Received 1 April 2013; revised 11 July 2013; accepted 15 July 2013. Date of publication 18 July 2013; date of current version 20 September 2013.

Digital Object Identifier 10.1109/TETC.2013.2274043

CONSUMER: A Novel Hybrid Intrusion Detection System for Distribution Networks in Smart Grid

CHUN-HAO LO (Student Member, IEEE) AND NIRWAN ANSARI (Fellow, IEEE)

Advanced Networking Laboratory, Department of Electrical and Computer Engineering,
New Jersey Institute of Technology, Newark, NJ 07102 USA

CORRESPONDING AUTHOR: C.-H. LO (CL96@njit.edu)

ABSTRACT Smart meters have been deployed worldwide in recent years that enable real-time communications and networking capabilities in power distribution systems. Problematically, recent reports have revealed incidents of energy theft in which dishonest customers would lower their electricity bills (aka stealing electricity) by tampering with their meters. The physical attack can be extended to a network attack by means of false data injection (FDI). This paper is thus motivated to investigate the currently-studied FDI attack by introducing the combination sum of energy profiles (CONSUMER) attack in a coordinated manner on a number of customers' smart meters, which results in a lower energy consumption reading for the attacker and a higher reading for the others in a neighborhood. We propose a CONSUMER attack model that is formulated into one type of coin change problems, which minimizes the number of compromised meters subject to the equality of an aggregated load to evade detection. A hybrid detection framework is developed to detect anomalous and malicious activities by incorporating our proposed grid sensor placement algorithm with observability analysis to increase the detection rate. Our simulations have shown that the network observability and detection accuracy can be improved by means of grid-placed sensor deployment.

INDEX TERMS Smart grid, cyber-physical security, state estimation, false data injection attack, energy theft, intrusion detection, sensor placement, observability.

I. INTRODUCTION

Integration of state-of-the-art information and communications technology (ICT), control, and computing is a critical enabler to facilitate grid modernization and optimization for the existing electric power systems. During the evolutional movement in smart grid development, the conventional critical infrastructure is gradually exposed to the public such that part of the systems especially the distribution networks involving smart metering communications along with controls of distributed generation and demand responses at consumption sites will potentially pose a number of security risks. Advanced Metering Infrastructure (AMI) in the distribution network is essentially comprised of endpoint-based home area networks (HANs), grid-based wireless sensor networks (WSNs), and access point-based neighborhood/field area networks (NANs and FANs) [1], [2]. Recently, a middleware architecture design has been proposed to consolidate heterogeneous quality of service/experience (QoS/QoE)-oriented smart grid applications, such as spectrum efficiency, power scheduling, and security protection [3], [4]. In the meantime, several surveys and tutorials have elaborately addressed a number of security issues in terms of confidentiality, integrity, and availability (CIA), from passive attacks to active attacks [5]–[14], such as eavesdropping, jamming, tampering, spoofing, altering, and other attacks against the protocol stacks of the OSI model; these attacks are foreseen inevitable and nontrivial within the context of the cyber-physical smart grid. Among which some literatures have emphasized the interrelationship between *cyber* and *physical* securities [5], [6], [15].

There are two primary research directions in smart grid security: 1) *a breach of network availability*: a power system involves real-time models that perform state estimation to observe the current state conditions in the power network by obtaining real-time measurement data from network meters and devices—without these data, state estimation cannot be effectively executed in real time, thus hampering the decision making of network operators—if the network communications is intruded by denial of service (DoS) attacks or other schemes against data availability, the services will be interrupted in both communications and power systems; and 2) *a breach of measurement data confidentiality and integrity*: due to the cause-effect attribute, if measurement data are further altered by intruders in a way that the attack is hard to be detected, not only customer privacy may be compromised, but the undetectability may also cause utilities to lose revenues and potentially result in severe power outage and equipment damages. Countermeasures relied on cryptographic mechanisms, secure communications architecture and network designs, device security, and intrusion detection systems (IDS) are anticipated options for securing the future power system against malicious intrusions and attacks from all perspectives in a complementary manner. The implementation of various strategic approaches will be based on different smart grid applications as well as communications requirements throughout the networks.

According to the Institute for Electric Efficiency (IEE) [16], one-third of households in the U.S. have had a smart meter (i.e., approximately 36 million smart meters) as of May 2012, and approximately 65 million smart meters will have been deployed by 2015. While the deployments continue to rise, a few *energy theft* incidents have been discovered that illegal customers intended to lower their electricity bills via meter tampering, bypassing, or other unlawful schemes regardless of traditional or smart meters in places such as Ireland, Hong Kong, and Virginia U.S. [17]. Notably, energy theft is one dominant component of non-technical losses, which account for 10%–40% of energy distribution [18], e.g., $1–6 billion losses due to energy theft yearly for utilities in the U.S. Moreover, the report [19] revealed that the current installations of smart meter communications protocols and associated infrastructure do not have sufficient security controls to protect the electric power system against false data injection attacks, not to mention older meters which were not designed to adequately cope with such attacks. In addition to the physical attacks, network attacks by compromising meters can also introduce malicious measurement data and cause degradation of grid operation [20], [21]. While some protection schemes against malicious network traffic have been proposed for smart grid communications networks monitoring [22]–[24], detection mechanisms and analyses for identifying malicious measurement data and energy theft have been investigated explicitly in [18] and [25]–[40]. The main contributions of this paper are summarized in the following.

- The common DC (direct current) model for state estimation in a power network and traditional techniques for processing bad measurement data are reviewed. Meanwhile, a false data injection (FDI) attack and associated impacts on the power network are illustrated, followed by a discussion of existing countermeasures and studies.
- An attack model related to FDI, *combination sum of energy profiles* (CONSUMER) attack, is defined and formulated into one type of coin change problems that minimizes the number of compromised smart meters without being revealed by maintaining a cumulative load at the aggregation point to which multiple households are connected in today's radial tree-like distribution network.
- A hybrid anomaly intrusion detection system framework, which incorporates *power information* and *sensor placement* (POISE) along with grid-placed sensor (GPS) algorithms using graph theory to provide network observability, is proposed to validate the correctness of customers' energy usage by detecting anomaly activities at the consumption level in the distribution network.
- Simulations for analyzing the proposed attack model as well as grid sensor implementation in terms of network observability and detection rate are conducted and discussed.
- Several potential research directions for furthering the proposed framework in the smart grid context are presented.

The remainder of this paper is structured as follows: Section II reviews the background and state-of-the-art studies related to this work. Section III illustrates the system measurement model prior to the discussion of the proposed detection designs. Section IV presents the problem formulation, attack model, and countermeasures. Section V analyzes the simulation results of the proposed detection framework and discusses the findings. Finally, Section VI summarizes the focal points, draws a conclusion, and presents the future works.

II. BACKGROUND AND RELATED WORKS

An electric power system is a feedback loop control system that relies on measurement data obtained from network measurement units such as meters and sensors. Based on the available data, the control center executes a series of tasks such as topology processing, network observability analysis, state estimation, and bad measurement data processing in order to identify the current status of the power network [41]. Consequently, the decision-making processes of controlling actuators, optimizing power flows, and analyzing possible contingencies are performed to ensure network stability and security, in accordance with what the system observes or estimates. In reality, the measurement data may not be always accurate because of errors in measurements, failures in telemetry and equipment, noises in communications channels, and possibly breached integrity by intentional intrusions or attacks. If the accuracy of measurement data is not as precise as it gets, the decision making can be mistaken in consequence of misguided state estimation.

For simplicity, the common formulation of the state estimation problem is to consider a DC power flow model [41]:

$\mathbf{z} = \mathbf{Hx} + \mathbf{e}$, where $\mathbf{H}$ is the m-by-n Jacobian matrix representing m independent network equations with n state variables related to the network topology, $\mathbf{x}$ is the n-vector ($n \times 1$ matrix) of the true states (unknown and to be estimated for bad data detection), $\mathbf{z}$ is the m-vector of measurements (observed by data collection; in this case, a macro grid power generation reading and $m - 1$ household energy consumption readings), and $\mathbf{e}$ is the m-vector of random errors. The state estimate $\hat{\mathbf{x}}$ can be obtained by calculating $\mathbf{G}^{-1}\mathbf{H}^{\mathsf{T}}\mathbf{Wz}$, where $\mathbf{G} = \mathbf{H}^{\mathsf{T}}\mathbf{WH}$ is the state estimation gain matrix, $(.)^{\mathsf{T}}$ is the transpose of (.), and $\mathbf{W}$ is a diagonal matrix whose entities are most commonly the reciprocals of the variance of measurement errors based on historical statistics, which may represent meter accuracy. In order to detect bad measurement data, the measurement residual $\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}$ is computed and its L_2-norm $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||$ is compared with a predetermined threshold δ; common techniques including normalized residuals and hypothesis testing are sufficient to detect anomalies, e.g., $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}|| > \delta$.

