

Article

Transformer-Based User Charging Duration Prediction Using Privacy Protection and Data Aggregation

Fei Zeng *, Yi Pan, Xiaodong Yuan, Mingshen Wang and Yajuan Guo

Electric Power Research Institute of State Grid Jiangsu Electric Power Co., Ltd., Nanjing 211103, China; panyi_seu@163.com (Y.P.); yuanxiaodong_2024@163.com (X.Y.); wmshtju@163.com (M.W.); mandy_guo@163.com (Y.G.)

* Correspondence: zf_2024@sina.cn

Abstract: The current uneven deployment of charging stations for electric vehicles (EVs) requires a reliable prediction solution for smart grids. Existing traffic prediction assumes that users' charging durations are constant in a given period and may not be realistic. In fact, the actual charging duration is affected by various factors including battery status, user behavior, and environment factors, leading to significant differences in charging duration among different charging stations. Ignoring these facts would severely affect the prediction accuracy. In this paper, a Transformer-based prediction of user charging durations is proposed. Moreover, a data aggregation scheme with privacy protection is designed. Specifically, the Transformer charging duration prediction dynamically selects active and reliable temporal nodes through a truncated attention mechanism. This effectively eliminates abnormal fluctuations in prediction accuracy. The proposed data aggregation scheme employs a federated learning framework, which centrally trains the Transformer without any prior knowledge and achieves reliable data aggregation through a dynamic data flow convergence mechanism. Furthermore, by leveraging the statistical characteristics of model parameters, an effective model parameter updating method is investigated to reduce the communication bandwidth requirements of federated learning. Experimental results show that the proposed algorithm can achieve the novel prediction accuracy of charging durations as well as protect user data privacy.

Keywords: data aggregation; privacy protection; prediction method; transformer



Citation: Zeng, F.; Pan, Y.; Yuan, X.; Wang, M.; Guo, Y. Transformer-Based User Charging Duration Prediction Using Privacy Protection and Data Aggregation. *Electronics* **2024**, *13*, 2022. <https://doi.org/10.3390/electronics13112022>

Academic Editor: Fabio Corti

Received: 24 April 2024

Revised: 15 May 2024

Accepted: 17 May 2024

Published: 22 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, electric vehicles (EVs) have been vigorously promoted across various provinces and cities in China due to their environmental friendliness, cleanliness, and energy efficiency advantages [1–3]. Electric vehicle charging stations, as an integral part of the EV product chain and the electric vehicle industry, are rapidly being developed. However, due to factors such as geography and policies, there has been a long-standing imbalance in the deployment of EV charging stations [4–6]. On the user distribution station side, load forecasting for individual charging stations' aggregation has become a focal point of research [5–8].

On the grid side, smart grids primarily consist of three components—intelligent infrastructure systems, intelligent management systems, and intelligent protection systems [3,9,10]. The intelligent infrastructure system includes intelligent energy subsystems, intelligent information subsystems, and intelligent communication subsystems. The widespread deployment of EV charging stations impacts power energy allocation and optimization, which are crucial factors to consider in building intelligent infrastructure. Compared to load forecasting, there has been relatively less work carried out on predicting charging duration, with average charging duration often chosen as the prediction result, resulting in poor accuracy [11]. On the one hand, factors such as the age of vehicle batteries, charging battery power, and remaining battery capacity differences can affect charging duration. Ignoring the actual charging duration differences of each vehicle and simply

using averages can lead to significant cumulative errors and severely affect the reliability of load data aggregation [12–15]. On the other hand, since charging stations are connected to the power grid, data related to electric vehicles, charging stations, and the power grid are susceptible to attacks, leading to data anomalies and decreased prediction accuracy, thereby causing issues such as unreasonable power energy allocation and uneven power load [16].

Therefore, by jointly considering the electric grid network architecture and user data from different charging stations, the prediction of user charging time becomes practically significant. On one hand, the state-of-the-art of this area generally relies on the data from single or few charging stations without fully utilizing electric grid network. Moreover, the data privacy has to face the high risk when conducting data aggregation. On the other hand, traditional prediction models including CNNs (convolutional neural networks) or recurrent neural networks would require extensive retraining or complex transfer strategies when migrating to significantly different applications. The main advantage of the Transformer model for charging duration prediction is its strong generalization ability to facilitate its direct application across various types of charging stations. Specifically, the Transformer-based mode can mitigate the impact of unstable current data on the prediction results.

In this paper, a charging prediction scheme is proposed by jointly considering data privacy and data aggregation. Firstly, it employs the Transformer-based model at an individual charging station for charging duration prediction. By incorporating a truncated attention mechanism, active and reliable sequential nodes are dynamically selected and effectively mitigate the abnormal fluctuation on prediction accuracy. Moreover, a data aggregation framework based on federated learning is proposed by taking data privacy into account. Considering the substantial parameterization of the Transformer model, the proposed privacy-preserving data aggregation framework selectively updates model parameters online. This approach ensures the efficient learning of model parameters, while protecting user privacy.

The remainder of this paper is organized as follows: In Section 2, we summarized some typical related works. In Section 3, the transformer-based charging duration prediction scheme is proposed including the framework and adaptive truncated sparse self-attention scheme. In Section 4, an effective parameter update and aggregation is discussed. The performance evaluation is presented in Section 5 and the paper is concluded in Section 6.

