SAI: A Suspicion Assessment-Based Inspection Algorithm to Detect Malicious Users in Smart Grid

Xiaofang Xia, Yang Xiao, *Senior Member, IEEE*, and Wei Liang, *Senior Member, IEEE*

***Abstract*— Integrated with cutting-edge equipment and technologies, smart grid takes prominent advantages over traditional power systems. However, hardware and software techniques also bring smart grid numerous security concerns, especially various cyberattacks. Malicious users can launch cyberattacks to tamper with smart meters anytime and anywhere, mainly for the purpose of stealing electricity. This makes electricity theft much easier to commit and more difficult to detect. Researchers have devised many approaches to identify malicious users. However, these approaches suffer from either poor accuracy or expensive cost of deploying monitoring devices. This paper aims to locate malicious users using a limited number of monitoring devices (called inspectors) within the shortest detection time. Before inspectors conduct any inspection, suspicions that users steal electricity are comprehensively assessed, mainly through analyzing prior records of electricity theft as well as deviations between the reported and predicted normal consumptions. On the basis of these suspicions, we further propose a suspicion assessment-based inspection (SAI) algorithm, in which the users with the highest suspicions will be first probed individually. Then, the other users will be probed by a binary tree-based inspection strategy. The binary tree is built according to users' suspicions. The inspection order of the nodes on the binary tree is also determined by the suspicions. The experiment results show that the SAI algorithm outperforms the existing methods.**

***Index Terms*— Smart grid, electricity theft, suspicion assessment, malicious meter inspection, security.**

I. Introduction

INTEGRATED with cutting-edge equipment and technologies (e.g., advanced metering infrastructure and modern information and communication technologies), smart grid significantly improves power systems' performance in reliability, economics, and efficiency [1], [2]. However, advanced hardware and software techniques also bring smart grid many vulnerabilities, among which security concerns garner the most attentions [3]. As a significant concern of utility companies, electricity theft[1] has been notorious since the establishment of traditional power systems.

Many factors influence users into stealing electricity. These factors include higher energy prices, unemployment, weak economic situations, weak accountability and enforcement of laws, and even corruption of employees in utility companies [4]. Some people are reported to commit electricity theft for the purpose of masking their illegal operations (for example, growing marijuana). On the whole, electricity theft is much more severe in developing countries, such as India and South Africa, than in developed countries, such as USA and UK. It was reported that the revenue losses of worldwide electricity theft was about $89.3 billion as early as 2014 [5].

In smart grid, malicious users[2] can steal electricity by either physical attacks or cyber attacks. The most popular physical attacks include bypassing a feeder, inverting a meter, and employing a strong permanent magnet. These can also be used in traditional power systems. The most common cyber attacks include modifying firmware/storage and stealing credentials of smart meters [6]. It is reported that malicious users can compromise smart meters with only a moderate level of computer knowledge. With the assistance of some low-cost tools and software readily available on the Internet [7], malicious users can easily launch cyber attacks. In general, cyber attacks are much more popular than physical attacks in smart grid.

Conventionally, utility companies employ personnel to physically check the tamper-evident seals door-to-door for targeting malicious users. However, this approach is labor-consuming and cannot deal with electricity theft caused by cyber attacks. Data mining techniques such as extreme learning machine and support vector machine are leveraged to analyze the fine-grained electricity consumption data, aiming at an exposure of abnormality [8]–[12]. These techniques have a poor accuracy, which is about 60% $\sim$ 70% [13]. It will raise controversy between users and utility companies. Different

Manuscript received October 6, 2018; revised February 17, 2019 and April 7, 2019; accepted May 29, 2019. Date of publication June 5, 2019; date of current version September 16, 2019. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFE0101300, in part by the U.S. National Science Foundation under Grant CNS-1059265, in part by the National Natural Science Foundation of China under Grant 61374200, Grant 71661147005 and Grant 61702403, in part by the Key Research and Development Plan of Jiangxi Province under Grant 20181ACE50029, and in part by the National Natural Science Foundation of Shaanxi Province under Grant 2019ZDLGY13-09 and Grant 2019CGXNG-023. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Guofei Gu. *(Corresponding authors: Yang Xiao; Wei Liang.)*

X. Xia is currently with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China. She was with the Key Laboratory of Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China, also with the Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110016, China, and also with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA.

Y. Xiao is with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA (e-mail: yangxiao@cs.ua.edu).

W. Liang is with the Key Laboratory of Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China (e-mail: weiliang@sia.cn).

Digital Object Identifier 10.1109/TIFS.2019.2921232

[1]Electricity theft is defined as using electricity from utility companies through totally or partially bypassing metering system or interfering this system to adulterate its measurement.

[2]The users stealing electricity are referred to as "malicious users".

from the above works, the authors in papers [14]–[22] propose to install redundant monitoring devices such as inspectors and sensors to detect malicious users. Nonetheless, these works suffer from either prohibitively expensive deployment cost or long detection time.

In this paper, we employ a limited number of inspectors to locate malicious users. The inspectors are function-enhanced smart meters with stronger computation capability and larger storage space. Our goal is to locate all malicious users within the shortest detection time. After detecting the existence of malicious users, we assess suspicions that users commit electricity theft before inspectors conduct any further inspections. The suspicions are comprehensively assessed through analyzing prior records of electricity theft from the perspective of criminology, as well as deviations between reported and predicted normal electricity consumptions. On the basis of these suspicions, we propose an inspection algorithm, called Suspicion Assessment based Inspection (SAI), in which, inspectors first probe users with the highest suspicions individually. After this process, the remaining users are probed by a binary tree based inspection strategy. The binary tree, whose leaf nodes represent the remaining users, is built in line with the suspicions. Users with larger suspicions have shorter distances from the root. We apply this particular binary tree as a logical structure to facilitate the inspection process, with each node representing one possible inspection step. The inspection order of the nodes on the binary tree is also determined by the suspicions. Specifically speaking, between two sibling nodes, the inspector will always first inspect the subtree of the node where users' average suspicion is larger. The SAI algorithm can deal with static cases where new malicious users appear during the inspection process as well as dynamic cases in which new malicious users do not appear. The major contributions of this paper are highlighted as follows: (1) We assess suspicions that users steal electricity through analyzing prior records as well as consumption deviations; (2) We propose the SAI algorithm, by which users with the highest suspicions will be first probed individually. Then, the remaining users will be probed by a binary tree based inspection strategy; (3) Experiment results show that the SAI algorithm outperforms existing methods.

The rest of this paper is organized as follows: Section II reviews the related work. Section III introduces the problem. Section IV assesses suspicions that users commit electricity theft. Section V demonstrate the working strategy of the SAI algorithm. Section VI reports experiment results and Section VII concludes the paper.

II. Related Work

Extensive works have been done on detection of electricity theft in smart grid. Defensive techniques vary from hardening smart meters to applying various inspection algorithms.

In papers [23], [24], the authors design several new types of smart meters that have an extra function of automatically detecting electricity theft. However, adding hardware, such as co-processors and tamper resilient memory, inevitably increases the price of smart meters. There are millions of smart meters that have already been installed around the globe. Billions of smart meters are expected to be deployed in the next few years. To replace all meters with newly designed smart meters, the cost will be extraordinarily huge. Moreover, they cannot deal with the cases where malicious users intercept communications to block or alter consumption readings during transmission. Thus, in both industrial and scientific communities, the envisioned meters are not recommended as a priority.

In papers [8]–[10], [25], the authors try to address the electricity theft issue by applying various machine learning approaches. Among these approaches, a classifier is first trained with a historical dataset. It is then applied to find irregularities or deviations in the customer energy consumption profiles. For example, in papers [8], [9], a support vector machine algorithm and a genetic algorithm are jointly employed to analyze meters' load profile information and additional attributes. Similarly, an extreme learning machine algorithm and its online sequential version [10] are utilized to classify meters and reveal whether any significant irregularities emerge in their electricity consumptions. These papers aim to exposing abnormal behaviors that are highly correlated with non-technical loss activities. However, they have a relatively low detection rate but a relatively high false positive rate. This will raise controversy between users and utility companies.

Another radically different line of work is to install specific monitoring devices. In papers [18]–[20], a central observer meter is employed to register the total electricity in a neighborhood. Malicious users are identified through modeling users' behaviors with different mathematical approaches. For example, in papers [18], Lagrange polynomial interpolation is used. In papers [29], the authors propose a Binary-Coded Grouping-based Inspection (BCGI) algorithm which groups users based on digit 1 of users' binary notations. The BCGI algorithm can exactly locate a unique malicious user with just one inspection. However, the BCGI algorithm can only deal with the case where there is one malicious user. In papers [15]–[17], an "inspector box" is installed at the distribution room in each neighborhood area network (NAN). Malicious users are identified by comparing the inspectors' own readings with the users' reported readings. The authors in papers [16], [17] adopt a binary tree as a logical structure to facilitate the inspection process. They propose a series of inspection algorithms, by which inspectors dynamically traverse on the binary tree to detect malicious users. Among these algorithms, Adaptive Tree Inspection (ATI) algorithm [16] is the most practical. It is a heuristic approach by which inspectors skip some internal nodes on the binary tree to directly inspect nodes at lower levels. Difference-Comparison-based Inspection (DCI) algorithm [17] can skip a large amount of nodes on the binary tree, and hence can locate malicious users faster. However, it assumes that inspectors can check malicious users' logs to know their true electricity consumptions. In the real world, malicious users will be smart enough to tamper with all logs when compromising smart meters. This makes the DCI algorithm less practical.

Some researchers apply state estimation-based approaches to address electricity theft detection issues. However, as indicated in [26], [27], electricity theft using state estimation can only be detected at the sub-station level instead of at the end user level, i.e., theft detection of middle-to-low voltage

transformers serving malicious users. To explicitly identify malicious users, further analyses or inspections are needed. For example, the paper [28] follows state estimation results to localize electricity usage irregularity at the distribution transformers and proposes an analysis of variance (ANOVA) method to create a suspect list of customers with metering problems. In [27], the results calculated by a state estimator and the data captured by filed devices are used as inputs of a multivariate procedure of monitoring and control to detect a possible power loss at distribution transformer terminals; afterwards, the authors apply a pathfinding procedure based on an A-star (A*) algorithm to locate the consumption points with power loss, whose coordinates in a geographical map are further determined using a geographical information system.

