# Importance of Using Load Profiling and Data Mining to Localise and Detect Deviation of the Customers


Salome Colleen. Tshikomba – B-Tech Power Eng. BSc.MOT

City Of Tshwane, S.A.
colleent@tshwane.co.za

**Abstract**

The paper will present the method of load profiling and data mining to detect and minimize stealing of electricity. The paper spotted generation, transmission and distribution of electrical energy as the most affected areas in terms of losses. Losses that occur in generation can be technically defined whereas losses that occur in transmission and distribution cannot be precisely quantified with the sending end information. Therefore this is illustrating the involvement of non- technical losses in transmission and distribution network [1].

In order to achieve the goal of mitigating electricity illegal connections the point of departure is to determine load profiles for different types of customers, this will assists in measuring the extent to which the Utility is losing revenue so that proper mitigation steps can be taken. Data mining and load profiling methods are classified as the one that can give accurate information about all the Utility's customer's load but it also need skilled human resources.

Nowadays technology that is used to curb stealing of electricity depend on proper data collection, analysing, verification etc. the objective of this paper is to use the profiling methods and data mining techniques to classify, detect and predict non-technical losses in the distribution sector[2]. It will also assist in gaining knowledge about the customer behaviour.

# 1 Introduction

Power theft is the most worrying factor in all South African Utility companies. Electricity fraud is a dishonest or illegal use of electricity equipment or service with the intention to avoid billing charge [1]. Losses encountered due to stealing of electricity is called energy losses, divided into Technical and non-technical losses of which no-technical losses are uncontrollable and affects collection of revenue in Utility companies. In this paper, load-profiling method is preferred as a means of reducing power theft or stealing of electricity. Many Countries identify load profiling as an alternative solution that provides a satisfactory and cost effective approach compared to the interval metering solution that is known to be expensive and impractical for small low-voltage commercial and domestic customers [2]. Data mining techniques is also preferred on fraud identification and detection in electricity businesses including rough sets, decision trees, Artificial Neutral Network (ANN), statistical-based outlier detection approach, wavelet-based feature extraction and multiple classifiers. Most of these studies use data mining technique directly to the customer database as the input. The advantage of load profiling is that the Utility knows the customers and can provides better marketing strategies and improves efficiency [3]. The disadvantage is that it needs lot of skilled staff to work on it, and it needs proper monitoring. It can only mitigate the stealing but it cannot stop the stealing due to customers' movement example relocation, stealing of meter boxes or exchange of meter boxes, self-changing of tariffs etc.

# 2 Power system losses

Consists of Technical and Non-Technical losses

Technical Losses are defined as caused by power dissipation due to internal electrical resistance and the affected components include generators, transformers and transmission lines [4] [5].

Non-Technical Losses are mainly cause by power theft or stealing of electricity due to unmetered energy used by the customers and non-customers.

Defined in a formula below as follows:

**ENERGY LOSSES**

$$E_{Loss} = E_{Delivered} - E_{sold} \quad \text{.....1}$$

Whereby the difference between the quantity of the energy delivered and the quantity of the energy recorded is defined as the power losses, and they are sold to the customer [5].

The amount of energy can be represented as $P_{Loss}$ and the amount of energy delivered can be expected as $P_{Delivered,}$ , whereas the amount of energy either recorded or sold will be determined as $P_{sold}$ [5].

**REVENUE LOSS DUE TO TECHNICAL LOSSES**

$$C_{Com\ Loss} = U_{Elec\ Cost} \times E_{Loss} + M_{Maintenance\ Cost} \quad \text{.....2}$$

NON – TECHNICAL LOSSES

$$C_{NTL} = C_{Com\ Loss} - C_{Technical.\ Losses} \quad \text{.....3}$$

# 3 Non- technical losses include the following:

- Non-payment of electricity bills
- Unauthorised line tapping and diversion
- Losses due to faulty meters and equipment
- Inadequate or faulty metering
- Poor revenue collection techniques
- Inadequate and inaccurate of meter reading
- Inaccurate customer electricity billing
- Loss/damage of equipment/hardware e.g. protective equipment, meters, cables/conductors and switchgear
- Inaccurate estimation of non-metered supplies, e.g. public lighting, agricultural consumption, rail traction
- Inefficiency of business and technology management systems

# 4 Load Profiling and data mining for NTL

Load profiling is defined as an electricity load consumption pattern for a customer or a group of customers over a given period [2].

