FAULT DETECTION IN LOW VOLTAGE NETWORKS WITH SMART METERS AND MACHINE LEARNING TECHNIQUES

Tania Maria Vazquez Sanchez
EdP – Spain
Tania.VazquezSanchez@edpenergia.es

Pablo Pérez Núñez
Universidad de Oviedo
UO232368@uniovi.es

Jorge Díez
Universidad de Oviedo
jdiez@uniovi.es

Jesús Fernández-Lopez
EdP-Spain
jesusfdez@edpenergia.es

ABSTRACT

Smart grid data analytics and artificial intelligence techniques are playing an increasingly critical role, becoming the focal point to understanding low voltage real-time grid performance. This new point of view, (advanced analytics in combination with electrical knowledge expertise), makes flexibility and efficiency in electrical grid management approach real.

HDCE (Hidrocantábrico Distribución Eléctrica) is the Electrical Distribution System Operator for EdP (Electricity of Portugal) around Spain who supplies energy to 650.000 customers. Starting from 2012, this company has nowadays replaced 99% of traditional meters by smart meters.

Based on the analysis of smart metering voltage alarms, recorded from EdP LV distribution network, an automatic learning system has been implemented that groups and orders these alarms helping the grid distribution operator to drive the network technicians to the right and more urgent places where a grid failure is happening, starts to happen or will happen.

INTRODUCTION

As the energy transition gathers pace, Distribution System Operators (DSOs) will need to increasingly perform a more active role in developing, managing and operating their networks. Clearly, the on-going transformation places new requirements on distribution networks in terms of system reliability and operational security, but it also offers opportunities for DSOs to manage their grids in a more **flexible and efficient** manner. In order for the European electricity sector to become carbon-neutral by 2050 will be a significantly higher share of highly volatile renewable energy sources (with most of it likely connected to the distribution networks) alongside new loads such as electric vehicles or heat pumps, introduces new challenges to the design and operation of the distribution system. In this respect, increasing controllability and flexibility of the variable supply and demand, provides a key pathway towards a more robust distribution system[1].

[1] SOURCE: Flexibility in the energy transition. A toolbox for Electricity DSOs. EDSO February 2018

DSOs are facing increased challenges in adapting the distribution network to this new reality, one of the main challenges will be constraints and distribution congestion. As a first step, EdP Spain has been committed to **smart meters data management, as an essential part of control and monitoring the grid**, to get future flexibility and efficiency on network topology and load-adjustment/load-balancing.

Since the installation of these smart meters and associated communications (Power Line Communication, or PLC technology), the vision of the distribution network has expanded and smart meters have been understood as electrical grid sensors. Although, initially their use was only for consumption recording, the smart meters have been updated with the aim of obtaining additional information. Smart meters at this moment do not register electrical variables in internal memory, but they send **events** (action or alarm initiated by a smart meter, when a success occurs), and especially quality service events: overvoltage (OVE) and undervoltage events (UVE). These events are part of the standard set defined by the PRIME association, made up of the main electricity distributors and the main smart meter manufacturers.

This article focuses on the operation, control and protection of the distribution network from the Low Voltage (LV) busbars in secondary substations through the LV grid to the smart meter or client. Networks quality criteria can be improved through the proactive detention of quality service events (OVE and UVE) based on start of the art artificial intelligence techniques.

SENSORIZATION. SMART METER EVENTS

EdP smart meter events can be categorized by 160 types that can be organized in two groups:

a) Quality service events: OVE and UVE (Figure 1)
b) Others: tampering, remote access with wrong key, firmware actualization, neutral loss, critical internal failures, etc

For the case study of this article, to improve the quality of service and proactively detecting breakdowns, we studied quality service events which are defined as those event records that last 3 minutes or more, either above or below the voltage threshold set by law (+-7%) Most electrical equipment is designed to operate properly when supplied with acceptable voltages.

