Timely Detection and Mitigation of IoT-based Cyberattacks in the Smart Grid

Yasin Yılmaz[a,*], Suleyman Uludag[b]

[a] *Department of Electrical Engineering, University of South Florida, Tampa, FL, USA.*
[b] *Department of Computer Science, Engineering and Physics, University of Michigan - Flint, MI, USA.*

Abstract

The ongoing changes, updates, and upgrades of the Smart Grid infrastructure open up new cybersecurity challenges whose successful and satisfactory handling is a vital necessity for a viable future of these initiatives. The characteristic of the Smart Grid that leads to physical damage and cascading power failures amplifies the severity of security breaches. A set of recent successful Distributed Denial-of-Service (DDoS) attacks on the Internet, facilitated by the proliferation of the Internet-of-Things (IoT) powered botnets, shows that the Smart Grid may become the target and likely victim of such an attack, potentially leaving catastrophic outage of power service to millions of people. In this paper, under a hierarchical data collection infrastructure we propose a general and scalable mitigation approach, called Minimally Invasive Attack Mitigation via Detection Isolation and Localization (MIAMI-DIL), based on an online and nonparametric anomaly detection algorithm which is scalable and capable of timely detection. We provide a proof-of-concept by means of simulations to show the efficacy and scalability of the proposed approach.

1. Introduction

The vision of the Smart Grid brings about enhanced automation, computing, communications and control characteristics. At the same time, a vital need emerges to address the plethora of security and privacy related challenges. The essential nature of the Smart Grid cybersecurity spans availability, integrity, and confidentiality of computing, communications, and/or control devices from intentional or accidental harm and damage. The severity of cybersecurity consequences in the Smart Grid is generally exasperated due to the complexity, sheer volume of the devices and stakeholders, and highly time sensitive operational

*Corresponding author
Email addresses: yasiny@usf.edu (Yasin Yılmaz), uludag@umich.edu (Suleyman Uludag)

constraints. For instance, a cascading blackout may leave thousands, if not millions, of customers without power for long time periods. A recent report by the University of Cambridge details a severe but plausible cyberattack against the US grid where about 100M people may be left without power with up to $1 trillion of monetary loss [1].

The recent proliferation of the Internet-of-Things (IoT) devices (8.4B in 2017 and expected to hit 20B by 2020 [2]) significantly expands the attack vectors of adversaries. Across the industries, and specifically in the power grid, a new genre of attack vectors and malicious capabilities are emerging as a result of this constant connection of objects, sensors, and services, commonly referred to as IoT. A prolific example is the Mirai malware [3]. The Mirai botnet, composed of a huge army of compromised IoT devices worldwide (mostly IP cameras, DVRs, and consumer routers with default passwords), and its variants, initiated tens of thousands of distributed denial-of-service (DDoS) attacks against high-profile targets such as Dyn, a popular DNS provider [4]. Such recent attacks demonstrate the ease and effectiveness of this IoT-based security threat by means of small, resource constrained, and hard-to-patch devices. What is even more dreading is the fact that Mirai botnet is highly customizable through its open source code, and it is available for sale as a service in the Dark Web, significantly improving the attack capabilities of even unsophisticated malicious actors [5]. It is just a matter of time before these capabilities are employed against critical infrastructures and cyber-physical systems, like the Smart Grid. Examples of vulnerable IoT devices in the Smart Grid include smart meters, smart light bulbs, smart thermostats, connected vehicles, electric vehicles, smart street lights, smart home appliances, etc. For instance, through compromised devices, adversaries can generate false data injection (FDI) attacks by sending manipulated energy consumption data. On the second installment of the Quadrennial Energy Review (QER) report titled "Transforming the Nation's Electricity System", released in January 2017, evolving cyber threats from botnets, especially via DoS attacks and the associated cybersecurity risks are highlighted. Finally, in the hands of more sophisticated parties, such as state actors, such attacks may become even more lethal.

The effects of cyberattacks on the Smart Grid have been manifested recently in the real-world attacks in Ukraine, both December 2015 and 2016, where the city of Ivano-Frankivsk with 100K people was cut from power for 6 hours as a result of a cyberattack. There is also the physical dimension of attacks on the Smart Grid. Cyber-physical attacks, also called blended attacks, may cause a greater damage when combined than the individual attacks separately [6]. An experiment demonstrating cyberattacks that lead to physical harm was conducted in 2007 in by Idaho National Lab (Aurora Test) where a pure cyberattack resulted in a diesel-generator going up in smokes and exploding[1]. A recent study also demonstrated, through simulations, the feasibility of launching IoT botnet-initiated attacks leading to disruption of power delivery through

[1]See the video of the experiment at `https://youtu.be/rTkXgqK1l9A`.

three different categories of vulnerabilities based on demand manipulation [7].

Even without any cybersecurity risks, it is generally agreed that the utilities should continuously monitor and proactively detect abnormal conditions to prevent disruption in power delivery and to improve grid reliability. Real-time analytics is becoming a promising approach for an effective solution to this end [8]. For example, a key factor for successful deployment of smart meter infrastructure is reported to be its data analytics [9]. The need is even more pronounced with the cybersecurity threats.

While there has been some cybersecurity related work on the Smart Grid related areas, to the best of our knowledge, except the conference version of this paper [10], there is no other study in the literature to mitigate the aforementioned new genre of IoT-initiated cyberattacks against the Smart Grid, which seems to be a prime target by many categories of adversaries at the first opportunity they get, such as nation states, curious/motivated eavesdroppers, terrorists/cyber-terrorists, organized crime, disgruntled employees, etc.

In this paper, we present several attack scenarios that can be initiated via IoT devices. We then provide a framework, called MIAMI-DIL (Minimally Invasive Attack Mitigation via Detection Isolation and Localization). The algorithmic underpinning of MIAMI-DIL's anomaly detection is based on an online, and non-parametric approach with a distributed statistical inference methodology that scales well to high-dimensional systems and provides small average detection delay. We specifically note that when the DoS attack is changed from a pure brute-force method to a more stealth one by trying to hide the attack, the detection speed of our approach is several orders of magnitude faster than the parametric alternatives (see Section 5). Compared to the conference version [10] we provide a much more comprehensive discussion of the topic with asymptotic performance analysis of the proposed detection method, a detailed mitigation approach and analysis, as well as significantly expanded simulation results for average detection delay in different attack scenarios.

The rest of the paper is organized as follows: Related work is presented in Section 2. The system and threat models are explained in Section 3. Our anomaly-based intrusion detection system (IDS) formulations with analytical details are provided in Section 4. Simulation results are given in Section 5. Finally, we conclude the paper in Section 6.

2. Related Work

Threats and vulnerabilities targeting the availability dimension of security, under the umbrella name of DoS attacks, are not new. They have been studied for the Internet for a while with many proposed defense mechanisms, e.g., [11, 12, 13, 14, 15]. Yet, providing efficient and effective solutions and mitigation techniques for the Internet DoS attacks are still challenging and elusive.

When it comes to the Smart Grid, the potential damage of such attacks is even more profound and due to the peculiar features of the infrastructure, the DoS attacks pose an even harder challenge [16]. A specification-based IDS is proposed in [17] tailored for the application layer protocol of ANSI C12.22 to catch

violations of the specified security policy. However, there are other protocols used in industry and even the same protocol might be deployed with proprietary implementations to make such an approach infeasible for all cases [18]. Another packet-level inspection for intrusion detection is proposed in [19] on encrypted traffic, albeit with the same application layer protocol. A fourth order Markov Chain is used to model the event logs of data aggregators in [20], which can only scale to a small number of aggregators and cannot be deployed on the smart meters. A hierarchical distributed IDS for the Smart Grid is proposed in [21], designed for specific wireless mesh network technology assumptions. An anomaly-based IDS is presented in [22] that can only be deployed at headend and the data aggregators due to its computational complexity. Recently, the false data injection attacks [23, 24, 25] and the jamming attacks [26, 27, 28, 29] against the smart grid are extensively studied in the literature. While conventional detectors classify a measurement as anomalous if the measurement residual exceeds a certain threshold [24, 30, 31, 32], there are also several online detectors based on the quickest detection theory that improve the timely and reliable detection of cyber-attacks more reliably [33, 34, 35, 36].

