# A Cryptographic-Based Approach for Electricity Theft Detection in Smart Grid

**Khelifi Naim[1, *], Benahmed Khelifa[2] and Bounaama Fateh[3]**

**Abstract:** In order to strengthen their security issues, electrical companies devote particular efforts to developing and enhancing their fraud detection techniques that cope with the information and communication technologies integration in smart grid fields. Having been treated earlier by several researchers, various detection schemes adapted from attack models that benefit from the smart grid topologies weaknesses, aiming primarily to the identification of suspicious incoming hazards. Wireless meshes have been extensively used in smart grid communication architectures due to their facility, lightness of conception and low cost installation; however, the communicated packets are still exposed to be intercepted maliciously in order either to falsify pertinent information like the smart meter readings, or to inject false data instead, aiming at electricity theft during the communication phase. For this reason, this paper initiates a novel method based on RSA cryptographic algorithm to detect electricity fraud in smart grid. This new method consists of generating two different cryptograms of one electricity measurement before sending, after which the recipient is used to find the same value after decrypting the two cyphers in a normal case. Otherwise, a fraudulent manipulation could occur during the transmission stage. The presented method allows us to kill two birds with one stone. First, satisfactory outcomes are shown: the algorithm accuracy reaches 100%, from one hand, and the privacy is protected thanks to the cryptology concept on the other hand.

**Keywords:** Electricity consumption, fraud detection, RSA algorithm, security, smart grid.

## 1 Introduction

In the field of information and communication technologies, the conventional electricity delivery network was significantly enhanced to support new features and smart technologies, leading to the introduction of a new concept called Smart Grid. By dint of

[1] Laboratory of Energetic in Arid Zones (ENERGARID), Department of Electrical Engineering, Faculty of Technology, Tahri Mohammed University of Bechar, Bechar, 08000, Algeria.

[2] Department of Mathematics and Computer Science, Faculty of Exact Sciences, Tahri Mohammed University of Bechar, Bechar, 08000, Algeria.

[3] Laboratory of Energetic in Arid Zones (ENERGARID), Department of Electrical Engineering, Faculty of Technology, Tahri Mohammed University of Bechar, Bechar, 08000, Algeria.

[*] Corresponding Author: Naim Khelifi. Email: khna0883@gmail.com.

Received: 09 December 2019; Accepted: 06 January 2020.

CMC. doi:10.32604/cmc.2020.09391 www.techscience.com/journal/cmc

an Advanced Metering Infrastructure (AMI), a bidirectional communication flow is ensured between the end-users (smart meters) and the utility companies (providers). However, the advanced technologies have their own weak spots that present a fertile environment for malicious users to affect the smart grid functionalities. Non-Technical Losses (NTL) are one of the most serious hazards that have been identified since the smart grid appearance. They refer mainly to energy theft that could be performed by means of illegal data manipulation (fraud) or direct connections to the grid (theft), adding to secondary facts as flawed equipment or billing errors [Kosut, Santomauro, Jorysz et al. (2015)]. A recent study from [LLC (2015)] reported that the top 50 emerging market countries lose 58.7 billion dollars per year due to the NTL fraud. Moreover, these countries will likely invest $168 billion over the next decade to improve the reliability of Smart Grid infrastructure and to combat the NTL fraud. Hence, several studies were conducted to overcome this challenge, aiming to improve fraud detection mechanisms to go in parallel with the worsening of attacks. Numerous fraud detection techniques are grounded on machine learning and data analysis algorithms, which are very promising and suitable approaches capable of determining user profiles [Diffie and Hellman (1976); Abreua, Pereira and Ferrão (2012); Huang, Tang, Cheng et al. (2014); Zanetti, Jamhour, Pellenz et al. (2016); Zheng, Chen, Wang et al. (2018)], etc. However, these techniques require large amounts of data to build robust consumer profiles. In other words, the more the consumer data base is important, the more its consumption pattern is significant. Moreover, the electricity consumption of one client changes depending on the season, period of day, holidays or vacation, which not only increases the difficulty of reporting significant client profiles, but also leads to a confusion in distinguishing the normal from the abnormal consumption patterns.

This paper presents a new concept in fraud detection schemes, whereby the smart meter reading is encrypted (smart meter side) before sending, where the recipient (company) decrypts the received codes to check the correctness of the electricity consumption value (Section 4). Thus, the RSA encryption algorithm is applied in this approach because of its easier conception algorithm and its strength against hacking. Based only on the instantaneous smart meter reading, the proposed technique is able to detect fraudulent changes to the real reported consumption, neglecting any external factors (seasons, period, etc.); in other words, no more data is required to learn the consumption profile. This main contribution allows the proposed approach to be practical in the real world due to the simplicity of its implementation and its promising results that show an accuracy of 100% (Section 7). As far as we know, this is the first time that this approach has been implemented with such impeccable results when compared with the existing works in that field.

## 2 Related works

The Smart Grid concept consists of handling the traditional power grid by an armada of information and communication technologies for the purpose of improving the electrical system profitability. Accordingly, new cyber-attacks that appeared led us to rethink new defence’s mechanisms able to strengthen the security of companies. In that way, numerous studies have taken place in order to come up with new fraud detection approaches that can deal with those hazards.

Aiming to maintain a high level of reliability and efficiency of smart grid systems, utilities and researchers focus on dealing with any potential threat by means of countermeasures and detection mechanisms in real time. In that field, Jian et al. [Jian, Lu, Wang et al. (2014)] reported a survey of some AMI energy theft detection schemes summarizing them into three categories: classification-based, state estimation-based, and game theory-based.

A recent review of various modelling techniques for the detection of electricity theft in smart grid environment has been reported in Ahmad et al. [Ahmad, Chen, Wang et al. (2018)], it focuses on the various modelling practices for the identification and apprehension of non-technical losses. Various data mining modelling approaches were deeply studied, showing their impact on expediting the investigators and scientists.

Habitual consumer consumption represents the client profile which is nearly the same over seasons, so each household has its usual consumption pattern that helps companies to identify the abnormal behaviour. Abreua et al. [Abreua, Pereira and Ferrão (2012)] highlighted a methodology for habitual electricity consumption detection using pattern recognition given the intrinsic characteristics of the family. Two main patterns are discovered: persistent daily routines and patterns of consumption or baselines typical of weather and daily conditions when nearly 80% of household electricity use can be explained.

Machine learning algorithms are also applied to separate secure from attacked measurements [Ozay, Esnaola, Vural et al. (2016)]. Well-known batch and online learning algorithms (supervised and semi supervised) are employed with decision and feature level fusion to model the attack detection problem. The relationships between statistical and geometric properties of attack vectors employed in the attack scenarios and learning algorithms are analysed to detect unobservable attacks using statistical learning methods.

In Nasim et al. [Nasim, Jelena, Vojislav et al. (2014)], the AMI security requirements and bugs have been discussed in a survey of recent studies of the threat detection solution before an Intrusion Detection System was proposed for the neighbourhood area network (NAN) in AMI in order to manage the smart metering communication network, taking into consideration several attacks against physical, MAC, transport, and network layers.