During the 1970s and early 1980s, evidence accumulated to indicate that a relatively small group of offenders committed most serious offenses [34]. These findings, coupled with increasing pressures on the budgets of criminal justice agencies, led to calls for more effective use of public expenditures for crime control by identifying and incarcerating the most serious and persistent offenders [34]. These calls challenge the research community to focus on the problem of predicting which individuals will commit crimes in the future. Many researches [34]–[39] are conducted to explore the issue of old prior records and their ability to predict future offending. In this paper, we will utilize some of these results and apply them to our applications.

In this paper, we propose the SAI algorithm to detect malicious users in the NAN. Before inspectors conduct inspections, suspicions that users steal electricity are assessed through analyzing prior records of electricity theft as well as deviations between reported and predicted normal electricity consumptions. In the SAI algorithm, users with the highest suspicion will be first probed individually. Then, the remaining users will be probed in line with a binary tree based inspection strategy. The inspection order of the nodes on the binary tree is also determined by the suspicions.

III. Problem Statement

In smart grid, electrical grids and communication networks overlay with each other [30]. In this paper, we consider a neighborhood area network (NAN) in a smart grid, which is defined as the utility companies' last-mile, outdoor access network that connects smart meters and distribution automation devices to a wide-area network (WAN) [31], as shown in Fig. 1, in which the solid and dashed lines with double-ended arrows represent two-way electrical flows and communication flows, respectively. The end users' smart meters are connected with the distribution automation devices which are usually installed at some places (e.g., on an electrical pole or in a distribution room) in the NAN via an individual power line. This implies that the NAN usually covers a small area (e.g., an apartment building). We suppose that there are a total number of n users in the NAN which are notated as $U = \{1, 2, \ldots, n\}$. At each user's premises there is a smart meter which records and reports electricity consumptions periodically. Let integers, $t = 1, 2, \ldots$, denote smart meters' reporting periods. We assume that smart meters start working from period 1. Based upon utility companies, users can be classified into two categories: (1) honest users who report electricity consumptions genuinely and (2) malicious users who report electricity consumptions less than what are actually consumed.

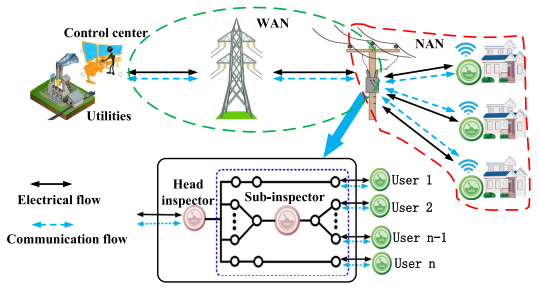


Fig. 1. A simplified architecture of smart grid.

This paper aims to identify all malicious users within the shortest detection time, which is formulated as the Malicious Meter[3] Inspection (MMI) problem in [16]. For this purpose, we install an inspector box [16] which serves as the distribution device and is assumed to be either secure or equipped with tamper-resistant components/functions. The inspector box contains a head inspector and several sub-inspectors. For example, in Fig. 1, we elaborate a possible design of the inspector box with one sub-inspector. The head inspector monitors all users in the NAN. It aims to detect whether there are reading anomalies in the NAN. Once the head inspector detects reading anomalies, sub-inspectors will try to locate malicious users exactly. We assume that users monitored by sub-inspectors can be arbitrarily changed manually or automatically, without interfering with any normal electricity services [16].

We now explain how the inspectors probe users. Let G denote a group of users monitored by an inspector, which can be the head inspector or a sub-inspector. For the head inspector, we have $G = U$; for a sub-inspector, we have $G \subseteq U$. When an inspector probes users in G, it works as follows: (1) Measuring the total amount of electricity distributed to the users in G; (2) Receiving the reported electricity consumptions of the users in G. If the dispute between the inspector's own reading and the summation of the reported readings exceeds a previously specified threshold, we can conclude that there are malicious users in G. The inspection result is correspondingly called "dirty". Otherwise, if the dispute is less than the threshold, all the users in G can be regarded as being honest. The inspection result is accordingly called "clean". In a real application, the threshold is equal to the summation of technical losses during the electricity transmission of all users being probed [16], [22]. Even though it is difficult to obtain the accurate value of technical losses of each user, we can estimate it using some existing mathematical models [53], which are not detailed here since they are out of the scope of this paper. To deal with the bias between the accurate and estimated technical loss, we can introduce a compensation value which

[3]In this paper, the two terms "user" and "meter" are exchangeable.

Algorithm 1 The Probe Operation

```
procedure PROBE(G)
    Measuring the total electricity consumed by users in G;
    Receiving the reported readings of users in G;
    if the dispute exceeds a specific threshold then
        There are malicious users in G; ▷ a dirty inspection result
    else
        Users in G are honest;  ▷ a clean inspection result
    end if
end procedure
```

can be carefully chosen by performing trial experiments before actually employing a specific inspection algorithm to locate malicious users in a specific NAN [22]. We conclude the probe operation in **Algorithm** 1. Note that in this paper, no matter the head inspector detects reading anomalies or the sub-inspectors locate malicious users, they conduct probing operations. When a sub-inspector performs the probing operation for one time, we say it conducts one inspection (step).

Note that each probing operation (i.e., inspection) lasts for one reporting period. Thus, the goal of this paper, which, as aforementioned, is to minimize the detection time, can be abstracted as minimizing the number of inspection steps. We consider both static cases and dynamic cases [16]. In static cases, no new malicious users will appear during the inspection process. On the other hand, in dynamic cases, new malicious users will appear during the inspection process.

In this paper, we assume that once a user is identified as being malicious, the utility company will immediately do the following two things: (1) noting down the period when the dirty meter is caught stealing electricity; (2) disconnecting this user from the service of electricity. The first assumption is obviously practical. We refer to these periods as users' prior records of electricity theft. With regard to the second assumption, it is consistent with the situation in the real world. As reported in [32], if the utility companies find that meter tampering has occurred, they will immediately disconnect the account and require the individual to come into the office to pay their whole balance and even a fine.

The main notations in this paper are listed in Table I. In general, we use lowercase letters to denote variables, uppercase letters to notate sets, and bold uppercase letters to represent vectors.

TABLE I
NOTATIONS

Notations	Descriptions
U	The set of all the users in the NAN. We have $U = \{1, 2, \ldots, n\}$, where n denotes the total number of users.
G	A group of users monitored by an inspector.
t, t^*	The integers $t = 1, 2, \ldots$ are used to number the reporting periods of smart meters. Specially, period t^* denotes the period when the head inspector detects reading anomalies.
$R(i,t)$	A set of periods prior to period t when user i is caught committing electricity theft. For the users who have never stolen electricity until period t, we have $R(i,t) = \emptyset$. For the users who have committed electricity theft for multiple times, we have $R(i,t) = \{t_{i,1}, t_{i,2}, t_{i,3}, \cdots\}$, where $t_{i,j}$ denotes the period when user i is caught committing electricity theft for the j-th time.
$r(i,t)$	User i's recidivism risk at period t. Generally, users with more prior records have larger recidivism risks. The more recently users commit the last electricity theft, the larger recidivism risks the corresponding users have.
$q(i,t)$, $q'(i,t)$, $q''(i,t)$	The terms $q(i,t)$, $q'(i,t)$, $q''(i,t)$ denote user i's actual, reported, and predicted electricity consumption at period t, respectively.
$d(i,t)$, $d_r(i,t)$, $\bar{d}_r(i,t_0,t)$	$d(i,t)$ means user i's consumption deviation at period t and is defined as the difference between $q''(i,t)$ and $q'(i,t)$; $d_r(i,t)$ denotes user i's relative consumption deviation at period t and is defined as the ratio of $d(i,t)$ to $q''(i,t)$; $\bar{d}_r(i,t_0,t)$ denotes user i's mean relative consumption deviation during the time interval $[t_0, t]$.
$h(i,t)$	User i's deviation risk at period t, which is obtained by analyzing user i's consumption deviation: the larger deviations between the reported and predicted normal readings, the higher risks the corresponding users have.
$s(i,t)$	User i's suspicion to commit electricity theft at period t, which is defined as a weighted value of $r(i,t)$ and $h(i,t)$.

IV. Suspicion Assessment

In this section, we assess suspicions that users commit electricity theft, through analyzing prior records of electricity theft as well as deviations between reported and predicted normal consumptions.

A. Prior Records

In this subsection, we assess suspicions that users steal electricity through analyzing prior records of electricity theft from the perspective of criminology. This is reasonable because electricity theft is essentially a particular form of economic crimes [33]. Many researches [34]–[39] are conducted to explore the issue of old prior records and their ability to predict future offending. Several well-documented empirical facts are summarized as follows [35], [36]: (1) Compared to individuals who have never offended, individuals who have offended in the past are relatively more likely to offend in the future. (2) Offenders with more prior criminal records are more likely to recidivate in the future. As shown in Fig. 2(a), with the increase of the number of prior criminal records, the risk of recidivism first increases and then stays stable. (3) The risk of recidivism declines as the time since the last criminal act increases. As shown in Fig. 2(b), individuals with criminal records do exhibit significantly higher risks of future criminal conducts than individuals without criminal records. However, the difference weakens dramatically and quickly with the increase of time since the last criminal act. At last, the risks of new offenses begin to approximate (but not match).

