

Electricity Theft Detection System

Snehal Sanjay Vipradas^{1*}, Prathamesh Ramesh Bendre², Apurva Shripad Upasani³, Geetha Chillarge⁴,
Anisha Shitole⁵

^{1,2,3,5}Student, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, India

⁴Assistant Professor, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, India

Abstract: Most of theft detection issues are observed in multiple areas. So, this system is proposed to detect the electricity theft, in which reading of electric meter through images are used and preprocessing done with the help of OCR and XGboost. User's monthly electric meter is captured in the camera and stored into the dataset. Upon successful completion of comparison between the meter unit values and its monthly meter average if we get difference between them then theft is detected. The values which actually goes under computation are trained and the results are displayed over the screen.

Keywords: OCR (Optical Character Recognition), Machine Learning, Smart grid meter.

1. Introduction

Stealing electrical power is a criminal practice called as electricity theft. The theft can be detected by reading the electric meter's recording, counting, and performing manual analysis and calculation. Theft can be done by neighbour, landlord, family member or anyone. Electricity theft increases the cost of electricity and it also creates the risks. Tampering the meter is one of the common types of electricity theft it can be occurred when violator interfered into internal mechanism of electric meter. Electrical theft is a third largest form of theft in normal lives. Nowadays efficient human-computer interaction has gained a lot of importance in developing solutions to real world problems and intelligent computing. Machine Learning plays a very important role in this theft detection. Popular algorithms to detect the electricity theft are XGboost and Optical character recognition (OCR).

2. Literature Review

[1] In the ZigBee based Monitoring theft detection and automatic electricity meter reading, we have observed it requires a large amount of man power and is the time energy consumption. To overcome the limitations of this system, proposal of a prototype module which includes advanced wireless technology.

[2] Theft detection in AMI Networks, the author used the DNN algorithm, which is the part of deep learning. DNN algorithm consist of many fully connected layers, that layers produce the final output.

[3] Theft detection in AMI network with random tuning of hyper parameters the theft prediction is done using RNN i.e.,

Recurrent Neural Network. It takes the sequential data as a input to produce the better result.

[4] In the cost-effective electricity theft detection based on IoT, the Arduino MKR thousand micro controller board is employed as a by function element, to co-ordinate the activity of the system and to connect the system to the Wi-Fi network. And to detect the sensitive part of the meter passive infrared sensor is used.

[5] Theft detection with intelligent agent meter based on IEM, which provides power quality, and also control billing system and power theft. Here, Arduino and Raspberry pi is used for validating IEM.

[6] GSM prepared electricity system consist PIC software, which is used as a programming controller. Serial ports are used to send the data to user energy meter.

[7] In Theft detection with smart meter data, temperature dependent predictive model is proposed which uses smart meter data from distribution transformer to detect electricity theft. This model was tested for varying amount of power thefts.

3. Proposed System

A. Problem Definition

In a distribution system power theft is a non-technical loss, this power theft is a crime. It is against to the rule of electricity authority. Due to power theft the appropriate bills are not read by the meter readings and electricity authority cannot take the actual charges of electricity usage. There is no any automatic system is built for a power theft detection and manually, it is not possible so that this system is proposed for the detection of the theft.

B. System Architecture

The proposed system is used to detect electricity theft with the help of optical character recognition (OCR) algorithm. For this system, we required one dataset which contains images of the meter. For the storage of login credentials, we have used one database called as SQLite. SQLite is a C language library which implements small, fast high reliable, SQL reliable, SQL database engine. It is built into most of phones and computers. SQLite file format is cross platform and stable. This proposed system also required one dataset which contains images of electric meters.

*Corresponding author: snehalvipradas2019.comp@mmcoe.edu.in

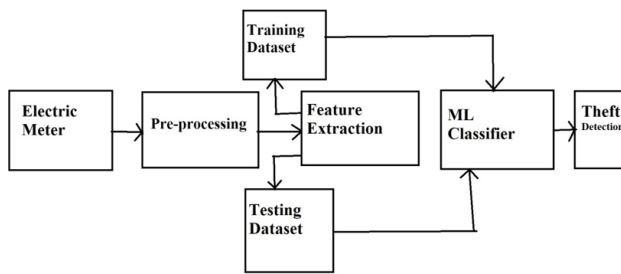


Fig. 1. The system architecture of theft detection

The images of meter are given to the system as an input. Initially we are providing some credential to handle user's account. We are taking users data at the time of registration and after successful login, we can select meter image from the given dataset and perform OCR algorithm to read meters unit and convert it into a text unit. For the prediction of actual result, we have calculated average of monthly meter values and after that comparing this result with meter units we are predicting if there is theft occurred or not.

C. Algorithms Used

1) OCR (Optical Character Recognition)

OCR algorithm analyzed printed or handwritten documents and converts it into the text format. Basically, it converts images of documents captured by camera into a readable, editable format. Below are the steps to implement OCR algorithm,

1. **Image acquisition:** This is a first step. Here, Images of paper document are acquired with the help of optical scanner. Most of paper documents are in the coloured of black and white. So, OCR helps to recognize it.
2. **Pre-processing:** To make a raw data as a usable to computers is done by pre-processing. Optimization of noisy level of image can be done and outside text area is also removed.
3. **Segmentation:** It is a process of grouping character into meaningful chunks.
4. **Feature extraction:** It is used to splitting the input data into a set of features.
5. **Training a neural network:** When all the features get extracted, they are fetched to a neural network to train and then it will recognize the characters.
6. **Post-processing:** This a refinement stage of OCR model. Post-processing is done manually and it is the last step in the algorithm.

