Aggregator Zone Selection for EV Smart Controls based-on ML Clustering of Grid Strength, Distance, and Charging Homogeneity

Rosemary E. Alden, Sam H. Lowe II, and Dan M. Ionel
SPARK Laboratory, Stanley and Karen Pigman College of Engineering, University of Kentucky, Lexington, KY, USA
rosemary.alden@uky.edu, samuel.lowe2@uky.edu, dan.ionel@ieee.org

***Abstract*—Smart electric vehicle (EV) charging control methods from a central utility hub often require communication infrastructure over a large service area of electric power distribution systems with a large number of nodes. Industry standards such as Open Charge Point Protocol (OCPP) 2.1 have evolved to include topologies for local controllers to the individual chargers, i.e. EV aggregator zones. A machine learning (ML) application of k-means clustering is proposed to establish zones for coordination of EV charging based on grid strength and EV owner decision-making to charge per day. Very large-scale distribution networks including the IEEE 123 and 8500 benchmark feeders are characterized by their distances from the substation and line reactance to resistance ratios (X/R). Then, a sensitivity study is performed with six different statistical distributions of "daily homogeneity", i.e. overlapping EV owner decision to charge that day between all homes on each line. Following increased voltage violations in scenarios with high homogeneity, a case study under the uniform statistical distribution shows how the development of the EV zone selection process based on stochastic inputs such as decision homogeneity, line X/R ratio, and house quantity may intelligently identify groups of EVs for localized controls.**

***Index Terms*—Machine learning, electric vehicle (EV), EV aggregator, power distribution system (PDS), smart grid**

I. Introduction

As distributed energy resources (DERs) and electric vehicles (EVs) become more prevalent, modeling improvements of devices and electric distribution systems are necessary for the smart grid transition [1]. Within the general topic of EV charging and DER grid impact, assessment of the electric power distribution grid strength and its influence on the voltage behavior warrants detailed study in the future the smart grid transition. Such grid strength evaluation studies have been conducted for DERs such as solar PV and wind by varying the grid strength as represented by the line reactance to resistance, X/R, ratio and evaluating voltage fluctuations [2], [3]. It has been further confirmed to influence voltage response to increased EV charging load by principal component analysis (PCA), and clusters of feeder types based on their X/R grid-strength and voltage response proposed [4]. An expansion to quantify grid impact of different portions of residential and work charging has also been proposed for clustering feeders into classes [5].

Smart controls for DERs and EVs have become a large subfield in electric grid research with a focus on the grid effects of EV charging and vehicle-to-grid operation based on stochastic human behavior and wide-spread data collection [6], [7]. While big data for EV-related human behavior exists from large sources such as the National Household Travel Survey, gaps remain regarding residential charging decisions at home including the time duration of which owners wait between charges. Within this paper, a sensitivity study was conducted with various distributions for overlap or homogeneity in EV charging decisions to assess grid impact possibilities then a machine learning (ML) method is proposed to cluster zones within a feeder for smart controls considering behavior, grid strength in X/R ratio, and number of EVs per line (Fig. 1).

II. Representative and Experimental Distribution System Modeling

Within this work two example IEEE benchmark circuits, the 123 and 8500 node test feeders were selected to represent a strong and weak grid type [8], [9]. These systems have a peak power of 3.6 and 12MW, respectively, and were designed to represent typical medium to large circuits in North America. The IEEE 123 node test feeder has been modified to include residential load profiles from a large private smart meter dataset in Glasgow, KY for 1,765 homes [10]. An EV with a charger of 5kW rated power has been assigned to each house on both circuits, and they are solved for the peak load time instance in the remaining case studies.

Minutely steady state power flow analysis was conducted in the open-source OpenDSS electric distribution system software across a simulation time window of one full day. OpenDSS calculates the power and voltages across nodes on the system based on load currents through a system admittance matrix as follows:

$$\overbrace{\begin{bmatrix} I_0 \\ I_1 \\ \vdots \\ I_N \end{bmatrix}}^{\mathbf{I}} = \overbrace{\begin{bmatrix} Y_0 + \cdots + Y_{1N} & \dots & Y_{1N} \\ \vdots & \ddots & \vdots \\ -Y_{1N} & \dots & Y_{01} + \cdots + Y_{1N} \end{bmatrix}}^{\mathbf{Y_S}} \overbrace{\begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_N \end{bmatrix}}^{\mathbf{V}}, \quad (1)$$

where $\mathbf{I}$ is a vector of node currents, Y_S the system admittance matrix based on system equipment impedance, $\mathbf{V}$ a vector of node voltages, and N the number of nodes on the distribution system. The solver uses a fixed point numerical method to solve based on an initial voltage vector, V_0, as follows:

$$\mathbf{V}_{i+1} = [\mathbf{Y}_S]^{-1} \cdot \mathbf{I} \cdot \mathbf{V}_i, i = 0, 1, 2, ..., i_c \text{ or } i_{max}, \quad (2)$$

