EVCSeer: An Exploratory Study on Electric Vehicle Charging Stations Utilization via Visual Analytics

Yutian Zhang, *School of Intelligent Systems Engineering, Sun Yat-sen University, Guangdong, China*
Shuxian Gu, *School of Intelligent Systems Engineering, Sun Yat-sen University, Guangdong, China*
Quan Li, *School of Information Science and Technology, ShanghaiTech University, Shanghai, China*
Haipeng Zeng, *School of Intelligent Systems Engineering, Sun Yat-sen University, Guangdong, China*

Abstract—Promoting the development of electric vehicles requires the widespread deployment of charging infrastructure, which poses numerous technical and financial constraints. Despite extensive research focusing on optimizing charging station locations, few studies have accounted for charging station utilization and the factors that influence it. This study aims to evaluate charging station operations and explore charging station utilization to inform planning and facilitate better utilization of funds for expanding charging infrastructure. We present EVCSeer, a visual analytics system that utilizes representative predictive models and well-designed visualizations to analyze factors affecting charging station utilization, compare deployment strategies, and optimize utilization. The system also enables "what-if" analysis of charging station deployments. Two case studies, expert interviews, and a qualitative user study support the validity and usefulness of EVCSeer.

Due to depleting oil reserves and deteriorating environmental pollution, the future of electric vehicles (EVs) as a major source of new energy consumption is increasingly promising. Facilitating the development of EVs requires widespread deployment of charging infrastructure [18]. Governments and industries [7] around the world plan to provide millions in investments to deploy EV charging stations in the near future.

However, deploying charging infrastructure is extremely costly and subject to a number of technical and financial constraints [19]. In order to quickly identify and expand new charging stations, a large amount of research has focused on optimizing the location of charging stations (e.g., how many charging stations are needed in a given area and where they should be placed) [17, 3]. Although previous work has proposed a series of representative approaches [4, 14], they tend to solve complex mathematical problems, leading to unintuitive and time-consuming results. In addition, they focused only on optimizing the location of charging stations without considering the ultimate utilization rate[1], which is a posterior indicator of charging station effectiveness. Specifically, a high utilization rate indicates that existing charging stations are well utilized, leading to increased profits. A low utilization rate may indicate a large number of redundant charging stations that waste resources. Therefore, charging station utilization can provide feedback for adjusting the deployment of charging stations.

As the number of EVs and charging stations increases, the utilization of public charging facilities around the world remains low [6]. Researchers have conducted a series of studies on charging station utilization, which generally fall into two categories, *prediction* and *exploration* [23, 6]. However, existing works do not explicitly explore the importance of the different factors affecting utilization. Although Adenaw et al. [1] derived a measurable set of influencing factors from the literature and explored their impact on charging

XXXX-XXX
Digital Object Identifier 10.1109/XXX.0000.0000000

[1] Utilization rate/Utilization: the percentage of in-used chargers in a charging station.

FIGURE 1. Our visualization system, *EVCSeer*, supports evaluating charging station operations and exploring charging station utilization. The data overview (A) shows an overview of charging stations, including charging clusters and utilization distribution. The spatial view (B) shows the spatial distribution of charging stations and detailed information with glyphs. The temporal view (C) provides temporal information of charging stations and prediction of utilization, as well as exploration of feature importance. The planning view (D) provides "what-if" analysis, where users can adjust different factors and compare corresponding results.

demand, what they fail to address is how we can use these factors to tailor deployment strategies (i.e., where to deploy charging stations, what types of chargers to install, and how many chargers are needed) to improve charging station utilization, which is our focus.

Visual analytics is an effective approach for location selection problems, which can naturally consider various influencing factors for deployment strategies. Previous studies have proposed representative visual analytics-based interfaces to select facility locations, such as home finding [21], billboards [13] and warehouses [11]. Unlike previous studies, our study applies visual analytics to charging station location problems, which remain unexplored and challenging. We need to consider the posterior indicator (i.e., utilization) and the impact of different factors on charging station deployment, as well as people's preferences in terms of where and when they charge their EVs [6]. More importantly, optimizing the utilization of charging stations is a nontrivial task due to three major challenges. **(C.1) Pattern exploration.** The utilization of charging stations contains large-scale spatiotemporal data. Charging stations in different regions and at different times may exhibit different performance and patterns. Simple statistical approaches cannot provide an intuitive overview of the spatiotemporal patterns of large-scale charging station utilization data. Without a comprehensive understanding of spatiotemporal patterns, it is difficult to guide how to improve the deployment of charging stations. **(C.2) Influencing factors.** The utilization of charging stations is influenced by numerous factors [1], such as temporal variation in charging demand, spatial geographic information, and attributes of charging stations. It is difficult to identify and understand the most significant influencing factors and how the influencing factors affect the utilization of charging stations. **(C.3) Deployment strategies.** There may be a large number of charging station deployment strategies in a given region. Different deployment strategies have different advantages and disadvantages, which may lead to different utilization rates of charging stations. To obtain a better deployment strategy, it is necessary to consider different influencing factors and compare different deployment strategies, which is challenging.

