ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.3

An Artificial Intelligent Algorithm for Electricity Theft Detection in AMI

R.Sowndarya[1], Dr.P.Latha M.E, Ph.D[2]
PG Scholar[1], Assistant Professor, Head of the Department[2]
Department of Computer Science and Engineering
Government College of Engineering, Tirunelveli, India

Abstract:
Electricity theft is a major challenging problem faced by the utilities. Apart from technical loss (TL), most of the developing countries are facing huge financial losses because of Non-technical loss (NTL). With the introduction of smart meter, the frequency of reporting energy consumption data to the utility company has been increased more. Thus it paves the way for advanced data analysis. Among the two smart meters, one is installed in distribution transformer and the other in each and every house. So we can detect the NTL in the particular area by taking the difference between these two meter readings. But to ping-point the particular fraudulent in the particular suspicious area is a tedious process in the traditional approach. In this paper, we present a novel framework called ETD (Electricity Theft Detection), which comprises of an intelligent algorithm such as ETD and k-NN classifier to detect fraudulent consumer from the normal consumer based upon the consumer's consumption pattern. Since k-NN employs retraining capability, it makes the framework robust against non-malicious changes in the usage pattern. We use a real time smart energy dataset from the Irish Smart Energy Trial to validate our approach.

Keywords: Advanced Metering Infrastructure (AMI), Electricity theft detector (ETD), Energy theft, k-nearest neighbors (k-NN).

I. INTRODUCTION

A MI (Advanced Metering Infrastructure) incorporates the modernization of metering system with the replacement of mechanical meters by smart meters. Smart meters are becoming famous among the electricity consumer that provides two-way communication between electrical utilities and the consumers thus it eliminates the manual intervention to read the meter reading, and providing various advantages over the traditional meter such as, the ability to monitor the consumption detail of every consumer, billing them etc. Apart from its advantages behavior, it faces some problems in the recent years. Electricity theft has been increasing gradually in various countries. Smart meters may not be tamper able, but electricity theft could still be possible by bypassing the smart meter. It has been addressed by checking the tamper-evident seals using balance meter by field personnel. Therefore, an efficient energy theft detection system against AMI is urgently required. Smart meter theft has been discussed in detail [1]. Preventing the smart meter theft from various vulnerabilities is the most welcomed in the day to day world [2],[3]. There were several approaches attempted for AMI to stand against Non-technical loss (NTL) such as game-theory based methods [4]-[5], classification based approach[6]-[12]. Electricity theft is the crime practice done by the fraudulent to evade the payment and it is punishable by fines. According to the annual Emerging Markets Smart Grid study by the Northeast Group, LLC, it was estimated that around $89.3 billion losses annually to electricity theft. It includes tapping energy directly from the distribution feeder, tampering with the meter to evade the payment [13], [14]. Improper and illegal calibration of smart meter during their design [15] can cause NTL. Nizar et al [17] employ a data mining approach to classify the consumption patterns. By comparing the results to the user database, the algorithm yields a list of users who might be stealing energy. Investigations are undertaken by the utility companies to access the impact of technical losses and overall performance [18]-[21]. Depuru et al [22] proposes another data mining approach. Nagi et al [23] proposes a genetic algorithm and SVM instead of data mining. But unfortunately the above three approaches fails to sort out the fraudulent with absolute certainty. Li et al design a protocol [24] which works similarly to those works like [25]. The most popular method which is prevailing upto date to reduce electricity theft is by using smart meter and intelligent algorithm that makes more difficult for fraudulent activities. The upcoming section of this paper is structured as follows: section II reviews the background details of electricity theft detection system studies to this work. Section III presents the overall methodology. Finally, the section IV and V analyzes the result and draws a conclusion.

