CONTRIBUTED PAPER

# Detection Methods in Smart Meters for Electricity Thefts: A Survey

*This article surveys the electricity theft issue and the existing detection methods and provides insight for shaping future research directions.*

By Xiaofang Xia, *Member IEEE*, Yang Xiao, *Fellow IEEE*, Wei Liang, *Senior Member IEEE*, and Jiangtao Cui, *Member IEEE*

**ABSTRACT** | For accommodating rapidly increasing power demands, power systems are transitioning from analog systems to systems with increasing digital control and communications. Although this modernization brings many far-reaching benefits, the hardware and software newly incorporated into the power systems also incur many vulnerabilities. By taking advantage of these vulnerabilities, adversaries can launch various cyber/physical attacks to tamper with electricity meter readings, i.e., to steal electricity. It is reported that total worldwide annual economic losses caused by electricity theft reached up to almost one hundred billion dollars in recent years. With methods to tamper with meter readings becoming more versatile, secret, and flexible, electricity theft tends to get even more serious in modernized power systems. For preventing adversaries from stealing electricity, researchers have done a lot of works. Although some related surveys on these works exist, they are not updated or just discuss electricity theft in a specific region. This survey aims to gain a comprehensive and in-depth understanding of the electricity theft issue. After investigating how adversaries tamper with meter readings, we systematically survey all existing detection methods up to date, which is classified into machine learning- and measurement mismatch-based methods. Adverse effects and political and socioeconomic factors of electricity theft are also provided. This survey can help relevant researchers to shape future research directions, especially in the area of developing new effective electricity theft detection methods.

**KEYWORDS** | Binary trees; cyber–physical systems; detection methods; electricity theft; Internet of Things (IoT); machine learning; measurement mismatch; security; smart grid; smart meters.

Manuscript received January 5, 2021; revised November 24, 2021; accepted December 21, 2021. Date of publication January 19, 2022; date of current version February 3, 2022. The work of Xiaofang Xia, Wei Liang, and Jiangtao Cui was supported in part by the Special Fund for Strategic Pilot Technology of Chinese Academy of Sciences under Grant XDC02020600; in part by the National Key Research and Development Program under Grant 2019YFB1707401 and Grant 2021YFB3301001; in part by the National Natural Science Foundation of China (NSFC) under Grant 61902299, Grant 61976168, and Grant 62022088; in part by the Liaoning Provincial Natural Science Foundation of China under Grant 2020JH2/10500002 and Grant 2019-YQ-09; in part by the Liaoning Revitalization Talents Program under Grant XLYC1902110; in part by the Natural Science Basic Research Program of Shaanxi Province under Grant 2019CGXNG-023; in part by the S&T Program of Hebei under Grant 20310102D; in part by the International Partnership Program of the Chinese Academy of Sciences under Grant 173321KYSB20200002; and in part by the China Postdoctoral Science Foundation under Grant 2019TQ0239 and Grant 2019M663636. *(Corresponding authors: Yang Xiao; Jiangtao Cui.)*

**Xiaofang Xia** is with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China, and also with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA (e-mail: xiaofangxia89@gmail.com).

**Yang Xiao** is with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA (e-mail: yangxiao@ieee.org).

**Wei Liang** is with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China, with the Key Laboratory of Networked Control System, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China, and also with the Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China (e-mail: weiliang@sia.cn).

**Jiangtao Cui** is with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China (e-mail: cuijt@xidian.edu.cn).

Digital Object Identifier 10.1109/JPROC.2021.3139754

## I. INTRODUCTION

Nowadays, people around the world are witnessing more and more power outages whose frequency and duration rise constantly [1]. This implies that antiquated electric power systems, which were built over one century ago, are having more difficulties in accommodating our rapidly increasing power demands [2]. Thus, many countries, such as the USA, Japan, and China, are currently making every effort to upgrade their existing

power systems. For example, at the supply side of power systems, we not only use traditional energy sources, such as coal or hydro, to generate electricity in large and centralized power plants that are usually located far away from customers' premises, but also apply renewable and environment-friendly energy resources, such as solar and wind near electricity-consumption locations. On the other hand, various new types of loads, such as electric vehicles and heat pumps, are connected to power systems at the demand side.

To provide reliable and high-quality electricity services, a balance between supply and demand needs to be achieved. This is mainly accomplished by incorporating modern information and communication technologies (ICTs) into power systems. Given the fact that nearly 90% of all power outages and disturbances have their roots in distribution networks, revolution toward modern power systems starts from distribution networks of power systems [3]. To be more specific, it begins with the metering infrastructure of distribution systems. As well known, customers' electricity consumption is first measured by electromechanical meters. At that time, utility companies have to send personnel to check meter readings door-to-door monthly for calculating customers' billings. Then, by introducing automated meter reading (AMR) systems into distribution networks, utility companies can automatically get metering data, such as consumption records, alarms, and status from customers' premises remotely. However, due to its one-way communication system, the AMR system does not allow utility companies to take corrective actions based on the received information.

To solve the issue of demand-side management, advanced metering infrastructure (AMI), which provides two-way communications between customers and utility companies, is further incorporated into power systems. Through AMI, utility companies can get instantaneous information about individual and aggregated demand and can enact various revenue models to control their costs. AMI is the milestone that traditional power systems are transitioning into smart grids where pervasive control at all levels is a basic premise. Apparently, during the process of modernizing power grids, communication and control systems of power grids are transitioning from analog systems where a handful of control parts are mainly deployed at central stations to systems with increasing digital control and communications where potentially millions of control points are deployed to almost every level of power systems.

Although the modernized power systems bring enormous and far-reaching technical and social benefits, they also have a variety of vulnerabilities [4]. The vulnerabilities may be inherited from existing ICTs, incurred by poor accommodation of emerging hardware and communication technologies, or caused by insufficient standards and regulations in modernized power systems [5]. By taking advantage of these vulnerabilities, adversaries can launch a variety of cyber–physical attacks against widely deployed device components and/or heterogeneous network components in modernized power systems [6]. One of these adversaries' main purposes is to manipulate their electricity consumptions into smaller values such that their billings can be lowered down, i.e., to achieve the goal of stealing electricity [7].

As a notorious phenomenon, electricity theft has harassed almost all utility companies around the globe for a long time. One of the earliest electricity theft events was documented in 1886, which was accomplished by unscrupulous persons tapping into Edison Electricity in New York [8]. In traditional power systems, adversaries usually leverage physical attacks, such as bypassing meters and directly hooking from power lines, to steal electricity. In contrast, after ICTs are introduced into power systems (especially in smart grids), besides these physical attacks, adversaries can also launch cyberattacks (such as spoofing attacks and man-in-the-middle attacks) to tamper with readings of electronic meters in both AMR and AMI. Since these electronic meters are usually equipped with low tamper-resistant components and are even based upon system-on-chip designs with security back doors [9], adversaries can launch these cyberattacks to falsify meter readings almost anywhere and anytime. The Federal Bureau of Investigation (FBI) reports that adversaries can hack into electronic meters with just a moderate level of computer knowledge, using low-cost tools and software readily available on the Internet [10]. Physical attacks are more common in old days or underdeveloped regions, and cyberattacks are more common in modernized power systems, such as smart grids. An attacker normally finds an achievable and convenient stealing method based on the current system setup.

Electricity theft has many negative effects, among which huge economic losses drive utility companies to figure out ways to reduce electricity theft greatly. Total worldwide revenue losses were total about $89.3 billion in 2014 [11], which rose to $96 billion in 2017 [12]. In Table 1, we summarize recent statistics regarding ratios of stolen electricity to total power generation and revenue losses in several countries. The statistics in Table 1 indicate that electricity theft is much more serious in developing countries than in developed countries [13]. Particularly, India loses the most money among all countries in the world. Since utility companies cannot afford the enormous economic losses alone, they usually pass on these losses to all customers via higher tariffs. It is reported that each customer in the U.K. has to pay extra €30 for electricity theft [14].

In addition, electricity theft induces utility companies to underestimate customers' power demands. When electricity theft occurs, adversaries report lower electricity consumption than what is consumed. The incorrect information would mislead utility companies to generate less electricity than what is demanded, resulting in power quality degradation. As reported, in regions where electricity theft is pervasive (e.g., India), customers often experience voltage sags and intermittent power disruptions, especially during peak load periods [24]. The power disruptions

**Table 1** Sample Electricity Theft Statistics

| Country | % power stolen | Revenue losses |
|---|---|---|
| USA [15] | 0.5% $\sim$ 3.5% | \$1 $\sim$ 10B |
| India [11][16] | 30% | \$16.2B |
| South Africa [17] | 33% | 20B Rand (\$1.5B) |
| Netherlands [18] | 23% | €114M(\$123.49M) |
| Brazil [18] [19] | 20 $\sim$ 30% | 8B reais(\$3.7B) |
| Bangladesh [15] | 14% | 396B TK(\$50.86M) |
| Malaysia [20] | 20% | \$229M |
| Turkish[21] | 15% | \$1B |
| Jamaica[22] | 18% | \$46M |
| Canada [23] | – | 100M CAD |

harm and even damage electrical appliances at homes and halt the normal production process in firms. Besides, due to the huge revenue losses, utility companies have to decrease the investment in advanced equipment and technologies. This slows down the expansion of generation capacity and the development of smart grids. Furthermore, some electricity theft behaviors, such as illegal tapping, raise safety concerns, such as electric shocks and even casualties. Under extreme weather conditions, these behaviors may cause wires to start sparking and even cause fire disasters, which threaten the whole community.

In many countries, electricity theft is characterized as a special form of economic crime [25], and stringent laws are issued to punish adversaries by fines and/or incarceration. For example, the Theft Act 1968 in the U.K. says that persons who dishonestly use electricity are liable to imprisonment for a term not exceeding five years; and the Electricity Act 2003 in India says that adversaries should be imprisoned from six months to five years and imposed a fine not less than the financial gain from electricity theft. However, it seems that these laws alone cannot refrain customers from stealing electricity effectively. Driven by immediate and profitable economic benefits, adversaries committing electricity theft can still be found in almost every region worldwide. These adversaries can be roughly categorized into: 1) principal adversaries, which include residential, business, and government customers whose goal is to decrease their electricity bills and 2) accomplice adversaries, including organized hackers and utility insiders, whose goal is to profit from principal adversaries by providing tools/services.

Since electricity theft has so many adverse effects and laws/regulations are not powerful enough to prevent its occurrences, researchers develop various electricity theft detection methods. These methods can be roughly classified into machine learning- and measurement mismatch-based methods. The machine learning-based methods [26] apply currently popular machine learning methods to analyze meter readings and/or other customer-related data, aiming to find abnormal electricity consumption patterns that are highly related to electricity theft. In contrast, the measurement mismatch-based methods require either to deploy advanced sensors or to utilize existing advanced sensors in distribution networks. By analyzing measurements from sensors and readings of customers' smart meters, researchers can constantly narrow down the search zone of adversaries until finally pinpoint them. Some of these works have been surveyed in several reviews [27]–[29]. However, the paper [27] only surveys research up to 2013. Since numerous new electricity theft detection methods have been proposed during recent years, the paper [27] is obviously out of date. The paper [28] mainly focuses on the electricity theft issues in India, and hence, it is not general enough to deal with electricity theft issues across the world. The paper [29] has the shortcoming that analyses about state-of-the-art electricity theft detection methods are too simple. The limitation of the review paper [30] is that it considers only some machine learning-based electricity theft detection methods. We compare our review with the above reviews in Table 2, which shows that our review has more comprehensive coverage and depth than the above reviews.

To address the above limitations of the existing reviews, in this article, we aim to provide a comprehensive and in-depth understanding of the electricity theft issue. We achieve this goal by answering the following questions: 1) why do we need to address electricity theft? 2) how do adversaries usually commit electricity theft? 3) how do we currently combat electricity theft? and 4) what are the reasons why adversaries steal electricity? We answer the first question by analyzing the adverse effects of electricity theft (as presented earlier). To answer the second question, we first analyze the working principles of different types of electricity meters and then summarize methods to tamper with meter readings. To answer the third question, we systematically survey both existing machine learning- and measurement mismatch-based electricity theft detection methods. For better understanding, we first present an overall workflow of each category of detection methods and then dive into more detailed analyses for each branch of both categories. We answer the fourth question by investigating political and social–economical factors impacting electricity theft, based upon which practical institutional policy recommendations are summarized. Our contributions are highlighted as follows.

1) First, by analyzing the adverse effects of electricity theft, we analyze the motivations of utility companies in detail combating electricity theft.
2) Second, by investigating how meter readings are tampered with, we know more about adversaries. By systematically surveying existing detection methods, we know how these methods work and their strengths, as well as drawbacks. These two aspects of information contribute greatly to designing more effective detection methods in the future and help us a lot to win the battle finally.
3) Third, with political and social–economical factors impacting electricity theft, policy-makers can

**Table 2** Comparison Between Our Review and Other Existing Reviews

| Coverage | Our review | Review [31] | Review [30] | Review [29] | Review [28] |
|---|---|---|---|---|---|
| Impacting factors | Both political and socio-economical factors are analyzed in detail. | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| Adverse effects | Economical losses, power quality degradation and safety concerns are summarized in detail. | Some economical losses and power quality degradation are simply mentioned. | Some economical losses are briefly mentioned. | Some economical losses are briefly mentioned. | Some economical losses are briefly mentioned. |
| Electricity theft methods | Analyzed in detail from both technical and data-levels. | Not mentioned | Some methods like billing irregularities and unpaid bills are simply mentioned | Not mentioned | Analyzed from technical levels; But not as detaily as ours. |
| Detection methods | Machine-learning based& measurement -mismatch based methods; Analyzed deeply; *Papers up to 2021* | Data oriented & network oriented methods; Analyzed simply; *Papers up to 2016* | Only machine learning-based methods are analyzed; *Papers up to 2016* | Not mentioned | Classification based & state based methods; Most analyses are simple; *Papers up to 2013* |
| Policy recom -mendations | Provided | Not provided | Not provided | Provided; but only for India | Not provided |

formulate more targeted institutional policies to refrain from electricity theft.

The remainder of this article is organized as follows. In Section II, factors impacting electricity theft are analyzed. In Section III, we discuss how adversaries tamper with readings of electric meters. We analyze machine learning and the measurement mismatch-based electricity theft detection methods in Sections IV and V, respectively. We compare the above two categories of detection methods in Section VI. We provide some future work directions in Section VII. We conclude this article in Section VIII. For enhancing the readability of this article, we show the entire structure of this article in Fig. 1.

## II. FACTORS IMPACTING ELECTRICITY THEFT

As aforementioned in Section I, in many countries, electricity theft is characterized as a special form of economic crimes [25]. On the one hand, stringent laws are issued to punish adversaries stealing electricity by fines and/or incarceration, as summarized in Table 3. On the other hand, researchers have developed a lot of electricity theft detection methods to prevent users from stealing electricity. However, there are still a lot of adversaries stealing electricity across the world, especially in developing countries. In this section, we analyze the political and socioeconomic factors behind electricity theft. Based upon existing works [21], [24], [31]–[39], we summarize these factors in Fig. 2. We next demonstrate these political factors and socioeconomic factors, respectively.

### A. Political Factors

Political factors also called governance factors are usually closely related to the government. They mainly include voice and accountability, political instability, government effectiveness, regulatory burden, the rule of law, corruption, tax-GDP (gross domestic product) ratio, and collection efficiency, as explained in Table 4.

By analyzing sample data of 102 countries from the World Bank, Smith [31] examines correlations between

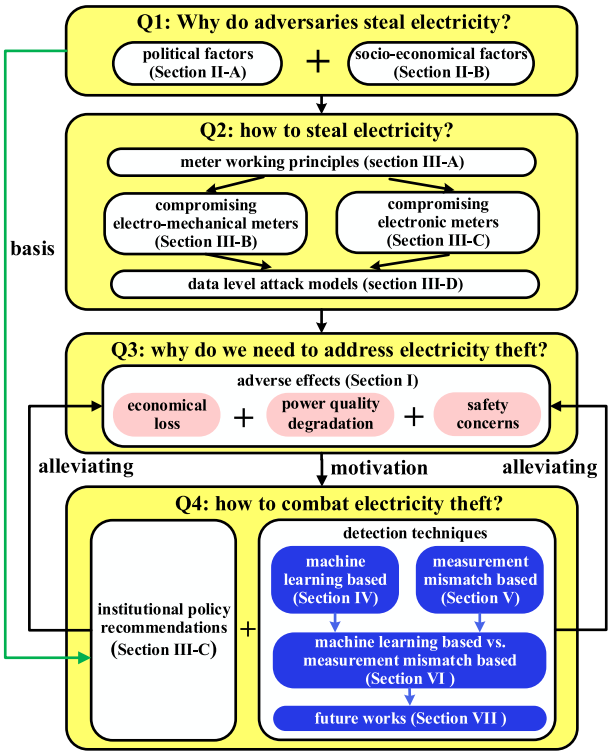


**Fig. 1.** *Entire structure of this article.*

**Table 3** Laws Relevant to Electricity Theft

| Countries | Laws | Punishment |
|---|---|---|
| UK | Theft Act 1968 | Imprisonment not exceeding five years [40] |
| India | The Electricity Act 2003 | Imprisonment from six months to five years; or/and fine not less than three times the financial gain by electricity theft [41] |
| China | Electricity Law | A fine of up to five times the amount of the electricity fees that should be paid or prosecuted for criminal liability |
| Pakistan | Pakistan Penal Code Electricity Theft Amendment 2016 | Imprisonment up to three years or/and with fines up to ten million rupees [42] |
| Turkey | Criminal Code | Imprisonment from 1 to 5 years [43] |
| Algeria | Code Penal Art. 350 | Imprisonment from 1 to 5 years and fines from 500 to 20,000 dinars [44] |
| Nigeria | Nigerian Electricity Regulatory Commission | Fines from N50,000 to N200,000 [45] |

the first six political factors and power losses during the transmission and distribution, of which electricity theft is the main cause. The analysis results imply the following.

1) There are usually more electricity theft events in countries with poorer civil rights, democratic institutions, and accountability.
2) When the political environment remains unstable, electricity theft tends to become severe. This is consistent with the fact that electricity theft in India is much more severe in election years than in other years [39].
3) Electricity theft occurs a lot when government effectiveness or the rule of law is weak or when regulatory burden is heavy.
4) Electricity thefts are closely related to corruption. In fact, in developing countries, such as India and Pakistan, corruption is a crucial determinant of electricity theft [24], [38], [39], [49].

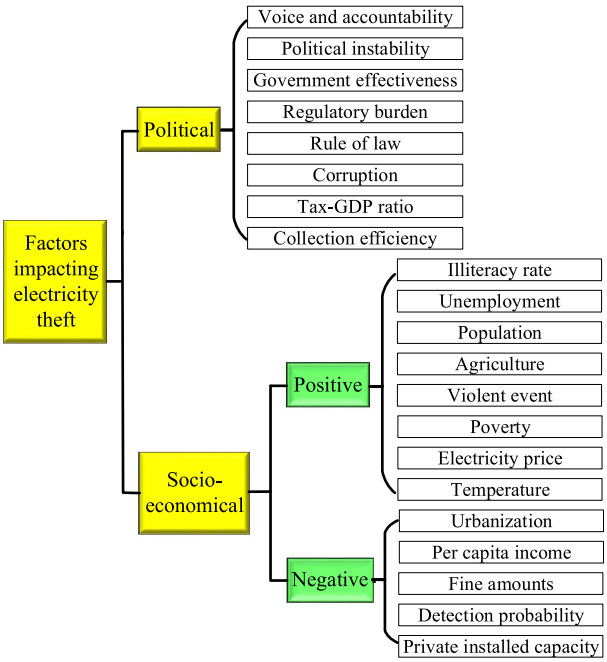


**Fig. 2.** *Factors impacting electricity theft [21], [24], [31]–[39].*

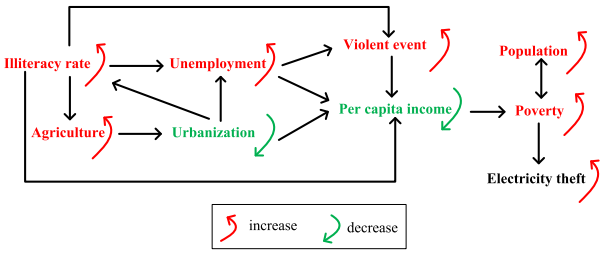


**Fig. 3.** *Relationship among several socioeconomic factors, among which illiteracy rate, unemployment, terrorist event, agriculture, population, and poverty are positive factors, and urbanization and per capita income are negative factors of electricity theft [21], [24], [33]–[38], [50], [51].*

Drawing data from 28 states of India over five years, Gaur and Gupta [24] investigate the determinants of electricity theft, among which the tax-GDP ratio and the collection efficiency are included. The tax-GDP ratio indicates people's honesty in revealing their true incomes and paying their obligations. On average, when the tax-GDP ratio and collection efficiency increase by 1%, electricity losses in transmission and distribution decrease by about 1.2% and 0.2 units, respectively [24].

## B. Socioeconomic Factors

As shown in Fig. 2, socioeconomic factors can be classified into negative factors and positive factors, which alleviate and deteriorate electricity theft, respectively [21], [24], [33]–[38], [50], [51]. Negative factors mainly include per capita income, urbanization, private installed capacity, detection probability, and fine amounts. Positive factors mainly include illiteracy rate, unemployment, population, poverty, violent event, agriculture, electricity price, and temperature. Socioeconomic factors are often related to each other and influence each other [24], [33], [36], [37], [50], [51]. These factors impact electricity theft, more or less. For example, since per capita income impacts electricity theft negatively [21], [24], [35], [38], electricity theft deteriorates with higher illiteracy rates, higher unemployment, more violent events, a greater share of agriculture, larger population, and more serious poverty, but alleviates with larger urbanization, as shown in Fig. 3.

We next explain how electricity prices, temperature, private installed capacity (which measures how much electricity is generated at the user's household), detection

**Table 4** Governance Factors Influencing With Electricity Theft

| Governance factors | Explanation |
|---|---|
| voice and accountability | It captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media [46]. |
| political instability | It is defined as the propensity of member changes in the executive government [47], usually achieved by democratic elections, violences and other constitutional or unconstitutional means. |
| government effectiveness | It mainly encompasses the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies [48]. |
| regulatory burden | It mainly involve the incidence of unfriendly market policies such as price controls, and perceptions of burdens imposed by excessive regulation [31]. |
| rule of law | It primarily refers to the influence and authority of laws within society, particularly as a constraint upon behaviors, including government officials' behaviors. |
| corruption | It means the unscrupulous use of public power for personal gains and is closely related to bribery. |
| tax-GDP ratio | It is the ratio of tax collected to national gross domestic product (GDP). It reflects the effectiveness of the respective government in collecting these taxes through proper enforcement mechanisms for ensuring public compliance. [24] |
| collection efficiency | It is the percentage of electricity bills that utility companies are able to recover [24]. It is an important indicator of the governance and management at the utility level [24]. |

probability, and fine amounts impact electricity theft. Economic theory suggests that crime should be committed only if the gain from offense exceeds the expected punishment cost. In our case, the crime is electricity theft. The studies in [21], [24], [35], and [38] indicate that electricity price and temperature positively affect electricity theft and private installed capacity, detection probability, and fine amounts impact electricity theft negatively. Specifically, customers are more likely to commit electricity theft under a higher electricity price, and a greater difference between the outdoor temperature and the assumed comfortable temperature [21], [35]. Customers are less likely to steal electricity if they have the more private installed capacity and, hence, not so dependent on the electricity delivered [24]. When the detection probability is low, and the amount of fine is small, customers are more likely to steal electricity because the expected punishment cost is low [38].

## C. Policy Recommendations

In line with the above political and socioeconomic factors, some targeted institutional policy recommendations for alleviating electricity theft are summarized as follows, and they may be applied to some countries.

First, since electricity theft is closely related to corruption, the following policies are recommended.

1) The government should encourage the privatization of electricity distribution companies since bribe is found more prevalent in public-owned utility companies [24], [34], [35].
2) Competitive strategies should be introduced among different distribution companies since bribe is negatively related to the level of competition [35].
3) Utility companies should organize training regularly to improve employees' professional ethics and to remind employees that it is illegal to accept bribes from customers.
4) An anonymous or nominal reporting mechanism should be applied such that customers aware of electricity theft behaviors can report the information to utility companies [32]. Some rewards can be provided to customers for encouragement after the adversaries or corrupted employees are finally checked.
5) Regulations/laws clearly describing punishments of bribery and corruption should be publicly issued [32].

Second, since electricity theft is much more severe during electoral periods, power company officials are suggested to be sheltered from political influence [32], [39]. By doing this, incumbent legislators cannot give illegal special priorities to particular categories of customers or supply more power to them than being allocated during election years in exchange for more votes [39].

Third, greater efforts must be made to brand power theft as a devastating economic crime that is strictly punished rather than a socially acceptable norm [34]. For the governments that do not have special laws exclusively for electricity theft, the relevant legislation process should be accelerated. For the governments that already have such laws, the crackdown on offenders stealing electricity should be strengthened.

Fourth, since illiteracy rates have a positive relationship with electricity theft, the government should increase its educational investment to meet broader social and economic objectives [21], [24]. Particularly, citizens should be educated about the disadvantages of electricity theft and the great toll it takes on the economy [21].

Fifth, since electricity theft is closely linked with poverty, the following policies are recommended.

1) The government should provide low-income families with special subsidies that are used to pay utility bills exclusively [31], [38]. This can greatly reduce the number of adversaries who commit electricity theft just because they cannot afford the power to meet their basic needs.
2) Utility companies should adopt a block rate tariff that charges customers a different price depending on how much electricity they have used [21]. According to

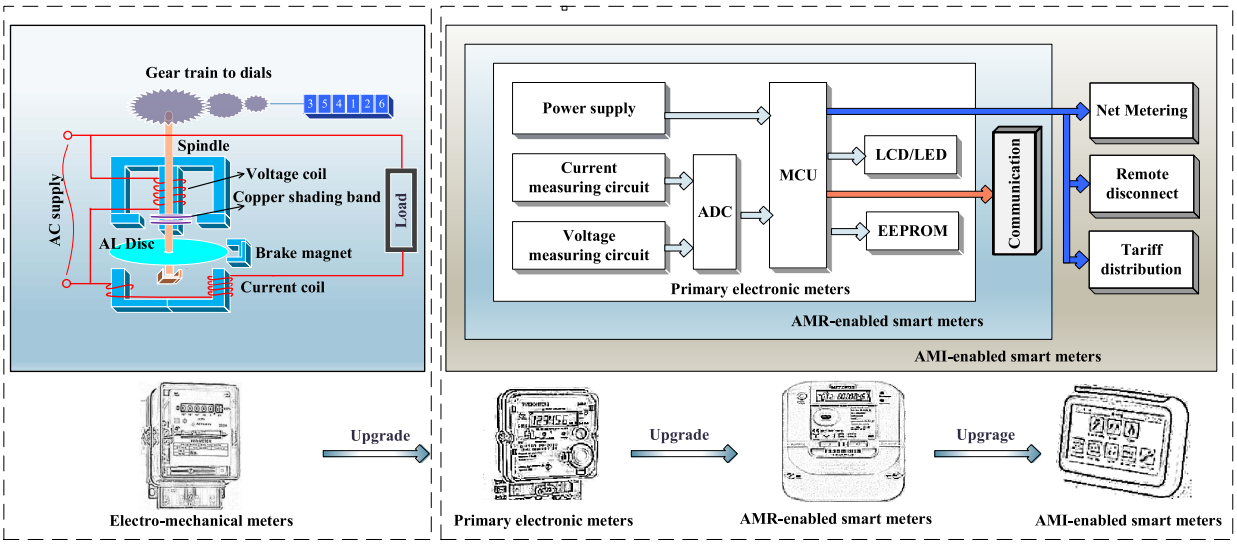


**Fig. 4.** *Structure of different kinds of electricity meters [52]–[54].*

international standards, the first 100 kWh of electricity per month is used for maintaining basic life necessities [21]. This block should be offered with very low prices or even free of charge [21]. The losses in revenue can be recovered by charging the following blocks with higher tariffs to meet operational maintenance and investment expenditures [21].

Sixth, regional tariffs should be adopted [21]. In other words, when deciding tariffs, policymakers should consider the region's economic conditions, and electricity prices should be set lower in regions with less per capita incomes [21]. This is inspired by the fact that electricity theft occurs more frequently in poorer districts than in more industrialized districts [21].

