A Novel Smart Energy Theft System (SETS) for IoT based Smart Home

Weixian Li, *Member, IEEE*, Thillainathan Logenthiran, *Senior Member, IEEE*, Van-Tung Phan, *Senior Member, IEEE*, and Wai Lok Woo, *Senior Member, IEEE*

***Abstract*—In the modern smart home, smart meters and Internet of Things (IoT) have been massively deployed to replace traditional analogue meters. It digitalises the data collection and the meter readings. The data can be wirelessly transmitted that significantly reduces manual works. However, the community of smart home network is vulnerable to energy theft. Such attacks cannot be effectively detected since the existing techniques require certain devices to be installed to work. This imposes a challenge for energy theft detection systems to be implemented despite the lack of energy monitoring devices. This paper develops an energy detection system called Smart Energy Theft System (SETS) based on machine learning and statistical models. There are 3 stages of decision-making modules, the first stage is the prediction model which uses multi-model forecasting System. This system integrates various machine learning models into a single forecast system for predicting the power consumption. The second stage is the primary decision making model that uses Simple Moving Average (SMA) for filtering abnormally. The third stage is the secondary decision making model that makes the final stage of the decision on energy theft. The simulation results demonstrate that the proposed system can successfully detect 99.96% accuracy that enhances the security of the IoT based smart home.**

***Index Terms*—Smart homes, Smart grid, Internet of things, Energy theft, Machine learning techniques**

I. Introduction

In the modern smart grid, massive deployment of advanced metering infrastructures (AMI) facilitate the efficient and reliable information exchange. The AMI can be divided into different sectors depending on the location which is crucial to end consumer. AMI includes smart meters and Internet of Things (IoT) monitoring devices that were able to collect data in large volumes and fast speed.

Smart home innovators today focus on system development, system architecture, communication protocols, and forecasting tools [1], [2]. These innovations provide home consumers with a better technology in terms of energy monitoring, control, and reliability. For example, Demand Side Management System (DSMS) was introduced to better manage and control power consumption for the smart homes [3]. This power conservation concept increased the research on improving DSMS methods like load-shifting, dynamic price management, forecasting demand, and demand response systems [4]–[6].

These advancements improved through the use of machine learning and statistical modelling. Algorithms such as Simple Moving Average (SMA), Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU) have been used in the energy efficiency sector [7]–[10]. However, it is still vulnerable to malicious behaviour such as energy theft.

Energy theft has been a rising issue for various countries around the world. Despite this, only a few preventive energy theft methods were created to combat the issue. Zhou, Y. et al. proposed a dynamic programming algorithm for leveraging probabilistic detection of energy theft in the smart home [11]. This proposed method requires the deployment of Feeder Remote Terminal Unit (FRTU) on top of a smart meter which incurs high costs for consumers. Additionally, it works only under the assumption that a smart meter is available.

Liu, Y. and Hu, S. proposed a detection technique that has a detection accuracy of 92.55% on average [12]. This proposed detection technique integrated Bollinger-bands-based detection with the partially observable Markov-decision process (POMDP). However, it does not reflect on all conditions of a house environment. Firstly the house demand data has consistent energy consumption throughout the entire 24 hours. It does not include any zero energy consumption for a particular hour. Another condition on the Bollinger Band method, the deviation can only be done in a consistent range of energy usage. However, if the range of energy usage became large, the Bollinger Band method could not be used due to its deviation.

This paper proposes a novel idea of Smart Energy Theft System (SETS) for the smart home. This energy theft detection algorithm is more efficient and reliable compared to previous methods. As a result of a non-intrusive method of data collection, the energy monitoring system was implemented in a real house in Singapore. The collected data includes Time series data power consumption from a non-controlled real-life house environment.

The remaining paper is organised as follows: Section II presents background information about the foundation of the Smart Energy Theft System (SETS). Section III shows the proposed methodology for Smart Energy Theft System (SETS). Section IV provides the simulation results of the proposed system. Finally, the paper is concluded in section V.

II. Background Information

A. Smart Homes

Smart Homes are created through implementation of Internet of Things (IoT) and smart meters [13]–[16]. In order to monitor and control the Advanced Metering Infrastructure

W. Li, T. Logenthiran, V.-T. Phan and W. L. Woo are with the School of Electrical and Electronic Engineering, Newcastle University, Singapore Campus, e-mail: (e-mails: w.li17@newcastle.ac.uk, t.logenthiran@ncl.ac.uk, vantung.phan@ncl.ac.uk and lok.woo@ncl.ac.uk.)

(AMI), Energy Management System (EMS) was an essential integration of the system infrastructure [17]–[20].

Demand Side Management System (DSMS) is included as a function of EMS [21]. Its functionality focuses mainly on managing the demand response and loads. It collects the demand information to dictate the optimal power usage such as implementing load-shifting to enable the use of electricity markets during peak and off-peak hours.

It allows users to conveniently dictate their smart appliances within the home area by using mobile devices. More advanced and developed systems could further analyse the data collected and make its own decision for the smart homes to operate in a cost-effective and energy-efficient method based on users' consumption patterns.