2. Related Work

Blockchain has been employed as a privacy protection framework for EVs within smart grids. A novel study [16] introduces a localized point-to-point (P2P) electricity trading model based on consortium blockchain for plug-in hybrid electric vehicles (PHEVs) within smart grids. This model incentivizes discharging PHEVs to balance local electricity demand for their own benefit, thus achieving a demand response. Numerical results based on real maps of Texas illustrate that the double auction mechanism can maximize societal welfare, while safeguarding PHEV privacy. In a similar study [17], a blockchain-based charging payment scheme is proposed to preserve privacy and traceability for electric vehicle (EV) users. This scheme utilizes smart contracts to ensure data transmission reliability and authentication and employs bilinear pairings for anonymous credential generation and validation. It achieves user identity privacy and transaction unlinkability, while allowing for the tracking of anomalous transactions or entities in specific scenarios. Reference [18] presents a blockchain-based, city-level intelligent charging station platform reconciliation application to address the inefficiencies in such platforms. The article compares three different fully homomorphic encryption techniques, including a novel polynomial homomorphic encryption algorithm, to enhance performance. The new algorithm demonstrates significant improvements in performance for all operations. Furthermore, the authors of [19] design an information security protection scheme for charging station backend service management centers based on Advanced Encryption Standard (AES) and HMAC-SHA256 encryption algorithms. They also construct a charging station communication simulation

platform to test the real-time communication of different types of data under plaintext transmission, encrypted transmission, authenticated transmission, and authenticated encrypted transmission scenarios.

Blockchain combined with fog computing is widely utilized for privacy protection in the data transmission of EV charging stations. In reference [20], the authors propose a decentralized privacy protection charging scheme using a combination of blockchain and fog computing to address the vulnerability of user data in electric vehicle charging systems to network attacks. This scheme leverages fog computing to provide localized services to reduce communication latency and employs a flexible consortium blockchain architecture to achieve decentralized and secure storage environments. Addressing privacy risks in charging station detection data transmission, reference [21] introduces a blockchain-based method for uploading and sharing charging station detection data to solve issues of data loss, difficulty in querying, and ensuring the uniqueness and integrity of data. The proposed method effectively reduces transmission latency and improves management efficiency. For vehicles within the same administrative region, the authors of [22] propose a bi-directional anonymous identity authentication protocol based on elliptic curve cryptography. For vehicles running in different administrative regions, the authors design a cross-domain identity authentication protocol using a consortium blockchain, where the regulatory agencies of each domain join the same consortium blockchain, and vehicle and charging station registration information are written into the chain. Reference [23] presents an improved k-anonymity privacy protection algorithm for electric vehicles, utilizing the k-anonymity principle to group electric vehicle users in nearby locations into the same virtual location for recommendation purposes. Additionally, the algorithm provides charging station recommendation services to users based on electric vehicle user driving distance and charging station utilization rate.

In the field of electric vehicle (EV) charging station data aggregation, reference [24] proposes a trustworthy aggregation method for EV charging privacy data based on a two-tier blockchain. This method employs a hierarchical aggregation framework, where, at the charging station level, a privacy-preserving local charging plan sharing algorithm is implemented, requiring vehicle owners to upload only encrypted charging plan data. Addressing the network architecture of vehicle-to-grid (V2G), reference [25] introduces a lightweight authentication and anonymous tracking scheme for V2G, leveraging non-singular elliptic curves to construct authentication protocols and pseudonymous techniques, thereby achieving a V2G network data sharing scheme based on member-level access. In terms of intra-vehicle communication, an anonymous communication link is established using onion routing technology to achieve trustworthy communication and data aggregation for electric vehicles.

In the domain of electric vehicle charging management, reference [26] presents energy blockchain technology, constructing a charging consortium chain algorithm suitable for shared EV charging [27], along with a charging pile management model based on energy blockchain. Regarding transaction platform management, reference [28] utilizes blockchain technology to design an aggregation scheme for storing charging transaction data. For electric vehicle location privacy protection, reference [29] employs local differential privacy techniques and Bayesian random multi-pseudonym mechanisms. This method, combined with the reconstruction algorithm of random multi-pseudonym [30], partitions the location domain to reduce the privacy domain and enhance the aggregation results.

In general, a significant amount of work focuses on the encryption and data protection of charging stations and their users. For data aggregation, numerous studies concentrate on employing multi-tier aggregation architectures and various signature and anonymity techniques to achieve effective privacy protection during data transmission. However, there are few works focusing on predicting charging durations. A highly accurate prediction of charging time would be expected to improve the charging station management for electric grid companies and operators. A commonly used method of prediction employs average historical data that may be influenced by various factors, leading to performance

deterioration. It is obvious that using data from different charging stations will increase the prediction accuracy. However, exchanging data from different charging stations should be carefully considered without losing privacy. Therefore, achieving the high-precision prediction of charging durations while ensuring data privacy is a key challenge.

3. Transformer-Based Charging Duration Prediction

3.1. The Framework of Charging Duration Prediction

As illustrated in Figure 1, the charging duration prediction scheme is proposed, which can be divided into two phases [15,31]. In the training phase, each charging station utilizes local data provided for model training and subsequently uploads the relevant model parameters to a central cloud server. The local data originates from a data session of user interactions with the charging stations. The data session, with its corresponding features, is presented in Table 1. The cloud server aggregates the model parameters uploaded from various charging stations. In the prediction phase, each charging station employs the model parameters disseminated from the cloud service, combined with local data, to predict charging durations. The prediction results can be transmitted via deployed infrastructure such as power and communication networks to charging station operators and utility companies, enabling the acquisition of the charging duration distribution for electric vehicles across the entire region.