II. RELATED WORK

In this section, we review the existing approach prevailing in the list which uses consumption detail of every consumer to find the fraudulent consumer. In [26], abnormal behavior of the consumers was detected by using historical consumption data, SVM along with a data mining method. But this approach requires a daily consumption data of every consumer over a period for two years and it is capable to detect theft only when it has abrupt changes in load profile. Traditionally, the number of theft samples from the consumers is much less than the number of normal samples, it might lead to ignore the minority class leaving the algorithm worthless in some scenario. In [27], the class imbalance problem was addressed by combining the one-

class SVM, optimum path forest and C4.5 decision tree method. The problem with this approach is that it incorporates a high-computational load. We address this by creating a theft samples synthetically from the normal samples. Both SVM and a rule engine approach were proposed in [28] to improve the classification accuracy and it reduces the detection time by parallelizing the algorithm. But still this approach faces some difficulties as in class-based methods. Multi-sensor approach was incorporated in [29] to detect the energy theft. This approach requires high-sampling rate and it reveals the information about the consumer. In this paper, we use only the consumption pattern to detect the energy theft and it promises the consumer's privacy. Salinas et al [30] incorporates a peer-to-peer approach and it addresses the problem of [29]. This approach acts as a privacy preserving scheme and the shortcoming is that it is only effective in theft detection with constant reduction rate. However, there are several range of theft it could be able to detect only small range. Our approach is capable of detecting more diverse attack types. Here, we short list areas with high-probability of theft with the help of transformer meter and it overcomes the limitations of classification-based techniques.

III. METHODOLOGY

A.TRAINING OF k-NN CLASSIFIER

The training samples in a multidimensional feature space are vectors, each with a class label. The training phase of the algorithm consists of storing the class labels and feature vectors of the training samples. The Fig.I.shows the sample Feature Space that clearly separating two classes with the help of Decision Surface and Voronoi spaces. In the classification phase, K is a user-defined constant, and an unlabeled vector or test vector. The test vector is classified under the label which is most frequently occurring among the K training samples nearest to that query (given) point. **Euclidean distance** plays a vital role in calculating the distance metric for continuous variables. Whereas, hamming distance can be used for discrete variables, such as for text classification.

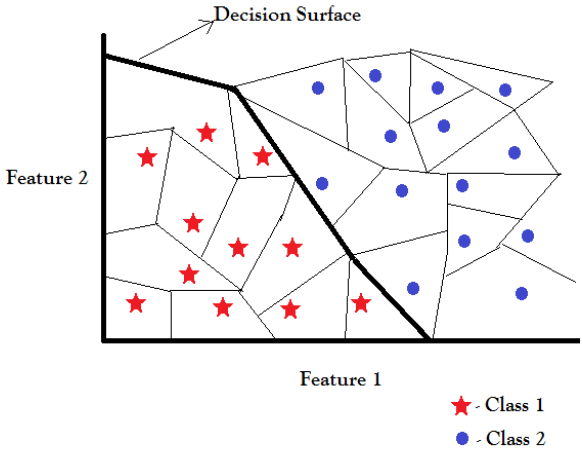


Figure.1. Sample Feature Space

K - VALUE SELECTION

The best choice of K-value selection depends upon the data. Generally, the effect of noise can be reduced greatly in the classification with the larger values of K, but it fails to make clear boundaries between classes. The case in which the data sample is predicted to be grouped under the class of the nearest training sample is called the nearest neighbor algorithm where the K value will be 1.Whenever the accuracy of the k-NN algorithm degraded by the presence of more number of nearest points, the classification process might be turned into tedious. This scenario is tackled by increasing the k value beyond the level. When the K value is 3, the classifier attempts to choose top three nearest neighbors and deciding to which majority class the data point belongs to.