B. Energy Theft

Energy theft has become a serious issue in the smart grid community [22]. It has caused massive losses for many countries that exceed billions of dollar. Nowadays, a smart meter will be placed at the end of every distribution network to record power consumption and generates the energy reports remotely. An example of the home distribution network is shown in Fig.1.

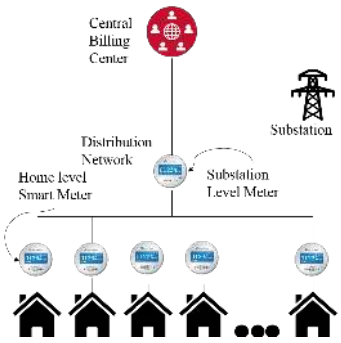

Fig. 1: Home distribution network

Energy theft methods involve hacking smart home appliance and most commonly direct hooking on other households electricity supplies. Other methods involved are tampering with the smart meter's software, mechanism, and manipulating data through cloud storage [23]. Thus, attackers can reduce their own electricity usage by manipulating other households through tampering and hacking to increase their electricity usage as the aggregate bill for all customers in the community remains the same [24]. Fig.2 shows an example of energy theft situation.

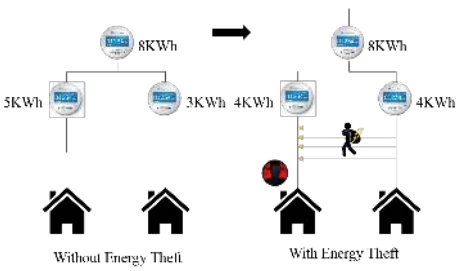

Fig. 2: Energy theft situation

The example shows that through energy theft, the higher consumption household can reduce their own power consumption through tapping on another household. It increases the electricity bills for the other household victim while reducing the energy theft culprit bills.

III. Proposed Smart Energy Theft System (SETS)

Fig. 3 shows the overall design of the proposed Smart Energy Theft System (SETS) for the smart homes. SETS is designed for detecting energy theft and alerting the consumers. It collects information from monitoring devices and analyses the data to detect energy theft.

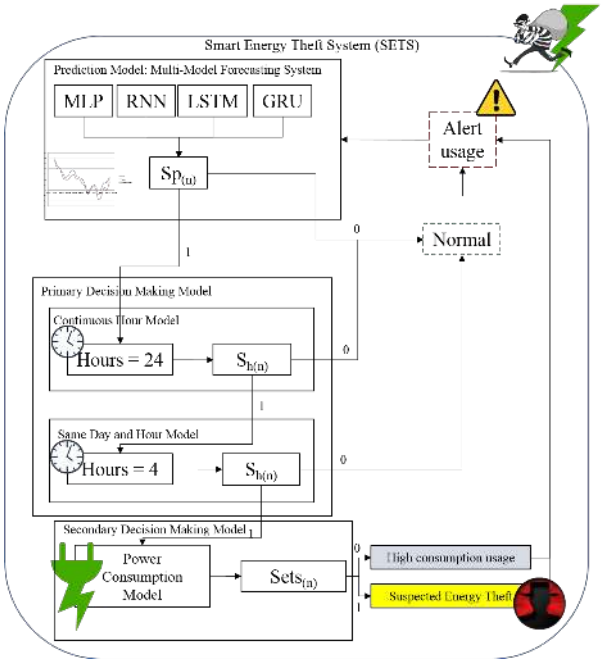

Fig. 3: Overall SETS architecture

The overall architecture comprises the following modules:

- Data Collection Module
- Prediction Model
- Primary Decision Making Model
  - Continuous Hour Model
  - Same Day and Hour Model
- Secondary Decision Making Model
  - Power Consumption Model

The data collection module collects the data for SETS. The first stage of SETS is the prediction model. The prediction model uses Multi-Model Forecasting System that comprises different machine learning methods: Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU). It predicts and compares the actual data to detect abnormally. Second stage of SETS is the primary decision making model. This stage uses a statistical model called Simple Moving Average (SMA) to filter the abnormally from the first stage.

Third stage of SETS is the secondary decision making model. This stage further filter from the second stage and decides whether energy theft had occurred. After taking the

final decision, the whole process will be repeated for the next incoming data. SETS is best implemented with an independent hardware system directly at the smart meters, this is because any interferences for energy theft regardless of tampering hardware or manipulation of data can be detected. It is more accurate compared to just monitoring the data from cloud or operator's database as many other factors may affect the analysis.

A. Data Collection Module

Demand Side Management System (DSMS) collates the information from various real-time monitoring smart devices in the house. The data collection module for setting up Smart Energy Theft System (SETS) is to get the real-time monitoring ready. Data collection module used a set of smart plugs called Aeon Labs Z-Wave UK Plug-in Switch plus Power Meter and the main controller was a VeraEdge Home Controller. Connectivity for data collection is shown in Fig. 4.

Fig. 4: Data collection system architecture

This system was placed on a Singapore smart home for collecting data through a non-invasive method of energy monitoring.