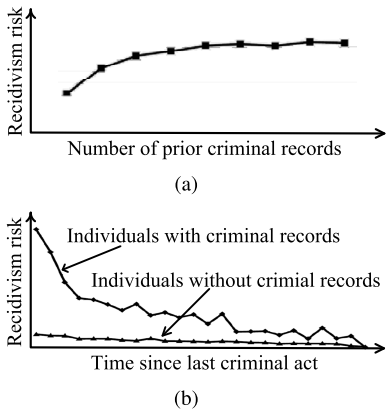


Fig. 2. Two well-documented empirical facts [35]. (a) Offenders with more criminal prior records are more likely to recidivate in the future. (b) The risk of recidivism declines as the time since the last criminal act increases.

In our case, the offenders are the users stealing electricity. As aforementioned, in this paper, we assume that once a user is identified as being malicious, the utility company will note down the period. Let $R(i,t)$ denote user i's prior records before period t, which are a set of periods when user i is caught committing electricity theft prior to period t. Particularly, for users who have never stolen electricity until period t, we have $R(i,t)=\emptyset$. For users who have committed electricity theft for multiple times, we have $R(i,t)=\{t_{i,1}, t_{i,2}, t_{i,3}, \cdots\}$, where $t_{i,j}$ denotes the period when user i is caught committing electricity theft for the j-th time. Notably, for $t_{i,j} \in R(i,t)$, we have $1 \leqslant t_{i,j} \leq t$.

As discussed earlier, the prior criminal records can, to a large extent, predict the recidivism risk of offenders. Specially, the total number of prior criminal records and the time since the last criminal act are the most significant. Thus, we assess users' risks of stealing electricity based upon the following two characteristics: (1) the total number of times that a user has been caught stealing electricity; for user i with prior records $R(i,t)$, it can be denoted as $|R(i,t)|$, where $|\cdot|$ represents the cardinality of a set; (2) the time interval between the current period and the period when a user commits the last electricity theft; for users who have committed electricity theft before period t, it can be denoted as $t - \max(R(i,t))$, where $\max(\cdot)$ returns the maximum value; particularly, for users with $R(i,t)=\emptyset$, we set this time interval as $+\infty$.

In the real application, when we assess users' recidivism risk, it is more practical to measure the time interval since the last electricity theft with the time unit such as days, weeks, months, etc., than using the time unit of reporting periods (which is usually set as 15 minutes as an example). To be consistent with the criminology researches, in this paper we adopt the time unit of months. Assume that smart meters report a total number of T readings to the utility companies during one month. Let $y(i,t) = \lfloor \frac{t-\max(R(i,t))}{T} \rfloor$. Then, we use $|R(i,t)|$ and $y(i,t)$ for assessing users' recidivism risk. According to the discussion before, we know that the recidivism risk increases monotonically with the total number of prior electricity thefts and decreases monotonically with the time interval since the last electricity theft. Let $r(i,t)$ denote user i's recidivism risk at period t. Then, we have $r(i,t) = f(|R(i,t)|, y(i,t))$. According to the above analysis, we can derive $\frac{\partial f}{\partial |R(i,t)|} > 0$ and $\frac{\partial f}{\partial y(i,t)} < 0$.

In this paper, we would like to make the recidivism risk $r(i,t)$ distributed in the interval $(0,1)$. Since the sigmoid function is quite similar to the real world thought process and adds the element of fuzziness to a conventional linear process [41], we in this paper use it to assess users' recidivism risks, as follows:

$$r(i,t) = f(|R(i,t)|, y(i,t)) = \frac{1}{a + b\exp(-w|R(i,t)| - (w-1)y(i,t))} + c(t), \quad (1)$$

where $c(t)$ is the risk of users with no prior records to commit electricity theft at period t. The constant a, b are subject to $\frac{1}{a} + c(t) < 1, b > 0$, respectively. The coefficient w is a weight factor satisfying $0 < w < 1$.

B. Consumption Deviation

In this subsection, we assess suspicions that users commit the electricity theft by analyzing deviations between reported and predicted normal consumptions.

With regard to the prediction of users' electricity consumptions, many kinds of technologies, such as the artificial neural networks and support vector machines, can be applied [42]. These technologies depend heavily on the computation capability of the devices. In our case, the inspectors are embedded devices whose computation ability is not strong enough. Thus, they are not suitable here. In the real life, people usually have similar routines on different days. For example, in the USA, on weekdays, most people leave home for work at about 8:00 a.m. and return home at about 7:00 p.m. This implies that users' load curves have a trend to repeat themselves. This trend, according to the studies on time series [43], is called "seasonality". For time series with "seasonality" characteristic, the Holt-Winters method which belongs to the exponential smoothing forecasting technique shows excellent prediction performance [43]. Thus, it is applied here to predict users' normal electricity consumptions; and the details are given in subsection IV-D.

Let $q'(i,t)$ and $q''(i,t)$ denote user i's reported and predicted electricity consumptions at period t, respectively. Let $d(i,t)$ denote user i's consumption deviation at period t, which is defined as the difference between $q'(i,t)$ and $q''(i,t)$. Mathematically speaking, we have $d(i,t) = q''(i,t) - q'(i,t)$.

When analyzing which users are more likely to steal electricity, the consumption deviation cannot be simply used. This can be easily understood when we consider the users whose consumption deviations are almost the same but whose predicted normal electricity consumptions differ a lot. Obviously, we cannot say that they are equally likely to commit electricity theft. Hence, we introduce the concept of relative consumption deviation which is defined as the ratio of users' consumption deviation to predicted normal electricity consumption. Let $d_r(i,t)$ denote user i's relative consumption deviation at period t. Then, we have $d_r(i,t) = \frac{d(i,t)}{q''(i,t)} = \frac{q''(i,t)-q'(i,t)}{q''(i,t)}$.

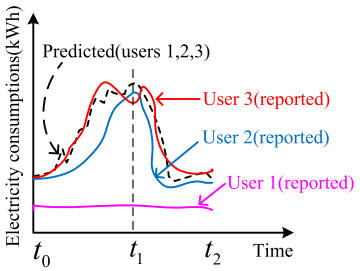


Fig. 3. An example: users 1, 2, and 3 have the same predicted readings but different reported readings.

Now, we are faced with the following problem: when we analyze which users are more likely to commit electricity theft, is it appropriate for us to consider the relative consumption deviation just at a certain period? The answer is obviously not. For example, in Fig. 3, we assume that during the time interval $[t_0, t_2]$, users 1, 2, 3 have the same predicted electricity consumptions; user 1's reported readings are much smaller than predicted normal readings; user 2's reported readings always lie in between user 1's reported readings and the predicted normal readings; user 3's reported readings are very close to the (predicted) normal readings for most of the time, but much smaller than that around period t_1. Empirically, during the time interval $[t_0, t_2]$, among the three users, user 1 is the most likely to commit electricity theft, while user 3 is the most likely to be honest. However, when we analyze users' possibility to commit electricity theft at period t_1, if we just consider the relative consumption deviation at period t_1, this conclusion will not be reached. A more reasonable way is to observe users' reported readings and predicted normal readings during the time interval $[t_0, t_1]$. This inspires us to involve the relative consumption deviation for a while when analyzing users' possibility of committing the electricity theft. Let $\bar{d}_r(i, t_0, t)$ denote user i's average relative consumption deviation during time interval $[t_0, t]$. Then, we have

$$\bar{d}_r(i, t_0, t) = \frac{1}{t - t_0 + 1} \sum_{\tau=t_0}^{t} d_r(i, \tau) = \frac{1}{t - t_0 + 1} \sum_{\tau=t_0}^{t} \frac{q''(i, \tau) - q'(i, \tau)}{q''(i, \tau)}.$$

We say a larger $\bar{d}_r(i, t_0, t)$ implies that user i is more likely to commit the electricity theft at period t. As aforementioned, the sigmoid function is quite similar to the real world thought process and adds the element of fuzziness to a conventional linear process [41]. Thus, we apply it to assess the risk of users to steal electricity through analyzing consumption deviations, as follows. Let

$$h(i, t) = \frac{1}{1 + \exp\left(-\bar{d}_r(i, t_0, t)\right)}, \quad \forall t \geqslant t_0$$

In the following context, we call $h(i, t)$ as user i's deviation risk at period t. Combining the above two equations, we can derive the deviation risk

$$h(i, t) = \frac{1}{1 + \exp\left(\frac{1}{t - t_0 + 1} \sum_{\tau=t_0}^{t} \frac{q'(i,\tau) - q''(i,\tau)}{q''(i,\tau)}\right)}.$$

Apparently, we have $0 < h(i, t) < 1$. For two users i and j, if $h(i, t) > h(j, t)$, we say that user i is more likely to commit the electricity theft at period t than user j. Let t^* denote the period when the head inspector detects reading anomalies. In application, we need to evaluate users' suspicions only after period t^*. Thus, we usually set $t_0 = t^*$.

C. Electricity Theft Suspicion

In the previous two sub-sections, we first assess the recidivism risks that users recommit electricity in line with their prior records from the perspective of criminology. Then, we further analyze the deviation risks based upon the deviations between reported readings and predicted normal readings. By integrating the above two aspects together, we in the following assess suspicions that users steal electricity.

Let $s(i, t)$ denote the suspicion that user i commits electricity theft at period t. Obviously, the suspicion $s(i, t)$ increases monotonically with both the recidivism risk $r(i, t)$ and the deviation risk $h(i, t)$. However, when we analyze the suspicion $s(i, t)$, should the weight between $r(i, t)$ and $h(i, t)$ be set statically or dynamically? Since the recidivism risk $r(i, t)$ is based on prior records, we can say that it represents how user i behaves in the past. By contrast, the deviation risk $h(i, t)$ reflects how user i behaves currently. Usually, when making a decision, although the information related to the past can give us some clues, we should value more on the present. Based on this principle, we conclude that the deviation risk $h(i, t)$ should weigh dynamically more and more with inspection process going on. Thus, we define the suspicion $s(i, t)$ as

$$s(i, t) = u^{\lfloor \frac{t - t^*}{g} \rfloor} r(i, t) + (1 - \mu) h(i, t), \tag{2}$$

where $0 < \mu < 1$ is a weight factor, g is a positive integer. Due to $0 < r(i, t) < 1$ and $0 < h(i, t) < 1$, we can derive $0 < s(i, t) < 1$.