With the above literature review and the pertinent features of the Smart Grid infrastructure as discussed in Section 3, a pure centralized intrusion detection would not yield acceptable performance due to the heterogeneity of the constituent and independent networks [17, 37]. A distributed and computationally efficient technique is needed to facilitate deployment at many system devices. Finally, to make the deployment feasible to as many systems as possible, it should not be tied to a specific protocol or data type. We address all of these features in our approach as explained in the following sections.

3. System and Threat Model

We consider a hierarchical system where a set of electrical devices[2] is connected to the Smart Grid by means of a smart meter in a Home Area Network (HAN)[3] as shown in Fig. 1. Neighborhood Area Network (NAN) or Field Area Network (FAN) represents a logical association of these smart meters. Data aggregators collect, summarize, and report the data from HAN through the Wide Area Network (WAN) to the utility's headend or the control center.

The smart meters may report a variety of different data back to the utility, from pricing to consumption data to power quality monitoring [38]. In this study, we do not assume anything about the specific data type and we call this approach *protocol and data type agnostic*.

The proliferation of the IoT devices connected to the Smart Grid coupled with the ease of launching such attacks, as recently exemplified by Mirai bot-

[2]Smart appliances (e.g., smart bulbs, smart thermostat, electrical vehicle, and other relevant IoT devices) at home, connected machinery in an industrial setting, business equipment in commercial environment, etc.

[3]While we have only depicted HAN, our ensuing approach is equally applicable to Industrial Area Networks (IANs).

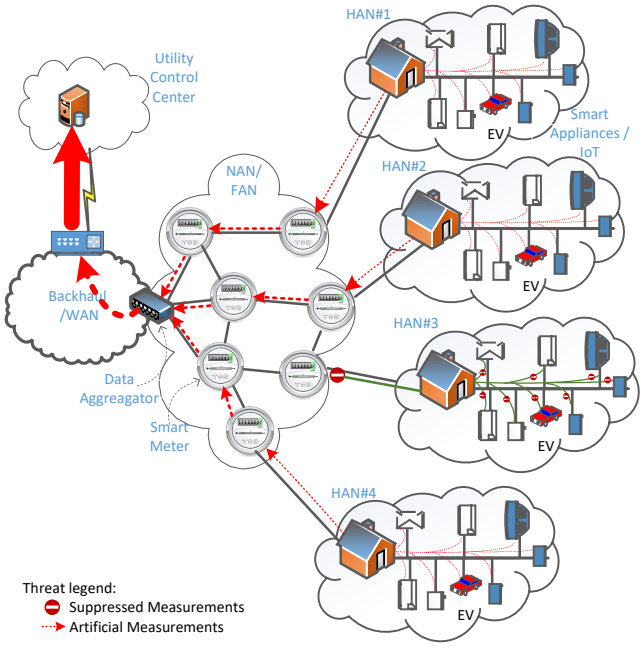

Figure 1: Threat model in our Smart Grid Model: In HAN #3, the attacker(s) try to mislead the data collection process by means of a classical DoS mechanism, such as jamming, to prevent data delivery. IoT appliances in HAN #1, 2, and 4, on the other hand, inject superfluous data to mislead the data collection in order to compromise the availability.

net, presents significantly expanded attack vectors. Thus, our threat model comprises any legitimate but superfluous traffic or false data packets injected into the Smart Grid at HAN, NAN, or FAN levels in order to compromise the system, such as resource exhaustion, routing, reflector, de-synchronization attacks, etc.

More specifically, we first consider a false data injection (FDI) attack through compromised devices for the purpose of misleading the state estimation subsystem of the power grid at the headend [33], as shown by the dotted lines in Fig. 1. Another threat we consider is a DoS attack through the prevention of data transmission from the lower levels to the control center, as shown by HAN#3 in Fig. 1, possibly by means of simple physical layer jamming attack [39]. Although FDI attacks are much harder to perform, they are at the same time much more effective when successful than jamming attacks since they deceive the receiver and may cause wrong decisions systemwide (i.e., wrong data caused by FDI vs. no data due to jamming). Specifically, in smart grid, FDI attacks may cause significant errors in state estimation, which may result in wide-area blackouts. In FDI attacks, an internal intervention to the transmitted data is required. For example, the transmitted signal can be manipulated

either by compromising the transmitting device or by hacking into the communication channel (i.e., knowing the frequency band and mimicking the modulation scheme). Whereas for jamming external intervention on the transmitted signal is sufficient by continuously transmitting noise in a range of frequency bands. In the considered threats, attackers can manipulate the data content or suppress the data communication, but not the statistics used for the defense mechanism (see Fig. 2). We assume that the statistics are transmitted through secure channels, or the attackers are not aware of the details of defense mechanism, and thus does not attack the defense mechanism itself. We reserve the case where the defense mechanism is vulnerable to attacks as a future work.

4. Anomaly-Based IDS and Attack Mitigation

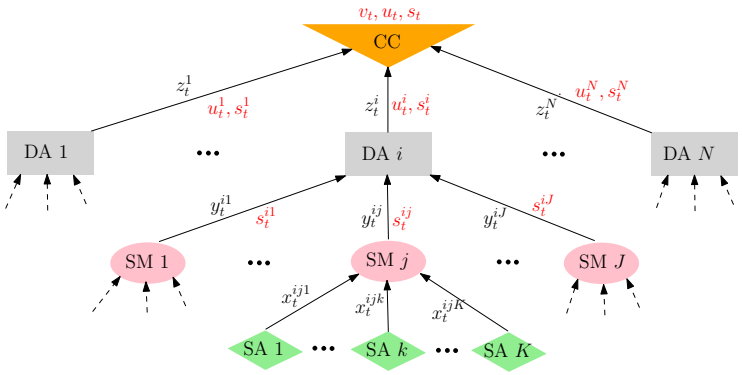

Figure 2: Proposed hierarchical IDS. Data are shown in black, e.g., x_t^{ijk}, and statistics are shown in red, e.g., s_t^{ij}.

Considering the hierarchical topology of the Smart Grid and security threats at each level of such a hierarchy, as shown in Figure 1, we propose a hierarchical and distributed IDS that consists of several subsystems. Specifically, each smart meter j in each NAN i monitors the streaming data $\{x_t^{ijk} : \forall k, t\}$ from each smart home appliance (i.e., IoT device) k in its HAN, and computes a statistic s_t^{ij} at each time $t = 1, 2, \ldots$, as shown in Figure 2. Similarly, each data aggregator i monitors the streaming data $\{y_t^{ij} : \forall j, t\}$ from each smart meter j in its NAN, and computes a statistic u_t^i. Furthermore, it gathers the statistics $\{s_t^{ij} : \forall j, t\}$, from its smart meters, and combines them in s_t^i. Finally, the control center monitors the data $\{z_t^i : \forall i, t\}$ from each data aggregator i, using which it computes a statistic v_t, and also combines the statistics $\{s_t^i\}$ and $\{u_t^i\}$ in s_t and u_t, respectively (see Figure 2).

Using the computed statistics s_t, u_t and v_t the control center performs the detection and mitigation procedure shown in Figure 3. If an attack is detected using the method presented in Section 4.3.1, then the mitigation procedure described in Section 4.3.2 is used to localize the problematic nodes and isolate the malicious data traffic.

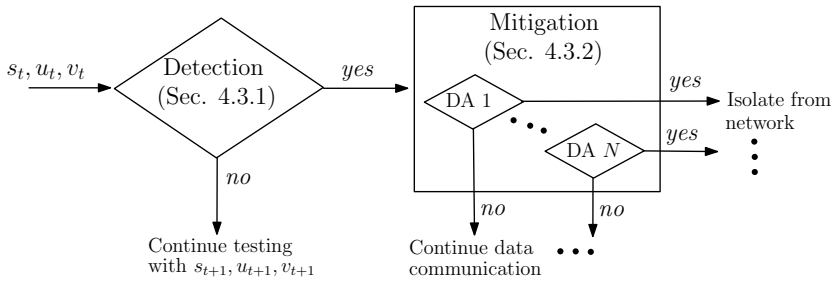

Figure 3: Flow chart of the proposed detection and mitigation mechanism at the control center.