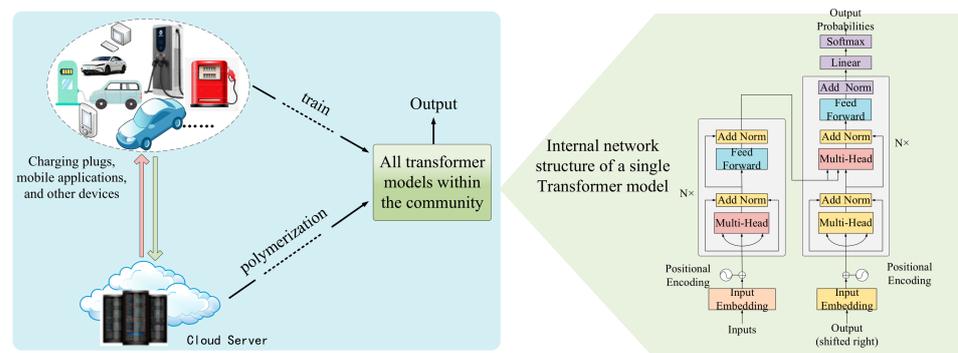


Figure 1. The structure of the charging duration prediction.

Table 1. Dataset information.

Features	Quantity
The number of locations where charging station user data is sourced	125
The total number of charging stations	124,548
The dimensionality of each sample feature	24
The number of samples in the test set	414,812
The number of normal data samples in the test set	228,644
The number of abnormal data samples in the test set	104,168
The number of incomplete data samples in the test set	82,000

The federated learning objective function designed in this study is presented in Equation (1); all charging stations follow this objective function for charging duration prediction.

$$\min_b \frac{1}{N_i} \sum_{n \in N_i} f(b, x_{i,n}, y_{i,n}) \tag{1}$$

where N_i is the total number of training samples. M represents the total number of charging stations and b denotes a vector composed of parameters of the Transformer model utilizing the adaptive truncation sparse self-attention mechanism described in Section 3.2. These parameters are delivered to the central server for aggregation. For charging station i , there are N_i training data samples, where the input vector of the n -th data sample is represented as $x_{i,n}$. This vector contains the data attributes reported in Table 2. The output vector is

denoted as $y_{i,n}$, which is the estimated charging duration. $f(\cdot)$ is the local objective function of the charging station and it is the loss function used in the Transformer.

Table 2. Main data attributes in each data session.

Features
Session ID
Unique identifier of EV
Time zone of charging station
Charging capacity
Time of connecting to charging station
Demanded charging energy
Time of the last non-zero charging rate
Time of disconnecting charging station
Measured energy delivered
Time of charging current becoming zero
Time of estimated departure time
Time of charging voltage becoming maximum

The objective function of federated learning is shown in Equation (1) through iteratively distributed optimization coordinated by a central controller, denoted as the Center Control (CC), which typically corresponds to a cloud server. A common optimization algorithm for federated learning is the local stochastic gradient descent method [2]. In each iteration of federated learning, each charging station, i , updates its o_i local model through τ iterations of stochastic gradient descent. Specifically, during the t iteration update process, the central coordinator shares the global model, b_t , with the devices. Each charging station initializes its local model as $o_{i,t+1}^0 = b_t$, then updates its local model parameters locally according to Equation (2).

$$o_{i,t+1}^k = o_{i,t+1}^{k-1} - \frac{\lambda_{t+1}^k}{N_{i,t+1}^k} \sum_{n \in N_{i,t+1}^k} \nabla f(o_{i,t+1}^{k-1}, x_i, n, y_i, n) \quad (2)$$

Here, $o_{i,t+1}^k$ represents the model of the local neural network for device i at the k -th local update during the $(t + 1)$ -th iteration, which corresponds to the parameters of the Transformer model utilizing the adaptive truncation sparse self-attention mechanism. λ_{t+1}^k denotes the learning rate and $N_{i,t+1}^k$ represents the training dataset for charging station i at the k -th step of the local stochastic gradient descent during the i -th iteration, consisting of a subset of N_i randomly selected $N_{i,t+1}^k$ training samples from the local training dataset. $\nabla f(\cdot)$ denotes the gradient of the objective function with respect to the model parameters, b . After each charging station completes τ steps of local training according to Equation (2), the locally trained model parameters $o_{i,t+1}^\tau - b_t$ are transmitted to the central coordinator. The central coordinator aggregates the parameters sent by all participating charging stations to form the global model parameters, following the specific process outlined below:

$$b_{t+1} = \frac{1}{N} \sum_{i=1}^M N_i (o_{i,t+1}^\tau - b_t) + b_t \quad (3)$$

The local stochastic gradient descent (SGD) algorithm enables multiple iterations of local model updates, as well as global federated learning model updates. When convergence is achieved, the global model parameters are represented as $b = o_1 = o_2 \dots = o_M$.

The algorithm proposed in this paper as shown in Algorithm 1, consists of a training phase and a prediction phase. For operators and power companies, accurate charging duration data for each charging station can be obtained through pre-deployed power or telecommunications networks; this algorithm supports such extensions. The next section provides a detailed analysis of the charging station selection and global model parameter aggregation in steps 2 and 5, along with corresponding effective methods.

Algorithm 1: Charging Duration Prediction

Step 1: All charging station systems register with the central cloud server.
 Step 2: The central cloud server generates initial global model parameters.
 Step 3: The cloud server randomly selects charging station systems to participate in training and sends the initialized global model parameters to the selected charging stations.
 Step 4: Each charging station trains and updates the received global model parameters.
 Step 4: Each charging station system sends local model parameters to the cloud server.
 Step 5: The cloud server aggregates the received parameters and updates the model.
 Step 6: The cloud server sends the aggregated model parameters to all charging stations.
 Step 7: Each charging station system inputs collected measurement data into the trained model to predict the charging duration.
Steps 2 to 5 are repeated L times.