To address the above challenges, we propose a visual analytics approach to support the exploration and understanding of charging station utilization. We first conduct an observational study of the methods currently used by domain experts to analyze charging station utilization and identify their primary needs and concerns. Then, based on literature and the observational study, we identify a set of measurable influencing factors, including *spatial information* and *temporal variation* in charging demand. Specifically, we develop a predictive model to forecast charging station utilization, taking into account both spatial and temporal factors. In addition, we use the SHAP (SHapley Additive exPlanations) model [15] to analyze the importance of features (**C.2**), which can provide guidance for improving charging station utilization. Next, based on the identified factors, we evaluate the utilization of charging stations with the predictive model. Finally, we craft a visualization system (Fig. 1), named *EVCSeer* to better explore the spatiotemporal patterns of large-scale charging station data and to effectively compare different deployment strategies. Specifically, our system provides different views, from city-level to location-level, to support users in exploring spatiotemporal patterns (**C.1**). Further, our system provides “what-if” analysis that allows users to adjust different influencing factors, which can help to explore better deployment strategies and improve charging station utilization (**C.3**). Afterwards, we conduct two case studies, expert interviews and a qualitative user study to demonstrate the efficacy of our approach. The major contributions of this paper can be summarized as follows:

- We consider different spatiotemporal influencing factors on charging station utilization and compare different deployment strategies, which provides better guidance for designing deployment strategies.
- We propose a visual analytics system to support the exploration, understanding, and “what-if” analysis of EV charging station deployment problems, aiming to improve the utilization of EV charging stations.
- We demonstrate the usability and effectiveness of *EVCSeer* through two case studies, expert interviews, and a qualitative user study.

Related Work

EV Charging Station Siting

With the popularity of EVs in recent years, the siting problem of charging stations has received increasing attention. According to the representation of charging demand [10], models for solving charging station siting can be broadly classified into two categories: node-based and flow-based approaches. The node-based approach assumes charging demand rises directly at a location, which is similar to classic facility sitting problems [3]. Researchers often establish correlations between charging demand and various factors at different locations, such as traffic, POI, population and so on [4, 1, 24]. On the contrary, the flow-based approach suggests charging demand occurs when EVs are traveling [10]. In recent years, the charging station siting problem has seen significant research advancements, attracting researchers from multiple fields such as transportation and power systems [17, 20].

Although prior studies contribute significantly to the charging station location problem, they still have some limitations. First, the generation of charging demand is based on some assumptions, which may not reflect the situation in the real world. Only a few studies such as [24] consider the real utilization rate to better demonstrate the connection between charging demand and influencing factors. In addition, despite computational complexity and simplification of many factors [10], existing models for charging station siting cannot replace experts to make decisions. However, little work has considered how to better integrate domain knowledge into decisions about the operational status of EV charging stations. In this work, we focus on predicting utilization and exploring the factors that influence utilization. We further develop a visual analytics system that can help experts make more detailed and comprehensive decisions based on adjusting different factors and comparing results.

Charging Station Utilization Analysis

Analyzing charging station utilization can provide feedback for adjusting deployment strategies and deploying appropriate charging stations. Researchers have conducted a series of studies on charging station utilization, which generally fall into two categories, exploration and prediction. In terms of exploration, many studies explored spatial and temporal patterns of EV charging station utilization via visualization techniques [16, 6, 8]. However, these efforts mainly focused on displaying the utilization data directly, which overlooked explicitly exploring different influencing factors that may affect utilization and addressing how these influences can be used to tailor deployment strategies. In terms of prediction, it is important to understand the various factors that affect the utilization of EV charging stations. The charging demand at a given location is influenced by activities associated with that location [1], such as POIs (Points of interest), weather, traffic, and nearby events. In addition, some attributes of charging stations, including charging price and geographic location, are important features that need to be considered. In recent years, machine learning has been increasingly used to predict charging demand and compute feature correlations [23, 26].

In this work, we consider the three aspects mentioned above (i.e., temporal variation in charging demand, spatial geographic information, and attributes of charging stations) and explore the influencing factors. Moreover, we propose a visual analytics system to better explore utilization and influencing factors in spatiotemporal dimensions.

Visual Location Selection

Location selection is an essential issue in modern urban planning. Previous studies have proposed representative visual analytics-based interfaces to select facility locations, such as home finding [21], billboards [13], retail stores [9], ambulance stations [12], and warehouses [11]. Unlike previous studies, our study applies visual analytics to the charging station location problems, where we need to consider the posterior indicator (i.e., utilization) and the impact of different factors on the utilization of a charging station over time. More importantly, people have certain preferences in terms of where and when they charge their EVs [6]. For example, people are likely to charge during their lunch break. They are also likely to charge after work at a public charging station close to where they live. Second, people are likely to leave their vehicles during charging [22]. The environment around the charging station (especially POI) also affects the location of the charging station.

In general, the change in charging demand is closely related to time, space and the surrounding environment (e.g., POI, nearby charging stations). It necessitates to demonstrate the influence of different factors on the utilization over time. Therefore, existing visualization solutions cannot be directly applied to our problem. In this work, we integrate several well-established techniques to explore the spatiotemporal patterns and the influential factors of utilization data for large-scale charging stations.

