To Evaluate Rural Power Consumers Consumption to Detect Power Thieving Activities

Bharat Dangar, S. K. Joshi

Abstract: Power Thieving Activities Evaluation To Rural Areas And Direct Point Connections To Small Rural Consumers Of Weaker Parts Of The Rural Society Is One Of The Major Reasons For Electricity Losses. Poor Quality Of Equipment Used In Rural Areas. Large Scale Rural Electrification Through Long 11kv And LT Lines. Artificial Neural Network Can Be Used To Classification And Detection Of Power Thieving In Rural Areas. Train Algorithm According Standard Features And Conditions Employs In Rural Area. Testing Will Give You Better Idea About Datasets According We Can Take New Decision.

Keywords: Power Thieving, consumers, Data mining, Support Vector Machine, Sigmoid, power loss, behavior.

I. INTRODUCTION

Data mining is employed to meet the above challenges in this research. Various data mining analysis on power thieving activities identification and detection in electricity profession are already invented, including artificial neural network, decision trees, rough sets, wavelet-based feature extraction and statistical-based outlier detection and multiple classifiers. Consumer databases as inputs, we can directly have applied in most of these analyses. In most of these cases, data mining has been applied as a tool that enables the detection and prediction of power thieving activities. An above all these applications applied data mining techniques to determine of power thieving activities directly from their databases of rural consumer. The present work is some small different in terms of its study of power thieving activities from time-series case derived from a load profiling case [1].

II. BASIC RURAL CONSUMERS DATA ANALYSIS

A consumer load data is described as “the pattern of electricity load demand of a consumer or a group of consumers over a given period”, in which time interval could be day, week, month, or year. The idea behind is to use a time slots as the effective tool for system planning, tariff rate formulation, devising marketing strategies and load management. The analysis has been categorized as power utility consumers which depends on consumer’s electricity consumption behavior [2]. In many countries electricity consumption profiles recognized as a substitute, price-signal approach to the mean time metered solution which is inappropriate and costly for lower & medium voltage, domestic & trading consumers. Power consumers load profiles helps power utilities to determine the power cost, it helps Power utilities to improve efficiency, planning & trading approach [4]. Typical important of fraud perpetration are as shown in Table 1. In all such cases, electricity consumers intentionally ignore paying their bills or are involved in pilferage, power thieving activities, and unauthorised use. The detection and prediction of Unmetered Power Losses activities on the distribution level is the aim of the current study is to focus where deviations in consumer behaviour are found to exist.

Table 1- Power thieving activities type basis of components identified

Components	Power Utilities	Electricity Consumers
Meter	Inadequacies and inaccuracies of meter reading.	Unauthorized line tapping and diversion.
	Losses due to faulty meters and equipment.	Stealing by bypassing the meter or otherwise making illegal connections.
	Inadequate or faulty metering.	Tampering with meters to ensure the meter recorded a lower consumption reading
	Loss/damage of equipment/hardware, e.g.: protective equipment, meters, cables/conductors, and switchgear.	Faulty meters not reported.
Bills	Inaccurate consumer electricity billing.	Non-payment of electricity bills.
	Inefficiency of business and technology management systems.	Adapting billing irregularities with the help of internal employees.
	Arrangement of billing irregularities with the help of internal employees.	Manipulating readings by bribing meter readers.
	Poor revenue collection techniques.	Inaccurate estimation of non-metered supplies, e.g.: public lighting, agricultural consumption, rail traction.
	Making out lower bills, adjusting the decimal point position on bills.	Avoiding unpaid bills.

In this research, “According to user population in order off power thieving activities detection [5] involves monitoring the behavior of to estimate, detect or avoid abnormal

Revised Manuscript Received October 05, 2019

Bharat Dangar, Electrical Engineering Department, Faculty of Technology and Engineering, The M S University of Baroda
S. K. Joshi , Electrical Engineering Department, Faculty of Technology and Engineering, The M S University of Baroda

behavior". Literature examined based on, an analysis of thieving activities detection and identification techniques is listed below, with professions categorized into two types as follows.

- Electricity professions.
- Type of profession like telecommunications, risk management, credit card provision and insurance.

For the first types, very large numbers of data mining research review on fraud detection and identification in electricity professions are analysed, artificial neural network [7], decision trees, accommodating rough set, statistical-based outlier detection, multiple classifiers and wavelet-based feature extraction. Several of this paper used data mining algorithm by directly determine them to consumer databases as inputs. A combination of multiple classifiers and wavelet techniques have been applied to identify power thieving activities in a power distribution network. Maximum accuracy is obtained with wavelet technique over easy methods because of its capacities in multi and localization resolution study. As another option, rough sets & decision tree were used respectively for the classification of power consumers. Study also conduct using statistical based outlier mining and the artificial neural network [6], where both studies review a method employing a common approach that had consumer databases as its input data source.

III. SUPPORT VECTOR MACHINE

SVM [8] has appeared as one of the most popular and useful techniques for data classification and it has been receiving increasing attention in many areas of research due to its remarkable generalization performance. SVM [9] originates from Vapnik's [10] statistical learning theory and has been observed as being useful for robust outlier detection. The simplest form of SVM classification is the maximal margin classifier. It is used to solve the most basic classification problem, namely the case of a binary classification with linear separable training data. Most generally, the objective of SVM is to generate a model that determines the target value of data instances in the testing set in which only properties are described. The classification target in SVM is to separate the two classes by means of a function derived from available data and thereby design a classifier that will work well on further unseen data.The most appropriate model from among these alternatives is to be selected based on estimations of the predictive accuracy of the model and the fastest processing speed. For the present purpose, the predictive accuracy of a model is defined as "the percentage of test samples that are correctly classified by the model", while the processing speed is referred to in terms of the resulting "computation costs involved in generating and using the model." As show in table 2 selected customers for power thieving activities. Figure 1 shows normalized data between [0,1].