Nevertheless, a recent study [21] observed that the traditional detection is not able to differentiate between natural anomalies and malicious intrusions attributed to *false data injection* (FDI) such that $\mathbf{z_b} = \mathbf{z} + \mathbf{a}$ and $\hat{\mathbf{x}}_\mathbf{b} = \hat{\mathbf{x}} + \mathbf{c}$, where $\mathbf{a} = \mathbf{Hc}$ is an $m \times 1$ attack vector injected to the system that is designed to be a linear combination of the column vectors of $\mathbf{H}$ in order to bypass the detection, i.e., $||\mathbf{z_b} - \mathbf{H}\hat{\mathbf{x}}_\mathbf{b}|| = ||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}|| \leq \delta$. The authors further showed that the attacker is required to compromise a number of meters (i.e., 30%–70% of meters in IEEE 9, 14, 30, 118, 300 bus test systems) in order to bypass detection and takes less than 10 seconds. This type of attack is interchangeably called an *unobservable*, *undetectable*, or *stealth* attack that needs to be launched in a *coordinated* manner [5], [27], [42] with knowledge of the network configuration matrix $\mathbf{H}$ while not violating the physics of power flow. Having knowledge of $\mathbf{H}$ by the attacker has been assumed in most of the current studies. Although a full knowledge of the entire system gained by the attacker may be improbable, it is worth studying and developing a detection framework to identify the malicious attack in case of the attacker possibly having acquired partial knowledge and considerable capability and resource. In fact, the attacker being able to launch FDI without prior knowledge of $\mathbf{H}$ has been studied in [37], that is, if the network topology remains static and the independent loads vary insignificantly for a period of time, $\mathbf{H}$ can be inferred.

Several works have rigorously investigated the FDI attack by proposing various detectors or analyzing the damage effects on the power system. For examples, Kosut *et al.* [26] proposed a detection scheme based on generalized likelihood ratio test while comparing with other two detectors based on the residual error $\mathbf{r}$ derived from the state estimation that uses minimum mean square error technique. The authors studied the outcomes of maximizing the residual error and minimizing the detection rate for the attack. Yuan *et al.* [32] identified the attack launched in two different time periods (i.e., immediate and delayed attack) in which the former may lead the system to perform unnecessary load shedding whereas the latter may cause power overflows on some transmission lines. However, the authors only modeled the immediate attack and showed that the attack leads to a high economic loss. Lin *et al.* [31] studied the effectiveness of the attack in terms of transmission cost and power outage rate by deceiving the amount of energy request and supply as well as the status of transmission lines by claiming a line is valid to deliver a certain amount of power while it is not and vice versa. Giani *et al.* [27] proposed countermeasures by utilizing known-secure PMUs (phasor measurement units) placement and illustrated that $p + 1$ PMUs are enough to detect p k–sparse attacks for $k \leq 5$ while assuming all lines are metered. Qin *et al.* [30] illustrated a case where the attack is detected but still *unidentifiable* in such a way that it is difficult for operators to know which set of meters are truly compromised. The authors proposed a three-step search process that firstly identifies the meter with the largest residual (which exceeds a predetermined threshold) after state estimation, secondly locates a feasible attack region associated with the meter, and finally checks a set of suspicious meters located in the region by using a brute-force search.

Most of the existing works have focused on the cyber-physical attacks at the transmission/distribution level, and very few have analyzed the attack at the distribution/consumption level where smart meters are deployed, e.g., [39], [40]. While motivated by [30], [40], to the best of our knowledge, this is the first work to investigate the CONSUMER attack in the distribution network where a dishonest customer intends to lower his or her electricity bill by compromising some of its neighbors' smart meters in a neighborhood. We will show that the attack can be neither detectable nor identifiable at a certain period of time for the defined scenario. This work can also be extended to an aggressive collusion attack that compromises a group of smart meters and intentionally causes service disruption or equipment damages throughout the distribution network.

III. SYSTEM MEASUREMENT MODEL

AC (alternating current) and DC power flow models are essentially used for studying state estimation. Nevertheless, the DC power flow model is often assessed due to its inexpensive computation and simplicity [43]. Moreover, a DC power grid is a foreseeable approach for the future distribution network [44] because 1) many distributed generators (e.g., household/neighborhood-based solar power systems) supply DC power, 2) AC grid-connected inverters are not needed, and 3) overall costs and power losses can be reduced. The ability to perform state estimation relies on the sufficiency of measurement data available in a network. In other words, the observability of a network has to be analyzed before state estimation can be processed.

Definition 1: A network is said to be **observable** [41] if all flows in the network can be observed by obtaining information in a set of sufficient measurement data such that no power flows in the network for which $\mathbf{Hx} = \mathbf{0}$, $\forall P \in \mathbf{x}$

(where P is an element of the state vector $\mathbf{x}$); otherwise, there is (are) **unobservable** state(s) where non-zero flows exist in the network. In other words, whenever there is a non-zero flow in the network, at least one of the measurements should be nonzero.

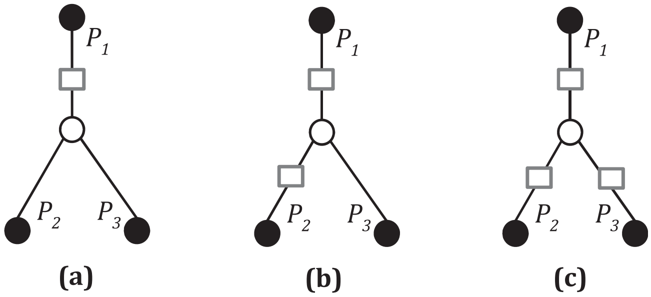

FIGURE 1. **Observability of a network comprised of generation and load nodes (black circle), bus node (white circle), lines (representing power connectivity), and meters/sensors (gray rectangle) in three cases: (a) underdetermined and partially observable, (b) observable and sufficient, and (c) observable but overdetermined.**

Consider a DC network model that has three state variables as shown in Fig. 1: to ensure that the power network is balanced, there is at least one state that acts as a generation or load node, i.e., $P_1 + P_2 + P_3 = 0$. Fig. 1(a) shows an underdetermined and partially observable case where only state P_1 is observable, and one of the states P_2, P_3 is unobservable, and another dependent state is indeterminate. Fig. 1(b) shows an observable and sufficient case where both states P_1 and P_2 are observable, and dependent state P_3 can be computed from the network model equation with the other two known state variables. Fig. 1(c) shows that all states P_1, P_2, P_3 are observable and form an overdetermined system, but can be solved as a least-squares problem. We will use this model to study the proposed CONSUMER attack model as well as grid sensor placement for the distribution network of smart grid in this paper. Additionally, we also consider the characteristics of the emerging smart grid network as follows.

- Nodes (e.g., smart meters, grid sensors) strategically deployed throughout distribution grids are *static*; in other words, grid operators have full knowledge of network topologies in terms of geographical locations and coordinates.
- Nodes are *wire-powered* while attached to power lines and taking various measurements such as voltage, current, frequency, and metering.
- The majority of data traffic generated at the nodes are *periodic* for real-time monitoring and control.
- Each measurement data generated at the nodes (representing individual customer energy consumption and grid line conditions for state estimation) *cannot be fused* at aggregation nodes as opposed to traditional sensor network scenarios where data of sensors tracking their surrounding environmental conditions (e.g., temperature) are aggregated at cluster nodes to generalize the current network status by determining the correlation of the multiple obtained measurements.