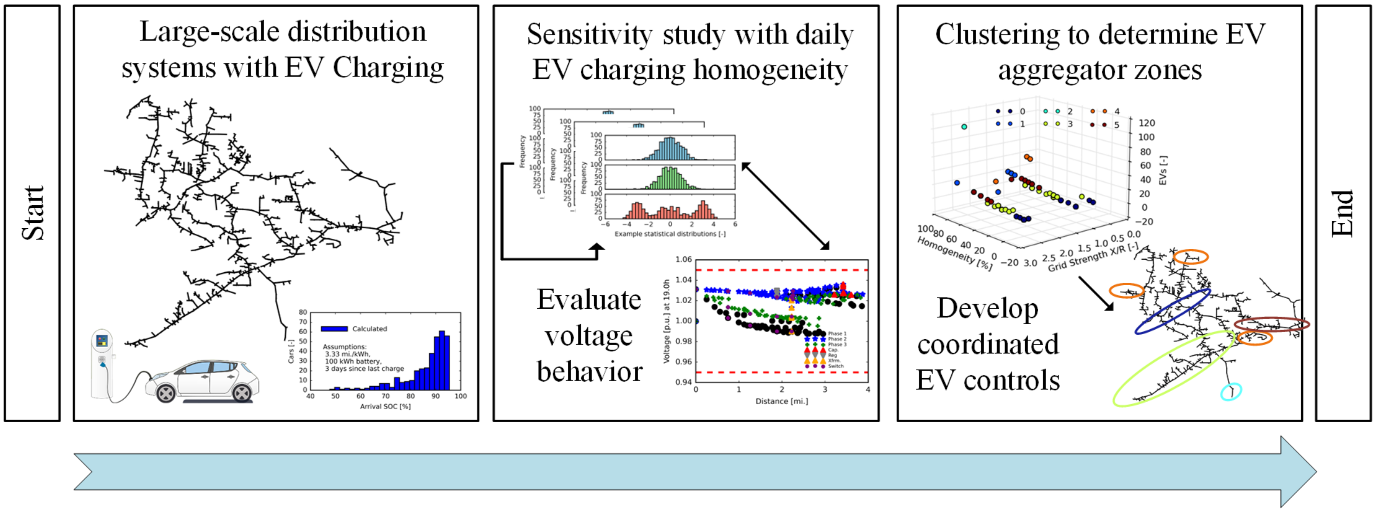

Fig. 1. Schematic of the proposed contributions to characterize large-scale distribution systems with hundreds to thousands of nodes by their grid strength, assess the impact of varied statistical distributions of EV owner homogeneity or overlap in daily charging decision, and cluster the lines into zones for use with smart controls through industry communication protocols such as OCPP.

where i_C is the iteration of convergence and i_{max} a stopping criterion.

In addition to the power flow, the grid strength, X/R, of each system line code was evaluated as follows: $\frac{X}{R} = \sum \mathbf{X} \, / \sum \mathbf{R}$, where the $\mathbf{X}$ is the reactance and $\mathbf{R}$ the resistance matrix portions of the impedance matrix, $\mathbf{Z}$, with self and mutual inductances for each multi-phase line unique to cable material and geometric configuration in the OpenDSS electric power distribution system simulator software. Distribution systems have low X/R ratios in general compared to transmission, with typical ranges between 0.5 and 3; within this range, the higher the X/R ratio, the more likely voltage fluctuations become [11]. Almost half of the IEEE 123 node test feeder lines have a high distribution X/R ratio, whereas very few of the IEEE 8500 do, making them examples of strong and weak low-voltage systems (Fig. 2).