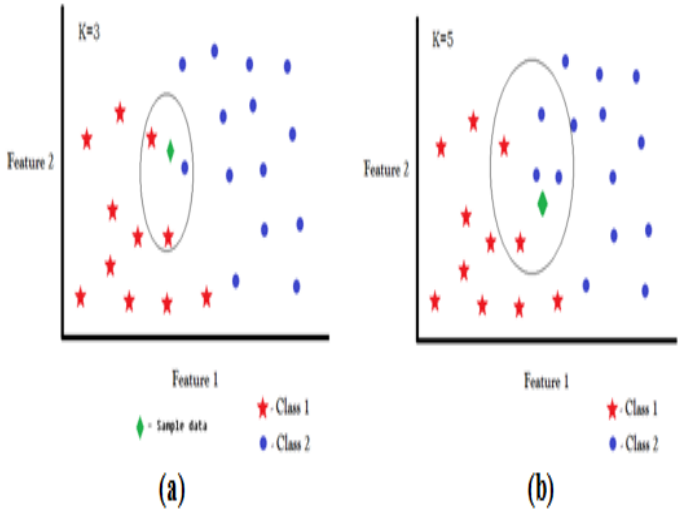


Figure.2. Boundary Selections In Feature Space: (A) For K=3 (B) For K=5

The Fig.II.(a) & (b) shows the boundary selection for two cases K=3 & K=5 respectively. In first case, the sample data is said to be Class 1 because the boundary area is having class 1 as majority nearest neighbor and by the same way in second case the sample data is related to class 2.

TRAINING AND TESTING OF K-NN CLASSIFIER

In pattern recognition, the ***k*-Nearest Neighbors algorithm** is a non-parametric method most commonly used for classification and regression. The input contains the *k* closest training examples in the feature space in both the cases. The output of the algorithm depends on whether *k*-NN is used for classification or regression. *k*-NN is also a type of instance-based learning and it is also called as lazy learning, where the function is only approximated locally and all computation is deferred until classification. The *k*-NN algorithm is the simplest algorithm among all machine learning algorithms. The good and synthetic malicious customers data are given as a training data for the KNN classifier. The K-NN classifier is responsible for classifying the data as good or bad by studying the predefined training data. The online dataset (for every 30 minutes) is given for classification. This pattern is checked against both the pattern, if any malicious is detected. It is kept as a suspected and given as a input for further verification. Then it is given to ETD algorithm.

B. ETD Algorithm

In ETD, the total consumption of each neighborhood is measured by transformer meters, and is compared with the total amount of usage reported by the smart meters. If at this level a nontechnical loss (NTL) is detected, customers in the area with

abnormal patterns will be selected as suspicious users. For each customer, a K-NN classifier is trained using historic data of the user as well as a synthetic attack dataset. The classifier is then used to decide whether a new sample is normal or malicious. The main contributions of this algorithm are as follows.

1) The design of new algorithm for detecting energy theft attacks. ETD employs transformer meters along with monitoring of abnormalities in customers' consumption patterns to provide a cost effective and high-performance solution for energy theft detection. Through application of appropriate clustering techniques and transformer meters, unlike existing classification- based methods, ETD is robust against contamination attacks and non-malicious changes in consumption patterns, and therefore, achieves a higher DR and a lower FPR.

2) The problem of imbalanced data and zero-day attacks can be addressed by generating a synthetic attack dataset, benefiting from the fact that theft patterns are predictable. Through extensive experiments we show that this significantly improves the DR and enables the detection of a wide range of attack types.

3) ETD provides a higher performance with a lower sampling rate. Hence, it has a reduced effect on customers' privacy. One or more transformer meters measure the total electricity supplied to the customers in each and every neighborhood, TMR *(t)*. This calculated value from the area is compared with the total amount of consumption indicated by the smart meters from the corresponding distribution transformer, Σ_i SMR$_i$ *(t)*. NTL is reported if, for any time *t*, during a day,

$$\mathbf{TMR}\ (t) > \Sigma_i\ \mathbf{SMR}_i(t) + \mathbf{ETL}(t) + \varepsilon$$

Where

ε is the calculation error for TL.
TMR(t) is the transformer meter reading.
SMR(t) is the smart meter reading.