Observational Study

Experts' Conventional Practices

To understand the current deployment of public charging stations in the industry, we worked with two domain experts from two different charging station deployment companies, denoted as E1 and E2. E1 (male, age: 36) is an engineer who has been involved in charging station deployment for approximately 8 years. E2 (male, age: 43) is currently the manager responsible for deploying charging stations and has 7 years of experience in charging station planning. Both experts have extensive experience and have helped us gain a deeper understanding of charging station planning and insight into the limitations of traditional methods. Overall, experts believed that the need to deploy charging stations to meet the growing demand for EVs remains high. In particular, they shared the typical process that most charging station companies use to deploy charging stations, as follows. First, they need to find some candidate locations. Most of the time, the candidate locations are usually chosen from public parking lots or vacant lots. In the next step, they mainly use the service provided by mobile navigation software to check the traffic conditions and the surroundings of the candidate locations (e.g. POI), both of which are important aspects to ensure user stability. After the construction is completed, they will monitor and evaluate the operation of the charging stations.

While domain experts have considered different strategies for deploying charging stations, they believe that the current approach is largely based on crude data analysis and past deployment experience. To sum up, they have encountered the following major problems. First, experts only use limited data to analyze and evaluate different charging stations. E2 mentioned that "*Assessing future returns is hard with limited data, such as historical utilization and profit. Therefore, we usually estimate future returns based on averages and assume the worst-case scenario.*" Second, there is no analysis of the importance of each factor. E1 said "*It is difficult to know how each factor affects the final utilization rate. Once I estimated it would take two years to recover the investment of a charging station, but it was achieved in less than one year, and I'm not sure why.*" Third, it is difficult to choose an optimal deployment option from multiple deployment schemes. It would be useful to provide some methods for selecting different deployment options.

To better understand how charging stations are used in daily life, we further interviewed 10 EV drivers (8 taxi drivers and 2 ride-hailing drivers). Although they are not direct target users of our system, they can inform our study from the perspective of charging station users. Their driving experience ranged from 2 years to 25 years. In general, they charge their cars when charging prices are low (midday, afternoon, evening) and prefer to use fast charging. When charging, they do other things such as having dinners, going home for a break, taking a lunch break, using the bathroom and fetching water. The reasons for choosing a particular charging station are proximity to home, ease of eating, no queues, and easy access to transportation. Overall, the utilization of charging stations is influenced by drivers' preferences and daily activities.

Experts' Requirements

To ensure that our approach meets the needs of the field, we further interviewed E1 and E2 to identify their key concerns for improved charging station deployment. Based on the experts' feedback, we have summarized the specific design requirements below.

R.1 Explore the spatiotemporal patterns of charging station utilization. After completing the construction of a charging station, experts need to monitor and evaluate the operation of the station. Utilization data can provide direct information about the usage of charging stations. Therefore, experts expect our approach to provide an overview of utilization rates and help them easily explore spatiotemporal patterns, which can provide valuable feedback guidance for subsequent charging station deployments.

R.2 Predict the utilization rate of a charging station. As a posterior indicator, the utilization reflects the effectiveness of charging stations, helping to decide whether to build charging stations at candidate locations if knowing the future utilization. Experts want to predict the utilization of planned charging stations.

R.3 Identify influencing factors and measure their importance for utilization. Experts mentioned that they need to consider many different factors before deciding on the final location of a charging station. In order to predict the final utilization rate, it is crucial to identify those factors that have the greatest impact on utilization. Since land use and the grid are relatively uncontrollable and almost fixed factors, E1 suggested that we consider other factors and measure their importance to utilization.

R.4 Compare the operational status of different charging stations. Charging stations in different areas and at different times behave differently, which may be influenced by the surrounding environment and other factors. Experts said that it would be helpful to compare the operational status of different charging stations.

R.5 Support "what-if" analysis to select a better deployment strategy. After identifying a number of candidate locations for building new charging stations, some efforts are needed to select a better deployment strategy. The experts wanted to visualize the performance of different deployment strategies. They felt that if we could adjust the different factors and then immediately see some of the effects, this would help provide valuable guidance for selecting a better deployment strategy.

Approach Overview

As shown in Fig. 2, our system pipeline consists of two parts: the back-end engine and the front-end visualization. In the back-end engine, we first collect OD (Origin-Destination), POI (Point of interest) and utilization data and store them in a database. Then, we determine the influencing factors (**R.3**) to predict the utilization (**R.2**). After that, we analyze the importance of different factors (**R.3**) and build a "what if" analysis framework (**R.5**). For the front-end visualization, we develop four coordinated views to facilitate the analysis of charging station utilization. The data overview provides an overview of charging stations, such as the distribution, average, and utilization details for each charging station (**R.1**). The spatial view displays the glyphs of charging stations on a map, allowing easy comparison of charging station details (**R.4**). The temporal view shows the actual and predicted utilization rate as well as feature importance (**R.2 – R.3**). The planning view allows users to make some adjustments to different factors and explore possible outcomes (**R.5**). Rich interactions are also provided to enhance our system.

Back-end Engine

Data Description

Our study focuses on Shenzhen, a city with the widespread adoption of electric vehicles. Experts provided us with taxi OD data and charging station utilization data, which have been stripped of identifying information. We further extract the POI data for Shenzhen through a public map API[2]. Specifically, the data can be categorized into three groups as follows: **1) OD Data.** OD data records the origin and destination of a trip, which can display the traffic demand in different regions of the whole city. The raw data records the GPS data generated by Shenzhen taxis from September 2019, which contains more than 39 million OD pairs. To better capture the traffic demand, we divide the whole city into many 1 km by 1 km grid cells, and

