Advanced Monitoring Based Intrusion Detection System for Distributed and Intelligent Energy Theft: DIET Attack in Advanced Metering Infrastructure

Manali Chakraborty(✉)

Department of Computer Science and Engineering,
University of Calcutta, Kolkata, India
manali4mkolkata@gmail.com

Abstract. Power grid and energy theft has an eternal relationship. Though we moved towards Smart Grid, with an expectation for a more efficient, reliable and secure service, so does the attackers. Smart Grid and AMI systems incorporate a good number of security measures, still it is open to various threats. Recent attacks on Smart Grids in U.S., Gulf State and Ukraine proved that the attacks on the grid have become more sophisticated. In this paper we have introduced a new, distributed and intelligent energy theft: DIET attack and proposed an advanced Intrusion Detection System to protect AMI system. The proposed IDS can perform a passive monitoring on the system as well as detect attackers. This features make this IDS more robust and reliable.

Keywords: Intrusion detection system · Non technical loss
Energy theft · AMI · Smart Grid · Trust

1 Introduction

Smart Grid is meant to modernize traditional power grids with the two-way data communication along with energy supply. In order to enhance the functionalities of the network, Smart Grid offers several applications to help both customers and utilities to optimize the energy usage and billing. Advanced Metering Infrastructure or AMI is one of the most important features of Smart Grid [1]. It establishes a direct communication between customers and utilities, including, meter readings at periodic intervals (sometimes on demand) to the Data Collection Units or DCUs, updated electricity tariffs at regular intervals to smart meters, electricity outage alert messages and sometimes it upgrades the meter firmware [2]. However, due to the unique characteristics of AMI, such as complex network structure, resource-constrained smart meter, and privacy-sensitive data, it is an especially challenging issue to make AMI secure. Energy theft is one of the most important concerns related to the smart grid implementation. It is estimated that utility companies lose more than $25 billion every year due

M. L. Gavrilova et al. (Eds.): Trans. on Comput. Sci. XXXI, LNCS 10730, pp. 77–97, 2018.
https://doi.org/10.1007/978-3-662-56499-8_5

to energy theft around the world [1]. Energy theft may become an even more serious problem since the smart meters used in smart grids are vulnerable to more types of attacks compared to traditional mechanical meters. The unique challenges for energy theft in AMI call for the development of effective detection techniques.

Generally, energy theft can be accomplished using three ways [1,3]. Firstly, by interrupting smart meters from recording correct electricity usages, secondly, by forging demand information in smart meters and lastly by injecting false/bad data in the communication line. Now, the first attack is the only one that exists for the traditional meters, the other two attacks are exclusively for smart meters. Besides, premise of energy theft in AMI can also be classified in three categories:

- Type-1: where the attacker modifies its own smart meter to maximize its individual gain.
- Type-2: where the attacker modifies numerous smart meters in his neighborhood to either maximize its personal gain or penalize the utilities.
- Type-3: the cooperative attack, where a bunch of attackers create a chain of attacks on a large scale to immobilize the system in a short time interval.

In the first scenario, it is quite easy to detect the attack by analyzing the electricity usage pattern. Classification based detection schemes are quite suitable for these types of attacks. Besides, there exist several works to detect energy thefts which belong into the first two category. However, these types of attacks can be made more intelligent and difficult to detect, if implemented wisely. In this paper, we have proposed an attack model for distributed and intelligent electricity theft and proposed a two tier trust based intrusion detection system. The third part: collaborative attack is much complicated to implement as well as detect. We have considered these types of attacks as a part of our future extension of this paper.

There exist a good number of research works addressing solutions towards energy theft problem in Smart Grid. In [8] authors proposed a Support Vector Machine (SVM) based detection model to construct users' load profile pattern and then detect deviations from the standard pattern in order to identify abnormal behavior. Besides, to improve the performance of this model, authors incorporate fuzzy systems. The complete detection model identifies abnormal behaviors in the grid by comparing current load with recorded load profile and other additional information. Authors in [19] proposed an Auto Regressive Moving Average (ARMA) based model to analyze the probability distributions of the normal and malicious consumption patterns of users. They have applied the generalized likelihood ratio (GLR) test to detect energy theft attacks. The proposed work is heavily dependent on the data capturing accuracy of the ARMA model. Besides, it is based on the assumption that the attacker would always choose to decreases the mean value of the real consumption. Works presented in [9,10] also proposed a detection mechanism based on pattern matching and data classification. First, they proposed an classification method based on SVM and Rule based systems [9]. Then in [10], they introduce High Performance Computing (HPC) based algorithms to enhance the performance of their previous model. They have implemented some parallelized encoding algorithms to speed up the data classification, analyze and detection process. They have been able to differentiate the behavior of fraud customers from genuine users, using this model. AMIDS [11] is another AMI Intrusion Detection model, where a data mining technique based Non Intrusive Load Monitoring (NILM) system can collect data from three different sensors. These sensors gather data to identify cyber attacks, physical attacks and power measurement based anomaly. Authors of this paper claim that the proposed intrusion detection system (IDS) can detect several attacks by using information fusion from different sensors and correlation of different alert triggers. A Radio Frequency IDentification (RFID) based theft detection technique is proposed in [12]. The proposed system is divided in two parts: ammeter inventory management and ammeter verification control. RFID tags are attached with meters and used to detect energy theft. In addition, the reader acquires the information transmitted from the tag and sends it to the company's ERP system through the network to determine whether it is the approved tag or a different one placed by electricity thieves. Although the RFID technology can be used to detect energy theft, the utility companies have to pay extra cost to install the system. In order to find out whether implementing RFID technology is beneficial for the utility company, cost-benefit theory is used to analyze different value changes caused by the proposed system. Authors of paper [13] proposed a rather simple approach. They compare the meter readings of users with utilities reading. If the difference exceeds a threshold value, then that meter is marked as malicious and the connection will be terminated immediately.

The detection methods for energy theft can be broadly categorize into three types [1,20]: classification or statistical methods based detection techniques, monitoring based detection techniques and game theory based detection techniques. Classification based methods apply data mining methods and machine learning to energy usage patterns, collected from smart meters. They detect attacks by finding the deviation from the original data. These methods are cost effective and can be implemented easily. However, due to its lack of consideration for innovative and adaptive attack techniques, often some intelligent and minute attacks remain undetectable. Besides the false positive rates are on a higher side for these type of methods. Monitoring based techniques use sensor nodes, RFIDs and sometimes other smart meters to monitor the state of the network to detect the attack. This method has a better detection rate and lower false positive rate than the previous one. Continuous monitoring ensures the detection of very minute changes in the system. However, the implementation and maintenance cost of such system can become a disadvantage for implementation. Lastly, game theory based methods [14,15] are new in this domain, very few works have been done to detect energy theft. Planning the strategies for each player and formulated their goals can be a bit tricky in Smart Grid environment. The rules of the games should update simultaneously according to the change of situations in network and characteristics of players. Besides, these types of methods have greater false positive rates than monitoring based methods. It may be summarized that the monitoring based methods are best suitable to detect intelligent and minute attacks, providing the implementation and maintenance costs are minimized. Thus the main goal is to propose an effective monitoring system which ensures the trade-off between cost optimization and detection efficiency.

