NUEJ
Naresuan University
Engineering Journal

Electricity Theft Detection in Electrical Distribution System Using Long Short-Term Memory

Chintana Xayalath[1], Sutthichai Premrudeepreechacharn [2]* and Kanchit Ngamsanroaj[3]

[1] Department of Mechanical, Faculty of Engineering, Chiang Mai University, Chiang Mai 50200, Thailand
[2] Department of Electrical Engineering, Faculty of Engineering, Chiangmai University, Chiang Mai 50200, Thailand
[3] Electricity Generating Authority of Thailand (EGAT), Nonthaburi, Thailand

* Corresponding author e-mail: suttic@eng.cmu.ac.th

(Received: 26 November 2021, Revised: 25 August 2022, Accepted: 05 September 2022)

Abstract

This paper presents the method of detection of power theft in the electrical distribution system by using the real data from industry consumers that use meter AMR. The real data were feature extracted by labeling event types by the number demonstrated to normal, voltage theft, and current theft. Then, the data were fed to Long Short-Term Memory (LSTM) for the created model by training and testing. The accuracy results were shown the model can be classified accurate archive to 99%.

Keywords: Non-technical loss, Power theft, Automatic Meter Reading, Long Short-Term Memory classification.

1. INTRODUCTION

Electricite Du Laos (EDL) is a state own enterprise responsible for distribution and transmission systems in the country. The municipal is one of the branches in Vientiane and has a large distribution system with a total of 164,126 customers. The Automatic Meter Reading (AMR) is installed in 135 customers as reported by (EDL, 2020). This area is a centralized country that has power use variety so that power consumption has large.

Along with power distribution, power losses have appeared in the power system. It is contained technical and non-technical losses. Which indicated to health index of the electrical distribution system and related revenue of the Organization.

Table 1 Summarize power consumption and power loss in Municipal Vientiane

No	Description	Energy (kWh)	Percentage
1	Energy demand sent out substation	1,130,046,682	100%
2	Energy from the billing system	981,961,993	86.9%
3	Technical loss	36,781,606	3.25%
4	Non- Technical loss	111,304,083	9.85%

As shown in Table 1, has demonstrated the ratio of non-technical losses (NTL) is coverage of the power sent from the substation more than technical losses. it is happened from power theft, meter fault, human error, unregistered meter, meter tampering, etc., as reported in the (CIRED, 2017) report.

The major problem of EDL is power theft which occurred by a fraudulent customer by going to adjust the meter. It made the power consumption in the billed less than actual. The most issue has appeared in commercial customer as shown in Table 2.

Table 2 Summarize non-technical loss in Municipal Vientiane

No.	Description	Non–technical losses (kWh)	Percentage
1	Household	988,215	0.88%
2	Commercial	109,625,838	98.49%
3	Other	690,030	0.61%

In previous years ago many methods related to power theft have been conducted. The traditional method of detecting power theft inspects the abnormal power consumption in the electrical bill and upgraded the measuring equipment in power distribution, such as the smart meter. Monitoring the power usage of customers in the smart grid, when using the smart meter for a long time have very large data in the data analytics (Bula et al., 2016). Due to the large amount of historical data being stored in the database. The data on power consumption is applied to the algorithms of machine learning. To detect NTL, Principal Component Analysis (PCA) can help to improve support vector machine performance and predictive accuracy (Toma et al., 2019). (Glauner et al., 2016) used the neighborhoods feature and master data of customers from successive consumption to apply logistic regression, k-nearest neighbors, linear support vector machine, and random forest to apply with logistic regression, k-nearest neighbors, linear support vector machine, and random forest. The result was compared to the accurate detection of NTL. (Micheli et al., 2019) investigated NTL detection using a variant of Peer-to-Peer (P2P) computing to detect fraud in the case of unreliable smart meters. This research applied data collected from smart meters and data in the same neighborhood area for verification by NTL. This data has been implemented in the application of multi-linear regression. (Nabil et al., 2018) presented a Gradient Boosting Theft Detector (GBTD) based on several Gradient Boosting classifiers to improve the detection rate and False Positive Rate (FPR) of Gradient Boosting classifiers (GBCs). This research focuses on feature engineering-based improvements to the classifiers' ability to improve. Furthermore, a stochastic feature was used to generate the standard deviation, mean, minimum, and maximum values of daily electricity usage. There are 6 theft cases sampled from real-world theft patterns and applied with the dataset for validation of the proposed algorithm. (Buzau et al., 2019) used a hybrid deep neural network by combining the architecture of a long-short term memory with a multi-layer perceptron. It used feature detection anomalies and fraud in smart meters to train and test the algorithm. The performance classifiers of hybrid neural networks significantly outperform the state-of-the-art. (Long et al., 2020) used the method of data-driven combined algorithm to systematically identify anomalies of power loss in the distribution network. It consisted of abnormal feeder, time, and position detection.