A proposed method for state estimation in smart grid that uses a revolutionary algorithm such as bat algorithm with a hybrid approach based on a weighted least square technique offers an opportunity to omit and detect fake data [Khorshidi and Shabaninia (2015)].

Non-Technical Losses Fraud Detection techniques (NFD) were known as accessible and practical schemes useful to identify NTL fraud in Smart Grid. They allow differentiating tampered-with meters from normal meters using the approximated difference between the billing electricity and the actually consumed electricity. In that way, a novel detector has been proposed in order to detect colluded Non-Technical Losses in Smart Grid (CNTL) where its main contribution focused on identifying four types of that sort of NTL which are: segmented CNTL frauds, fully overlapped CNTL frauds, partially overlapped CNTL frauds, and combined CNTL frauds [Han and Xiao (2017)].

Another solution has been developed to address the issue of false data injection called the adaptive Cumulative Sum (CUSUM) algorithm [Huang, Tang, Cheng et al. (2014)]. The presented scheme consists of two interleaved steps:

- Introduces the unknown variable solver technique based on the Rao test.
- Applies the multithread CUSUM algorithm in order to determine quickly the eventual adversary respecting the given constraints.

Additionally, the Markov-chain-based analytical model was used to characterize the behaviour of the proposed method.

In Liu et al. [Liu and Hu (2016)], authors explore the social pattern of the networked smart homes studying the fact of falsifying the smart meter readings inductive to energy theft. The simulation results show that the hacker's energy billing could be reduced by 208% at the cost of other consumers; for that, they also performed a detection approach based on Bollinger bands and partially observable Markov decision process (POMDP). In order to ameliorate the efficiency and to cope with the high complexity of POMDP, they propose a probabilistic belief-state-reduction-based adaptive dynamic programming method. The accuracy detection reached the 92.55%.

Tariq et al. [Tariq and Poor (2016)] used as a Stochastic Petri Net SPN formalism is used to detect and localize the occurrence of theft in grid-tied MGs. Singular Value Decomposition (SVD) is used to calculate the accurate line resistance for theft detection in distribution systems.

Jokar et al. [Jokar, Arianpoo and Leung (2016)] presented a consumption pattern-based energy theft detector, which leverages the predictability property of customers' normal and malicious consumption patterns. Using distribution transformer meters, areas with a high probability of energy theft are short listed, and by monitoring abnormalities in consumption patterns, suspicious customers are identified. Application of appropriate classification and clustering techniques, as well as concurrent use of transformer meters and anomaly detectors, make the algorithm robust against non-malicious changes in usage pattern, and provide a high and adjustable performance with a low sampling rate.

Most of existing schemes either rely on pre-defined thresholds, or require external knowledge. This may lead to low detection accuracy when the thresholds are improperly defined, and when there is a lack of external knowledge. To deal with these challenges, a Gaussian-Mixture Model-based Detection (GMMD) scheme is proposed in Yang et al. [Yang, Zhao, Zhang et al. (2016)] to counteract data integrity attacks. That scheme operates through narrowing the range of normal data, which can be obtained through clustering the historical data and learning minimum and maximum values, or distance values to the centre of each individual cluster. The historical data is partitioned into the appropriate number of clusters K leveraging the Gaussian-Mixture model and measurement values in each cluster $Clu_k$ (where k=1; 2 . . .K) giving the narrowed range of normal data that could increase the probability of detecting malicious data.

An anomaly detection method that combines Principal Component Analysis (PCA) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to verify the integrity of the smart meter measurements was proposed in Varun et al. [Varun , Gabriel and William (2015)].

Zheng et al. [Zheng, Yang, Niu et al. (2017)] proposed an electricity-theft detection method based on Wide & Deep Convolutional Neural Networks (CNN) model. The Deep CNN component aims to identify the non-periodicity of electricity-theft and the periodicity of normal electricity usage based on two dimensional (2-D) electricity consumption data. Meanwhile, the Wide component can capture the global features of 1-D electricity consumption data.

Zanetti et al. [Zanetti, Jamhour, Pellenz et al. (2017)] introduced a fraud detection system (FDS) for AMI based on anomaly detection on the energy consumption reports from smart meters. A short-lived patterns have been proposed in this paper, in which a small set of recent measures could clearly defined the consumer behaviour. This approach allows the FDS to account for natural changes in the consumption behaviour of users and also helps to preserve their privacy.

Zheng et al. [Zheng, Chen, Wang et al. (2018)] aims to overcome the problem of poor accuracies confronted in the approaches that are based on the labelled data set or additional system information which is difficult to obtain in reality. Two novel data mining techniques are combined in order to solve the problem. One technique is the Maximum Information Coefficient (MIC), which can find the correlations between the non-technical loss (NTL) and a certain electricity behaviour of the consumer. MIC can be used to precisely detect thefts that appear normal in shapes. The other technique is the clustering technique by fast search and find of density peaks (CFSFDP) that finds the abnormal users among thousands of load profiles, making it quite suitable for detecting electricity thefts with arbitrary shapes.

Comparing to existing works, our proposed approach distinguish itself by:

- A new and different fraud detection paradigm is presented by applying the RSA algorithm as our basic detection tool, not only its habitual purpose as encryption algorithm. As far as we know, this idea is the first in the fraud detection concept.
- No prior data, neither external knowledge are required to build the consumer's patterns. The only mandatory data is the instantaneous smart meter reading, whereby other fraud detection technics may require daily, weekly, monthly or even more, to report the consumer consumption profile.
- The present scheme reports an excellent result comparing with existing works, adding to the simplicity of FDS consumption, promote this technic to be easily partible in the real world.
- The privacy is extremely protected thanks to the RSA concept.

## 3 System and attack model

Improved by an advanced metering infrastructure (AMI), smart grid communication networks are able to gather, ensure flowing, and report a large amount of data optimally when compared with the legacy power grid. The supported data are as follows: electricity consumption amount, measurements sensing, commands monitoring and so on; aiming generally at real time pricing, at leading electricity distribution on the right way, and at enabling a near-instantaneous balance of supply and demand management.

