POWER FACTOR SIGNATURE ANALYSIS FOR DISAGGREGATION OF EV CHARGING LOADS FROM AGGREGATED POWER

Manuel L. Nunes
Instituto Superior Técnico,
Universidade de Lisboa, Portugal
manuelpsnunes@gmail.com

João F. P. Fernandes
IDMEC, Instituto Superior Técnico,
Universidade de Lisboa, Lisboa, Portugal
joao.f.p.fernandes@tecnico.ulisboa.pt

José Oliveira
Eneida.io, Portugal
Affiliation - Country
joliveira@eneida.io

P. J. Costa Branco
IDMEC, Instituto Superior Técnico
Universidade de Lisboa, Lisboa, Portugal
pbranco@tecnico.ulisboa.pt

ABSTRACT

With the increasing number of electric vehicles (EV), an overload of the low-voltage power distribution networks is expected, overload reflected in the power distribution transformers. Monitoring of EVs charging plays a vital role in monitoring and predicting how power loading patterns may affect the lifetime of the power transformers. The possibility of monitoring and predicting the charging profile of EVs will help the Distribution System Operator plan the integration of charging stations. This work proposes four algorithms for the first disaggregation stage of the charging EV profiles from aggregated power. These use only active/reactive powers acquired in real-time at the secondary side of a transformer feeding EV charging station(s). The most valuable feature of these algorithms is their capability of disaggregating an EV charging load from the aggregated power of many other loads in the same line.

INTRODUCTION

With the increasing number of electric vehicles (EV), an overload of the low-voltage power distribution networks is expected, reflected in the power transformers. This problem is starting to be reflected in the power distribution transformers, creating adverse effects on the grid [1]-[3]. Monitoring of EVs charging plays a vital role in monitoring and predicting how power loading patterns may affect the lifetime of the power transformers. The possibility of monitoring and predicting the charging profile of EVs will help the Distribution System Operator plan the integration of charging stations [4]-[7].

This work proposes four algorithms for the first disaggregation stage of the charging EV profiles from aggregated power. These use only active/reactive powers acquired in real-time at the secondary side of a transformer feeding EV charging station(s). The most valuable feature of these algorithms is their capability of disaggregating an EV charging load from the aggregated power of many other loads in the same line (commercial buildings, industrial consumers, residential houses, etc.). Of course, in some cases, this method fails to predict the EV charging, thus requiring future additional features to improve its capacity, such as using other features to increase the algorithm's precision [8]-[9].

This R&D project between Instituto Superior Técnico and *Eneida* uses Eneida's DeepGrid platform and their DTVI smart sensor-g installed to acquire data from a distribution transformer substation [10]. This substation feeds an EV charging station, among other loads, operating in the city center of Coimbra (GPS: 38,7436111 -9,1325), Portugal. The charging station (Fig. 1) uses a Type 2-62196-2 charging station socket of 3.7 kWh (slow charging), alternated current (AC).

Data acquired consists of daily 5-minute active and reactive power samples for 141 days distributed along the year from line 2 of the substation transformer. For the development and validation of the proposed algorithm, power consumption data were compared with that from Mobi.e operator, responsible for managing the charging station. With this, it is possible to determine the exact time of an EV load and compare it with the algorithm results using acquired data from the transformer. Four different algorithms are proposed and tested based on the confusion matrix [11] with the indexes: accuracy, miss-classification rate, true-positive rate, false-negative rate, specificity, and precision.

Fig. 1. Charging station used for this case study.

PROPOSED METHODOLOGY FOR SIGNATURE ANALYSIS

This methodology uses the active and reactive power data obtained at the transformer's secondary to calibrate a segregation algorithm capable of identifying two different clusters: with EV charging event and without EV charging event. This methodology was developed using a smart sensor installed at the secondary of a distribution transformer and, by assessing the active and reactive power in each feeder, the EV charging event is characterized, and its key features are extracted.

The exact time instant for an EV charging event was provided by Mobi.e operator, which is responsible for managing the charging station. With this, it is possible to evaluate the active and reactive power curves (*P*-*Q*) changes obtained at the transformer's secondary. The data acquired is obtained for each day (24 hours) with a sampling period of 5 minutes, between the 5th of March to the 26th of July 2018. One example of the active and reactive power evolution by April 4th, 2018, is shown in Fig. 2. This figure is highlighted in the zone where an EV charging event was registered by Mobi.e operator.