The NTL calculation is performed at each time whenever the new samples are collected. The next step is collection of history of consumption samples for the particular consumer. The comparison results into producing states whether the customer belongs to the benign or attack class. If the calculation of NTL did not detect any malicious attack and the new sample was classified as normal by the comparison, the new sample is added to the normal dataset and the corresponding attack patterns will be generated and added to the theft dataset.

If NTL was detected (i.e.the non-technical loss happened in the particular), then it takes the K-NN output into account. Both of them was compared to detect the attack, a suspicious behavior of the smart meter is reported to the utility if it is repeated *m* times over a certain period. During this time new samples are stored in a temporary database. Once an energy theft is detected, an action is taken accordingly such as on-site inspection. Based upon the amount of NTL calculated, Onsite inspection is made for the smart meter which has higher amount of NTL. If a theft is verified, samples in the temporary database are added to the theft dataset. Otherwise, they will be added to the normal dataset and their corresponding attack patterns to the theft dataset. If the NTL calculation shows there were no theft has happened but the K-NN recognizes the anomaly. This scenario might have three reasons. It might be due to K-NN misclassification or error in NTL calculation, changes in consumption habits, and changes of residents or appliances. Whenever the contradiction happens between the comparisons (i.e. there is no indication of NTL but the K-NN indicates the theft), the particular sample is stored in a temporary database. If the same thing happens in the following days, the old dataset will be dropped and the new dataset is created by taking the current consumption samples from the temporary dataset. Once it becomes large enough to train the K-NN, training process takes place. A flag (f_i) parameter is assigned for each and every smart meter, which is a binary variable initialized to one. When a non-malicious anomaly is detected, as described above, this flag is set to zero and when the situation is resolved it will be set back to one. This facility of the algorithm makes ETD robust against non-malicious changes in consumption pattern.

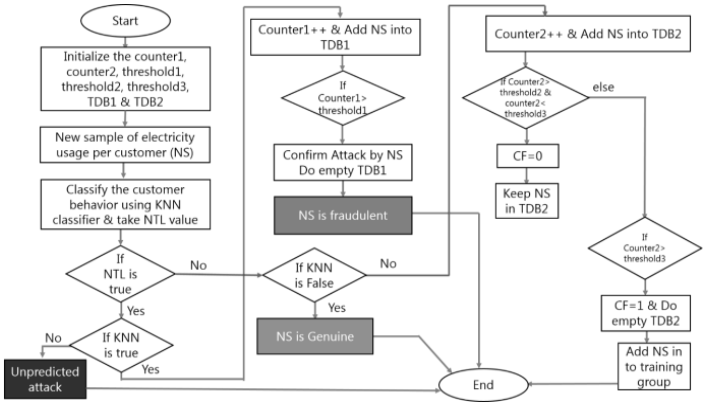


Figure.3. Flow of ETD Algorithm

If NTL was detected, but comparison did not recognize any anomaly and this scenario persists for a certain period, it shows that an attack might be happening, but comparison fails to identify it. In this case, the normal dataset of the consumers is analyzed for sign of data contamination attack, in which by gradual changes in data makes the learning machine to accept a malicious pattern as a normal one. The daily consumption of the customer is studied in a long-term trend manner. If the analysis of historic data does not show any contamination attack, ETD alerts the operator to indicate that an attack might be happening but unfortunately the algorithm could not able to detect it, and the algorithm continues its normal operations for new samples. Hardly, this scenario happens, e.g., when a new load with high consumption is directly connected to a feeder. This step of the algorithm makes ETD robust against contamination attacks.