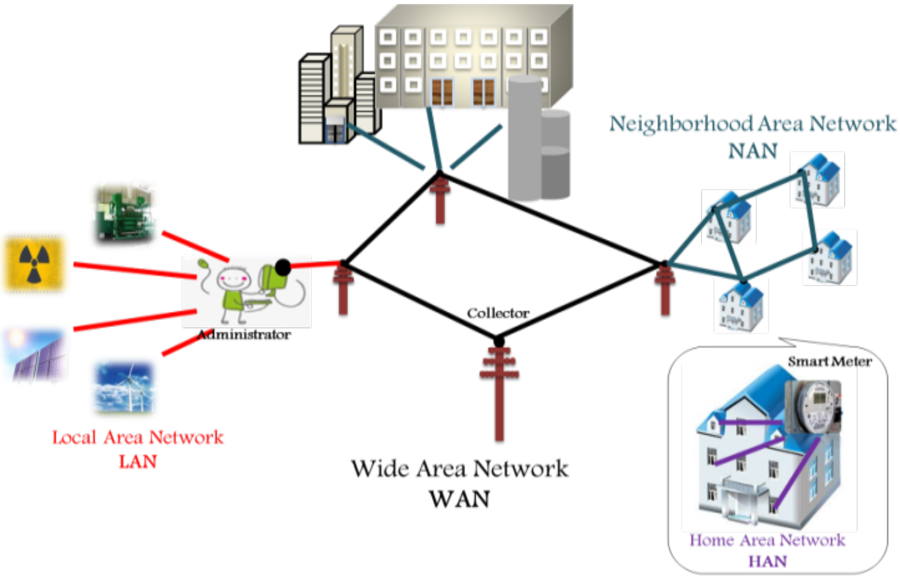

**Figure 1:** Basic AMI architecture

Fig. 1 presents the basic architecture of an AMI [Jian, Lu, Wang et al. (2014)], including the different kinds of the communication networks along with their requirements. However, this substantial enhancement is accompanied by important security issues about the potential vulnerabilities of the new technologies being introduced, especially, the wireless mesh implemented in the different network ranges (HAN, LAN or WAN); thus, communicated data are potentially vulnerable to adversaries' manipulations during any of these stages. Tab. 1 [Mahmud, Vallakati, Mukherjee et al. (2015)] recapitulates the common wireless communication technologies, adding to their vulnerabilities and the recommended solutions.

An understood framework of attack vector of Metering data was stated in Skopik et al. [Skopik and Ma (2012)], where the analyses of the threats and vulnerabilities are structured in three levels: the first one refers to the HAN stage's hazards that cope with threats to electric appliances, smart meters and their uplink to concentrator nodes. The second level refers to the NAN coverage that deal with vulnerabilities of the uplink from smart meters over concentrator nodes to data centres. And the last level deals with Web-based applications and community networks that use gathered meter data.

Furthermore, one common threat that is noticed within the three tiers reported in that study [Skopik and Ma (2012)] is the smart metering data fraudulent manipulations, serving to inject false data instead the legitimates ones, erase the measurement to behave as vacancy period, or report negative readings for acting as if that building provides the electricity power (e.g., solar panel), etc. Subsequently, the miss-reported readings lead to severe consequences ensuing from the attacker's goals such as financial gains, personal revenge, etc.

**Table 1:** Vulnerabilities and solutions to types of wireless communication technologies

| Technology | Advantage | Vulnerabilities | Solutions |
|---|---|---|---|
| **Wi-Fi** | Open Standard, High throughput Strong Home market penetration Low cost Relatively secure communication | Traffic Analysis, Passive and active eavesdropping, Man-in-the-middle attack, session hijacking, and replay attacks. | Two way authentication, encryption. |
| **ZigBee [Batista, Melício and Mendes (2014)]** | high reliability, self-configuration and self-healing, Low power consumption, low cost | Jamming, Message capturing and tampering, Exhaustion | A utility gateway device between HAN and SM, authentication, encryption |
| **Mobile Communications and Femtocells** | Consistent coverage in office or home, less power consumption | Network and service availability disruption, Fraud and service theft, Privacy and confidentiality disruption | Two way authentication, encryption |
| **WiMAX [Bian, Kuzlu, Pipattanasomporn et al. (2014)]** | High data rate (1 Gbps for stationary users), Low latency, Advanced Quality of Service (QoS), Sophisticated security | Ranging Attack (DoS attack, downgrading attack, water torture attack), Power Saving Attack, man-in-the-middle attack, Replay theft of service attack, Traffic analysis techniques | Encryption, Intrusion detection schemes, access control to specific applications |
| **Long Term Evolution (LTE)** | Less Interference, Resource efficient [Yaacoub and Abu-Dayya (2014)] | Attacks on the air interface, Attacks against the core network | Two way authentication, encryption, introduction of mobile virtual network operator (MVNO) |

This paper addresses mainly the false data injection (FDI) issue that occurs during the communication stage, resulting on erroneous data of smart meter readings. FDIs (false data injection) are defined as the attacks that modify how the smart meter computes and report energy consumption [Chen, Yang, McCann et al. (2015); Zanetti, Jamhour, Pellenz et al. (2017)]. Tab. 2 presents a formal definition of how FDI may modify consumption reports. The table includes proposals from [Jokar, Arianpoo and Leung (2016); Zanetti, Jamhour, Pellenz et al. (2017); Zheng, Chen, Wang et al. (2018)]. In Tab. 2, $\mathrm{x_t}$ is the original consumption reported at time *t* and $\tilde{\boldsymbol{x}}_{\boldsymbol{t}}$ is the tampered one.

**Table 2:** FDI types defined in [Zanetti, Jamhour, Pellenz et al. (2017)]

| Types | Modification |
|---|---|
| FDI1 | $\tilde{x}_t \leftarrow \alpha.x_t\ where\ 0 < \alpha < 1$<br>$\alpha$: same for all reports |
| FDI2 | $\tilde{x}_t \leftarrow f(x_t)$<br>$f(x_t) = \begin{cases} x_t\ if\ x_t \leq c \\ \tilde{c}\ if\ x_t\ > c \end{cases}$<br>$c$: cut-off point<br>$c_{min} < \tilde{c} < c$: randomly defined |
| FDI3 | $\tilde{x}_t \leftarrow \max\,(x_t - c, 0)$<br>$c$ fixed value in kwh |
| FDI4 | $\tilde{x}_t \leftarrow f(t).x_t$<br>$f(t) = \begin{cases} 0\ if\ \ t_i < \ t < t_f \\ \ 1\ otherwise \end{cases}$<br>$t_f - t_i$ randomly defined each day |
| FDI5 | $\tilde{x}_t \leftarrow \alpha_t.x_t$ **where** $0 < \ \alpha_t < 1$<br>$\alpha_t$ randomly defined for each report |
| FDI6 | $\tilde{x}_t \leftarrow \alpha_t.\bar{x}$ **where** $0 < \ \alpha_t < 1$<br>$\bar{x}$ average consumption for each selected period<br>$\alpha_t$ randomly defined for each report |

In FDI1, consumption reports are reduced by a constant percentage throughout the entire fraudulent period. In FDI2, reports above a threshold are clipped. In FDI3, a constant value is subtracted from all reports. FDI4 replaces by zero all reports during a random period defined each day. In FDI5, every consumption report is modified by a different percentage. Finally, FDI6 generates synthetic reports by multiplying the average consumption of previous month by a random percentage defined for each report.

For this reason, efficient detection mechanisms are strongly required in order to identify clearly the miss-reported values of smart meter readings, which is the subject of the present paper.

## 4 RSA cryptographic algorithm fraud detection system (RSA-FDS)

### *4.1 Fraud detection model (FDM)*

As stated in the previous section, smart metering data that play a vital role in the AMI functioning represent the main target for adversaries aiming to disturb the efficient processing of the smart grid tasks. So, it is very important to identify any illegal manipulation on the communicated metering data the soonest possible to prevent the perturbation of the system in responding as expected. The proposed model could be applied on various sorts of data flowing in the smart grid. In that case, the fraud in electricity consumption based on smart meter readings is picked out by high-lightening the model concept.