To sum up, both political and socioeconomic factors influencing electricity theft are heterogeneous. Statuses of politics, economy, and legislation may vary a lot across different countries. Therefore, the main political/socioeconomic factors influencing electricity theft differ in different countries. For example, the difference between developing and developed countries may be huge. In this article, we discuss electricity theft across the whole world instead of specific countries/regions. Thus, the institutional policy recommendations provided in this section may be a little general. In applications, different countries and even different utility companies in one country/region should choose appropriate institutional policies customized to themselves based upon their own political and socioeconomic developing statuses.

## III. ELECTRICITY THEFT METHODS

In this section, we review common electricity theft methods. As aforementioned, adversaries usually steal electricity by manipulating meter readings into smaller values. Methods for tampering electricity meters with different types vary a lot. Thus, in this section, we explain the working principles of electricity meters and then demonstrate how adversaries steal electricity for meters of different types.

### A. Meter Working Principles

Electricity meters keep evolving with technological advancements. In line with display types (analog or digital), electricity meters can be roughly classified into electromechanical and electronic meters.

The structure of electromechanical meters is given at the left-hand side of Fig. 4. As shown, an electromechanical meter consists of two silicon steel laminated electromagnets, which are called shunt magnet and series magnet, respectively. The shunt magnet carries a highly inductive voltage coil connected across the supply and some copper shading bands on the middle limb. The series magnet carries two current coils with a few turns connected in series with the load. The voltage coil and current coils produce two time-varying (sinusoidal) fluxes, respectively, which lag at 90° due to the inductive nature and the calibration of the copper shading bands. The two fluxes further induce eddy currents in the aluminum disk, which interacts with the magnetic fields of the two laminated electromagnets, exerting a driving torque in the disk to make it rotate. On the other hand, a braking torque on the disk is produced by the interaction between eddy currents and the magnetic field of a brake magnet (which is a kind of permanent magnet) over one side of the disk. Whenever the braking and driving torques become equal, the speed of the disk becomes steady, which is proportional to the power consumed by the load. A shaft or vertical spindle of the aluminum disk is associated with the gear arrangement that records a number proportional to the revolutions of

**Table 5** Comparison of Different Kinds of Electricity Meters

| Electricity meters | Electro-mechanical | Primary electronic | Smart meters | |
|---|---|---|---|---|
| | | | AMR-enabled | AMI-enabled |
| power flows [53] | unidirectional | unidirectional | unidirectional | bidirectional |
| communication flows [53] | × | × | unidirectional | bidirectional |
| net metering [53, 54] | × | × | × | ✓ |
| remote disconnect [53] | × | × | × | ✓ |
| tariff distribution [53] | × | × | × | ✓ |
| display [53] | mechanical dial | LCD/LED | | |
| accuracy class standards [55–57] | IEC 62053-11, ANSI C12.1 | IEC 62053-21, IEC 62053-22, ANSI C12.20 | | |
| measurement [53, 54] | kwh | kwh, MDI | kwh, MDI, voltage, current, power factors | kwh, MDI, voltage, current, power factors, net metering |
| reading frequency [53] | manual, monthly | manual, monthly | automatic, monthly | automatic, on demand (usually every 15min) |

the disk. This gear arrangement sets the number in a series of dials and indicates energy consumed over time.

The structure of electronic meters is depicted at the right-hand side of Fig. 4. As can be seen, electronic meters have three generations. The first generation is called primary electronic meters, which consists of current and voltage measuring circuits connected in serial and parallel with loads, respectively. Current transformers are usually applied for measuring the current due to its economic price and efficiency. For measuring the voltage, we usually apply a voltage divider in the case of 220 V but a potential transformer under the case of high voltage (like more than 500 V) to isolate the sensitive circuits. Analog signals from current and voltage measuring circuits are sampled and converted into digital values using an analog-to-digital converter (ADC). These digital values are immediately sent to the microcontroller unit (MCU) and multiplied for instantaneous powers. By integrating the instantaneous powers over a specified period, the MCU can obtain electricity consumption (kWh). A liquid crystal display (LCD) or light-emitting diode (LED) panel displays the electricity consumption. The data recorded by sensors or calculated by the MCU regularly (for example, every 30 min) are stored in an electrically erasable programmable read-only memory (EEPROM). Note that the MCU serves as the brain of the whole system by performing all the necessary operations, such as storing and retrieving data from EEPROM, operating the meter (e.g., pushing buttons) to view electricity consumptions, calibrating phases, and clearing readings. Also, the MCU drives the LCD/LED panel using decoder integrated circuits (ICs). There is also a power supply unit providing 5 V to the whole system. Usually, to energize LCD/LED panels and EEPROM, even in the case of power outages, a nonrechargeable lithium battery that can last for at least four to five years is installed.

The second and third generations of electronic meters are smart meters applied in AMR and AMI, respectively. Compared to primary electronic meters, which support just one-way power flows from utility companies to customers, AMR-enabled smart meters are equipped with a one-way communication module, allowing them to send electricity consumptions to utility companies automatically. Compared to AMR-enabled smart meters, AMI-enabled smart meters support two-way power and communication flows, i.e., from customers to utility companies and vice versa, which endows AMI-enabled smart meters with functions of net metering, remote disconnect, and tariff distribution, as shown in Fig. 4. The net metering unit allows AMI-enabled smart meters to measure the surplus power, which is the difference between power generated by customers' electricity generation equipment and their demands, exported back to the power grid. The remote disconnect can be realized by using certain communication technology and a latching relay and then clicking a button on specific meter data management software. The tariff distribution can guide customers to move their energy-intense activities from peak load periods with higher rates to off-peak load periods with lower rates for lowering electricity bills.

We summarize the differences between electromechanical meters and electronic meters in Table 5, where the maximum demand indicator (MDI) means the maximum amount of electrical energy required by a specific consumer during a given period. MDI mainly aids utility companies in load forecasting. For meters with an accuracy class $x$, the acceptable range for 100-W load is $[100 - x, 100 + x]$. For electromechanical meters, the International Electrotechnical Commission (IEC) Standard 62053-11 covers accuracy classes 0.5, 1.0, and 2, and the American National Standards Institute (ANSI) Standard C12.1 mainly covers accuracy classes 0.5 and 1.0 [55]–[57]. For electronic meters, the IEC Standard 62053-21 covers accuracy classes 1.0 and 2, the IEC 62053-22 covers accuracy classes 0.2S and 0.5S, and the ANSI Standard 12.20 covers accuracy classes 0.1, 0.2, and 0.5 [55], [57]. On the whole, with electricity meters

evolving, they are gradually equipped with more functions and can measure more variables with higher accuracy.

### B. Compromising Electromechanical Meters

As previously discussed, when the driving torque and braking torque are equal, the disk of electromechanical meters rotates steadily, with speed proportional to the load power. The revolutions of the disk are counted via a gearing mechanism and displayed on dial-in watthours. Thus, adversaries can lower electromechanical meter readings by decreasing the driving torque, increasing the braking torque, slowing the rotation speed of the disk, and/or lowering the gear ratio.

Since the driving torque is proportional to instantaneous load power, which equals the product of voltage, current, and power factor, we can decrease the driving torque by reducing the voltage in voltage coil, decreasing the current in current coils, and lowering the power factor, with more details provided in the following.

1) For decreasing the voltage, adversaries can disconnect the neutral wire from the return path and then connect a high-value resistor between the output end and the ground (or the neutral wire of a neighbor) [58], as shown in Fig. 5(a). Also, adversaries can open-circuit the loop of the voltage coil by either loosening its terminals/connecting links or just breaking the fuse/voltage coil violently. Besides, adversaries can poor-contact the loop of voltage coil by either unscrewing its terminals/connecting links or artificially developing an oxygen layer over terminals/connecting links.
2) About decreasing the current, adversaries can apply directly hooking from power lines, i.e., tapping into power lines from a point ahead of the electricity meter such that some/all currents do not go through the meters and, hence, are unmeasured [58]. It is reported that 80% of global electricity theft is committed through directly hooking from line [59]. Also, adversaries can prevent electricity from being registered by short-circuiting the loop of current coils through "bypassing meters," which means inserting an electric wire directly between meters' input and output terminals. Another commonly used method is to open-circuit the loops of current coils by either loosening its terminals/connecting links or just breaking current coils violently. In addition, adversaries can change the circuit connection by exchanging live wires with neutral wires and, in the meantime, connect the neutral wire to the ground, as shown in Fig. 5(b).
3) Since the power factor equals the cosine of the difference of phase angles between voltage and current in the circuit, the power factor can be lowered by shifting the phase angle of voltage/current. To achieve this goal, adversaries can exchange the input terminals with output terminals of the voltage/current coils' loop. They can also apply inductors or capacitors to change the difference of phase angles between the voltage and the current. To turn down or reverse electromechanical meters, adversaries can also apply a converter whose primary and secondary windings are not electrically connected, as shown in Fig. 5(c). Another possible way to reverse an electromechanical meter is to plug an external power supply, such as a hand-cranked generator into it.

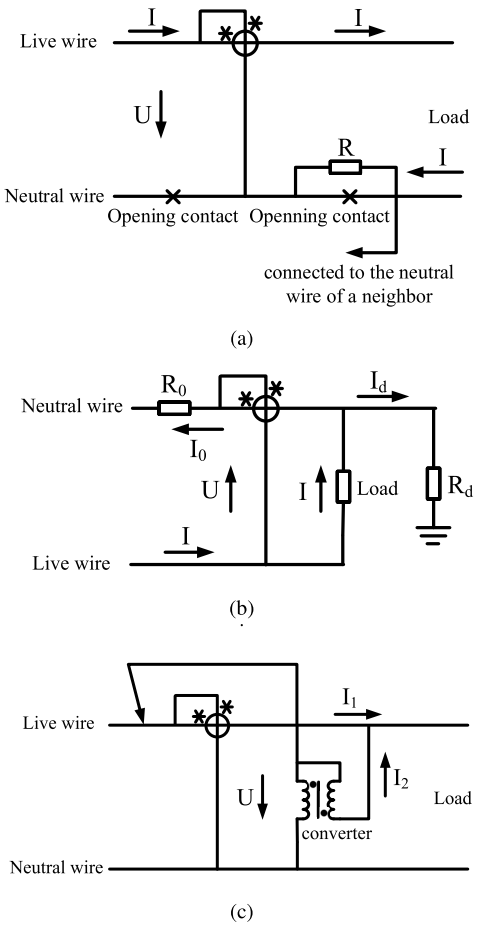


**Fig. 5.** *Circuits for tampering with electromechanical meters [58]. (a) Circuit for decreasing the voltage in the voltage coil. (b) Circuit for decreasing the current in the current coils. (c) Circuit for shifting the phase.*

Adversaries can increase the braking torque by moving the poles of the brake magnet away from the center of the disk. To lower the rotating disc's speed, adversaries can first drill a small hole on the top of electromechanical meters and then insert iron nails or other objects, such as high viscous fluid to stick the disc's rotation. Besides, placing magnets outside meters can drag the disk to slow and even stop the rotation. To lower the gear ratio, adversaries can replace the counter of the large-capacity electricity meter with the counter of the small-capacity one, which reduces readings dramatically. For example, an adversary can replace the counter of a 10(20)-A 900-rpm/kWh meter

with the counter of a 5(10)-A 1800-rpm/kWh meter [58]. In this case, the counter of the original meter increases by one for every 900 revolutions, whereas, after the replacement, it increments by one for every 1800 revolutions. This results in that electricity consumption measured by electromechanical meters is only one-half of what is consumed.

### C. Compromising Electronic Meters

Due to radically different structures, many electricity theft methods applied successfully to electromechanical meters fail in electronic meters. For this reason, electronic meters (especially smart meters) were once regarded as an effective tool for reducing electricity theft [37]. Nevertheless, many reports reveal that adversaries with a background in electronics and software engineering can easily manipulate smart meter readings using some cheap software and/or hardware available on the Internet [10], [60], [61].

Basically, for primary electronic meters that do not have communication functions, adversaries can tamper with electricity consumption data into smaller values during measurement/calculation/storage to steal electricity. For smart meters, which can automatically send electricity consumption data to utility companies across networks, adversaries can also manipulate the data in transmission. Particularly, for AMI-enabled smart meters, since a net metering unit in them (as aforementioned in Section III-A) registers the power generated by customers' own electricity generation equipment and exports back to the power grid, adversaries can also steal electricity by tampering with net metering data in storage into reasonably larger values. More details are provided as follows.

We first discuss ways to tamper with data in measurements. As aforementioned, modern electronic meters generally use current transformers for current measuring. The cores of current transformers are magnetic components that can be saturated by strong external magnets. Thus, adversaries can also place strong magnets outside electronic meters to reduce bills by 50%–75% [62]. Also, adversaries can lower down electronic meter readings by decreasing voltage/current/power factors through compromising circuits outside electronic meters. The ways are the same as those for compromising outside circuits of electromechanical meters, as presented in Section III-B. This is mainly because electronic and electromechanical meters are connected to the circuit in the same way. In other words, components measuring the voltage (voltage coil/voltage measuring circuit in the electromechanical/electronic meters) are connected in parallel with loads. In contrast, components measuring the current (current coil/current measuring circuit in the electromechanical/electronic meters) are connected in series with the loads, as aforementioned in Section III-A.

We then explain how adversaries tamper with data in the calculation. Since calculations are mainly performed in the MCU, the simplest way is to damage it physically. For adversaries with sophisticated computer science-related skills or reverse-engineering capability, they have various smarter ways. For example, these adversaries can masquerade themselves as control devices, such as a data collector in a neighborhood area network (NAN), which have legal access authorities to smart meters [63]. Also, they can inject malware (such as computer viruses and worms) or overwrite the firmware to interfere or redefine the normal calculation process of the MCU. It is reported that adversaries can establish a connection between a smart meter and a computer using an optical converter device (e.g., an infrared light) such that adversaries can tamper with power consumption recording settings using software downloaded from the Internet [10].

Adversaries can tamper with data in storage by methods of physically damaging storage units, masquerading as control devices, injecting malware, and overwriting firmware, similar to those used in the case of manipulating data in the calculation. Besides, adversaries can first extract passwords and then log in electronic meters to reset the data (e.g., electricity consumptions) to any reasonable values. To remove manipulation traces, adversaries usually clear corresponding audit logs immediately. McLaughlin *et al.* [62] use monitoring software installed on a laptop to capture communications over the optical port protocol between utility companies and smart meters. They find that passwords are transmitted in the clear and can be captured by placing a reader device on the optical port pins or near the optical lens. For escaping possible detections, if adversaries steal electricity by physically damaging the MCU or the storage unit, adversaries would further intercept relevant alarming signals; and if they apply other ways to tamper with data in storage, adversaries would pay attention to making compromised data consistent with common senses and physical laws. For example, when tampering with electricity consumptions (in kWh) in primary electronic meters, adversaries should also update the value of MDI correspondingly. When tampering with electricity consumption (in kWh) in smart meters, adversaries should also tamper with voltage, current, and/or power factors such that instantaneous load power still equal products of voltage, current, and power factors.

We next explain methods to tamper with data in transmission. Generally, smart meter readings are transmitted via various communication technologies before finally reaching utility companies. For example, to successfully transmit AMI-enabled smart meter readings from users' homes to control centers of utility companies, we can first apply Wi-Fi or power line communication (PLC) in the NAN and then apply fifth-generation wireless systems (5G) or optical fiber in the wide-area network (WAN). During the data transmission, sophisticated adversaries can take advantage of the inherited security vulnerabilities of the communication technologies to launch different cyberattacks. On the whole, cyberattacks usually involve passively interposing and actively injecting falsified usage data

into communications in the home area network (HAN), NAN, and/or WAN. The most commonly used methods to intercept communications include eavesdropping, traffic analysis, and radio frequency (RF) interception [64]. The most commonly used ways to inject usage data include session hijacking, man-in-the-middle attacks, replay attacks, and spoofing attacks.

We next introduce two electricity theft methods that are independent of the types of electricity meters. The first one is to bribe employees in utility companies with a certain amount of money such that corrupted employees would not report adversaries' electricity theft behaviors to utility companies or even directly manipulate adversaries' electricity bills to lower values [31]. An often-used strategy is to move the decimal point on electricity bills to the left so that adversaries, for example, pay \$47.48 instead of \$474.80 [31]. In India, 7% of households experienced corruption in public electricity service in 2017 [65]. The second one is unpaid bills, which means that some persons/organizations do not pay what they owe to utility companies for using electricity [31]. Unpaid bills usually occur when customers move to a new city, or enterprises go bankrupt. As reported, there are a lot of unpaid bills in South Africa, and the nonpayment level of Armenia even reaches 80%–90% in the residential sector [31].

We summarize all electricity theft methods in Fig. 6. As can be seen, electromechanical meters can be tampered with only when the data are measured; primary electronic meters can be tampered with when data are measured/calculated/stored; and smart meters can be compromised when the data are measured/calculated/stored/transmitted. For electricity theft methods that decrease the values of voltage/current/power factors in measurements, they can be applied to any type of electricity meter. These methods include disconnecting neutral from the return path, hooking from the line, bypassing meters, and so on. Bribing employees and unpaid bills are also electricity methods that can be applied regardless of the types of electricity meters.

In practice, when electromechanical and primary electronic meters without communication capability are installed, adversaries usually perform physical attacks. When smart meters are installed, adversaries usually launch cyberattacks to commit electricity theft. In Fig. 6, we enclose the cyberattacks in dashed lines. This is because, compared with physical attacks, cyberattacks are are much more secret and flexible and can be launched almost anywhere and anytime. Thus, with smart meters being widely deployed around the globe, electricity theft tends to become even more serious and more difficult to be detected [66], [67]. In addition, by taking advantage of these methods, adversaries can compromise numerous smart meters simultaneously, which can incur serious consequences such as large-scale power outages. This motivates researchers to develop various electricity theft detection methods, especially those applicable in smart grids.

### D. Data-Level Attack Models Against Smart Meters

In this article, we use technical level attack models to refer to the physical attacks (such as hooking from lines, bypassing meters) and the cyberattacks (such as hijacking and man-in-the-middle attacks) that adversaries leverage to compromise the data of smart meters. In Sections III-B and III-C, we demonstrate how adversaries steal electricity from the technical level under different types of electricity meters.

We use data-level attack models to refer to the attacks that manipulate meter data in some ways other than just simply lowering down the meter readings, e.g., changing the load profiles temporarily to steal electricity without changing the total electricity consumption or compromising readings of smart meters of neighbors' customers to steal electricity. Since electromechanical meters and primary electronic meters are currently outdated in many regions and will be finally replaced by smart meters across the whole world, in the following, we demonstrate how adversaries commit electricity theft from the data level, under the assumption that all customers are equipped with smart meters.

As mentioned earlier, for adversaries committing electricity theft, their goal is to reduce electricity bills, which are calculated as the summation of products of electricity consumptions reported by users and electricity prices issued by utility companies over a certain period [68]. In practice, utility companies usually apply one of the following pricing schemes: 1) the flat-rate pricing scheme, which has a consistent price per unit regardless of the amount purchased; 2) the time-of-use (TOU) pricing scheme, which applies different prices for electricity at different times of the day; and 3) the real-time pricing scheme, whereby electricity prices vary over short time intervals (typically hourly) according to the real-time costs of electricity generation and electricity demand of users. We depict the above three pricing schemes in Fig. 7.

Based upon what information is tampered with by adversaries, data level models against smart meters can be roughly classified into *consumption attacks* in which adversaries compromise smart meter readings and *pricing attacks* in which adversaries tamper with electricity prices, as demonstrated in the following.

*1) Consumption Attacks:* One of the most intuitive consumption attacks is to tamper with smart meter readings into smaller values, which is called *reduced consumption attacks* in the following. As the most common electricity theft attack models, reduced consumption attacks can be employed under any pricing scheme in Fig. 7. According to the magnitude of differences between the actual and the falsified electricity consumptions, reduced consumption attacks can be further divided into the following two categories: 1) *major difference attacks* in which adversaries' reported readings are much smaller than their actual electricity consumptions and 2) *minor difference attacks* in


**Fig. 6.** ***Archetypal attack tree for tampering with meter readings. The cyberattacks are enclosed in dashed lines.***

which adversaries deliberately tamper with meter readings to values just a little smaller than the actual values, to escape detection [69].

Particularly, under the TOU or real-time pricing schemes, adversaries can also shift their load profiles to lower electricity bills [70], which is called *load profile*

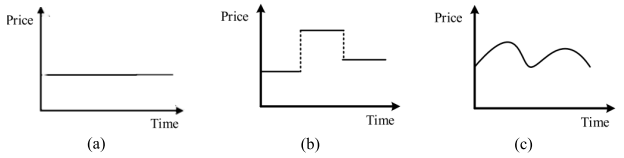


**Fig. 7.** *Curves describing the following electricity pricing schemes: (a) flat-rate pricing scheme, (b) TOU pricing scheme, and (c) real-time pricing scheme.*

*shifting attacks*. Specifically, adversaries report some electricity consumptions that happen when the electricity price is high as electricity consumptions that happen when the electricity price is low. For example, in Fig. 8, we assume that the electricity prices are low during time intervals $[t_0, t_1]$ and $[t_2, t_3]$ and high during time interval $[t_1, t_2]$. Adversaries can launch a load profile shifting attack by moving some electricity consumptions that actually happen during the time interval $[t_1, t_2]$ (represented by area A) as electricity consumptions that happen during time intervals $[t_0, t_1]$ and $[t_2, t_3]$, which are represented by areas B and C, respectively.

We next illustrate two types of collaborative attacks. The first type of collaborative attack is the collaborative NTL (CNTL) frauds, which is first proposed in [71]. A CNTL occurs when multiple malicious users commonly tamper with one meter so that the meter records less electricity than the consumed amount by the household [71]. The second type of collaborative attack is proposed in [72] explained as follows. When adversaries tamper with their readings to lower the meter values, they may simultaneously compromise neighbors' smart meters for increasing the readings to escape detection. These adversaries usually summate neighbors' increased value equal to their decreased value (i.e., the stolen electricity). In this way, adversaries can escape the detection of a central observer meter (or devices with similar functions), which registers the total amount of electricity supplied to all customers in the community.

In addition, most of the existing works assume that, once adversaries begin to launch attacks for tampering with meter readings, they would not stop this fraudulent behavior unless caught by utility companies. However, in reality, for interrupting the detection, adversaries may switch their behaviors constantly between committing electricity theft and honestly consuming electricity. Specifically, they may launch cyber/physical attacks to tamper with meter readings for a while and then stop these attacks for another while. These behavior patterns can be repeated many times. We define the above type of attacks as intermittent attacks.

Both collaborative attacks and intermittent attacks belong to reduced consumption attacks.

*2) Pricing Attacks:* As aforementioned, in pricing attacks, adversaries compromise the price information for reducing electricity bills. One of the most intuitive pricing attacks is to tamper with the electricity prices used for calculating electricity bills into smaller values, which is referred to as *reduced pricing attacks*. In the real world, electricity bills are usually calculated and stored at the utility company's control centers, usually high-security levels. Adversaries with profound knowledge in computer science, communication networks, and other related disciplines can still hack into the database systems of the utility companies' control centers to compromise the electricity prices. Adversaries can also bribe the employees of the utility companies who are authorized to log into the database systems to tamper with electricity prices into smaller values. Reduced pricing attacks can be applied under any pricing scheme.

In the following, we demonstrate how adversaries can launch another type of pricing attack, called neighbor pricing attacks, in which adversaries reduce electricity bills by affecting neighbors' electricity consumption through compromising electricity prices seen by neighbors. Note that the neighbor pricing attacks can only work under the real-time pricing scheme [70], [73], [74]. Assume that customers are equipped with automated demand response (ADR) interfaces by which customers can schedule their energy usage behaviors for saving electricity bills. The most well-known implementation of ADR, known as OpenADR, is based on the Energy Market Information Exchange specification [70]. To launch neighbor pricing attacks, adversaries tamper with the electricity prices seen by neighbors' ADR systems to larger values. Since electricity consumption is typically modeled as a monotonically decreasing function of the electricity prices, the ADR systems are programmed to automatically make schedules for less electricity consumption under a higher electricity price. Thus, when the electricity price seen by the ADR systems increases, the attacked neighbors tend to consume less electricity. If the number of attacked neighbors is large enough, this will decrease the electricity demand. Under the real-time pricing schemes, less electricity demand implies a lower electricity price. Thus, the adversaries' electricity bills are reduced.

To sum up, to achieve the goal of electricity theft, i.e., reducing electricity bills, adversaries can compromise

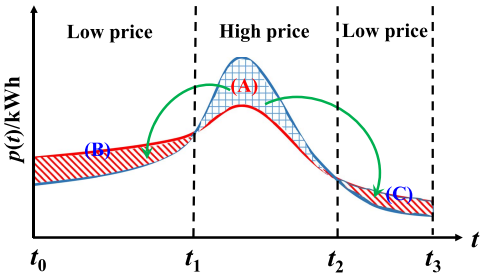


**Fig. 8.** *Curves describing the following electricity pricing schemes: 1) flat-rate pricing scheme; 2) TOU pricing scheme; and 3) real-time pricing scheme.*

**Table 6** Data-Level Attack Models

| data-level attack models | | | collaborative /intermittent attacks | pricing schemes | | | technical level | |
|---|---|---|---|---|---|---|---|---|
| | | | | flat rate | time-of-use | real-time | physical attacks | cyber attacks |
| consumption attacks | reduced consumption attacks | major difference attacks | both | ✓ | ✓ | ✓ | ✓ | ✓ |
| | | minor difference attacks | | | | | | |
| | load profile shifting attacks | | none | × | ✓ | ✓ | × | ✓ |
| pricing attack | reduced pricing attacks | | none | ✓ | ✓ | ✓ | × | ✓ |
| | neighbor pricing attacks | | none | × | × | ✓ | × | ✓ |

electricity consumption measurements or/and price information. The consumption attacks mainly include: 1) reduced consumption attacks that can be applied in any pricing scheme and 2) load profile shifting attacks that can be applied in TOU and real-time pricing schemes. For escaping detection, adversaries can also launch collaborative consumption attacks, in which adversaries launching consumption attacks can simultaneously tamper with their neighbors' electricity consumption measurements to make up for the differences. The pricing attacks mainly include reduced pricing attacks and neighbor pricing attacks, in which adversaries tamper with their own and neighbors' electricity price information, respectively. The reduced pricing attacks can be applied in any pricing scheme, whereas the neighbor pricing attacks can be applied only in a real-time pricing scheme. Among all the above four types of data-level attack models, the reduced consumption attacks are the only ones that can be launched by both physical attacks and cyberattacks. They are also currently the most widely used data-level attack models in practice. The other three types of data-level attack models can only be launched through cyberattacks, such as man-in-the-middle attacks or spoofing attacks to compromise smart meters' inputs or outputs. For better understanding, we summarize the above data-level attack models in Table 6.

## IV. MACHINE LEARNING-BASED DETECTION

### A. Overview

For this category of detection methods, the workflow can be divided into four phases: data selection, data preprocessing, model building, and model applications, as depicted in Fig. 9.

For data selection, *load profiles* that depict customers' electricity consumption during a certain period are almost involved in all of this category of detection methods since load profiles contain significant information of electricity consumption patterns [85]. *Prior records* that document, when adversaries are previously caught stealing electricity, are often used as auxiliary information for analyzing customers' tendency to recommit electricity theft [20]. Other information, such as geographical locations, seasons, and tariff categories (residential, agriculture, commercial, and industrial), may also be used to improve the detection accuracy [82], [84]–[86].

Data preprocessing can be further divided into data reduction, data cleaning, and data transformation, as follows.