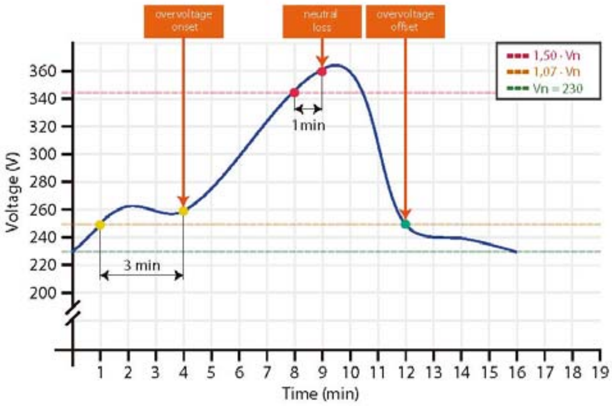

Figure 1. Quality service events (OVE and UVE)

It is important to say that at the begging of this project, GAMMA[2], OVE and UVE were non-spontaneous events. That means that the smart meter could register the event, but it was not able to send it to the master system in real-time. The way to achieve these events before this project were trough *cycle tasks*; that means than once per day the master system asks the devices or smart meters for these non-spontaneous events. This classification in spontaneous and no-spontaneous is defined as an OBIS set adjustment. So, the company change the setting remotely events in order to achieve real-time (or spontaneous) OVE and UVE.

RELATED WORK

An excellent collection of smart metering studies can be read in [1]. The authors explain the smart metering context and useful tasks that can be achieved with all the recorded data.

In [2] and [3] where presented two systems that use self-organized maps (SOM) [4] to build load profiles starting from smart data information. Classic clustering algorithms as k-means [5], have been used with this same objective [6], [7]. In [7], the presented system also carries out a load forecasting, as also [8]-[10] do.

Load data analysis is a usual study when we talk about smart meters. Data processing in combination with Support Vector Machines (SVMs) [11], can be found in tasks as the detection of power-theft [12], [13], or the identification of different types of home devices [14], [15]. Using traditional techniques as PCA [16], can be found a work trying to detect malicious grid tampering [17].

[2] GAMMA: Gestión de Alarmas Mediante Máquinas de Aprendizaje, EdP Project 2018

In [18], authors show a clustering based on finite mixture models (FMM), where they discover different customer behaviour, depending on its load profile and load variability.

As far as we know, there are no studies trying to make grid failure forecasting starting from smart meter events.

ANALYTICAL MODEL

The equipment (about 650.000 meters), send about 300-400 events per minute (any type) and about 100 quality events per minute. First question is, does electrical grid behaviour explain so many events…? Are they reliable? It is necessary to make a distinction among events due to grid failures (as low voltage neutral loss) or grid planning actions (transformer tap regulation, electric lines reinforcement, etc). Being able to find these failures could mean to save time in their resolution, achieving customer satisfaction and Network Technicians (NT) confidence, costs saving and security improvement. Furthermore, it is very important to develop specifically techniques for cleaning and debugging false-positives; this part it is vital to manage new sensors information. That is why artificial intelligence is necessary for electrical companies.

Descriptive proposed model

A descriptive model has been developed to understand what it is happening each hour per secondary substation, in two geographical zones in Asturias, what means about 300.000 smart meters and 5.000-7.000 events per day.

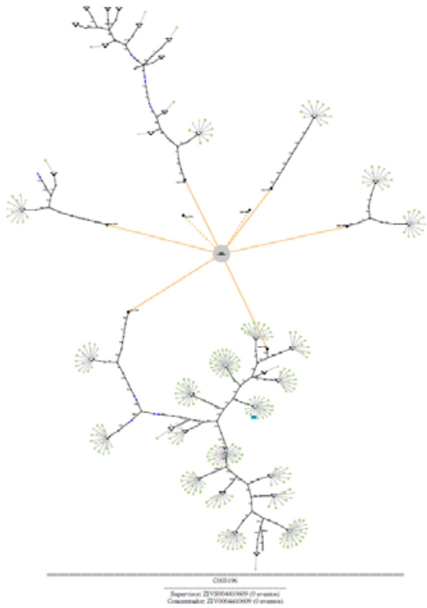

Figure 2: Network representation of each secondary substation and associated branches. Each subnetwork represents a branch and smart meters associated. The number of OVE and UVE are represented by different colours (in the figure green colour means no events). Main networks parameters as section or length, and

some measures as voltage at secondary substation are shown in each network.