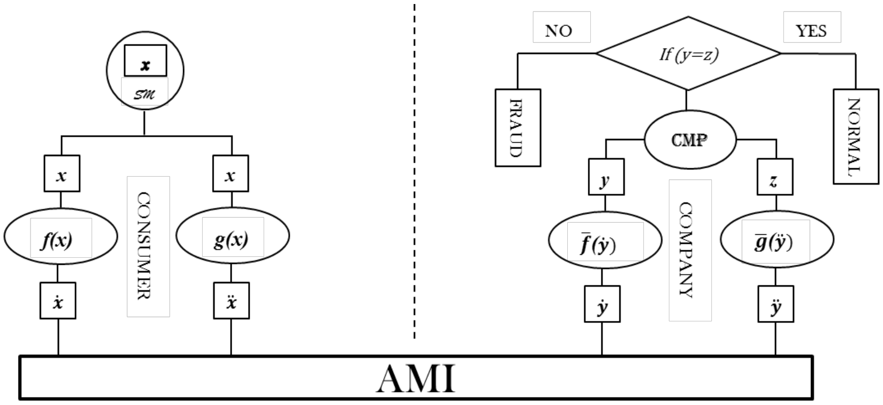

**Figure 2:** Fraud detection model (FDM)

Assuming that $x$ represents the normal measures of the electricity consumption of period of time $t$. By means of an encryption function $f$, a cypher message $\dot{x}$ is generated by the equation:

$$\dot{x} = f(x) \tag{1}$$

As it's known, the inverse function $\bar{f}$(decryption function) of the cryptogram $\dot{x}$ reports the original value x :

$$x = \bar{f}(\dot{x}) \tag{2}$$

So the basic concept of the proposed Fraud Detection Model (FDM) consists of encoding the smart meter reading before being communicated to the company, where this last one decodes the received value to obtain the original message. But, how can we confirm if the obtained value from the decryption function is the original one, or if it was altered, intentionally or unintentionally during the communication stages between the smart meter and the company?

In order to solve this dilemma, a reference value is needed for checking the correctness of the received measurement value. For that, one more code $\ddot{x}$ is also created at the smart meter side by virtue of a second encryption function $g$ where:

$$\ddot{x} = g(x) \tag{3}$$

As stated in the Eq. (2), the decryption function $\bar{g}$ of the code $\ddot{x}$ must return the original value $x$:

$$x = \bar{g}(\ddot{x}) \tag{4}$$

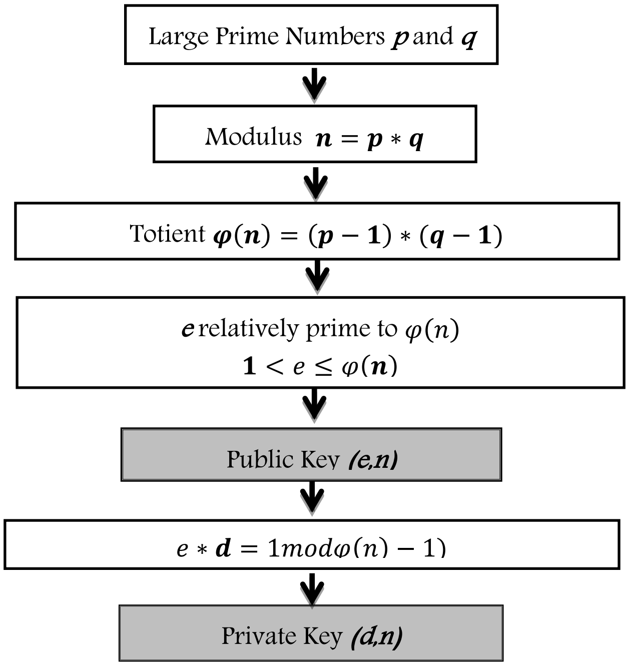
Large Prime Numbers $\boldsymbol{p}$ and $\boldsymbol{q}$
Modulus $\boldsymbol{n} = \boldsymbol{p} * \boldsymbol{q}$
Totient $\boldsymbol{\varphi}(\boldsymbol{n}) = (\boldsymbol{p} - \mathbf{1}) * (\boldsymbol{q} - \mathbf{1})$
$\boldsymbol{e}$ relatively prime to $\varphi(n)$
$\mathbf{1} < e \leq \varphi(\boldsymbol{n})$
Public Key *(e,n)*
$e * \boldsymbol{d} = 1 mod \varphi(n) - 1)$
Private Key *(d,n)*

**Figure 1:** RSA keys creation

The recipient (company) obtains two cryptograms, and after being decrypted, the same value $x$ should be reported by the two decryption functions $f$ and $g$ in normal cases. If the reported values are different, it means that the transferred data is altered during communication phase.

Quite simply, the measured amount of electricity consumption is encrypted twice at the consumer side (smart meter), before being sent through the AMI networks for the concerned stakeholders (e.g., billing services). So two codes are generated of one value at the first stage. The recipient (company) decrypts the two cryptograms and performs a comparison between the two. In the normal process, the two obtained values must be the same because the two codes are generated from the same value; otherwise, if the received messages are different, meaning that an illegal manipulation of the real amount has occurred during the communication stage between the consumer and the company. The fraud detection model (FDM) is described in Fig. 2.

### *4.2 RSA algorithm*

Utilizing various mathematic theorems, RSA cryptographic algorithm is promoted as one of the most effective encryption algorithms that have been widely applied in networks communication security protocols. Being introduced in 1978 by Ron Rivest, Adi Shamir, and Leonard Adleman, the RSA algorithm is inspired from the Diffie et al. theorems proved earlier [Diffie and Hellman (1976)]. The main idea implemented by the RSA algorithm is public-key encryption [Milanov (2009)], where two keys are required in order to encrypt and decrypt the messages: the encryption keys are public, where the decryption keys are private; in other words, only one who has the decryption keys can decipher the cryptograms.

Formally, if $A$ wants to send a message $M$ to $B$ , he must use $B$'s encryption key $E_B$ which is public (retrieving from the public file PF) to encrypt the message $M$ :

$$C = E_B(M) \tag{5}$$

And inversely,$B$ must use his private decryption key $D_B$ to decrypt the cypher text:

$$M = D_B(C) \tag{6}$$

In nutshell, the whole RSA process consists of three stages: keys creation (public and private), message encryption (using the public key) and the message decryption (by means of the decryption key). Fig. 3 summarises the full process of the public and private keys generation.