3.2. Adaptive Truncated Sparse Self-Attention Mechanism

As discussed in Section 3.1, each charging station trains the deep neural network using local data and then delivers the parameters of the deep neural network to the central server. The central server conducts the data aggregation to capture the global optimal parameters. Considering the unstable current data, we adapt the Transformer to the deep neural network. That is to say, each charging station employs the local data to train the Transformer. The parameters of the Transformer are delivered to the central server for data aggregation.

In contrast to the CNN and other traditional deep neural networks, the key aspect of the Transformer lies in its self-attention mechanism, which utilizes attention for temporal information propagation. The studied prediction of charging duration is precisely based on the advantages of self-attention. Transformers have demonstrated a remarkable generalization performance in natural language processing and computer vision domains. Specifically, Transformers employ self-attention mechanisms to explore the dependencies between different positions in the input sequence. This enables them to dynamically attend to the highly correlated charging states within the sequence when processing sequences. As a result, Transformers are good at exploring the time relation, leading to novel generalization ability. Therefore, it is suitable for handling long sequences and capturing contextual information. The traditional self-attention model is as follows, with a time and space complexity of $O(L_q L_k)$. L_q is the length of query nodes and L_k is the length of key nodes

$$\text{self-attention}(Q, K, V) = \text{Softmax}\left(\frac{QK}{\sqrt{d}}\right) \quad (4)$$

In Equation (4), Q presents the vector composed of query nodes, K is the vector composed of key nodes, and V is the vector powered by the value of information. In Equation (4), the $\text{self-attention}(Q, K, V)$ probability distribution exhibits sparsity. Sparse self-attention follows a long-tail distribution. The so-called long-tail distribution refers to the fact that not all charging status information from previous time points, t , is strongly correlated with the current time point, t . By ignoring historical nodes that are weakly correlated to the current time node, t , it is possible to effectively reduce the computational complexity. In addition to reducing complexity, it is important to note that weakly correlated queries, Q , and keys, K , are prone to noise interference, which can significantly impact the accuracy of charging duration prediction. Therefore, we select active query nodes to calculate attention. This approach not only avoids noise interference, but also enhances the model's ability to extract key features of the charging status data, thereby improving the prediction model's generalization ability.

Based on the above rationale, we utilize KL (Kullback–Leibler) divergence [32] to screen active query nodes. The screening function is as follows:

$$M(q_i, K) = \ln \sum_{j=1}^{L_K} e^{\frac{q_i k_j}{\sqrt{d}}} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q_i k_j}{\sqrt{d}} \quad (5)$$

In practical calculations, the following simplified formula is used:

$$\overline{M}(q_i, K) = \max_j \left\{ \frac{q_i k_j}{\sqrt{d}} \right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q_i k_j}{\sqrt{d}} \quad (6)$$

where $M(q_i, K)$ represents the filtering function and q_i denotes the Q -th element in i . The average active query nodes filtered by the \overline{M} function are represented as \overline{Q} . As the value of the filtering function increases, the difference between the attention probability distribution and the uniform distribution becomes larger and the q_i makes a greater contribution to the attention function. Therefore, when calculating attention, we only need to compute those with larger q_i values, reducing the computational complexity and ensuring the robustness of attention. Moreover, since only active queries are selected during computation, this method avoids anomalies in the data. Based on the above formula, the attention function proposed by us is represented as:

$$\text{attention}(\hat{Q}, \hat{K}, V) = \text{Softmax} \left(\frac{\hat{Q}\hat{K}}{\sqrt{d}} \right), \quad (7)$$

where $\hat{Q} = a_1 \overline{Q}$ and $\hat{K} = a_2 K$. Inspired by GATv2 (Graph Attention Networks v2), we added two learnable $d \times d$ -dimensional parameter matrices, \overline{Q} and K , to the query, a_1 and a_2 , to mitigate static interference. \overline{Q} is a sparse matrix, composed of the first u query nodes filtered by the \overline{M} function, where the number u is calculated as follows:

$$u = c \times \ln(L_q) \quad (8)$$

For the query scores that were not selected, we use the mean value of V to ensure that both the input and output sequence lengths are L .

4. Effective Parameter Update and Aggregation

The scale of the Transformer parameters adapted using the adaptive truncation sparse self-attention mechanism is relatively large. Relying on federated learning for data aggregation imposes high demands on communication bandwidth. The cost of transmitting all model parameters to the central node each time is high, posing significant challenges to the computational and storage capacity of the cloud server. Therefore, this paper proposes a low-complexity algorithm for selecting model parameter training nodes. The description is as follows:

A charger maintains a set of local model parameters denoted as $D_i = \{x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(K)}\}$, consisting of data for K model parameters. For each local model parameter x_i^k , charger i computes its variance based on historical data, as shown in Equation (9).

$$\text{var}(x_i^k) = \frac{1}{N} \sum_{q=1}^Q (x_i^{k,q} - \mu x_i^k)^2 \quad (9)$$

The local charging piles sort these K parameters based on variance. Subsequently, the top and bottom 20% of the parameters with the highest and lowest variance are removed. The remaining parameters are then transmitted to the central cloud node. Based on this method of updating charging pile model parameters, the central node categorizes and sorts all received model parameters by parameter type. Similarly, for each model parameter, the central node calculates the variance, removes the top and bottom 10% of parameters with the highest and lowest variance, respectively, and aggregates the remaining parameters by averaging. The aggregated parameters are then distributed to each charging pile. Algorithm 2 describes the proposed updated algorithm.