For generality, we do not specify the types of data represented by $\{x_t^{ijk}\}$, $\{y_t^{ij}\}$ and $\{z_t^i\}$. Some example data types communicated in the Smart Grid are energy consumption, voltage phase, current, active and reactive power, and power factor [38]. We only assume the observed data are numerical which can be normalized, e.g., to lie in $[0, 1]$ using upper and lower bounds, $\boldsymbol{x}_t^{ij} = [x_t^{ij1} \cdots x_t^{ijK}] \in [0, 1]^K$. We do not have assumptions on the probability distribution of K data dimensions. For instance, they can be correlated or follow different probability distributions (see Fig. 12). Normalization within each dimension is needed to deal with heterogeneity among data dimensions. Normalization by mean and standard deviation or some practical bounds such as 5th and 95th percentiles can be used if the absolute lower and upper bounds are unknown or have extreme values. The statistics s_t^{ij}, u_t^i and v_t are computed to detect anomalies in the data $\{x_t^{ijk}\}$, $\{y_t^{ij}\}$ and $\{z_t^i\}$, respectively. Anomalies might be caused by various types of threats, such as false data injection, man-in-the-middle, spoofing, and jamming [40]. We show how to compute the statistics s_t^{ij}, u_t^i and v_t, as well as s_t^i, s_t and u_t in Sections 4.2 and 4.3.

4.1. Online Nonparametric Anomaly Detection

Anomaly detection in our Smart Grid model is quite challenging due to the following reasons:

(C1) The *attack patterns are typically unknown* since there is a wide range of vulnerabilities for attackers, especially considering the lack of stringent security measures in the IoT devices (such as smart appliances). Hence, parametric anomaly detection-based IDSs that assume probabilistic models for anomalies, as well as conventional signature-based IDSs are not feasible in this emerging security threat.

(C2) The *problem is inherently high-dimensional* given the large number of IoT devices in a typical HAN (i.e., the dimension of $\boldsymbol{x}_t^{ij}$) and the number of smart meters in a NAN (i.e., the dimension of $\boldsymbol{y}_t^i = [y_t^{i1} \cdots y_t^{iJ}]$). Thus, computationally efficient algorithms that can scale well to high dimensionality are required.

(C3) *Timely and accurate detection is critical* given the broad societal impacts of a successful attack to the Smart Grid, and also due to the stringent response time requirements in the Smart Grid (e.g., real-time pricing).

Anomaly detection-based IDS has the capability of detecting unknown attacks under certain conditions. It typically needs to know a statistical description of the nominal (i.e., no attack) behavior, denoted as the baseline, and classifies each outlying instance that significantly deviates from the baseline as an anomaly. This conventional interpretation of anomaly detection is also called *outlier detection*. Ideally, with the nominal probability distribution f_0 completely known, an instance x is deemed an outlier if its likelihood under the nominal distribution is smaller than a predefined threshold. Equivalently, x is declared an outlier if it is outside the most compact set of data points under the nominal distribution, called the minimum volume set Ω_α given by

$$\Omega_\alpha = \arg\min_{\mathcal{A}} \int_{\mathcal{A}} \mathrm{d}y \;\; \text{subject to} \;\; \int_{\mathcal{A}} f_0(y)\mathrm{d}y \geq 1-\alpha, \tag{1}$$

where a data point is deemed nominal in the region $\mathcal{A}$, and α is the significance level, i.e., constraint on the false alarm probability. In high-dimensional problems like the one considered in this paper, even if f_0 is known, it is computationally very expensive (if not impossible) to determine Ω_α. Hence, in the literature, there are various methods for learning minimum volume sets [41]. One of them, called Geometric Entropy Minimization (GEM), is shown to be very effective with high-dimensional datasets [42] while asymptotically achieving the performance of minimum volume set [43].

The GEM approach provides a *scalable nonparametric* anomaly detector, addressing the first two challenges (C1) and (C2); however, it lacks the temporal aspect in the third challenge (C3), i.e., analyzes each time instance separately. As a result, it does not accumulate the anomaly evidences, as opposed to the sequential (i.e., online) detection techniques such as the Cumulative Sum (CUSUM) algorithm, which are tailored for *timely and accurate detection* [44]. Specifically, an outlying instance is a nominal tail event (i.e., false alarm) with probability α (e.g., $\alpha = 0.05$ is a typical value), but the probability of consecutive outliers being nominal is much lower.

4.2. Online Discrepancy Test (ODIT)

Recently a GEM-based *online and nonparametric* anomaly detector, called *Online Discrepancy Test (ODIT)*, was proposed in [45] to timely detect abrupt and persistent anomalies. ODIT combines the simplicity of the GEM approach with the timely and accurate detection capabilities of the CUSUM algorithm to enable online anomaly detection in high-dimensional problems. Hence, in this paper, we use ODIT to develop an effective and efficient IDS for mitigating IoT-based DoS attacks in the Smart Grid.

We next show the ODIT procedure for smart meter j under data aggregator i, which observes the data vector $\boldsymbol{x}_t^{ij}$ at each time t. ODIT assumes a training

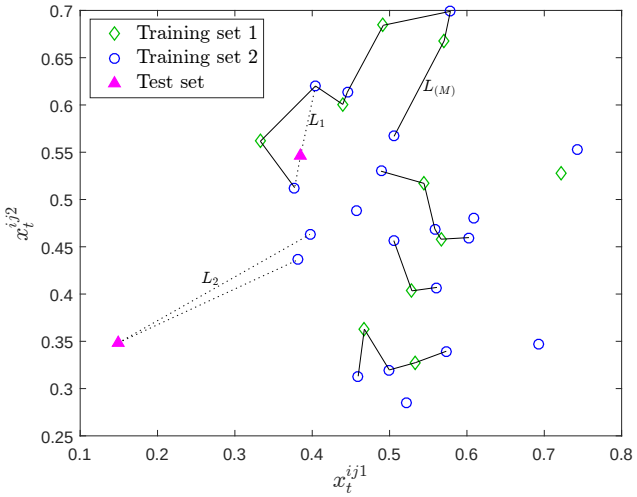

Figure 4: ODIT procedure with $N_1 = 10$, $N_2 = 20$, $M = 9$, $k = 2$, $s = 1$, $\gamma = 1$. $L_1 - L_{(M)}$ and $L_2 - L_{(M)}$ are used as in (3) for online anomaly detection (see also Figure 5). First test point is from the same nominal distribution as training points, which is a two-dimensional Gaussian with independent components with 0.5 mean and 0.1 standard deviation. Second test point is from uniform distribution over $[0, 1]$.

dataset $\mathcal{X}_N = \{\boldsymbol{x}_1^{ij}, \ldots, \boldsymbol{x}_N^{ij}\}$ that is free of anomaly, and randomly separates it into two subsets $\mathcal{X}^{N_1}$ and $\mathcal{X}^{N_2}$ for computational efficiency, as in the bipartite GEM algorithm [42]. Then, for each point in $\mathcal{X}^{N_1}$ it finds the k nearest neighbors from $\mathcal{X}^{N_2}$, and forms an M-point k-nearest-neighbor (M-kNN) Euclidean graph $G = (\mathcal{X}_M^{N_1}, E)$ by selecting the M points $\mathcal{X}_M^{N_1}$ in $\mathcal{X}^{N_1}$ with the smallest total edge length and their k closest neighbors in $\mathcal{X}^{N_2}$, where $E = \{e_{m(n)}\}$ is the set of edges with $e_{m(n)}$ denoting the edge between point m in $\mathcal{X}^{N_1}$ and its nth nearest neighbor in $\mathcal{X}^{N_2}$. The total edge length of a point m in $\mathcal{X}^{N_1}$ is given by

$$L_m = \sum_{n=k-s+1}^{k} |e_{m(n)}|^\gamma, \tag{2}$$

where $|e_{m(n)}|$ is the Euclidean distance between point m and its nth nearest neighbor in $\mathcal{X}^{N_2}$, $1 \leq s \leq k$ is a fixed number introduced for convenience, and $\gamma > 0$ is the weight. It is known [42] that $\mathcal{X}_M^{N_1}$ converges to the minimum volume set Ω_α as

$$\lim_{M, N_1 \to \infty} M/N_1 \to 1 - \alpha.$$

An example M-kNN graph is shown in Figure 4, where "Training set 1" and "Training set 2" denote $\mathcal{X}^{N_1}$ and $\mathcal{X}^{N_2}$, respectively, and the edges are shown with solid lines. In this M-kNN graph, 9 out of 10 points in $\mathcal{X}^{N_1}$ are connected to their 2 nearest neighbors in $\mathcal{X}^{N_2}$ with $L_{(M)}$ showing the longest edge in the graph. Two test points and their longest edges (L_1 and L_2) are also shown.