After the creation of the encryption key $(e, n)$ and decryption key $(d, n)$, the message $M$ is encrypted by being raised to the $e^{th}$ power modulo $n$, in order to compute the cypher text $C$ . The same operation is performed to deduce the plain text $M$ by raising the cypher text $C$ to the $d^{th}$ power modulo $n$.

$$C = M^{e} mod(n) \tag{7}$$

$$M = C^{d} mod(n) \tag{8}$$

### 4.3 Fraud detection system (FDS)

As stated in the fraud detection model (FDM), the basic idea is to create two detection codes from the smart meter reading before it is sent to the utility for the billing task; thus, the electricity amount is encrypted twice using two different encryption keys $(e; n)$ and $(\acute{e}; \acute{n})$, so the Eq. (1) and Eq. (3) become:

$$\dot{x} = f(x) \quad => \quad \dot{x} = x^{e} mod(n) \tag{9}$$

$$\ddot{x} = g(x) \quad => \quad \ddot{x} = x^{\acute{e}} mod(\acute{n}) \tag{10}$$

The recipient (utility company) gets two codes $\dot{x}$ and $\ddot{x}$. Performing the decryption operation by means of the different decryption keys $(d;\ n)$ and $(\grave{d};\ \grave{n})$ associated to their encryption keys $(e, n)$ and $(\grave{e}, \grave{n})$respectively, Eq. (2) and Eq. (4) report:

$$x = \bar{f}(\dot{x}) \quad => \quad x = \dot{x}^{d} mod(n) \tag{11}$$

$$x = \bar{g}(\ddot{x}) \quad => \quad x = \ddot{x}^{\acute{d}} mod(\acute{n}) \tag{12}$$

therefore, the final result of the two decryption operations must report the same value $x$ which is the original electricity measurement sent. Otherwise, the dissimilarity between the results of the two decryption processes means that one or both communicated values $\dot{x}$ and $\ddot{x}$ have been changed intentionally or unintentionally, which seems to be a fraudulent behaviour. Fig. 4 shows the full conception of the fraud detection scheme based on cryptographic RSA algorithm.

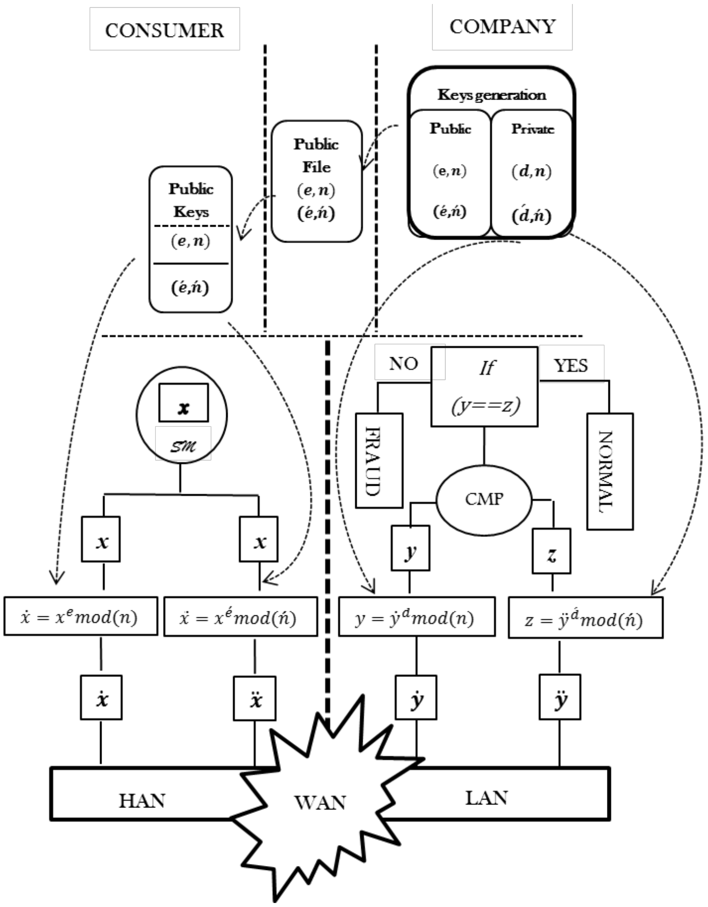

**Figure 2:** RSA based Fraud Detection System (RSAFDS)

### *4.4 RSA signature*

One more concern could be figured in this proposition, if an intruder attempts to use the company's encryption keys $(e;n)$ and $(\grave{e};\grave{n})$ in order to inject a false data pretending it comes from a legitimate smart meter. In that case, the receiver cannot guess the fraud attack because the final decrypted values are the same.

To address this issue, the signature concept of RSA algorithm is applied to ensure that the messages come from a trusted smart meter.

Basically, company generates a set of RSA keys couples (encryption and decryption keys), and allocates to each smart meter its own RSA encryption key $E_S$ that has been preconfigured into its soft program, but its associated decryption key $D_S$ remains stored at the company side. So, anyone could discover those keys because they are privates and not shared.

According to RSA proprieties cited in Milanov [Milanov (2009)], every message is the cipher-text of another message, and that every cipher-text can be interpreted as a message.

The principal idea consists of encrypting the returned values of RSA-FDS encryption algorithm with the smart meter encryption key $E_S$.

Formally, in order to explain simply the signature concept, we consider that $E_r$ (that stands for $(e; n)$ used previously) and $\grave{E}_r$(representing $(\grave{e}; \grave{n})$) represent the RSA-FDS encryption keys corresponding to the receiver (company) and $M$ is the meter reading:

$$\begin{gathered} E_r(M) = C \\ \grave{E}_r(M) = \grave{C} \end{gathered} \tag{13}$$

After that, we sign the messages $C$ and $\grave{C}$ by being encrypted with the smart meter (sender) encryption key $E_S$:

$$E_s(C) = S \text{ and } E_s(\grave{C}) = \grave{S} \tag{14}$$

This way, we can assure only that the receiver can decrypt the document. When it does, it gets the signature by:

$$\begin{gathered} D_s(S) = C \implies D_s(E_s(C)) = C \\ D_s(\grave{S}) = \grave{C} \implies D_s\left(E_s(\grave{C})\right) = \grave{C} \end{gathered} \tag{15}$$

Now, we know the message came from the smart meter, since only his encryption key $E_s$ could compute the signature.

After that, the following procedures of the RSA-FDS algorithm are effectuated as presented previously.

## 5 Evaluation

In order to evaluate the effectiveness of the proposed fraud detection technique, we use the Electricity consumption benchmarks [Department of Industry (2014)] that stands for a survey responses matched with household consumption data for 25 households. It contains includes half hourly electricity usage reports (in Watt Hours (WH)) for each household in the sample from 1 April 2012 to 31 March 2014 (or such time as data are available after the installation of a smart meter). Therefore, it is a reasonable assumption that all samples belong to honest users. The large number and variety of customers, long period of measurements and availability to the public make this dataset an excellent source for research in the area of analysis of smart meters data.

Then an abnormal vector of measures is created to simulate the six FDIs stated previously. The goal is not only to subject the two measurement vectors to the RSA-FDS, but also to evaluate the performance of the proposed algorithm by means of different metrics such as:

- True Positive (TP): refers to the positive instances classified as positive.
- False Negative (FN): refers to the misclassified positive instances.
- True Negative (TN): refers to the negative instances classified as negative.
- False Positive (FP): refers to the misclassified positive instances.

Using those different instances, the following metrics are defined as:

$$\mathrm{TPR} = \mathrm{TP}/\mathrm{P} \qquad \& \qquad \mathrm{FPR} = \mathrm{FP}/\mathrm{P} \tag{16}$$

Based on those two parameters (TPR, FPR), the ROC (Recipient operating characteristics) graph will be drawn in which TPR is plotted on the Y axis, and FPR is plotted on the X axis [Fawcett (2006)]. The ROC curve demonstrates a balance between the true classified class and the false alarm; in other words, it shows the classifier credibility. Hence, the more the TPR is higher and FPR is lower, the more the classifier is reliable, and inversely, the more the TPR is lower and the FPR is higher, the weaker the classifier becomes.