Algorithm 2: Transformer Parameter Aggregation Update

Initialization: All charging station systems register with the cloud server. The charging stations and the cloud server generate initial model parameters.

Model Training: Each charging station maintains model parameters, denoted as D_i , which are iterated Q times. For each model parameter, the charging station calculates the variance. The top and bottom 20% of the parameters with the largest variance are removed. The remaining parameters are sent to the central node

Update: The cloud server aggregates the received model parameters, calculates the variance for each parameter, removes the top and bottom 10% with the largest variance, computes the average for the remaining parameters, and then sends the aggregated model parameters to each charging station.

Iteration: The process is repeated M times until model parameters no longer exhibit significant changes.

5. Experimental Results and Analysis

5.1. Dataset and Experiment Configuration

The data used in this study were collected by the electric grid company and charging station operators from 125 different locations where charging piles are deployed. In the dataset, 124,548 data samples provided by charging piles are used as the training set, while 414,812 data samples measured by charging piles are used as the test set. Each sample is a 24-dimensional vector. The test set includes 228,644 normal measurement data samples and 104,168 abnormal data samples (the abnormal data means that some items in the data session are lost, or that some items in data session are out of normal ranges). The scale of the dataset is shown in Table 1 and the data features are shown in Table 2 (data session length).

To validate the accuracy of the electric vehicle charging duration prediction algorithm proposed in this paper, we examined the training loss of the prediction algorithm under different scenarios, as follows: varying numbers of charging piles participating in aggregation, different batch sizes, and different numbers of local training epochs. The algorithm runs on the following hardware specifications: Intel i9 14900KF processor, NVIDIA RTX 4070Ti GPU, and 32 GB RAM. Each charging pile internally utilizes a Transformer with the same structure for training; data aggregation is conducted under the federated learning framework. The relevant parameters are shown in Table 3.

Table 3. Parameter configuration of Transformer.

Parameter Name	Value
Self-Attention Mechanism and the Number of Layers in Feedforward Deep Neural Network	12
Dimension of the Hidden Layer	1024
Number of Attention Heads	16
Intermediate Layer Dimension in Feedforward Neural Network	4096
Dropout Rate	0.1
Learning Rate	10^{-4}
Maximum Sequence Length	2048 token

5.2. Experimental Results and Analysis

In general, we use training loss that is calculated using the MMSE (minimum mean square error) to evaluate the performance of the proposed algorithm. In Figure 2, different proportions of charging piles are selected for training, with the participation rates of charging piles being 10%, 40%, and 80%, respectively. Overall, the training loss decreases as the number of communication rounds increases. The number of communication rounds represents the number of times the charging piles send the model parameters to the cloud server after multiple iterations. On the one hand, this trend indicates the effectiveness of the proposed prediction algorithm. On the other hand, we can observe that, during the training, as the proportion of participating charging stations increases, the decrease in training loss with the increase in communication rounds becomes smoother. This indicates that the central cloud server can achieve accurate model parameters with more information.

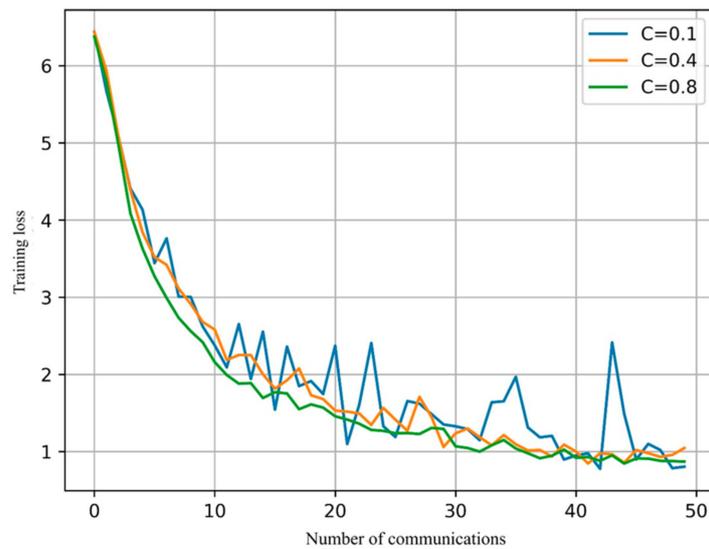


Figure 2. Training loss vs. number of communications, under different numbers of training charging stations.

In Figure 3, different batch sizes of charging pile datasets are selected for training, denoted by B , with B being 10, 50, and 100, respectively. Overall, the training loss decreases as the number of communication rounds increases. The number of communication rounds represents the number of times the charging pile sends the model parameters to the cloud server after multiple iterations. This trend is consistent with the trend shown in Figure 2. Additionally, we can observe that a smaller batch size leads to better training performance. For instance, the blue line in Figure 3 corresponds to $B = 10$, while the green line corresponds to $B = 100$. A smaller batch size indicates that a lower repetition rate of data are involved in each training iteration, allowing for a higher rate of utilization of new data. Consequently, the cloud server acquires more information, which is advantageous for training.

