24th International Conference & Exhibition on Electricity Distribution (CIRED)

12-15 June 2017

Session 6: Challenges for DSOs in new business environments

Non-technical losses: detection methods and regulatory aspects overview

ISSN 2515-0855
doi: 10.1049/oap-cired.2017.0825
www.ietdl.org

Christina Papadimitriou[1] ✉, Giorgis Messinis[1], Dimitris Vranis[2], Sophia Politopoulou[2], Nikos Hatziargyriou[2]

[1]*National Technical Univeristy of Athens-NTUA, Athens, Greece*
[2]*Hellenic Electricity Distribution Network Operator-HEDNO, Athens, Greece*
✉ *E-mail: chpapadi@mail.ntua.gr*

Abstract: Non-technical losses have been reported as one of the most serious problems faced by Electric Utilities. This study provides an overview of most recent methods for electricity fraud detection, based on concepts from data mining, state estimation, game theory and so on. Furthermore, metrics for evaluating such above methods will be discussed and evaluated. It also provides a comparative overview of the main regulatory aspects of frauds based on questionnaire circulated among ten EE countries. Results are discussed and evaluated.

1 Introduction

Several cases of high non-technical losses have been reported in many parts of the world showing that electricity fraud is a high dimensionality problem, implicating social, technical and financial factors, sometimes strongly related to both time and location. These characteristics compose a problem that requires extensive analysis, expert knowledge and state-of-the-art technical tools to be solved.

In this paper, an overview of the most recent methods for electricity fraud detection, based on data mining and power system analysis is given. Furthermore, metrics for evaluating fraud detection methods will be discussed.

It is easily understood that, the provision of expert tools alone for fraud detection will not solve the problem, if an appropriate regulatory framework is not defined. In particular, the following regulatory issues need to be addressed: the role and responsibilities of distribution system operators (DSOs) and other stakeholders, the methodology used to calculate the amount to be paid by the offender, the legal measures and penalties to combat and prevent frauds and so on.

This paper provides a comparative overview of the main regulatory aspects of frauds based on questionnaire circulated among ten EE countries. The questionnaire was formulated and sent by the Strategy and Regulatory Department of HEDNO to members of the WG-Distribution Regulation Policy of EURELECTRIC in the first months of 2016.

The questionnaire is divided into three parts:

(i) The first concerns the role and responsibilities of the stakeholders associated with electricity frauds, such as network operators, suppliers and regulators.
(ii) The second is the financial part and more specifically concerns the methodology used to calculate the amount to be paid by the offender.
(iii) The last module refers to the non-financial measures and penalties that apply in each country to combat and prevent the phenomenon of electricity frauds.

Ten (10) countries were involved:

- Greece (GR)
- Spain (SP)
- Bulgaria (BG)
- Poland (PL)
- Netherlands (NL)
- Austria (AU)
- French (FR)
- Italy (IT)
- Germany (GE)
- Latvia (LT)

The paper is structured as follows. Section 2 reviews the methods of electricity frauds while the Section 3 is a discussion of the evaluation metrics. In Section 4, the questionnaire results regarding the main regulatory aspects of frauds are given. Section 5 concludes the paper.

2 Overview of electricity fraud detection methods

A large number of fraud detection methods can be found in literature [1], although the characteristics of real fraud detection systems (FDSs) used in the field are not widely available. Such systems can be organised in the following three categories according to the resources used: data oriented, network oriented and hybrids.

Data oriented methods make use of data mining and data analytics methods on consumer related data, like time series of active energy consumption, consumer location and characteristics and so on. Such methods make no use of network topology or specialised distribution network infrastructure and they are typically agnostic of the grid. A number of data oriented algorithms can be used for classifying the behaviour of consumers regarding frauds. What differentiates these algorithms is the use (or not) of labelled data sets during training. In the case of supervised classification algorithms, the existence of labels for both classes (i.e. fraud and no-fraud) is assumed. This means that the DSO keeps a well-informed database which already includes a large number of verified fraud and no-fraud cases. The most common algorithm in this case is the support vector machine (SVM) which has also proven to be quite efficient, examples of supervised classification with SVM can be found in [2, 3]. Semi-supervised or anomaly detection methods can be used in case only one of the two classes is priory known. In this case, the DSO is sure that a group of consumers in not committing fraud, but this does not mean that the rest do (the opposite case is possible too). Examples for such methods can be found in [4, 5] where two entirely different statistical approaches have been used for detecting frauds. Finally, unsupervised methods are used in the more realistic case where the DSO has no prior knowledge of labels at all. In this case, the two classes can be separated by using

CIRED, Open Access Proc. J., 2017, Vol. 2017, Iss. 1, pp. 2830–2832

appropriate clustering algorithms. Fuzzy c-means clustering is used in [6], for example, for detecting frauds. The availability of class labels is one of the most important parameters when choosing a data oriented FDS.