[2] *AMap* API: https://lbs.amap.com/

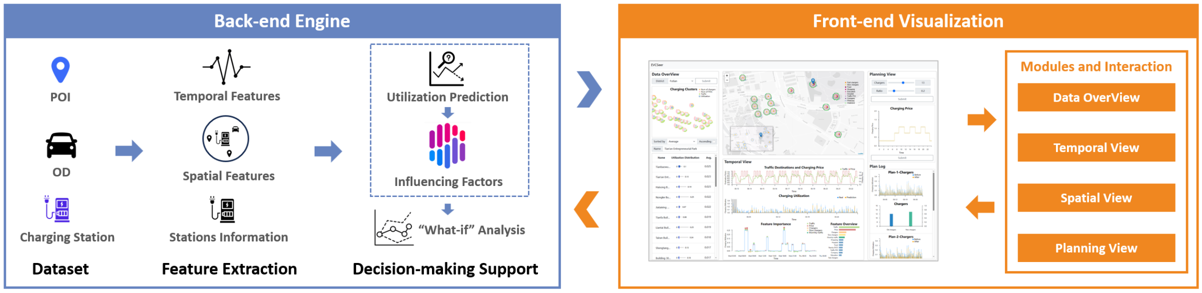


FIGURE 2. The system pipeline of *EVCSeer*: in the back-end engine, three types of data are collected, then some features are identified and used to train a prediction model. Afterward, we combine the prediction model with SHAP to analyze influencing factors and support "what-if" analysis; in the front-end visualization, four coordinated views with rich interactions are provided to support the exploration and optimization of charging station utilization.

then count the OD pairs in each grid cell. Over 2.5 thousand grid cells are used. **2) POI Data.** POI data are closely related to the life of citizens and the function of urban areas. For example, an area with many POIs of restaurants can indicate a commercial area where many people come to have lunch. In this study, we collect over 154,000 POI data. All POI data can be classified into 20 industry categories and 158 detail categories. **3) Charging Station Utilization Data.** The charging station utilization data used in this study contains 2,201,760 records from 1,529 charging stations in Shenzhen in June 2022. This data records the number of charging piles used by every 30 minutes. In addition to the charging station utilization rate, the data also contains basic information about each charging station (e.g., location, number of charging piles).

Feature Extraction

In order to predict the utilization, based on the feedback from two experts, we consider the factors affecting the demand for charging stations in three main aspects (i.e., temporal variation of utilization, spatial geographic information, and attributes of charging stations) and extract the corresponding features for prediction: **1) Temporal Features.** An important feature is the time-varying character of charging behavior. In our analysis, we slice the day by 30 minutes and encode each period. We also consider long-term temporal behavior, such as the day of the week and whether it is a weekday. **2) Spatial Geographic Information.** We define 7 variables concerning traffic conditions and POIs related to charging behavior (i.e., food, shopping, education, hospitals, transportation facilities and companies). For each charging station, we calculate the number of 6 types of POIs and destination points within 1 km for each time period. Nearby charging stations as competitors attract some of the demand. The number of charging stations within 1 km is denoted as the competition quantity ("*Nearby EVCS*"). **3) Attributes of Charging Stations.** First, based on the utilization data provided by experts, we extract some features related to charging stations, including utilization and the number of charging piles. In this study, utilization refers to the proportion of charging piles used at a charging station. The features of charging piles include the total number of charging piles ("*Chargers*"), the number of fast charging ("*Fast chargers*") and slow charging piles ("*Slow chargers*"), which are different for each charging station. In addition, we develop a web crawler based on *Python* to extract the charging price of different charging stations.

Utilization Prediction & Feature Explanation

To better predict the utilization, we use widely used models (e.g., *QRNN*, *LSTM*, *XGBoost* [2]) and conducted a series of experiments. Model performance is measured based on MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and EVS (Explained Variance Score). Lower MAE, lower RMSE, and higher EVS imply better performance. We choose *XGBoost* as the final prediction model because of its better performance (Table 1). To show how each feature affects the final utilization prediction, we apply *SHAP* [15], an interpretation method that computes Shapley values at the instance level that can be used to estimate the feature importance on the utilization prediction. Compared to other methods, the SHAP framework can determine whether a factor is beneficial for increasing utilization and quantify the feature importance of each factor. In addition, SHAP can also calculate feature importance for each sample, which allows users to observe changes in feature importance across time. Therefore, SHAP can better measure the influence of different factors, which supports "what-if" analysis and provide assistance in comparing different deployment strategies for charging stations.

TABLE 1. Performance comparison of different models.

Model	MAE	RMSE	EVS
Linear	0.15	0.80	42.71%
QRNN	0.15	0.75	51.42%
RandomForest	0.12	0.52	66.78%
LSTM	0.13	0.55	73.64%
XGBoost	**0.08**	**0.18**	**92.13%**