1) *Data reduction* is responsible for customer filtering and selection. During this step, customers with the following characteristics are removed:
   a) customers who have the same load profiles with others;
   b) customers whose electricity consumptions are zeros;
   c) customers who are not present in all recording periods;
   d) newly registered customers [20].
2) *Data cleaning* focuses on filling in missing values, smoothing out noise, identifying outliers, and correcting inconsistencies in the raw data. Due to various reasons, such as smart meter failures, unreliable data transmission, unscheduled system maintenance, and storage issues, electricity consumption raw data often contain missing or erroneous values [87]. Missing values are often recovered with an interpolation method or replaced with average consumption values [23], [75]. Erroneous data or outliers are usually restored or eliminated based on the "three-sigma rule of thumb" [23], [75]. This empirical rule mainly states the fact that percentages of values lying within one, two, and three standard deviations of the mean in a normal distribution are 68.27%, 95.45%, and 99.73%, respectively, and hence, it is also called "68–95–99.7 rule" [23], [75].
3) *Data transformation* involves data normalization and data discretization. Typically, data normalization scales all selected data to fall within a small and specified range. It not only prevents attributes with large ranges from outweighing those with small ranges but also speeds up the subsequent model-building process. Most existing machine learning-based methods employ a min–max normalization strategy [20], [23], [76], [77]. Data discretization divides the range of a continuous attribute into

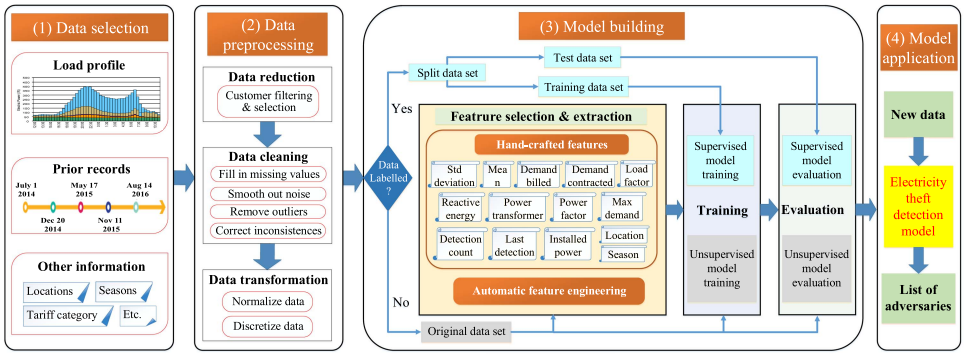


**Fig. 9.** ***Common work flow of machine learning-based detection methods [20], [23], [29], [75]–[88].***

**Table 7** Attributes Extracted From Selected Data

| Data types | Attributes | Comments |
|---|---|---|
| load profile | mean | mean of customers' daily electricity consumptions, in kWh [82, 86] |
| | standard deviation | standard deviation of customers' daily electricity consumptions, in kWh |
| | contracted demand | the load based on the customer requirement with power suppliers as per the supply agreement, in kW or kilovolt-amperes (KVA) [79–83, 86] |
| | maximum demand | the maximum load recorded at specific intervals (i.e., 15min, 30min and 1h) during the billing cycle (usually a month) as per the power supplier, in kW or KVA [78–81, 83]. The maximum power are not expected to cross the contracted demand; otherwise, customers are asked to pay penalties. |
| | billing demand | the demand charges billed by the supplier based on maximum demand recorded and contracted demand, in kW or KVA [78, 83] |
| | reactive energy | the electricity that flows through the electric and magnetic fields of an AC system, in kilovolt-amperes reactive hours (n kVArh) [78, 81, 83] |
| | power transformer | power transformer installed for customers, in kVA [78, 81, 83] |
| | power factor | the ratio of the active power used by customers' load to the apparent power which is the product of voltage and current in a circuit [78, 81, 83] |
| | installed power | the sum of the nominal power of all electrical equipment installed and ready to operate at the consumer unit, in kW [78, 80, 81, 83] |
| | load factor | the ratio of the average load to the maximum load, which shows how rational electricity is used [78–81, 83] |
| prior records | detection count | the total number of times that a customer was caught stealing electricity [20] |
| | last detection date | the last date when customer was detected for stealing electricity [20] |
| other information | geographical location | town or village [82, 85, 86] |
| | tariff category | residential, agricultural, commercial or industrial [82, 84, 85] |
| | number of phases | one, or three [82, 86] |
| | payment regularity | yes, or no [82] |
| | reading regularity | yes, or no [82] |

specified intervals. It is performed when inputs of classification algorithms (e.g., rough set) are categorical attributes. Equal width histogram analysis or equal frequency partitioning can be applied for data discretization [82], [88].

For model building, if the data at hand are labeled (not labeled) with correct information telling whether customers steal electricity or not, supervised (unsupervised) machine learning methods can be utilized. The model building phase can be further divided into stages of feature selection/extraction, model training, and model evaluation, as explained in the following.

1) *Feature selection/extraction:* The feature engineering can be done manually as in classical (or nondeep) machine learning methods or automatically as in deep learning methods, as shown in Fig. 9. In both cases, the selected/extracted features are highly related to the data chosen in the data selection phase. In deep learning, the feature extraction, and the classification are performed in a parallel manner, and there is no need to know what features are finally extracted from the input data. In contrast, in classical machine learning, there is a need first to extract and select features carefully to classify which customers are fraudulent. We demonstrate the handcrafted features in classical machine learning methods next. As summarized in Table 7, when the load profile is chosen, we can extract the following features:

**Table 8** Common Metrics Used for Model Evaluation

| Metrics | Explanation |
|---|---|
| detection accuracy | the ratio of the number of samples correctly classified as adversaries or honest customers to the total number of samples used for testing [20, 77] |
| FPR | false positive rate, which means the ratio of the number of samples incorrectly classified as adversaries to the total number of samples which are actually honest customers [77, 89] |
| precision | the ratio of the number of samples correctly classified as adversaries to the total number of samples classified as adversaries [90] |
| TPR | true positive rate (also called sensitivity, hit rate or recall), which means the ratio of the number of samples correctly identified as adversaries to the total number of samples which are actually adversaries [20, 90, 91] |
| F-measure | the harmonic mean of precision and recall, which is argued to be able to balance the ability of the detection system to detect frauds without generating excessive false positives [92] |
| AUC | the area under the receiver operating characteristic (ROC) curve. In the ROC curve, the true positive rate is plotted in function of the false positive rate for different cut-off points of a parameter [23] |

mean, standard deviation, contracted demand, billing demand, reactive energy, power transformer, power factor, installed factor, and load factor [78]–[83], [86]. When prior records are chosen, the detection count and last detection date can be selected [20]. Other features, such as geographical location, tariff category, number of phases, payment regularity, and reading regularity, can be extracted when related data are chosen [82], [84]–[86].

The principle component analysis (PCA) is a dimensionality-reduction method that is often used to transform a high-dimensional dataset into a smaller dimensional subspace while retaining most of the information. Before training machine learning algorithms, the PCA can be used to select the most informative features such that the data size fit into machine learning algorithms is reduced, and training burdens are lowered [93], [94]. For instance, Krishna *et al.* [94] apply the PCA to extract underlying daily or weekly repeated consumption trends before using the density-based spatial clustering of applications with noise (DBSCAN) approach to determine when customers steal electricity.

2) *Model training* aims to create a mathematical function whose inputs are the selected/extracted features and outputs are a list of adversaries. A machine learning method corresponds to a mathematical function whose parameters are optimized during the model training process.

   The machine learning algorithms can be classified into supervised and unsupervised learning methods. For supervised methods that have labels as supervisors, the overfitting problem should be carefully avoided. For this purpose, researchers usually split the data into a training set and a test set, which typically accounts for 70% and the other 30% of the whole dataset, respectively. Since the test set is withheld from the model training, it contributes to an unbiased evaluation of the model performance. Another antidote for overfitting is cross-validation, which involves partitioning data into multiple groups and then training and testing models on different group combinations.

   About unsupervised learning methods where the original data are kept as a whole, the parameter optimization process is usually achieved by producing relatively small intracluster distances (which is used to determine, for selected features, how close each user within each cluster is to every other user in its cluster) and relatively large intercluster distances (which is used to determine, for selected features, how close each cluster of customers is to other clusters).

3) *Model evaluation:* The commonly used metrics mainly include detection accuracy, false-positive rate (FPR), precision, true positive rate (TPR), F-measure, and the area under the curve (AUC), as summarized in Table 8.

In the phase of the model application, the established detection model is applied to new data samples to see whether corresponding customers are committing electricity theft or not. We summarize in Fig. 10 the commonly used machine learning methods for electricity theft detection, which mainly includes support vector machines (SVMs), optimum path forest (OPF), artificial neural networks (ANNs), statistical inference, clustering, rule extraction, and so on. Most of these methods are based on supervised models, with more details given as follows.

## B. Support Vector Machines

SVMs are often used for classification and regression analysis. Since electricity theft detection requires to classify customers into adversaries and honest customers, it is often dealt with by SVMs [20], [77], [78], [89], [92], [95]–[101]. The SVM-based methods employ a set of

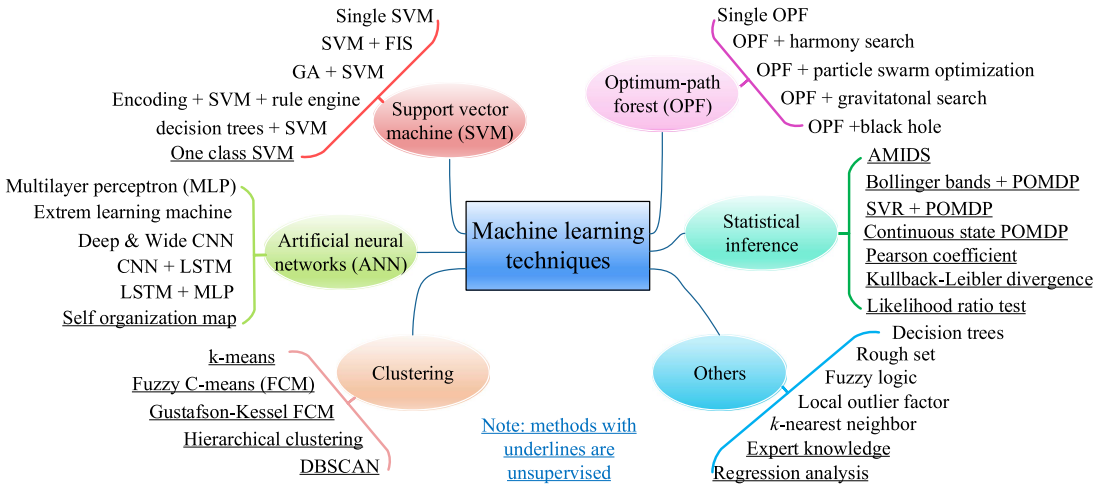


**Fig. 10.** ***Machine learning methods applied for electricity theft detection.***

mathematical functions (known as "kernels") to implicitly map the customers with selected features into high-dimensional spaces. In this way, the hyperplane that produces the largest margin between adversaries and honest customers can be found more easily. The most commonly used kernel functions include the radial basis function (RBF) and the sigmoid kernel [20], [77], [89], [96]–[98]. If prior knowledge of data structures and class boundary types is not available, RBF is usually employed as a general-purpose kernel, as done in most SVM-based electricity theft detection methods [20], [77], [89], [97], [98] When RBF is applied, for improving the classification accuracy, the following two hyperparameters should be carefully chosen: 1) the regularization parameter $C$, which trades off correct classification of training examples against maximization of the decision function's margin, with higher values meaning more penalty on misclassification and 2) the RBF kernel parameter $\gamma$, which defines how far the influence of a single training example reaches, with lower values meaning wider-spread influence [102]. These parameters can be optimized by asymptotical optimization, cross-validation, and grid search [20], [77].

The SVM-based detection model in [20] aims to find electricity theft-related irregularities in load profiles, and the TPR is only 60%. Specifically, input features of the SVM-based detection model in [20] include 24 daily average electricity consumption data and one credit worthiness rating value reflecting customers' monthly payment status; the output is a list of suspected customers for onsite inspection. However, the TPR is low, and it can only be applied to scenarios where abrupt changes appear in customers' load profiles. Through including human knowledge and expertise into the above SVM-based detection model via a fuzzy inference system, the TPR is improved to 72% [98]. Specifically, the fuzzy inference system acts as a post-processing scheme for short-listing customer suspects with higher probabilities of fraud activities [98]. Nagi *et al.* [97] combine the genetic algorithm (GA) with the SVM-based detection model in [20], where GA provides an increased convergence for globally optimizing SVM hyperparameters using a combination of random and prepopulated genomes.

To accelerate the detection process, Depuru *et al.* [96] propose to encode customer energy consumption data for quicker analysis as follows: 1) first, zero/nonzero consumption data are encoded into digit 0/1 and 2) then, every three bits of the binary (i.e., "000," "001," and so on) are regarded as a whole and transformed into decimal numbers ranging from 0 to 7. These data are then immediately fed into SVM and rule engine in parallel, for improving the detection accuracy. If the SVM module classifies customers as being honest/suspicious, then these customers are regarded as being genuine/suspicious. The suspicious customers' load profiles are further checked by some expert-specified rules, according to which suspicious customers are further divided into different groups. Groups of suspicious customers with a higher probability of illegal consumption are inspected immediately, while the others are reported to the utility companies for re-evaluation and periodical inspection. The limitation of this work lies in that the encoding procedure cannot reflect the change of electricity consumption when adversaries tamper with meter readings into smaller but nonzero values, which may result in a higher false-negative rate.

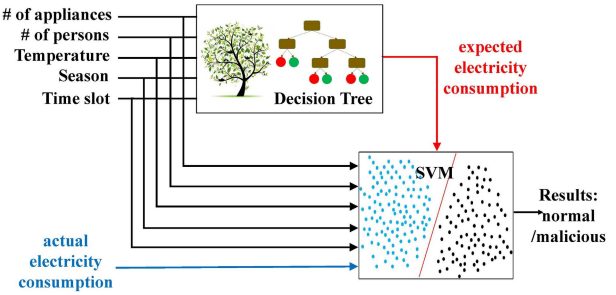


**Fig. 11.** ***Block diagram of the detection scheme based on the decision tree and the SVM [77].***

Jindal *et al.* [77] propose a decision tree and SVM-based detection scheme, as shown in Fig. 11, where the decision tree calculates customers' expected electricity consumptions, based on the following input parameters: the number of appliances/persons, temperature, season, and time during a day. Then, both the input and the output of the decision tree, together with users' actual electricity consumptions, are used as input parameters of the SVM module for classifying users into honest or malicious users. This reduces false positives to as low as 5.12% and improves detection accuracy to 92.5%.

For decreasing the interference of nonmalicious factors (such as changes of seasonality and appliances), which can also alter consumption patterns, before training the SVM model, Jokar *et al.* [89] apply the $k$-means algorithm on a benign dataset to remove clusters with fewer members. The detection accuracy and the false positive rate in [89] are 94% and 11%, respectively.[1]

For addressing the data imbalance issue (which means that there are many more benign samples than malicious samples in the training data), Zanetti *et al.* [92], Rodriguez *et al.* [101], and Martino *et al.* [103] propose to detect electricity theft using one-class SVM. Unlike classical SVM models, which need two classes of training samples, one class SVM only needs a specific class of training samples (in our case, normal consumption patterns are used). The one class SVM learns a decision function to classify new data, i.e., to find out whether it is alike or not like the training data [104].

In summary, most SVM-based electricity theft detection methods are supervised models that apply the RBF as kernels. As shown in Fig. 10, we can employ SVMs alone or together with other machine learning methods, such as fuzzy inference system, GA, rule engine, decision trees, and $k$-means. Particularly, the one class SVM is an unsupervised model [104].

## C. Artificial Neural Networks

ANNs are brain-inspired computing systems intended to replicate the way humans learn from examples [105]. Similar to our brains that consist of billions of biological neurons, an ANN is composed of a collection of connected units or nodes called artificial neurons. Each artificial neuron consists of inputs, weights, an activation function, and an output. After each input is multiplied by a corresponding weight, the inputs are summed up. The results are then computed in the activation function to determine the final output. The usually nonlinear activation function enables ANNs to learn complex relationships between the inputs and the output. The most commonly used activation functions include the sigmoid function, the hyperbolic tangent function, and the rectified linear unit (ReLU) function. Typically, artificial neurons are grouped into different layers that can be categorized into an input layer, at least one hidden layer and an output layer, each of which contains a user-specified number of neurons. For example, there is a single neuron and two neurons in the output layer of ANNs in [83] and [86], respectively. The output layer of the ANNs produces a probability or directly judges whether customers are committing electricity theft or not.

Deep neural networks (also called deep learning) are part of a broader family of machine learning methods based on ANNs with representation learning [106]. Deep learning has multiple (at least two) hidden layers, which progressively and automatically extract higher level features from the input data. In other words, deep learning automates the process of feature engineering, which makes it transcend classical (i.e., nondeep) machine learning methods (e.g., SVMs) dependent on handcrafted features [107]. In ANN-based electricity theft detection methods, the input can be either (a subset of) handcrafted features in Table 7 or simply customers' load profile (possibly together with other customer-related information) whose high-level features need to be further automatically extracted by deep learning.

In the existing literature, ANN models leveraged to identify adversaries mainly include feedforward neural networks, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), as explained in the following.

**Feed forward neural networks** are ANNs wherein connections between nodes do not form a cycle. The multilayer perceptron[2] (MLP) in both [83] and [86] is a widely employed feedforward neural network with an input layer, a hidden layer, and an output layer. Pereira *et al.* [83] and Costa *et al.* [86] first select features, such as mean consumption and maximum demand, and then feed these features into an MLP. Specifically, in [86], the MLP is trained with a backpropagation method [95], which aims to minimize the mean squared error between the desired and the obtained outputs of neurons in the output layer. For improving the training speed, Pereira *et al.* [83] train the MLP with the charged system search (which belongs to metaheuristics algorithms based on nature-social behavior). In [76], extreme learning machine (ELM), which is also a feed forward neural network with a single hidden layer, is applied to identify adversaries. The input of the ELM in [76] is users' daily load profile represented by 48 electricity consumptions that are captured every 30 min. Usually, ELM can learn much faster than MLP with backpropagation since hidden nodes in the ELM can be randomly assigned, and their parameters do not need to be tuned.

**CNNs** are a specialized type of ANNs that perform convolutions, instead of general matrix multiplication, in at least one of their hidden layers [105]. These layers are called convolutional layers, which usually shares an identical kernel (or filter) for fewer parameters to train and decrease the computational cost. For improving the

[1]It should be highlighted that the results above are hard to compare to each other since different metrics and datasets are used in the relevant works.

[2]The perceptron (i.e., neuron) is actually a threshold function that maps its input to an output of "0" or "1."

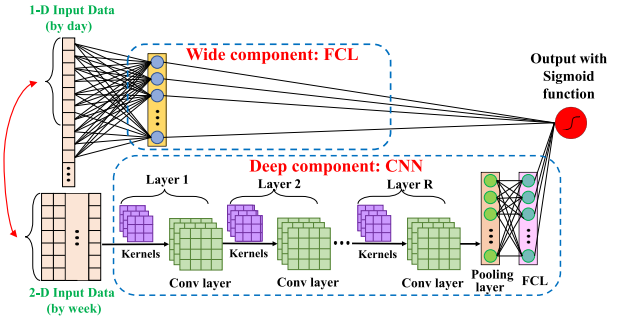


**Fig. 12.** ***Deep and wide conventional neural network for electricity theft detection [23].***

detection accuracy, Zheng *et al.* [23] propose a deep and wide conventional neural network to improve the detection accuracy, mainly by integrating the benefits of the wide component that can learn the global knowledge with the deep component that can capture the periodicity of electricity consumptions. As shown in Fig. 12, the deep component is a CNN whose inputs are customers' daily electricity consumption data folded by week in a 2-D matrix. By folding electricity consumptions into matrices, spatial features that refer to the arrangement of customers' daily electricity consumptions are introduced. The hidden layers of the deep CNN component consist of multiple convolutional layers, a pooling layer, and a fully connected layer (FCL). The pooling layer is used to reduce redundant features, and the FCL is used to obtain principle features [23]. On the other hand, the wide component is essentially an FCL fed with customers' 1-D daily electricity consumption. The wide component aims to learn features that frequently occur in the data from a global view. The weighted sum of the wide and deep component's outputs is finally fed into a logistic loss function for joint training and classification. The limitation of this work lies in that the FCL cannot learn temporal dependencies in the time series of customers' electricity consumption.

**RNNs** are ANNs wherein connections between nodes form a directed graph along a temporal sequence. Long short-term memory (LSTM) networks are a special class of RNNs designed to avoid the short-term memory problem of RNNs [108]. LSTM networks are usually stacked by multiple layers of LSTM nodes, each of which is composed of a cell, an input gate, an output gate, and a forget gate. The cell remembers values over arbitrary time intervals, and the three gates regulate the flow of information into and out of the cell. The LSTM networks are capable of remembering and propagating significant information from the initial stages of the network toward the final stage [108].

To overcome the limitation in [23], Hasan *et al.* [108] propose to integrate the CNN with the LSTM. Specifically, inputs of the CNN in [108] have the same structure as that in [23]. The CNN in [108] is used for automatically extracting high-level features. These features are then flattened and fed into an LSTM neural network for capturing temporal dependencies. Similarly, Buzau *et al.* [109] propose to combine an LSTM network with an MLP to detect electricity theft. As shown in Fig. 13, inputs of the LSTM module are sequential features extracted from users' electricity consumption data, including daily energy consumption of weekdays, the number of zero/missing measurements in each weekday, and the season of the week. In contrast, inputs of the MLP module are non-sequential features, such as contracted demand, contract type, and smart meter firmware versions [109]. Afterward, outputs of the LSTM and the MLP modules are used as inputs to a hybrid module that outputs the probability that the corresponding customer commits electricity theft.

Different from the above methods that are feedforward and supervised ANN models, a self-organizing map is an unsupervised ANN model and is applied in [84] for identifying groups of similar customers based on the graphical representation of time-variant inputs of electricity consumptions.

In summary, the feedforward neural networks, such as MLP and ELM, are the earliest ANN models used to detect electricity theft in smart meters. However, they cannot capture the temporal dependencies in the time series of customers' electricity consumption. In contrast, the LSTM networks can perform better in learning sequential information of the sequential data. Hence, they are recently used popularly in electricity theft detection. Spatial features, which refer to customers' daily electricity consumption, are introduced by folding these data weekly in a 2-D matrix. The CNNs that can capture these spatial features can successfully learn the periodicity of users' electricity consumption. Both CNNs and LSTM networks belong to deep learning methods. Most ANN-based methods are supervised models, which includes, but not limited to, MLP,[3] ELM, CNNs, and LSTM neural networks. As shown

[3]The perceptron is a threshold function that maps its input to an output of "0" or "1."

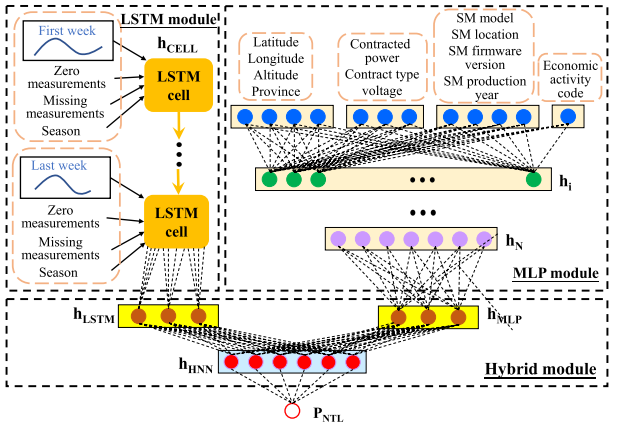


**Fig. 13.** ***Electricity theft detection framework combining the MLP and the LSTM [108].***

in Fig. 10, the self-organizing map is an unsupervised ANN used for electricity theft detection.

### D. Statistical Inference

To judge whether customers are adversaries, this category of detection methods apply different kinds of statistical tools to analyze customers' load profiles and/or other information [70], [73], [74], [95], [110]–[114]. The commonly used statistical tools are demonstrated as follows.

1) Bayesian models create a bridge between what we know and what we want to know by relating different conditional probabilities based on the Bayes' theorem. For electricity theft detection, researchers usually apply the following two Bayesian models.
   a) The Naive Bayes classifier [95], [110] that is a probabilistic classifier based on the assumption of strong independence between features. It requires prior knowledge of probabilities customers manipulate meter readings, which can be inferred from existing relevant data [110].
   b) The Bayesian networks [111] that represent a set of variables and their conditional dependencies via a directed acyclic graph. The Bayesian networks have the advantage of being easily interpreted by humans. That is to say, through observing a Bayesian network, we can easily pick out the most significant features that facilitate the detection of adversaries. Note that the Bayesian models not only can predict whether the customers are adversaries but also can produce a probability that the prediction is right [111].
2) Markov models are stochastic models used in randomly changing systems. Markov models assume that the Markov property that the conditional probability of future states depends only on the value of the current state and is independent of all previous states [115]. To address the electricity theft detection issue, researchers mainly apply the following two Markov models: 1) the hidden Markov model (HMM) that consists of invisible states and visible observations and 2) the partially observable Markov decision process (POMDP) that includes invisible states, visible observations, and actions, where states cannot be known perfectly [73], [74], [112]. The difference between HMM and POMDP is that in POMDP, we have control over the state transitions through the actions taken [116].
3) Bollinger bands are a statistical chart widely used in the financial area for data analysis. It consists of a middle band with two outer bands. The middle band is a simple moving average computed based on historical data. The upper and lower bands are several times the standard deviation of the historical data above and below the middle band, respectively. The Bollinger bands are leveraged to limit the normal range of customers' electricity consumption.
4) The Pearson correlation coefficient is the covariance of two variables divided by the product of their standard deviations, which ranges from $-1$ to 1. It is a measure of the linear correlation between two variables. The values of 1, 0, and $-1$ denote total positive linear correlation, no linear correlation, and total negative linear correlation, respectively. Monedero *et al.* [111] employ a means of windowed analysis with the use of Pearson correlation coefficient to detect whether there are anomalous drastic drops in customers' electricity consumptions.
5) The Kullback–Leibler divergence is a nonsymmetric measure of the difference between two probability distributions. It is useful for comparing the distribution of a set of measurements against the distribution of a baseline model and has the advantage that it does not assume any underlying parametric distribution. Krishna *et al.* [70] propose a Kullback–Leibler divergence-based detector. If consumption readings in a week deviate too much from the historic distribution, corresponding customers are regarded as adversaries. This detector can identify cleverly crafted electricity theft attacks that circumvent the checking performed at internal nodes of a radial network topology [70].
6) The likelihood ratio test is a statistical test used for comparing the goodness of fit of two statistical models—a null model against an alternative model. Based upon the likelihood ratio, the test expresses how the data follow one model better than the other. Under the assumption that smart meter measurements are independent and are drawn from identically distributed random variables, Amin *et al.* [113] and Lin *et al.* [114] apply the Neyman–Pearson theory to conduct likelihood ratio tests to identify adversaries.

In summary, the most commonly used statistical inference methods mainly include Bayesian models, Markov models, Bollinger bands, the Pearson coefficient, the Kullback–Leibler divergence, and the likelihood ratio test. As shown in Fig. 10, these methods are usually supervised. In practice, electricity theft detection models usually apply multiple statistical inference methods and/or other machine learning methods. For example, McLaughlin *et al.* [110] apply Bayesian models, HMMs, and $k$-means together to improve the detection accuracy of electricity theft.

We list several exemplary models to show how the above statistical tools are applied in the area of electricity theft detection as follows.

*1) AMI Intrusion Detection System (AMIDS):* McLaughlin *et al.* [110] propose an energy theft detection framework called AMIDS, which applies the naive Bayes classifier and the HMM together. As shown in Fig. 14, the

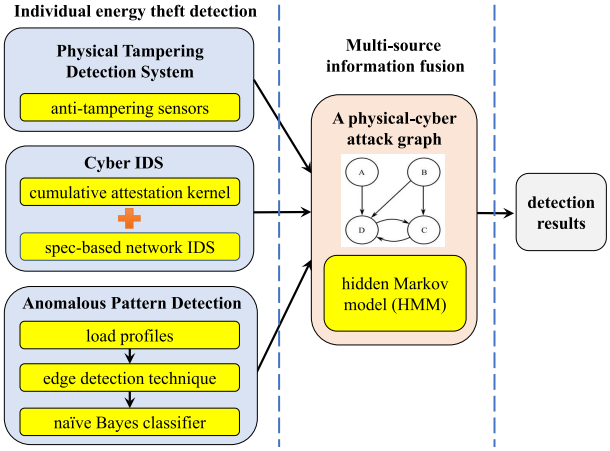


**Fig. 14.** ***AMIDS framework for energy theft detection [110].***

AMIDS consists of a physical tampering detection system, a cyber-intrusion detection system (IDS), and an anomalous pattern detection system, which are first individually applied to collect evidence of electricity theft. Specifically, the first detection system applies on-meter antitampering sensors to detect physical tampering events, such as cover removal and physical bumping. The cyber IDS consists of remote cumulative attestation kernels recording application firmware upgrades and specification-based network IDS monitoring meters' communications. The third detection system applies some edge detection methods to obtain daily usage frequencies of appliances, which are then fed into a naive Bayes classifier to determine whether a given day-long smart meter readings are normal. Outputs of the above three detection systems are then correlated using a directed physical–cyberattack graph formulated as an HMM. Particularly, each attack path in the graph is considered as a discrete-time hidden Markov process, in which an invisible state is a step for committing electricity theft, and the observations are the set of triggered alerts at that state. After obtaining the probability distributions over states, the AMIDS picks out the state with the highest probability. If the goal of electricity theft is achieved at this state, the AMIDS triggers an alert [110]. The AMIDS can significantly reduce FPRs since it combines several sources of information to help judge whether customers commit electricity theft.