In addition, the research paper by (Ghori et al., 2019) used real datasets from the company that supplies power in Pakistan. The study evaluated 15 algorithm classifiers of ML. It included the recent algorithm developed to compare the performance classifier in non-technical loss detection and the author, (Veerasamy et al., 2021) used a model recurrent neural network based on LSTM to classifier high impedance fault detection in solar PV.

This research aims to create a model to detect power theft in power distribution systems. The real data of case power theft has features extracted and applied with LSTM algorithms to establish a model.

This paper is organized as follows. The introduction of non-technical losses in the electrical distribution system was described in Section 1. Then, the power theft in AMR is described in Section 2. After that, the methodology for the detection of power theft is explained in Section 3. The simulation results are discussed in Section 4. Finally, the conclusion is described in Section

2. POWER THEFT IN AUTOMATIC METER READING (AMR)

2.1 Structure AMR system

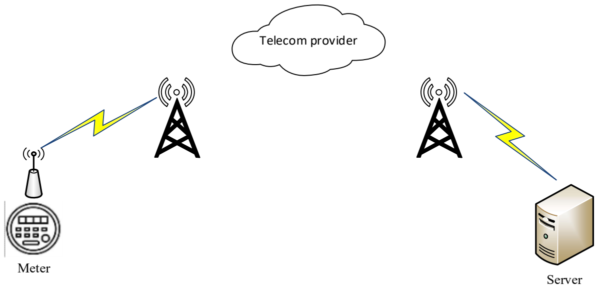

Figure 1 Automatic Meter Reading system.

The Automatic Meter Reading (AMR) is a similar meter measurement to the power parameter. It can be read, record values, and send the data to a server. The values include voltage, current, power consumption, energy import/export, etc. and the communication between meter and server uses the GPRS module. The data granularity is recorded every 15 minutes and the architecture system is shown in Figure 1.

2.2 Method AMR stealing electricity

The commercial customer is the customer category that uses power at high load capacity. While the customers consume large amounts of electricity, the electricity meter cannot support the power directly. It should have the equipment to scale the voltage and current appropriate with meter measurement range such as potential transformer (VT) and current transformer (CT) and has the meter diagram connection as shown in Figures 2 and 3.

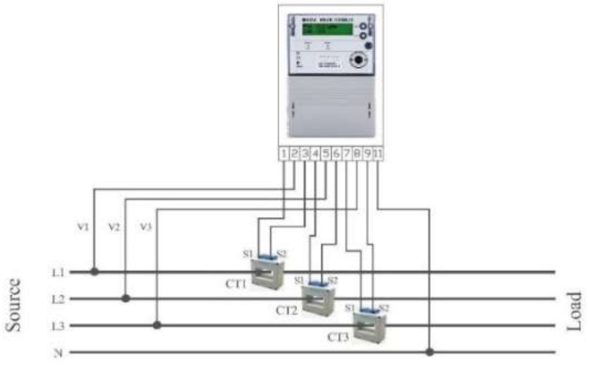

Figure 2 Single line diagram of meter 0.4 kV in EDL's standard.

