Data-driven, Internet-inspired, and Scalable EV Charging for Power Distribution Grid

Emin Ucer*, Mithat Kisacikoglu†, Murat Yuksel‡, and Ali C. Gurbuz†
* Dept. of Electrical and Computer Engineering, University of Central Florida, Orlando, FL
† Dept. of Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL
‡ Dept. of Electrical and Computer Engineering, Mississippi State University, Starkville, MS
E-mails: eucer@crimson.ua.edu, mkisacik@ua.edu, murat.yuksel@ucf.edu, gurbuz@ece.msstate.edu

Abstract—Electric vehicles (EVs) are finally making their way onto the roads. However, the challenges concerning their long charging times and their impact on congestion of the power distribution grid are still waiting to be resolved. With historical measurement data, EV chargers can take better-informed actions while staying mostly off-line. Proposed solutions that depend on heavy communication and rigorous computation for optimal operation are not scalable. The solutions that do not depend on power distribution topology information, such as Droop control, are more practical as they only use local measurements. However, they result in sub-optimal operation due to a lack of a feedback mechanism. This study develops a distributed and data-driven congestion detection methodology embedded in the Additive Increase Multiplicative Decrease (AIMD) algorithm to control mass EV charging in a distribution grid. The proposed distributed AIMD algorithm performs very closely to the ideal AIMD regarding fairness and congestion handling. Its communication need is almost as low as the Droop control. The results can provide crucial insights on how we can use data to reveal the inner dynamics and structure of the power grid and help develop more advanced data-driven algorithms for grid-integrated power electronics control.

Index Terms—Decentralized control; data-driven control; AIMD; smart charging; grid integration; EVs.

I. INTRODUCTION

HIGH penetration of electric vehicles (EVs) with uncontrolled charging will cause transformer and/or line congestion in the power distribution grid. Among some adverse effects of this congestion are severe voltage drops, increased peak loading, thermal overheating, and even failure of equipment [1]–[6]. Therefore, control of EV charging has become an important research effort to mitigate the impacts. Demand-side load management is a technique that is used to modify customer demand through various tools and methods. Curtailing this demand at peak hours is known as peak-shaving. It is used to eliminate short-term demand spikes by lowering and smoothing out peak loads; this prevents equipment overloading. In this study, we will investigate a peak-shaving methodology to reduce the substation peak loading caused by massive EV integration.

Commonly proposed methods for EV charging control require excessive system information, e.g., the grid topology, load forecasting, and customer preferences [7]–[11]. An uninterrupted connectivity is also required to communicate these information to the EV chargers and send/receive control commands to/from the chargers. Assuming these are available, various optimization problems have been formulated to achieve certain objectives (e.g., minimizing generation cost, losses and peak load, maximizing capacity utilization and total charging power) while respecting system constraints such as voltage limits and equipment overloading. Richardson et al. use linear programming to maximize the total EV charging power while operating within the grid network’s limits [7]. Sojoudi and Low solve an optimal power flow problem to minimize the generation and charging costs while satisfying the network constraints [8]. Restrepo et al. propose a three-stage control algorithm by solving an optimization problem in each stage to coordinate EV load under volt-var control devices [9]. Zhang et al. formulate the network-constrained EV charging control as a convex problem based on a linearized distribution grid model [10]. Liu et al. propose a valley-filling solution for EV charging as an optimization problem [11].

A traditional approach is to form an optimal power flow (OPF) problem and deterministically solve it using forecasted generation and demand values to determine the control set points. Recently, machine learning (ML) is also used to develop control algorithms that can evolve behaviors based on OPF simulations [12]. The idea is to learn local controller rules from a dataset generated by OPF solutions. In general, performing these optimization tasks is very challenging since the required system model and inputs are either missing or very hard to obtain completely. In [13] and [14], data-driven local controllers are proposed by solving an off-line OPF problem using ML regression models. Since these controllers are trained based on a historical and pre-defined dataset, they are vulnerable to uncertainties not covered in the dataset, and instances caused by policy and natural shifts [15].

A secondary low-voltage (LV) network is a part of power distribution grid which carries electric energy from distribution transformers to service meters of end customers. Grid topology, especially in LV distribution grids, is often unknown and not available [16]. Besides, the topology information is usually limited to primary network. The lengths (impedance) of the underground cables connected between houses and distribution feeder are often not available. Moreover, the information regarding phase connections are also often unknown [16]. Previously mentioned solutions heavily depend on this infor-
mation to perform as desired.

On the other hand, model-free approaches are generic and not dependent on the OPF model, i.e., system topology and variables. They usually use measurements (voltage, frequency, etc.) or historical data to obtain information regarding the grid’s dynamic behavior. In this line of work, in a previous study, we presented a decentralized EV charging solution that uses the local voltage measurements and adapts the Internet’s TCP/IP protocol for congestion detection [17]. Xia et al. proposed decentralized EV charging solutions using local and historical measurements [18], [19]. Geth et al. examine the adaptation of droop curve control for EV load management and their impacts on voltages [20]. Although these solutions are communication free, simpler, and faster, they often try to implicitly detect the peak loading of the grid via measurements and predefined set points and therefore result in sub-optimal control performance.

The Internet uses the Additive Increase and Multiplicative Decrease (AIMD) algorithm for its own congestion management. The only needed information for AIMD is a binary (yes/no) feedback that signifies the congestion event (CE) [21]. In the context of distribution grid, the CE corresponds to the overloading of substation transformer feeder that powers the distribution network. The CE is triggered when the substation is loaded over its capacity. Note that the CE can also be triggered when there is a need to shave the peak loading in a demand side load management scenario. Therefore, AIMD can be used to flatten the substation’s load.

The problem with the AIMD algorithm for EV charging is that its ideal implementation in the distribution grid requires explicit feedback of the CE (substation loading). The studies presented in [22]–[24] adapted AIMD for EV charging by using an explicit feedback signal of the CE broadcast to them in real-time. This, however, necessitates real-time connectivity for ideal operation. On the other hand, the Internet’s current TCP/IP protocol implements the AIMD algorithm based on local measurements, implicitly detecting the congestion without requiring explicit feedback signal or system knowledge. Our previous works [17], [25]–[27] and other studies [28], [29] investigated implicit ways of estimating the CE in the distribution grid with only local voltage measurements. However, they mostly stay limited and sub-optimal due to a lack of a feedback mechanism.