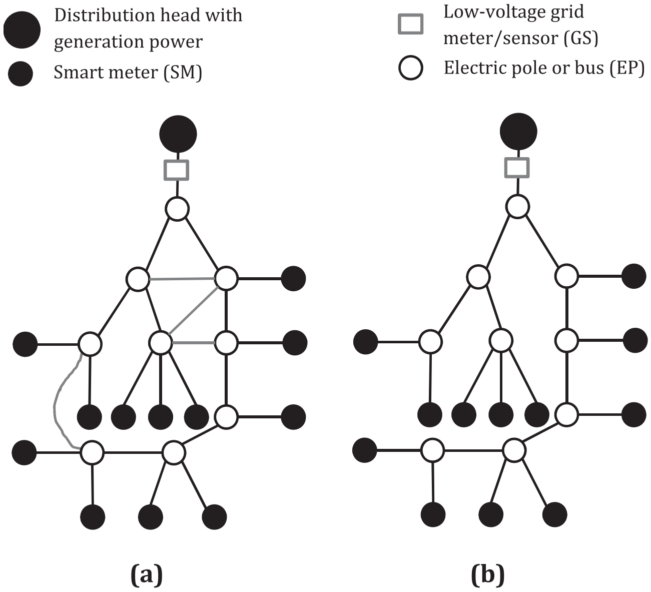
● Distribution head with generation power
● Smart meter (SM)
□ Low-voltage grid meter/sensor (GS)
○ Electric pole or bus (EP)
(a)
(b)

FIGURE 2. **A neighborhood distribution network (a) with loops, and (b) without loops.**

IV. PROBLEM DEFINITION AND FORMULATION

Most parts of the current distribution networks are characterized by radial tree-like topologies, which may or may not contain loops or cycles, as shown in Fig. 2. The *distribution network* consists of four components: 1) a root *aggregation* node at which power P_G is generated or delivered from other sources, such as macro grid or neighboring distribution networks (see [45]), and that supplies power P_G to customers' loads, 2) a *grid sensor* (GS) node that constantly measures the generating power P_G, 3) a set of *electric poles* (EPs) or *buses* as intermediate nodes, $\mathcal{N}_{EP} = (1, 2, \ldots, n_{EP})$ representing the indices of EPs, with distribution lines/feeders, transformers and capacitors (not shown) that construct a distribution grid and deliver power to customers, and 4) a set of household *smart meters* (SMs), $\mathcal{N}_{SM} = (1, 2, \ldots, n_{SM})$ representing the indices of SMs, that have two-way communications capability of reporting household energy consumption to the utility control center and receiving associate feedback messages in real time.

Notably, Fig. 2(a) shows a distribution network that has loops found among some EP nodes, whereas Fig. 2(b) depicts a network with no loops representing a spanning tree. Any spanning tree $G(V_T, E_T)$ from its originally connected graph $G(V, E)$ can be computed by using various algorithms, e.g., Prim's algorithm [46], where V is a collection of vertices, E is a collection of edges, and $V_T = V$. In other words, any connected distribution network $G(V, E)$ can have at least one spanning tree $G(V_T, E_T)$ (composed of $|V_T|$ nodes and $|V_T| - 1$ edges; $|.|$ is the cardinality) with the *fewest edges* among EP nodes[1] while the four *network properties* must be obeyed: 1) the network connectivity in terms of power and communications operations is maintained, 2) the spanning tree starts with the distribution head node, 3) the EP node cannot be a leaf node, and 4) the SM node must be a leaf node. Under these conditions, the spanning tree topology as illustrated in Fig. 2(b) can be discovered, and therefore considered in this study in order for us to determine the minimum number of grid sensors to be placed on edges such that the network is sufficiently observable (to be discussed in Sec. IV-B.3).

We further assume that power flow is unidirectional (in a traditional way) such that power is delivered from the root of the tree to the end leaves. We consider a practical scenario where utility operators currently have limited knowledge about the real-time conditions of distribution networks (e.g., the difficulty of exactly knowing how and how much power is delivered across feeders/lines as well as discovering how and where faults are caused if erroneous activities are present) in a geographically and temporally fine-grained manner due to lack of grid sensors along with effective coordinated monitoring. As shown in Fig. 2, for a power balance circumstance, a summation of individual loads must be equal to the amount of measurement metered at the aggregation GS node. If the aggregated load value exceeds or lessens the GS measurement for a tolerable amount, an anomalous activity is detected and alarmed, but somehow may not be identified easily whether it is caused by natural errors or malicious attacks.

A. THE CONSUMER ATTACK MODEL

In the CONSUMER attack model, we apply the FDI model (introduced in [21]) to construct our attack scenario at the smart meter level. The typical distribution network (shown in Fig. 2) is characterized by its own network topology and configuration matrix $\mathbf{H}$ and a set of observed measurements $\mathbf{z} = [P_G, P_1, P_2, \ldots, P_i]^\mathsf{T} \in \mathbb{Z}$, where $P_G \leq 0$ is the total amount of generated power, $P_i \geq 0, \forall i \in \mathcal{N}_{SM}$ indicates the energy consumption of household i, and $\sum_{\forall i \in \mathcal{N}_{SM}} P_i = P_G$ for a balanced system. We assume that no anomalies should be detected by traditional bad measurement detectors (i.e., $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}|| \leq \delta$) under a normal condition such that smart meters are functioning correctly and legitimately.

The attacker is assumed to have (partial) knowledge of $\mathbf{H}$ and estimation error whether they are obtained illegally or deduced by its own observation. By knowing them, the attacker is able to construct the attack vector $\mathbf{a}$ and associated $\bar{\mathbf{z}}$ such that $||\mathbf{z_b} - \mathbf{H}\hat{\mathbf{x}}_\mathbf{b}|| = ||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}|| \leq \delta$ is satisfied in order to bypass detection, where $\hat{\mathbf{x}}_\mathbf{b} = \hat{\mathbf{x}} + \mathbf{c}$ and $\mathbf{c}$ is a non-zero $n \times 1$ vector designed to derive the vector $\mathbf{a}$. The goal of the attacker is to launch the CONSUMER attack by fabricating $\bar{\mathbf{z}} = \mathbf{z} + \mathbf{a} = [\bar{P}_G, \bar{P}_1, \bar{P}_2, \ldots, \bar{P}_i]^\mathsf{T} \neq 0$ in which $\mathbf{a} = [a_G, a_1, a_2, \ldots, a_i]^\mathsf{T} \in \mathbb{Z}$ and $\sum_{\forall i \in \mathcal{N}_{SM}} a_i = a_G = 0$. There exists load alterations, i.e., $\exists a_i \in \mathbf{a}$, $a_i < 0$ for which the attacker compromises its meter $i \in \mathcal{A}$, and $\exists a_j \in \mathbf{a}$, $a_j > 0, j \neq i$ for which the attacker is able to compromise the victim's meter $j \in \mathcal{B}$. Note that the elements of $\mathcal{A}$ correspond a set of meters belonging to the attacker such that $1 \leq |\mathcal{A}| \leq |\mathcal{N}_{SM}| - 1$ and $\mathcal{A} \subset \mathcal{N}_{SM}$, whereas the elements of $\mathcal{B}$ correspond a set of meters belonging to the victim such that $1 \leq |\mathcal{B}| \leq |\mathcal{N}_{SM}| - 1$, $\mathcal{B} \subset \mathcal{N}_{SM}$, and $\mathcal{B} \cap \mathcal{A} = \emptyset$. The altered linear combination in the vector $\mathbf{a}$ cannot be easily detected by a traditional bad measurement detector. We consider $\mathbf{a} = [a_G, a_1\chi_1, a_2\chi_2, \ldots, a_i\chi_i]^\mathsf{T}$ where the indicator χ_i represents that the smart meter of household i is compromised if $\chi_i = 1$; otherwise, $\chi_i = 0$.

The objective of the attacker is to lower the reading of its own energy consumption level by raising others'. Owing to constrained resources, the attacker tries to minimize the number of compromised smart meters while achieving its objective subject to the inviolability of a total stealing value, $P_s \in \mathbb{N}$. The minimization problem for such attack is formulated as

$$\min \quad \sum_{i=1}^{n_{SM}} \chi_i$$

$$\text{s.t.}$$

$$\sum_{i=1}^{n_{SM}} a_i\chi_i = P_s, \qquad \chi_i \in \{0, 1\}, \qquad \forall i \in \mathcal{B}, \tag{1}$$

$$a_i \geq P_i^{\min}(t+1) - P_i(t), \qquad \forall i \in \mathcal{A}, \tag{2}$$

$$a_i \leq P_i^{\max}(t+1) - P_i(t), \qquad \forall i \in \mathcal{B}, \tag{3}$$

$$P_i(t), P_i^{\min}(t+1), P_i^{\max}(t+1) \geq 0, \qquad \forall i \in \mathcal{A}, \in \mathcal{B}, \tag{4}$$

where $P_s = -\left(\sum_{\forall i \in \mathcal{A}} a_i\right) \geq 0$ is the total amount of non-negative integer power that the attacker plans to steal, $P_{i\in\mathcal{A}}(t)$ is the energy consumption value of the attacker's smart meter i at time t, $P_{i\in\mathcal{B}}(t)$ is the energy consumption value of victim i at time t, $P_{i\in\mathcal{A}}^{\min}(t+1)$ is the minimum power value that the attacker with smart meter i is predicted to consume at time $t+1$, and $P_{i\in\mathcal{B}}^{\max}(t+1)$ is the maximum power value that the victim i is predicted to consume at time $t+1$.

