranipoor, "Trustworthy hardware: Identifying and classifying hardware trojans," *Computer*, vol. 43, no. 10, pp. 39–46, Oct 2010.
[10] M. Davis, "Recoverable advanced metering infrastructure," in *Black Hat Technical Security Conference*, 2009.
[11] M. A. Rahman, E. Al-Shaer, and R. G. Kavasseri, "A formal model for verifying the impact of stealthy attacks on optimal power flow in power grids," in *ACM/IEEE ICCPS*, 2014.
[12] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *CCS*, 2009.
[13] Y. Cai, M. Irving, and S. Case, "Iterative techniques for the solution of complex DC-rail-traction systems including regenerative braking," *IEE Proc. Generation, Transmission and Distribution*, vol. 142, no. 5, 1995.
[14] P. Arboleya, B. Mohamed, C. Gonzalez-Moran, and I. El-Sayed, "BFS algorithm for voltage-constrained meshed DC traction networks with nonsmooth voltage-dependent loads and generators," *IEEE Trans. Power Syst.*, 2015, http://dx.doi.org/10.1109/TPWRS.2015.2420574.
[15] C. Pires, S. Nabeta, and J. Cardoso, "ICCG method applied to solve DC traction load flow including earthing models," *IET Electric Power Applications*, vol. 1, no. 2, pp. 193–198, March 2007.
[16] L. Abrahamsson, *Optimal Railroad Power Supply System Operation and Design*. PhD Thesis, KTH Sweden, 2012.
[17] P. Arboleya, G. Diaz, and M. Coto, "Unified AC/DC power flow for traction systems: A new concept," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2421–2430, July 2012.
[18] "The train that powers its station," http://bbc.in/1KRROZK.
[19] S. Su, T. Tang, and C. Roberts, "A cooperative train control model for energy saving," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 622–631, Apr. 2015.
[20] A. U. Raghunathan, T. Wada, K. Ueda, and S. Takahashi, "Minimizing energy consumption in railways by voltage control on substations," in *Intl. Conf. Railway Engineering Design and Optimization*, 2014.
[21] A. A. Cárdenas, S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang, and S. Sastry, "Attacks against process control systems: Risk assessment, detection, and response," in *ACM AsiaCCS*, 2011.
[22] S. Amin, X. Litrico, S. Sastry, and A. M. Bayen, "Cyber security of water scada systemspart I: Analysis and experimentation of stealthy deception attacks," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 5, pp. 1963–1970, Sep. 2013.
[23] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1294–1305, Jul. 2013.
[24] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, 2011.
[25] S. N. Talukdar and R. Koo, "The analysis of electrified ground transportation networks," *IEEE Trans. Power Syst.*, vol. 96, no. 1, 1977.
[26] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security Privacy*, vol. 7, no. 3, pp. 75–77, 2009.
[27] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Electricity theft: Overview, issues, prevention and a smart meter based approach to control theft," *Energy Policy*, vol. 39, no. 2, pp. 1007–1015, Feb. 2011.
[28] A. Wood and B. Wollenberg, *Power Generation, Operation, and Control*. A Wiley-Interscience, 1996.
[29] "ERTMS/ETCS on-board ALSTOM solution," https://bit.ly/1OOb38f.
[30] "Technical report of this work," http://bit.ly/1RSVLkk.
[31] S. Acikbas and M. T. Soylemez, "Parameters affecting braking energy recuperation rate in dc rail transit," in *ASME/IEEE Joint Rail Conf. & Internal Combustion Engine Division Spring Technical Conf.*, 2007.
[32] The Economic Times – Railways, "Indian Railways to launch real-time train tracking via Google maps," https://bit.ly/1OIcMOe.
[33] K. Hartwig, M. Grimm, M. M. zu Hrste, and K. Lemmer, "Requirements for safety relevant positioning applications in rail traffic - A demonstrator for a train borne navigation platform called "DemoOrt"," http://elib.dlr.de/21252/1/wcrr.pdf.