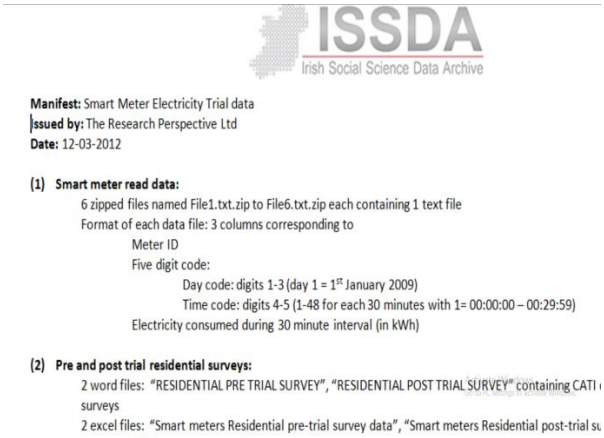
IV. RESULTS & DISCUSSION

This section contains the overall results obtained throughout the project and the outputs are discussed detailed manner. We used the smart energy data from the Irish Smart Energy Trial in our tests. The dataset was released by Electric Ireland and Sustainable Energy Authority of Ireland (SEAI) in March 2012. It includes half hourly electricity usage reports of over 5000 Irish homes and businesses during 2010 and 2011. Customers who participated in the trial had a smart meter installed in their homes and agreed to take part in the research. Therefore, it is a reasonable assumption that all samples belong to honest users.

The large number and variety of customers, long period of measurements and availability to the public make this dataset an excellent source for research in the area of analysis of smart meters data. For each customer there is a file containing half hourly metering reports for a 535 day period. We reduced the sampling rate to one per hour and for each customer divided the file into a dataset of 535 vectors, each with 24 components.

A. ISSDA Dataset

The Irish Social Science Data Archive's (ISSDA) mission is to ensure wide access to dataset and to promote the international comparative studies of the economy and it is the Ireland's leading Centre for data acquisition, preservation, and dissemination. The ISSDA provided the 6 Zipped files in text format (File1.txt to File6.txt) .


ISSDA
Irish Social Science Data Archive

Manifest: Smart Meter Electricity Trial data
Issued by: The Research Perspective Ltd
Date: 12-03-2012

(1) **Smart meter read data:**
6 zipped files named File1.txt.zip to File6.txt.zip each containing 1 text file
Format of each data file: 3 columns corresponding to
Meter ID
Five digit code:
Day code: digits 1-3 (day 1 = 1st January 2009)
Time code: digits 4-5 (1-48 for each 30 minutes with 1= 00:00:00 – 00:29:59)
Electricity consumed during 30 minute interval (in kWh)

(2) **Pre and post trial residential surveys:**
2 word files: "RESIDENTIAL PRE TRIAL SURVEY", "RESIDENTIAL POST TRIAL SURVEY" containing CATI surveys
2 excel files: "Smart meters Residential pre-trial survey data", "Smart meters Residential post-trial su

Figure.4. Dataset from Issda

B. CONVERSION OF TEXT FORMAT DATASET TO EXCEL FORMAT

The given dataset is in text format it should be converted into matrix format to ease the upcoming process. The text file is imported in matrix tool and each of the columns is separated by delimiter option as space. And then matrix format is selected for mat format. It contains three column first column indicates the unique consumer ID and the second column indicates the five digit code (1-3 indicates day code (ie)1 st jan 2011 and 4&5 indicate 1-48 for each 30 minutes and the third column indicates the consumption unit in KWh during 30 minutes.

C) CREATION OF SYNTHETIC DATASET

The historical data is considered as the good profile. Thus it is mandatory to create the malicious pattern. This malicious pattern is created by multiplying the random number to the good pattern. Now both the pattern can be given to train the K-NN classifier. Once it is trained, it will be able to classify the consumer's real time data by matching against the historical pattern. Finally, the ETD algorithm makes its decision about the classification. Now both the output such as K-NN and NTL is given to the ETD algorithm for testing. If both of the result says that there is theft, then it confirms the pattern as fraudulent. If the suspicious rate exceeded the given threshold (say 3), then it predicts the consumer is malicious.