B. SETS

SETS detects unexpected energy theft from any form of malicious attack. This proposed system is designed with the following stages:

1) Stage 1: Prediction model: Multi-Model Forecasting System: The Prediction Model forecast the next 24 hours by using Multi-Model Forecasting System. Measured data is used for predictions and comparison to determine the energy theft situation.

a) Stage 1: Multi-Model Forecasting Systems and Algorithms: The Multi-Model Forecasting System uses different machine learning methods and utilises the most accurate model through the state of prediction model decision making condition $sp_{(n)}$. The forecasting systems Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU) are used at this stage and a brief description is as follows:

- Multi-layer perceptron (MLP)
  Artificial neural networks (ANN) are often called neural networks or multi-layer perceptron (MLP) to represent the most useful type of neural network. It is inspired by the biological architecture of the brain which can be used to solve difficult computational tasks. The goal is developing robust algorithms and data structures that can be used to solve difficult problems [25].

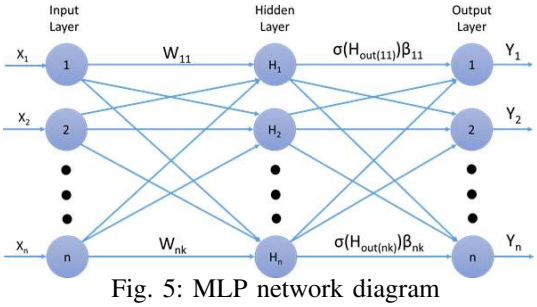

Fig. 5: MLP network diagram

Fig.5 shows the network of a typical MLP. The formulations [26] of the MLP are defined as follows:

$$H_{(out(nk))} = \sum_{i=1}^{n} \sum_{j=1}^{n=k} (X_n . W_{nk}) \tag{1}$$

$$Y_n = \sigma(\sum_{i=1}^{n} \sum_{j=1}^{n=k} (H_{(out(nk))} . \beta_{nk})) \tag{2}$$

Where, X_n : Input data, Y_n : Prediction output, $H_{(out(nk))}$: Hidden layer output, W_{nk} : Input-to-hidden layer weights, β_{nk} : Hidden-to-output layer weights, and σ : Activation function.

By using the hidden layer function, the best set of results can be found in the network. The power of MLP prediction capability comes from the ability to learn from training data and relating the best testing data to the given output data in a hierarchical or multi-layered structure of the network. It uses supervised learning technique called backpropagation for training the network. Due to its popular ability to solve difficult problems, a variety of MLP was created to optimise the result for different types of issue.

- Recurrent Neural Network (RNN)
  RNNs are a type of artificial neural network that was designed to learn patterns in data sequences such as numerical time series data, images, and text. It is a powerful type of neural network that has been used in industries such as sensors, the stock market, and government agencies.
  Fig.6 shows the RNN full network (unfolded) which is the complete sequence of the network. For example, if there is a sequence of three numerical values, the network would unfold into a three-layer neural network that supports a layer for each numerical value.

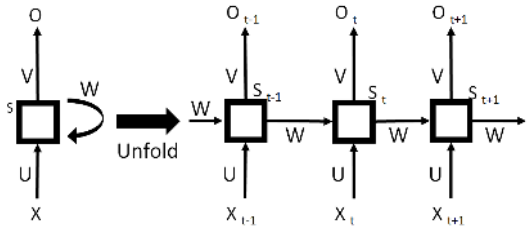

Fig. 6: Recurrent neural network and unfolding sequence diagram

The computational formulas [27] in an RNN happens as follows:

$$s_t = \sigma(s_{t-1}.W + x_t.U + b) \tag{3}$$

$$o_t = s_t.V \tag{4}$$

Where, t : Time step, x_t : Input data, o_t : Predicted output, s_t : Hidden state, U : Input-to-hidden weights, W : Hidden-to-hidden weights, V : Hidden-to-output weights, b : Bias value, and σ : Activation function.
Hidden state s_t is considered the memory of the network; it captures information about the situation in all previous time steps which was the main feature of an RNN. o_t is the output predicted solely based on the current memory at time step t. RNN weights U, V, W are constant throughout the process, unlike traditional neural network where it is different at each layer. This reduces the number of parameters required to be learnt by performing the same task at each time step but with different inputs.

- Long Short Term Memory (LSTM)
One of the appeals of RNNs is the idea that they might be able to connect previous information to the present task. In cases where the gap between the relevant information and the place which is required was small, RNNs is able to learn and utilise the past information [28]. However, if the gap is huge, RNN is unable to link the information for the learning process to kick in.
In order to solve long-term dependency issues, a special kind of RNN called Long Short Term Memory (LSTM) networks were created. It was introduced by Hochreiter & Schmidhuber [29] which was then popularised and refined by many people in various industries as it works extremely well on a variety of problems. Fig. 7 shows how each block of LSTM network interacts with each other.

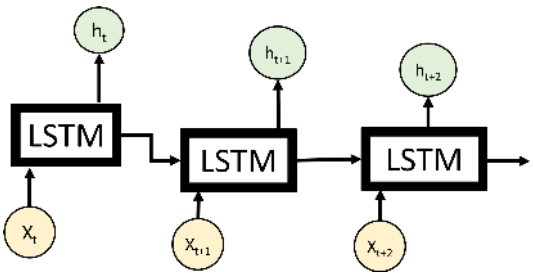

Fig. 7: LSTM network diagram

Fig.8 shows the details of the LSTM block [28]. In Fig. 8, each line carries an entire vector, from the output of one node to the inputs of the others. The grey circles represent pointwise operations, similar to vector addition, while the orange boxes are learned neural network layers. Lines (vector transfer) denote content going to different locations.