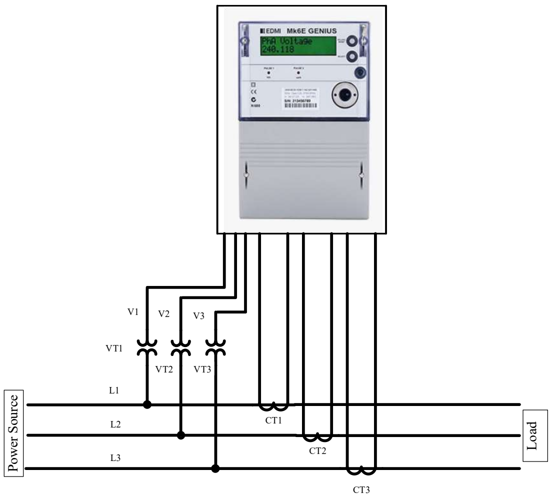

Figure 3 Single line diagram of meter 22kV in EDL's standard.

As Figures 2 and 3 are demonstrated the meter is connected with the power supplied. It has consisted of 2 different types connecting in the power distribution system. It also uses these types of meters in power distribution at low voltage and medium voltage. For the low voltage level meter, it should be used the current transformer (CT) to scale the range of current for measuring the electricity as shown in Figure 2. In the same way, in Figure 3 they also apply CT and installed voltage transformer (VT) to decrease voltage level from medium voltage level to applied at the range of meter requirement.

The power thefts on AMR all most is occurred from disturbing the measuring wire of voltage and current connect to the meter on site by cutting the measurement cable voltage and current 1 or 2 phases, short circuit the secondary winding of CT each other to make the measuring value is less than actual power utility.

3. METHODOLOGY

3.1 Long short-term memory

Long Short-Term Memory (LSTM) is a specialized Recurrent Neural Network. Which has architecture shown in Figure 4. It has the capability of learning long-term dependencies. Traditional RNNs have the problem of vanishing and exploding gradients. In vanishing gradients, weights in the early layers of the network get updated with very low values and, hence, train very slowly. This is because weight updates are proportional to the gradient of the error function, and in the initial layers, during backpropagation, gradients multiply with small activation values repeatedly, which further leads to even smaller values. Similarly, in exploding gradients, the gradients become too large and update the weight parameters by a large value and hence become harder to converge to an optimum value.

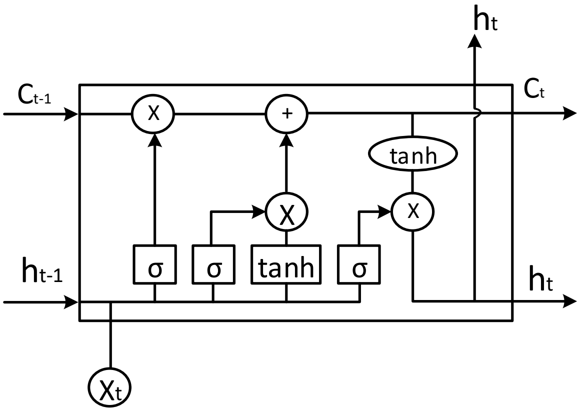

Figure 4 Long Short-Term Memory Architecture

The LSTM has three gates: the input gate i_t, the forget gate f_t and the output gate O_t. The forget gate takes the previously hidden state information (h_{t-1}) and current input x_t through a pointwise multiplication operation and decides what is memorized or forgotten the information from the cell state. This gate uses the sigmoid activation function to approve the predicts an output of either 0 or 1. A value of 1 shows that the relevant information should be kept in the cell state, while 0 represents irrelevant information, which is discarded from the cell state. The forget gate, input gate, and output gate are described in Equations (1) – (5).

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

where W_f represents the weight and b_f is the bias of the forget gate f_t. The σ is applied as the activation function on the forget gate.

The input gate decides what information is going to be stored in the cell state information C_t^i of input gate. It takes the input x_t and previous hidden state h_{t-1} and applies and tanh activation functions through a pointwise multiplication operation as follows:

$$i_t = \sigma\,(W_i[\,h_{t-1}\,,\,x_t\,]+ b_i) \tag{2}$$
$$C_t^i = \acute{C}_t * i_t \tag{3}$$
$$\acute{C}_t = \tanh\,(W_C[\,h_{t-1}\,,\,x_t\,]+ b_C) \tag{4}$$

where W_i and W_C represent the weights of the input gate i_t and new cell state, respectively. The b_i and b_C are the biases of the network, and $\acute{C}_t$ is the previously hidden cell state information. To update the information on the current cell state C_t, Equations (1) multiply by the previous cell state C_{t-1} are summed with The Equations (2) multiplied by Equations (3):

$$C_t = f_t * C_{t-1} + i_t * \acute{C}_t \tag{5}$$

Finally, in Equation (5), the output gate is determined. The output gate takes the current input x_t and previously hidden state h_{t-1} with the implication of the activation function σ. The b_o is added as a bias to the output network.