III. Sensitivity study for EV Owner Decision Scenarios

The importance of wide-spread data collection and release to the public must be balanced with data privacy concerns, resulting in few public sources. Within the United States of America (USA), the large majority of EVs charge at home, and workplace charging was listed as the second largest subsection in a large data collection effort of Chevrolet Volt and Nissan Leaf drivers [12]. Another example of the limited EV behavioral big data released to the public shows that out of 576 customers the highest spike in owners' charging start time occurs between 4 and 9pm with durations of 1-3 hours in Omaha, NE between 2020 and 2022. This represents a high chance for overlap of EV charging in time and additional load on a line leading to a residential transformer overload in a high EV penetration scenario [13].

It was also found that the percentage of total time durations between charges was strongly skewed to between 12 and 24h. Efforts to predict the EV charging behavior based on this data, found it to be a very difficult task with low R^2 values, even with substantial input data size. An additional example of EV predictions for duration and energy used found mean absolute percent errors (MAPEs) into the hundreds further indicating the difficult nature of predicting EV charging behavior in residences [14].

Due to the high difficulty and on-going efforts to obtain comprehensive and representative charging data with reliable forecasting results across the continent and globe, an effort to represent a range of EV charging behaviors has been made in the following sensitivity study. To generate six possible future scenarios in which voltage related issues may arise, EV charging behavior has been quantified through the proposed metric "daily EV charging homogeneity", H_N. This metric is defined as the similarity or overlap of all EV owners' decision at each distribution system node containing houses to charge on the given day at a peak load time in the evening. It is calculated follow:

$$H_N = \frac{\sum \mathcal{S}_{EV}}{E_N} \text{ with } \mathcal{S}_{EV} = \{0, 1, 0, ...\} \text{ of size } E_N, \quad (3)$$

where $\mathcal{S}_{EV}$ is a set of binary statuses to indicate whether the owner decided to charge for each EV at node N and E_N the total number of EVs at node N.

Ranging from 0-100%, the daily EV charging homogeneity indicates the percentage of homes that charge during peak load on each of the load containing nodes in a distribution system feeder. Common statistical distributions were sampled and randomly assigned to the lines of the strong grid example IEEE 123 node test feeder, including uniform, Gaussian, Poisson, bimodal, and two variations of a skewed curve toward low daily EV charging homogeneity (Fig. 3). Each statistical distribution represents different scenarios for EV adoption and utilization across a neighborhood: the uniform case an even spread, the gaussian and Poisson a moderately high adoption

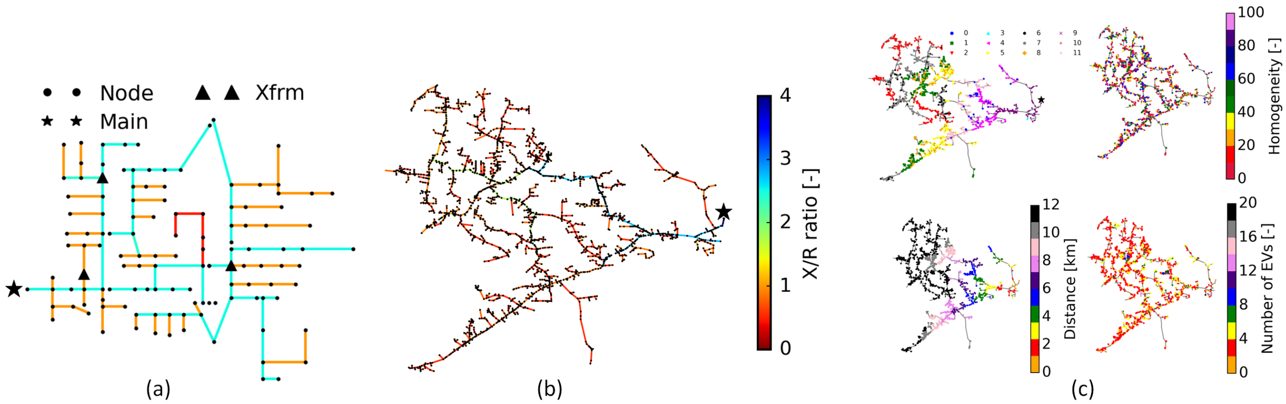

Fig. 2. The IEEE 123 (a) and 8500 (b) node test feeders with grid strength of the line reactance to resistance ratios shown by color. Regions of the circuits with similar ratios and characteristics are proposed to be grouped as EV aggregator zones to reduce signal congestion and the number of control variables for smart charging. The selected EV aggregator zones for the IEEE 8500 node system and the ML inputs proposed are visualized in (c) with distance as the most influential.