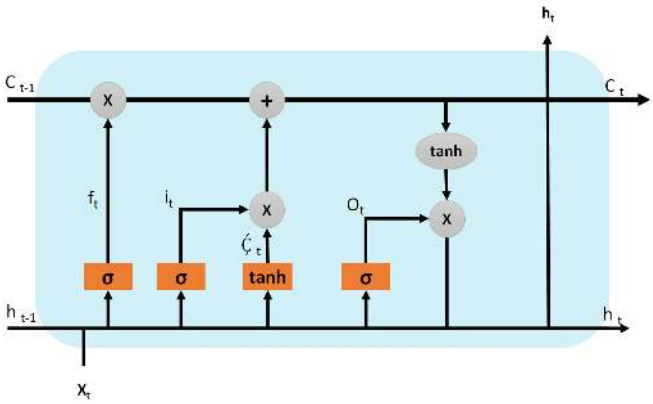

Fig. 8: LSTM block diagram

The computational formulas [30], [31] in an LSTM block are defined as follows:

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{5}$$

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{6}$$

$$\acute{C}_t = tanh(W_c.[h_{t-1}, x_t] + b_c) \tag{7}$$

$$C_t = f_t.C_{t-1} + i_t.\acute{C}_t \tag{8}$$

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{9}$$

$$h_t = o_t.tanh(C_t) \tag{10}$$

Where, t : Time step, x_t : Input value, h_t : Output value, o_t : Output gate, f_t : Forget gate, i_t : Input gate, C_t : Cell state, $\acute{C}_t$: Candidate value, W_o : Output gate weights, W_i : Input gate weights, W_f : Forget gate weights, W_c : Cell state weights, b_o : Output gate bias value, b_i : Input gate bias value, b_f : Forget gate bias value, b_c : Cell state bias value, and σ : Gate state.
There are three gates in the block that manage the block state and output:

- Forget Gate f_t: decides the information to throw in the block.
- Input Gate i_t: decides which input values to update the memory state.
- Output Gate o_t: decides the output depending on the input and memory state.

Each block represents a mini-state machine where gates have weights that are learned during the training procedure

[32]. This allows the creation of large LSTM to address complex sequence problems and achieve optimal results.

- Gated Recurrent Unit (GRU)
  A variation of the LSTM is the Gated Recurrent Unit (GRU) which was introduced by Cho, et al. [30]. This system has a single update gate which combines the input and output gate. It also merges the hidden and cell state which makes a simplified model than a standard LSTM model. Fig. 9 shows the details of the GRU model [28].

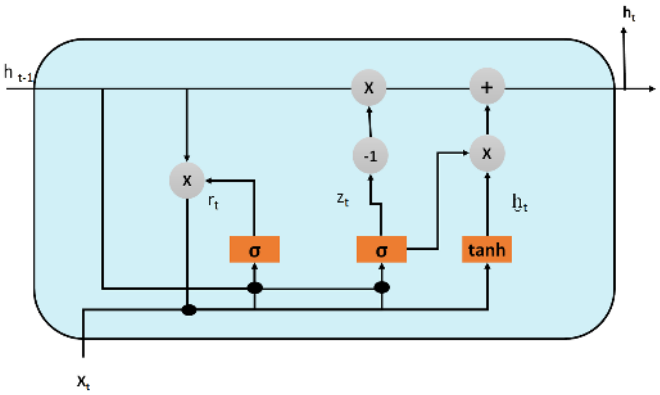

Fig. 9: GRU block diagram

The GRU layer is derived from the LSTM layer which results in similar equations:

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{11}$$

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{12}$$

$$\underset{\sim}{h}_t = tanh(W.[r_t.h_{t-1}, x_t]) \tag{13}$$

$$h_t = (1 - z_t).h_{t-1} + z_t.\underset{\sim}{h}_t \tag{14}$$

Where, t : Time step, x_t : Input value, h_t : Output value, r_t : Reset gate, z_t : Update gate, $\underset{\sim}{h}_t$: Candidate value, W_r : Reset gate weights, W_z : Update gate weights, W : Candidate gate weights, and σ : Gate state.
The reset gate determines the new input and previous memory combination and the update gate determines the amount of previous memory to be kept. The idea of using a gating mechanism is similar to LSTM with an objective to learn long-term dependencies. The key differences are:

  - GRU has two gates while LSTM has three.
  - GRU does not have output gate and internal memory.
  - GRU trains faster due to lesser parameters.

GRU and LSTM models had solved the long term dependencies issues but the trade-off of both system are not fully explored [32].