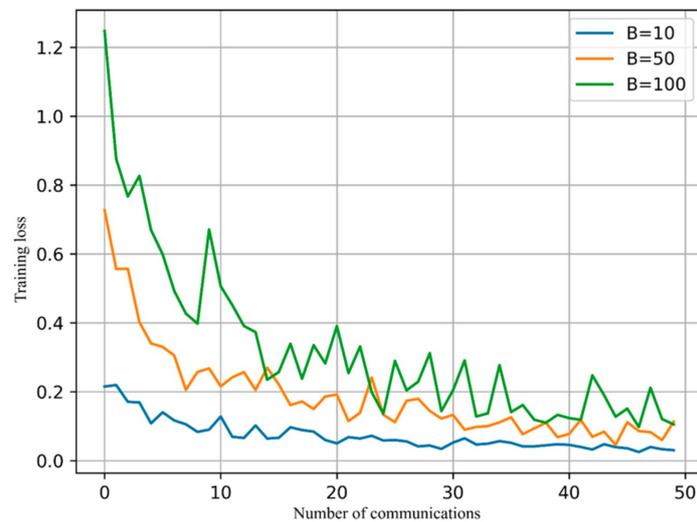


Figure 3. Training loss vs. number of communications, under different batch sizes.

In Figure 4, different numbers of local training epochs are selected for training, denoted by E , with E being 1, 5, and 10, respectively. Overall, the training loss decreases as the number of communication rounds increases. The number of communication rounds represents the number of times the charging pile sends the model parameters to the cloud server after multiple iterations. This trend is consistent with the trends shown in Figures 2 and 3. Additionally, we can observe that a larger number of local training epochs leads to a better

training performance. For instance, the green line in Figure 4 corresponds to $E = 10$, while the blue line corresponds to $E = 1$. A larger value of E indicates a greater number of local training rounds, resulting in more accurate local model parameters. This allows the cloud server to acquire more effective information, which is advantageous for training. We also report the estimation error of the charging duration when the training and aggregation are finished under the federated learning (Table 4).

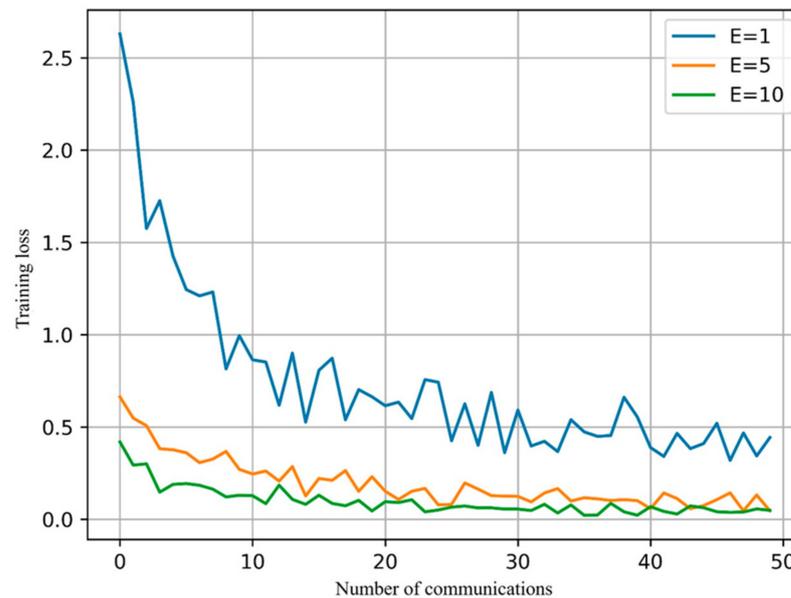


Figure 4. Training loss vs. number of communications, under different numbers of local epoch.

Table 4. Estimation error of charging duration.

EV ID	Actual Charging Duration (min)	Estimated Charging Duration (min)	Error Percentage (%)
1	307	300	2.28
2	200	210	5
3	79	75	5.06
4	221	213	3.62
5	216	211	2.31
6	100	95	5
7	366	361	1.37
8	191	180	5.76
9	423	435	2.84
10	370	363	1.89

To verify the Transformer parameter aggregation updating method, two types of attacks to affecting charging station data are considered. One is a sign-flipping attack and the other is a same-value attack. In the sign-flipping attack, the signs of the model parameters of the charging station are reversed, as are the values of the model parameters after amplifying the reversed signs by different degrees, denoted by ϵ . From Figure 5, it can be seen that the proposed algorithm achieves a success rate close to 100% after the charging pile is attacked at different proportions. In the same-value attack, the model parameters are amplified by different proportions, denoted by c . From Figure 6, it can be observed that after the charging pile is attacked at the same proportion, the smaller that c is, the higher that the success rate of the algorithm is, while the success rate decreases as the proportion of the charging pile attacked increases.

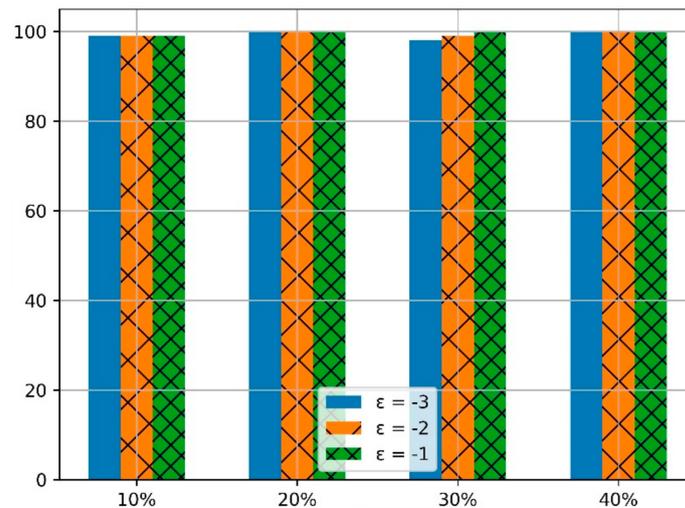


Figure 5. Success rate vs. different percentage of number of charging piles, under sign-flipping attack.