Network oriented methods make use of network related data and resources, including network topology, feeder remote terminal unit (FRTUs), observer meters [meters installed on medium-voltage (MV)/low-voltage (LV) transformer measuring aggregate consumption]. Most of these approaches include at some point the use of the observer meter (also known as balance meter), a smart meter installed in the MV/LV transformer which is used to check energy balances in parts of the grid. In case the reading of this meter substantially differs from the sum of the consumer smart meter measurements, fraud can be suspected in the area. The technical losses must also be included in the balancing check. Although this method is good for locating non-technical losses in a specific area, it cannot indicate specific consumers. In order to enhance such methods, power flow and state estimation techniques are used together with sensor placement algorithms. Sensor placement examines the installation of more sensors throughout the LV distribution grid in order to localise frauds. Examples of network oriented methods can be found in [7, 8]. Such methods can be much more accurate than data oriented methods but can be difficult to implement, since they require data that are not always available like the LV network topology. In addition, they are typically more expensive, since they require the use/installation of new devices.

Finally, hybrid methods are a mixture of the above two. Data oriented and network oriented methods are combined in an effort to reduce costs and improve accuracy. The observer meter may be used for indicating areas where technical losses occur while data oriented methods may be used to indicate specific consumers inside that area. Examples of such methods can be found in [9, 10] where SVMs and decision trees are used together with observer meters for locating frauds.

3 Evaluation metrics for fraud detection methods

Any FDS, data oriented, network oriented or hybrid, will produce a list of consumers suspected for frauds. These are marked as follows.

A true positive (TP) occurs in case a consumer is actually committing fraud and the FDS indicates fraud. FDSs should maximise TPs in contrast to false positives (FPs). A FP occurs when the consumer does not commit fraud but the system marks it as fraud. FPs must be kept low, in order to avoid costly inspections on site. True negatives (TNs), i.e. cases correctly predicted as not fraud should also be maximised, while keeping false negatives (FNs) low. These four categories are the main components for calculating evaluation metrics. A large number of metrics can be computed from the confusion matrix, among which the most important ones are the detection rate (recall), FP rate (FPR) and accuracy.

There are two aspects of the fraud detection problem that make the use of more metrics of utmost importance. The first one is the class imbalance problem and the second is the base rate fallacy phenomenon. Both problems are rooted in the fact that usually the fraud class contains much less members than the not-fraud class. This is typical to most fraud detection problems and not only for non-technical loss detection.

Class imbalance creates problems when training data oriented classifiers. A number of methodologies can be followed in order to solve this problem (like over-sampling the minority class or under-sampling the majority class), but still a metric like recognition rate [11] must be used. This metric includes both FP and FN cases, takes into account the size of the two classes and is calculated as follows:

$$\text{Rec.Rate} = 1 - 0.5\left(\frac{\text{FP}}{N} + \frac{\text{FN}}{P}\right)$$

The base rate fallacy phenomenon [9] can mislead a FDS designer in the following way: assume a classifier with DR = 95% and FPR = 1%, which is realistic given the results presented in various papers. Assume also that the probability of fraud in the population is 1%. The probability of fraud given an alarm produced by the FDS is called Bayesian detection rate (BDR) and calculated as follows:

$$\text{BDR} = \frac{P(I) \cdot \text{DR}}{P(I) \cdot \text{DR} + P(\neg I) \cdot \text{FPR}} = 48.97\%$$

where $P(I)$ is the probability of fraud and $P(\neg I)$ is the probability of no fraud.

This means that even with DR = 95% and FPR = 1% approximately half of the alarms produced by a FDS will be false alarms. Given the typically low values of $P(I)$, increasing DR will not significantly improve the BDR. Systems with low values of FPR (and high values of DR at the same time) must be thus implemented in order to obtain high values for BDR. Deciding the acceptable values for any metric is not easy and one must always take into account the following parameters: the compensation associated to detecting a fraud (e.g. in the form of monetary penalty), the cost of not detecting a fraud (energy not billed) and the cost of a false alarm (cost of manual meter inspection).

4 Regulatory aspects of frauds

In this section, the questionnaire results are comparatively provided.

4.1 Responsibilities and role of the DSO

In all countries, the DSO is in charge of the identification and certification of the electricity frauds. In Spain and Germany, it is indicated that together with the operator, other stakeholders may participate, such as the supplier.