This study explores how local historical measurement data can be utilized to learn a machine learning (ML) model that can predict the grid’s congestion level using local voltage measurements in real-time. We integrate this model into an AIMD-based EV charging control algorithm to improve distributed EV charging control performance. Our methodology is to learn an ML model that maps local voltage to feeder loading and use it as feedback for the AIMD controller to predict the substation load in real-time using local voltage measurements without requiring connectivity to a central controller. Hence, our approach differs from other studies because we do not require system topology and loading information for operation (model-free). We do not need a real-time feedback signal (reduced communication). We also do not learn rules for optimum controller actions using ML.

The proposed solution does not limit the total number of nodes. Therefore, it is scalable and can easily scale up to as many nodes as possible since the algorithm’s operation is independent of each node. The only limit we have is the requirement to occasionally get the substation loading information to train the end-node ML models against system changes/updates. The model parameter variances may also increase as the number of end-nodes increases in a larger network.

This study aims to develop a data-driven, local feedback mechanism for the AIMD controller that can estimate the CE offline. This way, our controller can operate more in a decentralized fashion and take advantage of its benefits. We will also provide a comparative study to test our proposed solution and compare it with two other solutions to evaluate the performance better.

Our main contributions are two folds:

- We devised an EV charging management solution that can work in a decentralized manner using solely end-node voltage without needing any real-time connectivity to receive charging commands.
- The proposed solution does not require the grid topology or to forecast the load profiles for solving an optimization problem. It uses historical end-node voltage and grid power measurements (i.e., data-driven) to determine critical voltage thresholds for charging control.

The rest of the paper is structured as follows: In Section-II, the AIMD algorithm is introduced and analyzed. Section-III investigates the relationship between the grid voltage and the demand power. A counterpart AIMD charging algorithm for EV charging is proposed in Section-IV. Section-V presents performance metrics and two published benchmark studies for comparison with the proposed solution. Section-VI presents the results, and Section-VII concludes the paper.

II. THE AIMD ALGORITHM

Today’s Internet owes its standards and protocols to years of debates and research around solving the congestion problem. Its early days suffered congestion control challenges as the number of endpoints drastically increased [30]. Observed congestion collapses [31], [32] revealed the need for a solution that assures both the stability of the system and the fair and efficient utilization of the network capacity. To that end, The AIMD algorithm [21] was proposed as a congestion avoidance solution at the endpoints. A detailed mathematical analysis of the AIMD algorithm was presented in a previous study [27].

AIMD takes actions each time a CE occurs on the network. However, there is no central authority on the Internet that detects CEs and regularly transfers a binary CE signal to end-nodes. Instead, the Internet’s existing transmission control protocol (TCP) uses a decentralized and indirect way to determine the degree of the congestion. Each end-node infers the congestion level by measuring how long it takes to get the acknowledgement of a transmitted data packet from its destination, i.e., round-trip time (RTT). This statistical learning of RTT provides some partial information of the congestion that might exist in the shared network paths. We
presented a detailed work on the explanation of this indirect approach and its adaptation for EV charging problem in [17]. Although this indirect approach stays sub-optimal, it has been quite successful. It was proved by Low and Lapsley that this straightforward solution can maintain stable operation and proportional fairness among end-nodes [33]. From the standpoint of allocating a limited resource among competing participants in a distributed fashion, AIMD stands out as a strong candidate. It can be modified into various fields with similar problems. Corless et al. provides a detailed mathematical analysis and modeling for AIMD and also discusses possible areas of applications [34].

The AIMD algorithm has two operational phases. The additive increase (AI) phase takes part when there is available capacity in the system. In this phase, agents are allowed to increase their shares linearly by a rate \( \alpha > 0 \). In the case of CE, the algorithm switches to the multiplicative decrease (MD) phase where agents scale down their shares by a factor \( 0 < \beta < 1 \). The capacity share of an agent at time \( t + 1 \) can be formulated as follows:

\[
\begin{align*}
  w_i(t+1) &= \begin{cases} 
    w_i(t) + \alpha_i & \text{if there is no congestion} \\
    w_i(t) \times \beta_i & \text{if congestion occurs}
  \end{cases} \\
\end{align*}
\]

where \( w_i \) denotes the share of the agent \( i \). (1) can be modified for a feeder-level peak-shaving problem for EV charging in the distribution grid as:

\[
\begin{align*}
  I_i(t+1) &= \begin{cases} 
    I_i(t) + \alpha_i & \text{if peak-shaving is not needed} \\
    I_i(t) \times \beta_i & \text{if peak-shaving is needed}
  \end{cases} \\
\end{align*}
\]

where \( I_i \) is the charging current of the \( i^{th} \) EV. CE here refers to overloading of some distribution grid assets, primarily the substation feeder. The effectiveness of AIMD to avoid a CE depends on how accurately the congestion or loading level is estimated by users.

Similar to the Internet case, there will be no central authority in our AIMD implementation that will notify end-nodes of CEs. Our proposed method will enable end-nodes to implicitly detect congestion based on their local voltage measurements. The proposed approach differs from our previous work [17] in that we develop a supervised ML model to map local voltage measurements to feeder loading level using historical data. Thus, we can improve the congestion detection accuracy and get closer to the optimum operation and more efficient capacity utilization. In the following section, we will investigate how the voltage of the network can be used to estimate the loading level implicitly.