2) XgBoost

Generally known as Extreme Gradient Boosting algorithm is one of most widely used in account of efficiency and performance. The algorithm itself has the parallel execution pattern which makes it somewhat different from other boosting algorithms. For the sake of accuracy and relevant results, the system which we have designed basically uses the XGBoost as for the prediction.

4. Results

The below snapshots define the working of the system. The performance is stated below.,

- Images are separate for every individual and initialization is done by choosing the image from the dataset which is of the individual's meter image.
- The uploaded Image gets pre-processed and Units is detected using OCR.
- The check does the summation and calculates the average.
- Prediction consists of XgBoost method. In which, it actually trains the values and perform the operation and predict the theft.
- Below in the fig. 2, theft is detected and it also displays the amount of theft it has detected.
- Looking at the fig. 3, the system tells the consumption is genuine.

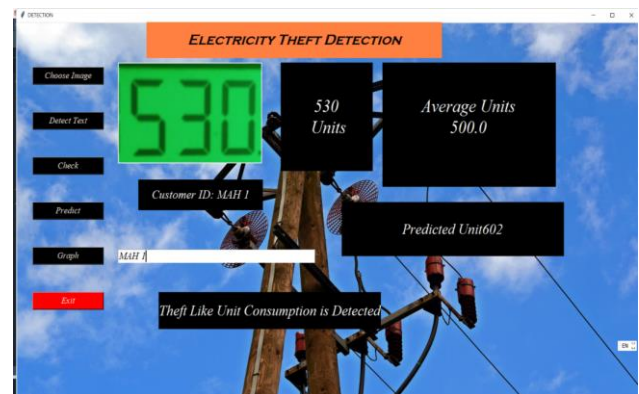


Fig. 2. Theft detected

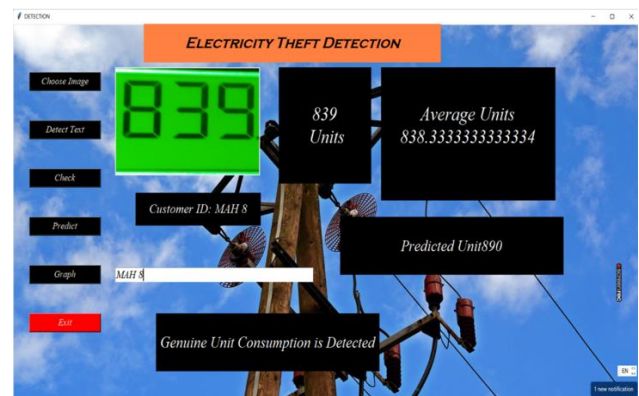


Fig. 3. Genuine consumption is detected

5. Conclusion

In the proposed system detects the electricity theft with the help of machine learning methods. It also helps to electricity utilities to detect electricity theft and they will not bare loss. This is important application of this project. Theft issues might be disastrous on every aspect. While working with OCR and after its implementation here, we came across the advance feature and effectiveness of it. XG-boost which actually enhances our system, leads out the better result for the execution.

6. Future Work

In future work, we are going to design android application. As well as we can add one more facility like alerting the user

immediately whenever theft is occurred. The system itself has the limit such as not having constant view, somewhat issue in predicting results. So, this system can be enhanced and upgrade or new creation is a great option.

References

- [1] J. Nagi, K. Yap, S. Tiong, S. Ahmed, and M. Mohamad, "Nontechnical loss detection for metered customers in power utility using support vector machines," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 1162–1171, 2010.
- [2] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6027 LNCS, pp. 176–187.
- [3] G. Tsekouras, N. Hatziaargyriou, and E. Dialynas, "Two-stage pattern recognition of load curves for classification of electricity customers," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1120–1128, 2007.
- [4] Y. Zhang, W. Chen, and J. Black, "Anomaly detection in premise energy consumption data," 2011.
- [5] S. Dua and X. Du, *Data Mining and Machine Learning in Cybersecurity*, 1st ed. Boston, MA, USA: Auerbach Publications, 2011.
- [6] V. Barnett and T. Lewis, *Outliers in Statistical Data*, ser. Wiley Series in Probability & Statistics. Wiley, 1994.
- [7] N. Billor, A. Hadi, and P. Velleman, "Bacon: Blocked adaptive computationally efficient outlier nominators," *Computational Statistics and Data Analysis*, vol. 34, no. 3, pp. 279–298, 2000.
- [8] E. Eskin, "Anomaly detection over noisy data using learned probability distributions," in *Proceedings of the Seventeenth International Conference on Machine Learning*, ser. ICML '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 255–262.
- [9] E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," in *Proceedings of the 24rd International Conference on Very Large Data Bases*, ser. VLDB '98. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998, pp. 392–403.
- [10] C. Aggarwal and P. Yu, "Outlier detection for high dimensional data," 2001, pp. 37–46.
- [11] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," *SIGMOD Record (ACM Special Interest Group on Management of Data)*, vol. 29, no. 2, pp. 427–438, 2000.
- [12] M. Brito, E. Chavez, A. Quiroz, and J. Yukich, "Connectivity of the mutual k-nearest-neighbor graph in clustering and outlier detection," *Statistics and Probability Letters*, vol. 35, no. 1, pp. 33–42, 1997.
- [13] V. Hautamki, I. Krkkinen, and P. Frnti, "Outlier detection using k-nearest neighbour graph," vol. 3, 2004, pp. 430–433.
- [14] S. Depuru, L. Wang, and V. Devabhaktuni, "Support vector machine-based data classification for detection of electricity theft," 2011.