To obtain comprehensive evaluation results in the unbalanced dataset, we use the AUC (Area Under Curve) and MAP (Mean Average Precision) values mentioned in Zheng et al. [Zheng, Chen, Wang et al. (2018)]. The two evaluation criteria have been widely adopted in classification tasks. The AUC is defined as the area under the receiver operating characteristic (ROC) curve, which is the trace of the false positive rate and the true positive rate. Define the set of fraudulent users $F$ as the positive class and benign users $B$ as the negative class. The suspicion *Rank* is in ascending order according to the suspicion degree of the users. AUC can be calculated using *Rank* as in the following equation:

$$AUC = \frac{\sum_{i \in F} Rank_i - \frac{1}{2}|F|(|F|+1)}{|F| * |B|} \tag{17}$$

Let $Y_k$ the number of electricity thieves who rank at top $k$, and define the precision $P@k = \frac{Y_k}{k}$.

Given a certain number of $N$, $MAP@N$ is the mean of $P@k$ defined in the equation:

$$MAP@N = \frac{\sum_{i=1}^{r} P@k_i}{r} \tag{18}$$

where $r$ is the number of electricity thieves who rank in the top $N$ and $k_i$ is the position of the *i-th* electricity thieves. In this paper we use $MAP@50$.

## 6 Experiments

In the experimental section, we chose 1000 smart meter readings as a normal set, whereas the fraudulent vector contains also 6000 of miss-reported measurements representing all the six false data injection (FDI) possibilities stated previously such as the off-set FDI, cut-off point, percentage, zero consumption and low profile FDI. Moreover, this experiment takes into account the fraud in energy production (the negative reported values), not just the energy stolen as the previous paper cited. Remembering the FDM concept that stands for encoding the real amount of electricity measurement twice at first, meaning that the pertinent data representing the consumption profile are the two cryptograms flowing into the AMI. Regardless of the stage where the attack could occur (HAN, WAN or LAN), three fraudulent scopes are expected from adversaries by manipulating the first value, the second one or both of them. Hence, three scenarios are accordingly studied in this way in order to evaluate the FDS-RSA algorithm performance for all possibilities. Using the cryptographic parameters reported in Tab. 3, an excerpt of 30 statements is revealed in Tab. 4, reporting the results of the whole process of the presented detection technique, including 13 fraud vectors representing the three scenarios stated previously. Notice that the fraud vectors have been generated randomly to simulate a real situation that could occur at any stage between the household and the controller (company). Recordings field stands for the smart meter readings; Code 1-1 and 1-2 stand for the two cryptograms obtained from the encryption phase at the consumer side; Code 2-1 and 2-2 represent the two received cryptograms at the utility side that will be decrypted to calculate the final communicated amount Decrypted 1 and 2; and finally the field Class shows the classification result of the received values where class '1' stands for the normal profile and class '0' represent the fraudulent behaviour.

**Table 3:** RSA keys

| | First keys | Second Keys |
|---|---|---|
| **Encryption key** | 709 | 347 |
| **Decryption key** | 17805 | 299 |
| **Modulo** | 33109 | 1191 |

**Table 4:** RSA-FDS demonstration

| Recordings | Code 1-1 | Code 1-2 | | Code 2-1 | Code 2-2 | Decrypted 1 | Decrypted 2 | Class |
|---|---|---|---|---|---|---|---|---|
| 386 | 10504 | 188 | **Fraud vector** | 10504 | 188 | 386 | 386 | 1 |
| 123 | 27666 | 117 | | 27666 | 117 | 123 | 123 | 1 |
| 1786 | 9666 | 1159 | | 7277 | 1090 | 26543 | 1111 | 0 |
| 655 | 14029 | 469 | | 14029 | 469 | 655 | 655 | 1 |
| 398 | 16817 | 398 | | 16817 | 398 | 398 | 398 | 1 |
| 354 | 20807 | 603 | | 20807 | 603 | 354 | 354 | 1 |
| 1149 | 4368 | 678 | | 4368 | 210 | 1149 | 303 | 0 |
| 436 | 29576 | 742 | | 29576 | 742 | 436 | 436 | 1 |
| 1732 | 10902 | 409 | | 2955 | 409 | 25063 | 541 | 0 |
| 1056 | 4670 | 165 | | 4670 | 165 | 1056 | 1056 | 1 |
| 324 | 10155 | 57 | | 10155 | 57 | 324 | 324 | 1 |
| 1034 | 27990 | 971 | | 27990 | 971 | 1034 | 1034 | 1 |
| 1964 | 22008 | 1052 | | 22008 | 1052 | 1964 | 773 | 0 |
| 1734 | 9303 | 1152 | | 9303 | 1152 | 1734 | 543 | 0 |
| 1468 | 8474 | 688 | | 8474 | 688 | 1468 | 277 | 0 |
| 580 | 8264 | 145 | | 8264 | 8 | 580 | 62 | 0 |
| 305 | 30290 | 650 | | 30290 | 650 | 305 | 305 | 1 |
| 1093 | 7291 | 346 | | 7291 | 346 | 1093 | 1093 | 1 |
| 635 | 20593 | 59 | | 20593 | 59 | 635 | 635 | 1 |
| 1418 | 25789 | 605 | | 14321 | 454 | 23375 | 721 | 0 |
| 1136 | 4604 | 17 | | 4604 | 17 | 1136 | 1136 | 1 |
| 236 | 9812 | 818 | | 9812 | 818 | 236 | 236 | 1 |
| 1971 | 24947 | 111 | | 24947 | 111 | 1971 | 780 | 0 |
| 1134 | 10606 | 192 | | 214 | 108 | 15457 | 945 | 0 |
| 1105 | 14875 | 1123 | | 14875 | 1123 | 1105 | 1105 | 1 |
| 480 | 9259 | 936 | | 9259 | 936 | 480 | 480 | 1 |
| 1907 | 2036 | 746 | | 2036 | 746 | 1907 | 716 | 0 |
| 1307 | 9351 | 179 | | 9351 | 179 | 1307 | 116 | 0 |
| 350 | 18902 | 368 | | 3197 | 358 | 14276 | 940 | 0 |
| 671 | 13157 | 677 | | 13157 | 677 | 671 | 671 | 1 |

**Table 5:** Zero consumption confusion

| Recordings | Code 1-1 | Code 1-2 | Code 2-1 | Code 2-2 | Decrypted 1 | Decrypted 2 | Class |
|---|---|---|---|---|---|---|---|
| 1820 | 1650 | 812 | 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

**Table 6:** Zero consumption fix

| Recordings | Amend | Code 1-1 | Code 1-2 | Code 2-1 | Code 2-2 | Decrypted 1 | Decrypted 2 | Class |
|---|---|---|---|---|---|---|---|---|
| 1820 | | 1650 | 812 | 0 | 0 | 0 | 0 | 0 |
| 0 | **6000** | 7370 | 23906 | 7370 | 23906 | 6000 | 6000 | 1 |

Although, the RSA-FDS is able to detect any kinds of FDI cited previously whether the target of the attack is stated in the first, second or third scenario, the proposed algorithm fails in detecting the zero-consumption FDI, where adversaries 'goal' is to erase the real communicated amount to report zero consumption amount.