In this paper, we have introduced a new attack type specific to Smart Power Grid. We propose to call it Distributed and Intelligent Energy Theft (DIET) attack. Further, we have proposed a two-tier solution to detect the proposed DIET attack and perform a passive monitoring on the system, to provide an additional level of security. We have simulated DIET attack and the proposed detection mechanism using Qualnet 5.2 simulator [16].

The rest of the paper is organized as follows: Sect. 2 describes the network infrastructure for our proposed solution, Sect. 3 elaborates the proposed attack scenario, the working mechanism of our proposed IDS is described in Sect. 4, whereas, the simulation results and performance analysis is presented in Sect. 5, and finally Sect. 6 concludes this work.

2 Smart Infrastructure for Communication

Figure 1 shows the communication architecture of Smart Grid. Smart-energy Utility Network (SUN) hierarchically consists of three components: Home Area

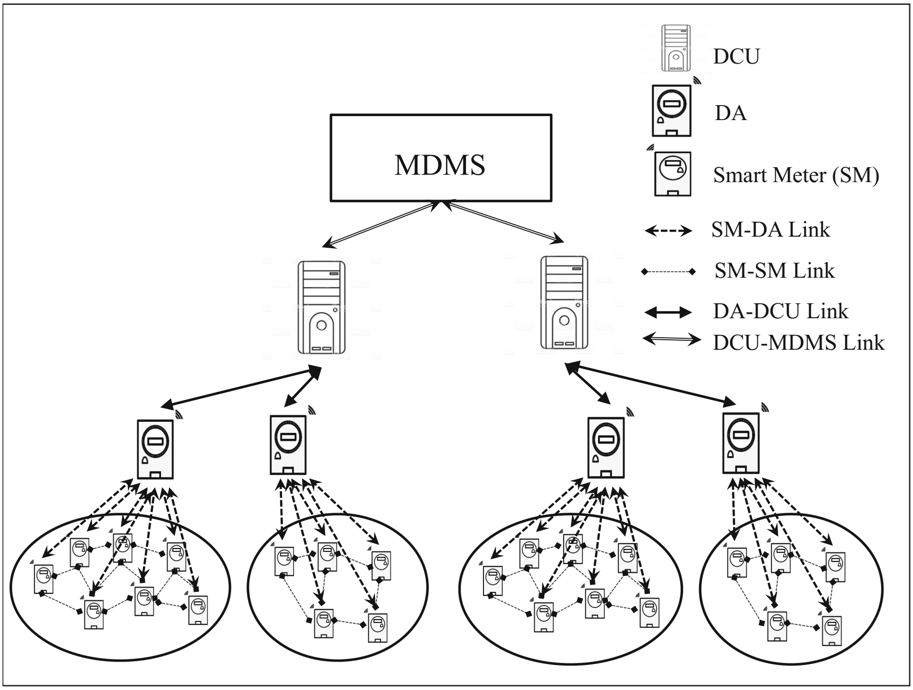

Fig. 1. Communication architecture of AMI in smart grid.

Network (HAN), Neighborhood Area Network (NAN), and Wide Area Network (WAN) [17]. The HAN provides the communication between the Smart Meters in a home and other appliances in that home. The NAN connects SMs to the Data Aggregators (DAs) and Data Collection Units (DCUs), and WAN provides access between the DCUs and Meter Data Management System (MDMS). DAs collect data from hundred of SMs registered under it and send them to DCUs. DCUs are responsible for communication with MDMS. Smart Grid has a quasi hierarchical structure, where the number of intermediate levels in the network varies with demographic and socio-economic condition of any particular region [7]. The smart meters act as hosts in a network, DCUs are the routers of the network and DAs are intermediate connectors. We assume that Smart Meters are managed by its immediate upper level DA or DCU, depending on the hierarchy. We assume that Smart Grid is a cluster based network, where each DA acts as a cluster head and can accommodate utmost 1000 of Smart meters. When a SM X is installed in a grid, it should find a DA to bind with. X will continue to communicate through DA in the network.

Firstly, we propose Near Term Digital Radio architecture [21] for the Clustering scheme used by our IDS. In this specific architecture, the cluster is divided into several physical subnets. Each subnet consists of the Cluster Head or DA and the Cluster Members or SMs which are at one - hop distance from the DAs. All SMs within a cluster can communicate with the DA using the same frequency and can communicate with other SMs within the same subnet using another frequency. These two channels are assumed to have different frequencies, so that there is no interference. DAs can also communicate with its upper level DCUs and DCUs with MDMS. The MDMS sends its messages to the DCUs, which then propagates the information through other DAs and finally it reaches to SM.

Communication between DAs, DCUs and MDMSs are supposed to be secure. However, SM to SM communication link may be compromised by attacker. In this paper, we have proposed a new Smart Grid specific energy theft (DIET) attack. In order to model this attack scenario we have used several network attacks, such as, extracting meter credentials and Man in the Middle (MITM) attack.

2.1 Information Stored at Smart Meters

The basic job of a smart meter is to track the energy usage of its customer and communicate with a DA. Generally a SM send their electricity usage after every 15 min of interval to the cluster head (DA) and receives various instructions from the DA. Now, in order to implement our IDS, we assume that,

- Every SM has an unique MAC address within its cluster.
- A smart meter can acquire its neighbor SMs' address through Neighbor Discovery procedure and can communicate with them over a wireless medium.

- A SM send a Electricity Usage (EU) message to the DA after every 15 min and broadcasts the same message at the same interval so that all of its neighbors can keep a tab of its electricity usage.
  $EU_{i,t}$ denotes the energy usage of SM i at time interval *t–(t–1)*, where i denotes it unique MAC address.
- Every SM maintains an array, Neighbor's Electricity Usage (NEU) for storing the information about the energy usage of its neighbors. $NEU_{i,t}[j]$ denotes the electricity usage of SM with unique MAC address j, at time t as stored at smart meter i.

Every time when a SM_i receives a broadcasted message from its neighbor SM_j at time t it adds the value of its electricity usage with the previously stored value for SM_j in the array, and updated its NEU as,

$$NEU_{i,t}[j] = NEU_{i,t-1}[j] + EU_{j,t}$$

Where, *t–(t–1)* = 15 min.
Thus, a SM stores its neighbor SMs' electricity usage in a cumulative array.