In *outlier detection* by bipartite GEM [42], each test point $\boldsymbol{x}_t^{ij}$ is classified as an outlier if its total edge length L_t is greater than that of the Mth point, which has the largest total edge length, in $\mathcal{X}_M^{N_1}$, i.e., $L_t > L_{(M)}$. On the other hand, in ODIT,

$$D_t = L_t - L_{(M)} \tag{3}$$

is treated as some positive/negative evidence for anomaly, which approximates the log-likelihood ratio $\ell_t = \log \frac{p(\boldsymbol{x}_t^{ij}|\mathsf{H}_1)}{p(\boldsymbol{x}_t^{ij}|\mathsf{H}_0)}$ between the alternative hypothesis H_1 claiming $\boldsymbol{x}_t^{ij}$ is anomalous ($\boldsymbol{x}_t^{ij} \notin \Omega_\alpha$) and the null hypothesis H_0 claiming $\boldsymbol{x}_t^{ij}$ is nominal ($\boldsymbol{x}_t^{ij} \in \Omega_\alpha$) [45].

Theorem 1. *As the training set size increases ($N_1, N_2 \to \infty$) we have the following asymptotic relationships*

$$\lim_{N_1,N_2\to\infty} D_t \overset{\text{monotonic}}{\sim} \log \frac{f_0(\boldsymbol{x}_\alpha)}{f_0(\boldsymbol{x}_t^{ij})}, \tag{4}$$

$$\textit{and} \quad \operatorname{sign}\left(\lim_{N_1,N_2\to\infty} D_t\right) = \operatorname{sign}\left(\log \frac{f_0(\boldsymbol{x}_\alpha)}{f_0(\boldsymbol{x}_t^{ij})}\right), \tag{5}$$

where $\overset{\text{monotonic}}{\sim}$ denotes a monotonic relationship between two variables, f_0 denotes the nominal probability distribution, $\boldsymbol{x}_\alpha$ is a boundary point of Ω_α (see (1)), N_1 and N_2 are the size of two partitions in the training set, and the test statistic D_t (given in (3)) is the difference between the total edge lengths of the new point $\boldsymbol{x}_t^{ij}$ and $\boldsymbol{x}_{(M)}^{ij}$, the Mth point in $\mathcal{X}_M^{N_1}$, which has the largest total edge length in $\mathcal{X}_M^{N_1}$.

Proof. The asymptotic properties in (4) and (5) follow from the asymptotic optimality of GEM. We start with (5). It is known [42] that the decision rule of GEM converges to that of the minimum volume set Ω_α, given by "choose H_0 if $f_0(\boldsymbol{x}_t^{ij}) \geq f_0(x_\alpha)$, i.e., $\log \frac{f_0(\boldsymbol{x}_\alpha)}{f_0(\boldsymbol{x}_t^{ij})} \leq 0$, and choose H_1 otherwise", hence the sign property in (5). To prove (4) assume that, as $N_2 \to \infty$, also $k \to \infty$ such that the total edge length $L_k(\boldsymbol{x}_t^{ij})$ of a point $\boldsymbol{x}_t^{ij}$ remains a constant. In that case, $L_k(\boldsymbol{x}_{t_1}^{ij}) < L_k(\boldsymbol{x}_{t_2}^{ij})$ for all $\boldsymbol{x}_{t_1}^{ij}$ and $\boldsymbol{x}_{t_2}^{ij}$ such that $f_0(\boldsymbol{x}_{t_1}^{ij}) > f_0(\boldsymbol{x}_{t_2}^{ij})$. Since

$$D_t = L_k(\boldsymbol{x}_t^{ij}) - L_k(\boldsymbol{x}_{(M)}^{ij}), \tag{6}$$

we have the monotonicity property stated in (4). Note also that $\boldsymbol{x}_{(M)}^{ij} \to \boldsymbol{x}_\alpha$ as $N_1, M \to \infty$ such that $M/N_1 = 1 - \alpha$. □

Theorem 1 shows the structural resemblance of D_t to the log-likelihood ratio between the boundary point $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}_t^{ij}$. To see the geometric relationship consider the case where f_0 is from the exponential family, i.e., $f_0 = e^{-\delta(\boldsymbol{x}_t^{ij},\boldsymbol{\theta})}$ where $\boldsymbol{\theta}$ is the parameter vector and $\delta(\boldsymbol{x}_t^{ij}, \boldsymbol{\theta})$ is a distance term causing the exponential decay in the probability density function. In this case, $\log \frac{f_0(\boldsymbol{x}_\alpha)}{f_0(\boldsymbol{x}_t^{ij})} =$

$\delta(\boldsymbol{x}_t^{ij}, \boldsymbol{\theta}) - \delta(\boldsymbol{x}_\alpha, \boldsymbol{\theta})$ is a distance metric that is similar to D_t as shown by (6). They also asymptotically share a very similar structure (see Theorem 1).

Assuming the data $\boldsymbol{x}_t^{ij}$ is independent over time, $\sum_{t=1}^{T} D_t$ gives the aggregate anomaly evidence until time T, similar to the running log-likelihood $\sum_{t=1}^{T} \ell_t$, which is the sufficient statistic for optimum statistical detection.

In the sequential change detection problem, it is assumed that the actual probability distribution of the data $\boldsymbol{x}_t^{ij}$ is initially the nominal distribution $f_0(\boldsymbol{x}_t^{ij})$, but after some random time τ (e.g., attack time), an abrupt and persistent change occurs and the actual distribution switches to an anomalous distribution $f_1(\boldsymbol{x}_t^{ij})$. A commonly used performance criterion is the minimax criterion in which the worst-case *expected detection delay* $\mathsf{E}[T_d - \tau | \{\boldsymbol{x}_1^{ij}, \ldots, \boldsymbol{x}_{T_d}^{ij}\}, T_d \geq \tau]$ is minimized while satisfying a false alarm constraint, where T_d is the detection time. More information on sequential change detection can be found in [44].

CUSUM is the optimum sequential change detection algorithm in terms of the minimax criterion when both the nominal distribution $f_0(\boldsymbol{x}_t^{ij})$ and the anomalous distribution $f_1(\boldsymbol{x}_t^{ij})$ are completely known [44]. This is not a realistic assumption for intrusion detection in the Smart Grid as the attack patterns are typically unknown, and even estimating the nominal distribution is intractable due to high-dimensionality, as noted in the challenges in Section 4.1. Generalized CUSUM, which is used in [33, 35] for detecting false data injection attacks against voltage phase estimation in the Smart Grid, assumes parametric distributions for f_0 and f_1, and estimates their parameters from data. CUSUM can be regarded as a clairvoyant detector in the considered Smart Grid security problem. Yet, the timely detection capability of CUSUM is very attractive here. Therefore, leveraging the analogy between the anomaly evidence D_t and the log-likelihood ℓ_t, shown in Theorem 1, ODIT mimics the CUSUM procedure for online and nonparametric anomaly detection (see [45] for a more technical discussion).

In particular, when the running log-likelihood $\sum_{t=1}^{T} \ell_t$ crosses a lower bound, say at time T_1, CUSUM decides that there is no change and restarts the test by considering $\sum_{t=T_1+1}^{T} \ell_t$. The test continues until $\sum_{t=T_n+1}^{T_d} \ell_t$ crosses an upper bound the first time, say after n restarts. In this case, CUSUM stops the test and decides for a change (e.g., an anomaly). Actually, in the CUSUM procedure, the lower bound is set to zero not to waste time to decide for a no change decision because it is known that initially there is no change. Hence, it is possible to recursively update the CUSUM statistic as $\mathcal{L}_t = \max\{\mathcal{L}_{t-1} + \ell_t, 0\}$, where $\mathcal{L}_0 = 0$. Then, the stopping time of CUSUM is given by $T_d = \min\{t : \mathcal{L}_t \geq h\}$, where $h > 0$ is a predetermined threshold. Similarly, the ODIT procedure is given by

$$T_d = \min\{t : s_t^{ij} \geq h\},\ s_t^{ij} = \max\{s_{t-1}^{ij} + D_t, 0\},\ s_0^{ij} = 0, \tag{7}$$

where D_t is given in (3), and $h > 0$ is a predetermined threshold. An example of s_t^{ij} and the detection procedure is shown in Figure 5.