*2) POMDP-Based Methods:*

*a) Bollinger-POMDP detector:* Liu and Hu [112] propose a Bollinger-POMDP detector that integrates the Bollinger bands and PODMP to detect collaborative reduced consumption attacks. The Bollinger bands are constructed with historical energy usage data, which are then used to detect whether each customer's energy usage is normal. Specifically, if electricity consumptions at a certain period are below the lower band, corresponding customers are considered as adversaries [112]. Results of the Bollinger bands are then used to obtain observations, i.e., the number of smart meters observed to be hacked. The states of the POMDP are the numbers of smart meters hacked, which cannot be perfectly known but estimated through observations. The POMDP model has two actions to choose from: 1) ignoring smart meters that are observed to be hacked (if any) and keeping monitoring when the impacts are negligible and 2) checking and fixing the hacked smart meters. The system losses or labor costs introduced by these actions are modeled as rewards in the POMDP model. The utility companies aim to take action at each time slot to maximize the long-term rewards. Suppose that the solution of the above POMDP model shows that the expected rewards associated with action (2) are larger than that with action (1). In that case, the smart meters identified by the Bollinger bands are checked and fixed. The simulation results demonstrate that the Bollinger-POMDP detector can successfully detect 92.55% of energy theft on average while effectively mitigating the impact to the community [112].

*b) SVR-PODMP detector:* Similarly, the same authors propose an SVR-PODMP detector in [73], which integrates the support vector regression (SVR) methods and POMDP to identify neighbor pricing attacks. Specifically, the SVR method is used to predict electricity prices seen by customers, called guideline prices, based on historical guideline prices. If the maximum difference between the predicted and the received daily guideline prices of a smart meter exceeds a threshold, this smart meter is identified as being hacked. The states, observations, actions, rewards, and goals of the POMDP model in [73] are the same as those in [112]. The simulation results show that the SVR-PODMP-based detector can identify neighbor pricing attacks with an accuracy of more than 97%.

*c) Continuous state PODMP-based detector:* From the above discussions, we know that detectors in both papers [73], [112] apply a discrete POMDP model whose states are discrete. The discrete POMDP model must take discrete observations, for obtaining which Liu *et al.* [73] and Liu and Hu [112] first apply Bollinger bands and SVR, respectively, to make a binary classification of smart meters (e.g., hacked or not hacked). However, the information contained in the binary classification result is much less than that contained in the customers' consumption measurements and electricity prices.

To address this limitation, the same authors further extend their work in [74] to develop a continuous state PODMP-based detector, which aims to identify adversaries that cooperatively launch both collaborative reduced consumption attacks and neighbor pricing attacks. The continuous state PODMP-based detector applies a POMDP model whose states are continuous. It takes continuous observations, such as the peak to average ratio of energy load and customers' electricity bills. The actions, rewards, and goals of the continuous state POMDP model are the same as those in [73] and [112]. However, the continuous state POMDP cannot be solved directly since this involves an infinite number of states, and there is no analytical

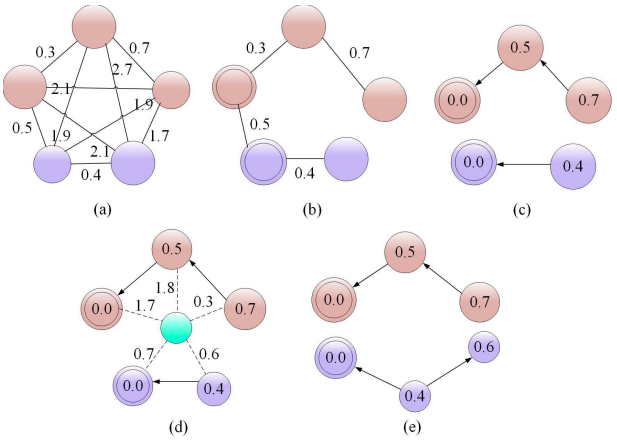


**Fig. 15.** *Example for illustrating the OPF model [119]: (a) training set modeled as a complete graph, (b) minimum spanning tree computation over the training set (prototypes are highlighted), (c) OPF over the training set, (d) classification process of a "green" sample, and (e) test sample is finally classified.*

solution. Although Liu *et al.* [74] introduce some efficient methods to solve the continuous state POMDP problem, the computational complexity is still high.

### E. Optimum-Path Forest (OPF)

OPF models the classification problem by partitioning a complete graph into optimum-path trees. Based upon the assumption that the elements in the same class are more similar than those in different classes, an OPF classifier can be constructed as follows [117], [118].

1) First, a complete graph is formed, with nodes being samples in the training set and arcs that link all pairs of nodes being weighted by the distances between feature vectors of their nodes, as shown in Fig. 15(a).
2) Second, a representative is chosen for each class, called a prototype. For finding prototype nodes, a minimum spanning tree is computed over the training set. The connected nodes are marked with different labels of classes, as shown in Fig. 15(b).
3) Third, an OPF is formed as the classifier, with roots being the prototype nodes. Each training sample belongs to one optimum-path tree rooted at its most strongly connected prototype, as shown in Fig. 15(c).
4) Finally, when classifying a new sample (for example, a test sample), the classifier evaluates the optimum paths from the prototypes to this sample, which is then labeled as the same of the most strongly connected root, as shown in Fig. 15(d) and (e).

The OPF classifiers have the following advantages: (a) they do not have parameters; (b) they have no assumption about the shape of separability of the feature space; and (c) they need a shorter time to train the samples than other methods. These make it possible to detect electricity theft in near real time. Ramos *et al.* [79] propose to apply the OPF classifier for electricity theft detection. Specifically, one node in the OPF represents a user sample with the following four features: contracted demand, maximum demand, load factor, and installed power, which are extracted from the corresponding user's load profiles. Definitions of the above four features can be seen in Table 7. Experiments are conducted on two datasets obtained from a Brazilian electric power company. Experiment results show that the OPF-based detector is 450–500 and 80–95 times faster than the SVM- and MLP-based detectors, respectively [79]. The detection accuracy of the OPF-based detector is about 88%, which is 10% and 30%–40% higher than that of the SVM-based and the MLP-based detectors, respectively. To address the limitation of the above work that the features are manually selected, multiple OPF-based methods are further proposed to apply evolutionary-based methods for feature selection. These evolutionary-based methods include the harmony search algorithm [78], [100], the particle swarm optimization [100], [120], the gravitational search algorithm [100], and the black hole algorithm [81], [121]. In [122], a probabilistic-based OPF classifier is used for electricity theft detection. For improving the detection accuracy, Trevizan *et al.* [123] propose to use the results of a distribution state estimator as additional input information of the OPF classifier.

As summarized in Fig. 10, OPF is a supervised learning method. When used for electricity theft detection, OPF can be applied by itself or together with other methods, which includes state estimation, various evolutionary algorithms, and so on.

### F. Clustering

Clustering involves partitioning a set of objects, usually represented by data vectors, into clusters such that objects in the same cluster are more similar than those in other clusters. The similarity is usually measured by distances between data vectors. In the case of electricity theft detection, the objects are customers consuming electricity [89], [90], [94], [124]–[127]. By and large, clustering methods can be classified into two categories: 1) hard clustering where each object belongs to exactly one cluster and 2) soft clustering, also called fuzzy clustering, where each object can potentially belong to multiple clusters, with a certain degree for each cluster.

The most commonly used hard clustering methods include $k$-means clustering [89], [90], hierarchical (agglomerative) clustering [124], and DBSCAN [94]. These methods are usually leveraged at the preprocessing level for removing outliers from benign training data [89], finding behavioral patterns of a single user [124], finding prototypes of a group of customers [90], and so on. With the outputs used for training classifiers, the detection accuracy can be improved. Krishna *et al.* [94] first apply the principal component analysis (PCA) to filter out noise in electricity consumption data and then leverage the DBSCAN method as the classifier to determine the weeks in which adversaries commit electricity theft.

On the other hand, the most commonly used soft clustering method is the fuzzy $C$-means (FCM) clustering, whose objective functions are the same as those in the $k$-means clustering. The only difference is the introduction of a vector that expresses the percentage of a given data point belonging to each of the clusters. As a variation, the Gustafson–Kessel FCM (GK-FCM) extends the FCM clustering algorithm by using cluster-specific fuzzy Mahalanobis distances instead of the previous Euclidean distance. As argued in [125], GK-FCM offers greater flexibility in terms of cluster shapes [125]. Different from [126] and [128] where FCM is used to find consumers with similar consumption profiles, GK-FCM is applied in [125] to extract the prototypes from benign data. Both FCM and GK-FCM finally get a score indicating the chances of customers stealing electricity.

As summarized in Fig. 10, all the clustering methods mentioned above, including $k$-means, hierarchical clustering, DBSCAN, FCM, and GK-FCM, are unsupervised methods. In almost all applications, clustering methods are used with other methods, such as SVM and PCA, for detecting adversaries stealing electricity.

### G. Others

Other machine learning-based methods are briefly introduced as follows.

1) Rule extraction-based methods draw a set of essential decision rules expressed in IF-THEN forms along with connectors OR and AND to distinguish adversaries from honest customers, given the values of selected features. The simplest way to define such rules is to employ expert knowledge, which can be acquired by personal interviews, structured interviews, observing experts at working sites, and protocol analysis [96], [129]. When relevant experts are absent, the rules are usually inferred by methods, such as the fuzzy logic or rough set theory [82], [98].
2) The local outlier factor is an unsupervised outlier detection method. For any given data point, the local density deviation is computed for its neighbors. Data points whose densities are substantially lower than their neighbors are considered outliers. It is usually used as one component of an electricity theft detector [124], [130]–[132]. For example, in [130], the local outlier factor and the clustering methods are combined together.
3) $k$-nearest neighbor is probably the simplest supervised classification algorithm for detecting malicious customers. It classifies a user by a majority vote of its neighbors, with the object being assigned to the class most common among its $k$ nearest neighbors. It is mainly used as a baseline for comparison with other algorithms [78], [95], [100].
4) A decision tree is a decision support tool using a tree-like graph. Internal nodes, branches, and leaf nodes of this graph represent tests on an attribute, outcomes of the tests, and class labels, respectively. The paths from the root to leaves represent classification rules, which can help understand the characteristics of adversaries' consumption patterns. Although there are various types of decision trees, C4.5 is mostly used, followed by CART [77], [95], [101], [111], [133].
5) Regression analysis is a tool for estimating the relationships among variables. Mashima and Cárdenas [131] use the autoregressive moving average to forecast customers' electricity consumptions, and adversaries are detected by comparing forecasting values with measured values. Monedero *et al.* [111] develop two regression analysis-based algorithms on the evolution of customer electricity consumption.

As summarized in Fig. 10, among all the methods mentioned above, fuzzy logic, rough set, the local outlier factor, $k$-nearest neighbor, and decision trees are supervised models, whereas the expert knowledge and regression analysis are unsupervised methods. Most of the above methods are used as one component of a complex electricity theft detector for improving the performance in terms of accuracy, efficiency, and so on.

In Tables 9 and 10, we summarize and compare the above machine learning-based methods in terms of the following aspects: 1) are the detection methods supervised or unsupervised? 2) are the features extracted manually or automatically? 3) what are the metrics used to evaluate these detection methods? 4) what are the data-level attack models adopted by the detection methods? 5) what are the outputs of these detection methods? 6) what are the kinds of datasets used to train and evaluate these detection methods? and 7) Are the detection methods general or user-specific? Note that a general detector regards all the users as a whole; it is trained by the dataset of all users; and it is then used to see whether there are adversaries among all the users. In contrast, a user-specific detector deals with each user individually. It is trained with only the dataset of one customer and then used to judge whether this user steals electricity or not.

## V. MEASUREMENT MISMATCH-BASED DETECTION

### A. Overview

Utility companies used to send personnel door-by-door to check whether customers compromise their energy meters, and this is very inefficient and labor-consuming [134]. For improving efficiency and reducing labor costs, Xiao *et al.* [135] and Xiao *et al.* [136] propose a mutual inspection strategy that requires installing a redundant smart meter at the utility company side for each user to monitor whether users are stealing electricity. However, this needs exorbitant deployment costs, especially in metropolitan areas, such as New York and Beijing [72]. In contrast, Xiao *et al.* [72] propose a scanning method to inspect each user's meter one by one. However, this is too time-consuming. The mutual inspection scheme and the

**Table 9** Comparing Different Machine Learning-Based Methods: Part I

| Algorithm | supervised (Yes\No?) | feature extraction | main metrics | attack model |
|---|---|---|---|---|
| single SVM[20], SVM+FIS[98], GA+SVM | supervised | manual | hitrate (60%∼72%) | reduced consumption (major difference) attacks |
| encoding +SVM+rule genine[96] | supervised | manual | accuracy (89%) | reduced consumption (major difference) attacks |
| Decision tree+SVM[77] | supervised | manual | accuracy (92.5%) FPR (5.12%) | reduced consumption (major difference) attacks |
| One-class SVM[92] | unsupervised | manual | TPR (85∼95%) FPR (25∼35%) F-measure (60∼75%) | reduced consumption (major difference) attacks |
| MLP +CCS[83] | supervised | manual | accuracy (about 92.5%) | reduced consumption (major difference) attacks |
| ELM[76] | supervised | manual | accuracy (20∼70%) | reduced consumption (major difference) attacks |
| deep & mide CNN[23] | supervised | automatic | AUC (about 80%) | reduced consumption (major difference) attacks |
| CNN +LSTM[108] | supervised | automatic | accuracy (about 90%) | reduced consumption (major difference) attacks |
| LSTM +MLP[109] | supervised | automatic | AUC (85∼85%) | reduced consumption (major difference) attacks |
| AMIDS[110] | unsupervised | manual | accuracy (87%) | reduced consumption (major difference) attacks |
| Bollinger +POMDP [112] | unsupervised | | accuracy (92.55%) | the 2nd type of collaborative attacks |
| SVR+POMDP [73] | unsupervised | | accuracy (97%) | neighbor pricing attacks |
| continuous state POMDP [74] | unsupervised | | computation time (40.75s) | the 2nd type of collaborative attacks +neighbor pricing attacks |
| OPF[78, 79, 81] [100, 121–123] | supervised | manual | accuracy (85∼98%) | reduced consumption (major difference) attacks |

scanning scheme are two extreme schemes: the former is the fastest but the most costly, and the latter is the cheapest but the slowest. Most of the methods that we survey next are some schemes between the above two extreme schemes. For example, for balancing deployment costs and detection efficiency, researchers further propose to install a limited number of advanced sensors at part of feeder nodes or some places, such as distribution rooms and electrical poles in distribution networks of power systems.

The advanced sensors are normally tamper-proof, and their measurement periods are synchronized with users' smart meters. They perform the following functions periodically: 1) receiving reported electricity consumptions of users under investigation; 2) measuring the total amount of electricity distributed to users under investigation; and 3) comparing their measurements with summations of the received readings. If the differences are large, there are adversaries among users under investigation; otherwise, all these users are honest. The above procedure lasts for one reporting period and is called one inspection step in [68], [72], and [137]–[141]. The basic idea of measurement mismatch-based methods is to constantly narrow down the search zone of adversaries by employing advanced sensors to simultaneously/successively perform inspections until all adversaries are finally found out.

For strategies employed to narrow down the search zones, existing methods of measurement mismatch detection can be roughly classified into the following three categories: 1) sensor deployment-based methods; 2) group change-based methods; and 3) behavior approximation-based methods, each consisting two or three branches, as shown in Fig. 16. We present the basic ideas, assumptions, and goals of the above three categories as follows.

*1) Sensor Deployment-Based Methods:* This category of detection methods assumes that users' historical load profiles are observable. Based on load profiles, users' attacking probabilities are calculated, where attacking probability is defined as the ratio of anomalous readings to the total number of readings during a certain period (e.g., a year) [134], [142]. Based upon users' attacking probabilities, pivotal locations near which electricity theft probably occurs can be identified. These pivotal locations are key monitoring areas where advanced sensors should be installed with higher chances. The aim is to find an optimal sensor deployment so that the number of deployed advanced sensors is as small as possible, and the ratio of anomalous users effectively monitored to all anomalous users is as large as possible. The focus of these detection methods is not to pinpoint all adversaries exactly but to

**Table 10** Comparing Different Machine Learning-Based Methods: Part II

| Algorithm | Outputs | data source | general/user-specific |
|---|---|---|---|
| single SVM[20], SVM+FIS[98], GA+SVM | list of adversaries | real data from Malaysia | general |
| encoding +SVM+rule genine[96] | classification results: potential adversary, adversary, genuine users | unclear | user-specific |
| Decision tree+SVM[77] | classification results: adversary or not | synthetic data | user-specific |
| One-class SVM[92] | classification results: adversary or not | real data from Australia | user-specific |
| MLP +CCS[83] | list of adversaries | real data from Brazilian | general |
| ELM[76] | list of adversaries | real data from Malaysia | general |
| deep & mide CNN[23] | classification results: adversary or not | real data from China | user-specific |
| CNN +LSTM[108] | classification results: adversary or not | real data from Chain | user-specific |
| LSTM +MLP[109] | classification results: adversary or not | Endesa,Spain | user-specific |
| AMIDS[110] | classification results: adversary or not | real data from University of Illinois | user-specific |
| Bollinger +POMDP [112] | optimal action: ignore or check & fix | synthetic data | general |
| SVR+POMDP [73] | optimal action: ignore or check & fix | synthetic data | general |
| continuous state POMDP [74] | optimal action: ignore or check & fix | synthetic data | general |
| OPF[78, 79, 81] [100, 121–123] | list of adversaries | real data from Brazilian | general |

narrow down the search zone of electricity theft to a certain number of customers.

*2) Group Change-Based Methods:* In these methods, an inspector box (containing a head meter and several submeters) is installed in each community, where the meter is also called an inspector [72]. The head inspector monitors all users in the community and is responsible for detecting the existence of electricity theft. The subinspectors can change the group of users under investigation arbitrarily and take charge of locating malicious users. The subinspectors are employed to perform inspection steps on different groups of users so that the search zone of electricity theft is constantly narrowed down. The goal of these detection methods is to find out all malicious meters [72].

*3) Behavior Approximation-Based Methods:* In these methods, a head meter called a central observer meter is installed in a community. Relationships between users' actual electricity consumptions and reported readings are modeled by linear or nonlinear functions, called users' behavior functions. Since the central observer meter's measurements are approximately equal to the summation of users' actual electricity consumption, the measurements can be further linked with users' reported readings via a linear or nonlinear system of equations.[4] By solving this system of equations, these detection methods can find out how much users' reported readings to deviate from their actual electricity consumptions. Besides identifying malicious users exactly, these detection methods also have other goals such as shortening detection time [144].

## B. Sensor Deployment-Based Detection

We show a typical topology of distribution networks in Fig. 17. As illustrated in the figure, many advanced sensors, called feeder remote terminal units (FRTUs), are deployed at some feeder nodes of the distribution networks for monitoring whether users commit electricity theft. However, in most cases, utility companies have limited budgets for deploying advanced sensors, usually expensive. For example, an FRTU is worth $9000, and a

[4]A system of equations is a collection of two or more equations with the same set of unknowns [143].

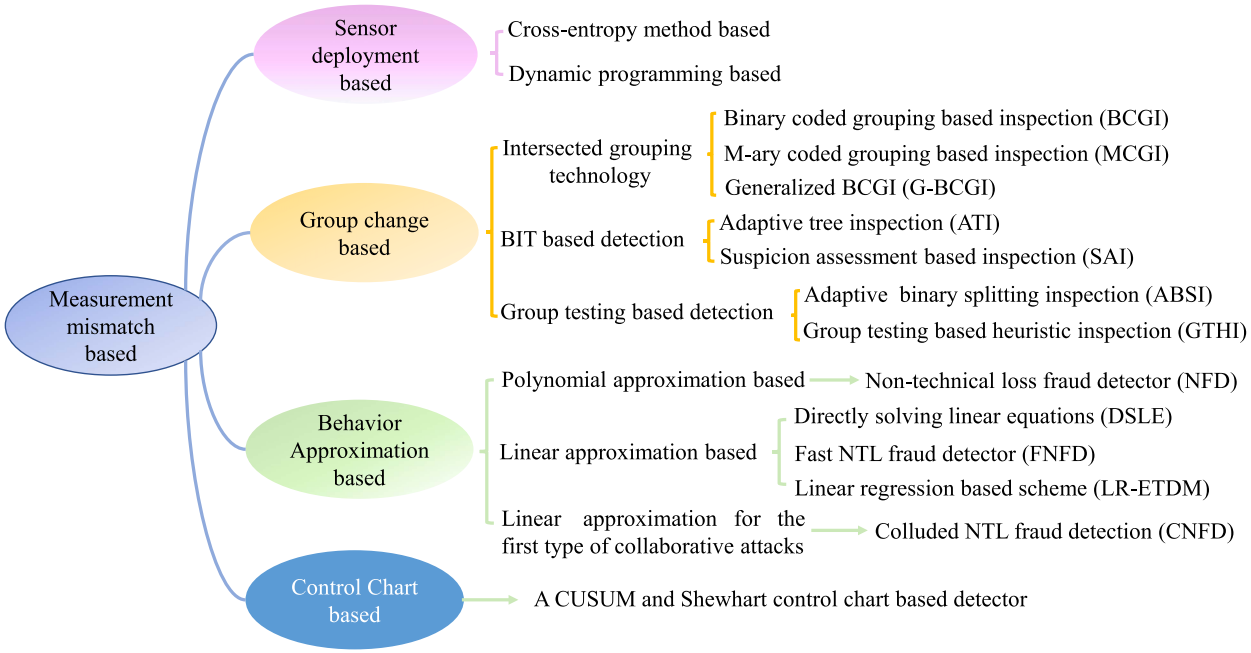


**Fig. 16.** ***Summary of measurement mismatch-based methods.***

digital protective relay (DPR) is worth $4750 [132], [134], [142]. Both FRTUs and DPRs are microprocessor-based electronic devices used for feeder fault detection in power systems. They can measure and analyze the following quantities (but not limited to): voltages, currents, power consumptions, and power factors [145], [146]. Due to budget limitations, researchers are motivated to find out an optimal sensor deployment solution that deploys the fewest advanced sensors and also can achieve a high anomaly coverage index (ACI), defined as follows.

Let $p_j$ denote the attacking probability of user $j$, which can be calculated based upon user $j$'s load profiles. Let $p^*$ and $n^*$ be two constants previously specified by utility companies, where $0 < p^* < 1$ and $n^*$ is a positive integer. Let $s_j$ denote the advanced sensor monitoring user $j$. Let $I(s_j)$ denote the set of users monitored by $s_j$. User $j$ is anomalous if $p_j > p^*$. User $j$ is effectively monitored if $|I(s_j)| \leq n^*$, where $|\cdot|$ denotes the cardinality of a set. Let $M_1$ denote the set of anomalous users, i.e., $M_1 = \{j|p_j > p^*\}$. Let $M_2$ denote the set of anomalous users being effectively monitored, i.e., $M_2 = \{j|j \in M_1, |I(s_j)| \leq n^*\}$. The ACI is the ratio of summation of attacking probabilities of anomalous users being effectively monitored to summation of attacking probabilities of anomalous users. Based on [132] and [142], we have

$$\text{ACI} = \frac{\sum_{j \in M_2} p_j}{\sum_{i \in M_1} p_i}.$$

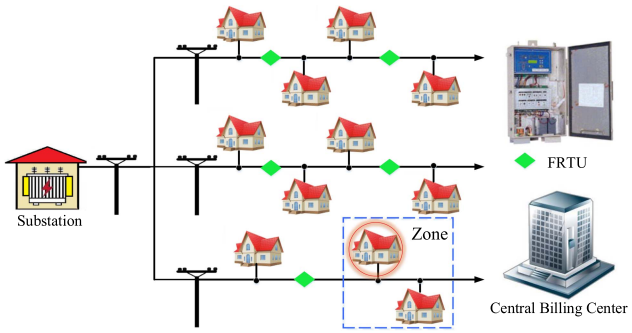


**Fig. 17.** ***Typical topology of distribution networks. At each user's premises, a smart meter is installed to report electricity consumption to utility companies' control centers for billing purposes. If FRTU's measurements are significantly larger than the summation of reported readings, then there is at least one malicious user among users monitored by this FRTU. For example, if the user highlighted in red is stealing electricity, only users in the highlighted zone are needed to be further examined [142].***

When the locations of advanced sensors are changed, the sets of users monitored by the sensors (i.e., $\{s|I(s) \leq n^*\}$) are changed, resulting in the change of $M_2$ and the ACI. Since a higher ACI implies that there are more users in $M_2$, a sensor deployment solution with a higher ACI can help narrow down the search zone to at most $n^*$ users with a higher probability. The goal of these detection methods is to find an optimal sensor deployment solution which is capable of: 1) achieving an ACI as high as possible and 2) requiring as few deployed advanced sensors as possible with the consideration of budget limitation. Existing related works mainly include cross-entropy (CE) algorithms and dynamic programming (DP) algorithms, explained as follows.

*1) CE Algorithms:* For optimizing the deployment of FRTUs, Liao *et al.* [132] model the topology of distribution

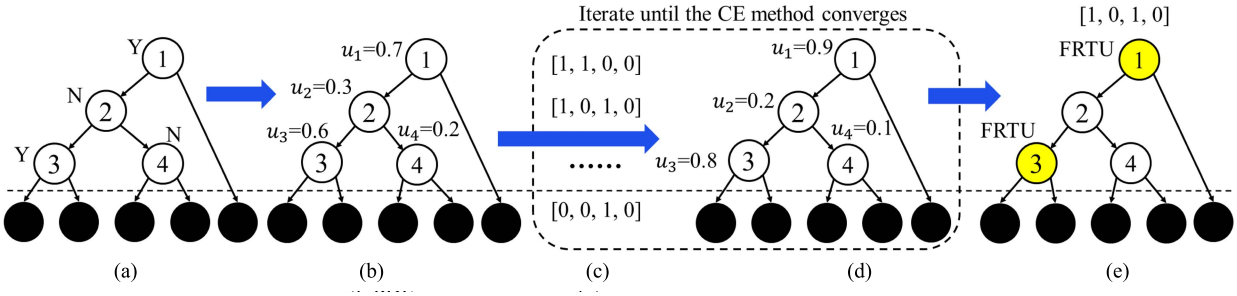


**Fig. 18.** *Example to illustrate the CE-based algorithm in [132] to optimize the FRTU deployment. The circles are internal nodes representing transformers in the distribution networks; the black nodes are leaf nodes representing users' smart meters. (a) Label nodes with Y/N. (b) Associate nodes with pdfs. (c) Generate possible solutions. (d) Update pdfs. (e) Output a solution.*

networks as a tree structure, as shown in Fig. 18. In this tree, leaf nodes and internal nodes represent smart meters and candidate locations (e.g., feeders) for installing FRTUs, respectively; lines connecting nodes represent power lines.

Assume that there are $n_0$ internal nodes in the tree. Then, there are a total number of $2^{n_0}$ possible FRTU deployment solutions. For avoiding enumerating all these solutions, Liao *et al.* [132] first apply a conditional random field (CRF) to assign each internal node with a label Y/N, as shown in Fig. 18(a). The CRF is an undirected graphical model whose nodes can be divided into two disjoint sets of observed and output variables. In [132], the observed variables are the network topology and users' historical load profiles; the output variables are a sequence of Y/N labels, which is used to initialize a Monte Carlo method, i.e., the CE method. Specifically, as shown in Fig. 18(b), a node with a label Y/N is associated with a Gaussian distribution's probability density function (PDF) with a larger/smaller mean $\mu$ and a smaller/larger variance, respectively, indicating that this node has a larger/smaller probability of being added with an FRTU. If there are some existing FRTUs in the distribution networks, these nodes are assigned with a probability of 1.

Afterward, the CE method proceeds iteratively with the following two steps.

1) Each node draws a sample $a$ from the associated pdf. If $a < 1 - \mu$, then the value of $a$ is reset as 0; otherwise, the value of $a$ is reset as 1. The binary number 0/1 indicates that an FRTU will/would not be installed on this node. As shown in Fig. 18(c), all nodes' binary numbers form an FRTU installation solution.
2) A set of such solutions are evaluated. Specifically, if the user-defined ACI constraint is not satisfied, the corresponding FRTU deployment solution is discarded. Otherwise, for the top 10% solutions with the highest ACIs, pdfs of candidate nodes are updated with an increasing mean and a decreasing variance, as shown in Fig. 18(d). The new pdfs are expected to produce better samples in the next iteration [147].