2.2 Information Stored at Cluster Heads

Cluster heads acts as a bridge between customers and utility. The main function of cluster heads, or DAs, or DCUs (depending on the hierarchical structure of the Grid) is to receive the electricity usage information of its SMs and provide billing information, electricity pricing and various informations to the SMs, depending on the applications. Besides cluster heads are also responsible for analyzing the data and detect any anomalies or abnormal behaviors in its cluster and report to the MDMS. We assume that,

- Each cluster head maintains a Smart Meter Connection (SMC) graph to store the topological information about its cluster. The graph is represented by an array of linked lists, where the size of the array defines the total number of SMs in the cluster and the size of each individual list represents the number of neighbors of that SM.
- Besides, every DA, i.e., cluster head will maintain a 2-D array ES(N,T) of energy supplied to each smart meters in its cluster, Where N is the total number of smart meters, registered with the cluster head and T denotes time. ES[i][t] denotes the energy supplied to the smart meter with MAC address i at t time-stamp.

Each cluster head communicates with the SMs within its cluster, and then transmits the aggregated data to its upper level DCU or MDMS.

Data structures, EU, NEU and ES are initialized after every 4 h. And the Smart meter Connection (SMC) graph is updated after the joining of every new SM in the cluster.

3 Proposed DIET Attack Model

In [3], the authors have addressed the issue of energy theft in smart grids using AMI. However, the authors elaborate on Type-1 attacks only and how they can be achieved in the AMI. The implementation of such attacks requires the attacker to hack into his Smart Meter using a Man-in-the-Middle attack. Once the attacker has his Smart Meter credentials, he can drop packets, inject new packets, as well as modify the usage information stored within the Smart Meters.

However, the situation is quite different for Type-2 attacks. Here, the attacker tries to hack and modify the data packets of its neighboring smart meters and then forward those malicious packets to DCU. Now, the first level of security can be easily provided by the AMI by implementing a modified version of the Needham Schroeder protocol [18] (like the Kerberos protocol or the Needham-Schroeder-Lowe protocol) to prevent replay attacks and Man-in-the-middle attacks. This prevents any eavesdropper from spoofing its neighboring Smart Meters during mutual authentication. Once this is ensured our only concern remains in securing the network communication. Whenever a Smart Meter tries to send a periodic update to the DCU about it's usage, the attacker may intercept this packet, drop it from the network and inject a new packet (with the neighbors credentials) with modified usage statistics.

In this paper, we have considered a special case of Type-2 energy theft attack. The Type-2 energy theft attack can be made quite difficult, if the attacker modifies the energy usage of a meter with a very negligible amount. 6% to 8% Technical Loss (TL) in transmission and distribution (T&D) is considered as normal in traditional power grid [4], but with Smart Grid the TL in T&D is reduced to 4% to 6% [5], i.e., if the allocated energy is 10 kW for a particular smart meter, then 9.4–9.6 kW of electricity is expected to be used by the smart meter. Now, suppose a smart meter registered 9.55 kW of electricity usage and sends it to the DCU. The attacker captures the packet and modifies the data to 9.45 kW. Apparently, the modified amount is so negligible to the customer that it would not bother him while billing. The DCU would also not be able to detect any anomaly. On the other hand, if the attacker modifies 100 smart meters like this, then it would create a 10 kW Non-Technical Loss (NTL) in the system.

There can be two intentions behind this type of attack: the attacker can either maximize its personal gain by reducing its electricity usage by the same amount as stole from the neighboring meters, or just minimize utility's gain by introducing a generous amount of NTL in the system.

- Personal Gain: The attacker may reduce its usage statistics by X% and uniformly distribute this power consumption value among its neighboring Smart Meters. From the DCU's perspective, the total power consumed by the Smart Meters appears to be proportional to the Power allocated to that DCU. The energy theft goes undetected. Here, the utility company does not bear the brunt of the attack.
- Utility Loss: In this other type of attack, the attacker does not look for personal gains; but is rather motivated by a more malicious intent of inflicting

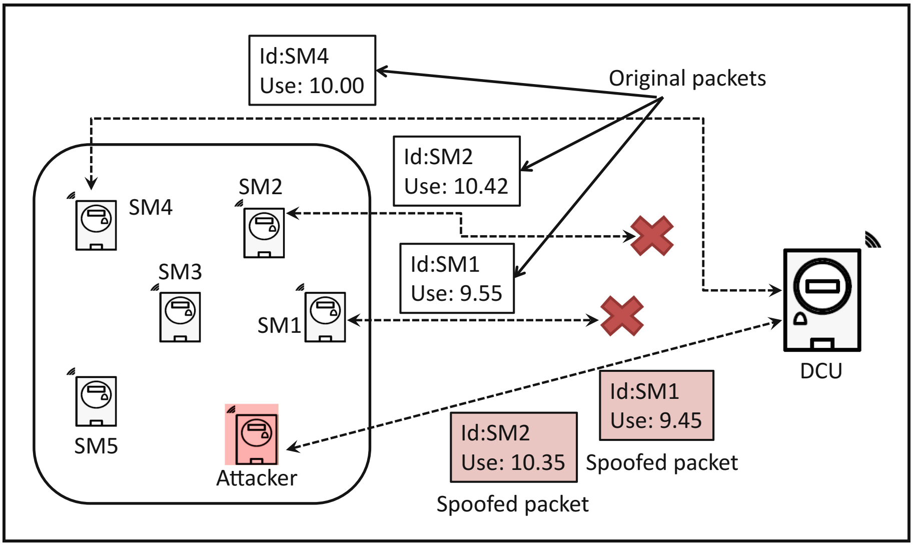

Fig. 2. DIET attack model.

financial losses to the utility. The attacker achieves this by considering the Technical Loss during Transmission and Distribution. It drops all packets from neighboring Smart Meters containing usage statistics and injects new packets having usage statistics slightly less than the original (within the TL threshold). The DCU interprets this as TL during T&D although consumers have consumed this power but have not been billed for the same.

Figure 2 explains our DIET attack scenario. Here, the attacker captures the data packets from SM 1 and 2, and modifies them slightly, so that it lies within the TL threshold and send them to the DCU. The transmission between SM 4 and DCU remains secure.

This attack will proved to be more effective in densely populated areas, where the attacker can have a huge number of smart meters as its neighbors Implementation of this attack become easier in a urban locality with numerous multi storied buildings, where a huge number of smart meters are placed across a long vertical line, but in a small horizontal section. As the attacker can access numerous SMs within its neighbor proximity, the scale of attack can be made more devastating in such scenarios.

4 Description of the Proposed IDS

In order to detect the DIET attack, we have proposed an IDS model. The detection mechanism will be performed in the cluster heads after certain time intervals. The IDS can detect Type-1 and Type-2 energy theft attacks. We assume that the Intrusion Detection System (IDS) will be running periodically with a time interval of 4 h.

4.1 Used Parameters in the IDS

We would first like to define the parameters used in the proposed IDS before explaining the working principle of our algorithm.