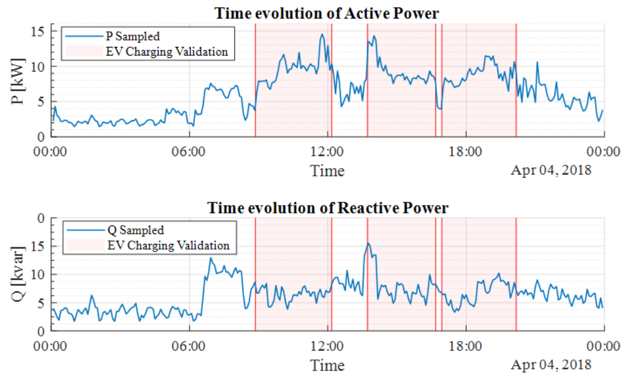

Fig. 2. Active power, *P*, and reactive power, *Q*, of transformer feeder 6, line 2, during 4th of April 2018. In red is highlighted the EV charging event identified by Mobi.e.

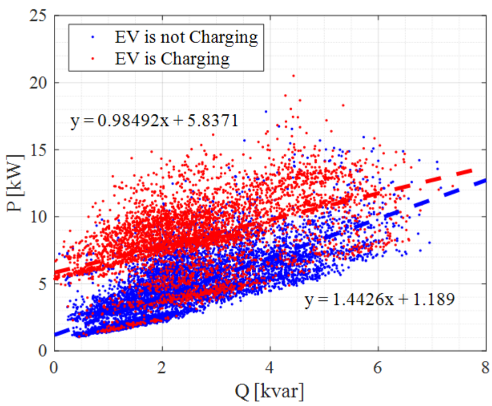

Fig. 3. P-Q data containing all weekdays samples, where EV charging events are marked as red and the remaining ones as blue.

Next, data is divided into two clusters: a) with no EV charging and b) with EV charging. One example is shown in Fig. 3, containing the *P*-*Q* samples between March and May, and where the EV charging events are marked in red. After the disaggregation of the EV charging events, a linear fitting ($P = mQ + b$) was used to characterize the two clusters (red and blue dashed lines in Fig. 3).

With these linear fitting curves as reference, the Gaussian Membership Function (GMF) was used, establishing two relative probabilities to each *P*-*Q* sample, associated with each cluster: p_{gmf}^{C} is the probability of belonging to the "EV is Charging" cluster, and p_{gmf}^{NC} is the probability of belonging to "EV is not Charging."

For each cluster, and considering the linear fitting expression in the form of $aQ + bP + c = 0$, the distance d_i between each sample point and the cluster fitting line, and distance to the corresponding standard deviation are given by (1) and (2),respectively, where i corresponds to the sample point and n to the total number of samples.

$$d_i = \frac{|aQ_i + bP_i + c|}{\sqrt{(a^2 + b^2)}} \tag{1}$$

$$\sigma_{gmf} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} \tag{2}$$

Then, the Gaussian membership function is defined by (3), with d the distance from one *P*-*Q* point to the correspondent cluster fitting line and $\mu = 0$.

$$f_{gmf}(d) = e^{\frac{-(d-\mu)^2}{2\sigma_{gmf}^2}} \tag{3}$$

Fig. 4 represents the GMFs computed using the data from the two clusters: "EV is not Charging" and "EV is Charging." With this, it is possible to obtain a probability for each sample point associated with each cluster. The baseline solution of the proposed classifier method considers the comparison of the probabilities associated with the cluster "EV is Charging" and "EV is not Charging."

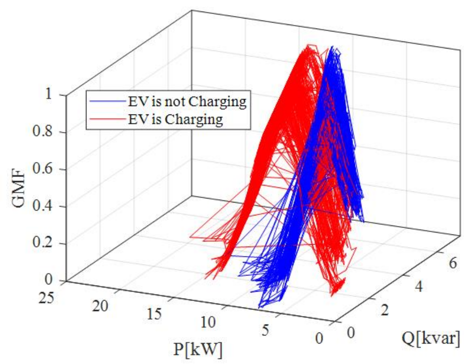

Fig. 4. Gaussian Membership Function for classes "EV is not charging" and "EV is charging" using unconstrained linear data fitting.

Finally, different algorithms were formulated to effectively classify an extracted sample as belonging to the cluster “EV is Charging” or “EV is Not Charging.” These take into account a purely analytical analysis proposed as the baseline solution and a contextualized reasoning, taking into consideration the EV charging panorama, namely the temporal relations between consecutive samples and a threshold establishment to validate a prediction. The following algorithms were evaluated:

A. Without temporal filter, binary (baseline solution);
B. With temporal filter, binary;
C. Without temporal filter, weighted;
D. With temporal filter, weighted.

The temporal filter and binary classification processes are explained in the following sections. Data were divided into training and testing sets. The training set corresponds to data from March to May, and the testing set from June to July.