The detection threshold h manifests a trade-off between minimizing the detection delay and minimizing the false alarm rate, as can be seen in Figure 5.

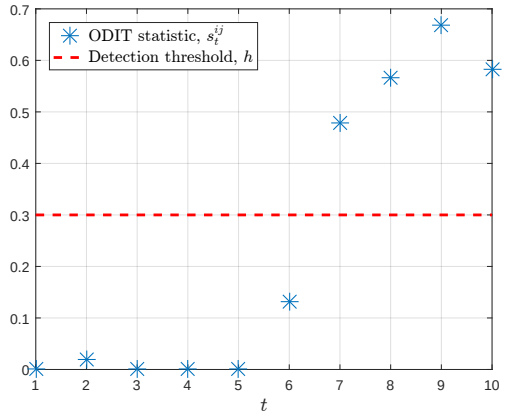

Figure 5: ODIT statistic and decision procedure using the setup in Figure 4 and anomalous test points from uniform distribution over $[0, 1]$. Anomaly starts at $t = 6$, and detected at $t = 7$ with the shown threshold.

Particularly, smaller threshold facilitates early detection, but also increases the probability of false alarm. In practice, h can be chosen to satisfy a given false alarm rate.

4.3. Proposed IDS and Attack Mitigation

After detecting an anomaly in the system, which potentially corresponds to an attack, control center takes action to mitigate its effects in a minimally invasive fashion (i.e., with minimal service interruption). Specifically, it first *isolates* the information flow from suspected nodes, and then *localizes* the actual attack places via further investigation. We call this framework *MIAMI-DIL (Minimally Invasive Attack Mitigation via Detection Isolation and Localization)*, which is summarized below.

4.3.1. Detection

We presented the ODIT anomaly detector for a single smart meter in Section 4.2. Leveraging the spatial diversity that is inherent to the hierarchical structure as shown in Figure 2, we propose a system-wide IDS in which each smart meter j does not decide alone based on its data $\{\boldsymbol{x}_t^{ij}\}$, but instead cooperates with other smart meters by passing its test statistic s_t^{ij} to its parent node, data aggregator i. Gathering $\{s_t^{ij}\}$ data aggregator i fuses them into $s_t^i = \sum_{j=1}^{J} s_t^{ij}$, and passes it to the control center, together with the statistic u_t^i of data $\{y_t^{ij}\}$ it receives from its children. It computes u_t^i in the same way as s_t^{ij}, as shown in (7). Note that s_t^{ij} denotes the evidence for anomaly at smart meter j, and summing the independent evidences $\{s_t^{ij} : \forall j\}$ from different smart meters we get the total evidence s_t^i among smart meters. Summing $\{s_t^{ij} : \forall j\}$ coincides with summing the independent CUSUM statistics to obtain a global CUSUM statistic, which is known to be optimum when the change times at different

nodes (where independent CUSUM statistics are computed) are not restricted to be identical [46].

Finally, control center, receiving the statistics $\{s_t^i, u_t^i\}$ obtains $s_t = \sum_{i=1}^N s_t^i$ and $u_t = \sum_{i=1}^N u_t^i$, and computes, following (7), the statistic v_t of data it receives from data aggregators. The statistics s_t, u_t, and v_t measure the anomaly (i.e., attack) evidence at different levels of hierarchy, namely smart appliances, smart meters, and data aggregators, respectively; hence they potentially exhibit heterogeneity. For instance, when there is an attack to data aggregators, the statistical evidence for attack would appear in v_t only, whereas the evidence for an attack targeting smart appliances could be visible in all layers s_t, u_t, and v_t. Using each of them control center runs three separate ODIT procedures

$$T_s = \min\{t : s_t \geq h_s\}, T_u = \min\{t : u_t \geq h_u\},\ T_v = \min\{t : v_t \geq h_v\} \quad (8)$$

decides for an anomaly the first time one of them stops, i.e.,

$$T_d = \min\{T_s, T_u, T_v\}. \quad (9)$$

Control center gathers data and statistics from the network, and sequentially *detects* possible anomalies in the system using the IDS given by (9). The mitigation approach, which consists of *localization* and *isolation*, is explained next.

4.3.2. Mitigation

To protect the control center from going offline due to data flooding (DoS) and making wrong decisions due to falsified data, the components under attack should be *localized*, and the data coming from them should be disregarded, i.e., the components under attack should be *isolated* from the data communication infrastructure until they are cleared from the attack.

After detection, control center identifies the data aggregators which positively contributed to s_t or u_t or v_t. Identifying data aggregators with highly positive s_t^i and u_t^i is straightforward. To identify data aggregators that contribute to highly positive v_t we compare, with a threshold, the average contribution of each dimension (i.e., data aggregator) to the total edge length L_t of data vector $\boldsymbol{z}_t$ (see Figure 2) since the last time v_t was 0 (denoted by $\hat{\tau}$) until the detection time T_d. That is, data from aggregator i is identified as problematic if

$$\frac{1}{T_d - \hat{\tau}} \sum_{t=\hat{\tau}+1}^{T_d} r_t^i > \beta, \quad (10)$$

where $\beta > 0$ is a threshold chosen to strike a balance between high detection probability (true positive rate) and small false alarm probability (false positive rate), and r_t^i is the sum of distances of the data z_t^i to the ith dimension of the considered nearest neighbors (see (2)).

Depending on the attack characteristics (e.g., change in data content or number of data packets in different sets of nodes) there might be different combinations of increase in these statistics. In order to have a general mitigation

mechanism for a variety of attacks, considering vulnerable data aggregators we propose that

- the control center *localizes* the attack by identifying the suspected data aggregators using (10), and
- the identified data aggregators are temporarily *isolated* from the network (i.e., data from them is disregarded) until further investigation.

After the attacked nodes are further *localized* through human investigation, data communication resumes immediately with the clean nodes (if the data aggregator itself is also clean from attacks) and after a cleaning/securing procedure with the attacked nodes.

Note that the security level increases in the higher levels of network, i.e., data aggregators are typically much more secure than smart meters and smart appliances. If the data aggregators are highly trusted to be secure by design, localization of attacked smart meters can be performed automatically without human investigation and there is no need to isolate the entire data coming from identified aggregators. Following the procedure given in (10)

- each identified data aggregator *localizes* the suspected smart meters, and
- the traffic coming only from these suspected meters are *isolated*.

In this case, human supervision is only needed to fix the attacked nodes. For a data aggregator using (10), the mitigation statistic r_t^j represents the sum of distances of the data y_t^{ij}. We should note that the procedure in (10) requires some memory to store the most recent $T_d - \hat{\tau}$ distance values for all dimensions local to the node performing the procedure.

5. Performance Evaluation

In this section, we numerically evaluate the performance of the proposed IDS. We consider a Smart Grid that consists of a control center, $N = 10$ data aggregators, $J = 100$ smart meters under each data aggregator, and $K = 10$ smart appliances under each smart meter, yielding a total of $10,000$ smart appliances and $1,000$ smart meters system-wide.

Baseline Model: In each home-area network, we assume data from appliances are independent and identically distributed (iid) with $x_t^{ijk} \sim \mathcal{N}(0.5, 0.01)$, similar to the IoT dataset in [47] [4]. Note that the nominal data traffic is commonly modeled as Gaussian, e.g., [48]. For generality, we do not specify the data type. An intuitive example is future energy consumption data, that can be predicted by the IoT using the scheduled usage and historical data. Smart

[4] This dataset can be found at `https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_N_BaIoT#`

meters can use this data for energy hedging to gain robustness to the volatility of real-time prices. In case of a successful attack, falsified data not only misleads pricing and hedging, but may also cause a demand-supply imbalance in the system, which might destabilize the grid and cause catastrophic outcomes such as a wide-area blackout. In line with this example, each smart meter sends the anticipated average energy consumption in the HAN, i.e., $y_t^{ij} = \frac{1}{K}\sum_{k=1}^{K} x_t^{ijk}$, and similarly each data aggregator reports $z_t^i = \frac{1}{J}\sum_{j=1}^{J} y_t^{ij}$.