III. ANALYSIS OF VOLTAGE VS. POWER RELATIONSHIP

The AIMD algorithm is based on detecting network congestion, which means the network infrastructure is close to being overloaded or it needs to reduce its peak loading. In a distribution system, the substation transformer could be a potential congestion point since it supplies power to the entire downstream network. Notifying the end-nodes of the substation congestion can be performed by a direct communication link, however, this requires an uninterrupted real-time connectivity. It is much less demanding in terms of communication needs if the local end-nodes can detect this congestion. In this regard, local voltage measurements can be an indicator for this congestion information. The voltage varies depending on the loading level on the grid. Therefore, each node can make an estimation of the substation loading by observing their voltage if they have a guideline that maps their voltage level to the total power drawn from the substation. This guideline can be referred as the voltage and power relationship, and it is investigated in various works [35–37].

We demonstrate this relationship between node voltage and total demand power on a radial distribution grid model shown in Fig. 1. In this model, \( V_o \) represents the substation voltage magnitude, \( \{V_1, \ldots, V_n\} \) are magnitudes of node voltages at nodes \( 1, \ldots, n \), and \( S^0_L = P^0_L + jQ^0_L \) (subscript \( L \) stands for line) denotes the complex power flowing from node \( i \) to node \( i+1 \) over a line impedance of \( Z^i_r = r_i + jx_i \). \( S^i_N = P^i_N + jQ^i_N \) (subscript \( N \) stands for node) is the complex power drawn from node \( i \). This model can be solved for an end-node voltage \( V_i \) using the distribution grid power flow (DistFlow) [38] equations (3) shown below:

\[
\begin{align*}
  P^i_L &= \frac{P^i_L + Q^i_L}{V_i^2} - P^i_{N} + r^i_i, \\
  Q^i_L &= \frac{P^i_L + Q^i_L}{V_i^2} - Q^i_{N}, \\
  V_{i+1} &= V_i - 2(r_i P^i_{L} + x_i Q^i_{L})/(r_i^2 + x_i^2 + 1), \\
  P^i_{L} &= P^i_{L}-r^i_i + r^i_i P^i_{L} + Q^i_{L} Q^i_{L}, \\
  Q^i_{L} &= Q^i_{L} - r^i_i, \\
  V_i^2 &= 2(r^{i+1}_i + x^{i+1}_i Q^i_{L}) + (r^{i+1}_i + x^{i+1}_i), \\
  P^i_{L} &= P^i_{L}-r^{i+1}_i + r^{i+1}_i P^i_{L} + Q^i_{L} Q^i_{L}, \\
\end{align*}
\]

By further simplifying the equations using the LinDistFlow approximation [38], we can derive an expression for the voltage of \( i^{th} \) node by recursively solving (3)

\[
V_{i+1} = V_i - 2(r^{i+1}_i P^i_{L} + x^{i+1}_i Q^i_{L})
\]

where \( P^i_{L} = P^i_{L-1} - P^i_{N} \) and \( Q^i_{L} = Q^i_{L-1} - Q^i_{N} \). Then the final relationship between end-node voltage \( V_i \) and total feeder active and reactive powers \( (P^0_L \text{ and } Q^0_L) \) can be formulated as

\[
V_i^2 = V_o^2 - 2P^0_L (R - \text{P}) - 2Q^0_L (X - \text{Q})
\]

where \( R = \sum_{k=1}^i r_k, \ X = \sum_{k=1}^i x_k, \ P = \phi^P U \phi^T, \ Q = \phi^Q U \psi^T, \ \phi^P = [\phi^P_1 \phi^P_2 \cdots \phi^P_{i-1} \phi^P_i], \ \phi^Q = [\phi^Q_1 \phi^Q_2 \cdots \phi^Q_{i-1} \phi^Q_i], \ \phi^P = [r^P_1 \cdots r^P_{i-1}], \ \phi^Q = [r^Q_1 \cdots r^Q_{i-1}], \ r^T = [r_2 r_3 \cdots r_i], \ x^T = [x_2 x_3 \cdots x_i], \ \text{and } U \text{ is an upper triangular matrix of size } (i-1) \times (i-1) \text{ with all the nonzero elements equal to one.} \ \phi^P \text{ and } \phi^Q \text{ are power consumption ratios that define the percentage of the active and reactive power drawn by each node to the total feeder power } (P^0_L/\phi^P \text{ and } Q^0_L/\phi^Q), \ \text{and they are stochastic variables by nature. (5) reveals that the end-node voltage magnitude square}
(V_i^2) is linearly proportional to total feeder power consumption (P_L^0 and Q_L^0). Utilizing the historical local voltage data (V_i) and total feeder power data (P_L^0 and Q_L^0), the problem of estimating grid loading level turns into a supervised learning problem. This is equivalent to mapping local voltage through a function f to substation powers such that f(V_i) → P_L^0, Q_L^0. Our goal is to learn this mapping function f(V_i) from historical data. One problem is that V_i is coupled with a linear combination of P_L^0 and Q_L^0 through (5), not each of them individually. This requires two mappings from V_i to both P_L^0 and Q_L^0. However, in practice, total reactive power consumption Q_L^0 in a distribution power grid is relatively much smaller compared to active power consumption P_L^0, and thus total apparent power S_L^0 = √(P_L^0^2 + Q_L^0^2) can be assumed to be very close to P_L^0. Therefore, we can construct the mapping from V_i to S_L^0 such that f(V_i) → S_L^0. This assumption reduces the mapping output to a single variable (S_L^0).

However, there are several factors that disturb this relationship in (5) and its parameters. First, V_i will have some variations throughout the day even though it is usually highly regulated at the feeder level. Second, there are a number of voltage regulation devices such as load tap changers (LTCs), voltage regulators (VRs), and capacitor banks operating at different points in the network. They can be very effective in changing the end-node voltage and thus affect the relationship. Third, distribution grid topology can change due to various reasons including system reconfiguration and expansion. This results in changes in r^T and x^T. Fourth, there are also stochastic variables (ϕ^P and ϕ^Q vectors) in (5) that change both in space and time. Fifth, significant reactive power consumption and generation will also deteriorate the relationship impacting the parameters. All these factors make the problem very challenging and therefore require a rigorous analysis. However, once the modeling and the relationship in (5) are resolved, this can truly unlock the ability of end-nodes to probe the grid on their own using local measurements. At this point, high function approximation capacity of ML becomes important. Data is the key to extract these hidden structures within the physical system and their associated time dependencies with the aid of ML.