The above two steps iterate until the CE method converges. The algorithm finally outputs a heuristically optimal FRTU deployment solution that satisfies the user-defined ACI constraint with the fewest number of FRTUs. The solution is shown in Fig. 18(e), where the output solution is denoted by a permutation $[1, 0, 1, 0]$, which means to install two FRTUs at nodes 1 and 3, respectively.

However, the algorithm is essentially a stochastic optimization approach. Since it considers the entire distribution network at each iteration, the runtime is prohibitively long for large-scale distribution networks. Thus, the above algorithm suffers from a scalability issue [142]. Experiment results show that the algorithm cannot handle a network of over 500 customers [142]. Furthermore, the nondeterministic nature (i.e., the Gaussian distribution assumption) makes it difficult to maintain solution quality. For a given distribution network, under different runs, the CE algorithm may output different solutions, which requires deploying a different number of FRTUs and achieves different ACIs. For distribution networks with different topologies, the output solutions of the CE algorithm may achieve different ACIs. Experiment results in [142] show that the solution quality of the CE algorithm can be improved greatly.

*2) DP Algorithm:* The DP simplifies a complicated problem by breaking it down into simpler subproblems in a recursive manner. The complicated problem is required to have overlapping subproblems and optimal substructure [148]. Zhou *et al.* [142] propose to apply the DP method to optimize the FRTU deployment solutions in the distribution networks. Through examining one part of the network at each step, the DP-based algorithm produces an optimal FRTU insertion solution in a bottom-up fashion within deterministic polynomial time, as demonstrated in the following.

First, users' attacking probabilities are calculated based on historical profiles during one year, as shown in Fig. 19(a). Then, to decide how FRTUs should be deployed on a given internal node's subtree, a merging method is applied to combine all FRTU insertion solutions of its left

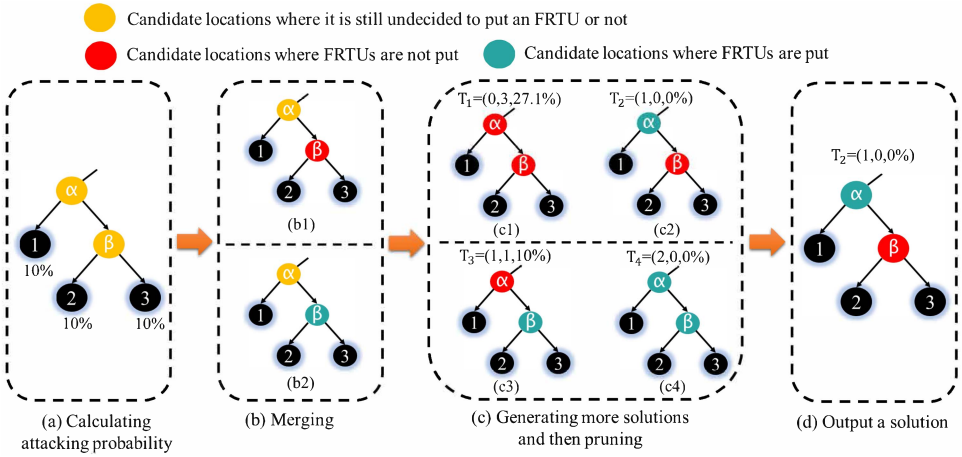


**Fig. 19.** ***Example illustrates the DP algorithm [142] for optimizing the FRTU deployment. (a) Calculating the attacking probability. (b) Merging. (c) Generating more solutions and then pruning. (d) Output a solution. Note that in (b1), the DP algorithm tries not to install an FRTU at node $\beta$, but it is still undetermined whether to install an FRTU at node $\alpha$ or not. Solutions (c1) and (c2) are derived from (b1). Specifically, in (c1), the DP algorithm tries not to install an FRTU at node $\alpha$, whereas in (c2), the DP algorithm tries to install an FRTU at node $\alpha$. In contrast, in (b2), the DP algorithm tries to install an FRTU at node $\beta$, but it is still undetermined whether to install an FRTU at node or not $\alpha$. Solutions (c3) and (c4) are derived from (b2). Specifically, in (c3), the DP algorithm tries not to install an FRTU at node $\alpha$, whereas in (c4), the DP algorithm tries to install an FRTU at node $\alpha$.***

and right subtrees. Taking Fig. 19 as an example, we consider how FRTUs should be deployed on node $\alpha$'s subtree, which consists of the following two subtrees: 1) a left subtree rooted by a leaf node representing user 1 and 2) a right subtree rooted by internal node $\beta$, which contains users $\{2, 3\}$. Since an FRTU cannot be put on a leaf node, but an internal node (e.g., node $\beta$), after the merging process of the above two subtrees, we can get two solutions, as shown in Fig. 19(b). In Fig. 19, yellow nodes represent candidate locations where it is still undecided to put FRTUs or not; green/red nodes represent candidate locations where FRTUs are/are not put, respectively.

After the merging process, from each merged solution, we can derive two solutions by installing or not installing an FRTU at the root of the merged solution. In this way, more solutions are generated. For example, by installing or not installing an FRTU at the root of solution (b1), we can derive solutions (c1) and (c2). Similarly, from solution (b2) we can derive solutions (c3) and (c4). Thus, from the two solutions in Fig. 19(b), we can derive four solutions in Fig. 19(c).

Let $T = (d, l, p)$ denote a solution at some internal node, where $d$ and $l$ are the number of FRTUs and the number of uncovered leaves (i.e., users that are not monitored by FRTUs), respectively, and $p$ denotes the attacking probability of the internal node. Let $p_l$ and $p_r$ denote the attacking probability of this internal node's left and right children, respectively. Then, we have $p = p_l + p_r - p_l p_r$. Particularly, if there exists an FRTU at an internal node, the attacking probability of this node is set to zero. For example, in Fig. 19(c3), since the attacking probability of user 1 is 0.1, we have $p_l = 0.1$. Since there is an FRTU installed on node $\beta$, we have $p_r = 0$ and $d = 1$. Thus, the attacking probability of root node $\alpha$ can be calculated as $p = 0.1 + 0 - 0.1 * 0 = 0.1$. It is obvious that, since only user 1 is not monitored, we have $l = 1$. Thus, the solution in Fig. 19(c3) can be denoted as $T_3 = (1, 1, 10\%)$.

For reducing computational overhead, a pruning method is then applied to remove invalid or inferior solutions, as explained in the following. A solution is said to be invalid if it meets either of the following criteria: 1) the number of users on the root's subtree is more than $n^*$ and 2) the attacking probability of the root node exceeds a predefined threshold $p^*$, but no FRTU is installed at the root node. As aforementioned, $n^*$ and $p^*$ are two previously specified parameters in the definition of ACI. For example, if we assume $n^* = 3$ and $p^* = 15\%$ in Fig. 19, since the root node of solution (c1) has an attacking probability of $27.1\% > p^* = 15\%$ and no FRTU is installed on the root node, we can infer solution (c1) is invalid. In addition, given two solutions $T = (d, l, p)$ and $T' = (d', l', p')$, $T$ is inferior to $T'$ if and only if the following criteria are all satisfied: 1) $d \geq d'$; 2) $l \geq l'$; and 3) $p \geq p'$. For example, in Fig. 19(c), since $T_2 = (1, 0, 0\%)$, $T_3 = (1, 1, 10\%)$, and $T_4 = (2, 0, 0\%)$, we can easily know that both solutions (c3) and (c4) are inferior to solution (c2). In the pruning process, all invalid and inferior solutions are discarded. If there are multiple valid solutions remain, only the one with the fewest installed FRTUs is the output. For example, in Fig. 19(c), since solution (c1) is invalid and solutions (c3) and (c4) are inferior, they are all discarded. As shown in Fig. 19(d), only solution (c2) is the output.

To sum up, in [142], the subproblem in DP is to obtain an optimal FRTU deployment solution for a certain part of the distribution network. The DP algorithm constantly performs the following three steps: merging solutions of two internal nodes' subtrees, generating more solutions based on the merged solutions, and pruning the invalid and inferior solutions. Then, the DP algorithm outputs FRTU deployment solutions, represented by subtrees with

more and more leaf nodes (users). In other words, solutions produced later cover a larger part of the distribution network than those produced earlier. The above procedure proceeds recursively until the algorithm obtains an approximately optimal FRTU deployment solution across the whole distribution network. Since, in each step, the algorithm always outputs a solution that satisfies the user-defined $n^*$ and $p^*$ constraints in ACI and requires the fewest FRTUs for that part of the distribution network, the final output solution satisfies the above constraints regarding ACI and requires the fewest FRTUs for the whole distribution network. In other words, the output solution is approximately optimal.

As aforementioned, in [132] and [142], users' attacking probabilities are calculated using readings across a whole year. It is obvious that, in this way, attacking probabilities remain constant and cannot accurately reflect changes in users' electricity theft patterns. For example, a user who seldom stole electricity before may increase the frequencies of electricity theft behaviors. To address this limitation, Zhou *et al.* [134] propose to calculate users' attacking probabilities with readings during moving time windows of fixed sizes. In this way, users' attacking probabilities vary over time. To capture this characteristic, the authors further adopt an anomaly rate range for each user, which is defined as an interval between minimum and maximum attacking probabilities of this user across all time slots [134]. Based on users' anomaly rate ranges, Zhou *et al.* [134] adapt the above DP-based algorithm to deploy the minimum number of DPRs into the distribution network of a typical multitenant data center to monitor the energy usages of users, i.e., tenants of the data center [149].

As summarized in Fig. 16, the sensor deployment-based detection solutions mainly include CE-based algorithms and DP-based algorithms. All of these algorithms aim to find out an approximately optimal sensor deployment solution, which requires to deploy the fewest advanced sensors while achieving an ACI as high as possible. A sensor deployment solution with a higher ACI implies that it can help narrow down the search zone of electricity theft to a certain number of users with a higher probability. Zhou *et al.* [142] claim that the DP-based algorithms can produce FRTU solutions more quickly and have a better performance in scalability than the CE-based algorithms.

## C. Group Change-Based Detection

Group change-based methods require installing an inspector box at some places (e.g., in a distribution room or on an electrical pole) in a community [72]. Xiao *et al.* [72] first propose a preliminary design of the inspector box (as shown in Fig. 20), which is also used in [139]–[141]. The inspector box contains one head inspector, and at least one subinspector, both of which are essentially function-enhanced smart meters with stronger computation capability and larger storage space [72], [139]–[141].

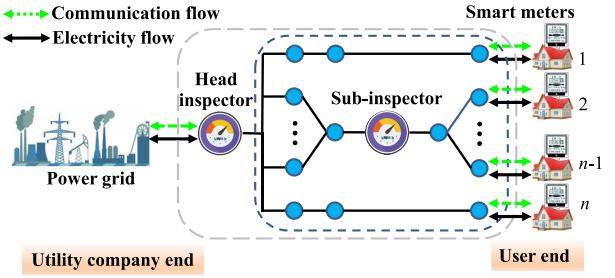


**Fig. 20.** *Preliminary idea of the inspector box [72].*

The head inspector monitors all users in the community and is responsible for detecting whether there are users stealing electricity theft. The subinspectors take charge of locating adversaries exactly and are featured with the following characteristics: 1) a power line is connected to an end-user at all times to prevent outages; 2) subinspectors can be effortlessly added or removed; and 3) users monitored by subinspectors can be arbitrarily changed manually or automatically [139].

The goal of this category of detection methods is to locate all adversaries within the shortest detection time by strategically changing the group of users inspected by subinspectors [68], [72], [137]–[141]. Since each inspection lasts for one reporting period (usually 15 min), this goal can be abstracted as minimizing the number of inspection steps. Existing related works can be roughly classified into the following three categories: 1) intersected grouping detection methods; 2) binary inspection tree (BIT)-based detection methods; 3) and group testing-based detection methods, as explained in the following.

*1) Intersected Grouping Detection:* The basic idea of the intersected grouping detection [68] is to use multiple inspectors to commonly monitor the intersection set of different groups of users [4]. As shown in Fig. 21(a), inspector A monitors users $\{a, b, c, d\}$, and inspector B monitors users $\{c, d, e, f\}$. If only inspector A detects reading anomalies, the search zone can be narrowed down to users $\{a, b\}$. However, Liu *et al.* [68] just propose a preliminary idea, and they do not clearly demonstrate how customers are grouped to different inspectors.

To address this limitation, Xia *et al.* [137], [138] propose a BCGI algorithm, which groups users to different inspectors according to users' unique decimal identifications (IDs), which are converted to binary IDs. These binary IDs' lengths are equal to the number of inspectors monitoring these users. Specifically, inspector $i$ is assigned with users whose binary IDs have a digit "1" at the $i$th bit from the rightmost. For example, in Fig. 21(b), users 1, 3, and 5 are assigned to inspector 1 due to the digit "1" at the rightmost of their binary IDs. If an inspector detects a reading anomaly, this inspector is called "dirty," and its state is denoted by digit "1"; otherwise, this inspector is called "clean," and its state is denoted by digit "0." Under

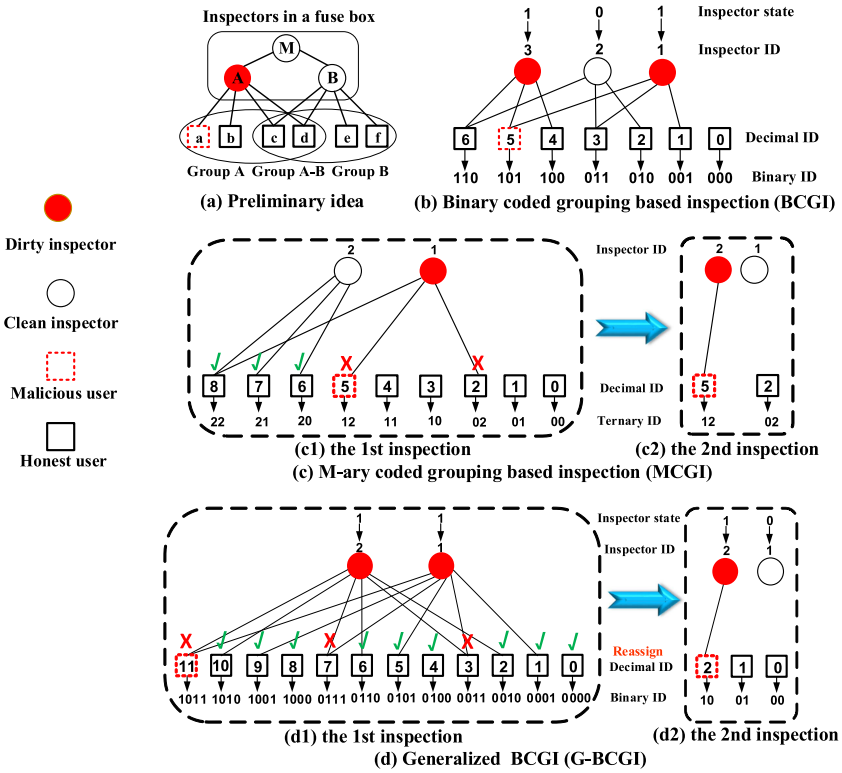


**Fig. 21.** *Examples to the intersected grouping detection methods: (a) Preliminary idea [68]; (b) BCGI [137], [138]; (c) MCGI [138]: (c1) first inspection and (c2) second inspection; and (d) G-MCGI [138]: (d1) first inspection and (d2) second inspection. Note that users with crosses X above will be inspected at the next inspection step, and users with checks √ above will not be inspected.*

the assumptions that: 1) there is only one malicious user among all users in a community and 2) there are enough subinspectors, and the unique malicious user is identified as the user whose binary ID is the same as inspectors' states. For example, in Fig. 21(b), states of inspectors "101" indicate that user 5 is malicious.

The BCGI algorithm has the advantage that it can identify the unique malicious user with just one inspection step. However, it requires $\Theta(\log_2(n))$ inspectors to monitor $n$ users [138]. To deal with cases where inspectors are inadequate (e.g., due to the budget limitation), Xia *et al.* [138] further propose an M-ary coded grouping-based inspection (MCGI) and a generalized BCGI (G-BCGI) algorithm, which takes multiple inspection steps to identify the unique malicious user, as introduced in the following.

In the MCGI algorithm, each user is assigned with a unique decimal ID, which is then converted to a number with base $m$, called the $m$-ary notation, where we have $m \geq 2$ and the $m$-ary notation's length equals to the number of inspectors monitoring these users. Let $\alpha$ denote the unique malicious meter. For any $i$, let $I_i$, $U_i$, and $U$ denote the state of inspector $i$, the users monitored by inspector $i$, and the set of users to be inspected, respectively. In different inspection steps, the MCGI algorithm groups users according to different digits of their IDs' $m$-ary notations. Specifically, at the $j$th inspection step, $U_i$ equals the set of users whose $i$th digit from the rightmost of $m$-ary notation is $m - j$. The following rule $R1$ holds.

*R1:* Based on inspector $i$'s state (i.e., $I_i =$ dirty or $I_i =$ clean), the search zone $U$ is constantly narrowed down until $\alpha$ can be identified; $I_i =$ dirty if only if $\alpha \in U_i$; and if there is at least one dirty inspector, $\alpha \in \cap\, U_i$, $I_i =$ dirty; otherwise, $\alpha \in (U - \cap\, U_i)$. For example, in Fig. 21(c), at the first inspection, since ternary notations of users $\{2, 5, 8\}$ and users $\{6, 7, 8\}$ have a common digit "2" at the first bit and the second bit from the rightmost, they are grouped to inspectors 1 and 2, respectively. Since inspector 1 is dirty and inspector 2 is clean, the unique malicious user must be among the users monitored by inspector 1, but not inspector 2. In other words, users $\{2, 5\}$ are potentially malicious. At the second inspection shown in Fig. 21(c), user 5 monitored by inspector 2 is identified as being malicious.

If there are not enough inspectors used in the BCGI algorithm, the BCGI algorithm can be used in multiple rounds, and this method is called the G-BCGI algorithm. The G-BCGI algorithm adopts the same user-inspector grouping strategy as in the BCGI algorithm. Let $s$ denote the number of inspectors available for monitoring users. When a new inspection begins, users in the search zone (denoted as $U$) are first assigned with new unique decimal IDs, then converted to new binary IDs. Unlike the BCGI algorithm in which the length (denoted as $L$) of users' binary IDs is set as $s$, in the G-BCGI algorithm, $L$ is determined by the number of users in the search zone (denoted as $|U|$). If $L > s$, inspectors are not adequate.

Specifically, the users whose new binary IDs have a digit "1" at the $i$th ($i \leq s$) bit from the rightmost are assigned to inspector $i$. For the users whose binary IDs have no digit 1 at the $s$ bits from the rightmost, they are not assigned to any inspectors. The rule $R1$ in the MCGI algorithm holds here too. For example, in Fig. 21(d), at the first inspection, $|U| = 12$, $L = 4$, and $s = 2$. Thus, the inspectors are inadequate since $L < s$. We have $U_1 = \{1, 3, 5, 7, 9, 11\}$ and $U_2 = \{2, 3, 6, 7, 10, 11\}$ according to the digit 1 of users at first and the second bits of binary IDs from the rightmost. For the users $\{0, 4, 8\}$, since there is no digit 1 at these two bits, they are not assigned to inspectors 1 and 2. Since both inspectors 1 and 2 are dirty, the unique malicious user must be among users $\{3, 7, 11\}$, which are commonly monitored by the two inspectors. Thus, at the second inspection step, only these three users are reprobed, which are reassigned with new unique decimal IDs 0, 1, and 2, respectively. At the second inspection step, binary IDs' lengths equal the number of inspectors (i.e., 2), which implies that the inspectors are adequate now. In this case, the G-BCGI algorithm is reduced to the BCGI algorithm, and thus, the unique malicious user is identified as the user whose newest binary ID is the same as inspectors' new states (i.e., "10").

To summarize, to identify a unique malicious user in a community, if there are enough inspectors, the BCGI algorithm can be applied, using just one inspection step; otherwise, the MCGI and G-BCGI algorithms can be applied using several more inspection steps. Essentially, both the MCGI and G-BCGI algorithms are generalizations of the BCGI algorithm, and the G-BCGI algorithm can identify the unique malicious user more quickly than the MCGI algorithm. We compare the BCGI, MCGI, and G-BCGI algorithms in Table 11 mainly regarding the times of assigning and coding users' decimal IDs and the times of grouping and inspecting these users. On the whole, the intersected grouping detection methods have the advantage that they can identify a unique malicious user very fast (even with just one inspection step). However, they cannot deal with cases where multiple malicious users coexist.

*2) BIT-Based Detection:* To deal with the cases where multiple malicious users coexist, Xiao *et al.* [72] propose a series of inspection algorithms based on a BIT whose leaf nodes and internal nodes represent users and potential inspection steps, respectively. If all users on a node's subtree are honest, the inspection on this node is called "clean"; otherwise, it is called "dirty." These inspection algorithms are generally performed in a top-down manner. Based upon previous inspection results (i.e., "clean" or "dirty"), these inspection algorithms determine how the following inspection proceeds. For example, in Fig. 22(a), after getting a "clean" inspection result at node $d$, we can infer that users $\{1, 2\}$ are honest, which means that no inspections should be further conducted on node $d$'s subtree. In addition, from the "dirty" inspection result on node $b$, we can infer that there is at least one malicious user among users $\{1, 2, 3, 4\}$. Since users $\{1, 2\}$ have been

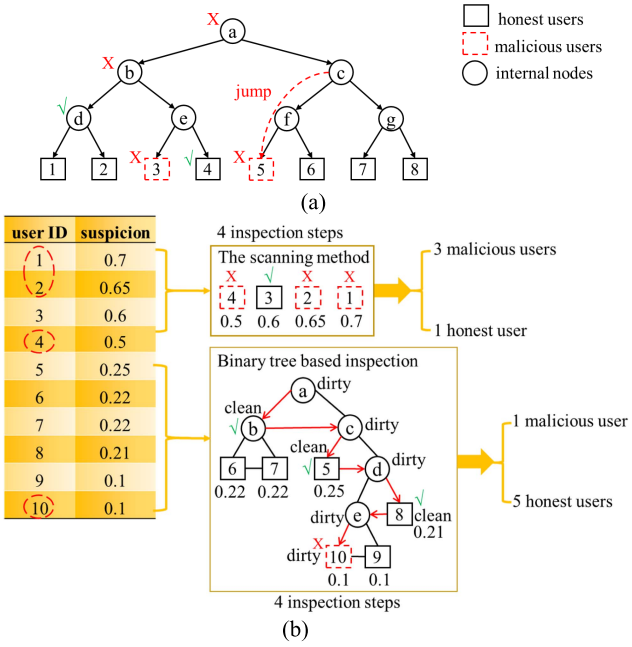


**Fig. 22.** ***Examples to illustrate the BIT-based inspection methods: (a) ATI algorithm [72] and (b) SAI algorithm [141]. Note that "dirty" and "clean" inspections are marked with "X" and "√," respectively.***

identified as being honest, we can further infer that at least one of users 3 and 4 is stealing electricity. Thus, the inspection on node $e$ can be skipped, and we can directly perform inspections on its children nodes. Xiao *et al.* [72] prove that the basic binary tree search algorithm is the fastest approach with complexity $O(\log_2(n))$ among all existing and future methods when there is one malicious meter and one inspector.

However, when the ratio of malicious users becomes larger, the binary tree method becomes worse due to the overhead of the logic tree structure and is worse than the scanning method when the ratio is very large [72]. There are two extreme cases: the binary search algorithm is the best when the ratio of malicious meters is small, and the scanning method is the best when the ratio is very high (such as near 100%). Thus, to achieve a balance between the scanning method and the binary search algorithm, Xiao *et al.* [72] further propose an adaptive tree inspection (ATI) algorithm to adaptively change between the scanning method and the binary search algorithm. Specifically, the ATI algorithm measures the ratio $\mu$ of malicious users and their arrangement, and if $\mu$ is smaller than a threshold, the ATI algorithm is the same as the binary tree scheme; otherwise, the algorithm proposes a jumping strategy to skip some tree nodes to go directly to lower levels of the tree to save the overhead. The number of skipped levels can be determined based on the heuristic information. For example, in Fig. 22(a), after node $c$, the jumping strategy is applied to skip a level containing nodes $f$ and $g$ and directly perform inspections at the bottom level containing users $\{5, 6, 7, 8\}$. Note that, in real applications, we believe that the ratio of malicious users

**Table 11** Comparison of BCGI, MCGI, and G-BCGI

| Algorithms | assigning decimal IDs | coding decimal IDs | grouping | inspecting |
|---|---|---|---|---|
| BCGI | one time | one time | one time | one time |
| MCGI | one time | one time | multiple times | multiple times |
| G-BCGI | multiple times | multiple times | multiple times | multiple times |

is usually small, and hence, the jumping usually does not occur. In this case, the ATI algorithm is simply reduced to the binary search algorithm. As mentioned before, when the ratio of malicious users is small, the binary search algorithms have the best performance in terms of efficiency.

Furthermore, the inspection algorithms in [72] are applied in both static cases where the malicious user set does not change and dynamic cases where the malicious user set changes during the inspection process. When applied in dynamic cases, the inspection methods have to run multiple rounds; and after finishing a round of inspection, the users that are so far honest need to be reprobed until there are no newly emerged malicious users.

Xiao *et al.* [72] treat all the users equally when conducting the inspection. However, electricity theft is essentially a particular form of economic crime [25]. Criminology studies [150]–[156] reveal the following facts: 1) more criminal prior records imply higher risks of recidivism in the future and 2) a longer time since the last criminal act implies a lower risk of recidivism. To further shorten the detection time, Xia *et al.* [141] propose a suspicion assessment-based inspection (SAI) algorithm, which assesses users' suspicions to steal electricity before conducting any inspection steps. On the whole, if users have more prior records of electricity theft, and the last electricity theft is committed more recently, these users have larger suspicions. In addition, larger deviations between users' reported and predicted normal electricity consumptions imply larger suspicions. For the users with the highest suspicions, they are probed first using the scanning method. For the remaining users, they are inspected with the help of a BIT, which is built according to the suspicions. Specifically, users with larger suspicions are closer to the root. The suspicions also determine the inspection order of the BIT's nodes. This means that, between two siblings, the node whose subtree has a larger mean of users' suspicions is first inspected. For example, in Fig. 22(b), since users $\{1, 2, 3, 4\}$ have the largest suspicions, they are inspected using the scanning method. On node $e$'s subtree, the average of users' suspicions equals 0.1, which is less than user 8's suspicion. Thus, the leaf node representing user 8 is inspected earlier than node $e$.

To sum up, compared with the intersected grouping detection methods, which can only deal with the unique malicious user setting, the BIT-based detection methods have the following advantages: 1) they can also be applied to multiple malicious user settings and 2) they can deal with dynamic cases. Thus, the BIT-based detection methods are more practical than the intersected grouping detection methods. Although the BIT-based detection methods have applied various strategies to save inspection steps, researchers are still finding more ways to shorten the detection time further.

*3) Group Testing-Based Detection:* The group testing problem was first introduced in World War II for accelerating and economizing the procedure of weeding out individuals infected with syphilis [157]. Electricity theft detection and group testing problems have many things in common: 1) the objects to be inspected/analyzed can be classified into two categories and 2) the goal is to conduct as few inspections/analyses as possible [139]. Specifically, the electricity theft detection problem aims to screen malicious users from honest users using as few inspections as possible. In contrast, the group testing problem aims to screen infected blood samples from pure blood samples with the fewest analyses [157].