This minimization problem is analogous to the coin change problem, which is NP-hard [49]. Both problems aim to match a given non-negative integer value (equality Constraint 1) while minimizing the number of components (Objective function) for the outcome. As opposed to the coin change problem, the CONSUMER attack problem considers inequality Constraints 2, 3, and 4 that determine $|\mathcal{A}| + |\mathcal{B}|$ sets of non-negative integer power corresponding to households (the attacker(s) and victim(s)) energy profiles at the present time slot (i.e., $P_{i\in\mathcal{A}}(t)$ and $P_{i\in\mathcal{B}}(t)$), as well as the sets of predicted values of energy consumption at the next time slot (i.e., $P_{i\in\mathcal{A}}^{\min}(t+1)$ and $P_{i\in\mathcal{B}}^{\max}(t+1)$). Given a set of a_i belonging to household i that are discovered under Constraints 2, 3, and 4, the problem can be solved as a coin change problem.

[1] How to find such a spanning tree of the distribution network is beyond the scope of this paper. Readers are referred to [47], [48] on how to compute spanning trees for topology control.

The total number of compromised smart meters is defined as $k_{SM} = |\mathcal{A}| + |\mathcal{B}| \leq |\mathcal{N}_{SM}|$.

Theorem 1: A CONSUMER attack can be launched successfully by compromising as few as **two** smart meters (one for the attacker and one for the victim; $k_{SM} \geq 2$) in any spanning tree of a distribution network (described in Sec. IV).

Proof: Consider a radial tree topology illustrated in Fig. 2(b) where there is only one grid sensor available near the supply node measuring the total amount of energy P_G consumed by end customers. We assume that the capacity of each edge in the network may sustain at least P_G during power transmission. A balanced system is maintained if $P_G + P_1 + P_2 + \cdots + P_i = 0, \forall i \in \mathcal{N}_{SM}, \forall P_i, P_G \in \mathbb{Z}, P_i \geq 0, P_G \leq 0$.

There exists combinations of various sets to satisfy the balance equation when $P_i \leq \mathsf{abs}(P_G)$, where $\mathsf{abs}(.)$ is the absolute value of (.). To capitalize on this property, the attacker may design a vector **a** such that $P_G + P_1 + a_1 + P_2 + a_2 + \cdots + P_i + a_i = 0, \forall i \in \mathcal{N}_{SM}, \forall a_i \in \mathbb{Z}$. No smart meter is compromised when $a_i = 0, \forall i$. However, if $\exists a_i : a_i < 0, a_i \in \mathcal{A}$, then $\exists a_j : a_j > 0, j \neq i, a_j \in \mathcal{B}$ without violating the balance equation. For $\mathsf{abs}(a_i) = a_j$ and $|\mathcal{A}| = |\mathcal{B}| = 1$, only two meters are compromised by the attacker; otherwise, more than two meters need to be compromised in order to evade detection. Hence, $k_{SM} = 2$ is the least number for the attacker to launch a successful CONSUMER attack. ■

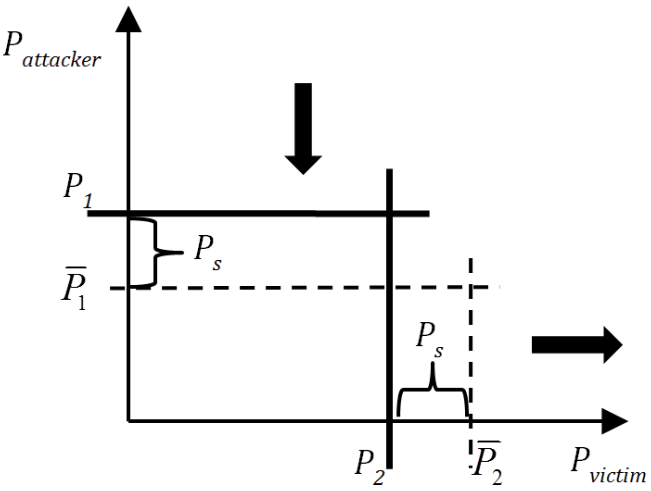

FIGURE 3. Impact of a CONSUMER attack on electricity reading of household 1 (the attacker) and household 2 (the victim).

In an one-attacker-one-victim scenario (depicted in Fig. 3), the attacker tries to decrease its consumption by increasing the victim's as much as it can. Under an *unconstrained* case (which excludes Constraints 2 and 3), the attacker can pick any arbitrary non-negative P_s and performs subtraction on its consumption amount and addition on the victim's to avoid detection as long as Constraint 1 is held; the minimization problem will be reduced to a simple linear programming problem.

On the other hand under a *constrained* case (which includes Constraints 2 and 3), the attacker cannot simply pick any number but needs to determine appropriate $P^{\min}_{i\in\mathcal{A}}$ and $P^{\max}_{i\in\mathcal{B}}$ in the next time slot in order to avoid detection as anomalous activities. In fact, utilities might implement various kinds of prediction methods to predict and monitor households' energy consumption from time to time, and that would complicate the problem. Any anomaly activity that deviates from the correspondingly estimated regression lines beyond a predetermined threshold will trigger an alarm in the intrusion detection system. Unless the attacker had prior knowledge of what the thresholds were, $P^{\min}_{i\in\mathcal{A}}$ and $P^{\max}_{i\in\mathcal{B}}$ could not be chosen too aggressively. Therefore, implementing Constraints 2 and 3 in the attack model would affect the outcome of a CONSUMER attack. In addition to these constraints, the costs of compromising smart meters via coordinated communications on the spatial and temporal scales are challenges from the attacker perspective.

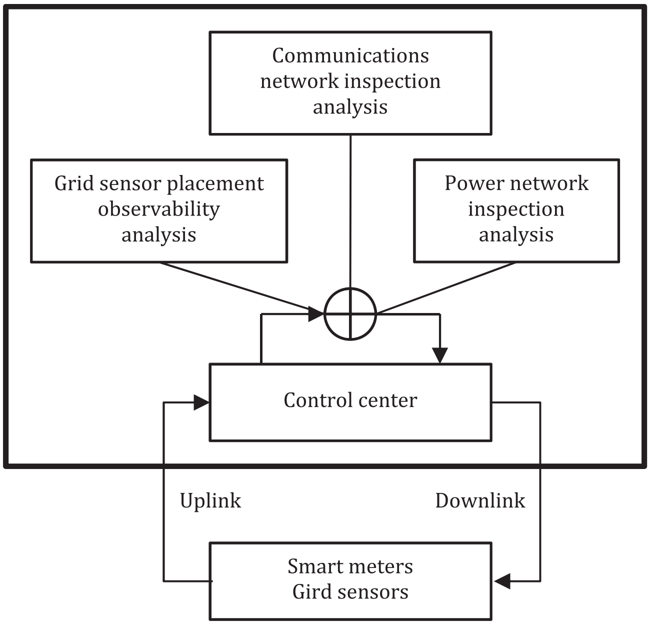
Communications network inspection analysis
Grid sensor placement observability analysis
Power network inspection analysis
Control center
Uplink
Downlink
Smart meters
Gird sensors

FIGURE 4. POISE: a hybrid intrusion detection system.

B. COUNTERMEASURES FOR THE UTILITY DEFENDER

It is unlikely to have an one-size-fits-all solution for detecting anomalous or malicious activities in smart grid. We develop a framework that integrates the characteristics of power network load consumption dynamics, communications network traffic dynamics, and network observability analysis via grid sensor placement for an evolutionary intrusion detection system, as shown in Fig. 4. The last item of the proposed framework is covered in this paper, and the first two items are left for our future works. In a cyber-physical smart grid AMI network, the uplink transmission from smart meters to control centers as well as downlink transmission in an opposite way is vulnerable to a breach of CIA. While a general FDI attack can be launched on the two way links, the CONSUMER attack is specifically instigated in the uplink transmission causing utility operators to make wrong decisions in consequence of receiving falsified measurement data which are hardly distinguishable from the legitimate ones. There are two fundamentally challenging questions in the context of the smart grid intrusion detection system design.

1) What is an adequate threshold for defining an anomaly activity, e.g., in the application of characterizing customers energy consumption behavior while they may be elusive to some extent? Does it even exist?
2) How to effectively distinguish between (unintentionally) anomaly and (intentionally) malicious activities?

While these intriguing questions require further research in the next few years, we provide some insights into the following first two detection methods based on power and communications networks dynamics analyses, and then propose a grid sensor placement mechanism to effectively enhance the intrusion detection process.