Attack Model: Note that the security level typically decreases as we go down the hierarchy of Figure 2, hence we consider a practical scenario where smart appliances (i.e., IoT devices) and smart meters could be under attack, but data aggregators are secure. In the considered scenarios, 3% of the $1,000$ HANs are attacked. In each attacked HAN, with probability 0.5 data from each smart appliance is manipulated by means of a DoS attack, for example through through jamming, or by a false data injection attack. In the former, variance is increased, $x_t^{ijk} \sim \mathcal{N}(0.5, (0.1\eta)^2), \eta > 1$, and in the latter, mean is changed, $x_t^{ijk} \sim \mathcal{N}(0.5 + \Delta, 0.01), \Delta \in \mathbb{R}$.

1) DoS via jamming: We first analyze the jamming-type DoS attack targeting the availability of the data at the smart meters or data aggregators. This can be realized through increasing the variance of transmitted data from compromised smart appliances or smart meters. Numerically we simulate this type of attack by significantly increasing the variance of x_t^{ijk}. In Figure 6, where the variance is 25 times the nominal value (i.e., $\eta = 5$), sample test statistics of the proposed ODIT detector and the parametric CUSUM detectors from a single trial are shown.

CUSUM is a clairvoyant detector, which is assumed to know the baseline and anomalous probability distributions exactly, and hence is not of practical interest. Specifically, it sequentially tests the baseline distribution $\mathcal{N}(\boldsymbol{x}_t^{ij}; \mathbf{1}_{[0.5]}, \boldsymbol{I}_{[0.01]})$ against the anomalous distribution $\mathcal{N}(\boldsymbol{x}_t^{ij}; \mathbf{1}_{[0.5]}, \boldsymbol{I}_{[0.25]})$, where $\mathbf{1}_{[a]}$ is the vector with all entries a, and $\boldsymbol{I}_{[a]}$ is the diagonal matrix with entries a. Generalized CUSUM (G-CUSUM), which is the practical version of CUSUM, on the other hand estimates the parameters of the baseline distribution and tests it against an assumed anomalous distribution. In Figure 6 and Figure 7, to show the effect of mismatch between the actual and estimated/assumed values for G-CUSUM we set the estimated baseline mean and variance as 0.505 and 0.0102, and the assumed anomalous mean and variance as 0.505 and 0.3672 (i.e., $\eta = 6$), respectively. Note that G-CUSUM still knows the true distributions – in practice, there might be also mismatch between the actual and estimated/assumed distributions.

It is clearly seen in Figure 6 that all three detectors detect the attack immediately after it occurs, i.e., there is a clear distinction between the pre-attack and post-attack behaviors of the test statistics. This is actually the case in all trials, thus exhibiting perfect detection with no false alarms for all three detectors. There is an obvious trade-off in selecting the jamming magnitude η. Large values such as the one in Figure 6 has the potential to completely make the data unavailable at the smart meter, but at the same time are easily detectable.

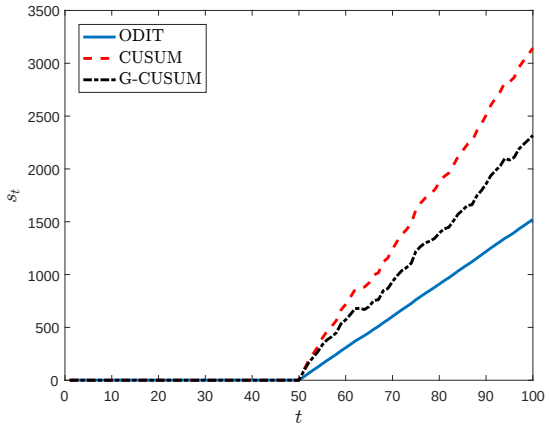

Figure 6: Sample test statistic s_t under jamming-type DoS attack ($\eta = 5$) for ODIT, CUSUM, and Generalized CUSUM. Attack starts at $t = 51$.

Alternatively, attacker might favor small η values to avoid detection (stealth attack) while still targeting the availability of data by increasing the noise level. We investigate this case in Figure 7, where average detection delay is plotted

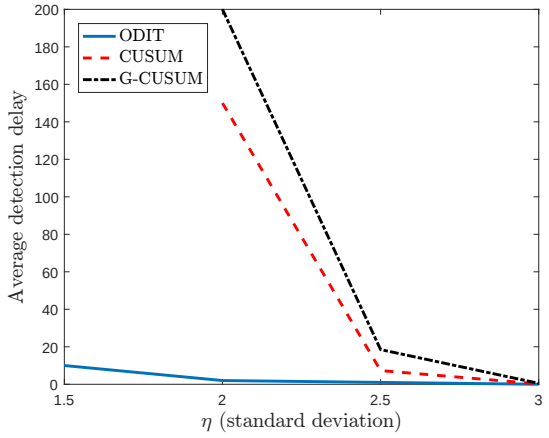

Figure 7: Average detection delay vs. jamming noise level in terms of the nominal standard deviation.

versus different η values. The detection delay in each trial refers to the time lag between the initialization of the attack and its detection. For all η values, average delay is measured considering the perfect detection case with no false alarm. It is observed that after $\eta = 3$ (i.e., noise level becomes 9 times the nominal variance), all three detectors achieve zero delay as also illustrated for $\eta = 5$ in Figure 6. However, for smaller attacks there is a huge performance gap between ODIT and the CUSUM detectors. Even the clairvoyant CUSUM

detector performs much worse than ODIT for $\eta < 2.5$.

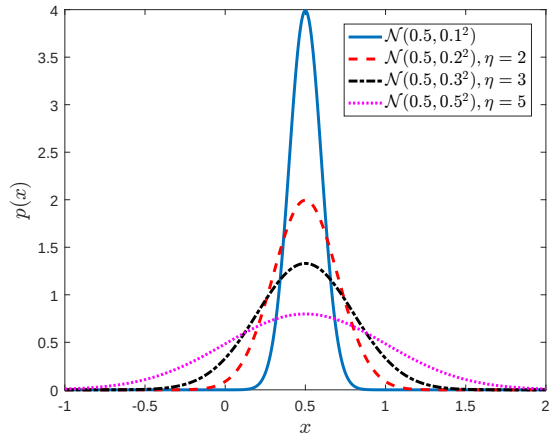

Figure 8: Probability density function for the baseline (variance 0.01) and three different jamming scenarios.

This result may sound counter-intuitive at first as CUSUM is typically known to be the optimum sequential change detector, however it should be noted that CUSUM is only *minimax* optimum, i.e., in terms of minimizing the worst-case average detection delay, not optimum per se in all cases. CUSUM compares the likelihoods under baseline and anomaly distributions, whereas ODIT measures only the discrepancy of data with respect to the nominal distribution. As shown for the univariate case in Figure 8, for $\eta = 2$, the region where the anomaly likelihood is smaller than the baseline likelihood holds a significant probability mass, hence the large detection delay in this case. As η increases (e.g., $\eta = 3$ and $\eta = 5$), the probability mass for which the anomaly likelihood is greater than the baseline likelihood increases, and as a result the attack becomes much more detectable. On the other hand, in ODIT, since there is no such comparison with the anomaly distribution, even the small discrepancies with respect to the baseline distribution accumulate and enable timely detection.

2) False Data Injection: We next investigate the false data injection attack scenario in which the mean of the data is altered by Δ. In Figure 9, $\Delta = 0.2$ (twice the nominal standard deviation) is considered. It is seen that ODIT significantly outperforms G-CUSUM, which estimates the baseline parameters with 1% error and tests the baseline against the 3 standard deviation mean-shifted versions in both directions. Estimation errors and mismatch in the assumed parameters for the unknown anomaly distribution causes a large performance degradation in G-CUSUM compared to CUSUM. The proposed nonparametric ODIT detector, on the other hand, achieves a close performance to the clairvoyant (non-practical) CUSUM detector.

We also analyze the combined attack scenario in which jamming and false data injection attacks are performed together. For the combination of previously

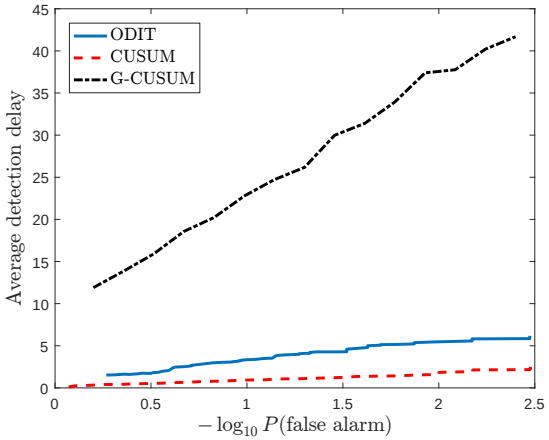

Figure 9: Average detection delay vs. false alarm probability under false data injection attack.

considered attack scenarios with $\Delta = 0.2$ and $\eta = 5$, Figure 10 similarly shows that ODIT achieves much lower average detection delays than G-CUSUM for the same false alarm rates.