As we can spot from Tab. 5, the two recordings were classified as normal consumption patterns in spite of the zero consumption attack that targets the first vector. So the proposed detector misclassifies the real zero readings from the zero consumption attack. To get rid of this shortcoming, before the encryption processes, the real zero electricity measurements will be raised by "$A_{max}$", which represents the multiple of the maximum reported value that will never be reached at any time. For example, the maximum value in the dataset is 2000 kwh, so the $A_{max}$=6000 kwh. Inversely, after decrypting the received cryptograms, if the results are equal to the predefined threshold $A_{max}$, that amount will be decreased to find finally the real zero amount of electricity consumption. And the received cryptograms that report zero values will be classified as abnormal consumption profile.

Tab. 6 clarifies the solution of the zero consumption confusion by setting two different cases where the first recording is subjected to a zero consumption attack, and the second one reports a real zero amount of electricity consumption. It is clear that the ambiguity is elucidated when the reported smart meter readings are well classified.

After performing the RSA-FDS on the full data set including the fraud vector that simulates the three scenarios with all the FDI cited previously, the results are as illustrated RSA-FDS performance metrics presented in Tab. 7.

## 7 Results and discussion

**Table 7:** RSA-FDS Performance Metrics

| Metrics (%) | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| **TPR** | 100 | 100 | 100 |
| **FPR** | 0 | 0 | 0 |
| **Accuracy** | 100 | 100 | 100 |
| **AUC** | 100 | 100 | 100 |
| **MAP@50** | 100 | 100 | 100 |

Great results are reported by applying the RSA-FDS in various fraudulent windows. The experimented dataset that is subjected to different FDI attacks is examined in order to check the performance of the proposed detection algorithm. In the FDM context, the goal is successfully reached when the normal smart meter readings are correctly classified (TPR=100%). In the meantime, the false alarms (FPR=0%) are avoided, meaning that any fraudulent manipulations are correctly classified as abnormal profiles of electricity consumption.

Furthermore, one serious concern confronted in previous works based in machine learning schemes [Jokar, Arianpoo and Leung (2016); Kosek (2016); Liu and Hu (2016); Ozay, Esnaola, Vural et al. (2016); Zanetti, Jamhour, Pellenz et al. (2016); Zheng, Yang, Niu et al. (2017); Zheng, Chen, Wang et al. (2018)], is the indispensable prior data required for the classifiers' training. Those should be as significant as possible so that they entirely present the consumers' profiles. Knowing that one consumer may present more consumption profiles according to period of day (the peak hours), weekends, holidays and seasons. In fact, it complicates the task of harmonizing different consumption windows in one training set. It is very difficult, if not impossible, to do so in reality in terms of a practical view of the obtained results in the proposed context.

One more challenge in the machine learning algorithm, is the sudden changes in normal consumption profiles because of the introduction of new appliances, or shutting down a functioning one. These behaviours could generally infer on the algorithm accuracy by classifying this new consumption patterns as abnormal which increases the false alarms [Zanetti, Jamhour, Pellenz et al. (2017)]. This constraint does not infer on the RSA-FDS algorithm at all, because it treats the real instantaneous reading neglecting any other prior data.

Furthermore, the privacy is highly protected in the proposed approach comparing with the existing works, thanks to the RSA cryptographic advantages.

Tab. 8 illustrates a comparative frame of the RSA-FDS approach with the best works in that field.

Fig. 5 demonstrates a comparative illustration between the proposed RSA-FDS technic with the best knowing machine learning based FDS approaches. The most popular SVM algorithm is qualified as an supervised learning approach that performs a good results [Jokar, Arianpoo and Leung (2016)], meanwhile, an unsupervised technics is applied in Zanetti et al. [Zanetti, Jamhour, Pellenz et al. (2017)] in which the well-known FCM (fuzzy c-mean) algorithm outperforms other clustering methods in that field.

The average results of SVM approach shows that the false alarms are slightly important comparing with the FCM and RSA-FDS, which could be linked with the short changes in the normal consumption profiles, that happens with the introduction of new appliance into household, or removing an existing one as discussed previously. The FPR has decreased with the [Zanetti, Jamhour, Pellenz et al. (2017)] proposition to 7%, that introduced the short lived (SL) patterns with the FCM clustering algorithm which improves the robustness against normal changing profiles.

The proposed RSA-FDS algorithm outperforms the existing technics by removing the false alarms, because it does not care about any disturbing in the electricity consumption profiles, hence, the only required data is just the instantaneous reading to find whether the real amount has been altered or no.

By means of the RSA-FDS, the confronted limits are successfully surpassed due to the FDM simplicity conception which does not require any prior data in order to perform the detection task. Moreover, the presented algorithm is able to detect any kinds of fraudulent attempts that may take place in any stage after recording the electricity consumption amount. One more benefit is that the original data still remain disguised when flowing through the AMI networks. It is impossible to guess the real amount without being decrypted. So, adversaries' attack objectives will very unlikely be reached with the FDM concept.

However, our technic discards the case where the energy fraud has occurred physically, or before reporting the real smart meter reading (e.g., shutting down the smart meter), that is will be the purpose of the future works.

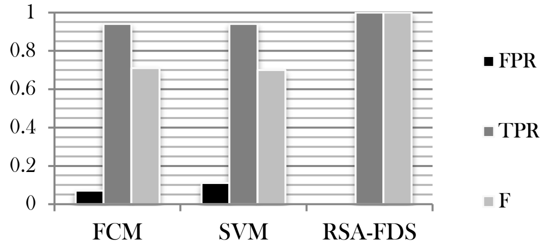

**Figure 5:** Metrics performances comparison between FCM, SVM and RSA-FDS

**Table 8:** Comparison with related works

| Parameters | Wide & Deep CNN [Zheng, Yang, Niu et al. (2017)] | SVM [Jokar, Arianpoo and Leung (2016)] | [Zanetti, Jamhour, Pellenz et al. (2017)] | RSA-FDS |
|---|---|---|---|---|
| **FPR** | NA | 0.11 | 2.7 | **0** |
| **ACC** | NA | 0.94 | 87.1 | **1** |
| **MAP@50** | 0.94 | NA | NA | **1** |
| **AUC** | 0.78 | NA | NA | **1** |
| **Detection delay** | daily | daily | Daily-weekly | **Instantaneous** |
| **Required prior data** | high | high | low | **No required data** |
| **Privacy preservation** | medium | low | high | **Extremely high** |
| **Robustness against normal changing pattern** | low | low | medium | **Extremely high** |

## 8 Conclusion

Numerous approaches have been performed aiming to combat electricity fraud in smart grid; hence, machine learning algorithms were widely applied in that context given their capability and performance in the classification of consumer's profiles, thus leading to the detection of abnormal behaviour in electricity consumption. However, the learning process requires prior data in order to build a meaningful learning database that should be able to dominate the client pattern variation over time. In fact, this task represents a serious challenge to researchers. This paper offers an opportunistic frame by introducing a new FDM concept using only the instantaneous measurement of electricity conception as data require. This approach stands for encrypting the smart meter recordings twice before they are sent to the utility, where they will be decrypted and checked to verify the correctness of the communicated values. Benefiting from its robustness and strength against hacking, the RSA algorithm is applied in this way to develop the Fraud Detection System (RSA-FDS). Any kind of fraud attempts aiming at illegal manipulation of the communicated data have been successfully detected where the FPR=0% with a precision=100%. The RSA-FDS' simplicity of conception and its exemplary results highlight the fact that it is one of the most promoted approaches performed in the real world.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

**Abreua, J.; Pereira, F.; Ferrão, P.** (2012): Using pattern recognition to identify habitual behavior in residential electricity consumption. *Energy And Buildings*, vol. 49, pp. 479-487.