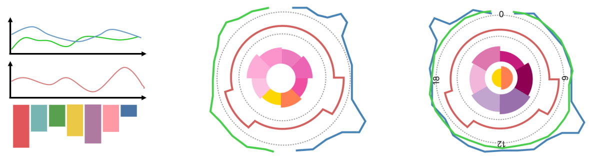
Front-end Visualization

Data Overview

To explore the utilization of charging stations (**R.1**), we present the corresponding information in the data overview (Fig. 1A). The data overview is divided into two parts: the upper part shows charging station clusters (Fig. 1A_1) and the lower part presents the utilization distribution of each charging station (Fig. 1A_2). After users select a district, the data overview will be updated accordingly. The charging stations in the selected district are first clustered according to their location and distance using the *DBSCAN* algorithm [5]. Then, using the *t-SNE* algorithm, the charging stations are projected onto a 2D plane based on similarity (i.e., number of chargers, number of POIs, and traffic conditions, location clusters generated by the *DBSCAN* algorithm). The opacity of the blue donuts indicates the mean utilization of the charging station and the similarity information is encoded as different sectors around the center. In this way, similar charging stations are placed close to each other on the plane. To avoid overlap, the sectors around the center will be hidden at first and only be displayed when users zoom in. Users can select the station of interest and then the corresponding information will be updated in the spatial view and temporal view. Also, users can easily link to the corresponding charging station in the lower part of the data overview (Fig. 1A_2). The utilization of each charging station is encoded with a horizontal box plot, and users can observe the average, minimum and maximum utilization values. Also, users can sort this view by these three values and easily compare different charging stations in this view.

Spatial View

To describe in detail the operation status of different charging stations and to compare different charging stations (**R.4**), we design the spatial view (Fig. 1B), which contains a map as a background. At a rough level, each charging station is encoded as a round glyph on the map (Fig. 1B_1). The opacity of the outer blue circle encodes the utilization. Higher opacity indicates higher utilization and lower opacity indicates lower utilization. Inside the outer blue circle, two semicircles encode the scale of the charging station. The



(A) (B) (C)

FIGURE 3. Glyph designs for showing detailed information about a charging station. (A) A design encodes the POI information as several bars and encodes the utilization and price as lines along time. (B) A circular design encodes the number of chargers and POI information as several sectors, the charging price as a dark red circular line, and the utilization on weekdays and weekends as two semi-circular lines (right and left circular lines). (C) A circular design encodes the number of chargers as semicircles, the POI information as several sectors, the charging price as a dark red circular line, and the utilization on weekdays and weekends as blue and green circular lines.

yellow semicircle encodes the number of slow chargers while the orange one indicates fast chargers. A larger radius indicates a larger scale of charging stations and vice versa. To display detailed information, users are allowed to zoom in, and then each round glyph becomes a detailed glyph (Fig. 1B_2), which makes it easier to compare detailed information about the charging stations. As shown in Fig. 3C, the glyph encodes four types of information. In the center, the radius of two semicircles encodes the number of slow and fast chargers; the radial pie chart around the two semicircles encodes different POI information around the charging station; in the middle part, a dark red line is used to encode the change in charging price along a 24-hours circular timeline; in the outer part, two different colored lines are used to encode the average utilization rate in weekdays (blue line) and weekends (green line). Users can easily compare the status of different charging stations.

Justification. Before finally adopting this glyph design (Fig. 3C), we considered other alternative designs. For example, we considered the design in Fig. 3A, where the utilization and price are encoded with lines, and POI information is encoded with several bars. However, the information in this design is relatively scattered, and it is difficult to demonstrate the specific location of the charging station on the map. Therefore, we used a circular design, as shown in Fig. 3B, where the number of chargers and POI information are encoded with several sectors, and the price is encoded with a circular line in the middle. We then used two semi-circular lines (the left and right circular lines) to encode the utilization on weekdays and weekends.

However, E1 said *"it is not easy to compare the utilization on weekdays and weekends."* Therefore, we considered two full circular lines of different colors to encode the utilization rate of weekdays and weekends (Fig. 3C). Moreover, since the number of chargers and POI are different kinds of factors, we moved the charger information from the sector to the center of the glyph. Experts appreciated this design as it facilitates comparison of the details of each charging station.

Temporal View

In order to display the prediction information of utilization and to explore the influencing factors that affect the prediction (**R.2 – R.3**), we design the temporal view (Fig. 1C), which can be divided into four parts. Fig. 1C_1 shows the traffic conditions and charging prices; Fig. 1C_2 presents the charging utilization; Fig. 1C_3 displays the feature importance; and Fig. 1C_4 shows the feature overview. In Fig. 1C_1, the X-axis represents the time information, the left Y-axis represents the number of traffic destinations, and the right Y-axis represents the charging prices. We use the green line to indicate the change of traffic destination count over time and the red line to indicate the charging prices over time, and we can easily observe the regular periodicity of the number of traffic destinations and the stable charging prices. In Fig. 1C_2, we use the blue line to encode the ground truth data and the orange line to encode the predicted data, which are used to observe whether the prediction results match the ground truth data. Fig. 1C_3 and C_4 are designed to illustrate the influence of factors and feature importance based on SHAP. In Fig. 1C_4, a rank bar chart is used to provide an overview of feature importance (whatever positive impact or negative impact on utilization), and we can observe that several top features contribute significantly to the prediction result, such as "*Traffic*", "*Price*" and "*Chargers*". Since we consider many features, matching each feature with a separate color can lead to confusion. Therefore, we only assign different colors to features with higher influence and provide further analysis. To better show how the first five important factors affect the forecasts over time, inspired by the design of [25], we design a step line with stacked bars (Fig. 1C_3), where the line denotes the real utilization change and the stacked bars denote the importance of the factors. As shown in Fig. 4, bars above the step line indicate factors that have positive impacts on utilization, while bars below the step line indicate factors that have negative impacts. The more influential the factor is, the higher the height of the bar. When users select a time range in Fig. 1C_2, Fig. 1C_3 will be updated with the corresponding time range.