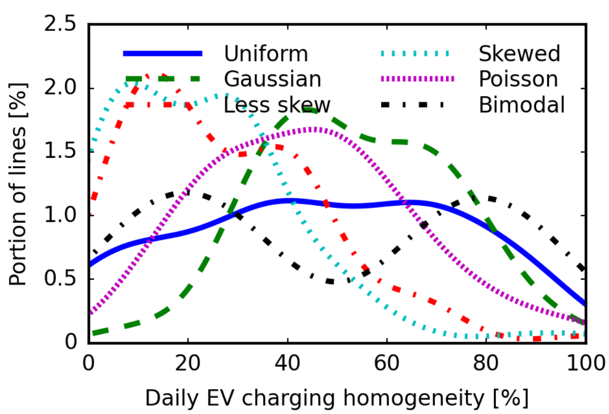

Fig. 3. Probability density functions for different scenarios displayed as percentages of the total lines on the IEEE 123 node test feeder representing continuous homogeneity in EV owner decision to charge per day.

Table I
SENSITIVITY STUDY OF EV OWNER CHARGING HOMOGENEITY DISTRIBUTIONS WITH IMPACT ON BUS VOLTAGE VIOLATION (VIO.) MEAN AND INTERQUARTILE RANGE (IQR).

Distribution sampled	EVs [%]	Peak P [kW]	Bus Vio. [1]	Mean [p.u.]	IQR [p.u.]
Gaussian	53	8728	77	0.018	0.014
Uniform	47	8165	70	0.016	0.011
Bimodal	50	8505	67	0.012	0.009
Poisson	45	7888	52	0.010	0.008
Less skewed	28	6211	22	0.005	0.011
Skewed	22	5692	8	0.005	0.001

[1] Buses with voltage violations (vio.) outside 0.95 and 1.05p.u.

and overlap in charging for the middle 60% of the lines, the bi-modal a split with both low and high adoption areas, and the skewed cases a low adoption rate where the majority of houses do not have an EV nor charge it during peak time.

In Table I, the total percentage of EVs on the network that charge, the peak substation aggregate power that occurs, the number of bus voltage violations as well as the violations mean and interquartile range (IQR) in p.u. were solved from the OpenDSS power flow simulation sensitivity study during peak load of 3.6MW at 7pm. In cases with a high number of violations, the mean p.u. bus voltage violations were fairly low with small IQRs indicating tight clustering of 50% of the distribution within less than 0.015 p.u., i.e. 15% of the acceptable p.u. range. It is important to note that in this sensitivity study, the main substation transformer rating of 5MW was exceeded in all cases, which supports the need for EV smart charging program incentives to change human behavior around charging.

The homogeneity distribution scenario impacts the number of voltage violations significantly as seen between the uniform and possion cases where a very similar percentage of EVs were charging, 47 and 45%, yet a substantially higher number of buses have voltage violations, i.e. 70 to 52 respectively. The skewed case shows the most favorable results with a lower percentage of EVs charging, 22%, and fewer violations, eight. Utilities could design charging programs to incentivize a low homogeneity and longer varied time durations between charges to mimic a heavily skewed EV daily homogeneity along with smart controls for the required grid performance of zero voltage violations.

IV. Cluster-based ML Method for OCPP and IEC 61850 Zonal Coordinated EV Controls

Utility programs to reduce EV daily homogeneity would rely on industry standards for communication from a central hub to the distributed residential EV chargers, some of which are both a great distance away and/ or densely located creating network congestion. Industry standards have responded to this problem by including the functionality of local EV controllers in the signal processing chain. Such examples include the Open Charge Point Protocol (OCPP) in the new 2.1 version [15]. Specifically, a model of IEC 81850 signals from the utility to a charge point operator (CPOs), then OCPP 2.1 to the fast, public, or at home chargers with ISO 15118 signals to the car itself has been described in a recent industry