$$O_t=\sigma\,(W_o[\,h_{t-1}\,,x_t\,]+b_o\,) \qquad (6)$$

The updated output gate O_t and the information from the cell state C_t are used to perform the pointwise multiplication operation to get the next hidden state h_t, given by Equation (7):

$$h_t= O_t * \tanh\,(C_t) \qquad (7)$$

The input of model LSTM can be applied with the data of research x_t is the input at time step data and a vector of 6 points that corresponds to the value of voltage and current and O_t is the output of the LSTM classifier to event-type (normal/theft by voltage/theft by current).

3.2 Data Collection

The simulation of this study use data from the meter AMR low voltage shown in Figure 2. which load profile dataset has been collected from the AMR database system at EDL. The data record during (01/03/2014 - 30/12/2019). it had been recorded every 15 minutes. The data consist of the value of voltage phase 1, 2, 3, current phase 1, 2, and 3 to show attribute behavior power usage of customers and have 28,382,400 data set.

3.3 Feature Extraction of Power Theft

The study used data from Vientiane municipal customers to create a feature extract by selecting Voltage L1(V_1), Voltage L2 (V_2), Voltage L3 (V_3), and Current L1 (I_1), Current L2 (I_2), and Current L3 (I_3) to demonstrate a characteristic pattern to describe power theft behaviour and the type of power theft below:

3.3.1 Theft Detection from Voltage Measurement (V_{theft})

This is present V_{theft} detection, the research applies the parameter of voltage recorded the database of meter AMR. This problem rarely appears in the pattern voltage value such as meter failure, voltage drop, external disturbance, power outage, etc. So, this study implements the voltage parameter to detect the abnormality caused by the fraud meter. according to Figure 5, there are 6 parameters including 3-phase voltage and 3-phase currents. From observation the pattern characteristics of the data which occur 2 periods time abnormal from 20 - 35 and 44 - 55, it causes of power theft on voltage phase 1.

As more as Vtheft 1 phase is directly impacted to NTL, Moreover the stealing power can be done more than 1 phase. So, Figure 6 demonstrates the feature of Vtheft 2 phases. It is affected by the power consumption record of the meter. This leads to an increase in power losses of NTL.

3.3.2 Theft Detection from Current Measurement (I_{theft})

The current is a variable to illustrate the attribute of load. It has the feature very complex for identifiers of the event that occurred in the data record. That the problem uses the current 3 phases is based on the detection of power theft. In addition, the study uses the pattern of voltage to help approved the event. Both parameters are the relation shown in Figures 7 and 8.

As Figure 7 can be found that the profile of voltage and current in periods time 1 – 16 demonstrates that voltage 3 phases are normal and current 3 phases have the value close up zero. Enough periods time 21 the has power use the current 3 phases gradually change go along with behaviors of load. But the pattern of phase 3 is more different between the 2 phases in periods time 17 – 21. that may illustrate the abnormal load.

In addition, Itheft1 phase is directly impacted by NTL, Furthermore, the stealing power can be done in more than 1 phase. So, Figure 8 demonstrates the feature of Itheft 2 phases. It is affected by the power consumption record of the meter. This leads to an increase in power losses of NTL.

From detection voltage theft and current theft to observations. The data of parameters is very difficult if using human inspection, the abnormal of meter as well as power theft. It is greater than when we have the tool to help inspection by the application of Artificial Intelligence. which uses data from above to help classifier issue that problem for more efficiency in detection.