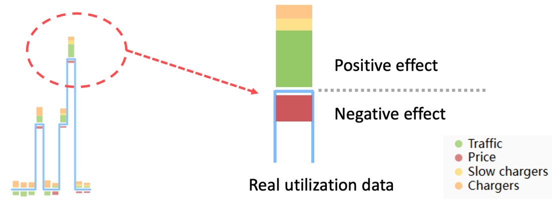


FIGURE 4. Visual design for showing how the top five factors affect utilization prediction. The step line denotes the real utilization and the stacked bars around the step line correspond to the factors affecting utilization at different times.

Planning View

To facilitate the selection of a better deployment strategy, the planning view (Fig. 1D) is designed to allow users to adjust the corresponding factors and compare different results (**R.5**). The planning view contains three parts: the upper part (Fig. 1D_1) and the middle part (Fig. 1D_2) show the control panel for adjusting the factors; the lower part (Fig. 1D_3) displays the prediction results for each deployment strategy. In Fig. 1D_1, users can adjust the number of charging stations and the ratio of fast and slow charging stations. In addition, users can also adjust the charging prices in Fig. 1D_2. After some adjustments are made by users, the final utilization is calculated by the predictive model. Users are allowed to observe the results of different deployment strategies. As shown in Fig. 1D_3, users can compare the previous charging utilization with the adjusted predicted charging utilization. Also, users can observe the corresponding changes after adjustment, such as the number of previous chargers and the number of current chargers, as well as the ratio of fast and slow charging stations.

Evaluation

We used two case studies, interviews with experts, and a user study to demonstrate how *EVCSeer* can help users gain insights into charging station utilization patterns and charging station deployment. These included experts in the collaborative areas mentioned above (E1 and E2) and two other areas (E3 and E4). E1 and E2 were involved in the design process, while E3 and E4 were not. E3 (male, age: 45) is a manager with 10 years of experience in deploying charging stations. E4 (male, age: 34) is an assistant professor whose research interests are mainly intelligent transportation. He has 3 years of research experience in optimizing charging stations. In the interviews, the experts explored the operational status of charging stations and deployment strategies of charging stations with

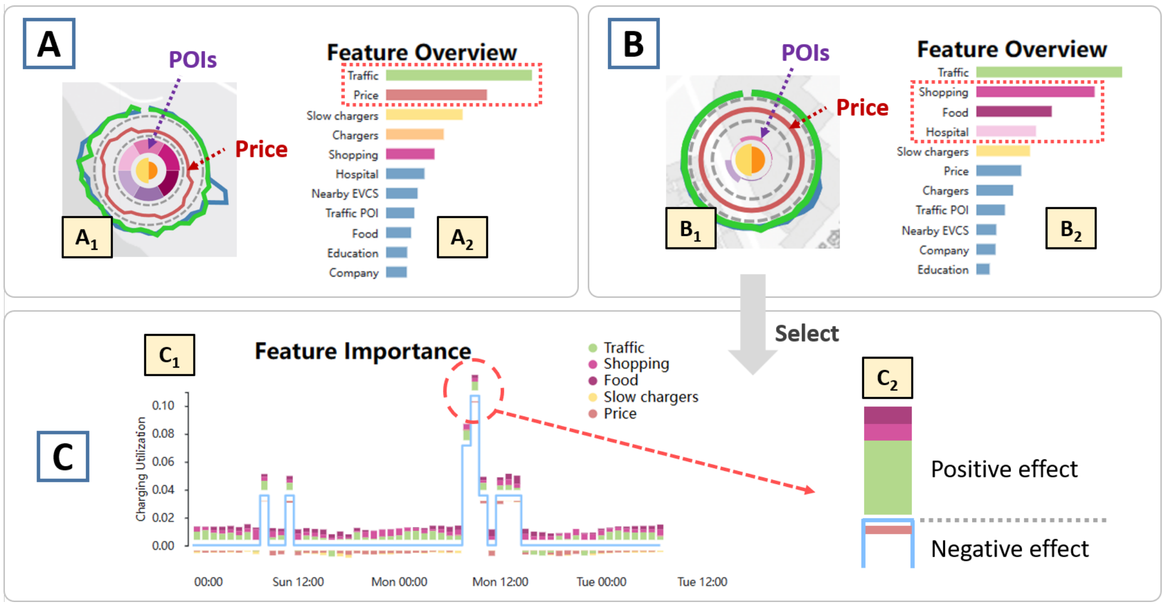


FIGURE 5. Comparison between charging stations with high utilization and low utilization. (A) The high-utilization charging station is surrounded by a relatively rich number and type of POIs, and the charging price has a strong influence on its utilization. (B) The charging station with low utilization is completely idle most of the time, and the number and type of surrounding POIs are not rich enough. (C) The number of shopping and food POIs has a positive effect on the improvement of utilization.

EVCSeer. Here we reported the cases conducted by E1 and E2 respectively, as well as the feedback from all experts. After that, we further reported a user study with 12 participants to evaluate *EVCSeer*'s performance in workflow, visual design and system usability.

Case Study

Case I: Explore Charging Station Operational Status In the first case, E1 wanted to explore the operational status of charging stations, in particular the difference between those that are highly utilized and those that are underutilized.