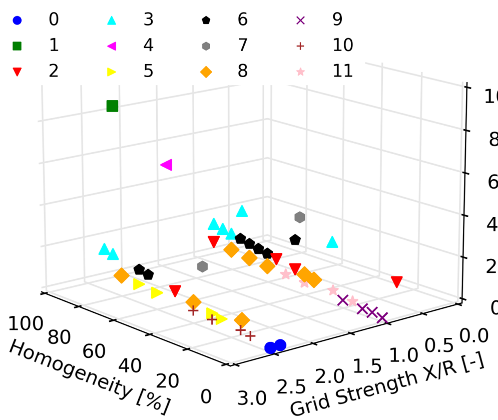

Fig. 4. Example k-means cluster results for the IEEE 123 node test feeder for twelve groups to be communicated to as EV aggregator zones. Groups with strong influence of grid strength were 0, 5, 9, 10, and 11.

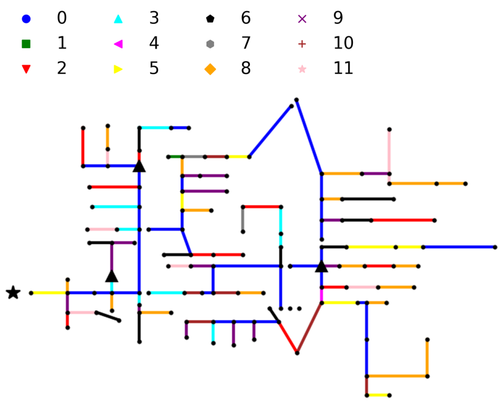

Fig. 5. Clustered EV zones selected by the k-means algorithm for controls. The node homogeneity was randomly assigned to lines irrespective of GIS coordinate, partially explaining the noncontinuous groups.

white paper [16]. With this configuration, questions on how to determine which EVs and lines should be grouped under a CPO, otherwise known in literature as an EV aggregator, or given the same control signals across different OPCs.

The application of ML unsupervised learning techniques such as the k-means clustering method to select zones for EV coordinated controls may serve as a potential solution to these highly topical and relevant questions. A case study has been conducted to assign the lines of the IEEE 123 and 8500 node test feeders to 12 zones in which similar EV coordinated controls may be beneficial to reduce voltage violations. The inputs to the k-means clustering method were homogeneity, total number of EVs connected per line, grid strength X/R ratio or distance from the substation (Fig. 6 and 7).

V. Conclusion

Large distribution systems with thousands of experimental smart meter profiles nodes have been characterized by grid strength and distance from the main substation for the purpose of an EV behavioral impact study on the severity of voltage fluctuations during high EV charging load. The sensitivity study with six scenarios based on common statistical distributions that may occur finds high influence of EV daily homogeneity, i.e. the percentage of total EVs on the line that charge during peak time, influential to the number of voltage violations that occur. Based on the study results and trends in industry communication protocols such as OCPP with IEC 61850 toward local controllers, an unsupervised ML method to cluster lines has been proposed and showed potential to group lines into EV aggregator zones to receive equivalent or similar controls on two benchmark IEEE test feeders with hundreds and thousands of nodes.

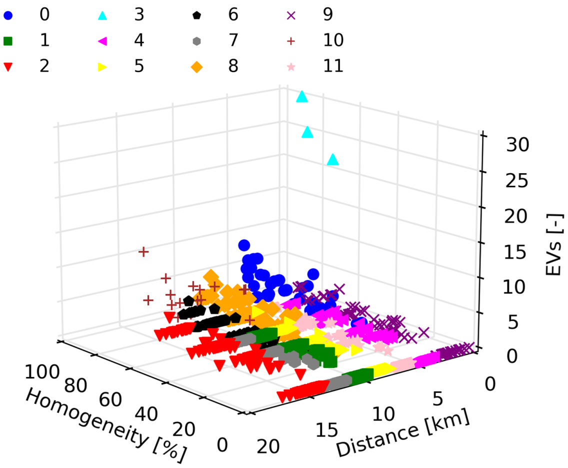

Fig. 6. The IEEE 8500 node test feeder with twelve groups based on the number of EVs on the lines, the homogeneity of the lines, and their distance from the substation. For most groups distance is the most influential with the exception of 10 and three.