IV. AIMD-BASED EV CHARGING ALGORITHM
The MD phase of the AIMD algorithm is triggered by a CE, where the shared capacity reaches its limit. In an ideal implementation of AIMD, every user is notified of CE, and they execute the MD phase of the algorithm reducing the total load simultaneously. Broadcasting CE information to every user requires constant and uninterrupted network connectivity, which increases the communication overhead and makes the system more vulnerable to cyber-attacks. A key goal in this study is to reduce the need for connectivity and operate the system without any information exchange as long as possible.

Following Section III, via providing substation loading information to end-users, every user can learn a voltage-substation loading relationship by building an ML regression model locally. For this reason, the historical substation loading data is made to be accessible to every end-node via a communication network, and end-nodes can download this data only when they need to train their ML model. We note that this is not a real-time feedback signal for our controller. Rather, it is historical power measurements collected at the substation feeder level. The transfer of this data to end-nodes is a unidirectional data flow from the substation to end-nodes, and it is not meant to be real-time. In fact, this information flow can take place as few as once per week or per month if the grid parameter distributions in (5) do not shift much. The infrequent and unidirectional information flow of the proposed method reduces the communication dependency of the proposed method making it less vulnerable to cybersecurity issues. This is illustrated in Fig. 2. End-users train their ML models where the inputs to the ML model are time-series local voltage measurements (V_i) and the outputs (labels) are time-series total substation feeder apparent power measurements (S_L^0).

Since (5) is a quadratic equation, we can use a second order model for our ML regression network to extract the voltage and total apparent power relationship based on our findings in Section III. This model defines a second order polynomial that can be fit to the data by:

\[
S_L^0(t) = \theta_{i,1} + \theta_{i,2}V_i(t) + \theta_{i,3}V_i^2(t) \tag{6}
\]

where the subscript i denotes the i\textsuperscript{th} end-node, and \(\theta_{i,1}, \theta_{i,2}, \text{and } \theta_{i,3}\) are the polynomial coefficients for the i\textsuperscript{th} node.

We will utilize the time-series end-node voltage measurements (V_i=[V_i(1) \cdots V_i(M)]^T) and time-series substation apparent power measurements (S_L^0=[S_L^0(1) \cdots S_L^0(M)]^T) to estimate the coefficient vector \(\theta_i = [\theta_{i,1}, \theta_{i,2}, \theta_{i,3}]^T\) for the duration [1 \ldots M]. We are investigating the dependencies of two variables (voltage at node i and total apparent power at the substation) and aim to learn a second-order model to build the relationship between the two. In other words, our model has one independent variable, which is the end-node voltage \(V_i\), and one dependent variable, which is the total substation power \(S_L^0\). Therefore, our model is a bi-variate model even though the actual power grid is a multi-variate system with many dependencies. This can reduce the ability to capture variations in other system parameters, and thus be considered the weakest point of the model. However, our approach assumes that the variations in the other system parameters are insignificant, which is likely valid especially in the short term. We also plan to include more variables (features) in our model to make it more robust to disturbances in a future study.
The three coefficients of the model can be estimated using the least squares regression (LSR). We define the following LSR formulation.

\[ S_L^0 = \Lambda_i \theta_i \]  

(7)

where \( S_L^0 \) is the substation apparent power measurements, \( \theta_i \) is the coefficient vector, and \( \Lambda_i \) is an \( M \times 3 \) matrix that is made of voltage measurements of the \( i \)th node and defined as follows

\[
\Lambda_i = \begin{bmatrix}
1 & V_i(1) & V_i^2(1) \\
1 & V_i(2) & V_i^2(2) \\
\vdots & \vdots & \vdots \\
1 & V_i(M) & V_i^2(M)
\end{bmatrix}
\]  

(8)

The estimate of parameters \( \hat{\theta}_i \) can now be calculated using the closed form solution of the least squares problem as in (9):

\[ \hat{\theta}_i = (\Lambda_i^T \Lambda_i)^{-1} \Lambda_i^T S_L^0 \]  

(9)

**Algorithm 1** AIMD algorithm for EV charging network

**Input:** Substation rated capacity: \( S_{Lr}^0 \)

**Input:** Voltage meas.: \( V_i = [V_i(1) \cdots V_i(M)] \)

**Input:** Substation power meas.: \( S_L^0 = [S_L^0(1) \cdots S_L^0(M)] \)

**Compute:** \( \theta_i \) using (9)

**Compute:** \( V_{i,th} \) using (10)

**Parameter:** Additive parameter: \( \alpha_i = 1 \)

**Parameter:** Multiplicative parameter: \( \beta_i = 0.5 \)

**Input:** Previous charging current: \( I_i(t) \)

**Input:** Node voltage: \( V_i(t) \)

**Output:** New charging current: \( I_i(t+1) \)

1: while SOC < 100% do
2: if \( V_i(t) > V_{i,th} \) and \( V_i(t) > V_{min} \) then
3: \( I_i(t+1) = I_i(t) + \alpha_i \)
4: else
5: \( I_i(t+1) = \beta_i \times I_i(t) \)
6: end if
7: end while

As we will demonstrate in the next section, the coefficient of the second order term usually turns out close to zero suggesting that a first order model could have also be used. The reasons a second-order method was preferred in this study are because (1) we tried to build a more generic model that complies with the derivations in Section-III regardless of the nominal voltage range, and (2) there is no additional significant cost of using a second order model in terms of computation. Besides, we also tested a linear regression (LR) model in one of our previous work [37].