1) POWER NETWORK INSPECTION

A power grid system obeys a series of control theories based on laws of physics. Data measurement collection does not only involve power loads but also voltage, current, power factor elements. Observation of phase differences on the transmission/distribution level studied in [27] can be further evaluated on the distribution/consumption level. Another useful metric for designing specification or rule-based anomaly detection systems is to deeply understand different classes of customer energy consumption patterns at different time scales, e.g., usage trends on weekdays, weekends, monthly, seasonal, and annual basis corresponding to individual activities and weather conditions. Many approaches for characterizing household electricity demands including Fourier series, Gaussian processes, neural networks, fuzzy logic, as well as regression and autoregression have been studied [50]. Meanwhile, the existing scheme of detecting illegal customers based on Support Vector Machine (SVM) learning and rule-based algorithms has also been investigated in [18]. These methods could be effectively incorporated in the intrusion detection system at the application level to improve detection accuracy. Computational intelligence [51] can also be readily applied for intrusion detection.

2) COMMUNICATIONS NETWORK INSPECTION

Along with the methods of power dynamics inspection, extensive studies on traditional low-power WSN attack scenarios [10] at the physical, MAC, and network layer levels are complementary intrusion detection tools to be integrated into the smart grid communications security environment, specifically the jamming, replay, and DoS attacks. Several dominant metrics such as data sending rate, receiving rate, packet loss rate, and signal strength will be tailored to effectively facilitate the detection of anomaly activities in smart grid communications in response to compromised circumstance.

3) INTRUSION DETECTION SYSTEM WITH POWER INFORMATION AND SENSOR PLACEMENT – IDS WITH POISE

Smart meter deployment has been initiated worldwide in the past few years. The rationale for replacing the traditional meters with smart meters is plentiful, but the fundamental one is to be able to monitor and control customer energy consumption more efficiently in real time through two-way communications by leveraging the state-of-the-art wire/wireless and power line communications technologies. By gaining knowledge of individual energy usage patterns, utilities can deal with primary issues easily such as peak demands alleviation, remote meter reading, and distributed renewable energy sources accommodation, in order to increase energy efficiency and reduce greenhouse gas emission. The entire smart grid AMI network consisting of a number of control centers and hundreds of thousands of smart meters is likely to operate using the IP Protocol with IPv6 addresses assignment connected to the Internet [1]. Smart meters support multiple communications protocols that facilitate smart energy management in HANs and mesh routing in NAN. Many have considered utilizing the existing networks such as WiFi and wireless mesh networks to communicate under unlicensed bands for economic reasons. This strategy creates network uncertainties by exposing security vulnerabilities of smart metering communications to the public.

In the meantime, we propose grid sensor placement across the distribution network in which these grid sensors with simpler design (than smart meters) are owned by utilities and construct grid sensor networks operating in dedicated or licensed bands specified in IEEE 802.15.4g Smart Utility Network (SUN), e.g., see [1], [2] for further studies. The grid sensor network is much less vulnerable to malicious attacks and is designed as surveillance guards in the distribution grid. Moreover, deploying grid sensors on lines/feeders (as low-voltage sensors) brings utilities a number of potential benefits: 1) greater transparency and stability can be achieved owing to the substantial observability of power flow conditions on each segment and portion of the network, 2) voltage fluctuation due to varying input of renewable energy sources (e.g., household/neighborhood-based PV solar systems) can be effectively monitored, and 3) optimization in volt-var control and optimal power flow operations can be intelligently performed. Hence, utility operators will have a full knowledge of their supervised network topologies in terms of geographical locations with coordinates of grid sensors as well as smart meters while monitoring the network quality and ensuring cyber-physical security. At this stage, we assume that all deployed grid sensors are intrusion resistant and their measurement data are trustworthy (i.e., false alarm rate is zero) so that the measurement data of smart meters can be compared with that of grid sensors to detect and identify any falsified data by compromised smart meters.

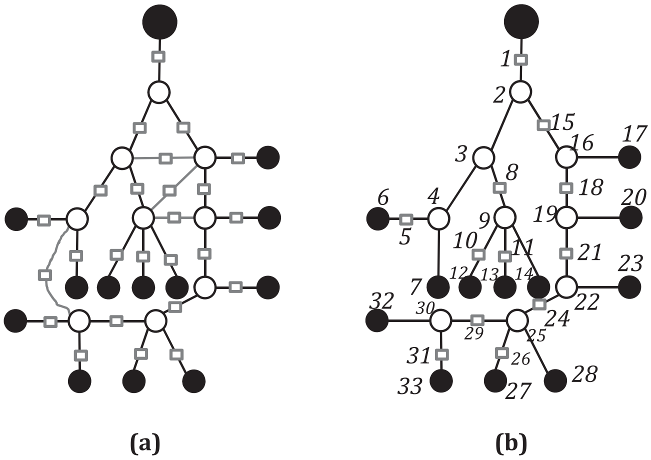

FIGURE 5. **A neighborhood distribution network deployed with a number of grid sensors in (a) overdetermined case, and (b) sufficient case.**

As discussed in Sec. IV, the existing distribution grid is not transparent to the utilities to a certain degree. The design of sensor grid placement can help provide topological observability by deploying a sufficient number of grid sensors to guarantee state estimation solvability. In Fig. 5(a), every grid line is placed with a sensor that results in an overdetermined system. In order to reduce the redundancy to a sufficient number while observability is still satisfied, a Grid-Placed Sensor (GPS) algorithm is proposed, as shown in Alg. 1. For the spanning tree illustrated in Fig. 5(b), the network graph $G(V_T, E_T)$ with depth $1, 2, \ldots, d \in D_T$ is constructed by a set of EP and SM nodes $v_1, v_2, \ldots, v_n \in V_T$ and a set of edges E_T, where $\mathcal{N}_{SM} \subset V_T, \mathcal{N}_{EP} \subset V_T, |V_T| = |\mathcal{N}_{SM}| + |\mathcal{N}_{EP}|$.

Algorithm 1 Grid-Placed Sensor (GPS) - Loop Free

```
Input: Given a connected, undirected spanning tree graph $G(V_T, E_T)$ and depth $D_T$.
Output: A $n \times n$ observability indicator matrix $I_O$ that represents observability status of each edge.
Place a GS node at the root node's edge.
for $i = 1$ to $d$ do
    Determine the number of child $u$ of $v(d), \forall v \in V_T$
    if $u = 1$ then
        No GS node is placed.
    else if $u > 1$ then
        A GS node is placed on any $(u - 1)$ of the $u$ edges connected to the child, and mark $1$ for the GS-placed edges in $I_O$.
    end if
    Repeat for other $v$ if having the same $d$.
end for
```

At the beginning, the GS node v_1 is directly placed on the edge between the generation source and distribution bus, i.e., v_2. The algorithm then starts with EP node v_2 and discovers that it has two children, which can be EP or SM nodes. Either $e(v_2, v_3)$ or $e(v_2, v_{16})$ placed with a GS node v_{15} in between will make both edges become observable, according to Def. 1 in Sec. III. Note that $e(w, v)$ or $e(v, w)$ denotes the edge e that connects both node w and v. Both edges becoming observable are then marked with *1* in I_O. Repeat the process for the right branch, the algorithm starts with EP node v_{16} and discovers that it also has two children, and therefore, either $e(v_{16}, v_{19})$ or $e(v_{16}, v_{17})$ placed with a GS node v_{18} will make both edges become observable; again, the two observable edges are marked with *1* in I_O. Notably, although SM node v_{17} has metering capability to make $e(v_{16}, v_{17})$ observable already, the GS node v_{18} is placed in order to later verify whether or not the measurement data of SM node v_{17} is legitimate. The process is repeated until it reaches the leaves with the largest d.

Theorem 2: A spanning tree of a distribution network is said to be (sufficiently) observable if $Y(G) - I_O = 0$, where $Y(G)$ is the $n \times n$ adjacency matrix of G and the entry $y_{i,j}$ in $Y(G)$ is the number of edges from node i to node j.

Proof: While $Y(G)$ represents the adjacency of edges where $y_{i,j} = 1$ if there exists an edge from node i to node j (otherwise, $y_{i,j} = 0$), the $n \times n$ matrix I_O specifies the observability of edges where the entry $\alpha_{i,j} = 1$ if $y_{i,j} = 1$ is observable and $\alpha_{i,j} = 0$ otherwise according to the GPS algorithm. Since both $Y(G)$ and I_O are determined by the status of edges between node i and node j for n nodes (i.e., edge existence/observability), both matrices are identical when the two conditions are true, i.e., $y_{i,j} - \alpha_{i,j} = 0, \forall i, j \in V_T$. ■

Theorem 3: In a spanning tree of a distribution network, the number of GS nodes placed on edges for the network to be observable is the same as the number of SM nodes.