Algorithm A: without temporal filter and binary

This optimization comprehends the baseline solution where the classifier determines whether a sample is labelled as belonging to the cluster “EV is Charging” or “EV is not Charging” by comparing its probabilities to be contained in each cluster. For every sample of the testing set, the probabilities to be associated with each cluster are compared, and the highest one will determine its classification. This is a binary classification: the sample belongs to the “EV is Charging” cluster if $p_{gmf}^{C} > p_{gmf}^{NC}$, or to the “EC is not Charging” cluster if $p_{gmf}^{C} < p_{gmf}^{NC}$.

Algorithm B: with temporal filter and binary

In this algorithm, chronological analysis of the extracted samples is considered when performing the classification. The reasoning behind this solution concerns the actual user’s behavior of that EV charger, in particular the charging time. Fig. 5 represents a histogram of the EV charging time during the workdays of approximately the 5 months corresponding to the full dataset.

Most EV charges have a time between 140 and 150 minutes, while the second most are between 0 and 10 minutes. However, by doing a separate inspection on the EV charges with less than 20 minutes, both the extracted P and Q evolution show very little variation, with almost no impact on the grid. With this information and considering the objective of predicting an increase of power load due to EV changing events, a temporal filter was applied to classify “EV charging” events. One sample can only be considered as belonging to this cluster if it remains in this cluster for more than 20 minutes.

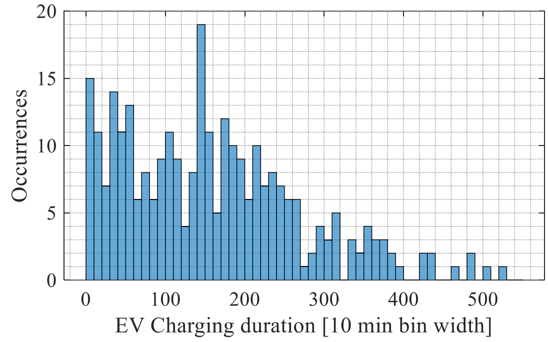

Fig. 5. Histogram of EV charging time during workdays with 10 minutes bin width.

Algorithm C: without temporal filter and weighted

Algorithm C is also an incremental modification to the baseline solution A, without temporal featuring, however considering a threshold (T) in terms of the difference between GMF probabilities to validate a prediction. In this solution, to obtain a positive prediction, one has to achieve the following: $p_{gmf}^{C} \geq p_{gmf}^{NC} + T\ , T\ \in\ ]0,1[$.

Now, to obtain a calibrated T value, a commitment between the truly predicted samples and falsely predicted ones had to be considered. In other words, T was chosen so that the true positive and true negative predictions are maximized, and false positive and false negative predictions are minimized. To accomplish that, a histogram comparing the true positive predictions with the positive ones was computed, and it is presented in Fig. 6.

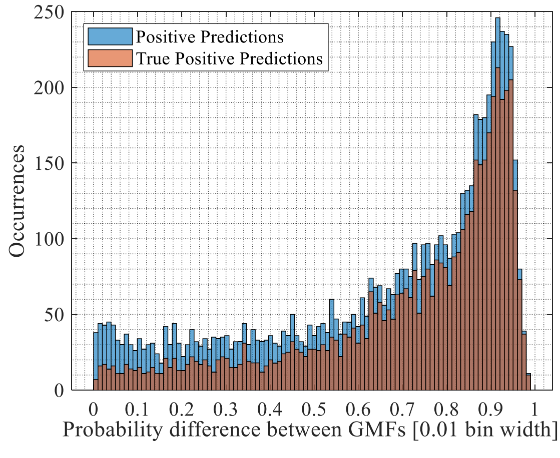

Fig. 6. Histogram representing the number of positive and true positive predictions with different probability differences between GMFs of classes ”EV is Charging” and ”EV is not Charging.

Based on the results presented in Fig. 6, the selected threshold T value was 0.1, given that the total percentage of true positive predictions out of the positive ones is 36.9%. This quantity means that there must be at least a 10 % difference between the GMF probabilities of each class.

Algorithm D: with temporal filter and weighted

Finally, this solution results from a confluence of the three previously algorithms B and C, using the baseline solution with both the temporal filter and the threshold value.

RESULTS

This section comprises the results concerning the proposed algorithms for the classifier. To have a realistic assessment of the performance of this classifier, the confusion matrix [11] was computed, and some efficiency indices were extracted. These indices comprise the overall *accuracy* and *miss-classification rate* (MCR) as well as the *true-positive rate* (TPR), *false-negative rate* (FNR), *specificity*, *and precision* [11]. Furthermore, to increase the impartiality of the assessment of the classifier performance, a stratified 5-fold cross-validation was done using the full dataset, and the displayed results, presented in Fig. 7, are the proceedings. Moreover, it is worth considering the prevalence level of that EV charger of 30.81%, indicating the usage percentage with the used case study data.