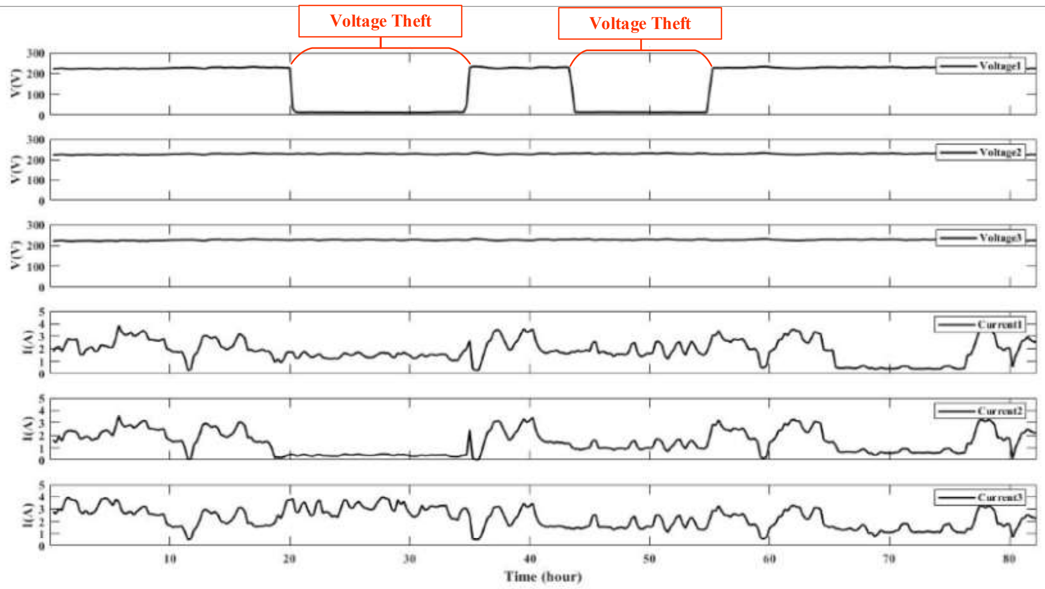

Figure 5 Voltage and current characteristics of power theft by voltage phase 1

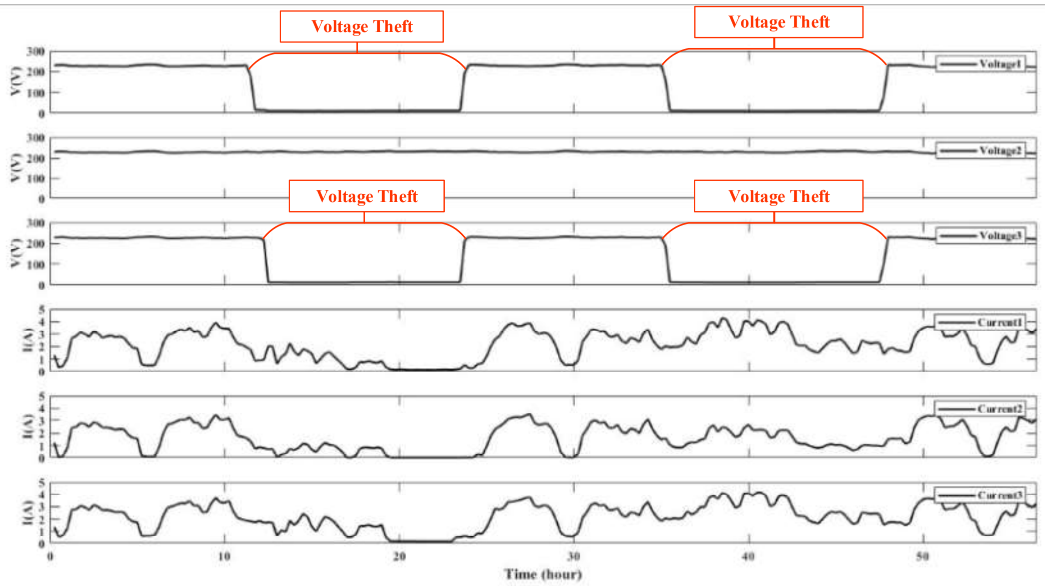

Figure 6 Voltage and current characteristics of power theft voltage phase 1 and 3