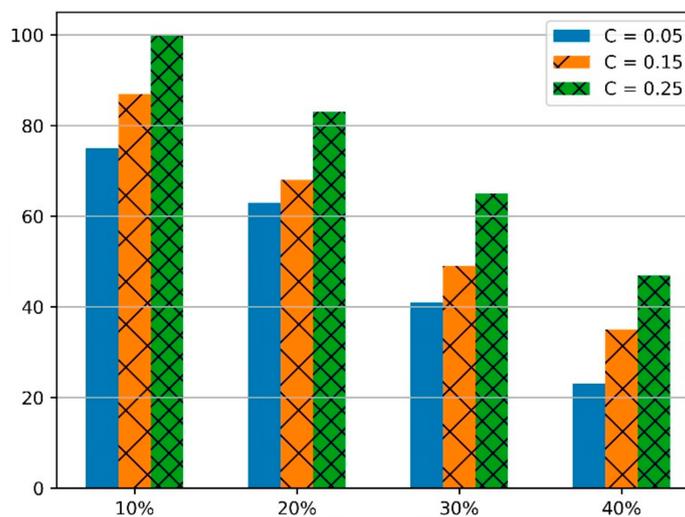


Figure 6. The success rate of the algorithm in detecting attacks on charging pile data at different proportions.

6. Conclusions

In this paper, a prediction scheme for the charging duration of electric vehicles is proposed. By leveraging federated learning, this scheme achieves novel prediction accuracy and protects data privacy based on the Transformer without the need for prior knowledge. More specifically, multiple charging station systems cooperate to train their own Transformer locally and then deliver the parameters of the Transformer to the central node. The central node aggregates multiple copies from different charging stations and sends the aggregated parameters to the charging stations. Moreover, the charging station employs a parameter update policy to reduce the bandwidth demand of federated learning without sacrificing data privacy. Experimental results demonstrate that the proposed scheme can accurately predict user charging durations, while protecting data privacy.

Author Contributions: Conceptualization, F.Z. and Y.P.; methodology, X.Y.; software, M.W.; validation, F.Z., Y.P. and Y.G.; formal analysis, Y.G.; writing—original draft preparation, F.Z.; writing—review and editing, F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the science and technology project of State Grid Jiangsu Electric Power Co., Ltd. (J2023015).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Authors Fei Zeng, Yi Pan, Xiaodong Yuan, Mingshen Wang, and Yajuan Guo were employed by the Electric Power Research Institute of State Grid Jiangsu Electric Power Co., Ltd. The authors declare that this study received funding from science and technology project of State Grid Jiangsu Electric Power Co., Ltd. (J2023015). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Wan, Y.; Cao, W.; Wang, L. A Prediction Method for EV Charging Load Based on Fuzzy Inference Algorithm. In Proceedings of the 2019 Chinese Control Conference (CCC), Guangzhou, China, 27–30 July 2019; pp. 2803–2808.
2. Mohammed, A.; Saif, O.; Abo-Adma, M.; Fahmy, A.; Elazab, R. Strategies and sustainability in fast charging station deployment for electric vehicles. *Sci. Rep.* **2024**, *14*, 283. [\[CrossRef\]](#)
3. Shoushtari, F.; Talebi, M.; Rezvanjou, S. Electric Vehicle Charging Station Location by Applying Optimization Approach. *Int. J. Ind. Eng. Oper. Res.* **2024**, *6*, 1–15.
4. Esmaili, A.; Oshanreh, M.M.; Naderian, S.; MacKenzie, D.; Chen, C. Assessing the spatial distributions of public electric vehicle charging stations with emphasis on equity considerations in King County, Washington. *Sustain. Cities Soc.* **2024**, *107*, 105409. [\[CrossRef\]](#)
5. Al-Ogaili, A.S.; Hashim, T.J.T.; Rahmat, N.A.; Ramasamy, A.K.; Marsadek, M.B.; Faisal, M.; Hannan, M.A. Review on Scheduling, Clustering, and Forecasting Strategies for Controlling Electric Vehicle Charging: Challenges and Recommendations. *IEEE Access* **2019**, *7*, 128353–128371. [\[CrossRef\]](#)
6. Fang, X.; Misra, S.; Xue, G.; Yang, D. Smart grid—The new and improved power grid: A survey. *IEEE Commun. Surv. Tutor.* **2011**, *14*, 944–980. [\[CrossRef\]](#)
7. Kotarela, F.; Rigogiannis, N.; Glavinou, E.; Mpailis, F.; Kyritsis, A.; Papanikolaou, N. Techno-Economic and Environmental Assessment of a Photovoltaic-Based Fast-Charging Station for Public Utility Vehicles. *Energies* **2024**, *17*, 632. [\[CrossRef\]](#)
8. Asamer, J.; Reinthaler, M.; Ruthmair, M.; Straub, M.; Puchinger, J. Optimizing charging station locations for urban taxi providers. *Transp. Res. Part A Policy Pract.* **2016**, *85*, 233–246. [\[CrossRef\]](#)
9. Deb, S.; Tammi, K.; Kalita, K.; Mahanta, P. Impact of electric vehicle charging station load on distribution network. *Energies* **2018**, *11*, 178. [\[CrossRef\]](#)
10. Capasso, C.; Veneri, O. Experimental study of a DC charging station for full electric and plug in hybrid vehicles. *Appl. Energy* **2015**, *152*, 131–142. [\[CrossRef\]](#)
11. Shariff, S.M.; Alam, M.S.; Ahmad, F.; Rafat, Y.; Asghar, M.S.J.; Khan, S. System design and realization of a solar-powered electric vehicle charging station. *IEEE Syst. J.* **2019**, *14*, 2748–2758. [\[CrossRef\]](#)
12. Wang, X.; Yuen, C.; Hassan, N.U.; An, N.; Wu, W. Electric vehicle charging station placement for urban public bus systems. *IEEE Trans. Intell. Transp. Syst.* **2016**, *18*, 128–139. [\[CrossRef\]](#)
13. Yuan, W.; Huang, J.; Zhang, Y.J.A. Competitive charging station pricing for plug-in electric vehicles. *IEEE Trans. Smart Grid* **2015**, *8*, 627–639. [\[CrossRef\]](#)
14. Timpner, J.; Wolf, L. Design and evaluation of charging station scheduling strategies for electric vehicles. *IEEE Trans. Intell. Transp. Syst.* **2013**, *15*, 579–588. [\[CrossRef\]](#)
15. Xiong, Y.; An, B.; Kraus, S. Electric vehicle charging strategy study and the application on charging station placement. *Auton. Agents Multi-Agent Syst.* **2021**, *35*, 3. [\[CrossRef\]](#)
16. Kang, J.; Yu, R.; Huang, X.; Maharjan, S.; Zhang, Y.; Hossain, E. Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains. *IEEE Trans. Ind. Inform.* **2017**, *13*, 3154–3164. [\[CrossRef\]](#)
17. Wu, Y.; Zhang, C.; Zhu, L. Privacy-preserving and Traceable Blockchain-based Charging Payment Scheme for Electric Vehicles. *IEEE Internet Things J.* **2023**, *10*, 21254–21265. [\[CrossRef\]](#)
18. Li, H.; Han, D.; Tang, M. A privacy-preserving charging scheme for electric vehicles using blockchain and fog computing. *IEEE Syst. J.* **2020**, *15*, 3189–3200. [\[CrossRef\]](#)
19. Jiang, L.; Li, T.; Jiao, R.; Zhou, W.; Zhang, Y.; Zhao, X. Blockchain-Based Charging Pile Detection Data Uploading and Sharing. In Proceedings of the 2022 IEEE 5th International Conference on Electronics Technology (ICET), Chengdu, China, 13–16 May 2022; pp. 1229–1233.
20. Aji, P.; Renata, D.A.; Larasati, A. Development of Electric Vehicle Charging Station Management System in Urban Areas. In Proceedings of the 2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP), Bandung, Indonesia, 23–24 September 2020; pp. 199–203.
21. Guo, J.; Zhao, H.; Shen, Z.; Wang, A.; Cao, L.; Hu, E.; Wang, Z.; Song, X. Research on Harmonic Characteristics and Harmonic Counteraction Problem of EV Charging Station. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018; pp. 1–5.
22. Arya, A.; Sridhar, S. Strategic Placement of Electric Vehicle Charging Stations Using Grading Algorithm. In Proceedings of the 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, 19–21 April 2023; pp. 99–104.