Table 2 Costumers for power thieving activities

City	Area	No. of Customer	Mean	Std. Deviation
BARODA O&M CIRCLE	Urban	10	102.89	9.078
ANAND O&M CIRCLE	Urban	22	102.67	4.789

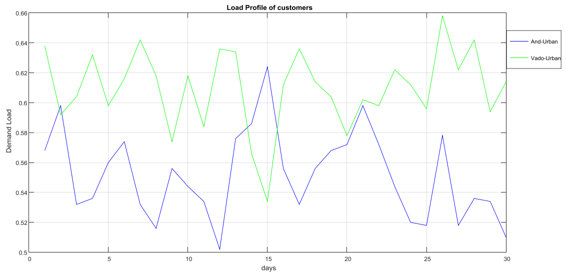

Figure 1 Customers' load normalized data

Table 3 shows the results for time processing durations in seconds applying SVM with the sigmoid function. Datasets with different cost parameters yield different time processing speeds. The cost parameters suitable for each type of days are different, including urban customers with 40 and 0.672 seconds. Table 4 shows the results for classification accuracies as percentages applying SVM with the sigmoid function. It can be seen that different types of days have different numbers of cost parameters for the highest classification accuracy. The best cost parameter for monsoon it is 5 with 94.21% accuracy.Table 3 shows the time processing durations in seconds. monthly training datasets took longer (and fluctuating) times than did testing datasets to complete the classification task due to the large datasets compared to other types of days. The Table 4 shows the classification accuracies as percentages. The monthly dataset classification accuracy for testing data is higher than that for training data. The classification accuracy results for testing datasets on other types of days are slightly lower than for training datasets for all types of days.

Table 3: Results for time processing durations using SVM

C Cost Parameters	SVM Sigmoid in seconds	
	Training (secs)	Testing (secs)
5	0.067	0.062
10	0.124	0.112
15	0.187	0.132
20	0.190	0.140
25	0.213	0.195
30	0.233	0.200
35	0.329	0.265
40	0.672	0.542
45	0.882	0.589
50	0.999	0.623

Table 4: Results for classification using SVM

C Cost Parameters	SVM Sigmoid success rate as percentages	
	Training (%)	Testing (%)
5	94.78	94.21
10	95.00	94.56
15	95.21	95.08
20	95.88	95.65

25	96.46	95.84
30	97.76	96.43
35	98.12	96.71
40	98.87	97.22
45	98.91	97.50
50	98.93	97.89

IV. CONCLUSION

From the above observed that the algorithm the support vector machine has the largest time duration required. Power thieving activities evaluation in urban areas of Gujarat classified using support vector machine. It's give excellent result in radial basis kernel compare to RBF kernel.

REFERENCES

1. M. Kantardzic, "Data Mining: Concepts, Models, Methods and Algorithms", Ed. IEEE Press, 2003.
2. D. Gerbec, F. Gubina, and Z. Toro, "Actual load profiles of consumers without real time metering," presented at the IEEE Power Engineering Society General Meeting, San Francisco, California, USA, June 2005.
3. J. Nagi, K.S. Yap, S.K. Tiong, A. M. Mohammad, and S. K. Ahmed, "Non-Technical Loss Analysis for Detection of Electricity Theft using Support Vector Machines", in Proc. of the 2nd IEEE International Power and Energy Conference (PECON) 2008, Dec. 1-3, 2008, Johor Bahru, Malaysia, pp. 907–912.
4. D. Gerbec, F. Gubina, and Z. Toro, "Actual load profiles of consumers without real time metering," presented at the IEEE Power Engineering Society General Meeting, San Francisco, California, USA, June 2005.
5. A.R. Ahmad, A.M. Mohamad, "Intelligent System for Detection of Abnormalities and Probable Fraud by Metered Customers", in Proc. of the 19th International Conference on Electricity Distribution (CIRED) 2007, Vienna, May 21–24 2007.
6. Inigo Monedero, Felix Biscarri, Carlos Leon, Jesus Biscarri, and Rocío Millan, "Detection of Non-technical Losses in Electrical Consumption Using Neural Networks and Statistical Techniques", ICCSA, LNCS 3984, pp.725 - 734, 2006.
7. C. Cortes and V. Vapnik, "Support-vector network, "Machine Learning, vol. 20, issue 3, pp. 273 - 97, Sept.1995.
8. K. Vojislav, Learning and Soft Computing—Support Vector Machines, Neural Networks and Fuzzy Logic Models. Cambridge, MA: MIT Press, 2001.
9. H.-C. Kim, S. Pang, H.-M. Je, D. Kim, and S. Y. Bang, "Pattern classification using support vector machine ensemble," presented at the 16th International Conference on Pattern Recognition, Quebec, Canada, Aug. 2002.
10. N. Vapnik, The Nature of Statistical Learning Theory, 2nd ed, New York: Springer, 2000, pp. xix, 314