Inspired from the above similarities, Xia *et al.* [139] propose a group testing method, called adaptive binary splitting inspection (ABSI) algorithm, to address the electricity theft detection. The ABSI algorithm has the assumption that the minimum upper bound of the number of malicious

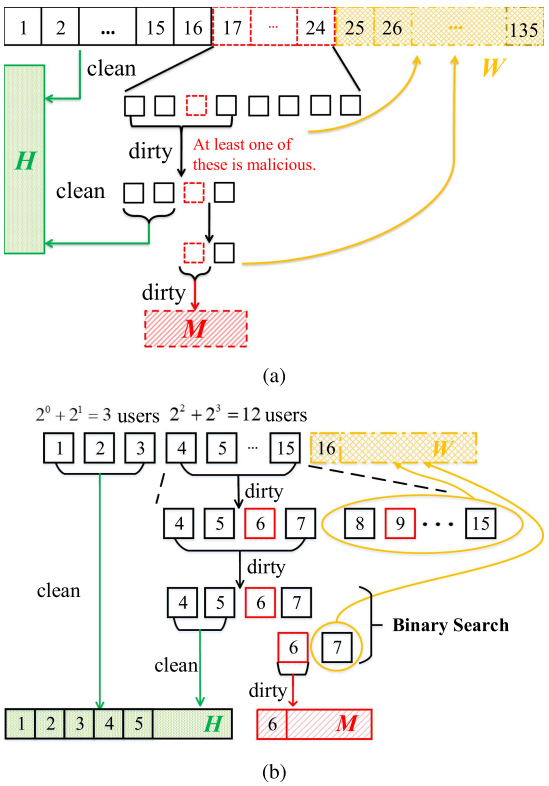


**Fig. 23.** *Examples to illustrate the group testing-based inspection methods: (a) ABSI algorithm [139] and (b) GTHI algorithm [140].*

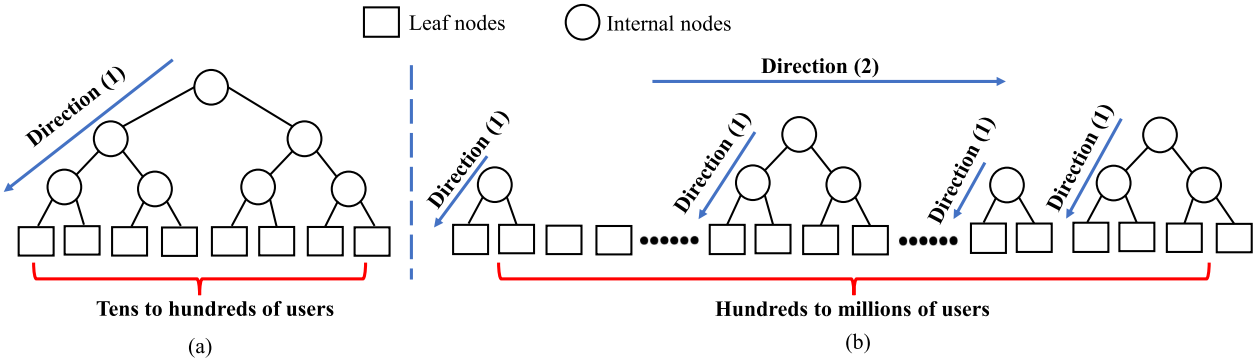


**Fig. 24.** *(a) BIT-based methods versus (b) group testing-based inspection.*

users can be known previously. Let $w$ denote the total number of users whose statuses (i.e., being malicious or honest) have not been determined yet. Let $m'$ denote the number of malicious users that have not been identified yet. When an inspection begins, $m'$ is constantly updated as the difference between the bound and the number of malicious users that have been identified. Among the users whose statuses have not been determined, if, on average, one user out of at least two users is malicious, i.e., $w \geq 2m'-1$, then $2^\theta$ users are extracted to be inspected using a binary search method until a malicious user are identified, where $\theta$ is a nonnegative integer determined by $w$ and $m'$. During this process, some honest users may also be identified. The remaining users are put back into the set of users whose statuses have not been determined yet and will be inspected later. If $w < 2m' - 1$, a scanning method is applied to probe users whose statuses have not been determined yet individually. Note that, after every round of binary search or scanning, the values of $w$ and $m'$ are updated immediately. Based upon the relationship between $w$ and $m'$, the ABSI algorithm adaptively adjusts its inspection strategy between the scanning method and the binary search method. For example, in Fig. 23(a), it is assumed that there are at most eight malicious users among a total number of 135 users. It is obvious that, at the first inspection, we have $w = 135$ and $m' \leq 8$, according to which the ABSI algorithm derives $\theta = 4$. Thus, 16 users are inspected at the first inspection step. Since the inspection result is "clean," these 16 users are identified as being honest, and the value of $w$ is updated as $135-16 = 119$. According to $w = 119$ and $m' \leq 8$, the ABSI algorithm derives $\theta = 3$. Thus, at the second inspection step, eight users are inspected. Since the inspection result is "dirty," these users are inspected using the binary search method. As shown in Fig. 23(a), with two more inspection steps, a malicious user is identified, and two honest users are identified. For the remaining five users, they are put back into the set of users whose statuses have not been determined yet.

To address the limitation that the ABSI algorithm requires to know the minimum upper bound of the number of malicious users, Xia *et al.* [140] propose a group testing-based heuristic inspection (GTHI) algorithm, which does not need to know any prior information. When the inspection process starts, the GTHI algorithm probes disjoint user sets of sizes $2^0 + 2^1$, $2^2 + 2^3$, $2^4 + 2^5, \ldots$, until an inspection result "dirty" is obtained. During the inspection process, it constantly estimates the malicious user ratio. If this ratio is larger than a specified threshold (usually chosen as $1/3$), then the GTHI algorithm adopts the scanning method. Otherwise, the GTHI probes $2^k + 2^{k+1}$ users, where $k$ is a nonnegative integer determined by the estimated malicious user ratio. If the inspection on these $2^k + 2^{k+1}$ users gets a "dirty" result, then the inspection proceeds as follows: 1) for the first $2^k$ users, they are further inspected using the binary search method and 2) for the remaining $2^{k+1}$ users, they are put back to the set of users waiting for further inspection. As shown in Fig. 23(b), at the first and second inspection steps, three and 12 users are inspected, respectively. Since a "dirty" inspection result is obtained at the second inspection step, for the users $\{4, 5, 6, 7\}$, they are further inspected by a binary search method. For the users $\{8, 9, \ldots, 15\}$, they are put back to the set of users waiting for further inspection.

We compare the ABSI algorithm with the GTHI algorithm. They both adopt two inspection strategies, i.e., a binary search method and a scanning method. Their differences are analyzed as follows. First, the ABSI algorithm requires the prior information of the minimum upper bound of the number of malicious users, whereas the GTHI algorithm does not. Second, in the ABSI algorithm, the group size is determined by the number of malicious users that have not been found out (i.e., $m'$) and the number of users whose statuses have not been determined yet (i.e., $w$). In contrast, in the GTHI algorithm, the group size is determined by the malicious user ratio.

Finally, we compare the BIT-based detection algorithms with the group testing-based detection algorithms in Fig. 24. First, the BIT-based algorithms group all users in one group (a big tree), and within the group (tree), the methods start the inspection process from the root node and run in a top-down manner, as direction (1) shown in Fig. 24(a). In applications, the BIT-based detection algorithms are usually applied when the total number of the inspected users in a community is not very large, such as tens to hundreds of users in a community. In contrast,

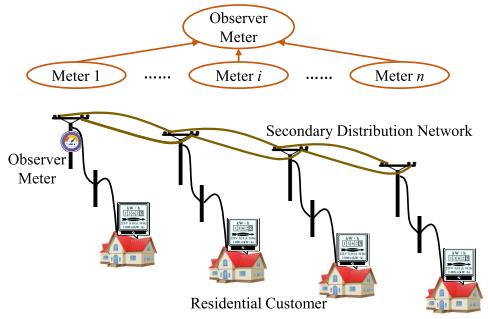


**Fig. 25.** ***Conceptual framework for behavior approximation-based methods [158], [159].***

the group testing methods group users into a series of groups to test one by one, as direction (2) shown in Fig. 24(b), and within each group, the methods also start the inspection process from the root node and run in a top-down manner, as direction (1) shown in Fig. 24(b). The group testing-based detection methods are usually applied when the total number of the inspected users in a community is large, such as hundreds to millions of users in the community. Second, comparing the searching methods inside one group/tree, we observe that both the BIT-based detection algorithms and the group testing-based detection algorithms adaptively adjust between the scanning method and the binary search method. The group testing-based detection algorithms [139], [140] were inspired by the BIT-based detection algorithms [72] in terms of searching algorithms. The BIT-based detection algorithms adopt a jumping strategy based on a threshold of the ratio of malicious meters so that the whole scheme is a true adaptive strategy in one binary tree. On the other hand, the group testing-based detection algorithms switch back and forth between two strategies (the scanning strategy without a tree and the binary search strategy with a tree) based on a threshold. Moreover, the binary searching in the BIT-based detection algorithms stops if finding all malicious meters in the tree. In contrast, the binary searching in the group testing-based detection algorithms stops if all the meters in the tree are clean or if one malicious meter is found and then puts undetermined meters to the later groups.

## D. Behavior Approximation-Based Detection

Behavior approximation-based methods require to use an existing central observer meter or install a new central observer meter in the community, which has similar functions with the head inspector, as shown in Fig. 25 [159]. In many situations, such a central observer meter already exists, e.g., a head meter in an apartment complex. In behavior approximation methods, relationships between users' reported and actual electricity consumptions are modeled by nonlinear or linear functions, called users' behavior functions [159]. The central observer meter's measurements, which approximately equal the summation of users' actual electricity consumptions, are then expressed as the summation of polynomial or linear functions whose input variables are users' reported readings. In this way, the measurements are related to users' reported readings via a system of equations whose unknown variables are the coefficients defining the polynomial functions or linear functions, respectively. Then, the problem of identifying malicious users is transformed into the problem of solving the system of equations for the unknown coefficients. Once the coefficients are obtained, the approximated behavior functions are also determined. Then, the malicious users can be identified by comparing these functions with some specified functions. Existing behavior approximation methods can be further classified into polynomial approximation and linear approximation methods as follows.

*1) Polynomial Approximation Based:* Han and Xiao [158], [159] propose a nontechnical loss (NTL) fraud detector (NFD), where the NTL fraud means electricity theft. Let $n$ denote the number of users in the community. Let $e_{i,j}$ and $x_{i,j}$ denote user $i$'s ($i = 1, 2, \ldots, n$) actual and reported electricity consumptions at period $j$, respectively. The NFD applies the Lagrange polynomial interpolation to approximate users' behavior functions, denoted by $e_{i,j} = f_i(x_{i,j}) = \sum_{k=q}^{0} a_{i,k} x_{i,j}^k$ for the $i$th user and the $j$th period. Let $E_j$ denote the central observer meter's measurement at period $j$ and $E_j$ equals to the summation of all users' actual electricity consumption at period $j$ as follows:

$$E_j = \sum_{i=1}^{n} e_{i,j} = \sum_{i=1}^{n} \sum_{k=q}^{0} a_{i,k} x_{i,j}^k \tag{1}$$

where $E_j$ and $x_{i,j}$ are available and coefficients $a_{i,k}$ are unknown. For $t$ periods with central observer meter's measurement and users' reported readings, a system of $t$ linear equations whose unknown variables are coefficients $a_{i,k}$ can be formed. If $t$ is large enough, the coefficients $a_{i,k}$ can be solved with various methods, such as matrix inversion either accurately or approximately. Afterward, all the polynomial functions can be determined. Let $\alpha_{\min}$ and $\alpha_{\max}$ denote two thresholds, where $0 < \alpha_{\min} < 1 < \alpha_{\max}$. Then, by comparing these polynomial functions with functions $y = \alpha_{\max} x$ and $y = \alpha_{\min} x$ on a coordinate plane, malicious users can be identified. Specifically, if a polynomial function is above $y = \alpha_{\max} x$, then the related user is malicious [e.g., users 2, 3, and 4 in Fig. 26(a)]; if the curve is between the lines of $y = \alpha_{\max} x$ and $y = \alpha_{\min} x$, then the related user is honest [e.g., user 5 in Fig. 26(a)]; and if the curve is below the line of $y = \alpha_{\min} x$, the related meter is defective [e.g., user 7 in Fig. 26(a)].

In general, a larger degree of the polynomial functions implies that users' behavior functions can be described more accurately. However, a large degree of the polynomial functions also implies a long detection time. Specifically, for a community with $n$ users, polynomial functions of

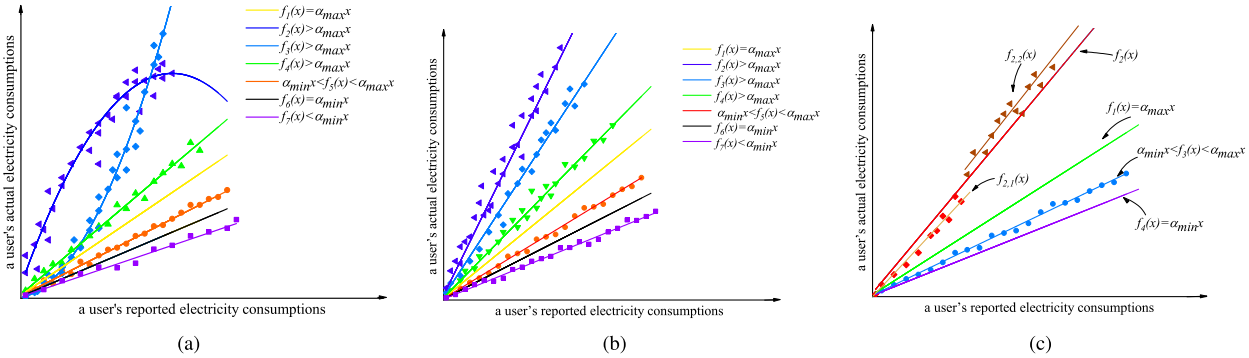


**Fig. 26.** ***Examples to illustrate behavior approximation-based detection. (a) Approximating users' behavior functions with polynomial functions [159]. (b) Approximating users' behavior functions with linear functions [144]. (c) Linear approximation for collaborative attacks [71]. Note that $f_i(x)$ denotes the polynomial or linear function approximating user $i$'s behavior function.***

degree $q$ imply a detection time lasting for at least $nq$ periods [158], [159].

*2) Linear Approximation Based:* To shorten the detection time, most existing works [144], [160], [161] choose to roughly approximate users' behavior functions with linear functions. Specifically, user $i$'s behavior function is approximated as $f_i(x_{i,j}) = \alpha_i x_{i,j}$. Based on how to solve the approximated users' linear functions, there are several different approaches as follows.

*a) Directly solving linear equations (DSLEs):* The central observer meter's measurement at period $j$, i.e., $E_j$, can be expressed as

$$E_j = \sum_{i=1}^{n} \alpha_i x_{i,j} \tag{2}$$

where $E_j$ and $x_{i,j}$ are available and coefficients $\alpha_i$ are unknown. With enough samples of $E_j$ and $x_{i,j}$, a system of linear equations can be formed based on the above equation and can be solved either accurately or approximately. With this idea, Bandim *et al.* [161] apply the matrix inversion to solve for the coefficients $\alpha_i$. Once the coefficients are known, the linear functions are determined.

*b) Fast NTL fraud detector (FNFD)-recursive least squares (RLSs):* Han and Xiao [144], [160] propose an FNFD to apply RLSs to find coefficients $\alpha_i$ such that the linear functions $f_i(x_{i,j}) = \alpha_i x_{i,j}$ can best approximate corresponding users' behavior functions. The RLS aims to minimize a weighted least-squares error function by appropriately selecting the coefficients $\alpha_i$ [162]. Specifically, the weighted least-squares error function for user $i$ is defined as the summation of squares of differences between user $i$'s actual electricity consumptions (i.e., $e_{i,j}$) and the corresponding approximated values (i.e., $\alpha_i x_{i,j}$) from the first period to the current period, i.e., $\sum_{j=1}^{\tau}(e_{i,j} - \alpha_i x_{i,j})^2$, where period $\tau$ is the current period. Assume that estimations of all coefficients $\alpha_i$ have already been obtained. Then, based on the theory of RLS, when a new period of central observer meter's measurement and users' reported readings are obtained, the estimation of $\alpha_i$ is updated. This process proceeds iteratively until the cost function for every user is minimized. Once the coefficients are known, the linear functions are determined. By comparing these linear functions with $y = \alpha_{\min} x$ and $y = \alpha_{\max} x$, malicious users can be identified. Specifically, if a line is above $y = \alpha_{\max} x$, then the related user is malicious [e.g., users 2, 3, and 4 in Fig. 26(b)]; if a line is between $y = \alpha_{\max} x$ and $y = \alpha_{\min} x$, then the related user is honest [e.g., user 5 in Fig. 26(b)]; and if a line is below $y = \alpha_{\min} x$, then the meter of related user [e.g., user 7 in Fig. 26(b)] is defective.

*c) Linear regression of energy theft and defective smart meters (LR-ETDM):* Yip *et al.* [163] propose a linear regression-based scheme for detection of energy theft and defective smart meters (LR-ETDM), to apply linear regression to find coefficients $\alpha_i$ such that the relationship between the central observer meter's measurement at any given period (i.e., $E_j$) and all users' reported readings at that period (i.e., $x_{1,j}, x_{2,j}, \ldots, x_{n,j}$) can be modeled as closely as possible. For achieving this purpose, the LR-ETDM chooses $E_j$ as the dependent variable and chooses $x_{1,j}, x_{2,j}, \ldots, x_{n,j}$ as the independent variables. The LR-ETDM aims to minimize the summation of squares of differences between measurement $E_j$ and the corresponding estimated value (i.e., $\sum_{i=1}^{n} \alpha_i x_{i,j}$) during a certain number of periods, i.e., $\sum_{j=1}^{T}(E_j - \sum_{i=1}^{n} \alpha_i x_{i,j})$, where $T$ is the number of periods for sampling enough measurements to solve for coefficients $\alpha_i$. Since the function is convex, coefficients $\alpha_i$ can be obtained by setting the gradient of the loss function to zero [164].

We compare the DSLE, FNFD, and LR-ETDM methods as follows. The DSLE method is an easy and fast approach, but the obtained linear functions are not optimal. On the other hand, both the FNFD and LR-ETDM methods try to optimize the linear functions to approximate the users' behaviors. The difference between the FNFD method and the LR-ETDM method is explained as follows. In the FNFD method, every user is considered individually with the user's behavior function and the cost/objective function. The FNFD method models the relationship between every user's actual electricity consumptions (i.e., $e_{i,j}$)

and reported readings (i.e., $x_{i,j}$) during different periods. In contrast, in the LR-ETDM method (which was published one year later than the FNFD method), all the users are considered together with the objective function designed for all users. The LR-ETDM method models the relationship between the central observer meter's measurement at any given period (i.e., $E_j$) and all users' reporting periods at that period (i.e., $x_{1,j}, x_{2,j}, \ldots, x_{n,j}$).

*3) Linear Approximation for Collaborative Attacks:* Han and Xiao [71], [165] consider a new potential type of electricity theft, called CNTL frauds. A CNTL occurs when multiple malicious users commonly tamper with one meter so that the meter records less electricity than the consumed amount by the household [71]. According to how much malicious users' electricity behaviors are overlapped, the CNTL frauds are classified into the following categories: 1) segmented CNTL frauds where malicious users tamper with the meter at different time segments; 2) overlapped CNTL frauds where different malicious users manipulate the same meter with totally and partially overlapped times, called fully overlapped CNTL frauds and partially overlapped CNTL frauds, respectively; and 3) combined overlapped CNTL frauds that consist of both segmented and overlapped CNTL frauds [71].

To detect the collaborating fraudsters, Han and Xiao [71], [165] propose a collaborative NTL fraud detection (CNFD) algorithm. The CNFD consists of the following two phases.

1) NTL fraud detection, which aims to identify the tampered meter. In this phase, the CNFD applies the theory of RLS to solve for the coefficients $\alpha_i$, as done in the FNFD. If $\alpha_i > \alpha_{\max}$, then user $i$ is identified as being malicious.
2) Fraudster differentiation that differentiates multiple malicious users in the tampered meter $t$.

The basic idea is to calculate the coefficients of a tampered meter at every period, denoted by $\alpha_{t,j}$, and analyze the similarity of these values within a user-defined threshold $e$ [71]. According to the similarity analysis results, the CNFD can judge whether these values belong to different malicious users or not [71]. Specifically, let $S_t$ contain the output coefficients of all fraudsters tampering with meter $t$, which is initialized as $S_t = \{\alpha_t\}$. Let $\min(S_t)$ and $\max(S_t)$ return the minimum and maximum elements in $S_t$. If $\alpha_{t,j} - \max(S_t) \geq e$ or $\alpha_{t,j} - \min(S_t) \leq -e$, this $\alpha_{t,j}$ is considered to be not similar to any values already in $S_t$ and is then added into $S_t$. If, finally, $S_t$ contains multiple elements, these values are identified as belonging to different malicious users. Otherwise, there is only one malicious user, and this means that it is not a CNTL fraud. For example, in Fig. 26(c), since $\alpha_2 > \alpha_{\max}$, meter 2 is identified as the tampered meter. After calculating user 2's coefficients at every period and then analyzing these values' similarities, the CNFD identifies that meter 2 is first tampered with by two malicious users successively.

To sum up, compared to the cases where users' behavior functions are approximated by linear functions, polynomial functions can approximate users' behavior functions more accurately. However, polynomial functions with higher degrees indicate that more periods of measurements and users' reported readings are needed to solve for the coefficients. This implies a longer detection time. Most existing works apply the linear function to approximate users' behavior functions. To solve for the coefficients in the polynomial or linear functions, various methods, such as matrix inversion, multiple linear regression, and RLS, can be applied.

## E. CUSUM and the Shewhart Control Chart-Based Detector

Following papers [159], [163], Xia *et al.* [69] assume that a central observer meter is installed in a community to measure the total electricity supplied to all users. Under this assumption, they propose a control chart-based detector that detects both minor difference and major difference attacks. Specifically, this detector jointly applies the cumulative sum (CUSUM) control chart and the Shewhart control chart, two commonly used statistical tools for change detection to analyze users' reported readings and the central observer meter's measurements.

Both the CUSUM and the Shewhart control charts have a centerline and two control limits. The CUSUM control chart plots the CUSUM of deviations between sample values and a target value of a process variable of interest in time order [166]. If the CUSUM of deviations exceeds one of the two control limits, the CUSUM control chart claims that this process variable is affected by some special causes of variation. Since the CUSUM control chart incorporates all the information in multiple consecutive samples, it can efficiently detect small changes in the process. Thus, Xia *et al.* [69] apply the CUSUM control chart to detect the minor difference attacks. In contrast, the Shewhart control chart plots the sample values of process variables of interest in time order. If the last sample value exceeds one of its control limits, it claims that this process variable is affected by some special causes of variation. Since the Shewhart control chart uses only one sample and ignores other information given by the entire sequence of samples, it is insensitive to detect small changes. Still, it can detect large changes in a process more quickly than the CUSUM control chart. Thus, Xia *et al.* [69] employ the Shewhart control chart to detect major difference attacks.

The control chart-based detector in [69] consists of an electricity theft detection phase and a malicious user ID phase. In both phases, the control charts' parameters need to be first estimated based upon historical readings of smart meters and measurements of the central observer meter. In the electricity theft detection phase, the goal is to detect the existence of electricity theft. In this phase, the Shewhart and the CUSUM control charts are applied to analyze the difference between the central

observer meter's measurements and the summation of users' reported readings. If the Shewhart control chart detects reading anomalies, it indicates the existence of at least one malicious user launching major difference attacks and/or several malicious users launching minor difference attacks. If the CUSUM control chart detects reading anomalies, it indicates the existence of at least one malicious user launching minor difference attacks. About the malicious user ID phase, it aims to locate malicious users exactly. In this phase, the above two control charts are combined to analyze every user's daily electricity consumption. On the whole, if the Shewhart/CUSUM control chart detects reading anomalies, the corresponding user is a malicious user launching major/minor difference attacks.

### F. Summary and Comparison

In Table 12, we compare the above measurement-mismatch-based methods mainly from the following aspects: 1) what advanced sensors do the detection methods require to deploy in the distribution network? 2) what are the input data of the detection methods? 3) what are the data-level attack models adopted by the detection methods? and 4) what are the metrics mainly used to evaluate these detection methods?

## VI. MACHINE LEARNING-BASED VERSUS MEASUREMENT MISMATCH-BASED METHODS

### A. Overall Comparison

As previously discussed, existing electricity theft detection methods can be roughly classified into machine learning-based and measurement mismatch-based methods. The machine learning-based methods apply currently popular machine learning methods to analyze meter readings and/or other customer-related data, aiming to find abnormal electricity consumption patterns that are highly related to electricity theft. In contrast, the basic idea of the measurement mismatch-based methods is to deploy advanced sensors in the distribution networks. By comparing measurements from sensors with reported readings of customers' smart meters, the methods can narrow down the search zone of adversaries until finally pinpoint them.

We compare machine learning-based and measurement mismatch-based methods in Table 13. As shown, the machine learning-based methods do not require to install extra advanced sensors. Hence, their deployment costs are moderate. The input is the data spontaneously generated in customers' daily life, mainly including load profiles, prior records of electricity theft, and some other user-related attributes, such as geographical locations and tariff categories. Usually, the output may be a list of adversaries, the classification results of a specific customer, or optimal actions (ignore or check/fix), according to the types of machine learning-based methods. The computation involved in the machine learning-based methods is usually intensive and performed in a centralized manner. That is to say, almost all the computation is performed in the control center of utility companies. Based upon the summary in Tables 9 and 10, the machine learning-based methods have been used to detect reduced consumption attacks (including both major and minor differences attacks), neighbor pricing attacks, and the second type of collaborative attacks. These methods usually have a moderate to high detection accuracy and a high FPR. The parameters of these methods are usually trained with users' historical data in an off-line fashion. The machine learning-based methods suffer from the limitations, such as data imbalance, data shifts due to nonmalicious factors, and novel attack models. We will explain these limitations in detail later.

In contrast, with regard to the measurement mismatch-based methods, they are usually required to apply different types of advanced sensors, which mainly includes, but are not limited to, redundant smart meters, FRTUs, DPRs, inspector boxes, central observer meters, and so on. Thus, the deployment cost of this category of detection methods usually ranges from moderate to high. The input mainly includes users' reported electricity consumptions, advanced sensors' measurements, and/or topology of distribution networks. The output of the measurement mismatch-based methods is either an optimal sensor deployment solution or a list of adversaries. Specifically, comparing the advanced sensors' measurements with summations of users' reported readings, these detection methods can either narrow down the search zone of electricity theft to a certain number of users (e.g., sensor deployment-based methods) or, finally, locate adversaries (e.g., group change- and behavior approximation-based methods) in an online fashion. The computation involved in the measurement mismatch-based methods is usually light-weighted and performed in a distributed manner (i.e., in the advanced sensors). Based upon the summary in Tables 9 and 10, the measurement mismatch-based methods have been used to detect reduced consumption attacks (major differences attacks). These methods usually achieve a high detection accuracy and a low FPR. The measurement mismatch-based methods suffer from the limitations such as high deployment cost, limited coverage of (data level) attack models, and threshold selection problems. We will explain these limitations in detail later.

### B. Limitations of Machine Learning-Based Methods

The machine learning-based suffer from the following limitations.

*1) Data Imbalance:* The data imbalance issue means that, among the collected samples, there are usually plenty of benign samples of honest customers but very few abnormal samples of malicious users [89], [92], [101], [103]. Since machine learning methods usually require that the number of abnormal samples should be

**Table 12** Comparing Measurement-Mismatch-Based Methods

| Algorithm | Advanced sensors | Input | Output | Main metrics | Notes |
|---|---|---|---|---|---|
| CE-based algorithm [132]; DP-based algorithm [142] | multiple FRTUs | topology of distribution networks; users' historical load profiles | an optimal sensor deployment solution | # of inserted FRTUs, ACI | The DP-based algorithm deploys 18.8% less FRTUs on average than the CE-based algorithm |
| BCGI, MCGI, G-BCGI [139] | multiple inspectors | inspectors' measurements; users' reported electricity consumptions | the unique adversary | # of inspection steps | Pros:very fast; Cons: only applicable to the unique-malicious-user setting |
| ATI [72] | an inspector box | inspectors' measurements; users' reported electricity consumptions | a list of adversaries | | SAI is faster than ATI with some overhead |
| SAI [141] | | prior records of electricity theft; historical local profiles; users' reported readings | | | |
| ABSI [139] | | minimum upper bound of the # of malicious users; users' reported readings | | | ABSI and GTHI have the same # of inspection steps; but GTHI is more practical. |
| GTHI [140] | | users' reported readings | | | |
| NFD [159], DSLE [161], FNFD [160], LR-ETDM [163] | a central observer meter | central observer meter's measurements; users' reported reading | | detection time; accuracy | NFD is polynomial approximation-based and others are linear approximation-based. |
| CNFD [71] | | | | | CNFD detects the first type of collaborative attacks; and the control chart based detector detects both minor and major difference attacks. |
| Control chart based detector [69] | | | | | |

Note that all existing measurement mismatch-based methods except the control chart method are used to detect reduced consumption attacks (major difference).

almost equal to the number of benign samples, the data imbalance issue adversely impacts the detection accuracy. As reported in [20], [77], and [101], many machine learning-based methods have a relatively low detection accuracy, i.e., about 60%–70%. To address this limitation, there have been several works conducted on the data-level and/or algorithm level in recent years to mitigate the adverse impacts incurred by the data imbalance issue. The data-level approaches resample the original dataset to make it more balanced; the algorithm-level approaches modify existing classification algorithms to make them more appropriate for imbalanced datasets. For example, Zanetti *et al.* [92], Rodriguez *et al.* [101], and Martino *et al.* [103] propose to apply one-class SVM, an algorithm level approach, which only needs a normal consumption pattern in the training phase to identify malicious users. However, its performance in terms of recall, precision, and F-measure is not very good. For further improving the performance, Qu *et al.* [168] and Gunturi and Sarkar [169] first apply different resampling methods such as synthetic minority oversampling technique (SMOTE) to preprocess the datasets and then apply ensemble learning methods, such as random forests to combine different classification results of multiple classifiers.