The monthly correlated data provides valuable information regarding the consumption characteristics of LPC customers, which helps to expose abnormal consumption behavior that is known to be highly correlated with NTL activities [1].

## 4.1 Loading conditions

It is crucial to set the loading conditions because it is obvious in the previous studies that different loading conditions yield different load shapes from one customer to another.

Loading conditions is explained using a table as follows:

| Loading conditions | Items |
|---|---|
| | Domestic |
| Type of customer | Commercial |
| | Industrial |
| Location | Urban |
| | Rural |

| | |
|---|---|
| Voltage level | Low Voltage |
| | Medium/High Voltage |
| Type of climate | Rainy/ windy |
| | Hot/Cold |
| Type of day | Week day |
| | Saturday (weekend) |
| | Sunday (weekend) |
| | Public holiday |

Table1: Loading condition [2].

## 4.2 Customer characterization

Data mining is fragmented into different steps with different degree of complexity and different periods of time [5].
Data mining based non-technical losses detection framework (DMNTL) is summarised below:

The representative load profiles gathered from the load-profiling module is used as a reference. In previous studies, most of the researchers rejected the untypical load profile due to insignificant behavior. However, in the current study the untypical load profile is used as a benchmark for investigation with outlier detection techniques.
The benchmarking is based on two types of load profiles, which is abnormal behavior pattern and normal behavior pattern. The abnormal behavior pattern is taken further for investigation.
Compare load profile gathered with the new load profile, and use it as a reference in detecting non-technical losses activities. New load profile is updated if no outlier found, but if there is deviation, then the new load profile be investigated using any outlier detection techniques including statistical based, density based, distance based, model based and deviation based.
If the outlier is detected is confirmed due to NTL, then the new load profile is updated for an anomalous reference for forecasting purposes. Then this load profile is used as a reference for forecasting NTL activities for new load profile using forecasting techniques such as SVM and time series [6].

## 4.3 Data selection

- When choosing customers it is very important to look at the following feature characteristics:
- Period of recorded invoices e.g. monthly or bimonthly.
- Geographical localization e.g. all customers are in Region 1 or Region 2.
- Contractual power: LPU or SPU (HT, 1ph, 3ph).
- Economic activity classification e.g. high rate of NTLs, the investigations will be

Centered in these sectors.

- Consumption range, 20Amps to 60Amps etc.
- History of customer inspection

## 4.4 Data Preprocessing

With respect to data cleaning, no data should be rejected from a set, customers with less than six monthly register per year were eliminated and who had negative values on consumption attributes. On the other hand, reflection exercise about lecture consumption data and billed consumption data are necessary. In a normal situation, the consumption billed depends on the results of the consumption read but is not always true. If the company get no access where consumption made. Estimation is possible based on historical consumption. Several and continuous differences between read data and billed data show abnormal behavior [6].

# 5 Conclusions

The load profiling technique is a unique method of checking the electrical load consumption pattern of the customer or customers over a given period. It is easy to use when detecting the outliers but it needs competent staff to acquire data , analyse it, interpret to detect outliers and send for audit on site, impose temper fee if any tempering or normalize the meter if malfunctioning.. Many Utility companies reported good results upon using this technique and is cost effective when compared to other techniques that uses licensed software. Load profiling produce good results where a large database with too many customers that are connecting illegally is experienced, the results can be good and systematically checked, to give more concentration on cluster of customers with a high rate of historical NTL, s.

# 6 References

[1] Nagi J. Yap K.S., Nagi F., “NTL Detection of electricity theft and abnormalities for large power consumers in TNB Malaysia”, IEEE Power Engineering Journal, 2010
[2] Nizar A.H., Dong Z.Y., Jalaluddin M., “Load profiling method in detecting non-technical loss activities in a power utility”, IEEE First International Power and Energy Conference PEC, 2006
[3] Gerbec D., Gasperic S., Smon I., Gubina F., “An approach to customers daily load profile determination”, IEEE Power Engineering Society, 2002
[4] Glauner P., Meira J. A., Valtcher P., State R., Bettinger F., “The challenge of non-technical loss detection using artificial intelligence: A survey”, International Journal of Computational Intelligence System, Vol.10, p760-775, 2017
[5] Ahmad T., “Non-technical loss analysis and prevention using smart meters”, International Journal, Huazhong University of Science and Technology, May 2017
[6] Biscari F., Leon C., Guerrero J. I., “A mining framework to detect non-technical losses in power Utilities”, University of Seville Spain, January 2009