1. CH = Cluster Head.
2. SM = Smart Meter, SM_i, SM_j denotes smart meters with MAC address i and j respectively.
3. $EU_{i,t}$ denotes the electricity usage of SM_i at time t.
4. δ defines the allowed technical loss margin for each SM.
5. $EURec_t[i]$ defines the total electricity usage of SM_i at time t as recorded by the CH.
6. $NEU_{i,t}[j]$, denotes the electricity usage of SM with unique MAC address j, at time t as stored at smart meter i.
7. DEP_VL_i denotes the *Dependability Factor* of SM_i.
8. DEP_TH denotes the threshold value for the *Dependability Factor*.
9. *Attacked Nodes* defines a list to store the SMs which have been attacked by the attacker.
10. *Negative Neighbors* of SM_i is the list of neighbour nodes which causes an anomaly in the detection phase of the IDS.
11. *Possible attacker Node* holds the MAC address of those SMs which show abnormal behaviour in terms of stored information of its neighbor SMs.
12. *Attacker Nodes* contains the nodes which are detected as attacker.

4.2 Working Principle of Proposed IDS

The working principle of our algorithm can be divided into two phases: Data processing phase and Detection phase.

Data Processing Phase: In data processing phase, each SM in a cluster send its electricity usage data to the cluster head at 15 min interval. Besides, they also broadcast the same message over a separate channel, meant for only SMs in a cluster. Upon receiving this messages, each SM update itself regarding its neighbors' usage history. Cluster heads perform a preliminary detection at every 15 min, to detect type-1 attacks. Besides, CH also stores this periodic usage values in order to maintain a consistent usage log of each SM. The flow diagram of information for data processing phase of IDS, among various components in the AMI communication hierarchy is depicted in Fig. 3.

Detection Phase: We assume that the detection phase will execute at every 4 h instead of 15 min. The reason behind this is to reduce the packet transmission overhead and network congestion. At every 4th hour CH will request its SMs to send their $NEU_{i,t}[j]$ array. Upon receiving this packets from all the SMs, CH

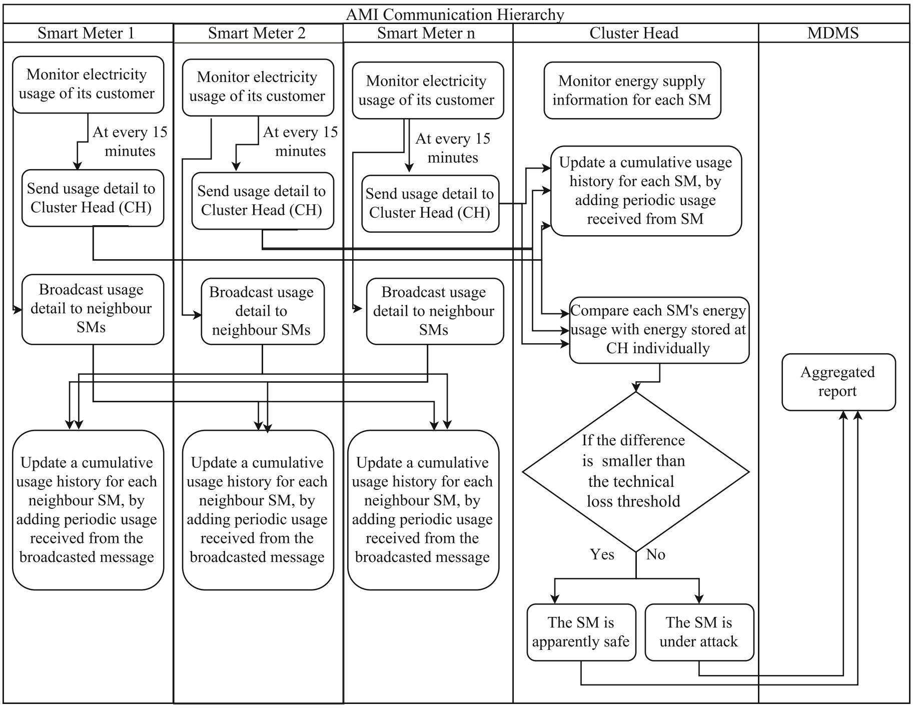
AMI Communication Hierarchy
Smart Meter 1
Smart Meter 2
Smart Meter n
Cluster Head
MDMS
Monitor electricity usage of its customer
At every 15 minutes
Send usage detail to Cluster Head (CH)
Broadcast usage detail to neighbour SMs
Update a cumulative usage history for each neighbour SM, by adding periodic usage received from the broadcasted message
Monitor energy supply information for each SM
Update a cumulative usage history for each SM, by adding periodic usage received from SM
Compare each SM's energy usage with energy stored at CH individually
If the difference is smaller than the technical loss threshold
Yes
No
The SM is apparently safe
The SM is under attack
Aggregated report

Fig. 3. Flow diagram for data processing phase of the proposed IDS.

extracts the neighbor information for each SM_i, and compares those values with $EURec_t[i]$. If all the neighbors' information match with $EURec_t[i]$, then SM_i is marked as safe, otherwise, the CH marks SM_i as *Attacked Node*. The CH then identifies the neighbors whose information matched with the stored information for SM_i and marks those as *Matched Nodes*, and the other neighbors of SM_i as, *Mismatched Nodes*. Now, the main idea of DIET attack is that the attacker is working alone. So, in case of an attack, the attacker's information will match with the recorded usage of SM_i, while the other neighbors will have a different information, but same collectively. Thus, the CH then compares the total numbers of *Matched Nodes* and *Mismatched Nodes* of SM_i and marks the minority group as *Negative Neighbors*. The intersection of *Negative Neighbors* of all SMs in a cluster is detected as *Attacker Nodes*. Whereas, the other *Negative Neighbors* are marked as *Possible Attacker Nodes* and the CH decreases the *Dependability Value* of these nodes at each detection cycle. These nodes can unmark themselves by showcasing good behavior and gaining *Dependability Value* at next detection cycles, or can be marked as *Attacker Nodes* if the *Dependability Value* goes under the threshold level.

The detailed procedure for attack detection is described in Fig. 4.