Branches description

Each event is sent individually, but its treatment can be grouped. Field knowledge says that if a smart meter detects an overvoltage, other smart meters in same branch should see it (Figure 3). So, first decision, is to group events by branches, and by secondary substations. The main attributes that characterize a branch are showed in Table 1

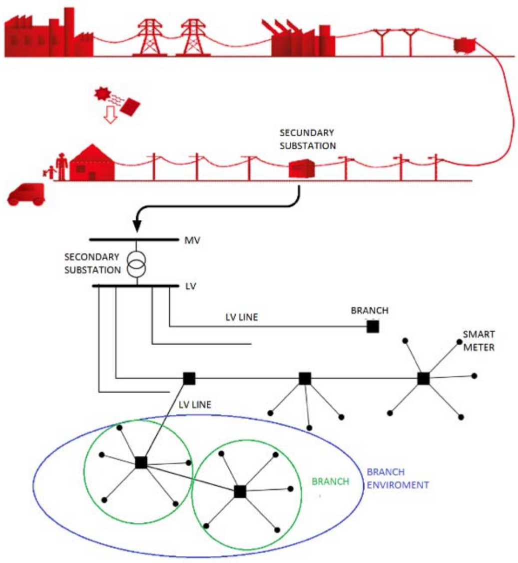

Figure 3: Branch identification and branch environment (line with neighbour branches).

Twice: 1. Per branch 2. Per line neighbours	Events (number)	Last_24_OVE
		Last_24_UVE
		Last_72_OVE
		Last_72_OVE
	Events (%)	Last_24_OVE
		Last_24-UVE
		Last_72_OVE
		Last_72_OVE
	Time interval	Max_UVE
		Max_OVE
		Average_UVE
		Average_OVE
	Smart meters (number)	
	Phases (%)	1-phase
		3-phase
		Other
	Section	
	Average contract power	
	Supplier	ZIV
		SOG
		SAG
		ORB
		Other
From Line	Branches (number)	
From Secundary Substation	Outdoors	

Table 1: Used attributes for branches representation

In total, 48 attributes are used; 23 for branch description, and other 23 for the branch environment (branches in same line), 1 for branch number in each line, and 1 that indicates the secondary substation type (outdoors/indoors). Branch description features are the following: number of OVE and UVE sent in the last 24 and 74 h, smart meters percentage that have sent OVE and UVE in the las 24 and 72 h, maximum and average time interval in the last 72h, smart meters amount per branch, smart meters type (1phase/3phase), branch section, contracted power in each branch, and smart meter supplier. For describing branch environment (all branches associated to the same line), events are aggregated per line, and same attributes have been calculated.

Labelling process

To proceed with the labelling process, during a month OVE and UVE were sent daily to an EdP expert. This expert selected a reduced valid group of branches for order a trip into the field to the NT. The NT, after checking in field, labelled the branches into three different values: i) immediate and necessary trip (class 2), ii) non-immediate but necessary trip (class 1) and iii) non-necessary trip (class 0).

Preferences learning

The aim of this study is to have field trips orders (caused by OVE and UVE) ranked by urgency; so, a preference learning algorithm has been chosen.

We started from labelling branches as has been explained in the previous section.

$$\mathcal{D} = \{(\boldsymbol{a}_i, y_i) : i = 1, \ldots, n\}, \qquad (1)$$

Being a_i a branch, y_i its label and n the total number of branches. Starting from that data set, *preference judgements* can be done, that is, data pairs that indicate that it is *better* choice to travel to a branch than another (*worse*)

$$\boldsymbol{a}_m > \boldsymbol{a}_p, \qquad (2)$$

where a_m represents the branch that it is better to travel, and a_p the one less urgent. Must be considered, that in these judgements it does not appear the branch class, just the preference between going to one or another.

Starting from that preference judgements, the goal will be to find a utility function f to give a higher score to the most interesting branch to visit. The purpose of this function will be to maximize the probability

$$Pr\left(\boldsymbol{a}_m > \boldsymbol{a}_p \iff f(\boldsymbol{a}_m) > f(\boldsymbol{a}_p)\right). \qquad (3)$$

For learning f, we will start from the preference judgements set D_{jp},

$$\mathcal{D}_{jp} = \{(\boldsymbol{a}_m, \boldsymbol{a}_p) : \boldsymbol{a}_m > \boldsymbol{a}_p, m, p = 1, \ldots, n\}, \qquad (4)$$

Where the branch a_m needs to be paid more attention than the a_p . These judgements are generated by making all the positive comparisons that can establish between the branches of the set D (1). In D_{jp}, the opposite comparisons could also have been included; this is, those in which $a_p < a_m$; however, this is not necessary, due to that negative comparisons are symmetrical respect to the positive ones, and when it is pretended to learn a lineal model, as this is the case, they are not necessary if the hyperplane that should be learned cross the coordinate origin.