With this ML model, end-users can estimate how much the substation is loaded by only using their local voltage levels. Similarly, users can calculate a particular threshold voltage value (\( V_{i,th} \)) by solving (6) for voltage given the estimated polynomial coefficients (\( \hat{\theta}_i \)) and substation rated capacity (\( S_{Lr}^0 \)) such that

\[ S_{Lr}^0 - \hat{\theta}_{i,1} - \hat{\theta}_{i,2} V_{i,th} - \hat{\theta}_{i,3} V_{i,th}^2 = 0. \]  

(10)

The MD phase of AIMD is triggered if the measured voltage \( V_i(t) \) drops below \( V_{i,th} \). If \( V_i(t) \) happens to be less than the minimum utilization voltage \( V_{min} \), then MD should also trigger again.}

The proposed charging algorithm implementing AIMD control using ML-based learned thresholds is presented in Alg. 1. The algorithm checks for any CE by monitoring the end-node voltage (\( V_i(t) \)) and comparing it with the threshold voltage (\( V_{i,th} \)) and minimum voltage (\( V_{min} \)). If the end-node voltage drops below either of these, it triggers the MD phase of AIMD. The period of CE check is defined as the algorithm period \( T_a \) (not shown in Alg. 1), and it decides how often the algorithm checks for the triggering conditions. We set it to 10 s considering the power distribution grid and EV charger dynamics. A faster response would cause a more oscillatory behaviour. We assume that the voltage is being recorded at high granularity at end-nodes in EV supply equipment (EVSE). The substation power data, on the other hand, does not have to be as granular. The algorithm needs connectivity only when it fetches the substation data. If we assume that the grid topology or rated capacity does not frequently change, the algorithm can autonomously operate for a long period of duration without exchanging any information. Every node can preserve its voltage-total power characteristics and continue using it until the grid is subject to a significant change/update in terms of supply capacity, loading level, or physical architecture. Distribution shift is also another problem of data-driven methods, which is discussed in detail in [15]. When there is a distribution shift in the stochastic parameters (e.g., EV charging needs, end-node power profiles, substation voltage, etc.) due to changes in seasons or other important events such as the one we experienced in COVID-19 pandemic, the algorithm will need to understand the changes via more frequently collecting data from the substation and comparing them to the past observations. Each learning phase will update the model parameters to the latest changes. Even though the data distribution is likely to change, the model itself is highly shaped by the topology parameters, which may not change so dramatically in the short term. The frequency of data exchange greatly depends on these changes and dynamics. End-nodes can decide themselves when they need to re-learn their new voltage-power models by taking longer timescale measurements and making comparisons, which are not studied here.

**V. COMPARATIVE PERFORMANCE EVALUATION**

**A. Description of the Test System**

The simulations are performed in IEEE-37 test distribution grid illustrated in Fig. 3. Each red circle in the figure represents a neighborhood that is connected to the primary network. There are a total of 26 neighborhoods each of which is modeled as a secondary network following a similar procedure described in [39]. Each neighborhood contains four 25 kVA transformers (26 × 4 = 104 transformers in total) powering four inner nodes. The transformers step down the primary feeder voltage of 4.8 kV to a secondary voltage level of split-phase 120/240 V. Each inner node consists of 4 houses making a total of 16 houses in a neighborhood, and a total of 416 houses in the overall distribution model. Each house is modeled as two separate nodes; one node is dedicated
only to EV connection and the other node is reserved for the uncontrollable household load. Hence, there are 416 × 2 = 832 separate nodes in the distribution network. Each end-node represents a residential customer that has a unique household power consumption profile and an EV that is modeled separately. The capacity of the substation transformer is rated at 2.5 MVA, and the grid operates slightly over 1.36 MVA at peak hours without any EV charging event, i.e. the base load.

For modeling residential loads, we developed a random household load generator. The load profile of each house is generated using 16-day of power consumption data of a real household collected using e-Gauge [40]. A new power consumption profile is generated using 16-day of power consumption data of a household load generator. The load profile of each house is typically allowed to change over a small range (0.9–1 pu).

### B. Performance Score Description

In addition to the well-known metrics of capacity utilization and fairness, we define five more performance scores for fair comparison with the available literature. These scores are

1. Voltage violation, $\text{VVS} = \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{T} \max[0, V_{\text{min}} - V_i(t)] dt \quad (11)$
2. Local Congestion Score ($\text{LCS}$)
3. Global Congestion Score ($\text{GCS}$)
4. Fairness
5. Communication overhead

Voltage Violation Score ($\text{VVS}$) is calculated for each algorithm using (11) in V-s as the time integral whose upper bound is the minimum allowed voltage level ($V_{\text{min}} = 216$ V), and the lower bound is the measured node-voltage ($V_i(t)$),

\[ \text{VVS} = \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{T} \max[0, V_{\text{min}} - V_i(t)] dt \quad (11) \]

where $N$ is the number of nodes, and $T$ is the simulation time horizon, which is 8 hours. If the node voltage never drops below $V_{\text{min}}$, then, the score will be zero.

Global Congestion Score ($\text{GCS}$) is the energy in MVAh consumed beyond the rated substation capacity $S_{\text{rated}}$ for the global congestion. If the power drawn from the substation at time $t$ is denoted by $S_i(t)$, then the GCS can be calculated as in (12).

\[ \text{GCS} = \int_{0}^{T} \max[0, S_i(t) - S_{\text{rated}}] dt \quad (12) \]

If the active power drawn from the substation always stays below the substation’s capacity, GCS will be zero.

Local Congestion Score ($\text{LCS}$) is the energy in kVAh consumed over the local distribution transformers’ capacity limit. Since there are 104 local transformers in the grid model, the LCS will be presented as an average value for the entire grid and for each 26 neighborhoods. In general, for a distribution system with $K$ distribution transformers with capacity $S_{\text{rated}}$ and apparent power $S_{\text{rated}}$ for the $i^{th}$ transformer ($i \in \{1, ..., K\}$), LCS can be evaluated as:

\[ \text{LCS} = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \sum_{t=0}^{T} \max[0, S_i(t) - S_{\text{rated}}] \right) \right) / K \]
charged at the maximum rated power of 10 kW. is linearly changed with voltage when the voltage is between when the voltage goes below 0.9 pu (216 V). Charging power droop function presented in [20] for the nominal charging functions are pre-defined. In this study, we adapt the LM2 type propose to use droop control where the power-voltage droop 1 A RMS in each AI phase is assumed. In [20], the authors α AIMD algorithm and therefore will serve as a benchmark. real time. This algorithm is the ideal implementation of the presented in [20] and [22]. [22] proposes an AIMD-based algorithm, we run the test scenario with two other algorithms implement AIMD.