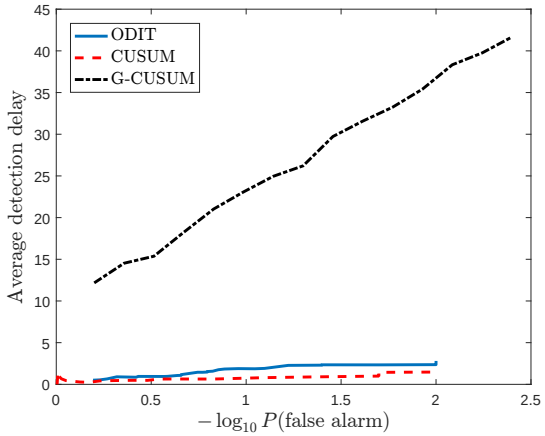

Figure 10: Average detection delay vs. false alarm probability under combined false data injection and jamming attack.

3) Identification of Compromised Nodes: In addition to detection, we now evaluate the mitigation performance of the proposed MIAMI-DIL framework (i.e., isolation and localization techniques) by quantifying the performance for identifying the compromised appliances and meters. The false data injection scenario above is considered. In Figure 11, the *Receiver Operating Characteristic (ROC)* curve of the identification procedure given by Equation (10) is shown. It shows that a very high detection probability (true positive rate) is achieved

even for very small false alarm probability (false positive rate) constraints.

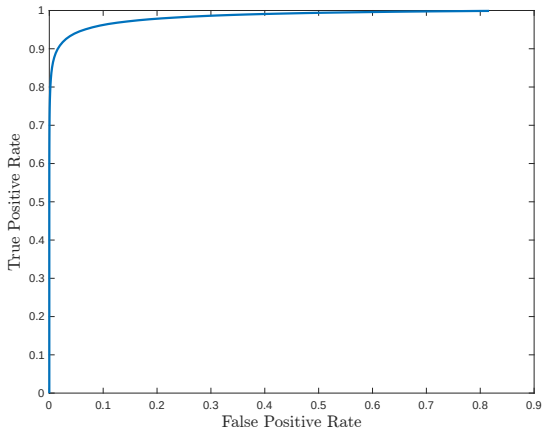

Figure 11: ROC curve for the mitigation (isolation and localization) procedure given by (10) under the false data injection attack scenario considered in Fig. 9.

4) Heterogeneous Data: We finally illustrate that our proposed method is capable of handling heterogeneous data after proper normalization thanks to its nonparametric (i.e., model-free) nature. Consider a two-dimensional highly heterogeneous system (e.g., two disparate IoT devices) with the first dimension following exponential distribution with mean 0.1, $x_t^{ij1} \sim \text{Exp}(10)$, and the second dimension following Gaussian distribution with mean $10,000$ and standard deviation $1,000$, $x_t^{ij2} \sim \mathcal{N}(10^4, 10^3)$. For simplicity, we consider a single-layer detector in which smart meter j runs ODIT using the statistic s_t^{ij}. After normalizing the data using the mean and standard deviation values estimated from the training data for each dimension, the proposed detection method timely and accurately detects an increase in the mean of first dimension from 0.1 to 0.5, as shown in Fig. 12. Since the first dimension takes much smaller values than the second dimension, such an increase in the first dimension would normally be invisible (buried under the large nominal values of the second dimension) unless data is properly normalized.

6. Conclusion

The elusive and challenging goal of providing effective and efficient solution to intrusion detection for the Smart Grid is poised to present more menacing, and thus more interesting research difficulties than in the Internet domain. Particularly, the mushrooming of the IoT devices, coupled with the ease of triggering cyberattacks even from unsophisticated malicious parties, make the challenge even more formidable. In light of these developments, we present two potential attack vectors, including a stealth one, facilitated by the expanded set of new IoT appliances and devices in the Smart Grid. Then, we develop

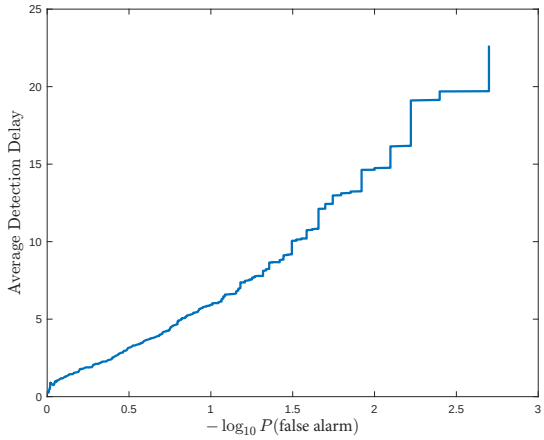

Figure 12: The average performance (average detection delay vs. false alarm rate) of the proposed detector in a highly heterogeneous system consisting of data following exponential distribution with mean 0.1 and Gaussian distribution with mean 10^4 and standard deviation 10^3. Attack increases only the mean of the exponential data to 0.5 without affecting the Gaussian data.

a novel intrusion detection framework, called Minimally Invasive Attack Mitigation via Detection Isolation and Localization (MIAMI-DIL) that employs a timely, distributed, scalable, and nonparametric intrusion detection system (IDS). Another important, distinguishing feature of MIAMI-DIL is that it is protocol-agnostic and free from data-type assumptions. We have numerically shown that in a challenging high-dimensional scenario, the proposed IDS is capable of timely and accurately detecting cyber-attacks, in some cases even more quickly than the optimally designed, clairvoyant Cumulative Sum (CUSUM) detector. Specifically, in detecting stealth DoS attacks, our approach is significantly faster than the clairvoyant CUSUM and its more practical version Generalized CUSUM (G-CUSUM), which estimates the model parameters. In this work, we did not consider the attacks targeting the proposed defense mechanism to focus on the novel framework. We reserve this interesting extension, as well as more challenging DoS attacks such as intermittent DoS and low-rate DoS to future works.

[1] Trevor Maynard and Nick Beecroft, Business Blackout, Lloyd's Emerging Risk Report (may 2015).

[2] Rob van der Meulen, Gartner Press Release February 2017 (feb 2017). URL `http://www.gartner.com/newsroom/id/3598917`

[3] C. Kolias, G. Kambourakis, A. Stavrou, J. Voas, Ddos in the iot: Mirai and other botnets, Computer 50 (7) (2017) 80–84.

[4] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis,

et al., Understanding the mirai botnet, in: USENIX Security Symposium, 2017.

[5] James Scott and Drew Spaniel, Rise of the Machines: The Dyn Attack Was Just a Practice Run, Institute for Critical Infrastructure Technology (ICIT) (dec 2016).

[6] Victoria Y. Pillitteri and Tanya L. Brewer, NISTIR 7628 Revision 1, Guidelines for Smart Grid Cybersecurity, Smart Grid Interoperability Panel (SGIP), Smart Grid Cybersecurity Committee (sep 2014).

[7] S. Soltan, P. Mittal, H. V. Poor, Blackiot: Iot botnet of high wattage devices can disrupt the power grid, in: 27th {USENIX} Security Symposium ({USENIX} Security 18), 2018, pp. 15–32.

[8] M. Chen, S. Mao, Y. Liu, Big data: A survey, in: Mobile Networks and Applications, Vol. 19, 2014, pp. 171–209. doi:10.1007/s11036-013-0489-0.

[9] D. Alahakoon, X. Yu, Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey, IEEE Trans. on Industrial Informatics 12 (1) (2016) 425–436. doi:10.1109/TII.2015.2414355.

[10] Y. Yilmaz, S. Uludag, Mitigating iot-based cyberattacks on the smart grid, in: IEEE International Conference on Machine Learning and Applications (ICMLA), 2017.

[11] R. K. C. Chang, Defending against flooding-based distributed denial-of-service attacks: A tutorial, IEEE Communications Magazine 40 (10) (2002) 42–51. doi:10.1109/MCOM.2002.1039856.