- State of Prediction Model ($sp_{(n)}$) The State of Prediction Model ($sp_{(n)}$) determines the abnormally for energy theft in stage 1. The following formulas were used for this stage:
  - The number of hidden layer [33] :

$$n_h = \frac{(n_i + n_o)}{2} + \sqrt{n_t} \tag{15}$$

    Where, n_h : Number of the hidden layer, n_i : Number of the input layer, n_0 : Number of the output layer, and n_t : Number of the training sets.
  - The Mean Absolute Percentage Error (MAPE):

$$MAPE_n = \frac{100}{n}\sum_{i=1}^{n}|\frac{A_i - F_i}{A_i}|, where\ A_i \neq 0 \tag{16}$$

    Where, n : Number of data, A_i : Actual output data, and F_i : Forecast output data.
  - The Absolute Percentage Error (APE):

$$APE_n = 100(|\frac{A_n - F_n}{A_n}|), where\ A_n \neq 0 \tag{17}$$

    where APE_n= Absolute Percentage Error for n.
  - The state of prediction:

$$sp_{(n)} = \begin{cases} 0, & \text{if } APE_n \leq MAPE_n \\ 1, & \text{otherwise} \end{cases} \tag{18}$$

    Where, $sp_{(n)}$: State of prediction model decision making condition.

b) Stage 1: Procedures: The following steps are taken for this stage:

- Step 1: Pre-process the data to accumulative data.
- Step 2: Using prediction model to predict the data.
- Step 3: Using Mean Absolute Percentage Error (MAPE) to dictate the best prediction model.
- Step 4: Use the updated MAPE to compare with Absolute Percentage Error (APE) for every hour.
- Step 5: If $sp_{(n)} = 1$ then go to the next stage, otherwise go to the next iteration.

2) Stage 2: Primary Decision Making Model: This stage uses Simple Moving Average (SMA) to determine the energy theft predictions.

a) Stage 2: Algorithms: The following formulas are used for this stage:

- The Simple Moving Average (SMA):

$$SMA_{(n)} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{19}$$

  Where, n: The number of hours for SMA and x: The variable for the hour in the list.
- The Maximum SMA difference algorithm:

$$SMA_{(md)} = \max_{i \in n} f(|SMA_{(i)} - SMA_{(i-1)}|), \quad where\ n \neq 0 \tag{20}$$

  Where, $SMA_{(md)}$: Maximum of the SMA difference between before and after.
- The state of hours:

$$s_{h(n)} = \begin{cases} 0, & \text{if } (SMA_n - SMA_{n-1}) \leq \frac{3}{4}SMA_{(md)}, \\ 1, & \text{otherwise} \end{cases} \tag{21}$$

Where, $s_{h(n)}$: State of hours algorithm decision making condition.

b) Stage 2: Procedures: The following steps are taken for this stage:

- Stage 2.1: Continuous Hour Model:
  - Step 1: Calculate Simple Moving Average (SMA) using 24 hours period.
  - Step 2: Find the difference between the SMA calculation for the last hour and the current hour after 25 hours of measured data.
  - Step 3: Use the Maximum SMA difference algorithm and proceed to the state of hours algorithm.
  - Step 4: If $s_{h(n)} = 1$ then start the Same Day and Hour Model, otherwise go to the next iteration.
- Stage 2.2: Same Day and Hour Model:
  - Step 1: Rearrange the data according to the day and hour.
  - Step 2: Calculate SMA using 4 hours of data from the same day and hour from different dates.
  - Step 3: Find the difference between the SMA calculation for the last point and the current point after 5 points of measured data.
  - Step 4: Use the Maximum SMA difference algorithm and proceed to the state of hours algorithm.
  - Step 5: If $s_{h(n)} = 1$ then go to the next stage, otherwise go to the next iteration.

3) Stage 3: Secondary Decision Making model: This stage uses the user's history to find the occasional maximum power usages.

a) Stage 3: Algorithms: The following formulas are used for this stage:

- The Maximum wattage:

$$P_{(md)} = \max_{i \in n} f(|P_{(i)}|) \tag{22}$$

Where, $P_{(md)}$: The maximum power from the list of measurement.

- The state of energy theft:

$$sets_{(n)} = \begin{cases} 0, & \text{if } \frac{3}{4}P_{(md)} \leq P_n \leq P_{(md)} \\ 1, & \text{otherwise} \end{cases} \tag{23}$$

Where, $sets_{(n)}$: State of energy theft algorithm decision making condition.

b) Stage 3: Procedures: The following steps are taken for this stage:

- Step 1: Find the Maximum watt and proceed to the state of energy theft algorithm.
- Step 2: If $sets_{(n)} = 1$ then possible energy theft, otherwise unexpected high consumption usage from consumers.
- Step 3: Proceed to next iteration.

After all the stages are completed, it will move to the next period and repeat the process from stage 1. However, SETS requires at least 5 weeks of non-malicious data collection at every hour in order for the system to learn from the historical data. This learning will be constantly updated for real-time monitoring and it can increase its accuracy with more data coming in.

IV. Simulation Studies and Results

A. Experiment Setup and Data Collection

The Aeon Labs Z-Wave UK Plug-in Switch plus Power Meter were installed on every available energy consumption devices in the experimental house. Then, the data was collected through a centralised smart device called VeraEdge Home Controller. Fig. 10 shows the demand data collected from the experimental house. The data collected from 04/12/2016 – 02/04/2017 were in kilowatt (kW) and timestamp (DD/MM/YYYY HH:MM).