D) TESTING

To test the overall process, the good pattern is taken from the benign samples and then converted into malicious sample by multiplying 0.5 with the benign pattern. Here the consumer-ID 5 is selected for testing and changing the pattern for the time period from 11 to 15. Since the K-NN classifier is trained with both the benign and malicious pattern and by now it is ready to test the real time data. Now the consumer-ID 5 real time data is given to the K-NN classifier. It checks the given pattern by plotting them against the historical pattern. Then it classifies them whether it belongs to benign or malicious pattern. It produces the result by giving zeroes and ones. Indication of ones is genuine whereas zeroes indicate suspicious pattern. The Non-technical Loss is detected by taking the difference between the total energy supplied to the area and the measured energy from every house for the particular time period. If the difference is greater than 0, it indicates there is a non-technical loss.

TEM=sum(X);
diffe=TES-TEM;
diffe(diffe>0)=1;
where
TEM = Total Energy Measured
TES = Total Energy Supplied
X = Consumed energy

Now both the output such as K-NN and NTL is given to the ETD algorithm for testing. If both of the result says that there is theft, then it confirms the pattern as fraudulent.

F) PERFORMANCE ANALYSIS

The performance measure of both the k-NN (Fig.V) and the ETD (Fig.VI) have been analyzed and their accuracy levels are calculated respectively.

N=100		Predicted	
		Genuine	**Fraudulent**
Actual	**Genuine**	40(TP)	10 (FP)
	Fraudulent	2 (FN)	48 (TN)

Figure.5. Performance analysis of k-nn classifier

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{40+48}{40+48+10+2} = 0.88 * 100 = 88\%$$

N=100		Predicted		
		Genuine	**Fraudulent**	**Undetectable**
Actual	**Genuine**	50 (TP)	0 (FN)	0 (FN)
	Fraudulent	0 (FP)	43(TN)	7 (FN)
	Undetectable	0 (FP)	0 (FN)	0 (TN)

Figure.6. Performance analysis of ETD

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{50+43}{50+43+0+0} = 0.93 * 100 = 93\%$$

V. CONCLUSION

In this paper, we have introduced ETD (Electricity Theft Detector), a novel algorithm for detecting the customer who is engaging in energy theft using AMI data. It relies on the predictability of consumers' normal and fraudulent usage pattern. Along with the application of k-NN anomaly detector, the algorithm considers all the cases of Energy theft. ETD also provides the retraining facilities that may occur due to change in residents, changes of appliances, alteration in consumption habits. 93% of accuracy is achieved through ETD implementation whereas only 88% of accuracy with KNN classifier. This makes the ETD algorithm robust against any non-malicious changes in consumption pattern.

VI. REFERENCES

[1]. Energy-Theft Detection Issues for Advanced Metering Infrastructure in Smart Grid TSINGHUA SCIENCE AND TECHNOLOGY ISSNll1007-0214ll01/12llpp105-120 Volume 19, Number 2, April 2014 Rong Jiang, Rongxing Lu, Ye Wang, Jun Luo, Changxiang Shen, and Xuemin (Sherman) Shen http://ieeexplore.ieee.org/document/6787363/

[2]. J. Wright. (2010). Smart Meters Have Security Holes. [Online]. http://www.msnbc.com/id/36055667

[3]. IOActive's Mike Davis to Unveil Smart Grid Research at Black Hat USA, IOActive Press Release, Seattle, WA, USA, Jul. 2009.http://www.marketwired.com/press-release/ioactives-mike-davis-to-unveil-smart-grid-research-at-black-hat-usa-1214671.htm

[4]. S. Amin, G. A. Schwartz, and H. Tembine, "Incentives and security in electricity distribution networks," in Decision and Game Theory for Security. Berlin, Germany: Springer-Verlag, 2012, pp. 264–280.https://link.springer.com/chapter/ 10.1007/ 978 -3-642- 34266-0_16