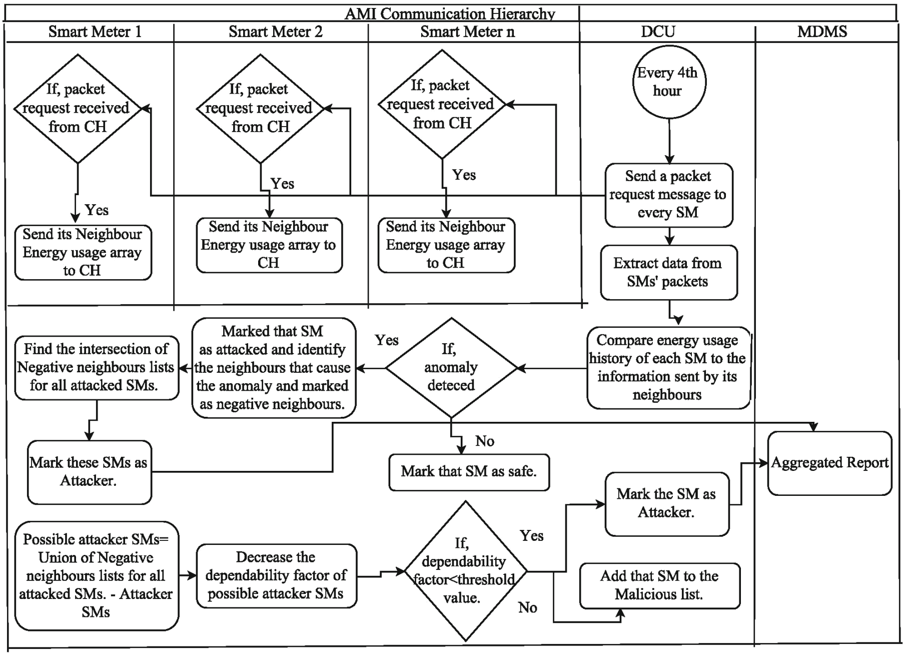
AMI Communication Hierarchy
Smart Meter 1
Smart Meter 2
Smart Meter n
DCU
MDMS
Every 4th hour
If, packet request received from CH
If, packet request received from CH
If, packet request received from CH
Yes
Yes
Yes
Send a packet request message to every SM
Send its Neighbour Energy usage array to CH
Send its Neighbour Energy usage array to CH
Send its Neighbour Energy usage array to CH
Extract data from SMs' packets
Find the intersection of Negative neighbours lists for all attacked SMs.
Marked that SM as attacked and identify the neighbours that cause the anomaly and marked as negative neighbours.
Yes
If, anomaly deteced
Compare energy usage history of each SM to the information sent by its neighbours
No
Mark these SMs as Attacker.
Mark that SM as safe.
Aggregated Report
Mark the SM as Attacker.
Possible attacker SMs= Union of Negative neighbours lists for all attacked SMs. - Attacker SMs
Decrease the dependability factor of possible attacker SMs
If, dependability factor<threshold value.
Yes
No
Add that SM to the Malicious list.

Fig. 4. Flow diagram for detection phase of the proposed IDS.

4.3 Algorithm for the Proposed IDS

The algorithm for both the phases of our proposed algorithm is described in this section. Step 1 of this algorithm defines the data processing phase and the rest of the part is used to detect DIET attack. Here n is the total number of SMs under a Cluster Head and m is the maximum number of neighbors for each individual SMs among n SMs in the cluster

1. At every 15 min, when CH receives a $EU_{i,t}$ message from SM_i, it checks if $(ES_{i,t-1} - \delta <= EU_{i,t} <= (ES_{i,t-1}))$
   {
   then, apparently SM_i is safe;
   $EURec_t[i] = EURec_{t-1}[i] + EU_{i,t}$;
   }
   else, SM_i is under attack;
2. After every 4 h, CH will ask its SMs to send their NEU array;
3. After receiving all the arrays, CH checks
   for(i = 1 to n)
   for(j = 1 to m)

```
   {
   if ($EURec_t[i]$ == $NEU_{j,t}$[i])
   {
   match_count++;
   Add $SM_j$ to the list of *Matched Nodes*;
   }
   else
   {
   mismatch_count++;
   Add $SM_j$ to the list of *Mismatched Nodes*;
   }
   }
4. if(match_count==m)
   then, $SM_i$ is safe;
   else,
   {
   add $SM_i$ to the list of *Attacked Nodes*;
   if(mismatch_count > match_count)
   Mark *Matched Nodes* as *Negative Neighbors*;
   else
   Mark *Mismatched Nodes* as *Negative Neighbors*;
   }
5. Let, $r$ be the total number of attacked nodes;
6. *Attacker Nodes* = Intersection of *Negative Neighbors* for all the $r$ SMs in the
   list *Attacked Nodes*;
7. *Possible Attacker Node* = Union of *Negative Neighbors* for all the $r$ SMs in
   the list *Attacked Nodes* − *Attacker Nodes*;
8. Decrease the DEP_VL of every SMs in *Possible Attacker Node* list.
   if, $DEP\_VL_x$ < DEP_TH
   Mark $SM_x$ as *Attacker Node*;
   else
   Add $SM_x$ to the list of *Malicious Nodes*;
9. End.
```

Once, the algorithm detects the attacker nodes and put them in *Attacker Nodes* list, the CH, then checks the *Attacked Nodes* list and replace their forged energy usage values by the original usage statistics with the help of its neighbors' information.

5 Simulation Results

We have implemented The DIET attack and the proposed IDS in Qualnet 5.2. The simulation settings and the used scenario are described in Table 1.

Table 1. Parameter settings for simulation environment

Parameter	Value
Experimental area	$1500 * 1500m^2$
Running time for each simulation	100 s
Mac layer protocol	DCF of IEEE 802.11b standard
Network layer protocol	AODV
Traffic model	CBR
Number of CBR traffics	10% of the total number of nodes
Cluster head : Smart meter	1 : 10

5.1 DIET Attack Simulation

Firstly, the DIET attack is implemented and the results are analyzed. We have considered varying node density of 11 to 55 nodes for our experiment. Data has been collected for every variation and then the averaged values are plotted in the graph. In order to implement DIET attack, we assumed that each Cluster Head can have atmost 10 SMs under its surveillance, and the energy supplied to these SMs is remained fixed at 10 kW. δ is defined as 5% of supplied energy, i.e., 0.5 kW.

Figure 5 depicts the allowed, original and registered TL for implemented DIET attack. Energy theft = (registered TL − original TL). Now, the registered TL is still under the threshold of allowed TL, thats why the theft can not be detected by the system.

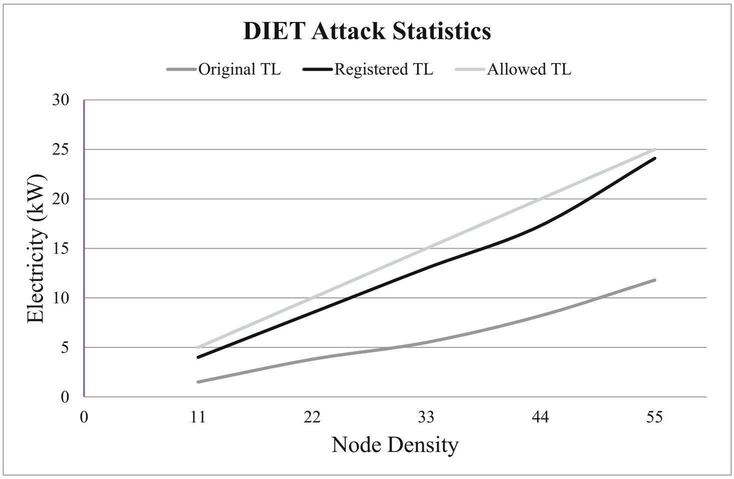

Fig. 5. Original technical loss and registered technical loss for DIET attack with varying node density.

5.2 IDS Implementation

In order to evaluate the performance of our proposed algorithm, we have considered three metrics: false positive, false negative and detection efficiency. False negative and false positive are both very important metric towards Smart Grid. Identify a legitimate SM as an attacker (false positive) can harm the customers', as well as utility's reputation and cause for temporarily disruption of service for that innocent customer. On the other hand, not being able to identify an attacker can lead to financial loss, malicious billing and can even cause havoc devastation.