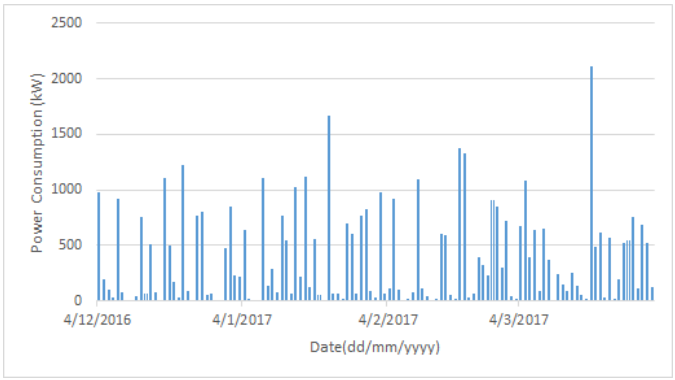

Fig. 10: Plot of experimental house demand data

B. Smart Energy Theft System (SETS) Results

The SETS was tested using simulated energy theft scenarios. The scenario was created by randomly stealing energy on 50 different periods. Fig.11, 12, 13, and 14 show the respective prediction results for MLP, RNN, LSTM, and GRU.

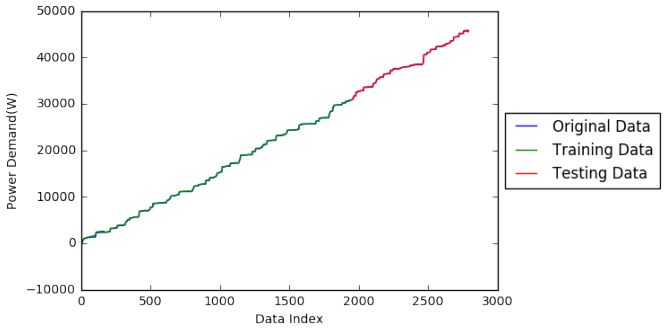

Fig. 11: MLP prediction result

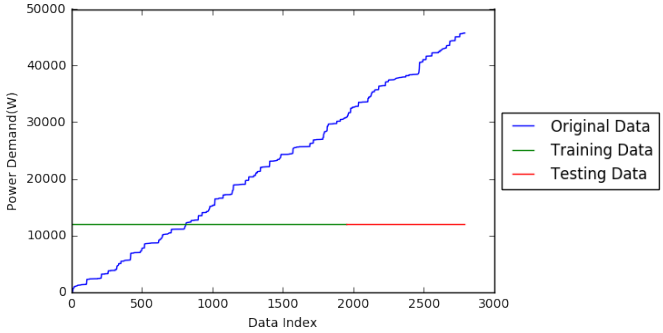

Fig. 12: RNN prediction result

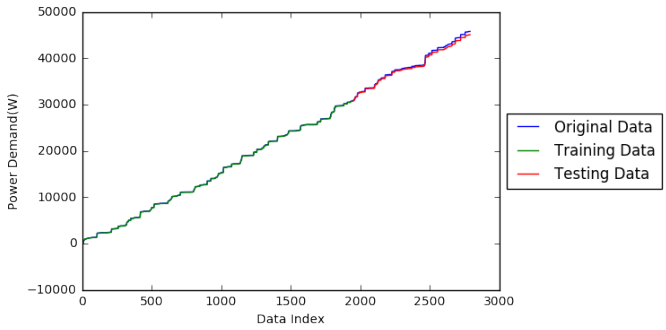

Fig. 13: LSTM prediction result

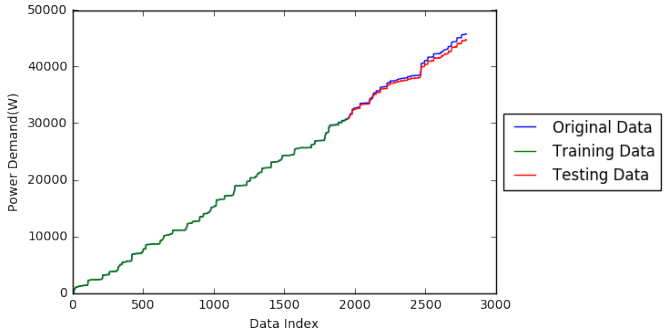

Fig. 14: GRU prediction result

Table I shows the MAPE results for different forecasting systems. The best MAPE result was 0.18% which was considered most suitable method as compared to other methods tested.

TABLE I: SETS: Prediction model MAPE results

Prediction model	MLP	RNN	LSTM	GRU
MAPE(%)-Train	33.99	2353.23	5.48	11.20
MAPE(%)-Test	0.18	68.83	0.81	1.32

Fig.15 shows the stage 2 alert system for Smart Energy Theft System (SETS). These results were obtained after the data processed through stage 2 in SETS.

```
Normal Usage:'5/3/2017 22:00'
Normal Usage:'12/3/2017 22:00'
Possible Sudden High consumption usage:'19/3/2017 22:00'
Normal Usage:'26/3/2017 22:00'
Normal Usage:'18/12/2016 23:00'
Normal Usage:'25/12/2016 23:00'
```

Fig. 15: SETS: Stage 2 alert notifications

In Fig.15, the alert notifications were made after processing through stage 2. It filters the abnormally from stage 1 and proceeded to stage 3 if it is not able to make a decision.