Proof: By performing a breadth-first search, one can determine the number of children m_i for every $v_i \in V_T$; $\mathcal{M} = (m_1, m_2, \ldots, m_n)$ is thus a set of non-negative integers and $\sum_{\forall m_i \in \mathcal{M}} m_i = |V_T| - 1$ is the total number of edges. Since the spanning tree imposes that the SM node must be a leaf and the EP node cannot be a leaf, there are γ ($< |V_T|$) SM nodes (i.e., $m_i = 0$ for a leaf node) while there are $|V_T| - \gamma$ EP nodes (i.e., $m_i > 0$).

According to the GPS algorithm, $(m_i - 1)$ of m_i children of the EP node i (at each depth and branch) are placed with a GS node on the associated edges in order for the network to be observable; this implies that one edge connecting the EP node i and one of its associated children does not require a GS node. Since there are totally $|V_T| - \gamma$ EP nodes with $|V_T| - \gamma$ edges that do not need the GS nodes, the total number of GS nodes is derived as $(|V_T| - 1) - (|V_T| - \gamma) + 1 = \gamma$ (which equals the total number of SM nodes), where the third term, 1, accounts for the GS node near the root. ■

Theorem IV-B.3 allows utility network operators to rapidly discover the observability status of a distribution network upon the collections of the adjacency matrix $Y(G)$ and observability matrix I_O. Theorem IV-B.3 helps operators realize the total number of grid sensors required for a distribution network to be observable given the total number of smart meters.

V. NUMERICAL RESULTS AND ANALYSES

We conducted three types of simulations in this paper in order to analyze the outcomes of the proposed CONSUMER attack model as well as grid sensor placement for detecting the attack.

A. DETERMINATION OF SUCCESSFUL CONSUMER ATTACKS IN DIFFERENT CONSTRAINT SCENARIOS

In the first simulation, we set 5 kWh for the actual amount of power that the attacker consumes at a certain time period, and four different values (4 kWh, 3 kWh, 2 kWh, and 1 kWh) to which the attacker aims to reduce and for which it actually pays; this means that the difference in energy consumption between the real and fabricated values has to be compensated by a number of chosen neighboring victims in order to evade detection. From Fig. 6, we consider three conditions in terms of constraint level while the attacker performs such action. In an unconstrained scenario, there are no boundaries for the attacker to steal. As a result, it only needs to compromise as low as one smart meter from the neighbors (in addition to its own meter) for stealing the four different amounts of power. In more practical cases where there are boundaries (predetermined at utility control centers), we set an expected amount of 2 kWh that can be tolerated in fluctuation of customers energy consumption for a loosely constrained case, and 1 kWh for a strictly constrained case. From the results, we discover that more smart meters need to be compromised to achieve the targets while bypassing detection.

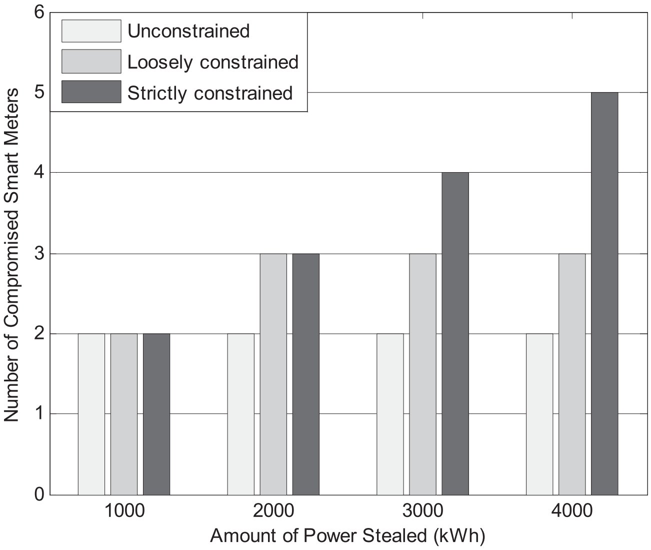

FIGURE 6. Requirements for a successful CONSUMER attack under different constraints.

Remarkably, compromising a large number of smart meters is believed to be an improbable scenario because the attacker needs to know the upper and lower bounds for the victims' and its energy consumption patterns upon which the utility control center constantly monitors. However, a probable case should be emphasized, that is, the attacker may change its strategy to launch a p–cluster $\mathbf{k}$–sparse attack throughout the network and still, without being detected, where p is the number of clustered attacks and $\mathbf{k}$ is a set of distinct element values that represent the number of compromised smart meters consisted in one of p clusters. Such attack is studied in the next subsection, and note that the term *sparse* introduced here somewhat differs from that in [27].

B. STUDY OF p–CLUSTER $\mathbf{k}$–SPARSE CONSUMER ATTACK

In the second simulation, we study how the multiple CONSUMER attacks, referred to as a multi-CONSUMER attack, may be potentially launched by designating p number of clustered attacks in which each cluster may be composed of a number of compromised smart meters based on any unique value in the set $\mathbf{k} = \{2, 3, \ldots, k_{SM}\}$. Conditionally, each cluster must contain at least two meters: one for the attacker and one for the victim (see Theorem IV-A). If all p clusters are composed of two meters (i.e., $\mathbf{k} = \{2\}$), a combination of $\mathcal{C} = (2, 2, \ldots, 2)$ is formed where $|\mathcal{C}| = p \leq \lfloor \frac{k_{SM}}{2} \rfloor$. Note that the *order* in $\mathcal{C}$ does not matter and the summation of all element values in $\mathcal{C}$ equals k_{SM}. Given k_{SM} and $\mathbf{k}$, one can determine a number of possible combinations $\mathcal{C}$'s involved in a multi-CONSUMER attack.

Suppose there were six smart meters in a distribution network and the attacker was able to compromise all six meters. Originally, *one* attack cluster with all *six* meters could be formed as a single CONSUMER attack, namely, a 1–cluster {6}–sparse attack with the combination of $\mathcal{C} = (6)$. Alternatively, there might be two other possibilities of launching a multi-CONSUMER attack: 1) *two* attack clusters and *three* meters in each cluster—a 2–cluster {3}–sparse attack with the combination of $\mathcal{C} = (3, 3)$, and 2) *three* attack clusters and *two* meters in each cluster—a 3–cluster {2}–sparse attack with the combination of $\mathcal{C} = (2, 2, 2)$. Similarly, if the attacker was to compromise eight meters with $\mathbf{k} = \{2, 3, 4\}$, a total of four cluster combinations can be generated: 1) $p = 4$, $\mathcal{C} = (2, 2, 2, 2)$, 2) $p = 3$, $\mathcal{C} = (2, 3, 3)$, 3) $p = 3$, $\mathcal{C} = (2, 2, 4)$, and 4) $p = 2$, $\mathcal{C} = (4, 4)$. Again, the order in $\mathcal{C}$ does not mater.

We set $k_{SM} = 50$, 100, and 300. In Table 1, 14 sets of element values in $\mathbf{k}$ for producing different sizes of clusters are discovered under a multi-CONSUMER attack scenario, while $\mathbf{k}$ contains a set of values ranging from 2 to 5. Take $k_{SM} = 50$ as an example, in order for the attacker to maintain 2 compromised smart meters in each attack cluster, 25 clusters can be launched at once. On the contrary, when the attacker launches a multi-CONSUMER attack that is collectively formed by 2, 3, 4, and/or 5 compromised meters, the number of combinations of cluster formation can be as high as 258; this can also be solved as a coin change problem. As more different sizes of clusters are involved in the

TABLE 1. Number of Ways of Forming A Multi-CONSUMER Attack.