We have considered four different attack scenarios to evaluate our algorithm.

- **Attack Scenario A** has 100% of DIET attacks, i.e., where an attacker modifies its neighbor SMs usage data, but within the TL threshold. We assume that typical Type-1 and Type-2 attacks are not present for this situation.
- **Attack Scenario B** has 50% of DIET attackers and another 50% of both Type-1 and Type-2 attackers.
- **Attack Scenario C** has 25% of DIET attackers and another 75% of both Type-1 and Type-2 attackers.
- **Attack Scenario D** has equal share of all the three attackers, i.e., 33.33% of DIET, Type-1 and Type-2 attackers.

False Positive: Our proposed algorithm does not identify any false positives for our entire simulation time with various node density and number of attackers It only detects genuine attackers and put them in *Attacker Nodes* list. However, it adds some genuine SMs in *Possible Attacker Nodes* list and decreases the DEP_VL for those nodes at some iterations.

False Negative: False negatives are used to measure the accuracy of the system. If the total number of attackers present in the system is x, and the IDS detects y of them, then the false negative can be calculated as:

$$\text{False Negative} = ((x-y)/x) * 100\%$$

Figure 6 depicts the total percentage of false negatives against varying node density for various attack scenarios. Number of attackers in the system are also increased proportionally with the number of nodes. Figure 6 shows that for attack scenario A and C, the false negative is null for the first two instances. However, it started increasing gradually thereafter, and reaches its peak when the node density is 55. After that the false negative tend to decrease in scenario A. Now, while analyzing the graph, we find that, for every instances, our proposed IDS either successfully marked every attacker node or add them to the list of *Possible attacker Nodes*. With 55 nodes in the scenario, the IDS is able to identify every attacker node as a possible attacker, however, due to the lack of enough neighbor support, it cannot mark the attackers immediately Though, the trust evaluation process will help them detect gradually. Thus, we can confirm that our proposed system can eventually detect all the attackers.

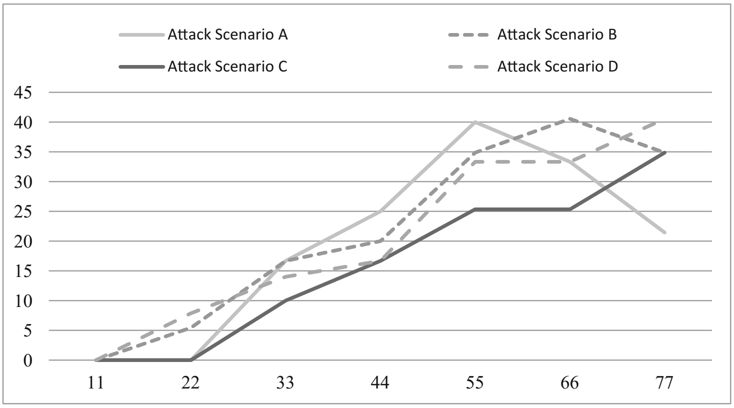

Fig. 6. False negatives vs Node density for different attack scenarios.

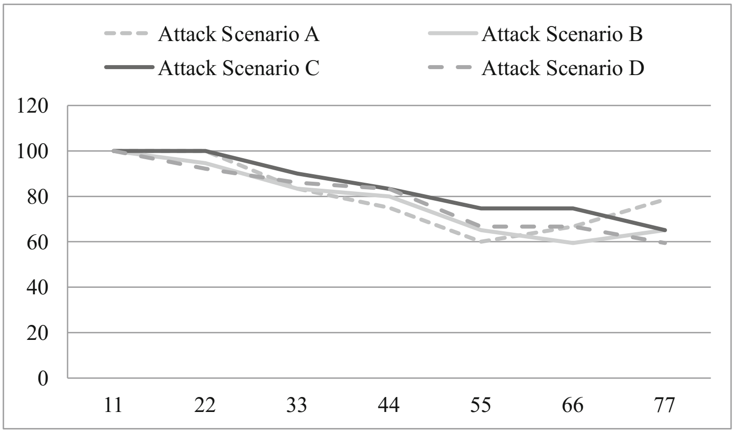

Fig. 7. Detection efficiency vs Node density for different attack scenarios.

Detection Efficiency: Detection efficiency can simply be calculated as, (100 - false negative), provided there is no false positives in the system.

Figure 7 shows the detection efficiency of our IDS. Since there is no false positives in our IDS, the graph for detection efficiency is simple reciprocal of false negative.

The detection efficiency for our proposed IDS remains 100% for smaller number of nodes (i.e., up to 22 nodes for our simulation scenario.). However, with increasing number of nodes and attackers, the detection efficiency tends to decrease gradually. The detection mechanism of our proposed IDS depends heavily on the anomalies in the data provided by SMs in a neighborhood. With fewer nodes in the scenario, it will be easier to analysis the data and hence the detection of an attacker. On the contrary, when the node density increases, it affects

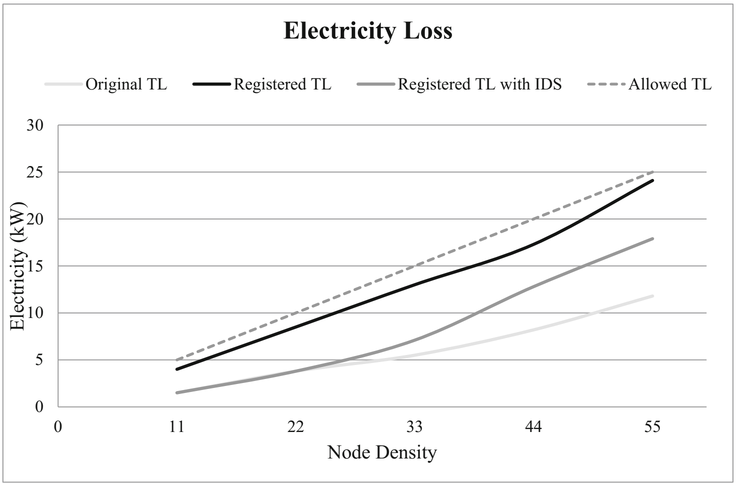

Fig. 8. Electricity loss in proposed IDS.

the complexity of data analysis and hence the detection efficiency. However, to address this situation and provide a stability in the system, our IDS marked all suspicious nodes as *Possible Attacker Nodes*, and decreases their dependability factor as well, so that the system can be aware of that nodes and do not let those nodes to further affect the decision making process. When the dependability factor goes beyond the threshold value, then only a node will be marked as attacker.

Energy Loss: Finally we measure the energy loss for our IDS, and Fig. 8 demonstrated that the proposed IDS is successfully able reduce the electricity loss due to DIET attack.