Fig.16 shows the stage 3 final stage alert system for Smart Energy Theft System (SETS). These results were obtained after the data processed through stage 2 and 3 in SETS.

```
Possible Sudden High consumption usage:'4/1/2017 13:00'
Possible Sudden High consumption usage:'1/3/2017 21:00'
Possible Energy Theft:'29/3/2017 21:00'
Possible Sudden High consumption usage:'12/1/2017 13:00'
Possible Sudden High consumption usage:'2/2/2017 15:00'
```

Fig. 16: SETS: Stage 3 alert notifications

In Fig.16, the final stage alert notifications were made from filtering stage 2 and using stage 3 algorithms. This results in 99.96% accuracy of classifications using SETS with all stages implemented.

C. Discussion

Table II shows classification results for different cases with the same energy theft scenario. The cases in Table II were done by randomly stealing the energy of 50 different periods. These conditions were maintained to present a fair environment for the detection capability of Smart Energy Theft System (SETS).

TABLE II: Summary of classification results in different stages

SETS Case Studies	Classification Accuracy (%)
Case 1: Stage 1	56.39
Case 2: Stage 2	99.46
Case 3: Stage 3	0.68
Case 4: Stage 1 & 3	56.87
Case 5: Stage 2 & 3	99.89
Case 6: Stage 1 & 2	99.89
Case 7: All Stages	99.96

Table III shows classification results for different sub-cases with the same energy theft scenario.

TABLE III: Summary of classification results for sub-cases

SETS Sub-Case Studies	Classification Accuracy (%)
Sub-Case 1: Stage 2.1	2.04
Sub-Case 2: Stage 2.2	19.39
Sub-Case 3: Stage 2.1 & 3	99.39
Sub-Case 4: Stage 2.2 & 3	99.32
Sub-Case 5: Stage 1 & 2.1	99.4
Sub-Case 6: Stage 1 & 2.2	99.4

Case 1, 2, and 3 were a single stage detection system. Case 4, 5, and 6 were 2 stages detection systems. Case 7 represents the Smart Energy Theft System (SETS).

Case 1, 2, and 3 achieved classifications accuracy of 56.39%, 99.46%, and 0.68%. Among the single stage detection systems, case 3 had the worst accuracy result while case 2 had the best accuracy results. However for case 2, further findings were found by separating stage 2 into stage 2.1 (Continuous Model) and stage 2.2 (Same Day and Hour Model). Sub-cases 1 and 2 achieved just 2.04% and 19.39% respectively. Case 2 had further demonstrated that by integrating the 2 models, it shows tremendous improvements for detection techniques.

Case 4, 5, and 6 achieved classifications accuracy of 56.87%, 99.89%, and 99.89%. Among the 2 stages detection systems, case 4 had the worst accuracy result while case 5 and 6 had the best accuracy results. 2 stages integration results show improvements compared to single stage detection systems. Case 5 was further analysed in sub-case 3 and 4. Sub-case 3 had a 99.39% accuracy and sub-case 4 achieved 99.32%. Case

6 was also further analysed in sub-case 5 and 6. Sub-cases 5 and 6 had both achieved 99.4%. Case 7 was done using SETS to achieve a classification accuracy of 99.96%.

After reviewing all the cases, it shows significant increment by integrating the different stages in SETS. By using a single detection system, detection accuracy results like Case 1 and 3 would not be efficient enough for energy theft situations. By integrating 2 detection systems, although case 4 was still not efficient but case 5 and 6 had shown considerable improvements on its classification accuracy. Ultimately, this led to an integration of all 3 detection techniques with the best classification accuracy among all cases.

V. Conclusions

In this paper, an innovative Smart Energy Theft System (SETS) is proposed for energy theft detection. A Multi-Model Forecasting System based on the integration of machine learning models such as Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU) was developed as part of SETS. Additionally, a statistical model called Simple Moving Average (SMA) was also further developed into SETS. These algorithms enable SETS to efficiently detect energy theft activities. The evaluation of its system carried out in a Singapore home environment. Stage 1 has an energy theft accuracy result of 56.39%, by adding stage 2 has 99.89% and all 3 stages present the evidence of its energy detection algorithm accuracy of 99.96%. In conclusion, SETS enhances the security of the Internet of Things (IoT) based smart home systems from energy theft and can be further implemented in commercial and industrial sectors.

References

[1] W. Li, T. Logenthiran, V.-T. Phan, and W. L. Woo, "Intelligent multi-agent system for power grid communication," in *Region 10 Conference (TENCON), 2016 IEEE*. IEEE, 2016, pp. 3386–3389.

[2] ——, "Housing development building management system (hdbms) for optimized electricity bills," *Transactions on Environment and Electrical Engineering*, vol. 2, no. 2, pp. 64–71, 2017.