Lineal model that have been chosen as function *f* is:

$$f(\boldsymbol{a}) = \langle \boldsymbol{w}, \boldsymbol{a} \rangle = \boldsymbol{w} \cdot \boldsymbol{a}^{\boldsymbol{T}} \quad (5)$$

Where *f* represents a *w* scalar product, parameter that we should learn, and a branch *a*. Must be said that there is not independent term to force a hyperplane to cross the coordinate origin, so negative comparisons are not needed.

We assume that all the equation examples (4) are independent and identically distributed (i.d.d.). Furthermore, using *maximum likelihood estimation* and *margin maximization*, the parameter *w* should maximize:

$$\mathcal{L} = \prod_{(\boldsymbol{a}_m, \boldsymbol{a}_p) \in \mathcal{D}_{jp}} Pr\left(f(\boldsymbol{a}_m) > f(\boldsymbol{a}_p) + 1 | \boldsymbol{w}\right). \quad (6)$$

This optimization can be accomplished with an algorithm based on *gradient descent* [19] and applying regularization to the learned parameters for acquiring numeric stabilization. In this way, optimal *w* value will be obtained through the calculation of

$$\begin{aligned} \boldsymbol{w}^* &= \underset{\boldsymbol{w}}{\operatorname{argmax}}\ log(\mathcal{L}) + \nu \|\boldsymbol{w}\|^2 \quad (7) \\ &= \underset{\boldsymbol{w}}{\operatorname{argmin}} -log\left(\prod Pr\left(f(\boldsymbol{a}_m) > f(\boldsymbol{a}_p) + 1 | \boldsymbol{w}\right)\right) - \nu \|\boldsymbol{w}\|^2 \end{aligned}$$

for all the (a_m, a_p) and loss function will be:

$$L = \sum_{(\boldsymbol{a}_m, \boldsymbol{a}_p) \in \mathcal{D}_{jp}} max\left(0, f(\boldsymbol{a}_p) - f(\boldsymbol{a}_m) + 1\right). \quad (8)$$

The *w* learning stage using gradient *descent* will be made in this way:

$$\boldsymbol{w} \leftarrow \boldsymbol{w} - \gamma \left[\frac{\partial L}{\partial \boldsymbol{w}} + \nu \frac{\partial \|\boldsymbol{w}\|^2}{\partial \boldsymbol{w}}\right], \quad (9)$$

where

$$\frac{\partial L}{\partial \boldsymbol{w}} = \frac{\boldsymbol{w} \cdot \boldsymbol{a}_{\boldsymbol{p}}^{\boldsymbol{T}}}{\partial \boldsymbol{w}} - \frac{\boldsymbol{w} \cdot \boldsymbol{a}_{\boldsymbol{m}}^{\boldsymbol{T}}}{\partial \boldsymbol{w}} = \boldsymbol{a}_{\boldsymbol{p}} - \boldsymbol{a}_{\boldsymbol{m}} \quad (10)$$

$$\frac{\partial \|\boldsymbol{w}\|^2}{\partial \boldsymbol{w}} = \frac{\boldsymbol{w} \cdot \boldsymbol{w}^{\boldsymbol{T}}}{\partial \boldsymbol{w}} = 2\boldsymbol{w}. \quad (11)$$

RESULTS

The Network Technicians (NT) achieved the right results in 2 out of 3 grid failures forecasting based on smart meter quality service events (overvoltage and undervoltage events). The right results were: **high priority intervention (10%)**, such as network failures (generally neutral losses or lost connections), **medium priority intervention (44%)**, such as situations in which the transformer tapping at the head-end was high, or **low priority (46%)** as the need for reinforcement on some lines.

Figure 4: Neutral loss detected (high priority intervention)

CONCLUSIONS

Based on the analysis of smart meter events, an automatic learning system has been implemented that groups and orders overvoltage and undervoltage events helping the grid distribution operator to drive the network technician to the more urgent place where a grid failure is happening, starts to happen or will happen. For this study, data from 227 branches is used. This data is categorized and labelled using 48 attributes: most of them associated to branches but others associated with the secondary substation or low voltage lines.