**Ahmad, T.; Chen, H.; Wang, J.; Guo, Y.** (2018): Review of various modeling techniques for the detection of electricity theft in smart grid environment. *Renewable And Sustainable Energy Reviews*, vol. 82, pp. 2916-2933.

**Batista, N. C.; Melício, R.; Mendes, V. M. F.** (2014): Layered smart grid architecture approach and field tests by zigbee technology. *Energy Conversion And Management*, vol. 88, pp. 49-59.

**Bian, D.; Kuzlu, M.; Pipattanasomporn, M.; Rahman, S.** (2014): Analysis of communication schemes for advanced metering infrastructure (ami). *IEEE PES General Meeting Conference & Exposition*, pp. 1-5.

**Chen, P. Y.; Yang, S.; McCann, J. A.; Lin, J.; Yang, X.** (2015): Detection of false data injection attacks in smart-grid systems. *IEEE Communications Magazine*, vol. 53,no. 2, pp. 206-213.

**Department of Industry, I. a. S.** (2014): Electricity consumption benchmarks. http://data.gov.au/dataset/0f3d60db-bd63-419e-9cd9-0a663f3abbc9.

**Diffie, W.; Hellman, M.** (1976): New directions in cryptography. *IEEE Transactions On Information Theory*, vol. 22, no. 6, pp. 644-654.

**Fawcett, T.** (2006): An introduction to roc analysis. *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874.

**Han, W.; Xiao, Y.** (2017): A novel detector to detect colluded non-technical loss frauds in smart grid. *Computer Networks*, vol. 117, pp. 19-31.

**Huang, Y.; Tang, J.; Cheng, Y.; Li, H.; Campbell, K. A. et al.** (2014): Real-time detection of false data injection in smart grid networks: an adaptive cusum method and analysis. *IEEE Systems Journal*, vol. 10, no. 2, pp. 532-543.

**Jian, R.; Lu, R.; Wang, Y.; Luo, J.; Shen, C. et al.** (2014): Energy-theft detection issues for advanced metering infrastructure in smart grid. *Tsinghua Science And Technology*, vol. 19, no. 2, pp. 105-120.

**Jokar, P.; Arianpoo, N.; Leung, V. C.** (2016): Electricity theft detection in ami using customers' consumption patterns. *IEEE Transactions On Smart Grid*, vol. 7, no. 1, pp. 216-226.

**Khorshidi, R.; Shabaninia, F.** (2015): A new method for detection of fake data in measurements at smart grids state estimation. *IET Science Measurement & Technology*, vol. 9, no. 6, pp. 765-773.

**Kosek, A. M.** (2016): Contextual anomaly detection for cyber-physical security in smart grids based on an artificial neural network model. *Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*, pp. 1-6.

**Kosut, J. P.; Santomauro, F.; Jorysz, A.; Fern´andez, A.; Lecumberry, F. et al.** (2015): Abnormal consumption analysis for fraud detection: Ute-udelar joint efforts. *IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM)*, pp. 887-892.

**Liu, Y.; Hu, S.** (2016): Cyberthreat analysis and detection for energy theft in social networking of smart homes. *IEEE Transactions On Computational Social Systems*, vol. 2, no. 4, pp. 148-158.

**LLC, N. G.** (2015): Emerging markets smart grid: outlook 2015. Northeast Group, LLC.

**Mahmud, R.; Vallakati, R.; Mukherjee, A.; Ranganathan, P.; Nejadpak, A.** (2015): A survey on smart grid metering infrastructures: threats and solutions. *IEEE International Conference on Electro/Information Technology*, pp. 386-391.

**Milanov, E.** (2009): The rsa algorithm. RSA Laboratories.

**Nasim, B. M.; Jelena, M.; Vojislav, B. M.; Hamzeh, K.** (2014): A framework for intrusion detection system in advanced metering infrastructure. *Security and Communication Networks*, vol. 7, no. 1, pp. 195-205.

**Ozay, M.; Esnaola, I.; Vural, F. T. Y.; Kulkarni, S. R.; Poor, H. V.** (2016): Machine learning methods for attack detection in the smart grid. *IEEE Transactions On Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773-1786.

**Skopik, F.; Ma, Z.** (2012): Attack vectors to metering data in smart grids under security constraints. *Computer Software and Applications Conference Workshops (COMPSACW)*, pp. 134-139.

**Tariq, M.; Poor, H. V.** (2016): Electricity theft detection and localization in grid-tied microgrids. *IEEE Transactions On Smart Grid*, vol. 9, no. 3, pp. 1920-1929.

**Varun Badrinath, K.; Gabriel, A. W.; William, H. S.** (2015): Pca-based method for detecting integrity attacks on advanced metering infrastructure. *12th International Conference on Quantitative Evaluation of Systems*, pp. 70-85.

**Yaacoub, E.; Abu-Dayya, A.** (2014): Automatic meter reading in the smart grid using contention based random access over the free cellular spectrum. *Computer Networks*, vol. 59, pp. 171-183.

**Yang, X.; Zhao, P.; Zhang, X.; Lin, J.; Yu, W.** (2016): Toward a gaussian-mixture model-based detection scheme against data integrity attacks in the smart grid. *IEEE Internet of Thing Journal*, vol. 4, no. 1, pp. 147-161.

**Zanetti, M.; Jamhour, E.; Pellenz, M.; Penna, M.** (2016): A new svm-based fraud detection model for ami. *International Conference on Computer Safety, Reliability, and Security*, pp. 226-237.

**Zanetti, M.; Jamhour, E.; Pellenz, M.; Penna, M.; Zambenedetti, V. et al.** (2017): A tunable fraud detection system for advanced metering infrastructure using short-lived patterns. *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 830-840.

**Zheng, K.; Chen, Q.; Wang, Y.; Kang, C.; Xia, Q.** (2018): A novel combined data-driven approach for electricity theft detection. *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1809-1819.

**Zheng, Z.; Yang, Y.; Niu, X.; Dai, H. N.; Zhou, Y.** (2017): Wide & deep convolutional neural networks for electricity-theft detection to secure smart grids. *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1606-1615.