D. Predicting Normal Electricity Consumptions

In the following, we explain how the Holt-Winters method works when predicting users' normal electricity consumptions. Since the inspectors only can measure the sub-total amount of a group of users and cannot measure the amount of a particular user, measurements from the inspectors cannot be used as baseline. Let $\mathbf{Q}(i, t)$ denote user i's historical electricity consumptions before period t, which is a series of user i's actual electricity consumptions. Technically speaking, we have $\mathbf{Q}(i, t) = (q(i, 1), q(i, 2), \ldots, q(i, t-1))$. Before period t^*, all the users in the NAN are honest. Thus, we have $\forall t < t^*, q(i, t) = q'(i, t)$. However, after period t^*, malicious users appear. This means users' reported readings are no longer trustworthy and that the values $q(i, t), \forall t \geqslant t^*$ cannot be determined. Hence, in this paper, we define $\mathbf{Q}(i, t)$ as follows:

$$\mathbf{Q}(i, t) = \begin{cases} \left(q'(i, 1), q'(i, 2), \ldots, q'(i, t-1)\right), & \forall t < t^* \\ \left(q'(i, 1), q'(i, 2), \ldots, q'(i, t^* - 1)\right), & \forall t \geqslant t^*, \end{cases}$$

Specifically, let $\mathbf{Q}_\rho(i, t)$ denote user i's historical electricity consumptions in the latest ρ days before period t, where

$\rho \in \mathbb{N}^+$. Let υ denote the total number of readings generated by the smart meters every day. Then, we have

$$\mathbf{Q}_\rho(i,t) = \begin{cases} \left(q'(i,t-\rho\upsilon),\ldots,q'(i,t-1)\right), & \forall t < t^* \\ \left(q'(i,t^*-\rho\upsilon),\ldots,q'(i,t^*-1)\right), & \forall t \geqslant t^*, \end{cases}$$

To obtain user i's predicted electricity consumption at period t, i.e., $q''(i,t), \forall t \geqslant t^*$, for the purpose of reducing the computation complexity, we in application choose $\mathbf{Q}_\rho(i,t)$ rather than $\mathbf{Q}(i,t)$. To guarantee the prediction accuracy, as a rule of thumb, we usually have $\rho \geqslant 2$ [43]. Considering that in reality we only need to predict users' normal electricity consumption after the head inspector detects reading anomalies, we in the following simply let $\mathbf{Q}_\rho(i,t) = \left(q'(i,t^*-\rho\upsilon),\ldots,q'(i,t^*-1)\right)$.

The Holt-Winters method assumes the users' electricity consumptions comprise of level, trend and seasonal index [49] and that the forecasting value is obtained through the above three components. Specifically, with the historical data in $\mathbf{Q}_\rho(i,t)$, Holt-Winters method predicts user i's electricity consumption η ($\eta \in \mathbb{N}^+$) periods ahead of period t as follows [43] [49]: $q''(i,t+\eta) = (d(i,t)+\eta e(i,t))\, o(i,t+1-\upsilon+\lfloor(\eta-1) \bmod \upsilon\rfloor)$, where $d(i,t) = \alpha\frac{q'(i,t)}{o(i,t-\upsilon)} + (1-\alpha)(d(i,t-1)+e(i,t-1))$, $e(i,t) = \beta(d(i,t)-d(i,t-1)) + (1-\beta)e(i,t-1)$, and $o(i,t) = \varphi\frac{q'(i,t)}{d(i,t)} + (1-\varphi)o(i,t-\upsilon)$.

The parameters α, β and φ are the level, trend, and seasonal smoothing factors, respectively, which satisfy $0 < \alpha < 1$, $0 < \beta < 1$ and $0 < \varphi < 1$. They could be estimated by minimizing the root mean square error [49] of the normal electricity consumption prediction, i.e., $\min : \sqrt{\frac{1}{(\rho-1)\upsilon+1}\sum_{t=t^*-(\rho-1)\upsilon}^{t^*-1}(q''(i,t)-q'(i,t))^2}$. The initialization of $d(i,0)$, $e(i,0)$, and $o(i,k), \forall k = 0,1,\ldots,m-1$ is as follows [50]: $d(i,0) = q'(i,t^*-\rho\upsilon)$, $e(i,0) = \frac{1}{\upsilon^2}\sum_{j=0}^{\upsilon-1}\left(q'(i,t^*-\rho\upsilon+\upsilon+j)-q'(i,t^*-\rho\upsilon+j)\right)$, and $o(i,k) = \frac{1}{\rho}\sum_{j=0}^{\rho-1}\frac{q'(i,t^*-\rho\upsilon+j\upsilon+k)}{\xi_j}, \forall k = 0,1,\ldots,\upsilon-1$, where $\xi_j = \sum_{k=0}^{\upsilon-1}\frac{q'(i,t^*-\rho\upsilon+j\upsilon+k)}{\upsilon}, \forall j = 0,\ldots,\rho-1$ is the average electricity consumption in the j-th day from the period $t^*-\rho\upsilon$.

The electricity consumption patterns on weekdays are usually different from that at weekends. Thus, to predict the normal electricity consumptions on weekdays (at weekends), the latest ρ weekday (weekend) electricity consumption data are used.

V. The SAI Algorithm

In this section, we first demonstrate the working strategy of the Suspicion Assessment based Inspection (SAI) algorithm and then explain how to implement it in real applications. We assume that we have obtained all suspicions that users commit electricity theft.

A. Working Strategy

We define a round of inspection as the inspection process during which the users whose statuses are first unclear are identified as being malicious or honest. In static cases where no new malicious users appear, sub-inspectors conduct one round of inspection. In dynamic cases where new malicious users do appear, they usually conduct multiple rounds of inspection. Let W denote the set of users whose statuses are not clear. When the head inspector detects reading anomalies, we initiate $W = U$ and start the first round of inspection. The notation U, as defined earlier, represents the set of all users in the NAN. In the inspection process, W will be updated constantly.

Intuitively, for the purpose of shortening detection time, each round of inspection should start from the users with the highest suspicions. Explicitly, we cannot say that the users with highest suspicions are definitely malicious. On the contrary, in the real world, some of these users may be actually honest. This is mainly because many non-malicious factors can cause users' electricity consumptions to be much lower than normal readings. For example, when users are traveling out, they will obviously consume much less electricity. Thus, sub-inspectors should conduct inspections on these users to confirm whether they commit electricity theft or not. It is reasonable to reckon that among the users with the highest suspicions, there are more malicious users than honest users. In this case, the individual inspection strategy which probes users individually is efficient to identify malicious users. Hence, we apply it to probe users with the highest suspicions, as follows.

Let m denote the number of malicious users in the NAN. Clearly, in the real world, we do not know the value of m in advance. However, we can roughly estimate it as the average number of malicious users identified in the past. Let m_j denote the number of malicious users that have been found out when the head inspector detects reading anomalies at the j-th time, with j being a positive integer. Let $\tilde{m}$ denote the estimated value of m. Then, for the l-th time that the head inspector detects reading anomalies, we estimate $\tilde{m}$ as

$$\tilde{m} = \frac{1}{l}\sum_{j=0}^{l-1} m_j. \tag{3}$$

where m_0 is previously set by the utility companies. For example, the utility company may set it as $m_0 = 1\%n$, since it was reported that 1% of users were stealing power in 1984 in the USA [44].

After obtaining $\tilde{m}$, the sub-inspectors will individually probe the $\tilde{m}$ users with the highest suspicions. For each inspection step, we will first update suspicions that users in W commit electricity theft. Then, the user with the highest suspicion is probed. If the inspection result is dirty, this user is identified as being malicious; otherwise, this user is honest. The above process is concluded in lines $4 \sim 13$ in **Algorithm** 2. Now, we take the example in Fig. 4 to illustrate the above process. In Fig. 4, we assume that there are ten users in the NAN. When the inspection starts, we initiate $W = \{1,2,\cdots,10\}$. We assume $\tilde{m} = 4$. Then, the four users $1,2,3,4$ with the highest suspicions are probed individually. Users $1,2,4$ are identified as being malicious, whereas user 3 is identified as an honest user. At this time, the set W is updated

Algorithm 2 The Suspicion Assessment Based Inspection (SAI) Approach

Require: $U = \{1, 2, ..., n\}$
Ensure: M, H ▷ the set of malicious and honest users, respectively
Initialization: $M \leftarrow \emptyset, H \leftarrow \emptyset, W \leftarrow U$ ▷ Initialization
1: **while** the head inspector detects reading anomalies **do**
2: $k \leftarrow 1$; ▷ start one round of inspection
3: **while** the head inspector detects reading anomalies **do**
4: **if** $k \leq \tilde{m}$ **then** ▷ individual inspection
5: Update suspicions for users in W;
6: user $i \leftarrow$ the user in W with the highest suspicion;
7: probe(user i);
8: **if** the inspection result is dirty **then**
9: $M \leftarrow M \cup \{\text{user } i\}$, $W \leftarrow W \setminus \{\text{user } i\}$;
10: **else**
11: $H \leftarrow H \cup \{\text{user } i\}$, $W \leftarrow W \setminus \{\text{user } i\}$;
12: **end if**
13: $k \leftarrow k + 1$;
14: **else** ▷ binary tree based inspection
15: $z \leftarrow$ build_bit(W);
16: Update node $z.sta \leftarrow$ "dirty"; ▷ skip the root node
17: $z \leftarrow$ next_node(z);
18: **while** node z **do**
19: probe(leaf(z));
20: **if** the inspection result is dirty **then**
21: Update node $z.sta \leftarrow$ "dirty";
22: **if** node z is a leaf **then**
23: $M \leftarrow M \cup \text{leaf}(z)$; $W \leftarrow W \setminus \text{leaf}(z)$;
24: **break**;
25: **else** ▷ node z is an internal node
26: $z \leftarrow$ next_node(z);
27: **end if**
28: **else** ▷ the inspection result is clean
29: $H \leftarrow H \cup \text{leaf}(z)$; $W \leftarrow W \setminus leaf(z)$;
30: Update node $z.sta \leftarrow$ "clean"
31: **if** node $z.par.sta$ is "dirty" **then**
32: $z \leftarrow z.sib$;
33: Perform inspections by lines 21 $\sim$ 27;
34: **end if**
35: **end if**
36: **end while**
37: **end if**
38: **end while**
39: $W \leftarrow H$; $H \leftarrow \emptyset$; ▷ End one round of inspection
40: **end while**