Also in all countries, the amount of energy loss estimation is done by the DSO.

Responsible for calculating the payable amount of money by the offender is:

- 'The operator' in Bulgaria, Poland, Latvia and Netherlands
- 'The supplier' in Italy, France and Greece. In Italy, the DSO calculates the amount of network damages that occurred and are added to the amount of the debt.
- 'The supplier/operator' in Spain, Austria, Germany. The role of the DSO here is to determine the network charges, while the supplier calculates the energy charges. In case of no contract with the supplier, the calculation is made entirely by the DSO.

Fig. 1 shows schematically the role of the DSO in European countries as far as the economic calculation of electricity frauds is concerned.

The methodology for the economic calculation is determined by the Regulatory Authorities in Spain, Bulgaria, France and Italy with the participation of the DSO (volume of losses). In other countries (Poland, Holland, Austria, Germany, Greece, Latvia), the methodology is specified by the operator. Especially in Latvia, the Government sets maximum value of the recalculated amount of electricity according to the maximum discharge capacity of the connection within a 24 h period.

Regarding the collection of payments:

- In Spain and Austria, 'the DSO along with the supplier' (in case there is a legal contract with a supplier) are in charge, each one for the corresponding part of the bill. Otherwise, only the DSO is in charge. In Bulgaria, for the clients with regulated tariffs, DSO is responsible; otherwise the supplier is in charge.
- 'Only the operator' is responsible in Poland, Latvia and Netherlands.
- 'Only the supplier' is responsible in France, Germany, Greece and Italy (Fig. 2).

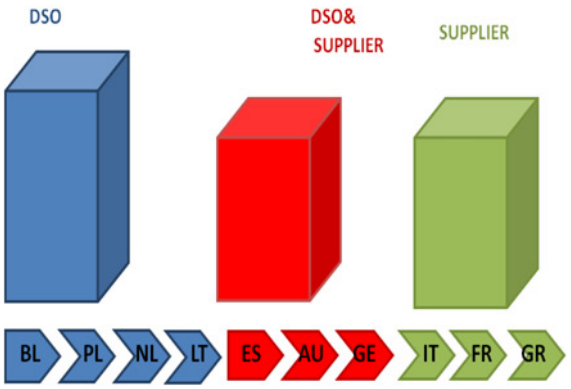

Fig. 1 *Role of stakeholders in the economic calculation [based on HEDNO questionnaire]*

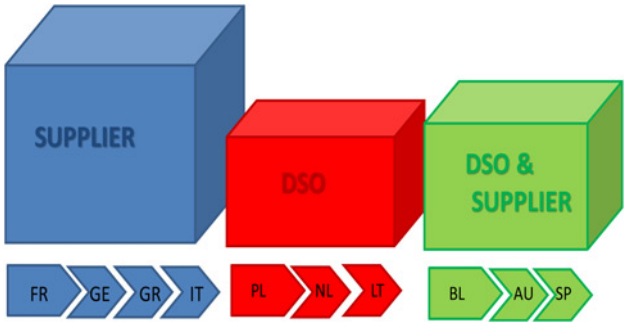

Fig. 2 *Role of stakeholders in collecting the revenues [based on HEDNO questionnaire]*

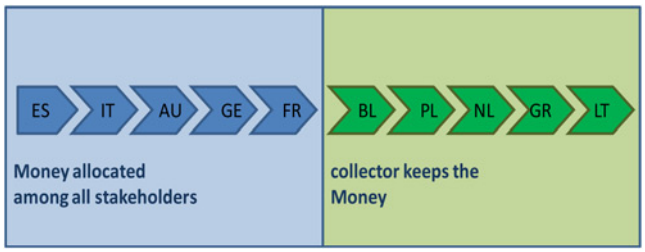

Fig. 3 *Revenues allocation [based on HEDNO questionnaire]*

Finally, the amount recovered is shared among all players accordingly (Spain, Austria, France, Italy, Germany). More specifically, the energy part stays at the supplier, while the network charges are shared among the rest beneficiaries.

Revenues are not shared in five countries (Bulgaria, Poland, Netherlands, Greece, Latvia) (Fig. 3).

4.1.1 Methodology for calculating the amount of recovery: All countries calculate the amount according to the conventional formula ($E \times P$) (E is the energy used estimation, P is the cost). Additional amount to be paid on top is considered in Poland, Latvia, Italy and Germany.

- In Poland, the above amount ($E \times P$) is multiplied by 5 in the absence of prior agreement and by 2, otherwise.
- In Germany, Italy and Latvia, additional costs are calculated if there is a need of a new meter, reconnection, network damages and so on.
- In Latvia, electricity price is set the same as for last resort supplier.