	$k_{SM} = 50$		$k_{SM} = 100$		$k_{SM} = 300$	
k	p	# of $\mathcal{C}$'s	p	# of $\mathcal{C}$'s	p	# of $\mathcal{C}$'s
2	25	1	50	1	150	1
2,3	17–25	9	34–50	17	100–150	51
2,3,4	13–25	65	25–50	234	75–150	1951
2,3,4,5	10–25	258	20–50	1703	60–150	40191
2,4	13–25	13	25–50	26	75–150	76
2,4,5	10–25	42	20–50	146	60–150	1186
2,5	10–25	6	20–50	11	60–150	31
3	–	0	–	0	100	1
3,4	13–16	4	25–33	9	75–100	26
3,4,5	10–16	26	20–33	94	60–100	781
3,5	10–16	4	20–32	7	60–100	21
4	–	0	25	1	75	1
4,5	10–12	3	20–25	6	60–75	16
5	10	1	20	1	60	1

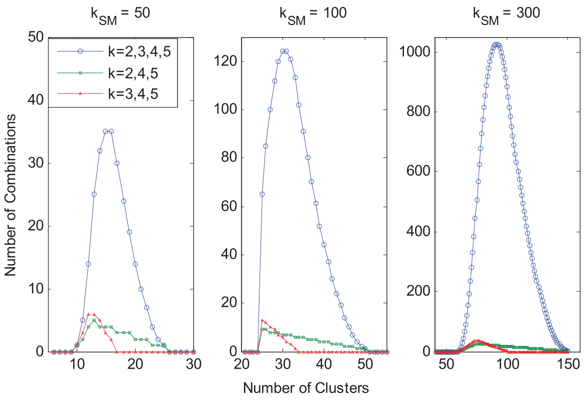

FIGURE 7. Distribution of cluster quantities for launching a multi-cluster attack for different total number of compromised smart meters as (a) 50, (b) 100, and (c) 300.

attack, more combinations are generated. This outcome may complicate the detection performance for utility operators because p–cluster attack may instigate p CONSUMER attack problems at the same time.

More explicitly, Fig. 7(a) shows that there can be 10–25 clusters formed for $\mathbf{k} = \{2, 3, 4, 5\}$ and $k_{SM} = 50$ in which the average number of clusters is 16.29 with variance 8.48; also 10–25 clusters formed for $\mathbf{k} = \{2, 4, 5\}$ but 16.31 averaged number of clusters with 14.66 variance; and 10–16 clusters formed for $\mathbf{k} = \{3, 4, 5\}$ and 13.08 averaged number of clusters with 2.47 variance. It is worth noting that the value of elements as well as quantity of elements in $\mathbf{k}$ will determine the number of combinations, which may increase dramatically when the total number of compromised meters increases, as depicted in Fig. 7(b) and 7(c).

C. ANALYSIS OF NETWORK OBSERVABILITY AND CORRESPONDING DETECTION RATES

In the last simulation, we investigate how the detection rate varies with different levels of network observability in terms of the number of grid sensors placed in the network. From an attacker point of view, it can have $\binom{|\mathcal{N}_{SM}|}{k_{SM}}$ of ways to compromise k_{SM} out of $|\mathcal{N}_{SM}|$ smart meters. Similarly, from a utility defender point of view, the operator has to determine $\binom{n_{GS}}{k_{GS}}$ of possible ways that k_{GS} out of n_{GS} grid sensors may become unavailable and cause partial unobservability of the network when n_{GS} is a sufficient number for the network to be observable. In the worst case, the detection rate can be as low as zero when compromised smart meters are next to each other (whether they are connected to the same parent node or connected to their parents whose edge is shared by each other) and where exactly the grid sensor becomes unavailable. Two examples may be depicted from Fig. 5(b): 1) *the worst undetectable and unidentifiable cases*: consider the case that SM nodes v_{27} and v_{28} are compromised and at the same time GS node v_{26} is unavailable, thus causing unobservability on $e(v_{25}, v_{27})$ and $e(v_{25}, v_{28})$ – the CONSUMER attack on these two smart meters is undetected; also consider the case that SM nodes v_{17} and v_{20} are compromised, in which case the unavailability of GS node v_{18} can cause $e(v_{16}, v_{17})$ and $e(v_{19}, v_{20})$ to be unobservable, and hence undetectable on SM nodes v_{17} and v_{20}; and 2) *the unidentifiable but detectable case*: consider the case that SM nodes v_{17} and v_{23} are compromised and GS node v_{18} becomes unavailable, in which case SM node v_{23} is detected as an attacked node by observing GS nodes v_{21} and v_{24} but SM nodes v_{17} and v_{20} cannot be identified whether one or all of the smart meters are attacked. Hence, SM nodes v_{17} and v_{20} must be further inspected by the utility and considered as a detected case.

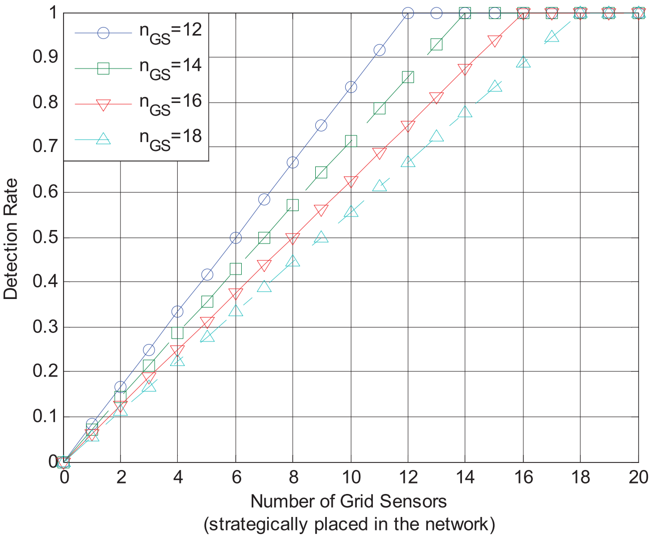

FIGURE 8. Detection rate versus the quantity of GS nodes installation.

Fig. 8 shows how the average rate of detecting the CONSUMER attack(s) can be improved by increasing the number of grid sensors. Since the number of smart meters and grid sensors are identical (proven in Theorem IV-B.3), at the same time the number of times the smart meters to be attacked and the number of times the grid sensors to become unavailable are equally likely, the outcomes of the detection rate and grid sensor availability shown in Fig. 8 exhibit a linear relationship. From the results, we notice that the slope of the detection rate is steeper when the number of grid sensors (as well as smart meters) is smaller. On the other hand, the slope of the detection rate declines when the number of grid sensors increases. This means that a smaller network with a smaller number of sufficient n_{GS} deployed is more vulnerable to unobservability as compared to a larger network with the same number of GS nodes becoming unavailable.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we have investigated a breach of data integrity attributed to false data injection attacks for the future power grid environment. We have formulated an attack model (CONSUMER) to illustrate that by compromising smart meters, an illegal customer "can steal" electricity by lowering the reading of its energy consumption and raising others' in a neighborhood distribution network. A novel hybrid intrusion detection system framework that incorporates power information and sensor placement has been developed to detect malicious activities such as CONSUMER attacks while the traditional bad measurement detectors cannot. An algorithm for placing grid sensors on lines or feeders strategically throughout a spanning-tree distribution network is proposed to provide sufficient network observability to enhance detection performance. We have shown that compromising a large number of smart meters may be improbable and indicated that the attack may be turned into multiple clustered attacks with a few compromised smart meters. We have also shown that while the detection rate can be improved by the proposed grid sensor placement with sufficient observability, it can be degraded by the unavailability of grid sensors as well.

Intrusion detection for the smart grid system (deployed with a large number of smart meters and grid sensors) will attract further investigation for the coming years. Meanwhile, we provide a few insights into some potential research topics associated with the proposed intrusion detection framework:

- Grid sensors in this paper are considered fully trustable. For practical scenarios, trustworthiness of meters and sensors can be explored to determine possible impacts on the proposed intrusion detection framework by addressing uncertainties of network dynamics in the context of smart grid security, e.g., the attacker can launch an observability attack by compromising or disabling some of the grid sensors, thus making intrusion detection more challenging.
- Grid sensor localization and associated observability studies can be further extended to grid isolation designs. For example, grid isolation may be employed to prevent catastrophic failures from cyber-physical attacks, but the grid in islanded mode must remain observable as well.
- The proposed CONSUMER attack design, which is currently limited to a one-player attack, can be extended to multi-player CONSUMER warfare where more than one attacker tries to steal electricity at the same time period. The attack can be redesigned as a (non)cooperative game based on the CONSUMER attack model to broadly explore its variants against the integrity of the power distribution grid system.
- The complementary detection methods of utilizing power and communications networks inspection incorporated in the proposed framework can be developed elaborately to improve detection performance.
- Further development of effective and efficient countermeasures are desired to cope with variants of the CONSUMER attack.

REFERENCES

[1] C.-H. Lo and N. Ansari, "The progressive smart grid system from both power and communications aspects," *IEEE Commun. Surv. Tuts.*, vol. 14, no. 3, pp. 799–821, Sep./Dec. 2012.

[2] C.-H. Lo and N. Ansari, "IEEE 802.15.4-based wireless sensor network design for smart grid communications," in *Handbook on Green Information and Communications Systems*, M. S. Obaidat, A. Anpalagan, and I. Woungang, Eds. New York, NY, USA: Academic, 2013, ch. 4.

[3] L. Zhou and J. Rodrigues, "Service-oriented middleware for smart grid: Principle, infrastructure, and application," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 84–89, Jan. 2013.