23. Lin, Y.J.; Chen, Y.C.; Zheng, J.Y.; Shao, D.W.; Chu, D.; Yang, H.T. Blockchain-Based Intelligent Charging Station Management System Platform. *IEEE Access* **2022**, *10*, 101936–101956. [[CrossRef](#)]
24. Santos, G.; Pina, J.M.; Belém, R. B2G (Buggy-to-Grid): Vehicle-to-Grid (V2G) concept in microgrids with strong penetration of electric vehicles. In Proceedings of the 2022 International Young Engineers Forum (YEF-ECE), Caparica/Lisbon, Portugal, 1 July 2022; pp. 106–111.
25. Odeh, Y.S.; Elkahlout, I.S.; Naeimi, P.V.; ElGhanam, E.A.; Hassan, M.S.; Osman, A.H. Planning and Allocation of Dynamic Wireless Charging Infrastructure for Electric Vehicles. In Proceedings of the 2022 9th International Conference on Electrical and Electronics Engineering (ICEEE), Alanya, Turkey, 29–31 March 2022; pp. 306–310.
26. Zheng, Q.; Zhao, J.; Xiang, J. Research on Blockchain Reconciliation Applications of Intelligent Charging Pile Platforms. In Proceedings of the 2021 IEEE 3rd International Conference on Power Data Science (ICPDS), Harbin, China, 26 December 2021; pp. 62–68.
27. von Hoffen, M. Enrichment and context-based analytics of Electric Vehicle charging transaction data. In Proceedings of the 2016 6th International Electric Drives Production Conference (EDPC), Nuremberg, Germany, 30 November–1 December 2016; pp. 94–100.
28. Ju, Z.; Li, Y. Local Differential Privacy-Based Privacy-Preserving Data Range Query Scheme for Electric Vehicle Charging. *IEEE Trans. Netw. Sci. Eng.* **2024**, *11*, 673–684. [[CrossRef](#)]
29. Li, Y.; Ma, Y.; Zhu, Z.; Qi, G.; Zhou, Z.; Huang, X. Blockchain-based EVs Charging Right Trading Model. In Proceedings of the 2023 IEEE 6th International Electrical and Energy Conference (CIEEC), Hefei, China, 12–14 May 2023; pp. 1107–1111.
30. Wang, W.; Peng, X.; Yang, Y. Self-Training Enabled Efficient Classification Algorithm: An Application to Charging Pile Risk Assessment. *IEEE Access* **2022**, *10*, 86953–86961. [[CrossRef](#)]
31. Han, K.; Xiao, A.; Wu, E.; Guo, J.; Xu, C.; Wang, Y. Transformer in transformer. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 15908–15919.
32. Johnson, D.H.; Sinanovic, S. Symmetrizing the kullback-leibler distance. *IEEE Trans. Inf. Theory* **2001**, *1*, 1–10.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.