Eight out of ten countries calculate the amount of money with no energy related charges, e.g. VAT (Spain, Bulgaria, Holland, Austria, France, Italy, Germany, Greece, Latvia). In Netherlands, the VAT is not calculated in the absence of agreement. In Poland, there is no such prediction.

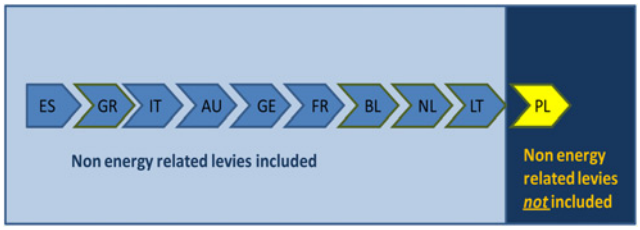

Fig. 4 *Non-energy related levies [based on HEDNO questionnaire]*

4.1.2 Non-financial penalties and measures: All countries except Poland disconnect the offender and make parallel use of extrajudicial remedies.

Reconnecting the offender to another supplier is feasible without paying debts in Spain, Poland, Netherlands, Greece and Italy. Reconnection is not possible without paying for Bulgaria, Austria, France, Latvia and Germany. Exceptions are recognised in Austria, in case of vulnerable customers (Fig. 4).

5 Conclusions

Electricity fraud is a widely spread phenomenon which must be tackled in order to reduce costs, ensure individuals safety and infrastructure integrity and prevent subsidies from the honest costumers. A large number of fraud detection methods have been presented in literature borrowing concepts from various fields including data mining, network analysis, cyber security and so on. When choosing a specific methodology, the available resources and costs for implementing the FDS must be taken into account. In addition, a variety of metrics must be evaluated not forgetting to account for the data imbalance and base rate fallacy problems.

The key role of the DSO to identify, certify and manage the electricity frauds is recognised in all European countries. In some cases, close cooperation with suppliers is needed.

6 References

1 Jiang, R., Lu, R., Wang, Y., *et al.*: 'Energy-theft detection issues for advanced metering infrastructure in smart grid', *Tsinghua Sci. Technol.*, 2014, **19**, (2), pp. 105–120
2 Nagi, J., Yap, K.S., Tiong, S.K., *et al.*: 'Nontechnical loss detection for metered customers in power utility using support vector machines', *IEEE Trans. Power Deliv.*, 2010, **25**, (2), pp. 1162–1171
3 Messinis, G., Dimeas, A., Rogkakos, V., *et al.*: 'Utilizing smart meter data for electricity fraud detection'. SEERC Power Conf., Portoroz, Slovenia, 7–8 June 2016
4 McLaughlin, S., Holbert, B., Fawaz, A., *et al.*: 'A multi-sensor energy theft detection framework for advanced metering infrastructures', *IEEE J. Sel. Areas Commun.*, 2013, **31**, (7), pp. 1319–1330
5 Spirić, J.V., Dočić, M.B., Stanković, S.S.: 'Fraud detection in registered electricity time series', *Int. J. Electr. Power Energy Syst.*, 2015, **71**, pp. 42–50
6 Dos Angelos, E.W.S., Saavedra, O.R., Cortés, O.A.C., *et al.*: 'Detection and identification of abnormalities in customer consumptions in power distribution systems', *IEEE Trans. Power Deliv.*, 2011, **26**, (4), pp. 2436–2442
7 Salinas, S.A., Li, P.: 'Privacy-preserving energy theft detection in microgrids: a state estimation approach', *IEEE Trans. Power Syst.*, 2015, **31**, pp. 1–12
8 Lo, C.-H., Ansari, N.: 'CONSUMER: a novel hybrid intrusion detection system for distribution networks in smart grid', *IEEE Trans. Emerg. Top. Comput.*, 2013, **1**, (1), pp. 33–44
9 Jokar, P., Arianpoo, N., Leung, V.C.M.: 'Electricity theft detection in AMI using customers' consumption patterns', *IEEE Trans. Smart Grid*, 2016, **7**, (1), pp. 216–226
10 Jindal, A., Dua, A., Kaur, K., *et al.*: 'Decision tree and SVM-based data analytics for theft detection in smart grid', *IEEE Trans. Ind. Inf.*, 2016, **12**, (3), pp. 1005–1016
11 Ramos, C.C.O., Rodrigues, D., de Souza, A.N., *et al.*: 'On the study of commercial losses in Brazil: a binary black hole algorithm for theft characterization', *IEEE Trans. Smart Grid*, 2016, **PP**, (99), p. 1