41: **procedure** BUILD_BIT((W))
42: Update suspicions that users in W commit electricity theft;
43: Create a set of leaf nodes, Z, to represent the users in W;
44: **while** $|Z| > 1$ **do** ▷ Build a BIT
45: Allocate a new node z_0;
46: $z_0.lch \leftarrow z_1 \leftarrow$ extractMin (Z);
47: $z_0.rch \leftarrow z_2 \leftarrow$ extractMin (Z);
48: $z_1.sib \leftarrow z_2$, $z_2.sib \leftarrow z_1$;
49: $z_0.sp \leftarrow z_1.sp + z_2.sp$, $z_0.sta \leftarrow$ null;
50: $Z \leftarrow Z \cup \{z_0\}$;
51: **end while**
52: **Return** $z \leftarrow$ the unique node remaining in set Z;
53: **end procedure**

54: **procedure** NEXT_NODE((z))
55: **if** aver_susp($z.lch$) $\geq$ aver_susp($z.rch$) **then**
56: **Return** node $z \leftarrow$ node $z.lch$;
57: **else**
58: **Return** node $z \leftarrow$ node $z.rch$;
59: **end if**
60: **end procedure**

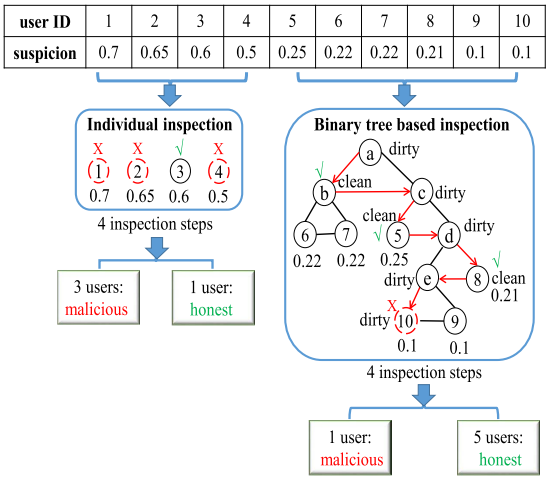


Fig. 4. An example to illustrate the SAI algorithm. The symbols ✓ and X represent a clean inspection and a dirty inspection, respectively. For the convenience of illustration, we assume that users' suspicions do not change during the inspection course.

as $\{5, 6, \cdots 10\}$. Note that in Fig. 4, for the convenience of illustration, we assume that users' suspicions do not change during the inspection course.

Since the account of users who are identified as being malicious will be immediately disconnected, they can be regarded no longer being monitored by the head inspector. Thus, during the inspection course, if the head inspector stops detecting reading anomalies, we can know that all malicious users have been located. In other words, if there are any users still remaining in set W, all of them are honest. At this time, the inspection process can be ceased immediately.

For the purpose of demonstrating the complete SAI algorithm, in the following context, we assume that after the individual inspection process, there are malicious users still remaining in W. It happens a lot in the real application, mainly due to the following two factors: (1) the estimate of the number of malicious users is less than itself, i.e., $\tilde{m} < m$; (2) users with low suspicions may also be malicious. For example, a user who consumes much more electricity for holding a party is likely to report normal readings to utility companies. In this case, the suspicion is low, but this user actually steals electricity. Among the users with low suspicions, there should be more honest users than malicious users. As indicated in paper [16], when the ratio of malicious users is small, the binary tree can be applied as a logical structure to facilitate and accelerate the inspection process. This inspires us to apply a binary tree based inspection strategy on the remaining users in W.

We next demonstrate how to establish the particular binary tree. After updating the suspicions that the remaining users in W commit electricity theft, the binary tree is built in line with the refreshed suspicions. The building process is stated as follows.

(1) First, we create a set of leaf nodes, which is denoted by Z, to represent the users in W. For any leaf node $z \in Z$, it has four attributes: suspicion, left child, right child, and parent, which are denoted as $z.sp$, $z.lch$, $z.rch$, $z.par$, respectively. For any leaf node z, its left child, right child and

parent will be initiated as empty; and its suspicion will be set as the suspicion value of the corresponding user.

(2) Then, we repeat merging two nodes in set Z which have the lowest suspicions. Specifically, we will first allocate a new node z_0. We denote the two nodes by z_1, z_2 which have the lowest suspicions in set Z. These two nodes are extracted from set Z and then assigned as the left and right children of the new node z_0, respectively. The suspicion value of the new node z_0 is set as the summation of the suspicion values of nodes z_1 and z_2. Technically, we have $z_0.sp = z_1.sp + z_2.sp$. We then add the new node z_0 into set Z. At this time, the set Z is updated as $Z = Z \setminus \{z_1, z_2\} \cup \{z_0\}$.

Obviously, after such a merging process, the number of nodes in Z will be reduced by one. Thus, after it is repeated for $|Z|-1$ times, there will be only one node in set Z. This unique node is exactly the root of the binary tree. Obviously, on this binary tree, the users with larger suspicions have shorter distances from the root. The above process is summarized as the BUILD_BIT procedure in lines 41 $\sim$ 53 in **Algorithm** 2, where the function $extractMin\,(Z)$ extracts the node with the largest suspicion from set Z.

We now take Fig. 4 as an example to illustrate how we build the BIT. As shown, for the users remaining in W, i.e., users $\{5, 6, \cdots 10\}$, we build a binary tree. Since users 9, 10 have the same lowest suspicion, i.e., 0.1, the leaf nodes representing them are merged into the internal node e whose suspicion is set as 0.2. Then, node e and the leaf node representing user 8 have the lowest suspicions. Thus, they are merged into the internal node d whose suspicion value is set as 0.41. Next, the two leaf nodes representing users 6, 7 are merged into the internal node b whose suspicion is set as 0.44. With two more merging operations, node c and the root node a are subsequently created. The established BIT is shown in Fig. 4.

After the BIT is built, we will apply it as a logical structure to facilitate the inspection process. For convenience of illustration, for all nodes on the BIT, we now assign them two more attributes: (1) sibling node. For any given node z, it is notated as $z.sib$. Specifically speaking, if node z is a left child, it has a right sibling node; otherwise, it has a left sibling node. (2) status. For any given node z, it is notated as $z.sta$. Each node represents one possible inspection step. Let $\text{leaf}(z)$ return the set of users on the subtree of node z. Obviously, if the inspection result on node z is dirty, there are malicious users in $\text{leaf}(z)$. In this case, we set $z.sta$ as being malicious. Otherwise, if the inspection result on node z is clean, there are malicious users in $\text{leaf}(z)$. In this case, we set $z.sta$ as being clean.

Since it's known that there are still malicious users remaining in W, we can skip the inspection step on the root node and directly set its status as being malicious. The sub-inspectors will conduct the next inspection step on the child node which has a larger average suspicion. Specifically speaking, assume that the status of node z is dirty. Then, the sub-inspectors will conduct the next inspection step on node $z.lch$ if we have $\text{aver\_susp}(z.lch) \geq \text{aver\_susp}(z.rch)$, where $\text{aver\_susp}(z)$ returns the average value of the suspicions of all the users in $\text{leaf}(z)$. Otherwise, if $\text{aver\_susp}(z.lch) < \text{aver\_susp}(z.rch)$, the next inspection step will be performed on node $z.rch$. We conclude the above process as the procedure NEXT_NODE in lines 54 $\sim$ 60 in **Algorithm** 2.

Let H denote the honest user set. During the inspection process, if node z is probed as being clean, we can know that the users in set $\text{leaf}(z)$ are honest. In this case, we will add these users into set H and remove them from set W. The status of node z will also be updated as being clean. In this case, if the status of node z's parent is dirty, we can infer that there are malicious users on the subtree of node z's sibling node. This means that the status of node z's sibling node must be dirty. We then focus on the inspection process on the subtree of this sibling node.

We next consider the cases where node z is probed as being malicious. If node z is an internal node, we only know that there is at least one malicious user on node z's subtree. In this case, more inspections will be further conducted on it. If node z is a leaf, the user represented by it is apparently malicious. Let M denote the malicious user set. This user will be added into set M, but removed from set W. At this time, we end the current binary tree based inspection process. If the head inspector still detects reading anomalies, We will build a new binary tree for the users remaining in W, based upon the refreshed suspicions. A new binary tree based inspection process will then be started.

We now take the example in Fig. 4 to illustrate the above process. As shown in Fig. 4, due to $\text{aver\_susp}(b) > \text{aver\_susp}(c)$, the sub-inspectors first conduct inspection step on node b. As shown, the inspection result is clean. Since the status of node a (i.e., the parent of node b) is dirty, we can infer that node c must be dirty. We then focus on the inspection on node c's subtree. Due to $\text{aver\_susp}(d) = 0.137 < 0.25$, the leaf node representing user 5 is probed next. Since the inspection result is also clean, we can refer that there are malicious users on node d's subtree. Due to $\text{aver\_susp}(e) = 0.1 < 0.21$, the leaf node representing user 8 is probed next. With one more inspection step on the leaf node representing user 10, the sub-inspectors identify user 10 as being malicious. After the account of user 10 is disconnected, the head inspector stops detecting reading anomalies. Thus, we know all malicious users have been located.

Now let's consider the following dynamic cases: after all the users in W are identified as being either malicious or honest, the head inspector can still detects reading anomalies. This implies that new malicious users appear among the users who have already been identified as being malicious. Because the users identified as being malicious are disconnected from the service of electricity, we do not need to consider them in the new rounds of inspections. To locate the new malicious users, the sub-inspector will incur a new round of inspection among the users in set H.