*2) Data Shifts:* Data shifts are mainly due to nonmalicious factors, which mainly includes, but are not limited to, the change of electrical appliances and seasonality, and the normal moving in/out of residents [89]. These nonmalicious factors can make consumption patterns of honest users shift greatly. Since the machine learning-based methods identify malicious users by analyzing whether users' electricity consumption patterns vary significantly, the nonmalicious factors may mislead these detection methods to classify honest users as malicious users mistakenly. In other words, nonmalicious factors are one of the main reasons for relatively high FPRs [77], [89]. For decreasing the influence of nonmalicious factors, Jokar *et al.* [89] apply the $k$-means algorithm on a benign dataset to remove clusters with fewer members before training the SVM model. For reducing false positives, Jindal *et al.* [77] propose to use a decision tree to calculate customers' expected electricity consumptions, which are then used as an input for the SVM classifier. However, there is still a large space for improvement.

*3) Novel Attack Models:* On the other hand, indicate that with time going on, adversaries are likely to adopt new malicious consumption patterns that are not involved in the original training dataset. Since machine learning classifiers cannot recognize these new malicious consumption

**Table 13** Machine Learning- Versus Measurement Mismatch-Based Methods [27], [67], [77], [89], [139], [167]

| Detection methods | Machine learning-based | Measurement mismatch-based |
|---|---|---|
| advanced sensors | not required | required, mainly including smart meters, FRTUs, DPRs, inspector boxes, central observer meters |
| deployment cost | low to moderate | moderate to high |
| main input | data from customers, mainly including users' (historical and present) reporting electricity consumptions, prior records, and other user related attributes (e.g., geographical locations, tariff categories, etc.) | data from customers, mainly including users' (present) reporting electricity consumptions; measurements of advanced sensors; topology of distribution networks |
| possible output | a list of adversaries; classification results of a specific customer; optimal action | an optimal sensor deployment solution; a list of adversaries |
| computation characteristics | centralized; almost all computations are intensive and performed in the control center of utility companies | distributed; almost all computations are light-weighted and performed in the advanced sensors; |
| (data-level) attack models | reduced consumption (both major and minor difference) attacks, neighbor pricing attacks, the second type of collaborative attacks | reduced consumption (major difference) attacks |
| limitations | data imbalance, data shifts due to nonmalicious factors, novel attack models, data poisoning attacks, long detection time, parameter optimization, computational burden, threshold selection problems | high deployment cost, threshold selection problems |
| detection accuracy | moderate to high | high |
| FPR | high | low |
| mode | offline | online |

patterns, novel attack models apparently hurt the TPR on the machine learning-based methods. Therefore, adversaries are likely to escape the detection. Jokar *et al.* [89] assume that electricity theft patterns are predictable and apply several types of attack models to generate synthetic malicious user datasets. However, the synthetic attack dataset cannot cover all types of abnormal electricity consumption patterns, and new novel attack models are likely to appear.

*4) Data Poisoning Attacks:* Almost all machine learning-based methods implicitly assume that the detector is trained using samples with correct labels [170]. Assume that, in a community, there are customers committing electricity theft who has never been caught before. In this case, the data of these malicious customers will be mistakenly labeled as benign samples and then used for training the machine learning-based methods [170]. Such training that is based on false labels is referred to as a data poisoning attack, which results in a shift of the detector's decision boundaries and, hence, a deterioration of the detector's ability to distinguish benign from malicious data [170].

*5) Long Detection Time:* It usually takes a relatively long time (which may last even for several years) to collect customers' load profiles, which are then used to train the machine learning-based methods in an off-line fashion. Even after being trained, these detection methods still need multiple days of load profiles to determine whether customers are adversaries or not. In addition, inherited from the applied machine learning algorithms, these detection methods usually suffer from heavy computational burdens and parameter optimization problems in training phases [79].

*6) Threshold Selection Problems:* No matter what machine learning methods, intrusion detection methods, or statistical methods to be used, in the end, thresholds are used to judge good meters or bad meters. It is difficult to choose the best threshold for all these methods, and there is no optimal threshold since false-negative rates and false negative rates conflict with each other. Threshold selection is complex and often based on experiences or heuristic methods. Inappropriate thresholds end up with large false-negative rates or false-negative rates.

## C. Limitations of Measurement Mismatch-Based Methods

The measurement mismatch-based detection methods suffer from the following limitation.

*1) High Deployment Costs:* As shown in Table 13, these detection methods require to deploy advanced sensors, such as FRTUs, DPRs, inspector boxes, or central observer meters in distribution networks. Usually, these advanced sensors are expensive, for tens to thousands of dollars per device. Since there are many feeder nodes/communities in power systems across the world, it takes an astounding deployment cost to install advanced sensors across the whole distribution network. Some works [132], [134], [142] have been developed to optimize the sensor deployment such that the fewest number of advanced sensors is needed to be deployed. Other methods [72], [141] are on-depend features, and they can be used when the utilities feel to investigate the frauds. However, there is still some space for improvement to reduce the cost.

*2) Threshold Selection Problems:* For measurement mismatch-based methods, the advanced sensors perform the following functions periodically: 1) receiving reported electricity consumptions of users under investigation, 2) measuring the total amount of electricity distributed to users under investigation, or 3) comparing their measurements with summations of the received readings. Define

the thresholds in the measurement mismatch-based methods as the technical losses from the advanced sensors to all the users under investigation. If the differences between the advanced sensors' measurements and the summations of users' reported readings are larger than the thresholds, then there are adversaries among users under investigation, and more inspections will be performed on these users to identify at least one adversary. Otherwise, if the differences are smaller than the thresholds, all the users under investigation are honest. In applications, the technical losses are usually estimated based upon some existing mathematical models, some results of trial and error experiments, or experts' experiences [139].

However, there are always some biases between the estimated values and the true values of technical losses. If the technical losses (i.e., the thresholds) are estimated to be too large, the advanced sensors may not detect the existence of electricity theft when there are adversaries among users under investigation. In this case, the false-negative rate (i.e., the ratio of users mistakenly identified as honest users to the total number of actually honest users) increases. On the other hand, if the technical losses (i.e., the thresholds) are estimated to be too small, the advanced sensors may mistakenly detect the existence of electricity theft when there are no adversaries among users under investigation. In this case, the false positive rate (i.e., the ratio of users mistakenly identified as adversaries to the total number of real adversaries) increases. Thus, selecting appropriate thresholds is very important to the measurement mismatch-based methods to have good performances.

Our recommendation of detection methods for real applications is stated as follows. Utility companies can learn the tradeoff of the different methods in this article and, based on their budget, history handling, and current methods, choose an appropriate one customized to themselves.

## VII. FUTURE WORKS

We have analyzed how malicious users compromise their electricity meters. We have also reviewed existing electricity theft detections, which are classified into machine learning- and measurement mismatch-based methods. We also compare these two categories of detection methods, in terms of the input, output, deployment cost, detection accuracy, false rates, and so on. To complete our overview of electricity theft detection methods, we provide several future work directions in the following and then analyze technical challenges for developing efficient and accurate electricity theft detection methods.

### A. Enhanced Machine Learning- or Measurement Mismatch-Based Detection

As aforementioned, existing machine learning-based methods mainly suffer from a data imbalance issue and data shifts due to various nonmalicious factors, novel attack models, data poisoning attacks, long detection time, and threshold selection problems, whereas existing measurement mismatched-based methods suffer from a high deployment cost. Thus, in the future, we can develop better machine learning- or measurement mismatch-based detection algorithms that can overcome their limitations for achieving a higher detection accuracy and a lower false rate. Some specific directions are listed as follows.

1) Machine learning-based methods against the data imbalance issue should be further studied at the data level and/or the algorithm level. For example, generative adversarial networks that can generate new data with the same statistics as the training set may be a potentially effective tool at the data level. In addition, some algorithms that implicitly or explicitly preferentially favor the rare classes and, hence, tend to perform well on classifying rare classes are promising choices at the algorithm level. Specifically, cost-sensitive learning algorithms that place higher weights on the abnormal classes and/or increase/decrease the weights associated with the incorrectly/correctly classified examples are potential choices.
2) Machine learning-based methods against data shifts should be further studied. Specifically, to mitigate the adverse impacts of data shifts, some change-detection tests can be applied to explicitly detect changes in the statistics of customers' reported electricity consumptions. When data shift is detected, the current model is no longer up-to-date and must be substituted with a new one to maintain the prediction accuracy. Alternatively, the model should be continuously updated by retraining it with the most recently observed samples or enforcing an ensemble of classifiers. In addition, contextual information, such as the season information, the number of persons, and appliances, when available, can be added into the model for better explaining the causes of the data drift.
3) Machine learning-based detectors with fast detection time should be further studied. To achieve this purpose, few-shot learning-based detectors can be adopted since few-shot learning is an emerging method that aims to build accurate machine learning models with less training data.
4) Measurement mismatch-based methods to reduce the deployment cost of advanced sensors should be further studied. Although the CE method- and DP-based methods have been proposed in [132], [134], and [142] to optimize the deployment of advanced sensors (as introduced in Section V-B), the studied scenarios are limited, and there is still a space to improve the performance. In other words, reducing the cost due to the number of advanced sensors inserted in the distribution networks while increasing the ACI is worth studying in various system configurations.

5) As summarized in Table 9, existing machine learning-based methods have only been used to detect reduced consumption attacks (mainly major difference attacks), neighbor pricing attacks, and the second type of collaborative attacks. Thus, machine learning-based methods should be further studied against the first type of collaborative attacks, the intermittent attacks, the load profile shifting attacks, and even more complex attack models (e.g., the adversaries jointly launch several attack models simultaneously).
6) As summarized in Table 9, existing measurement mismatch-based methods have only been used to detect reduced consumption attacks (both major and minor difference attacks) and the first type of collaborative attacks. Thus, measurement mismatch-based methods should be further studied for detecting the second type of collaborative attacks and other types of attack models (e.g., the load profile shifting attacks).

### B. Hybrid Detection Methods

We can also try to integrate the above two kinds of detection methods, for achieving a balance between performance in terms of detection accuracy, as well as false rates and costs of deploying advanced sensors [123], [128], [167].

Some existing works integrate the above two categories of detection methods as follows [69], [89]. Assume that each community has an advanced sensor that can register the total amount of electricity supplied to the distribution networks (e.g., the central observer meter in [159]). Then, the community containing adversaries can be detected by comparing the advanced sensor's measurements with the summation of users' reported readings. Some machine learning-based methods can be applied to analyze user load profiles in the community to find out adversaries exactly.

Compared to existing machine learning-based methods, this type of solution has the advantage that the control center of the utility companies does not need to apply the machine learning methods to analyze all users' load profiles all the time. Instead, it does this only when the advanced sensor detects reading anomalies and only for the users in the community of which the advanced sensor detects the existence of reading anomalies. However, it also has the disadvantage of inheriting some limitations of both machine learning- and measurement mismatch-based methods.

Hybrid detection methods should be studied further to leverage the advantages of both the measurement mismatch- and machine learning-based methods.

### C. Detection Methods for Advanced Attacks

Some adversaries may have a sense of counter-reconnaissance. Thus, they may first figure out the working principles of detection methods and then launch some special advanced attacks to escape the detection. We list some possible antidetection attack models as follows.

*1) Intermittent Attacks:* Most existing works assume that, once adversaries begin to launch attacks for tampering with meter readings, they would not stop this fraudulent behavior unless caught by utility companies. However, in reality, for interrupting the detection, adversaries may switch their behaviors constantly between committing electricity theft and honestly consuming electricity. Specifically, they may launch cyber/physical attacks to tamper with meter readings for a while and then stop these attacks for another while. This behavior patterns can be repeated many times. To the best of our knowledge, most proposed methods are not designed for this kind of attack. Developing detection methods against this special kind of attack can be future work directions.

*2) Collaborative Attacks:* Almost all existing works assume that adversaries only compromise their meters. However, in real applications, adversaries may be cunning enough to launch collaborative attacks.

As aforementioned, the first type of collaborative attacks is the CNTL frauds, which is first proposed in [71]. As aforementioned in Section V-D, a CNTL occurs when multiple malicious users commonly tamper with one meter so that the meter records less electricity than the consumed amount by the household [71]. To the best of our knowledge, CNTL is only studied in [71] in which the authors propose the CNFD to identify the tampered meter and differentiate the malicious users. However, there is a big space to improve the accuracy and efficiency. Designing better methods to detect this kind of attack can be future work directions.

The second type of collaborative attack is proposed in [72] explained as follows. When adversaries tamper with their readings to lower the meter values, they may simultaneously compromise neighbors' smart meters for increasing the readings to escape detection. These adversaries usually summate neighbors' increased value equal to their decreased value (i.e., the stolen electricity). Lo and Ansari [171] propose an algorithm to deploy advanced sensors on edges of a tree structure modeling the topology of distribution networks of the power systems. In this way, every user's electricity consumption behavior is observable. The mutual inspection strategy [135], [136] can also be used to prevent collaborative attacks. However, the above two algorithms require that the number of advanced sensors installed is the same as the number of users in the distribution network. The cost is too high. More effective algorithms to detect collaborative attacks can be future work directions.

*3) Other Advanced Attacks:* We list some of these attacks as follows.

Both intermittent attacks and collaborative attacks are antidetection attacks belonging to reduced consumption attacks, which can be further classified into a major

difference and minor difference attacks. Most of the existing works focus on detecting major difference attacks, and few works target detecting minor difference attacks with the limited scenarios [69]. Minor difference attacks should be studied further in various system configurations.

As aforementioned in Section III-D, under the time of use pricing scheme and real-time pricing scheme, adversaries can also launch load profile shifting attacks, reduced pricing attacks, and neighbor pricing attacks to achieve the purpose of stealing electricity. There are no detection methods designed for the load profile shifting attacks and reduced pricing attacks. Few works target detecting neighbor pricing attacks with the limited scenarios [73], [74]. Thus, detection methods for load profile shifting attacks, reduced pricing attacks, and neighbor pricing attacks should be studied further in various system configurations.

### D. Privacy-Preserving Detection Methods

Most electricity theft detection methods require knowing some customers' private information, such as fine-grained load profiles or meter readings at certain times. However, by applying nonintrusive load monitoring methods to analyze customers' load profiles, the third parties can infer a lot of private information of customers, which includes, but are not limited to, customers' daily routines, electronic appliances, and even whether there are medical devices and alarm systems used in their houses. This violates customers' privacy and potentially poses risks to customers' properties and even life. For instance, burglars can carefully select the most vulnerable targets based on the daily routine and alarm information. If insurance companies know that there are medical devices used in some customers' homes, they may strategically increase the premiums of these customers. Based upon the information of customers' electrical appliances, marketing companies may incite customers to buy their products with targeted advertisement. On the other hand, industrial customers' load profiles even contain proprietary information about equipment and logistics. If the opponents obtain such information, the companies will lose advantages in commercial competitions. Limited research works are focusing on privacy-preserving electricity theft detection [67], [172]. Privacy-preserving electricity theft detection should be studied further in various system configurations.

### E. Technical Challenges

In this section, we summarize the technical challenges for developing efficient and accurate electricity theft detection methods as follows.

1) The motivation for stealing electricity has been strong all the time around almost all countries. We can observe recent cases listed as follows.
   a) It is revealed that, between 2014 and 2016, 354 people in Northern Ireland were convicted on charges related to electricity theft [173]. Among the 354 convictions, 47 people were sent to prison, 90 people received suspended sentences, and 55 people received community sentences [173]. In some areas of Northern Ireland, as many as six out of every 10 m are tampered with for electricity theft [173].
   b) The Energy and Minerals Regulatory Commission in Jordan says that a total number of 8836 electricity thefts were documented during the first half of 2019 [174].
   c) According to Netbeheer Nederland, the illegal cannabis farms discovered and closed down in the Netherlands in 2019 steal a total of around 60 million euros of electricity [175].
   d) In March 2020, police in California also busted three illicit marijuana grow-ops that have stolen electricity worthing approximately $120 000, $88 000, and $11 000, respectively [176]. Marijuana growers steal power mainly due to the following two reasons.
      i) Growing marijuana needs a huge amount of electricity, which causes a high energy cost; stealing power helps them to save electricity bills.
      ii) One important indicator for growing marijuana is that someone's power is much more significant than their neighbors. Thus, marijuana growers steal electricity to escape the detection of police.
   e) Since May 2021, many regions of China have suffered outages because of soaring prices and the tight supply of coal, a key fuel for its power generation, restricted power plants' operations [175]. With winter coming, the power outages become even more serious since hydropower generation is expected to fall, while consumption picks up during the winter heating season [175]. In November 2021, China's State Grid Corporation warned of a "tight balance" between power supply and demand through winter until spring [175]. The power shortages usually result in more customers stealing electricity.
2) Users' electricity consumption is measured by electricity meters installed near customers' premises. Regardless of the types of electricity meters, the measurements can be tampered with by adversaries through either physical attacks or cyberattacks, particularly when meters are replaced by smart meters. However, the readings reported by electricity meters are the only information that utility companies can obtain for knowing how much electricity users consume over a certain period. When these meters are hacked and report wrong values, it becomes very challenging for utility companies to identify which users in the community are stealing electricity.
   Some of the existing detection methods predict users' normal electricity consumption with their historical load profiles, assuming that daily electricity consumption shows some periodical characteristics. However,

in the real world, users' daily electricity consumption is impacted by various factors, including seasons, weather, holidays, the number of persons/appliances in a family, and even daily routines. Many of the above factors change randomly and suddenly from day to day. In other words, the reported readings can be true when they differ a lot from the predicted normal electricity consumption. On the contrary, the reported readings can be false when they are the same as the predicted normal electricity consumption.

3) To be smart systems, smart grids include more advanced devices and software to enable computation, communication, sensing, and intelligent decision making. Therefore, power systems are introduced with tremendous and various hardware, such as smart meters, distributed energy resources, and electric vehicles. At the same time, different kinds of software are introduced to modernize the power systems. While hardware and software empower the power systems with numerous new functions, such as demand response, they, at the same time, introduce many vulnerabilities into smart grids. Leveraging these vulnerabilities, adversaries can launch various cyberattacks to compromise the input (i.e., the pricing information) and the output (i.e., the energy usages) of the smart meters of theirs and neighbors, to steal electricity and even escape regular detection. This makes electricity theft in smart grids much more versatile and covet than traditional power systems and, hence, much more difficult to detect.
4) As aforementioned, for identifying which users are adversaries stealing electricity, both machine learning- and measurement mismatch-based methods require the input of users' fine-grained electricity consumptions and even other customer-related data, such as prior records of electricity theft and geographical information, which belongs to customers' private data. Currently, people around the world are becoming increasingly cautious about releasing their private data. Moreover, for protecting users' private data, many countries and regions issue various laws, which includes, but are not limited to, the California Consumer Privacy Act in California, United States of America, the General Data Protection Regulation in European Union, the Personal Data Protection Act in Singapore, and the Personal Information Protection Law of the People's Republic of China.

As aforementioned, privacy-preserving detection methods should be studied further in various system configurations to protect users' privacy. However, to a large extent, the goal of identifying malicious users accurately and preserving users' privacy conflicts with each other. Specifically, for identifying malicious users accurately, we prefer to have more comprehensive, more detailed, and more accurate information about the customers' electricity consumptions and other aspects, such as the tariff category and even the number of persons/appliances at home. This violates the goal of protecting users' privacy. For privacy-preserving purposes, some types of data, such as prior records and tariff categories, may not be accessible anymore; other types of data, such as users' fine-grained electricity consumptions, may not be provided in an accurate way (for example, these data may be added with noise). In many cases, data with less diversity and lower quality usually mean lower accuracy and higher FPRs for detecting adversaries. Hence, it is a technical challenge to identify users accurately and protect their privacy when developing privacy-preserving detection methods.

## VIII. CONCLUSION

Although many countries issue stringent laws to punish adversaries for stealing energy, electricity theft is still very common due to immediate and profitable economic benefits. Thus, in this article, we provide a comprehensive and in-depth understanding of electricity theft. We not only analyze its adverse effects such as enormous economic losses, degraded power quality, and even safety concerns but also summarize political and socioeconomic factors impacting electricity theft. We analyze physical/cyberattacks that adversaries usually launch to manipulate meter readings. Overall, methods for tampering with smart meter readings are more versatile, secret, and flexible than those for electromechanical meter readings. We explore existing electricity theft detection methods, which are classified into machine learning- and measurement mismatch-based methods. The machine learning-based methods suffer from a data imbalance issue and some nonmalicious factors, leading to a moderate detection accuracy and a high FPR. The measurement mismatched-based methods mainly suffer from high deployment costs. They usually have higher detection accuracies and lower FPRs than the machine learning-based methods. Some future research directions are also provided in this article. ■

## REFERENCES


[1] C. L. Stimmel, *Big Data Analytics Strategies for the Smart Grid*, 1st ed. New York, NY, USA: CRC Press, 2014.
[2] 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.
[3] A. B. M. S. Ali, Ed., *Smart Grids: Opportunities, Developments, and Trends*. Cham, Switzerland: Springer, 2013.
[4] 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.
[5] S. Sagiroglu, A. Ozbilen, and I. Colak, "Vulnerabilities and measures on smart grid application in renewable energy," in *Proc. Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2012, pp. 1–4.
[6] 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, 4th Quart., 2012.
[7] E. McCary and Y. Xiao, "Home area network accountability with varying consumption devices in smart grid," *Secur. Commun. Netw.*, vol. 9, no. 10, pp. 977–995, Jul. 2016.
[8] Daily Yellowstone Journal. (1886). *People Who Steal Edison's Electricity*. [Online]. Available: https://chroniclingamerica.loc.gov/lccn/sn86075021/1886-03-27/ed-1/seq-2/print/image_681x648_from_493%2C4711_to_2380%2C6508/#Cite_Brian_Cohen
[9] T. Instruments. (2018). *Smart Meters: Electricity Meters Solutions From Texas Instruments*. [Online].

Available: http://www.ti.com/solution/smart-e-meter-amr-ami
[10] B. Krebs. (2012). *FBI: Smart Meter Hacks Likely to Spread*. [Online]. Available: https://krebsonsecurity.com/2012/04/fbi-smart-meter-hacks-likely-to-spread/
[11] 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
[12] L. Northeast Group. (2017). *96 Billion Dollars Is Lost Every Year to Electricity Theft*. [Online]. Available: https://www.prnewswire.com/news-releases/96-billion-is-lost-every-year-to-electricity-theft-300453411.html
[13] P. Antmann. (2009). *Reducing Technical and Non-Technical Losses in the Power Sector*. [Online]. Available: https://openknowledge.worldbank.org/handle/10986/20786
[14] H. Arkell. (2014). *How Middle-Class Families Are Turning to Crime by Getting Specialist Gangs to 'Hotwire' Their Gas and Electricity Supplies to Beat Soaring Energy Bills*. [Online]. Available: http://www.dailymail.co.U.K./news/article-2542487/Energy-theft.html
[15] 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. 5, pp. 1007–1015, 2001.
[16] (2011). *Fighting Electricity Theft With Advanced Metering Infrastructure*. ECI. Telecom. [Online]. Available: http://www.ecitele.com
[17] Metering & Smart Energy International. (2017). *Analysis: Electricity Theft in South Africa. Metering & Smart Energy International*. [Online]. Available: https://www.metering.com/features-analysis/electricity-theft-south-africa/
[18] P. Kadurek, J. Blom, J. F. G. Cobben, and W. L. Kling, "Theft detection and smart metering practices and expectations in The Netherlands," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur. (ISGT Eur.)*, Gothenberg, Sweden, Oct. 2010, pp. 1–6.
[19] M. Place. (2013). *Energy Theft in Brazil Out of Control*. [Online]. Available: http://www.bnamericas.com/news/electricpower/energy-theft-in-brazil-out-of-control-expert
[20] 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.
[21] Ç. Yurtseven, "The causes of electricity theft: An econometric analysis of the case of Turkey," *Utilities Policy*, vol. 37, pp. 70–78, Dec. 2015.
[22] F. B. Lewis, "Costly 'throw-ups': Electricity theft and power disruptions," *Electr. J.*, vol. 28, no. 7, pp. 118–135, Aug. 2015.
[23] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.
[24] V. Gaur and E. Gupta, "The determinants of electricity theft: An empirical analysis of Indian states," *Energy Policy*, vol. 93, pp. 127–136, Jun. 2016.
[25] K. A. Seger and D. J. Icove, "Power theft: The silent crime," *FBI Law Enforcement Bull.*, vol. 57, pp. 20–25, Mar. 1988.
[26] P. Lynggaard, "Controlling interferencesin smart building IoT networks using machine learning," *Int. J. Sens. Netw.*, vol. 30, no. 1, pp. 46–55, 2019.
[27] 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.
[28] T. Sharma, K. K. Pandey, D. K. Punia, and J. Rao, "Of pilferers and poachers: Combating electricity theft in India," *Energy Res. Social Sci.*, vol. 11, pp. 40–52, Jan. 2016.
[29] 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.
[30] T. Ahmad, H. Chen, J. Wang, and Y. Guo, "Review of various modeling techniques for the detection of electricity theft in smart grid environment," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 2916–2933, Feb. 2018.
[31] T. B. Smith, "Electricity theft: A comparative analysis," *Energy Policy*, vol. 32, no. 18, pp. 2067–2076, Dec. 2004.
[32] S. Saini, "Electricity theft—A primary cause of high distribution losses in Indian state," *Int. Res. J. Manage. Commerce*, vol. 5, pp. 187–203, Jan. 2018.
[33] A. Park. (2018). *Rural-Urban Inequality in China*. [Online]. Available: http://web.worldbank.org/archive/website10208/WEB/PDF/CHAPTE-2.PDF
[34] N. Onat, "Techno-economic analysis of illegal electricity usage in Turkey and policy proposals," *WSEAS Trans. Power Syst.*, vol. 5, pp. 213–222, Jul. 2010.
[35] F. M. Mirza, M. S. Hashmi, and S. M. Mirza, "Long run determinants of electricity theft in Pakistan: An empirical analysis," *Pakistan J. Social Sci.*, vol. 35, no. 2, pp. 599–608, 2015.
[36] D. Meadows. (1986). *Poverty Causes Population Growth Causes Poverty*. [Online]. Available: http://donellameadows.org/archives/poverty-causes-population-growth-causes-poverty/
[37] C. Marangoz, "Illegal electricity use in Turkey: Causes and policy implications," M.S. thesis, Graduate School Vanderbilt Univ., Nashville, TN, USA, Dec. 2013. [Online]. Available: https://etd.library.vanderbilt.edu/available/etd-11272013-094111/unrestricted/Marangoz.pdf
[38] F. Jamil and E. Ahmad, "An empirical study of electricity theft from electricity distribution companies in Pakistan," *Pakistan Develop. Rev.*, vol. 53, no. 3, pp. 239–254, 2014.
[39] M. Golden and B. Min, "Theft and loss of electricity in an Indian state," Int. Growth Center, London, U.K., Tech. Rep., Feb. 2012. [Online]. Available: https://www.theigc.org/wp-content/uploads/2014/09/Golden-Min-2012-Working-Paper.pdf
[40] (2018). *Theft Act 1968*. [Online]. Available: https://www.legislation.gov.U.K./ukpga/1968/60/contents
[41] (2003). *The Electricity Act*. [Online]. Available: https://www.vakilno1.com/bareacts/theelectricityact/theelectricityact.html#135_Theft_of_Electricity
[42] (2017). *Pakistan Penal Code Electricity Theft Amendment 2016*. [Online]. Available: http://www.lawsofpakistan.com/pakistan-penal-code-electricity-theft/
[43] (2017). *Turkey—Criminal Code (Unofficial English Translation)*. [Online]. Available: https://ppp.worldbank.org/public-private-partnership/node/508/
[44] (2017). *Theft/Non-Technical Losses (Water and Electricity)*. [Online]. Available: https://ppp.worldbank.org/ppp/legislation-regulation/laws/theft-nontechnical-loss#Turkey
[45] O. Nnodim. (2017). *Electricity Theft May Attract N200,000 in Fines*. [Online]. Available: http://punchng.com/electricity-theft-may-attract-n200000-in-fines/
[46] *Voice and Accountability*. [Online]. Available: http://info.worldbank.org/governance/wgi/pdf/va.pdf
[47] A. Alesina, S. Özler, N. Roubini, and P. Swagel, "Political instability and economic growth," *J. Econ. Growth*, vol. 1, no. 2, pp. 189–211, Jun. 1996.
[48] *Government Effectiveness*. Accessed: Apr. 5, 2020. [Online]. Available: https://info.worldbank.org/governance/wgi/pdf/ge.pdf
[49] S. K. Katiyar, "Political economy of electricity theft in rural areas: A case study from Rajasthan," *Econ. Political Weekly*, vol. 40, no. 7, pp. 644–648, 2005.
[50] J. Mincer, "Education and unemployment," National Bureau of Economic Research, Cambridge, MA, USA, White Paper 3838, Sep. 1991. [Online]. Available: http://www.nber.org/papers/w3838
[51] S. Ajimotokin, A. Haskins, and Z. Wade. (2015). *The Effects of Unemployment on Crime Rates in the U.S.* [Online]. Available: https://smartech.gatech.edu/handle/1853/53294
[52] *Types of Energy Meters and Their Working Principles*. Accessed: Jun. 8, 2020. [Online]. Available: https://www.elprocus.com/watt-hour-meter-circuit-working-with-microcontroller/
[53] I. Sarwar. (2017). *Basic Types of Energy Meters*. [Online]. Available: http://engineerexperiences.com/basic-types-energy-meters.html#Digital_Energy_Meters
[54] Wikipedia. (2019). *Smart Meter*. [Online]. Available: https://en.wikipedia.org/wiki/Smart_meter
[55] SATEC (Australia) Pty Ltd. *Electricity Metering Accuracy Explained*. Accessed: Jun. 10, 2020. [Online]. Available: https://www.ecdonline.com.au/content/electrical-distribution/article/electricity-metering-accuracy-explained-372339275#axzz651mEkAx3
[56] Electric Power Research Institute. (2010). *Accuracy of Digital Electricity Meters*. [Online]. Available: https://skyvisionsolutions.files.wordpress.com/2015/06/epri_-accuracy-of-digital-meters.pdf
[57] P. Przydatek. (2018). *Updated ANSI Standards Address New Power Metering Capabilities and Challenges*. [Online]. Available: https://blog.se.com/power-management-metering-monitoring-power-quality/2018/08/21/updated-ansi-standards-address-new-power-metering-capabilities-and-challenges/
[58] *Common Methods for Electricity Theft*. [Online]. Available: http://www.sohu.com/a/121685601_470565
[59] Wikipedia. (2018). *Theft of Electricity*. [Online]. Available: https://en.wikipedia.org/wiki/Theft_of_electricity
[60] J. Hamill. (2014). *U.K. Smart Meters Arrive in 2020. Hackers Have Already Found a Flaw*. [Online]. Available: https://www.theregister.co.U.K./2014/10/30/smart_meter_hackable_for_free_electricity_say_security_reserachers/
[61] J. Chris Foreman and D. Gurugubelli, "Identifying the cyber attack surface of the advanced metering infrastructure," *Electr. J.*, vol. 28, no. 1, pp. 94–103, Jan. 2015.
[62] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," in *Proc. Int. Workshop Crit. Inf. Infrastruct. Secur. (CRITIS)*, Bonn, Germany, Oct. 2009, pp. 176–187.
[63] 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.
[64] F. M. Cleveland, "Cyber security issues for advanced metering infrasttructure (AMI)," in *Proc. IEEE Power Energy Soc. Gen. Meeting-Convers. Del. Electr. Energy 21st Century*, Jul. 2008, pp. 1–5.
[65] The Centre for Media Studies. (2017). *CMS-India Corruption Study 2017 Perception and Experience With Public Services & Snapshot View for 2005–2017*. [Online]. Available: http://cmsindia.org/sites/default/files/Monograph_ICS_2017.pdf
[66] S. Amin, G. A. Schwartz, and H. Tembine, *Incentives and Security in Electricity Distribution Networks*. Berlin, Germany: Springer, 2012, pp. 264–280.
[67] S. A. Salinas and P. Li, "Privacy-preserving energy theft detection in microgrids: A state estimation approach," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 883–894, Mar. 2016.
[68] J. Liu, Y. Xiao, and J. Gao, "Achieving accountability in smart grid," *IEEE Syst. J.*, vol. 8, no. 2, pp. 493–508, Jun. 2014.
[69] X. Xia, J. Lin, Y. Xiao, J. Cui, Y. Peng, and Y. Ma, "A control chart based detector for small-amount electricity theft (SET) attack in smart grids," *IEEE Internet Things J.*, early access, Sep. 17, 2021, doi: 10.1109/JIOT.2021.3113348.
[70] V. B. Krishna, K. Lee, G. A. Weaver, R. K. Iyer, and W. H. Sanders, "F-DETA: A framework for detecting electricity theft attacks in smart grids,"