First, E1 chose a familiar area, the “Futian” district, the data overview updated correspondingly. “*Futian is the center of Shenzhen, with a large number of electric vehicles and a high demand for charging.*”, said E1. He observed Fig. 1A_1 and found a group of charging stations with good traffic and POI conditions nearby, indicated by large green and pick sectors. The opacity of some blue donuts is high, which means higher utilization of these charging stations (highlighted by red dash rectangles in Fig. 1A_1). By clicking on the glyph, the spatial view locates the corresponding location, which is indicated by a blue marker. As shown in Fig. 1B_1, some charging stations have relatively high utilization, which is indicated by the high opacity of the blue circles. Intrigued by those charging stations with high utilization, he zoomed in to see more details. He could then observe the glyphs in Fig. 1B_2. He immediately noticed that the glyphs showed large sectors for different POI (purple sectors) on the right of the map, indicating that these charging stations contain many different POIs around them. “*Charging stations with various POIs are easy to attract people to. For example, people tend to charge their cars while they are shopping and having lunch.*” said E1. In addition, E1 observed an interesting point in the glyphs. Most glyphs in this area have higher utilization on weekdays (blue lines) around 8 a.m. and not on weekends (green lines) around 8 a.m. E1 quickly understood the reason for this. “*People come to work at Futian around* 8 *a.m. Then, they charge their electric vehicles around the office.*”

To further observe the traffic conditions and charging prices, E1 shifted this attention to Fig. 1C_1. He observed a cyclical pattern of traffic conditions (green line) and charging prices (red line). When traffic volumes are high, the charging price tends to be high. As shown by the feature importance in Fig. 1C_3 and C_4, E1 found that some factors, such as the number of chargers, traffic conditions, and charging prices have a significant impact on utilization.

To compare the high-utilization charging stations with the low-utilization charging stations, E1 directly used Fig. 1A_2 to locate the high- or low-utilization charging stations. As shown in Fig. 5, he found two representative charging stations. This high-utilization

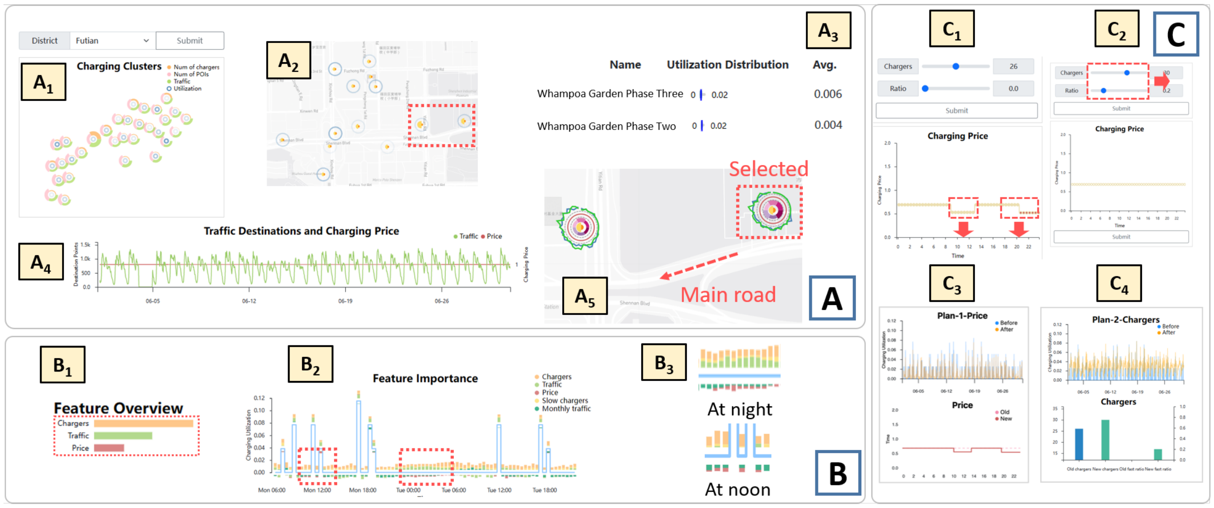


FIGURE 6. "What-if" analysis for improving utilization. (A) Basic information about charging stations. (B) Influencing factors exploration. (C) Modify charging station deployment.

charging station is surrounded by a relatively rich number and type of POIs (Fig. 5A_1). Charging prices also vary over time. Fig. 5A_2 showed that the traffic and price are important features that affect the utilization of this charging station. Conversely, as for the charging station with low utilization in Fig. 5B_1, the area of all purple sectors is small, which demonstrates the number and type of surrounding POIs were not rich enough. The price they charge (dark red line) was also fixed. Fig. 5B_2 also shows that shopping, food and hospitals have strong influences on the utilization of this charging station. If there are not many POIs around shopping and food, it is difficult to attract consumers or taxi drivers to charge here. Then E1 examined the changes in these influencing factors over time. In Fig. 5C_1, the blue step line shows that the low-utilization charging station (Fig. 5B_2) is completely idle most of the time. There are also many purple and magenta bars above the blue step line, which indicates the number of shopping and food POIs has positive effects on the utilization of this charging station. "*It seems that increasing the number of nearby shopping and food POIs can improve its utilization,*" said E1.

In summary, E1 found it interesting to explore the operational status of charging stations with *EVCSeer*. He could easily observe the spatiotemporal patterns of different charging stations. Also, he could easily identify those factors affecting the utilization most. What's more, he thought it is useful to compare different charging stations with high or low utilization.

Case II: Adjust Charging Station Deployment In the second case, E2 wanted to explore how different factors affect the final utilization and further improve it.