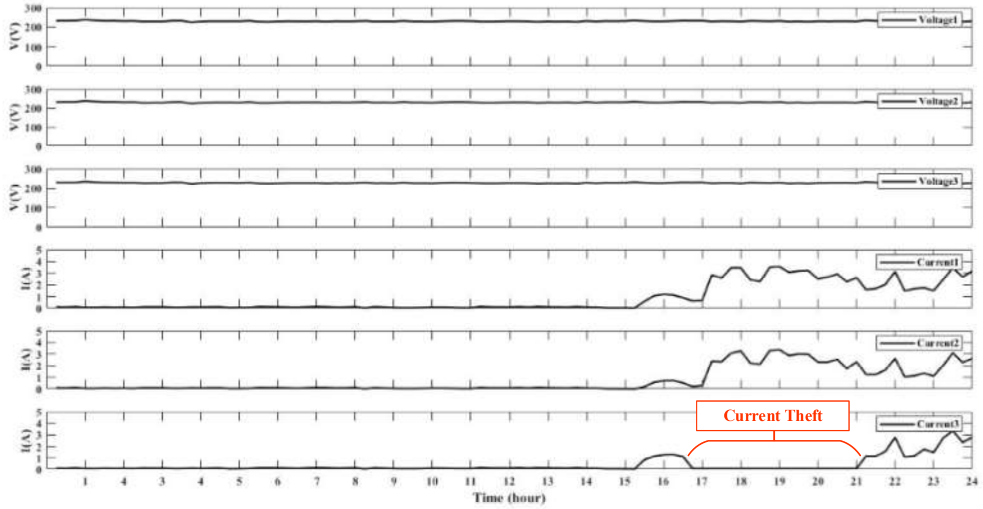

Figure 7 Voltage and current characteristics of power theft from current phase 3

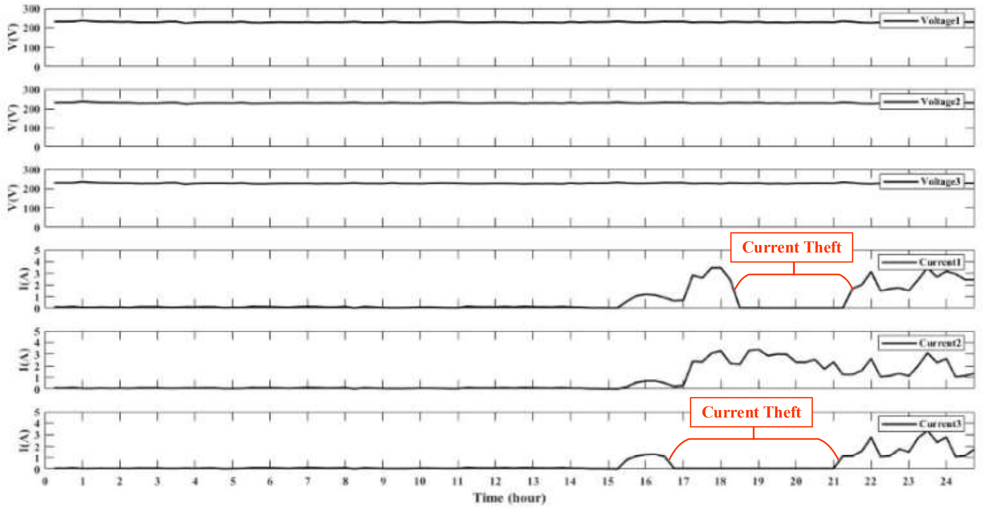

Figure 8 Voltage and current characteristics of power theft from current phase 1 and 3

3.4 Data Preprocessing

This study is using the real data of power theft from the AMR to feature extract. The dataset has included Voltage L1 (V1), Voltage L2 (V2), Voltage L3 (V3), Current L1 (I1), Current L2 (I2), and Current L3 (I3) to define the event type each time step of data. The event typically consists of normal, Vtheft, and Itheft, represented by numbers 1, 2, and 3, respectively. as shown in Table 3. The total dataset feature extracted contained the attribute of voltage and current to demonstrate normal, Vtheft, and Itheft in the numbers 28.237.278, 9,277.605, 52.791.

3.5 Experimental

The detection method of power theft is manipulating the feature extracted data to be applied with an LSTM algorithm which has a process as shown in Figure 9.