As a result, the new smart grid system has improved the quality of service, in terms of customer satisfaction and in terms of predictive maintenance. Smart meter grid management is an essential part of control and monitoring the grid, to get future flexibility and efficiency on network topology.

APPRECIATION
Pablo Mayordomo Vendrell, Leandro D’Angelo Galán, Oscar Álvarez Pérez, Jose Luis Rodríguez Pérez y Luis Miguel Arniella Cano (EdP)

REFERENCES

[1] D. Alahakoon and X. Yu, “*Smart electricity meter data intelligence for future energy systems: A survey*,” *IEEE* Transactions on Industrial Informatics, vol. 12, no. 1, pp.

425–436, 2016.

[2] J. Nagi, K. Yap, S. Tiong, and S. Ahmed, "*Electrical power load forecasting using hybrid self-organizing maps and support vector machines*" Training, vol. 99, no. 1, p. 31, 2008.

[3] D. De Silva, X. Yu, D. Alahakoon, and G. Holmes, "*A data mining framework for electricity consumption analysis from meter data*" *IEEE* Transactions on Industrial Informatics, vol. 7, no. 3, pp. 399–407, 2011.

[4] T. Kohonen, "*Self-organized formation of topologically correct feature maps*," Biological cybernetics, vol. 43, no. 1, pp. 59–69, 1982.

[5] S. Lloyd, "*Least squares quantization in pcm*" *IEEE* transactions on information theory, vol. 28, no. 2, pp. 129–137, 1982.

[6] M. Espinoza, C. Joye, R. Belmans, and B. De Moor, "*Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series*" *IEEE* Transactions on Power Systems, vol. 20, no. 3, pp. 1622–1630, 2005.

[7] F. L. Quilumba, W.-J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, "*Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities*" IEEE Transactions on Smart Grid, vol. 6, no. 2, pp. 911–918, 2015.

[8] K.-H. Kim, H.-S. Youn, and Y.-C. Kang, "*Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method*" *IEEE* Transactions on Power Systems, vol. 15, no. 2, pp. 559–565, 2000.

[9] S. V. Verdú, M. O. Garcia, C. Senabre, A. G. Marin, and F. G. Franco, "*Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps*" *IEEE* Transactions on Power Systems, vol. 21, no. 4, pp. 1672–1682, 2006.

[10] H.-T. Pao, "*Forecasting electricity market pricing using artificial neural networks*" Energy Conversion and Management, vol. 48, no. 3, pp. 907– 912, 2007.

[11] C. Cortes and V. Vapnik, "*Support-vector networks*" Machine learning, vol. 20, no. 3, pp. 273–297, 1995.

[12] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "*A multi-sensor energy theft detection framework for advanced metering infrastructures*" *IEEE* Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1319–1330, 2013.

[13] M. Anas, N. Javaid, A. Mahmood, S. Raza, U. Qasim, and Z. A. *Khan, "Minimizing electricity theft using smart meters in ami*" in P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2012 Seventh International Conference on, pp. 176–182, *IEEE*, 2012.

[14] M. Mittelsdorf, A. Hüwel, T. Klingenberg, and M. Sonnenschein, "*Submeter based training of multi-class support vector machines for appliance recognition in home electricity consumption data*" in SMARTGREENS, pp. 151–158, 2013.

[15] J. Z. Moro, L. F. Duarte, E. C. Ferreira, and J. A. S. Dias, "*A home appliance recognition system using the approach of measuring power consumption and power factor on the electrical panel, based on energy meter ics*" Circuits and Systems, vol. 4, no. 03, p. 245, 2013.

[16] S. Wold, K. Esbensen, and P. Geladi, "*Principal component analysis*," Chemometrics and intelligent laboratory systems, vol. 2, no. 1-3, pp. 37– 52, 1987.

[17] J. Valenzuela, J. Wang, and N. Bissinger, "*Real-time intrusion detection in power system operations*," *IEEE* Transactions on Power Systems, vol. 28, no. 2, pp. 1052–1062, 2013.

[18] S. Haben, C. Singleton, and P. Grindrod, "*Analysis and clustering of residential customers energy behavioral demand using smart meter data*," *IEEE* transactions on smart grid, vol. 7, no. 1, pp. 136–144, 2016.

[19] H. Robbins and S. Monro, "*A stochastic approximation method*" Ann. Math. Statist., vol. 22, pp. 400–407, 09 1951.