We conclude the above strategies in **Algorithm** 2, which is termed as the Suspicion Assessment based Inspection (SAI) approach. According to the SAI algorithm, the users with high suspicions will be first probed individually. After this process, the users with low suspicions will be probed by a binary tree based inspection strategy.

B. Implementation

In this subsection, we explain how our method can be implemented in an NAN of a smart grid system in two aspects: 1) *algorithm* and 2) *hardware configuration*.

1) Algorithm: to implement the SAI algorithm, the most important thing is to assess users' suspicions that they steal electricity. It is achieved by assessing users' prior records in line with existing criminology knowledge as well as by comparing users' reported readings with their predicted normal readings, as shown in Equation (2). To a large extent, prior records and consumption deviations represent users' past behaviors and current behaviors, respectively. As aforementioned, once a user is identified as being malicious, utility companies note down the time period when this user is caught stealing electricity. This is how we obtain users' prior records. For a newly established NAN where all the users' prior records are empty, users' suspicions are only determined by the consumption deviation. Since smart meters are spontaneously generating data, by applying Holt-Winters method on data generated during periods when the head inspector does not detect reading anomalies, we can always obtain the predicted normal readings and then the consumption deviation. Another important aspect to implement the SAI algorithm is to estimate the number of malicious users in the NAN, i.e., $\tilde{m}$. This is because the value of $\tilde{m}$ determines when we should change the inspection strategy from individual inspection to binary tree based inspection, as shown in Fig. 4. $\tilde{m}$ can be approximately estimated as the average number of malicious users identified in the past when the head inspector detected reading anomalies. Specifically, it can be determined according to Equation (3).

2) Hardware Configuration: We did not change the topology of power line. We just installed an inspector box [16] at some places (e.g., on an electric pole or at a distribution room) in the NAN. An exemplary installation case of smart meters and an inspector box is shown in Fig. 5(a), which is drawn based upon a real photo of an apartment complex in Berkeley in [51]. As shown in Fig. 5(a), an array of thirty-six smart meters are mounted together on a wall. The array is arranged in four rows, each containing nine meters. To the right side of users' smart meters, a meter-like device is installed which can be regarded as an inspector box in the paper. Apparently, for getting electricity, users' smart meters can be connected with the inspector box individually. Furthermore, we observed that in many apartment buildings in USA, tenants' smart meters are installed together at some place near the apartments, with each meter connecting one individual apartment. Therefore, it is more than feasible to install an inspector box (working as a distribution automation device) near the place where the smart meters are installed.

To sum up, our model agrees with the topology of distribution network in a real power system. For better understanding, we draw Fig. 5(b) based upon a figure in [40] which depicts one typical (radial) distribution system. As shown in Fig. 5(b), in some rural areas, users are connected as a star network, i.e., in the way that we described in Fig. 1. In these cases, we can install the inspector box at the places like point A

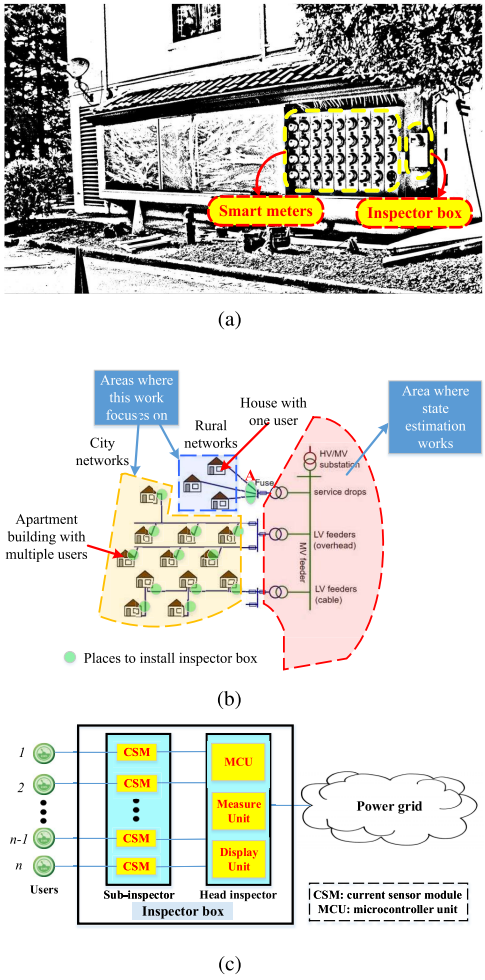


Fig. 5. Illustration of how to apply our schemes in real world. (a) An exemplary installation case of smart meters and the inspector box. (b) Locations to install the inspector boxes. (c) A possible block diagram of the inspector box.

in Fig. 5(b). In the city areas in which buildings usually consist of multiple apartments (as described in Fig. 5(a), we can install an inspector box at some places near the corresponding buildings. Fig. 5(a) also shows places where the state estimation approaches work.

As aforementioned, the inspector box contains one head inspector and at least one sub-inspector. The inspectors are essentially function-enhanced smart meters with stronger computation capability and larger storage space. The inspectors measure the total amount of electricity distributed to the users monitored by them. Head inspectors which monitor all the users in the NAN have already existed (maybe called different names, e.g., central observer meter, collector, substation level meter, etc.) and are widely deployed in the power systems, as indicated in the papers [18], [47], [48]. Different from the above papers, we have sub-inspectors in our paper which are functionally-identical with the head inspector. As explained in paper [16], the inspector box is able to automatically assign any user combination with any number to one sub-inspector, without interfering with any normal electricity services (i.e., without incurring power outages). The above

function can be achieved by carefully designing electrical circuits which connect the sub-inspectors.

We next demonstrate how an inspector box achieves the goal of assigning different groups of users to the sub-inspector without interrupting users' normal electricity consumptions. A possible (but not optimized) block diagram of an inspector box is shown in Fig. 5(c) in which the sub-inspector consists of multiple current sensor modules (CSMs), each being connected to a smart meter with an individual power line for measuring the smart meter's current. We expect that CSMs output digital signals such that these signals can be conveniently controlled in the microcontroller unit (MCU) of the head inspector. Otherwise, if a CSM outputs analog signals (e.g., ACS712 hall effect current sensors [52]), an analog-to-digital converter can be further employed. Specifically, if at a certain inspection step, the sub-inspector is signaled to inspect a group of users denoted by G, then the digital outputs of the CSMs connected to users in $U \setminus G$ are cancelled by setting them as zeros. Since residential electricity voltage in a certain country remains a standard value (e.g., 120V and 220V in USA and China, respectively), the MCU can easily calculate electricity consumptions of users in G using the unconcealed currents and the standard voltage. In other words, the sub-inspector measures only the electricity consumptions of the users in G effectively. Note that the head inspector also has a measure unit and a display unit. The measure unit measures the total electricity consumption of all users in the NAN.

VI. Experiment Results

This section reports experiment results. The experiments are conducted in Python 2.7.13 on an integrated development environment - the PyCharm Community Edition 2017.1.3. Users' actual electricity consumptions are generated based on the individual household electric power consumption data set in [45]. The data are measurements of electric power consumption in one household with a one-minute sampling rate over a period of almost four years. In the experiments, we set the reporting periods of users' smart meters as 15 minutes. Different users' actual electricity consumptions are proportional to the recorded individual household power consumption in [45], with the corresponding coefficients being randomly generated which distribute in the interval $[0, 2]$. The honest users report their electricity consumption genuinely. With regard to the malicious users, we assume the relationships between the actual and the reported electricity consumptions follow one as below: (1) $q'(i,t) = q(i,t) - c_0$; (2) $q'(i,t) = (1-c_1)\,q(i,t)$; and (3) $q'(i,t) = c_2$, where the constants c_0, c_1, c_2 satisfy $c_0 > 0$, $0 < c_1 < 1$, $c_2 < q(i,t)$, $\forall t \geqslant 1$. Note that the constants c_1, c_2, and c_3 maintain fixed for all simulation periods.

For users without prior records, the total number of electricity thefts is set $|R(i,t^*)| = 0$; and the time interval since the last electricity theft is set as $y(i,t^*) = 1000$ months. For users with prior records, $|R(i,t^*)|$ is randomly chosen between 1 and 10; and $y(i,t^*)$ is randomly chosen between 0 and 80 months. Let us define the random variable X_i by $X_i = 1$ if user i commits electricity theft at period t^* (where t^*, as defined earlier, is the period when the head inspector detects reading anomalies) and by $X_i = 0$ if user i does not. We assume that the random variable X_i follows the Bernoulli distribution [46] as below:

$$\begin{cases} \Pr(X_i = 1) \\ \quad = \dfrac{1}{2.5 + 7.5\exp(0.8|R(i,t^*)| - 0.2y(i,t^*))} + 0.05, \\ \Pr(X_i = 0) = 1 - \Pr(X_i = 1). \end{cases}$$

We assume this Bernoulli distribution because it is consistent with the criminology knowledge that the recidivism risk increases monotonically with the total number of prior criminal records and decreases monotonically with the time interval since the last criminal act. Note that the criminal act in our case is electricity theft.

In the experiments, for the parameters in Equation (1), we set $a = 1.25, b = 3.75, w = 0.8$ and $c(t) = 0.1$. For the parameters in Equation (2), we set $u = 0.4$ and $g = 3$. For the initial value m_0 in Equation (3), we set it as 10% of the total number of users in the NAN. Note that each piece of data in the following figures is averaged over 30 times of repeated experiments.

In Fig. 6, we define the prior record ratio as the ratio of the number of users with prior records to the number of all the users in the NAN. We explore how prior record ratios influence the performance of the SAI algorithm. For this purpose, we introduce a new metric[4], i.e., the inspection ratio, which is defined as the ratio between the number of inspection steps to the number of malicious users located. Obviously, in practice, a smaller inspection ratio implies a more effective inspection algorithm. As can be observed, on the whole, inspection ratios will be smaller when prior record ratios are larger. This demonstrate that the prior records do help us narrow down the searching area to a large extent, although the influence gets smaller during the inspection course.