in *Proc. 46th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Jun. 2016, pp. 407–418.

[71] 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.

[72] 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.

[73] Y. Liu, S. Hu, and T.-Y. Ho, "Leveraging strategic detection techniques for smart home pricing cyberattacks," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 2, pp. 220–235, Mar. 2016.

[74] Y. Liu, Y. Zhou, and S. Hu, "Combating coordinated pricing cyberattack and energy theft in smart home cyber-physical systems," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 3, pp. 573–586, Mar. 2018.

[75] V. Ford, A. Siraj, and W. Eberle, "Smart grid energy fraud detection using artificial neural networks," in *Proc. IEEE Symp. Comput. Intell. Appl. Smart Grid (CIASG)*, Dec. 2014, pp. 1–6.

[76] 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.

[77] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and SVM-based data analytics for theft detection in smart grid," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1005–1016, Jun. 2016.

[78] C. C. O. Ramos, A. N. Souza, G. Chiachia, A. X. Falcão, and J. P. Papa, "A novel algorithm for feature selection using harmony search and its application for non-technical losses detection," *Comput. Electr. Eng.*, vol. 37, no. 6, pp. 886–894, Nov. 2011.

[79] C. C. O. Ramos, A. N. de Sousa, J. P. Papa, and A. X. Falcao, "A new approach for nontechnical losses detection based on optimum-path forest," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 181–189, Feb. 2011.

[80] C. C. O. Ramos, A. N. Souza, J. P. Papa, and A. X. Falcao, "Fast non-technical losses identification through optimum-path forest," in *Proc. 15th Int. Conf. Intell. Syst. Appl. Power Syst.*, Nov. 2009, pp. 1–5.

[81] C. C. O. Ramos, D. Rodrigues, A. N. de Souza, and J. P. Papa, "On the study of commercial losses in Brazil: A binary black hole algorithm for theft characterization," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 676–683, Mar. 2018.

[82] J. V. Spirić, S. S. Stanković, M. B. Dočić, and T. D. Popović, "Using the rough set theory to detect fraud committed by electricity customers," *Int. J. Electr. Power Energy Syst.*, vol. 62, pp. 727–734, Nov. 2014.

[83] L. A. M. Pereira *et al.*, "Multilayer perceptron neural networks training through charged system search and its application for non-technical losses detection," in *Proc. IEEE PES Conf. Innov. Smart Grid Technol. (ISGT Latin America)*, Apr. 2013, pp. 1–6.

[84] J. I. Guerrero, I. Monedero, F. Biscarri, J. Biscarri, R. Millán, and C. León, "Non-technical losses reduction by improving the inspections accuracy in a power utility," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1209–1218, Mar. 2018.

[85] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Support vector machine based data classification for detection of electricity theft," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, Mar. 2011, pp. 1–8.

[86] B. C. Costa, B. L. A. Alberto, A. M. Portela, W. Maduro, and E. O. Eler, "Fraud detection in electric power distribution networks using an ann-based knowledge-discovery process," *Int. J. Artif. Intell. Appl.*, vol. 4, no. 6, pp. 17–23, Nov. 2013.

[87] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, and D. Coca, "Recovering missing data via matrix completion in electricity distribution systems," in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–6.

[88] J. E. Cabral, J. O. P. Pinto, K. S. C. Linares, and A. M. A. C. Pinto, "Methodology for fraud detection using rough sets," in *Proc. IEEE Int. Conf. Granular Comput.*, May 2006, pp. 244–249.

[89] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan. 2016.

[90] S. Y. Han, J. No, J. Shin, and Y. Joo, "Conditional abnormality detection based on AMI data mining," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 12, pp. 3010–3016, Sep. 2016.

[91] L. Wang, C. Feng, Y. Ren, and J. Xia, "Local outlier detection based on information entropy weighting," *Int. J. Sens. Netw.*, vol. 30, no. 4, pp. 207–217, 2019.

[92] M. Zanetti, E. Jamhour, M. Pellenz, and M. Penna, "A new SVM-based fraud detection model for AMI," in *Computer Safety, Reliability, and Security*, A. Skavhaug, J. Guiochet, and F. Bitsch, Eds. Cham, Switzerland: Springer, 2016, pp. 226–237.

[93] S. K. Singh, R. Bose, and A. Joshi, "Energy theft detection in advanced metering infrastructure," in *Proc. IEEE 4th World Forum Internet Things (WF-IoT)*, Feb. 2018, pp. 529–534.

[94] V. Badrinath Krishna, G. A. Weaver, and W. H. Sanders, "PCA-based method for detecting integrity attacks on advanced metering infrastructure," in *Quantitative Evaluation of Systems*, J. Campos and B. R. Haverkort, Eds. Cham, Switzerland: Springer, 2015, pp. 70–85.

[95] B. Coma-Puig, J. Carmona, R. Gavalda, S. Alcoverro, and V. Martin, "Fraud detection in energy consumption: A supervised approach," in *Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Oct. 2016, pp. 120–129.

[96] 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.

[97] 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 Conf.*, Nov. 2008, pp. 1–6.

[98] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and F. Nagi, "Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 1284–1285, Apr. 2011.

[99] D. R. Pereira *et al.*, "Social-spider optimization-based support vector machines applied for energy theft detection," *Comput. Electr. Eng.*, vol. 49, pp. 25–38, Jan. 2016.

[100] C. C. O. Ramos, A. N. de Souza, A. X. Falcao, and J. P. Papa, "New insights on nontechnical losses characterization through evolutionary-based feature selection," *IEEE Trans. Power Del.*, vol. 27, no. 1, pp. 140–146, Jan. 2012.

[101] F. Rodriguez, M. D. Martino, J. P. Kosut, F. Santomauro, F. Lecumberry, and A. Fernández, "Optimal and linear f-measure classifiers applied to non-technical losses detection," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, A. Pardo and J. Kittler, Eds. Cham, Switzerland: Springer, 2015, pp. 83–91.

[102] Scikit-Learn. (2020). *RBF SVM Parameters. Scikit-Learn.Org*. [Online]. Available: https://scikit-learn.org/stable/auto_examples/SVM/plot_rbf_parameters.html

[103] M. D. Martino, F. Decia, J. Molinelli, and A. Fernández, "Improving electric fraud detection using class imbalance strategies," in *Proc. ICPRAM*, 2012, pp. 1–7.

[104] Scikit-Learn. (2020). *One-Class SVM With Non-Linear Kernel (RBF). Scikit-Learn.Org*. [Online]. Available: https://scikit-learn:org/stable/auto examples/SVM/plotoneclass:html

[105] J. Song, R. Paul, J. Yun, H. Kim, and Y. Choi, "Cnn-based anomaly detection for packet payloads ofindustrial control system," *Int. J. Sens. Netw.*, vol. 36, no. 1, pp. 36–49, 2021.

[106] (2021). *Wikipedia. Deep Learning*. [Online]. Available: https://en.wikipedia.org/wiki/Deep_learning

[107] N. A. Alwan and Z. M. Hussain, "Deep learning techniques for noise-resilient localisation inwireless sensor networks," *Int. J. Sens. Netw.*, vol. 36, no. 2, pp. 59–67, 2021.

[108] M. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. M. Islam, and J.-M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," *Energies*, vol. 12, no. 17, p. 3310, Aug. 2019.

[109] M.-M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero, and A. Gomez-Exposito, "Hybrid deep neural networks for detection of non-technical losses in electricity smart meters," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1254–1263, Mar. 2020.

[110] 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.

[111] I. Monedero, F. Biscarri, C. León, J. I. Guerrero, J. Biscarri, and R. Millán, "Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees," *Int. J. Elect. Power Energy Syst.*, vol. 34, no. 1, pp. 90–98, 2012.

[112] 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.

[113] S. Amin, G. A. Schwartz, A. A. Cardenas, and S. S. Sastry, "Game-theoretic models of electricity theft detection in smart utility networks: Providing new capabilities with advanced metering infrastructure," *IEEE Control Syst.*, vol. 35, no. 1, pp. 66–81, Feb. 2015.

[114] C. H. Lin, S. J. Chen, C. L. Kuo, and J. L. Chen, "Non-cooperative game model applied to an advanced metering infrastructure for non-technical loss screening in micro-distribution systems," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2468–2469, Sep. 2014.

[115] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2392–2431, 4th Quart., 2017.

[116] O. C. Ibe, *Markov Processes for Stochastic Modeling*, 2nd ed. 2013.

[117] J. P. Papa, A. X. Falcão, and C. T. N. Suzuki, "Supervised pattern classification based on optimum-path forest," *Int. J. Imag. Syst. Technol.*, vol. 19, no. 2, pp. 120–131, Jun. 2009.

[118] A. S. Iwashita *et al.*, "Speeding up optimum-path forest training by path-cost propagation," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR2012)*, Nov. 2012, pp. 1233–1236.

[119] J. P. Papa, W. P. Amorim, A. X. Falc ao, and J. M. R. S. Tavares, "Recent advances on optimum-path forest for data classification: Supervised, semi-supervised and unsupervised learning," in *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 2016, pp. 109–123.

[120] C. C. O. Ramos, J. P. Papa, A. N. Souza, G. Chiachia, and A. X. Falcao, "What is the importance of selecting features for non-technical losses identification?" in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2011, pp. 1045–1048.

[121] D. Rodrigues, C. C. O. Ramos, A. N. De Souza, and J. P. Papa, "Black hole algorithm for non-technical losses characterization," in *Proc. IEEE 6th Latin Amer. Symp. Circuits Syst.*, Feb. 2015, pp. 1–4.

[122] S. E. N. Fernandes, D. R. Pereira, C. C. O. Ramos, A. N. Souza, D. S. Gastaldello, and J. P. Papa, "A probabilistic optimum-path forest classifier for non-technical losses detection," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3226–3235, May 2019.

[123] R. D. Trevizan *et al.*, "Non-technical losses identification using optimum-path forest and state estimation," in *Proc. IEEE Eindhoven PowerTech*,

Jun. 2015, pp. 1–6.

[124] E. Villar-Rodriguez, J. Del Ser, I. Oregi, M. N. Bilbao, and S. Gil-Lopez, "Detection of non-technical losses in smart meter data based on load curve profiling and time series analysis," *Energy*, vol. 137, pp. 118–128, Oct. 2017.

[125] J. L. Viegas, P. R. Esteves, and S. M. Vieira, "Clustering-based novelty detection for identification of non-technical losses," *Int. J. Elect. Power Energy Syst.*, vol. 101, pp. 301–310, Oct. 2018.

[126] E. W. S. Angelos, O. R. Saavedra, O. A. C. Cortés, and A. N. de Souza, "Detection and identification of abnormalities in customer consumptions in power distribution systems," *IEEE Trans. Power Del.*, vol. 26, no. 4, pp. 2436–2442, Oct. 2011.

[127] C. Richardson, N. Race, and P. Smith, "A privacy preserving approach to energy theft detection in smart grids," in *Proc. IEEE Int. Smart Cities Conf. (ISC2)*, Sep. 2016, pp. 1–4.

[128] Y. Guo, C.-W. Ten, and P. Jirutitijaroen, "Online data validation for distribution operations against cybertampering," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 550–560, Mar. 2014.

[129] J. I. Guerrero, C. León, I. Monedero, F. Biscarri, and J. Biscarri, "Improving knowledge-based systems with statistical techniques, text mining, and neural networks for non-technical loss detection," *Knowl.-Based Syst.*, vol. 71, pp. 376–388, Nov. 2014.

[130] D. D. Sharma, S. N. Singh, J. Lin, and E. Foruzan, "Identification and characterization of irregular consumptions of load data," *J. Modern Power Syst. Clean Energy*, vol. 5, no. 3, pp. 465–477, May 2017.

[131] D. Mashima and A. A. Cárdenas, *Evaluating Electricity Theft Detectors in Smart Grid Networks*. Berlin, Germany: Springer, 2012, pp. 210–229.

[132] C. Liao, C.-W. Ten, and S. Hu, "Strategic FRTU deployment considering cybersecurity in secondary distribution network," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1264–1274, Sep. 2013.

[133] J. P. Kosut, F. Santomauro, A. Jorysz, A. Fernandez, F. Lecumberry, and F. Rodriguez, "Abnormal consumption analysis for fraud detection: UTE-UDELAR joint efforts," in *Proc. IEEE PES Innov. Smart Grid Technol. Latin Amer. (ISGT LATAM)*, Oct. 2015, pp. 887–892.

[134] Y. Zhou, Y. Liu, and S. Hu, "Energy theft detection in multi-tenant data centers with digital protective relay deployment," *IEEE Trans. Sustain. Comput.*, vol. 3, no. 1, pp. 16–29, Jan. 2018.

[135] 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.

[136] Z. Xiao, Y. Xiao, and D. H.-C. Du, "Building accountable smart grids in neighborhood area networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–5.

[137] X. Xia, W. Liang, Y. Xiao, and M. Zheng, "BCGI: A fast approach to detect malicious meters in neighborhood area smart grid," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 7228–7233.

[138] 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.

[139] 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, pp. 445–458, Feb. 2019.

[140] X. Xia, Y. Xiao, W. Liang, and M. Zheng, "GTHI: A heuristic algorithm to detect malicious users in smart grids," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 805–816, Apr. 2020.

[141] X. Xia, Y. Xiao, and W. Liang, "SAI: A suspicion assessment-based inspection algorithm to detect malicious users in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 361–374, 2020.

[142] Y. Zhou, X. Chen, A. Y. Zomaya, L. Wang, and S. Hu, "A dynamic programming algorithm for leveraging probabilistic detection of energy theft in smart home," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 4, pp. 502–513, Dec. 2015.

[143] Sosmath. (2020) *Systems of Equations*. [Online]. Available: http://www.sosmath.com/soe/SE/SE.html

[144] 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.

[145] Provincial Electricity Authority. (2007). *Part A : Technical Specifications of FRTU for Remote Controlled Switches (FRTUS), no. ASD-FRTU-001/2007*. [Online]. Available: https://www.pea.co.th/Webapplications/tor/Attachments/50c306d8-2371-4e7b-8450-89b6dc5e6869/Spec.pdf

[146] Evercredit Enterprise Co., LTD. (2020). *Digital Protective Relay DPR 1000*. [Online]. Available: http://www.ecredit.com.tw/webfiles/62cfacd5-10d3-44f8-a5d6-fa22b0e8f43a.pdf

[147] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 19–67, 2005.

[148] Geeksforgeeks. (2020). *Dynamic Programming*. [Online]. Available: https://www.geeksforgeeks.org/dynamic-programming/

[149] L. Chen, J. Liu, and W. Ha, "Cloud service security evaluation of smart grid using deep belief network," *Int. J. Sens. Netw.*, vol. 33, no. 2, pp. 109–121, 2020.

[150] M. D. Maltz and R. McCleary, "The mathematics of behavioral change: Recidivism and construct validity," *Eval. Quart.*, vol. 1, no. 3, pp. 421–438, Aug. 1977.

[151] A. Blumstein and S. Moitra, "The identification of 'career criminals' from "chronic offenders" in a cohort," *Law Policy*, vol. 2, pp. 321–334, Jul. 1980.

[152] A. Blumstein, D. P. Farrington, and S. Moitra, "Delinquency careers: Innocents, desisters, and persisters," *Crime Justice*, vol. 6, pp. 187–219, Jan. 1985, doi: 10.1086/449107.

[153] D. P. Farrington and R. Tarling, *Prediction in Criminology*. Albany, NY, USA: State Univ. New York Press, 1985.

[154] P. Schmidt and A. Witte, *Predicting Recidivism Using Survival Models*. Cham, Switzerland: Springer, 1988.

[155] 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.

[156] M. C. Kurlychek, R. Brame, and S. D. Bushway, "Scarlet letters and recidivism: Does an old criminal record predict future offending," *Criminol. Public Policy*, vol. 5, no. 3, pp. 483–504, Aug. 2006.

[157] R. Dorfman, "The detection of defective members of large populations," *Ann. Math. Statist.*, vol. 14, no. 4, pp. 436–440, Dec. 1943.

[158] W. Han and Y. Xiao, "NFD: A practical scheme to detect non-technical loss fraud in smart grid," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 605–609.

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

[160] W. Han and Y. Xiao, "FNFD: A fast scheme to detect and verify non-technical loss fraud in smart grid," in *Proc. ACM Int. Workshop Traffic Meas. Cybersecurity*, May 2016, pp. 24–34.

[161] C. J. Bandim *et al.*, "Identification of energy theft and tampered meters using a central observer meter: A mathematical approach," in *Proc. IEEE PES Transmiss. Distrib. Conf. Expo.*, Dallas, TX, USA, Sep. 2003, pp. 163–168.

[162] Wikipedia. (2020) *Recursive Least Squares Filter*. [Online]. Available: https://en.wikipedia.org/wiki/Recursive_least_squares_filter#:~:text=Recursive%20least%20squares%20(RLS)%20is,reduce%20the%20mean%20square%20error

[163] S.-C. Yip, K. Wong, W.-P. Hew, M.-T. Gan, R. C.-W. Phan, and S.-W. Tan, "Detection of energy theft and defective smart meters in smart grids using linear regression," *Int. J. Electr. Power Energy Syst.*, vol. 91, pp. 230–240, Oct. 2017.

[164] *Linear Regression*, Wikipedia, San Francisco, CA, USA, 2020.

[165] W. Han and Y. Xiao, "CNFD: A novel scheme to detect colluded non-technical loss fraud in smart grid," in *Proc. Int. Conf. Wireless Algorithms, Syst., Appl. (WASA)*, Bozeman, MT, USA, Aug. 2016, pp. 47–55.

[166] B. Sun, X. Shan, K. Wu, and Y. Xiao, "Anomaly detection based secure in-network aggregation for wireless sensor networks," *IEEE Syst. J.*, vol. 7, no. 1, pp. 13–25, Nov. 2013.

[167] 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.

[168] Z. Qu, H. Li, Y. Wang, J. Zhang, and Y. Yao, "Detection of electricity theft behavior based on improved synthetic minority oversampling technique and random forest classifier," *Energies*, vol. 13, no. 8, p. 2039, 2020.

[169] S. K. Gunturi and D. Sarkar, "Ensemble machine learning models for the detection of energy theft," *Electr. Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 106904. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779620307021

[170] A. Takiddin, M. Ismail, U. Zafar, and E. Serpedin, "Robust electricity theft detection against data poisoning attacks in smart grids," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2675–2684, May 2021.

[171] C.-H. Lo and N. Ansari, "CONSUMER: A novel hybrid intrusion detection system for distribution networks in smart grid," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 1, pp. 33–44, Jun. 2013.

[172] 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.

[173] Cormac Campbell. (2018) *Dozens Jailed for Electricity Theft Over Three Years*. [Online]. Available: https://www.bbc.com/news/U.K.-northern-ireland-43318845

[174] The Jordan Times. (2019). *8,836 Cases of Electricity Theft Recorded in First Half of 2019*. [Online]. Available: https://energycentral.com/news/8836-cases-electricity-theft-recorded-first-half-2019

[175] Nltimes. (2020). *Illegal Cannabis Cultivators Stole 60 Million Electricity Last Year*. [Online]. Available: https://imabuds.com/illegal-cannabis-cultivators-stole-e60-million-electricity-last-year/

[176] Robert Bryce. (2020). *How Cannabis Farms Steal Megawatts to Grow Mega-Weed*. [Online]. Available: https://www.forbes.com/sites/robertbryce/2020/04/20/an-epidemic-of-stealing-watts-for-weed/#7d43ea9372a5

## ABOUT THE AUTHORS

**Xiaofang Xia** (Member, IEEE) received the Ph.D. degree in control theory and control engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2019.

She was a Visiting Student with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL, USA, from August 2016 to February 2018. She is currently an Assistant Professor with the School of Computer Science and Technology, Xidian University, Xi'an, China. Her research interests are mainly in cyber–physical systems, smart grid security, database management system, and anomaly detection.

**Yang Xiao** (Fellow, IEEE) received the B.S. and M.S. degrees in computational mathematics from Jilin University, Changchun, China, in 1989 and 1990, respectively, and the M.S. and Ph.D. degrees in computer science and engineering from Wright State University, Dayton, OH, USA, in 2000 and 2001, respectively.

He is currently a Full Professor with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL, USA. He had directed 20 doctoral dissertations and supervised 19 M.S. theses/projects. He has published over 300 Science Citation Index (SCI)-indexed journal papers (including over 60 IEEE/ACM TRANSACTIONS) and 300 Engineering Index (EI)-indexed refereed conference papers and book chapters related to these research areas. His current research interests include cyber–physical systems, the Internet of Things, security, wireless networks, smart grid, and telemedicine.

Prof. Xiao was a Voting Member of the IEEE 802.11 Working Group from 2001 to 2004, involving the IEEE 802.11 (Wi-Fi) standardization work. He is also an IET Fellow and an AAIA Fellow. He has served as a Guest Editor over 30 times for different international journals, including the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING in 2021, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING in 2021, *IEEE Network* in 2007, IEEE WIRELESS COMMUNICATIONS in 2006 and 2021, *IEEE Communications Standards Magazine* in 2021, and *Mobile Networks and Applications* (MONET) (ACM/Springer) in 2008. He also serves as the Editor-in-Chief of *Cyber-Physical Systems* journal, *International Journal of Sensor Networks* (IJSNet), and *International Journal of Security and Networks* (IJSN). He has been serving as an Editorial Board Member or an Associate Editor for 20 international journals, including the IEEE TRANSACTIONS ON CYBERNETICS since 2020, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS from 2014 to 2015, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2007 to 2009, and IEEE COMMUNICATIONS SURVEYS AND TUTORIALS from 2007 to 2014.

**Wei Liang** (Senior Member, IEEE) received the Ph.D. degree in mechatronic engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2002.

She is currently a Professor with the Shenyang Institute of Automation, Chinese Academy of Sciences. As a primary participant/a project leader, she developed the Wireless Networks for Industrial Automation–Process Automation (WIA-PA) and Wireless Networks for Industrial Automation–Factory Automation (WIA-FA) standards for industrial wireless networks, specified by IEC 62601 and IEC 62948, respectively. Her research interests include industrial wireless sensor networks and wireless body area networks.

Dr. Liang received the International Electrotechnical Commission 1906 Award in 2015 as a Distinguished Expert of industrial wireless network technology and standard.

**Jiangtao Cui** (Member, IEEE) received the M.S. and Ph.D. degrees in computer science from Xidian University, Xi'an, China, in 2001 and 2005, respectively.

From 2007 to 2008, he was with the Data and Knowledge Engineering Group, The University of Queensland, Brisbane, QLD, Australia, working on high-dimensional indexing for large-scale image retrieval. He is currently the Executive Dean and a Distinguished Professor with the School of Computer Science and Technology, Xidian University. He has published over 80 journal articles and conference papers, including Very Large Data Bases (VLDB), Special Interest Group on Management of Data (SIGMOD), International Conference on Data Engineering (ICDE), IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING (TKDE), *VLDB Journal* (VLDB J), IEEE TRANSACTIONS ON BIG DATA, and so on. His research interests include database management and core techniques, temporal–spatial data management, blockchain data management, and computer vision.

Dr. Cui is also a Distinguished Member and a Fellow of China Computer Federation (CCF), and a Committee Member of CCF Technical Committee for Database (TCDB), CCF Technical Committee for Computer Applications (TCAPP), and CCF Technical Committee for Block Chain (TCBC).