In model development to detect power theft. The study is using the feature extracted data in Table 3 to apply the LSTM algorithm. by defining the parameters of V1, V2, V3, I1, I2, I3 as input and event type output of the model. Then, the data has divided between 80% and 20% following the number of each class for the train and test model. The data were fed to LSTM and computed to the accuracy when receipts high accurate. the model will be utilized the next time.

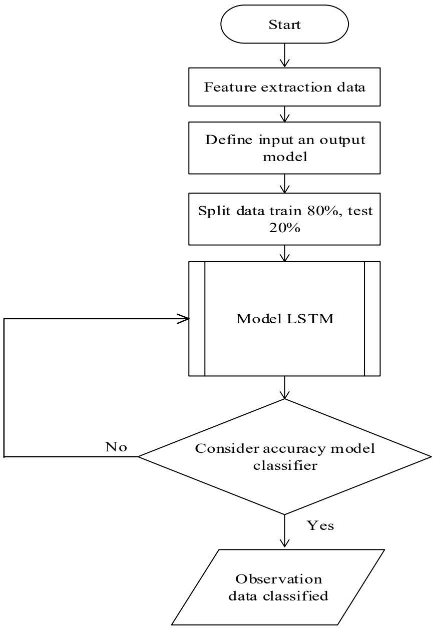
Start
Feature extraction data
Define input an output model
Split data train 80%, test 20%
Model LSTM
No
Consider accuracy model classifier
Yes
Observation data classified

Figure 9 Flowchart model of LSTM classifier

Table 3 Data feature extracted

V1	V2	V3	I1	I2	I3	Event type
9	224.2	11	1.468	0.95	1.497	2
10.6	224.2	11.9	1.665	0.946	1.663	2
12	224.2	12.7	1.31	0.568	1.24	2
11.9	222.8	12.6	1.388	0.868	1.388	2
11.7	220.2	12.4	1.641	1.129	1.467	2
208.1	222.3	206.1	1.042	0.572	1.031	1
223.2	223.6	222.5	0.764	0.315	0.657	1
223.6	224	223	0.579	0.168	0.557	1
224.6	224.5	223.9	0.453	0.026	0.293	1
230.7	230.6	229	0.09	0.079	0.089	1
228.3	227.4	226.6	0.086	0.079	0.086	1
230.4	229.5	228.7	0.042	0.033	0.05	1
230	229.3	228.4	0.038	0.024	0.045	1
230.7	229.8	229.1	0.035	0.021	0.041	1
233.6	232.8	232	0.025	0.007	0.03	1
230.6	229.3	228.8	0.587	0.19	0.827	1
228.7	227.7	227.2	1.057	0.569	1.14	1
229.1	228.2	227.5	1.185	0.705	1.27	1
230.7	229.7	229	1.121	0.715	1.293	1
232.8	231.8	230.8	0.902	0.523	1.107	1
232	231.2	230.2	0.631	0.211	0.211	1
232.4	231.1	230.6	0.684	0.299	0.670	1
228.2	227	226.3	2.839	2.393	0.067	3
229.8	228.6	227.9	2.588	2.343	0.067	3
229.5	228.3	227.8	3.457	3.084	0.067	3
227.8	226.5	225.7	3.457	3.277	0.067	3
231.4	230.4	229.6	2.425	2.218	0.067	3
230.5	229.9	228.5	2.27	2.131	0.067	3
229.4	228.7	227.6	3.499	3.306	0.067	3
231.2	230.6	229.3	3.563	3.386	0.067	3

4. RESULT

This section is describing the simulation results of model power theft detection in a power distribution system on AMR. The study used real data to extract features by labeling event data as normal, V_{theft} and I_{theft}, and then applied an LSTM classified model for training and testing the event. In this work, we used data to train 80% of the classifier algorithm and 20% to test it. The results of model classification are shown in the pattern of the confusion matrix and accuracy.