[5]. A. A. Cardenas, S. Amin, G. Schwartz, R. Dong, and S. Sastry, "A game theory model for electricity theft detection and privacy-aware control in AMI systems," in Proc. IEEE Allerton Conf. Commun. Control Comput., Monticello, IL, USA, 2012, pp.1830–1837.https://www.google.co. in/url?sa=t&rct=j&q= &esrc=s&source=web&cd=1&cad=rja&uact= 8&ved=0ahUKEwiX6biP2NXSAhUItY8KHVLxAQ8QFggZMAA&url=https%3A%2F%2Fpeople.eecs.berkeley.edu%2F~schwartz%2FAllerton2012Alvaro.pdf&usg=AFQjCNFfPI296pTMd8KIUUXsA3lxfcpOiA&bvm=bv.149397726,d.c2I

[6]. E. Angelos, O. R. Saavedra, O. A. Cortes, and A. N. de Souza,"Detection and identification of abnormalities in customer consumptions in power distribution systems," IEEE Trans. Power Del., vol. 26, no. 4,pp. 2436–2442, Oct. 2011.http:// ieeexplore.ieee. Org/document/5989884/

[7]. S. Depuru, L. Wang, and V. Devabhaktuni, "Support vector machine based data classification for detection of electricity theft," in Proc. IEEE Power Syst. Conf. Expo., Phoenix, AZ, USA, 2011, pp. 1–8.http://ieeexplore.ieee.org/document /5772466/

[8]. D. Mashima and A. A. Cárdenas, "Evaluating electricity theft detectors in smart grid networks," in Research in Attacks, Intrusions, and Defenses. Berlin, Germany: Springer-Verlag, 2012, pp. 210–229.https://link.springer.com/chapter /10.1007/978-3-642-33338-5_11

[9]. S. Depuru, L. Wang, V. Devabhaktuni, and P. Nelapati, "A hybrid neural network model and encoding technique for enhanced classification of energy consumption data," in Proc. IEEE Power Energy Soc. Gen. Meeting, San Diego, CA, USA, 2011, pp. 1–8.http://ieeexplore.ieee.or g/document/6039050/

[10]. S. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, "High performance computing for detection of electricity theft," Int. J. Elect. Power Energy Syst., vol. 47, pp. 21–30, May 2013. http://www.sciencedirect.com/science/article/pii/S0142061512005947

[11]. M. D. Martino, F. Decia, J. Molinelli, and A. Fernández, "Improving electric fraud detection using class imbalance strategies," in Proc. ICPRAM, vol. 2. Vilamoura, Portugal, 2012, pp. 135–141.https://www.semanticscholar.org/ paper/ Improving-Electric-Fraud-Detection-using-Class-Martino-Decia/cbbe3220b07377a1875dc0208bc576886cfcd7df

[12]. S. Salinas, M. Li, and P. Li, "Privacy preserving energy theft detection in smart grids: A P2P computing approach," IEEE J. Sel. Areas Commun., vol. 31, no. 9, pp. 257–267, Sep. 2013.https://pdfs.semanticscholar.org/6f64/85f5de473feea39d16f59c90820695f8c188.pdf

[13]. Dick AJ. Theft of electricity – how UK electricity companies detect and deter.In: Proc European convention security and detection. Brighton, UK; 1995. p.90–5.

[14]. Depuru SS, Wang L, Devabhaktuni V, Gudi N. Measures and setbacks for controlling electricity theft. In: Proc IEEE North American power symposium, Arlington, TX; 2010.http://ieeexplore.ieee.org/document/5619966/

[15]. Cespedes R, Duran H, Hernandez H, Rodriguez A. Assessment of electrical energy losses in the Colombian power system. IEEE Trans Power Apparat Syst 1983;102:3509–15.http://ieeexplore.ieee.org/document/5520188/

[16]. Depuru SSSR, Wang L, Devabhaktuni V. Electricity theft: overview, issues, prevention and a smart meter based approach to control theft. Energy Policy 2011;39:1007–15.http:// econpapers.repec .org/article/eeeenepol/v_ 3a39_3ay_ 3a2011_ 3ai_3a2_3ap_3a1007-1015.htm