In the following figures, we assume the prior record ratio as 50%.

In Fig. 7, we assume there are a total number of 250 users in the NAN. We present the first 30 experiment results. In Fig. 7, both the number of malicious users identified and the number of inspection steps performed by the sub-inspectors are involved. As we can see, when the number of malicious users gets larger, the number of inspection steps gets

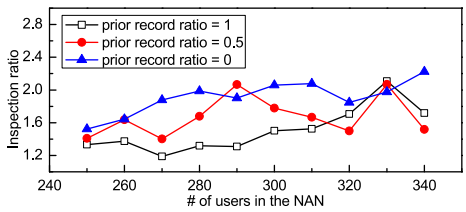


Fig. 6. Results under different prior record ratios. The inspection ratio is defined as the ratio between the number of inspection steps to the number of malicious users located.

[4]We here do not use the metric of the number of inspection steps as in papers [16], [17]. This is because the malicious users in this paper are generated based on a Bernoulli distribution. It implies that for a given total number of users in the NAN, the number of malicious users being generated/located may be different in several experiments.

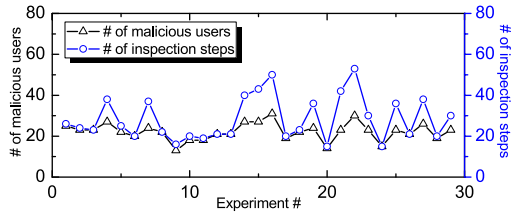


Fig. 7. Results of the first 30 experiments under the case $n = 250$.

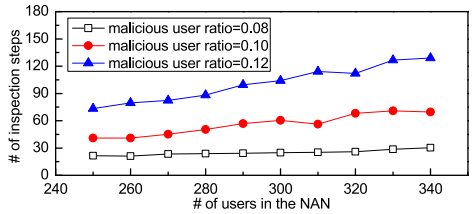


Fig. 8. Results under different malicious user ratios. The malicious user ratio is defined as the ratio of the number of malicious users to the number of all the users in the NAN.

larger correspondingly. Specially, if the number of malicious users increases suddenly, the number of inspection steps will accordingly reach a peak. This is because in these cases, more malicious users will be located through applying the binary tree based inspection strategy, which usually takes several inspection steps to locate a malicious user.

In Fig. 8, we display the experiment results when the malicious user ratio is $0.08, 0.10, 0.12$, respectively. Note that the malicious user ratio is defined as the ratio of the number of malicious users to the number of all the users in the NAN. We investigate how the number of inspection steps varies under different malicious user ratios. As we can see, for any given total number of users in the NAN, sub-inspectors perform the fewest inspection steps under the smallest malicious user ratio. When the ratio of malicious users is larger than 10%, the number of inspection steps has a tendency to increase. This is because we have set the initial value of the threshold as 10% of the total number of malicious users. When the total number of malicious users increases, more inspection steps will be performed during the binary tree based inspection process.

In Fig. 9, we explore the performance of the SAI algorithm in dynamic cases, in terms of the number of inspection steps. We show the results of the experiments where the ratio of malicious users is 0.1. We consider the cases where the number of new malicious users is $5, 10$ and 15, respectively. For comparison purposes, we also display the results of static cases, where the number of new malicious users is 0. As can be observed in Fig. 9, on the whole, for a given total number of users in the NAN, when the number of new malicious users is larger, the number of inspection steps will accordingly get larger.

In Fig. 10, we compare the SAI algorithm with the ATI algorithm [16] as well as the DCI algorithm [17] in static cases.[5] We assume that there are a total number of 250 users

[5]Since the inspection process in dynamic cases can be regarded as a repeated inspection process in static cases, the tendency of curves in dynamic cases is similar.

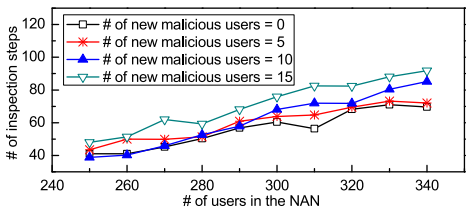


Fig. 9. Results under dynamic cases.

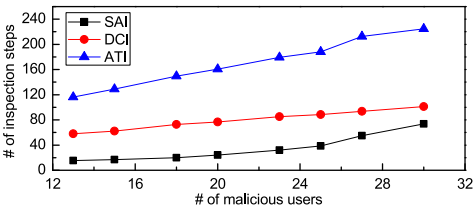


Fig. 10. Results of SAI vs. ATI under the case $n = 250$.

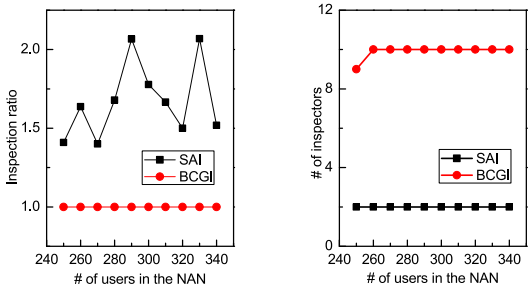


Fig. 11. SAI vs. BCGI: (a) inspection ratio; (b) number of inspectors.

in the NAN. In the real world, there will be few malicious users who start committing electricity theft simultaneously, i.e., at the same period t^*. Furthermore, it is pointed out in paper [16] that the ATI algorithm has better performance when the ratio of malicious users is small. Thus, in Fig. 10, we consider the number of malicious users varying from 12 to 30. As can be seen, the SAI algorithm obviously outperforms both the ATI algorithm and the DCI algorithm in terms of inspection steps. With the increase of malicious user number, the performance gap between the SAI algorithm and the DCI algorithm becomes smaller. Besides, under the condition that smart meters report users' electricity consumptions every 15 minutes, the investigation procedures of the SAI algorithm usually take hours. As shown in Fig. 10, for locating 13 malicious users from a total number of 250 users, the SAI algorithm takes about 15 inspection steps. This indicates that the inspection process lasts for less than four hours. With the number of malicious users increasing, the inspection process prolongs. As shown in the figure, when the number of malicious users reaches 30, by the SAI algorithm the sub-inspector takes approximately 60 inspection steps, which lasts for about 15 hours. However, we do not expect that we have so many malicious users in reality.

In Fig. 11, we compare the performance of the SAI algorithm with the BCGI algorithm, in terms of inspection ratio. As pointed out earlier, a smaller inspection ratio implies a more effective inspection algorithm. From Fig. 11(a), we can

observe that the BCGI is more effective than the SAI algorithm in locating malicious users. However, we cannot neglect the fact that the SAI algorithm needs only two inspectors - the head inspector and a sub-inspector, while the BCGI algorithm requires many more inspectors, as shown in Fig. 11(b). Furthermore, the application of the BCGI algorithm is very limited. More specifically, it can only be applied when there is only one malicious user in the NAN.

VII. Conclusion

This paper investigates the MMI problem which aims to detect the malicious users within the shortest detection time using a limited number of inspectors. Before inspectors conduct inspections, we assess users' suspicions to steal electricity: (1) Considering that electricity theft is a particular form of economic crime, we assess users' recidivism risks through analyzing electricity theft prior records; (2) We assess users' deviation risks based upon deviations between reported and predicted normal readings; (3) The suspicions are comprehensively assessed as a weighted value of recidivism risks and deviation risks. Note that with the inspection process going on, the deviation risks weigh more and more. On the basis of the suspicions, we further propose the SAI algorithm. According to it, the users with the highest suspicions will be individually inspected earlier. After this process, the remaining users will be inspected by a binary tree based inspection strategy. Experiment results show that the SAI algorithm outperforms both the ATI algorithm and the DCI algorithm in terms of the number of inspection steps. Although the SAI algorithm is less effective to locate malicious users than the BCGI algorithm, it has much wider application.

Acknowledgment

The authors would like to thank Dr. Charlene Lucky Coburn for comprehensive editing and mentorship in English writing. They thank for the reviewers for their comments which significantly helped in enhancing the quality of this paper.

References

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

[2] Z. Zhou, J. Bai, M. Dong, K. Ota, and S. Zhou, "Game-theoretical energy management design for smart cyber-physical power systems," *Cyber-Phys. Syst.*, vol. 1, no. 1, pp. 24–45, 2015.

[3] E. McCary and Y. Xiao, "Malicious device inspection home area network in smart grids," *Int. J. Sensor Netw.*, vol. 25, no. 1, pp. 45–62, 2017.

[4] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Electricity theft: Overview, issues, prevention and a smart meter based approach to control theft," *Energy Policy*, vol. 39, no. 2, pp. 1007–1015, Feb. 2011.

[5] Northeast Group, LLC. (2014). *World Loses $89.3 Billion to Electricity Theft Annually, $58.7 Billion in Emerging Markets*. [Online]. Available: http://www:prnewswire:com/news-releases/world-loses-893-billion-to-electricity-theft-annually-587-billion-in-emerging-markets-300006515:html

[6] J. Jow, Y. Xiao, and W. Han, "A survey of intrusion detection systems in smart grid," *Int. J. Sensor Netw.*, vol. 23, no. 3, pp. 170–186, 2017.

[7] D. Mashima and A. A. Cárdenas, "Evaluating electricity theft detectors in smart grid networks," in *Proc. Int. Workshop Recent Adv. Intrusion Detection (SAID)*, Amsterdam, The Netherlands, Sep. 2012, pp. 210–229.

[8] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and A. M. Mohammad, "Detection of abnormalities and electricity theft using genetic support vector machines," in *Proc. IEEE Region 10 Conf. (TENCON)*, Hyderabad, India, Nov. 2008, pp. 1–6.

[9] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, "Nontechnical loss detection for metered customers in power utility using support vector machines," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 1162–1171, Apr. 2010.

[10] A. H. Nizar, Z. Y. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 946–955, Aug. 2008.