[4] L. Zhou, J. J. P. C. Rodrigues, and L. Oliveira, "QoE-driven power scheduling in smart grid: Architecture, strategy, and methodology," *IEEE Commun. Mag.*, vol. 50, no. 5, pp. 136–141, May 2012.

[5] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[6] Y. Mo, T.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber-physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, Jan. 2012.

[7] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on cyber security for smart grid communications," *IEEE Commun. Surv. Tuts.*, vol. 14, no. 4, pp. 998–1010, Apr. 2012.

[8] Y. Yang, T. Littler, S. Sezer, K. McLaughlin, and H. Wang, "Impact of cyber-security issues on smart grid," in *Proc. 2nd IEEE PES ISGT Eur.*, Dec. 2011, pp. 1–7.

[9] J. Liu, Y. Xiao, S. Li, W. Liang, and C. Chen, "Cyber security and privacy issues in smart grids," *IEEE Commun. Surv. Tuts.*, vol. 14, no. 4, pp. 981–997, Apr. 2012.

[10] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: A survey," *IEEE Commun. Surv. Tuts.*, vol. 11, no. 2, pp. 52–73, May/Aug. 2009.

[11] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 24–29, Aug. 2012.

[12] R. Berthier, W. Sanders, and H. Khurana, "Intrusion detection for advanced metering infrastructures: Requirements and architectural directions," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 350–355.

[13] P. Sakarindr and N. Ansari, "Security services in group communications over wireless infrastructure, mobile ad hoc, and wireless sensor networks," *IEEE Wireless Commun.*, vol. 14, no. 5, pp. 8–20, Oct. 2007.

[14] P. Sakarindr and N. Ansari, "Survey of security services on group communications," *IET Inf. Security*, vol. 4, no. 4, pp. 258–272, Dec. 2010.

[15] C. Neuman and K. Tan, "Mediating cyber and physical threat propagation in secure smart grid architectures," in *Proc. 2nd IEEE Int. Conf. Smart Grid Commun.*, Oct. 2011, pp. 238–243.

[16] Inst. Electr. Efficiency. (2012, May). *Utility-Scale Smart Meter Deployments, Plans, and Proposals*, Washington, DC, USA [Online]. Available: http://www.edisonfoundation.net/iee/Documents/IEE_SmartMeter Rollouts_0512.pdf

[17] L. Enbysk. (2013, Jan.). *Energy Theft: From Bad to Worse (and what Some Utilities are Doing About it)* [Online]. Available: http://www.smartgridnews.com/

[18] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, "High performance computing for detection of electricity theft," *Int. J. Electr. Power Energy Syst.*, vol. 47, pp. 21–30, May 2013.

[19] nCircle, London, U.K. (2012). *Information Risk & Security Performance Management*, [Online]. Available: http://www.ncircle.com/index.php?s=resources_surveys_Survey-SmartGrid-2012

[20] M. Costache, V. Tudor, M. Almgren, M. Papatriantafilou, and C. Saunders, "Remote control of smart meters: Friend or foe?" in *Proc. 7th Eur. Conf. Comput. Netw. Defense*, Sep. 2011, pp. 49–56.

[21] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. CCS*, 2009, pp. 21–32.

[22] Z. Fadlullah, M. Fouda, N. Kato, X. Shen, and Y. Nozaki, "An early warning system against malicious activities for smart grid communications," *IEEE Netw.*, vol. 25, no. 5, pp. 50–55, Sep./Oct. 2011.

[23] X. Wang and P. Yi, "Security framework for wireless communications in smart distribution grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 809–818, Dec. 2011.

[24] Y. Zhang, L. Wang, W. Sun, R. Green, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 796–808, Dec. 2011.

[25] J. Valenzuela, J. Wang, and N. Bissinger, "Real-time intrusion detection in power system operations," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1052–1062, May 2013.

[26] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[27] A. Giani, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla, "Smart grid data integrity attacks: Characterizations and countermeasures," in *Proc. 2nd IEEE Int. Conf. Smart Grid Commun.*, Oct. 2011, pp. 232–237.

[28] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet-based load altering attacks against smart power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 667–674, Dec. 2011.

[29] C. Ma, D. Yau, X. Lou, and N. Rao, "Markov game analysis for attack-defense of power networks under possible misinformation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1676–1686, May 2013.

[30] Z. Qin, Q. Li, and M. Chuah, "Defending against unidentifiable attacks in electric power grids," *IEEE Trans. Parallel Distrib. Syst.*, to be published.

[31] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On false data injection attacks against distributed energy routing in smart grid," in *Proc. 3rd IEEE/ACM ICCPS*, Apr. 2012, pp. 183–192.

[32] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 382–390, Jun. 2011.

[33] S. Bi and Y. J. Zhang, "Defending mechanisms against false-data injection attacks in the power system state estimation," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2011, pp. 1162–1167.

[34] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 6, pp. 1108–1118, Jul. 2012.

[35] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 226–231.

[36] D. Wei, Y. Lu, M. Jafari, P. Skare, and K. Rohde, "Protecting smart grid automation systems against cyberattacks," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 782–795, Dec. 2011.

[37] Y. Huang, M. Esmalifalak, H. Nguyen, R. Zheng, Z. Han, H. Li, and L. Song, "Bad data injection in smart grid: Attack and defense mechanisms," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 27–33, Jan. 2013.

[38] Y. Zhang, W. Chen, and J. Black, "Anomaly detection in premise energy consumption data," in *Proc. IEEE Power Energy Soc. General Meeting*, Jul. 2011, pp. 1–8.

[39] S. Salinas, M. Li, and P. Li, "Privacy-preserving energy theft detection in smart grids," in *Proc. 9th Annu. IEEE Commun. Soc. SECON*, Jun. 2012, pp. 605–613.

[40] Z. Xiao, Y. Xiao, and D.-C. Du, "Exploring malicious meter inspection in neighborhood area smart grids," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 214–226, Mar. 2013.

[41] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*. Norwell, MA, USA: Kluwer, 1999.

[42] S. Cui, Z. Han, S. Kar, T. Kim, H. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 106–115, Sep. 2012.

[43] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *Proc. 8th IEEE Int. Conf. ACDC*, Mar. 2006, pp. 58–62.

[44] K. Kurohane, T. Senjyu, A. Yona, N. Urasaki, T. Goya, and T. Funabashi, "A hybrid smart AC/DC power system," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 199–204, Sep. 2010.

[45] C.-H. Lo and N. Ansari, "Decentralized controls and communications for autonomous distribution networks in smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 66–77, Mar. 2013.

[46] R. Sedgewick and K. Wayne, *Algorithms*. Upper Saddle River, NJ, USA: Pearson Edu., 2011.

[47] N. Ansari, G. Cheng, and R. Krishnan, "Efficient and reliable link state information dissemination," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 317–319, May 2004.

[48] K. Miyao, H. Nakayama, N. Ansari, and N. Kato, "LTRT: An efficient and reliable topology control algorithm for ad-hoc networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6050–6058, Dec. 2009.

[49] D. Pearson, "A polynomial-time algorithm for the change-making problem," *Operat. Res. Lett.*, vol. 33, no. 3, pp. 231–234, May 2005.

[50] F. McLoughlin, A. Duffy, and M. Conlon, "Evaluation of time series techniques to characterise domestic electricity demand," *Energy*, vol. 50, pp. 120–130, Feb. 2013.

[51] N. Ansari and E. Hou, *Computational Intelligence for Optimization*. Norwell, MA, USA: Kluwer, 1997.

CHUN-HAO LO (S'08) received the B.S. degree in electrical and computer engineering from Ohio State University, Columbus, OH, USA, in 2003, the M.S. degree in engineering from the University of Detroit Mercy, Detroit, MI, USA, in 2004, the M.S. degree in telecommunications from the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 2006, and the Ph.D. degree in electrical engineering from NJIT in 2013. His research interests include smart grid networks, sensor networks, wireless communications, network optimization, and intrusion detection.

NIRWAN ANSARI (S'78–M'83–SM'94–F'09) received the B.S.E.E. (*summa cum laude*) degree from the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 1982, the M.S.E.E. degree from the University of Michigan, Ann Arbor, MI, USA, in 1983, and the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1988. He joined the NJIT's Department of Electrical and Computer Engineering as an Assistant Professor in 1988 and has been a Full Professor since 1997.

He has contributed over 400 technical papers, over one third of which were published in widely cited refereed journals and magazines. His research focuses on various aspects of broadband networks and multimedia communications. He has also been granted over twenty U.S. patents.