[17]. A. Nizar, Z. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," IEEE Trans. Power Syst., vol. 23, no. 3, pp. 946–955, Aug. 2008.http://ieeexplore.ieee.org/document/4578740/

[18]. C. R. Paul, "System loss in a metropolitan utility network," Power Eng.J., vol. 1, no. 5, pp. 305–307, Sep. 1987.http://ieeexplore.ieee.org/document/4807765/

[19]. N. Tobin and N. Sheil, “Managing to reduce power transmission system losses,” in Transmission Performance. Dublin, Ireland: Publ. Electricity Supply Board Int., 1987.http://siteresources.worldbank.org/EXTESC/Resources/Background_paper_Reducing_losses_in_the_power_sector.pdf

[20]. R. L. Sellick and C. T. Gaunt, “Load data preparation for losses estimation,” in Proc. 7th Southern African Universities Power Engineering Conf. , Stellenbosch, South Africa, 1998, vol. 7, pp. 117–120.https://www.researchgate.net/publication/283730148_Load_data_preparation_for_losses_estimation

[21]. I. E. Davidson, A. Odubiyi, M. O. Kachienga, and B. Manhire, “Technical loss computation and economic dispatch model in T&D systems in a deregulated ESI,” Power Eng. J., vol. 16, no. 2, pp. 55–60, Apr. http://ieeexplore.ieee.org/document/1001937/

[22]. S. Depuru, L. Wang, and V. Devabhaktuni, “Support vector machine based data classification for detection of electricity theft,”in Proc. Power Syst. Conf. Exposition (PSCE), Mar. 2011.http://ieeexplore.ieee.org/document/5772466/

[23]. J. Nagi, K. Yap, S. Tiong, S. Ahmed, and A. Mohammad, “Detection of abnormalities and electricity theft using genetic support vector machines,” in Proc. IEEE Region 10 Conf., Nov. 2008.http://ieeexplore.ieee.org/document/4766403

[24]. F. Li, B. Luo, and P. Liu, “Secure information aggregation for smart grids using homomorphic encryption,” in Proc. 1st IEEE Int. Conf. Smart Grid Commun. (SmartGridComm), Oct. 2010. http://ieeexplore.ieee.org/document/5622064/
\

[25]. W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher, “PDA: Privacy-preserving data aggregation in wireless sensor networks,” in Proc. IEEE INFOCOM, May 2007. 2002. http://ieeexplore.ieee.org/document/4215819/

[26]. J. Nagi, K. S. Yap, K. Sieh, S. Ahmed, and M. Mohamad, “Nontechnical loss detection for metered customers in power utility using support vector machines,” IEEE Trans. Power Del., vol. 25, no. 2, pp. 1162–1171, Apr. 2010.http://ieeexplore.ieee.org/document/5286297/

[27]. M. D. Martino, F. Decia,J. Molinelli, and A. Fernández, “Improving electric fraud detection using class imbalance strategies,” in Proc. ICPRAM, vol. 2. Vilamoura, Portugal, 2012, pp. 135–141.http://www.bibsonomy.org/bibtex/ 1dc10af9c2bff6e6ea564957597c0d65a

[28]. S. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, “High performance computing for detection of electricity theft,” Int. J. Elect. Power Energy Syst., vol. 47, pp. 21–30, May 2013. http://www.sciencedirect.com/science/article/pii/S0142061512005947

[29]. S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, “A multi-sensor energy theft detection framework for advanced metering infrastructures,” IEEE J. Sel. Areas Commun., vol. 31, no. 7, pp. 1319–1330, Jul. 2013. http://ieeexplore.ieee.org/document/6486009/

[30]. S. Salinas, M. Li, and P. Li, “Privacy preserving energy theft detection in smart grids: A P2P computing approach,” IEEE J. Sel. Areas Commun.,vol. 31, no. 9, pp. 257–267, Sep. 2013. http://ieeexplore.ieee.org/document/6585887/