In order to evaluate the performance of the proposed algorithm, we run the test scenario with two other algorithms presented in [20] and [22]. [22] proposes an AIMD-based charging control assuming that the capacity CE is notified to each user over a unidirectional communication link in real time. This algorithm is the ideal implementation of the AIMD algorithm and therefore will serve as a benchmark. α and β parameters presented in Alg. 1 are set to 1 and 0.5, respectively. These values are chosen specifically based on their domain set (α>0 and 0<β<1). An increment current of 1 A RMS in each AI phase is assumed. In [20], the authors propose to use droop control where the power-voltage droop functions are pre-defined. In this study, we adapt the LM2 type droop function presented in [20] for the nominal charging power of 10 kW. The droop function of this controller is illustrated in Fig. 6. The algorithm cuts off the charging power when the voltage goes below 0.9 pu (216 V). Charging power is linearly changed with voltage when the voltage is between 0.9-1.0 pu. When the voltage is above 1.0 pu, the EVs will be charged at the maximum rated power of 10 kW.

VI. SIMULATION TEST RESULTS

We refer the algorithms as follows: D-AIMD for our proposed distributed AIMD algorithm, C-AIMD for the ideal centralized AIMD algorithm [22], Droop for the droop control algorithm [20], and No-Control for the rated charging power for all EVs. To test the proposed D-AIMD, we first run the simulation with 0% EV penetration and obtain the substation power data. Then, this data is sent to the end-nodes, and every node computes three coefficients for their voltage-demand polynomial (Fig. 5). Nodes find a threshold voltage using (10) and implement Alg. 1. We then repeat the simulation at 100% EV penetration and with all EVs implementing the algorithm. Finally, we compare the performance score of each algorithm. The results are listed in Table I.

A. Voltage Stability vs. Capacity Utilization

AIMD-based control successfully avoids voltage violations across the grid since its triggering condition also includes minimum voltage check. Fig. 7 shows the minimum of voltage values measured across all 416 end-nodes at every second for each algorithm. Without any proper control (i.e., No-control case), the minimum voltage drops way below the critical level of 216 V resulting in an average VVS of 21,938 V-s as shown in Table I. C-AIMD and D-AIMD algorithms operate close to the critical voltage level since they have a condition that triggers MD phase when voltage limit is about to be violated. Droop algorithm results in a much higher minimum voltage. Consequently, none of the three algorithms violated the voltage limits of the grid, and therefore, their VVS are all zero.

Ideally, AIMD ensures the full utilization of system capacity since it is using the capacity congestion as its triggering mechanism. Fig. 8 presents the resulting substation apparent power for each algorithm. The red horizontal line indicates the substation capacity (2.5 MVA). In No-Control, the capacity is exceeded by more than 70% with a CUS of 170.8% and GCS of 1.82 MVAh (Table I). C-AIMD triggers when the capacity is reached, and this information is sent to the end-

\[ LCS = \frac{1}{K} \sum_{i=1}^{K} \int_{0}^{T} \max(0, S_{X}^{i}(t) - S_{X}^{i}(\text{rated})) \, dt \]  

where in our case, K = 104.

**Capacity Utilization Score (CUS)** is quantified as the maximum percentage (%) utilization of the capacity as:

\[ CUS = \frac{S_{P}^{0}(\text{max})}{S_{P}^{0}(\text{rated})} \times 100\%. \]  

where \( S_{P}^{0}(\text{max}) \) is the maximum apparent feeder power observed. A higher CUS means a more utilized network system.

**Average Charging Power Score (ACPS)** is the average of all N charging powers (\( P_{i}^{\text{avg}} \)) in kW, and calculated as:

\[ ACPS = \frac{1}{N} \sum_{i=1}^{N} P_{i}^{\text{avg}}. \]  

**Fairness Score (FS)** is calculated based on the Jain’s fairness index presented in (16), where \( w = \{ \tilde{w}_{1}, \tilde{w}_{2}, \ldots, \tilde{w}_{N} \} \) is the vector of average user shares. The same fairness measure is also used in [21]. In the fairest case where everyone has the same exact share, the fairness score becomes 1, whereas in the worst case, the score approaches 1/N, where N is the total number of users.

\[ FS(w) = FS(\tilde{w}_{1}, \tilde{w}_{2}, \ldots, \tilde{w}_{N}) = \frac{(\sum \tilde{w}_{i}^{*})^2}{N \cdot (\sum \tilde{w}_{i}^{*})^2}. \]  

**Communication Overhead Score (COS)** is based on the number of information exchanges needed for the EVs to implement AIMD.

**C. Benchmarks**

In order to evaluate the performance of the proposed algorithm, we run the test scenario with two other algorithms presented in [20] and [22]. [22] proposes an AIMD-based charging control assuming that the capacity CE is notified to each user over a unidirectional communication link in real time. This algorithm is the ideal implementation of the AIMD algorithm and therefore will serve as a benchmark. α and β parameters presented in Alg. 1 are set to 1 and 0.5, respectively. These values are chosen specifically based on their domain set (α>0 and 0<β<1). An increment current of 1 A RMS in each AI phase is assumed. In [20], the authors propose to use droop control where the power-voltage droop functions are pre-defined. In this study, we adapt the LM2 type droop function presented in [20] for the nominal charging power of 10 kW. The droop function of this controller is illustrated in Fig. 6. The algorithm cuts off the charging power when the voltage goes below 0.9 pu (216 V). Charging power is linearly changed with voltage when the voltage is between 0.9-1.0 pu. When the voltage is above 1.0 pu, the EVs will be charged at the maximum rated power of 10 kW.