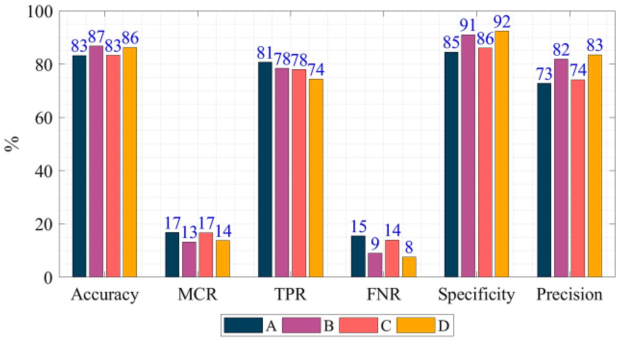

Fig. 7. Representation of the confusion matrix indices to evaluate the classifier performance of each algorithm A-D, being MCR the *miss-classification rate*, TPR the *true-positive rate*, and FNR the *false-negative rate*.

Analyzing the results represented in Fig. 7, the first thing to notice is the *accuracy* and *miss-classification rate*, respectively between 83%-87% and 13%-17% across solutions A to D. These results give an overall perspective on the classifier performance. However, they can be misleading, hence the need for the other indices.

Moreover, knowing that the *prevalence* level is around 31%, there is significantly more data concerning the class "EV is not Charging." This information might be related to the difference between the TPR and *specificity* - how often does the classifier predict that an EV charging event was not occurring when, in fact, it was not occurring - observable in Fig. 7, respectively 74%-81% and 85%-92% across solutions. In addition, the *precision* – the percentage of correct positive predictions - ranges from 73%-83% across solutions.

In conclusion, the four presented algorithms solutions have different performances, and they can be employed according to the DSO specification. If the TPR performance is more important, solution A is more suitable at the expense of a worse FNR. On the other hand, if the minimization of the FNR is more important than the maximization of the TPR, solution B is the best one. In addition, if the *specificity* and *precision* are preferable when performing the classification, solution D is more adequate.

Overall, an equalized solution might be optimization B, guarantying 78% of TPR, 91% of *specificity*, and a *precision* of 82%. In addition to the performance of each algorithm, the feasibility limits of the proposed algorithms were verified.

Up until this point, the research was conducted using the actual *P* of the used EV charger – 3.7kW – corresponding to *Slow Charging* mode [1]. In this subsection, different *P* charging values (P_k^{CH}) are tested to understand the feasibility of this classifier when it is applied to charging stations with various power modes.

In detail, the classifier feasibility is determined when the active charging power varies linearly by a multiplication factor *k*. To accomplish that, whenever a charge was occurring, the *P* load was subtracted by 3.7kW, corresponding to the EV charging, remaining with the original background load $\left(P^{bk} \text{ and } Q^{bk}\right)$. Then the active power during the EV charging events is multiplied by *k* and added to the background load. The *accuracy*, TPR, and *specificity* results of each algorithm classifier with different EV charging power magnitudes are presented in Fig. 8, being *k* the multiplication factor.

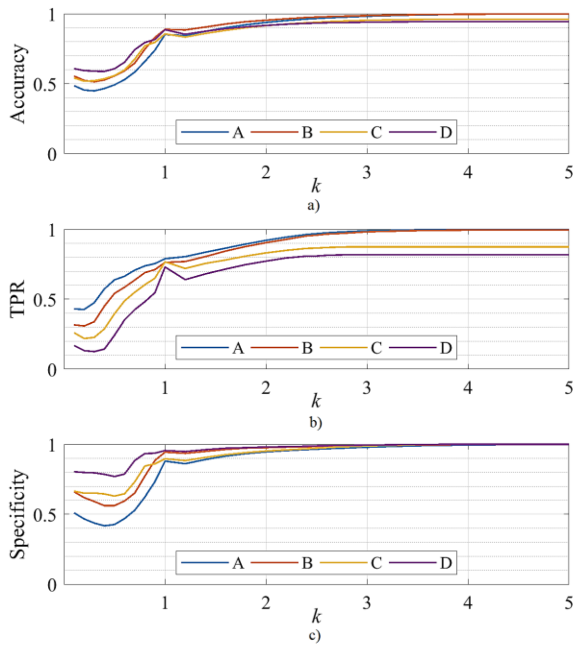

Fig. 8. *Accuracy*, a), TPR, b), and specificity, c), of algorithm A to D, for a *k* variation of EV charging load.