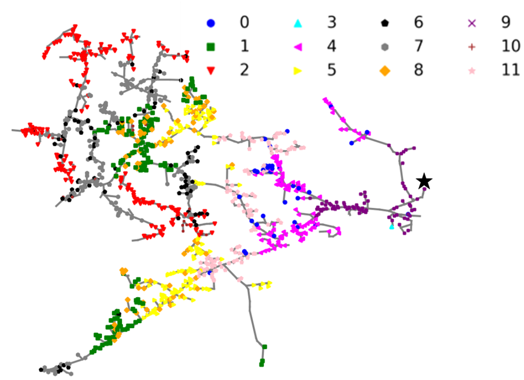

Fig. 7. The clustered EV aggregator zones were strongly influenced by distance with group 4, 9, and 0 dominated by a short distance from the substation. The groups 2, 6, and 7 have a wide range of distances.

Acknowledgment

This paper is based upon work supported by the National Science Foundation (NSF) under NSF Graduate Research Fellowship Grant No. 2239063. The support of the Lighthouse Beacon Foundation and L. Stanley Pigman Chair in Power Endowment at the University of Kentucky are also gratefully acknowledged. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

References

[1] F. Blaabjerg and D. M. Ionel, *Renewable energy devices and systems with simulations in MATLAB® and ANSYS®*. Boca Raton, FL: CRC Press, 2017.

[2] M. Pramanik, U. Roy, K. D. Bhattacharya, and C. K. Chanda, "Effects of X/R on the power dynamics of a rural distribution system and the economic implications," in *2022 IEEE Calcutta Conference (CALCON)*, 2022, pp. 313–317.

[3] S. M. Alizadeh, C. Ozansoy, and T. Alpcan, "The impact of X/R ratio on voltage stability in a distribution network penetrated by wind farms," in *2016 Australasian Universities Power Engineering Conference (AUPEC)*, 2016, pp. 1–6.

[4] C. B. Jones, M. Lave, and R. Darbali-Zamora, "Overall capacity assessment of distribution feeders with different electric vehicle adoptions," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.

[5] C. B. Jones, M. Lave, W. Vining, and B. M. Garcia, "Uncontrolled electric vehicle charging impacts on distribution electric power systems with primarily residential, commercial or industrial loads," *Energies*, vol. 14, no. 6, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/6/1688

[6] S. B. Poore, R. E. Alden, and D. M. Ionel, "Coordinated controls of residential EV chargers considering high power appliances," in *2024 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2024, pp. 1–5.

[7] R. E. Alden, A. Timilsina, S. Silvestri, and D. M. Ionel, "V2G optimization for dispatchable residential load operation and minimal utility cost," in *2023 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2023, pp. 1–4.

[8] K. P. Schneider, B. A. Mather, B. C. Pal, C.-W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W. Kersting, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3181–3188, 2018.

[9] R. F. Arritt and R. C. Dugan, "The IEEE 8500-node test feeder," in *IEEE PES T&D 2010*, 2010, pp. 1–6.

[10] H. Gong, V. Rallabandi, M. L. McIntyre, E. Hossain, and D. M. Ionel, "Peak reduction and long term load forecasting for large residential communities including smart homes with energy storage," *IEEE Access*, vol. 9, pp. 19 345–19 355, 2021.

[11] R. Walling, "Distribution system characteristics, distribution practices, and relevance to DER and composite load modeling," 2023.

[12] J. G. Smart and S. D. Salisbury, "Plugged in: How americans charge their electric vehicles," 7 2015. [Online]. Available: https://www.osti.gov/biblio/1369632

[13] A. Almaghrebi, K. James, F. Al Juheshi, and M. Alahmad, "Insights into household electric vehicle charging behavior: Analysis and predictive modeling," *Energies*, vol. 17, no. 4, 2024. [Online]. Available: https://www.mdpi.com/1996-1073/17/4/925

[14] M. Kreft, T. Brudermueller, E. Fleisch, and T. Staake, "Predictability of electric vehicle charging: Explaining extensive user behavior-specific heterogeneity," *Applied Energy*, vol. 370, p. 123544, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261924009279

[15] "What is new in OCPP 2.0.1," Open Charge Alliance, Tech. Rep. v1.0, 2023.

[16] "White paper: OCPP & IEC 61850: a winning team," DNV. Open Charge Alliance, Tech. Rep. 23-3107, August 2023. [Online]. Available: https://openchargealliance.org/wp-content/uploads/2024/01/Whitepaper-OCPP-IEC-61850-a-winning-team-Report-No.-23-3107-08-09-2023-Version.-1.0.pdf