![Fig. 7. Minimum voltage waveforms for all three algorithms.](Image)

| Algorithm  | VVS (V-s) | GCS (MVAh) | LCS (%) | CUS (%) | ACPS (kW) | FS | COS |
|------------|-----------|------------|---------|---------|-----------|----|-----|
| No-Control | 21.938    | 1.82       | 64.80   | 170.8   | 8.91      | 0.999 | 0   |
| D-AIMD     | 0         | 0          | 4.34    | 99.96   | 4.78      | 0.944 | 1   |
| C-AIMD     | 0         | 9.34 x 10^{-4} | 15.87 | 100.47 | 4.955 | 0.958 | 2880 |
| Droop      | 0         | 0.183      | 93.16   | 3.25    | 0.980     | 0   | 0   |

**TABLE I**

PERFORMANCE COMPARISON WITH DIFFERENT ALGORITHMS.
nodes. Therefore, it operates at full capacity (100.47% in Table I). D-AIMD, however, receives no information regarding the capacity, and makes its own estimation based on the derived voltage-power demand polynomial. It still manages to operate near the full capacity (99.96%), though slightly less than the ideal case (C-AIMD). Droop is utilizing less of the total capacity (93.16%). The reason that Droop is the least efficient in terms of utilization is that it uses the same droop function for all nodes. This causes the nodes experiencing low voltage to have also low charging power due to the linear droop function. The AIMD-based algorithms, however, keep increasing their charging power until the capacity limit is reached, which allows them to attain a higher utilization. On the positive side for Droop, it stays farther away from the critical voltage. Droop, essentially, provides a coarser trade-off between voltage stability and capacity utilization. Note that it is possible to tune the AIMD-based algorithms to a $V_{min}$ so that the distance to the critical voltage stays high in return of a drop in the capacity utilization. This can be the case when the substation peak loading is to be shaved below its rated capacity.

We present the current waveforms of a randomly chosen EV under four different algorithms in Fig. 9. Since the data-set used in the simulations are the same, the EV’s arrival time does not change per algorithm. In No-control, the EV is charged at the fixed current of 41 A until fully charged. D-AIMD and C-AIMD are dynamic in nature (due to AIMD) and constantly increase or decrease the charging current based on the network congestion. The current is more steady in Droop due to its pre-defined linear control droop function.

B. Fairness

AIMD allocates the available capacity among users in the fairest way if every end-user is ideally notified of the congestion at the same time. Fig. 10 demonstrates the distribution of average charging powers of 416 EVs for each algorithm in boxplots. No-control naturally results in the highest ACPS of 8.91 kW on average and highest FS of 0.999 since it charges at the maximum charging power for all EVs. The charging power distributions of D-AIMD and C-AIMD show that the resulting ACPSs are reduced below 5 kW (4.78 kW for D-AIMD and 4.95 kW for C-AIMD) to alleviate the grid stress. Although both algorithms generated nearly similar distributions, we can infer from the mean and the spread of the distributions that C-AIMD has a slightly higher mean value and less variation, and thus closer charging powers (i.e. fairer charging) due to its centralized mechanism. This leads to the FS of 0.944 for D-AIMD and 0.958 for C-AIMD as listed in Table I. Therefore, the proposed D-AIMD has almost the same performance as the ideal (C-AIMD) algorithm without needing any central information, and it starts operating only with local measurements after receiving the power demand information and extracting the voltage vs. power polynomial. Droop algorithm, on the other hand, resulted in the lowest average charging power (3.25 kW) but a narrower distribution spread indicating that it is fairer (0.980) than D-AIMD and C-AIMD.

C. Congestion

Table I shows that No-control results in an LCS of 64.80 kV Ah. By applying either of the three control algorithms, the substation congestion can be entirely avoided ($GCS \approx 0$ MVAh). However, despite its significant reduction, the local congestion is still present. We see that the lowest LCS of 0.183 kV Ah occurs with Droop, since its capacity utilization is not as much as the other two. Further, D-AIMD causes less local congestion than C-AIMD (i.e., an LCS of 4.34 kV Ah vs. 15.87 kV Ah in Table I). This is a direct result of higher capacity utilization of C-AIMD. Both AIMD-based algorithms do not directly take local congestion into account, and they operate based on global congestion information. C-AIMD can be greedier and causes more local congestion since it receives the global congestion information directly from the substation whereas D-AIMD acts upon an estimation of congestion at the end-node.

D. Communication Overhead

A key parameter to evaluate the algorithms is their need for central information. We defined COS to be the number Fig. 10. Distribution of average charging power of EVs for No-Control, D-AIMD, C-AIMD, and Droop algorithms.
of information exchanges. No-control and Droop charging require no additional information since they are fully decentralized, and therefore their COSs are both zero in Table I. D-AIMD receives substation power demand data only one time before charging and calculates a threshold based on it, which results in a COS of only 1 in Table I. C-AIMD, however, needs a central notification of the capacity event to operate. This means that the end-nodes must request this information at every algorithm period \((T_a)\), which is 10 s. This case study simulates only 8 hours between 16:00-24:00, and the information regarding the presence of the congestion is broadcast at 10 s intervals. Therefore COS for C-AIMD appears to be the highest (2,880) in Table I. So, using D-AIMD has significantly reduced the communication overhead and made the control operation almost fully decentralized.

We should note that using an external data (e.g. substation loading) certainly increases the amount of information and communication needs compared to fully-autonomous (stand-alone) solutions. However, our approach targets to minimize these dependencies by only requiring the substation loading, which technically carries no information regarding the grid topology and its parameters, and reducing the frequency of information exchange. This advantage highlights practical advantage of using D-AIMD with more secure and resilient operation against cyber attacks.

VII. CONCLUSIONS

In this paper, we introduced the AIMD algorithm that is adapted from the Internet and proposed an ML-assisted, distributed AIMD controller for EV charging. We exploited the relationship between voltage and power demand in a low-voltage distribution grid and presented a mathematical model of this relationship. Based on this relationship, we then proposed an ML-based method to estimate the grid capacity using only local voltage measurements.