As expected, the increase in the EV charging load ($k > 1$), leads to both the TPR and the *specificity* increase, resulting in an overall *accuracy* increase of the classifier

performance. These results are expected, given that with an increase in the EV charging power, the dissociation of *P-Q* samples when training and testing the model is higher, resulting in better classification performance.
On the other hand, the performance drop when the *k* value decreases and stabilizes around k = 0.5. In this matter, the data clusters “EV is Charging” and “EV is Not Charging” start to overlay, with no real distinction between them, making the respective linear fitting lines and GMF curves very close together, ensuring an unfeasible meaning when used to make a prediction. However, it is interesting to note that the accuracy of algorithm D presents the highest value, higher than 50%, for k=0, while the TPR is higher for algorithm A.

CONCLUSIONS

The current work presents a methodology for EV charging identification and classification in a low-voltage distribution system. The *P-Q* curves measured at the secondary power transformer are compared with the list of charging events provided by the Mobi.e operator to characterize two clusters: one without EV charging and the other with EV charging events). The clusters present distinct behaviors described by two gaussian membership functions centered in two distinct *P-Q* lines.
Based on the probability of each point to belong to each cluster, four classification algorithms are tested: with and without temporal filters and using a binary or weighted probability. The algorithms using a temporal filter with binary or weighted probabilities present the best overall performance. This temporal filter only considers a charging event if charging is identified for more than 20 minutes. The accuracy and precision increase, while the miss classification and false-negative rates decrease when the temporal filter is used. Compared with the base algorithm without a temporal filter, the accuracy increases from 83% to 87%, the precision increases from 73% to 83%, the miss classification rate decreases from 17% to 13%, and the false-negative rate decreases from 15% to 8%.
The proposed classification algorithms are simple to implement and can be used as a first disaggregation stage of the charging EV profiles from aggregated power. With the correct calibration of the clusters and their update along the time, they can provide additional information to the Distribution System Operators to plan the integration or redistribution of EV charging stations.

Acknowledgments

This work was supported by FCT, through IDMEC, under LAETA, project UIDB/50022/2020.

REFERENCES

[1] S. A. A. Rizvi, A. Xin, A. Masood, S. Iqbal, M. U. Jan, and H. Rehman, 2018, "Electric Vehicles and their Impacts on Integration into Power Grid: A Review," *2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1-6.

[2] S. Deb, K. Kalita, P. Mahanta, 2017, "Impact of electric vehicle charging stations on reliability of distribution network," *2017 International Conference on Technological Advancements in Power and Energy (TAP Energy)*, pp. 1-6.

[3] Deb, S.; Tammi, K.; Kalita, K.; Mahanta, P., 2018, “Impact of Electric Vehicle Charging Station Load on Distribution Network.” *Energies*, 11, 178.

[4] N. Leemput, F. Geth, J. V. Roy, P. Olivella-Rosell, J. Driesen, A. Sumper, 2015, ”MV and LV Residential Grid Impact of Combined Slow and Fast Charging of Electric Vehicles.” *Energies*, vol. 8, pp. 1760-1783.

[5] V. Monteiro, H. Gonçalves and J. L. Afonso, 2011, "Impact of Electric Vehicles on power quality in a Smart Grid context," *11th International Conference on Electrical Power Quality and Utilisation*, Lisbon, pp. 1-6.

[6] M. T. Hussain, N. Sulaiman, M. S. Hussain, M. Jabir, 2021, “Optimal Management strategies to solve issues of grid having Electric Vehicles (EV): A review,” *Journal of Energy Storage*, vol. 33.

[7] X. Ren, H. Zhang, R. Hu, Y. Qiu, 2019, “Location of electric vehicle charging stations: A perspective using the grey decision-making model,” Energy, vol. 173, pp. 548-553.

[8] S. Wang, L. Du, J. Ye, and D. Zhao, 2020, "A Deep Generative Model for Non-Intrusive Identification of EV Charging Profiles," in *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 4916-4927.

[9] A. S. Al-Ogaili et al., 2019, "Review on Scheduling, Clustering, and Forecasting Strategies for Controlling Electric Vehicle Charging: Challenges and Recommendations," in *IEEE Access*, vol. 7, pp. 128353-128371.

[10] DeepGrid, “Low Voltage Network Monitoring and Optimization,” available online at: https://eneida.io/deepgrid-2/

[11] K.M. Ting, 2011, “Confusion Matrix.” *In Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning*, Springer, Boston, MA.