[3] W. Li, T. Logenthiran, W. Woo, V. Phan, and D. Srinivasan, "Implementation of demand side management of a smart home using multi-agent system," in *IEEE World Congress on Computational Intelligence*. IEEE, 2016, pp. 1–8.

[4] C. Yang, J. Yao, W. Lou, and S. Xie, "On demand response management performance optimization for microgrids under imperfect communication constraints," *IEEE Internet of Things Journal*, 2017.

[5] F. L. Quilumba, W.-J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 911–918, 2015.

[6] T.-C. Chiu, Y.-Y. Shih, A.-C. Pang, and C.-W. Pai, "Optimized day-ahead pricing with renewable energy demand-side management for smart grids," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 374–383, 2017.

[7] T. G. Nikolaou, D. S. Kolokotsa, G. S. Stavrakakis, and I. D. Skias, "On the application of clustering techniques for office buildings' energy and thermal comfort classification," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 2196–2210, 2012.

[8] J. Siryani, B. Tanju, and T. J. Eveleigh, "A machine learning decision-support system improves the internet of things smart meter operations," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 1056–1066, 2017.

[9] W. Li, T. Logenthiran, V. T. Phan, and W. L. Woo, "Implemented iot based self-learning home management system (shms) for singapore," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

[10] R. C. Luo, C.-C. Yih, and K. L. Su, "Multisensor fusion and integration: approaches, applications, and future research directions," *IEEE Sensors journal*, vol. 2, no. 2, pp. 107–119, 2002.

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

[12] Y. Liu and S. Hu, "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, 2015.

[13] Q. Hu and F. Li, "Hardware design of smart home energy management system with dynamic price response," *IEEE Transactions on Smart grid*, vol. 4, no. 4, pp. 1878–1887, 2013.

[14] S. K. Viswanath, C. Yuen, W. Tushar, W.-T. Li, C.-K. Wen, K. Hu, C. Chen, and X. Liu, "System design of the internet of things for residential smart grid," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 90–98, 2016.

[15] D. Minoli, K. Sohraby, and B. Occhiogrosso, "Iot considerations, requirements, and architectures for smart buildings–energy optimization and next generation building management systems," *IEEE Internet of Things Journal*, 2017.

[16] W. Ejaz, M. Naeem, A. Shahid, A. Anpalagan, and M. Jo, "Efficient energy management for internet of things in smart cities," *IEEE Communications Magazine*, 2017.

[17] D. Niyato, L. Xiao, and P. Wang, "Machine-to-machine communications for home energy management system in smart grid," *IEEE Communications Magazine*, vol. 49, no. 4, 2011.

[18] S. D. T. Kelly, N. K. Suryadevara, and S. C. Mukhopadhyay, "Towards the implementation of iot for environmental condition monitoring in homes," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3846–3853, 2013.

[19] J. Han, C.-S. Choi, W.-K. Park, I. Lee, and S.-H. Kim, "Smart home energy management system including renewable energy based on zigbee and plc," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 198–202, 2014.

[20] A. Pratt, D. Krishnamurthy, M. Ruth, H. Wu, M. Lunacek, and P. Vaynshenk, "Transactive home energy management systems: The impact of their proliferation on the electric grid," *IEEE Electrification Magazine*, vol. 4, no. 4, pp. 8–14, 2016.

[21] W. Li, T. Logenthiran, and W. Woo, "Intelligent multi-agent system for smart home energy management," in *Innovative Smart Grid Technologies-Asia (ISGT ASIA), 2015 IEEE*. IEEE, 2015, pp. 1–6.

[22] northeast group llc, "World loses $89.3 billion to electricity theft annually, $58.7 billion in emerging markets," 2015. [Online]. Available: https://www.prnewswire.com/news-releases/world-loses-893-billion-to-electricity-theft-annually-587-billion-in-emerging-markets-300006515.html

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

[24] Y. Liu, Y. Zhou, and S. Hu, "Combating coordinated pricing cyberattack and energy theft in smart home cyber-physical systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017.

[25] M. Riedmiller and A. M. Lernen, "Multi layer perceptrons," 2014.

[26] T. Teo, T. Logenthiran, and W. Woo, "Forecasting of photovoltaic power using extreme learning machine," in *Smart Grid Technologies-Asia (ISGT ASIA), 2015 IEEE Innovative*. IEEE, 2015, pp. 1–6.

[27] J. L. Elman, "Distributed representations, simple recurrent networks, and grammatical structure," *Machine learning*, vol. 7, no. 2-3, pp. 195–225, 1991.

[28] C. Olah, "Understanding lstm networks," 2015. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[30] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[31] W. Zaremba, "An empirical exploration of recurrent network architectures," 2015.

[32] D. Britz, "Recurrent neural network tutorial, part 4 implementing a gru/lstm rnn with python and theano," 2015. [Online]. Available: http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-grulstm-rnn-with-python-and-theano/

[33] A. Zi Yang Adrian, W. Wai Lok, and M. Ehsan, "Artificial neural network based prediction of energy generation from thermoelectric generator with environmental parameters," *Journal of Clean Energy Technologies*, 2017.