We tested the proposed approach along with two other algorithms proposed in the literature on a test distribution grid. We concluded that all algorithms manage to avoid global power congestion and voltage violation. The performance of the proposed algorithm is very close to the ideal AIMD in terms of capacity utilization, average charging power, and fairness. The proposed D-AIMD method outperforms the droop control in capacity utilization by 6.8%, and in average charging power by 47% (+1.53 kW). Therefore, D-AIMD uses the available capacity more efficiently and performs charging faster compared to Droop control. As a result, it causes more average local congestion (+4.157 kV Ah) and slightly less fair charging (-3.6%) than Droop control. The proposed control significantly reduces the communication overhead need of the ideal C-AIMD (from a COS of 2880 to 1) and performs almost the same in terms of fairness and global congestion. In addition, the proposed algorithm outperforms the centralized one with respect to local congestion (from LCS of 15.87 kV Ah to 4.34 kV Ah).

We should note that this study does not consider the external factors that may impact the voltage and substation loading relationship. We can only provide some formal guarantees to the system operator that the system operate as expected after truly covering all these possible factors and worst case scenarios. This will be the subject of a future work where we will use more data that includes additional features such as frequency, weather, and end-node power injection or consumption to truly understand the source of variations and train our ML network accordingly. We will also investigate whether it is possible and to what extent to provide these formal guarantees.

REFERENCES

[1] F. Erden, M. C. Kısacıkoglu, and O. H. Gurec, “Examination of ev-grid integration using real driving and transformer loading data,” in Int. Conf. Elect. Electron. Eng., 2015, pp. 364–368.
[2] L. P. Fernandez, T. G. S. Roman, R. Cosent, C. M. Domingo, and P. Frias, “Assessment of the impact of plug-in electric vehicles on distribution networks,” IEEE Trans. Power Syst., vol. 26, no. 1, pp. 206–213, Feb. 2011.
[3] S. Shaﬁee, M. Fotuhi-Firuzabad, and M. Rastegar, “Investigating the impacts of plug-in hybrid electric vehicles on power distribution systems,” IEEE Trans. Smart Grid, vol. 4, no. 3, pp. 1351–1360, 2013.
[4] E. Veldman and R. A. Verzijlbergh, “Distribution grid impacts of smart electric vehicle charging from different perspectives,” IEEE Trans. Smart Grid, vol. 6, no. 1, pp. 333–342, 2015.
[5] N. Leemput, F. Geth, J. Van Roy, A. Delnooz, J. Buscher, and J. Driesen, “Impact of electric vehicle on-board single-phase charging strategies on a Flemish residential grid,” IEEE Trans. Smart Grid, vol. 5, no. 4, pp. 1815–1822, 2014.
[6] K. Clement-Nyns, E. Haesen, and J. Driesen, “The impact of charging plug-in hybrid electric vehicles on a residential distribution grid,” IEEE Trans. Power Syst., vol. 25, no. 1, pp. 371–380, Feb. 2010.
[7] P. Richardson, D. Flynn, and A. Keane, “Optimal charging of electric vehicles in low-voltage distribution systems,” in IEEE Power Energy Soc. General Meeting, Jul. 2012.
[8] S. Sojoudi and S. H. Low, “Optimal charging of plug-in hybrid electric vehicles in smart grids,” in IEEE Power Energy Soc. General Meeting, Jul. 2011, pp. 1–6.
[9] M. Restrepo, C. A. Caiuzares, and M. Kazerani, “Three-stage distribution feeder control considering four-quadrant EV chargers,” IEEE Trans. Smart Grid, vol. 9, no. 4, pp. 3736–3747, July 2018.
[10] L. Zhang, V. Kokatos, and G. B. Giannakis, “Scalable electric vehicle charging protocols,” IEEE Trans. Power Syst., vol. 32, no. 2, pp. 1451–1462, 2017.
[11] M. Liu, P. K. Phanivong, and D. S. Callaway, “Electric vehicle charging control in residential distribution network: A decentralized event-driven realization,” in IEEE 56th Annual Conf. Decision Control (CDC), 2017, pp. 214–219.
[12] M. S. Ibrahim, W. Dong, and Q. Yang, “Machine learning driven smart electric power systems: Current trends and new perspectives,” Applied Energy, vol. 272, p. 115237, 2020.
[13] S. Karagiannopoulos, P. Aristidou, and G. Hug, “Data-driven local control design for active distribution grids using off-line optimal power flow and machine learning techniques,” IEEE Trans. Smart Grid, vol. 10, no. 6, pp. 6461–6471, 2019.
[14] R. Dobbe, O. Sondermeijer, D. Fridovich-Keil, D. Arnold, D. Callaway, and C. Tomlin, “Toward distributed energy services: Decentralizing optimal power flow with machine learning,” IEEE Trans. Smart Grid, vol. 11, no. 2, pp. 1296–1306, 2020.
[15] R. Dobbe, P. Hidalgo-Gonzalez, S. Karagiannopoulos, R. Henriquez-Auba, G. Hug, D. S. Callaway, and C. J. Tomlin, “Learning to control in power systems: Design and analysis guidelines for concrete safety problems,” Electr. Power Syst. Res., vol. 189, p. 106615, 2020.
[16] S. Weckx, R. D’Huil, and J. Driesen, “Voltage sensitivity analysis of a laboratory distribution grid with incomplete data,” IEEE Trans. Smart Grid, vol. 6, no. 3, pp. 1271–1280, 2015.
[17] E. Ucer, M. C. Kısacıkoglu, and M. Yuksel, “Decentralized additive increase and multiplicative decrease-based electric vehicle charging,” IEEE Syst. J., to be published.
[18] L. Xia, T. Alpcan, I. Mareels, M. Brazil, J. de Hoog, and D. A. Thomas, “Local measurements and virtual pricing signals for residential demand side management,” Sustain. Energy Grids Netw., vol. 4, pp. 62–71, Dec. 2015.
[19] L. Xia, J. de Hoog, T. Alpcan, M. Brazil, I. Mareels, and D. Thomas, “Electric vehicle charging: A noncooperative game using local measures-ments,” IFAC Proc. Volumes, vol. 47, no. 3, pp. 5426–5431, 2014, 19th IFAC World Congress.