[12] T. Peng, C. Leckie, K. Ramamohanarao, Survey of network-based defense mechanisms countering the DoS and DDoS problems, ACM Computing Surveys 39 (1) (2007) 3–es. doi:10.1145/1216370.1216373.

[13] M. Abliz, Internet Denial of Service Attacks and Defense Mechanisms, Tech. rep. (2011).
URL `http://people.cs.pitt.edu/{~}mehmud/docs/abliz11-TR-11-178.pdf`

[14] S. T. Zargar, J. Joshi, D. Tipper, A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks, IEEE Communications Surveys & Tutorials 15 (4) (2013) 2046–2069. doi:10.1109/SURV.2013.031413.00127.

[15] M. Aamir, M. Arif, Study and Performance Evaluation on Recent DDoS Trends of Attack & Defense, I.J. Information Technology and Computer Science Information Technology and Computer Science 05 (08) (2013) 54–65. doi:10.5815/ijitcs.2013.08.06.

[16] A. Huseinovic, S. Mrdovic, K. Bicakci, S. Uludag, A Taxonomy of the Emerging Denial-of-Service Attacks in the Smart Grid and Countermeasures, in: 2018 26th IEEE Telecommunication Forum (TELFOR), Belgrade, Serbia, 2018, pp. 1–4.

[17] R. Berthier, W. H. Sanders, Specification-Based Intrusion Detection for Advanced Metering Infrastructures, in: IEEE 17th Pacific Rim Int'l Symposium on Dependable Computing, 2011, pp. 184–193. doi:10.1109/PRDC.2011.30.

[18] M. Q. Ali, E. Al-shaer, Configuration-based IDS for Advanced Metering Infrastructure, Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (2013) 451–462doi:10.1145/2508859.2516745.

[19] R. Berthier, D. I. Urbina, A. A. Cardenas, M. Guerrero, U. Herberg, J. G. Jetcheva, D. Mashima, J. H. Huh, R. B. Bobba, On the practicality of detecting anomalies with encrypted traffic in AMI, in: SmartGridComm 2014, IEEE, 2014, pp. 890–895. doi:10.1109/SmartGridComm.2014.7007761.

[20] M. Q. Ali, E. Al-Shaer, Randomization-Based Intrusion Detection System for Advanced Metering Infrastructure, ACM Trans. on Information and System Security 18 (2) (2015) 7:1—-7:30. doi:10.1145/2814936.

[21] Y. Zhang, L. Wang, W. Sun, R. C. G. Ii, M. Alam, Distributed Intrusion Detection System in a Multi-Layer Network Architecture of Smart Grids, IEEE Trans. on Smart Grid 2 (4) (2011) 796–808. doi:10.1109/TSG.2011.2159818.

[22] F. A. A. Alseiari, Z. Aung, Real-time anomaly-based distributed intrusion detection systems for advanced Metering Infrastructure utilizing stream data mining, in: Int'l Conf. on Smart Grid and Clean Energy Technologies (ICSGCE), IEEE, 2015, pp. 148–153. doi:10.1109/ICSGCE.2015.7454287.

[23] G. Liang, J. Zhao, F. Luo, S. R. Weller, Z. Y. Dong, A review of false data injection attacks against modern power systems, IEEE Transactions on Smart Grid 8 (4) (2017) 1630–1638.

[24] Y. Liu, P. Ning, M. K. Reiter, False data injection attacks against state estimation in electric power grids, ACM Transactions on Information and System Security (TISSEC) 14 (1) (2011) 13.

[25] S. Tan, D. De, W.-Z. Song, J. Yang, S. K. Das, Survey of security advances in smart grid: A data driven approach, IEEE Communications Surveys & Tutorials 19 (1) (2017) 397–422.

[26] S. Amin, A. A. Cárdenas, S. S. Sastry, Safe and secure networked control systems under denial-of-service attacks, in: International Workshop on Hybrid Systems: Computation and Control, Springer, 2009, pp. 31–45.

[27] A. Sargolzaei, K. Yen, M. Abdelghani, A. Abbaspour, S. Sargolzaei, Generalized attack model for networked control systems, evaluation of control methods, Intelligent Control and Automation 8 (03) (2017) 164.

[28] Y. Li, L. Shi, P. Cheng, J. Chen, D. E. Quevedo, Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach, IEEE Transactions on Automatic Control 60 (10) (2015) 2831–2836.

[29] D. Deka, R. Baldick, S. Vishwanath, Data attacks on power grids: Leveraging detection, in: Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society, IEEE, 2015, pp. 1–5.

[30] A. Gomez-Exposito, A. Abur, Power system state estimation: theory and implementation, CRC press, 2004.

[31] K. Manandhar, X. Cao, F. Hu, Y. Liu, Detection of faults and attacks including false data injection attack in smart grid using kalman filter, IEEE transactions on control of network systems 1 (4) (2014) 370–379.

[32] D. B. Rawat, C. Bajracharya, Detection of false data injection attacks in smart grid communication systems, IEEE Signal Processing Letters 22 (10) (2015) 1652–1656.

[33] S. Li, Y. Yılmaz, X. Wang, Quickest detection of false data injection attack in wide-area smart grids, IEEE Transactions on Smart Grid 6 (6) (2015) 2725–2735.

[34] Y. Huang, J. Tang, Y. Cheng, H. Li, K. A. Campbell, Z. Han, Real-time detection of false data injection in smart grid networks: an adaptive cusum method and analysis, IEEE Systems Journal 10 (2) (2016) 532–543.

[35] M. N. Kurt, Y. Yılmaz, X. Wang, Distributed quickest detection of cyber-attacks in smart grid, IEEE Transactions on Information Forensics and Security 13 (8) (2018) 2015–2030.

[36] M. N. Kurt, Y. Yılmaz, X. Wang, Real-time detection of hybrid and stealthy cyber-attacks in smart grid, IEEE Transactions on Information Forensics and Security 14 (2) (2019) 498–513. doi:10.1109/TIFS.2018.2854745.

[37] V. Gulisano, M. Almgren, M. Papatriantafilou, METIS: A Two-Tier Intrusion Detection System for Advanced Metering Infrastructures, Springer International Publishing, Cham, 2015, pp. 51–68.

[38] R. R. Mohassel, A. Fung, F. Mohammadi, K. Raahemifar, A survey on advanced metering infrastructure, International Journal of Electrical Power & Energy Systems 63 (2014) 473–484.

[39] H. Li, L. Lai, R. C. Qiu, A denial-of-service jamming game for remote state monitoring in smart grid, in: 2011 45th Annual Conference on Information Sciences and Systems, 2011, pp. 1–6. doi:10.1109/CISS.2011.5766137.

[40] Y. Lopes, N. C. Fernandes, T. B. de Castro, V. dos Santos Farias, J. D. Noce, J. P. Marques, D. C. Muchaluat-Saade, Vulnerabilities and threats in smart grid communication networks, Security Solutions and Applied Cryptography in Smart Grid Communications (2016) 1.

[41] C. Scott, R. Nowak, Learning minimum volume sets, Journal of Machine Learning Research 7 (2006) 665–704.

[42] K. Srichanran, A. O. Hero, Efficient anomaly detection using bipartite k-nn graphs, in: Advances in Neural Information Processing Systems (NIPS), 2011, pp. 478–486.

[43] A. O. Hero, Geometric entropy minimization (gem) for anomaly detection and localization, in: Advances in Neural Information Processing Systems (NIPS), 2006, pp. 585–592.

[44] M. Basseville, I. V. Nikiforov, Detection of abrupt changes : theory and application, Prentice Hall, 1993.

[45] Y. Yilmaz, Online nonparametric anomaly detection based on geometric entropy minimization, in: IEEE International Symposium on Information Theory (ISIT), 2017, pp. 3010–3014.

[46] Y. Mei, Efficient scalable schemes for monitoring a large number of data streams, Biometrika 97 (2) (2010) 419–433.

[47] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai, Y. Elovici, N-baiot: Network-based detection of iot botnet attacks using deep autoencoders, arXiv preprint arXiv:1805.03409.

[48] Y. Xiang, K. Li, W. Zhou, Low-rate ddos attacks detection and traceback by using new information metrics, IEEE transactions on information forensics and security 6 (2) (2011) 426–437.