Comparison with Existing Works: In this section, we have done a detailed comparative analysis of our proposed IDS with a specification based IDS proposed in [22]. In this paper, authors deployed sensors in NAN to monitor the communication network and detect malicious activities in the AMI based on formal verification of the specifications and monitoring operations. Authors claimed that the proposed IDS can detect both known and unknown attacks in network level, including MITM, black hole attack etc. Since, our proposed IDS handles DIET attack, which in turn associates with MITM and stealing of meter data credential attacks, we consider the IDS, proposed in [22] as an appropriate choice for comparison. We have implemented the IDS of [22] for different attack scenarios, as mentioned in Sect. 5.2.

Figure 9 provides the comparative analysis of our proposed IDS and specification based IDS proposed in [22]. We have considered the performance of both the algorithms for four different attack scenarios and with seven different node

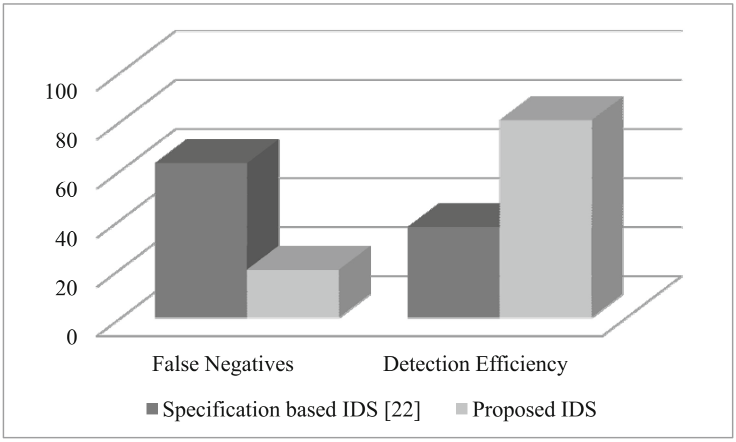

Fig. 9. Comparative results of false negatives and detection efficiency.

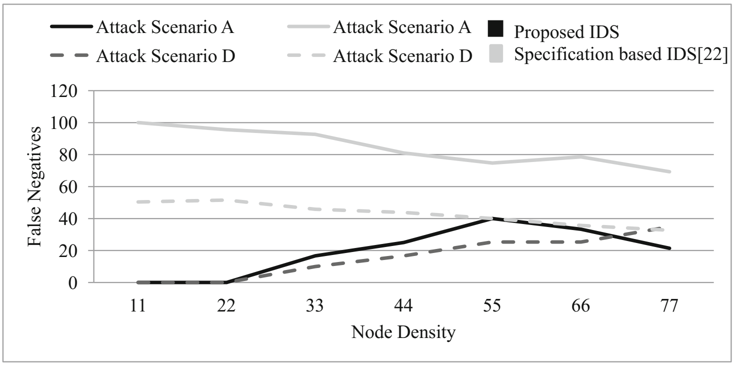

Fig. 10. Comparative analysis for different attack scenarios.

densities over 11 nodes to 77 nodes, and then averaged the results. The IDS in [22] monitors the state of communication network and verifies with existing specification rules. It does not consider the content of packets transmitted over the network, rather it keeps tab of total packet received and sent, time stamps of different events, frequency of packet transmission etc. Thus, it cannot detect Type-I attack, as in this case, a SM modifies its own packets and transmits them. On the other hand, in DIET attack, a SM steals its neighbor SMs' credentials, manipulates their meter data within tolerance level and retransmits to the DA. Thus, at the end points, the packet received and sent metrics for a particular SM will remain unaltered. Hence the attack will remain undetected. However, when node density increases, so does the traffic and if the sensors detect an abnormally large number of packet transmission on a particular channel, then it can detect the attack sometimes. However, for typical Type-II attack, the attack can be detected by considering cumulative energy usage parameters.

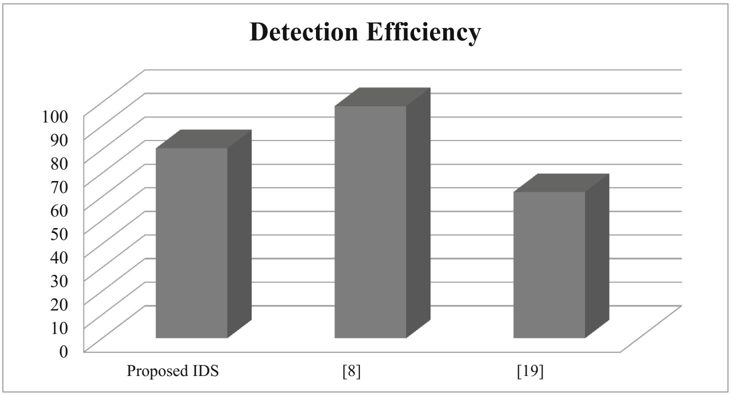

Fig. 11. Comparison of detection efficiency among different IDSs.

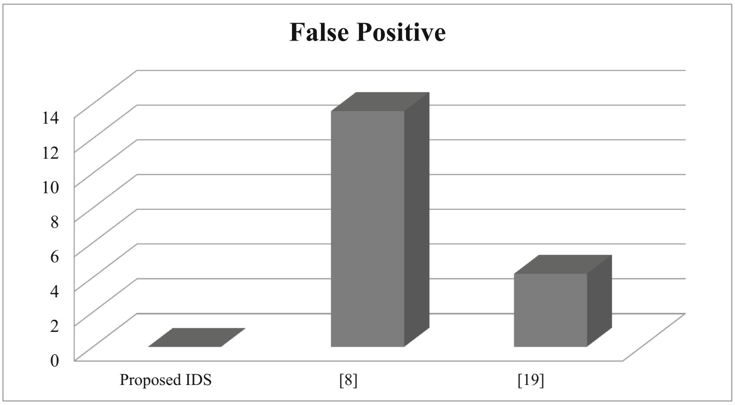

Fig. 12. Comparison of false positives among different IDSs.

Figure 10 gives the performance of two IDSs for attack scenario A and D. In attack scenario A, our proposed IDS performs much better than the other one. Specification based IDS [22] performs better with increased node density. Still it never performs like our proposed IDS. However, for attack scenario D, the IDS of [22] performs much better, but still the false negatives are much higher than our proposed IDS.

We have compared the performance our IDS with other existing IDSs. Figures 11 and 12 provide comparative analysis of our proposed IDS and two existing IDSs based on SVM [8] and ARMA-GLR [19] models respectively. Now, our proposed IDS offers better detection efficiency than [8], however, Fig. 11 shows that the [19] performs much better than our IDS. On the other hand Fig. 12 shows that both [8,19] have false positives, where our IDS has none.

As we already mentioned, false positives and false negatives are two important metrics to evaluate the performance of any IDS. Now, energy theft is an attack scenario which involves the customers directly. The attackers disguised themselves as customers. So we have to give extra care to detect attackers. False positives may harm the reputation of genuine customers, which incurs in bad reputation for the utilities. Thus, while dealing with energy theft attack, false positives are much more important than false negatives. Thus we designed our IDS in such a manner that it marks a node as attacker only after being 100% sure of that. Otherwise, it can mark suspicious nodes as possible attackers, and monitor their further behavior. This in turn justifies the results of Figs. 11 and 12.