[11] A. Jose, R. Malekian, and B. B. Letswamotse, "Improving smart home security; integrating behaviourprediction into smart home," *Int. J. Sensor Netw.*, vol. 28, no. 4, pp. 253–269, 2018.

[12] R. A. Nadi and M. G. H. A. Zamil, "A profile based data segmentation for in-home activity recognition," *Int. J. Sensor Netw.*, vol. 29, no. 1, pp. 28–37, 2019.

[13] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Sci. Technol.*, vol. 19, no. 2, pp. 105–120, Apr. 2014.

[14] Q. Liu, F. Chen, F. Chen, Z. Wu, X. Liu, and N. Linge, "Home appliances classification based on multi-feature using ELM," *Int. J. Sensor Netw.*, vol. 28, no. 1, 2018, pp. 34–42.

[15] Z. Xiao, Y. Xiao, and D. H. C. Du, "Non-repudiation in neighborhood area networks for smart grid," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 18–26, Jan. 2013.

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

[17] X. Xia, W. Liang, Y. Xiao, and M. Zheng, "Difference-comparison-based malicious meter inspection in neighborhood area networks in Smart Grid," *Comput. J.*, vol. 60, no. 12, pp. 1852–1870, Dec. 2017.

[18] W. Han and Y. Xiao, "NFD: Non-technical loss fraud detection in Smart Grid," *Comput. Secur.*, vol. 65, pp. 187–201, Mar. 2017.

[19] W. Han and Y. Xiao, "A novel detector to detect colluded non-technical loss frauds in smart grid," *Comput. Netw.*, vol. 117, pp. 19–31, Apr. 2017.

[20] W. Han and Y. Xiao, "Design a fast non-technical loss fraud detector for smart grid," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 5116–5132, Dec. 2016.

[21] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1319–1330, Jul. 2013.

[22] X. Xia, Y. Xiao, and W. Liang, "ABSI: An adaptive binary splitting algorithm for malicious meter inspection in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, Feb. 2019, pp. 445–458.

[23] B. Bat-Erdene, B. Lee, M.-Y. Kim, T. H. Ahn, and D. Kim, "Extended smart meters-based remote detection method for illegal electricity usage," *IET Gener., Transmiss. Distrib.*, vol. 7, no. 11, pp. 1332–1343, Nov. 2013.

[24] B. Bat-Erdene, S.-Y. Nam, and D.-H. Kim, "A novel remote detection method of illegal electricity usage based on smart resistance," *Future Inf. Technol.*, vol. 185, no. 26, pp. 214–223, Jan. 2011.

[25] C. Liu, S. Ghosal, Z. Jiang, and S. Sarkar, "An unsupervised anomaly detection approach using energy-based spatiotemporal graphical modeling," *Cyber-Phys. Syst.*, vol. 3, nos. 1–4, pp. 66–102, 2017.

[26] G. M. Messinis and N. D. Hatziargyriou, "Review of non-technical loss detection methods," *Electr. Power Syst. Res.*, vol. 158, pp. 250–266, May 2018.

[27] J. B. Leite and J. R. S. Mantovani, "Detecting and locating non-technical losses in modern distribution networks," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1023–1032, Mar. 2018.

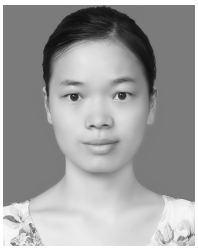
[28] S.-C. Huang, Y.-L. Lo, and C.-N. Lu, "Non-technical loss detection using state estimation and analysis of variance," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2959–2966, Aug. 2013.

[29] X. Xia, Y. Xiao, W. Liang, and M. Zheng, "Coded grouping-based inspection algorithms to detect malicious meters in neighborhood area smart grid," *Comput. Secur.*, vol. 77, pp. 547–564, Aug. 2018. doi: 10.1016/j.cose.2018.05.004.

[30] I. Hosni and N. Hamdi, "Distributed cooperative spectrum sensing with wireless sensor network cluster architecture for smart grid communications," *Int. J. Sensor Netw.*, vol. 24, no. 2, pp. 118–124, 2017.

[31] Lou Frenzel. (2014). *Smart Grid Neighborhood-Area Network Standard Charges Ahead*. [Online]. Available: https://www.electronicdesign.com/trends-amp-analysis/smart-grid-neighborhood-area-network-standard-charges-ahead

[32] Carolina Country. (2013). *Stealing Electricity—Another Way to Get Electrocuted or Land in Jail*. [Online]. Available: https://www.carolinacountry.com/your-energy/between-thelines/departments/between-the-lines/stealing-electricity
[33] K. A. Seger and D. J. Icove, "Power theft: The silent crime," *FBI Law Enforcement Bull.*, vol. 57, pp. 20–25, Mar. 1988.
[34] P. Schmidt and A. D. Witte, *Predicting Recidivism Using Survival Models*. New York, NY, USA: Springer-Verlag, 1988.
[35] M. C. Kurlychek, R. Brame, and S. D. Bushway, "Scarlet letters and recidivism: Does an old criminal record predict future offending?" *Criminology Public Policy*, vol. 5, no. 3, pp. 483–504, Aug. 2006.
[36] M. C. Kurlychek, R. Brame, and S. D. Bushway, "Enduring risk? Old criminal records and predictions of future criminal involvement," *Crime Delinquency*, vol. 53, no. 1, pp. 64–83, Jan. 2007.
[37] A. Blumstein and S. Moitra, "The identification of 'career criminals' from 'chronic offenders' in a cohort," *Law Policy*, vol. 2, no. 3, pp. 321–334, Jul. 1980.
[38] D. P. Farrington and R. Tarling, *Prediction in Criminology*. Albany, NY, USA: State Univ. of New York Press, 1985.
[39] A. Blumstein, D. P. Farrington, and S. Moitra, "Delinquency careers: Innocents, desisters, and persisters," *Crime Justice*, vol. 6, pp. 187–219, Jan. 1985.
[40] Electricaleasy.com. (2019). *Radial, Parallel, Ring Main And Interconnected Distribution Systems*. [Online]. Available: https://www.electricaleasy.com/2018/02/radial-parallel-ring-main-interconneted-distribution.html
[41] P. R. Mahalingam and S. Vivek, "Predicting financial savings decisions using sigmoid function and information gain ratio," *Procedia Comput. Sci.*, vol. 93, pp. 19–25, Sep. 2016.
[42] B. Dong, Z. Li, S. M. M. Rahman, and R. Vega, "A hybrid model approach for forecasting future residential electricity consumption," *Energy Buildings*, vol. 117, pp. 341–351, Apr. 2016.
[43] P. S. Kalekar, "Time series forecasting using Holt-Winters exponential smoothing," Kanwal Rekhi School Inf. Technol., Tech. Rep., 2004. Accessed: Mar. 2018. [Online]. Available: https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/
[44] W. King. (1984). *Utilities Say 1% of Users are Stealing Power*. [Online]. Available: http://www:nytimes:com/1984/03/26/us/utilities-say-1-of-users-are-stealing-power:html
[45] UCI Machine Learning Repository. (2012). *Individual Household Electric Power Consumption Data Set*. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption/
[46] V. Krishnan, *Probability and Random Processes*. Hoboken, NJ, USA: Wiley, 2006.
[47] S. Salinas, M. Li, and P. Li, "Privacy-preserving energy theft detection in smart grids: A P2P computing approach," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 257–267, Sep. 2013.
[48] Y. Liu and S. Hu, "Cyberthreat analysis and detection for energy theft in social networking of smart homes," *IEEE Trans. Comput. Social Syst.*, vol. 2, no. 4, pp. 148–158, Dec. 2015.
[49] C. Chatfield and M. Yar, "Holt-winters forecasting: Some practical issues," *J. Roy. Stat. Soc. D (Statistician)*, vol. 37, no. 2, pp. 129–141, 1988.
[50] Wikipedia. (2017). *Exponential Smoothing*. [Online]. Available: https://en.wikipedia.org/wiki/Exponential_smoothing
[51] Josh Hart. (2012). *Smart Meters Still Being Forced On Us-Speak Out Thursday*. [Online]. Available: https://stopsmartmeters.org/2012/06/19/smart-meters-still-being-forced-on-us-speak-out-thursday/
[52] Allegro MicroSystems, LLC. (2019). *ACS712: Fully Integrated, Hall-Effect-Based Linear Current Sensor IC with 2.1 kVRMS Voltage Isolation and a Low-Resistance Current Conductor*. [Online]. Available: https://www.allegromicro.com/en/Products/Current-Sensor-ICs/Zero-To-Fifty-Amp-Integrated-Conductor-Sensor-ICs/ACS712.aspx
[53] P. S. N. Rao and R. Deekshit, "Energy loss estimation in distribution feeders," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1092–1100, Jul. 2006.



Xiaofang Xia received the B.E. degree from Xiangtan University, China, in 2012, and the Ph.D. degree in control theory and control engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, China, in 2019. She was a Visiting Student with the Department of Computer Science, University of Alabama, USA, from 2016 to 2018. She is currently a Post-Doctoral Faculty with the School of Computer Science and Technology, Xidian University, China. Her research interests are mainly in communication networks, cyber physical system, smart grid security, and privacy preserving.

Yang Xiao is currently a Professor with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL, USA. He has published over 200 journal papers and over 200 conference papers. His current research interests include cyber-physical systems, the Internet of Things, security, wired/wireless networks, smart grid, and telemedicine. He was a Voting Member of the IEEE 802.11 Working Group from 2001 to 2004, involving the IEEE 802.11 (WIFI) standardization work.

Wei Liang received the Ph.D. degree in mechatronic engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, in 2002. As a Primary Participant/a Project Leader, she developed the WIA-PA and WIA-FA standards for industrial wireless networks, which are specified by IEC 62601 and IEC 62948, respectively. She is currently a Professor with the Shenyang Institute of Automation, Chinese Academy of Sciences. Her research interests include industrial wireless sensor networks and wireless body area networks. She received the International Electrotechnical Commission 1906 Award in 2015 as a Distinguished Expert of industrial wireless network technology and standard.