Table 4 Experimental results of classifier model by using training data

No.	Type of Data set	Data set	Detection Result			Detection Accuracy
			Normal	V_{theft}	I_{theft}	
1	Normal	**15,237,942**	**15,237,942**	0	0	**100%**
2	V_{theft}	**7,422,565**	0	**7,422,565**	0	
3	I_{theft}	**43,141**	0	0	**43,141**	

As Table 4 to illustrated that the number actual of each class can be learned the feature of event type accurately. It demonstrated the feature unconfigure. the accuracy of the model is achieved at 100%. That shows a model has high accuracy.

Table 5 Experimental results of classifier model by using test data

No.	Type of Data set	Data set	Detection Result			Detection Accuracy
			Normal	V_{theft}	I_{theft}	
1	Normal	**3,807,480**	**3,807,480**	0	0	**99.95%**
2	V_{theft}	**1,854,500**	**3,384**	**1,851,116**	0	
3	I_{theft}	**9,650**	0	0	**9,650**	

The result in Table 5 demonstrates that is some missed prediction from the actual power theft by voltage prediction to a normal 3,384 data set. Due to this, the attribute data between both is similar. When compare the total data is less. However, the overall accuracy of detection of the event is achieved at 99.95%. Which the accuracy detection is higher than previous research on these issues.

5. CONCLUSION

The paper presented the detection method for electricity theft in the power distribution system by using the data feature extracted from industry customers to be applied with the LSTM algorithm. The results illustrate that the model can classify the event type to appear in the data accurately. The accuracy of a model is achieved at 100% of training and testing, which demonstrates the model can distinguish between the pattern of power theft and normal activity accurately. However, the model also helps reduce the meter inspection onsite and reduce power losses for a long time.

This research can be a guideline for applying the model to detect the problem the same as this paper in the existing system. Before applying this model to detect and achieve higher efficiency the data must relate to a large dataset. The preparation data for the train and test model should be feature extracted and accurate. In divided data following each class to avoid model overfitting and the selection activation function of the model should be appropriated with the number of classes output.

6. References

Bula, I., Hoxha, V., Shala, M., & Hajrizi, E. (2016). Minimizing non-technical losses with point-to-point measurement of voltage drop between “SMART”meters. *IFAC PapersOnLine*, *49*(29),206-211.

Buzau, M. M., Tejedor-Aguilera, J., Cruz-Romero, P., & Gómez-Expósito, A. (2019). Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Transactions on Power Systems*, *35*(2), 1254-1263.

EDL (2020). Annual Report. EDL Vientiane.

Ghori, K. M., Abbasi, R. A., Awais, M., Imran, M., Ullah, A., & Szathmary, L. (2019). Performance analysis of different types of machine learning classifiers for non-technical loss detection. *IEEE Access*, *8*, 16033-16048.

Glauner, P., Meira, J. A., Dolberg, L., State, R., Bettinger, F., & Rangoni, Y. (2016, December). Neighborhood features help detecting non-technical losses in big data sets. In *Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies* (pp. 253-261).

Long, H., Chen, C., Gu, W., Xie, J., Wang, Z., & Li, G. (2020). A Data-Driven Combined Algorithm for Abnormal Power Loss Detection in the Distribution Network. *IEEE Access*, *8*, 24675-24686.

Micheli, G., Soda, E., Vespucci, M. T., Gobbi, M., & Bertani, A. (2019). Big data analytics: an aid to detection of non-technical losses in power utilities. *Computational Management Science*, *16*(1), 329-343.

Nabil, M., Ismail, M., Mahmoud, M., Shahin, M., Qaraqe, K., & Serpedin, E. (2018, August). Deep recurrent electricity theft detection in AMI networks with random tuning of hyper-parameters. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 740-745). IEEE.

Toma, R. N., Hasan, M. N., Nahid, A. A., & Li, B. (2019, May). Electricity theft detection to reduce non-technical loss using support vector machine in smart grid. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (pp. 1-6). IEEE.

Veerasamy, V., Wahab, N. I. A., Othman, M. L., Padmanaban, S., Sekar, K., Ramachandran, R., ... & Islam, M. Z. (2021). LSTM recurrent neural network classifier for high impedance fault detection in solar PV integrated power system. *IEEE Access*, *9*, 32672-32687.

Working Group on Losses Reduction CIRED WG CC-2015-2 (2017). Reduction of Technical and Non-Technical Losses in Distribution Networks. CIRED http://www.cired.net.