6 Conclusions

Smart grid and especially AMI system enhances the efficiency, reliability, stability, security and economic facilities of traditional power grid systems. Advanced metering infrastructure (AMI) is arguably the most important and critical part of Smart Grid. AMI deals with the most sensitive informations in the Grid and transmits them through the network. There already exist a good number of security solutions for AMI. However the percentage of security attacks are also increasing day by day, and so does the innovative and intelligent ideas behind those attacks.

Energy theft is always a serious concern for power industry. With traditional power grid, tapping, physical tampering of meters are the common sources to theft. Smart Grid and AMI can mitigate these attacks, however, with the recent advancement in the technology, the attackers also invent newer and sophisticated ideas to attack the grid. In this paper, we have proposed a new attack situation named, DIET attack. Simultaneously, we have simulated this attack in QUALNET and analyze the effect on the grid. In order to detect DIET attack, we have proposed an advanced IDS.

Our IDS can successfully detect Type-1 and Type-2 attacks. Moreover, for some scenarios, the IDS cannot detect the attacker primarily, but it is been able to mark all of them as *Possible Attacker* and take precautionary measures against them. If those nodes continue to being malicious, then eventually the proposed IDS detect that node as attacker, otherwise, in case of a genuine node, the dependability factor will be increased with positive behavior. Besides, there exists lots of works for detecting energy theft, many of them are only capable to detect whether a theft happened or not. On the contrary, the proposed IDS can not only identify an intelligent theft situation, but can detect the attackers and mark possible attackers in the network as well.

As a future extension of this paper, we would like to merge our idea with some secure routing protocols like [6], where trust based evolution of nodes are performed for route selection to ensure a secure communication system. The collaboration of the proposed IDS with this type of routing protocols will confirm security from DIET attack at transmission time and improve the performance of the system, in terms of detection efficiency and false negatives.

Acknowledgments. This work is a part of the Ph.D. work of the author, a Senior Research Fellow of Council of Scientific & Industrial Research (CSIR), Government of India. We would like to acknowledge CSIR, for providing the support required for carrying out the research work.

References

1. Jiang, R., Lu, R., Wang, Y., Luo, J., Shen, C., Shen, X.S.: Energy-theft detection issues for advanced metering infrastructure in smart grid. Tsinghua Sci. Technol. **19**(2), 105–120 (2014)
2. Grochocki, D., Huh, J.H., Berthier, R., Bobba, R., Sanders, W.H., Crdenas, A.A., Jetcheva, J.G.: AMI threats, intrusion detection requirements and deployment recommendations. In: Smart Grid Communications (SmartGridComm), pp. 395–400. IEEE (2012)
3. McLaughlin, S., Podkuiko, D., McDaniel, P.: Energy theft in the advanced metering infrastructure. In: Rome, E., Bloomfield, R. (eds.) CRITIS 2009. LNCS, vol. 6027, pp. 176–187. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14379-3_15
4. ABB Inc.: Energy Efficiency in the Power Grid. ABB Inc., Fort Smith (2007)
5. U.S. Energy Information Administration. www.eia.gov
6. Chakraborty, M., Deb, N., Chaki, N.: POMSec: Pseudo-opportunistic, multipath secured routing protocol for communications in smart grid. In: Saeed, K., Homenda, W., Chaki, R. (eds.) CISIM 2017. LNCS, vol. 10244, pp. 264–276. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59105-6_23
7. Chakraborty, M., Chaki, N.: An IPv6 based hierarchical address configuration scheme for smart grid. In: Applications and Innovations in Mobile Computing (AIMoC), Kolkata, pp. 109–116 (2015). https://doi.org/10.1109/AIMOC.2015.7083838
8. Nagi, J., Yap, K.S., Tiong, S.K., Ahmed, S.K., Mohamad, M.: Nontechnical loss detection for metered customers in power utility using support vector machines. IEEE Trans. Power Deliv. **25**(2), 1162–1171 (2010)
9. Depuru, S., Wang, L., Devabhaktuni, V.: Support vector machine based data classification for detection of electricity theft. In: IEEE/PES Power Systems Conference and Exposition (PSCE), pp. 1–8 (2011)
10. Depuru, S., Wang, L., Devabhaktuni, V., Green, R.C.: High performance computing for detection of electricity theft. Int. J. Electr. Power Energy Syst. **47**, 21–30 (2013)
11. McLaughlin, S., Holbert, B., Zonouz, S., Berthier, R.: AMIDS: A multi-sensor energy theft detection framework for advanced metering infrastructures. In: IEEE Third International Conference on Smart Grid Communications (SmartGridComm), pp. 354–359 (2012)
12. Khoo, B., Cheng, Y.: Using RFID for anti-theft in a Chinese electrical supply company: A cost-benefit analysis. In: IEEE Wireless Telecommunications Symposium (WTS), pp. 1–6 (2011)
13. Xiao, Z., Xiao, Y., Du, D.H.C.: Non-repudiation in neighborhood area networks for smart grid. IEEE Commun. Mag. **51**(1), 18–26 (2013)
14. Amin, S., Schwartz, G.A., Tembine, H.: Incentives and security in electricity distribution networks. In: Grossklags, J., Walrand, J. (eds.) GameSec 2012. LNCS, vol. 7638, pp. 264–280. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34266-0_16

15. Cardenas, A.A., Amin, S., Schwartz, G., Dong, R., Sastry, S.: A game theory model for electricity theft detection and privacy-aware control in AMI systems. In: IEEE 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 1830–1837 (2012)
16. QualNet 5.2 Simulator: Scalable network technologies Inc.
17. Kim, M.: A survey on guaranteeing availability in smart grid communications. In: Advanced Communication Technology (ICACT), pp. 314–317 (2012)
18. Stallings, W.: Cryptography and Network Security: Principles and Practice, 5th edn. Prentice Hall Press, Upper Saddle River (2010). ISBN: 0136097049 9780136097044
19. Mashima, D., Cárdenas, A.A.: Evaluating electricity theft detectors in smart grid networks. In: Balzarotti, D., Stolfo, S.J., Cova, M. (eds.) RAID 2012. LNCS, vol. 7462, pp. 210–229. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33338-5_11
20. Jokar, P., Arianpoo, N., Leung, V.C.M.: Electricity theft detection in AMI using customers consumption patterns. IEEE Trans. Smart Grid **7**(1), 216–226 (2016)
21. Ruppe, R., Griswald, S., Walsh, P., Martin, R.: Near Term Digital Radio (NTDR) system. In: MILCOM 1997, pp. 1282–1287 (1997)
22. Berthier, R., Sanders, W.H.: Specification-based intrusion detection for advanced metering infrastructures. In: IEEE 17th Pacific Rim International Symposium on Dependable Computing, pp. 184–193 (2011)