E2 also selected the "Futian" district for exploration. When he explored the data overview, he noticed a cluster of charging stations with light blue donuts in Fig. 6A_1, which indicates low utilization. The traffic demand (green sector) and the number of POIs (pink sector) near this charging station cluster are high, which indicates more potential charging demand. E2 then moved the map in the spatial view to the location of this cluster. In the spatial view (Fig. 6A_2), the expert found two charging stations located in the southeast of this area. Although these charging stations have more chargers than other charging stations around them according to the size of the yellow semi-circles, they are not as well utilized, as indicated by the light-stroked blue circles. This finding can also be demonstrated by the boxplot in Fig. 6A_3. The expert zoomed in to get more information. As shown in Fig. 6A_5, there are also many types of POIs (purple sectors) near these two charging stations. E2 noted that these two charging stations are close to a main road. The higher traffic volume may provide more charging vehicles (Fig. 6A_4). All these findings suggest the above stations have the potential to increase utilization. E2 chose the charging station with more POIs (highlighted by a red dash rectangle in Fig. 6A_5) to adjust its deployment strategy.

To understand what factors affect the utilization of this charging station, E2 shifted his attention to the temporal view. As shown in Fig. 6B_1, he found that "Chargers", "Traffic" and "Price" ranked in the top three in the feature overview. Fig. 6B_2 shows how the influencing factors change over time. In particular, the height of the dark red bar below the blue step line in Fig. 6B_3 became higher in the middle of the day and evening, implying that charging price has a higher negative impact on utilization during these times. Therefore, E2 was concerned with adjusting the charging price and the number of chargers.

E2 could easily observe the difference before and after the change in the planning view. The original charging utilization was shown by the blue line and the current prediction charging utilization was shown by the light orange line. E2 tried to reduce the charging prices slightly in Fig. 6C_1 and checked the results. However, the utilization didn't change significantly in Fig. 6C_3. Regarding the number of chargers, this factor ranks top one in Fig. 6B_1. There are many orange bars above the blue step line in Fig. 6B_2, which suggests the number of chargers has a positive effect on the utilization. In addition, E2 explained that the main users of public charging stations in Shenzhen are taxi drivers. "*Taxi drivers prefer shorter charging time.*", he said. For these reasons, he added the number of chargers and turned 20% of the slow chargers into fast chargers (Fig. 6C_2). In this case, the utilization finally improved. The final result can be seen in Fig. 6C_4.

To sum up, E2 believed that *EVCSeer* could help him quickly find interesting charging stations and make deployment adjustments. He could analyze how the influencing factors change over time and compare different deployment strategies, which could provide him with some guidance for adjustments.

Expert Interview

We collected experts feedback from individual interviews with the aforementioned experts (E1 – E4). Each interview lasted approximately one hour. First, we provided participants with a fifteen-minute tutorial to outline the functions, as well as the visual designs and interactions of *EVCSeer*. Then, participants were allowed to explore the system with a think-aloud manner for about thirty minutes. In this process, we provided several tasks to guide their exploration. 1) Find out the spatiotemporal patterns of charging stations (**R.1**). 2) Compare different charging stations (**R.4**). 3) Identify important factors affecting the utilization (**R.3**). 4) Adjust parameters to improve the utilization (**R.5**). After that, we collected and summarized post-study feedback on system designs, usability and suggestions for improvements.

System Designs. All experts appreciated our approach to exploring and optimizing the utilization of EV charging stations. Although E3 and E4 were not involved in the design process, they could easily understand the visual designs. Both E1 and E2 mentioned that the existing tools are either not tailored to the charging station operation scenarios, or have relatively simple functionality and lack exploration of the reasons for charging station utilization and comparison of different charging stations. In addition, E3 said that "*the system can give some reasonable guidance on charging station operation decisions and could help us understand the reasons for changes in charging station utilization.*" E4 mentioned that the visual design is different from traditional charts and is relatively novel and covers more information. He said that "*map glyphs can be a good representation of changes in utilization rates and charging prices over time.*"

Usability and Suggestions. All experts agreed that the data overview, spatial view and planning view are easy to understand and use. As for the design of Fig. 1C_3 in the temporal view, E3 mentioned that it was easy to understand but took some effort to master. He considered explaining the factors that affect the final utilization as an advanced feature. E2 suggested that the system can take into account more factors, such as land use and power grid, as well as different charging stations for different types of vehicles. Further, E3 said that "*the construction of charging stations is also an investment. I hope to include some calculation of the return on investment in the system*". E4 offered some improvements regarding interaction. He said that "*when a charging station is clicked on, a floating window can be displayed, showing basic information about that charging station. It is also helpful to add search function based on criteria.*"

User Study

We also conducted a user study to evaluate *EVCSeer*'s performance in terms of *workflow efficiency*, *visual design*, and *system usability*.

Participants. We recruited 12 participants (5 females, 7 males, age_{mean}=23.5, age_{std}=2.05) from a local university through word-of-mouth method. The participants all have background knowledge of urban transportation and have many years of study and research experience in the field of traffic engineering. They have a basic understanding of simple visualization techniques. Four of the participants have conducted relevant research on the charging station location problem or charging station utilization prediction and are aware of the state of research in the corresponding fields. Although they are not the direct target users of our system, we chose them because they could provide us with more comprehensible insights and help validate the system usability.

Tasks. Participants were asked to sequentially complete the following two tasks, which guided them through the pipeline of *EVCSeer*. **Task I** was to verify

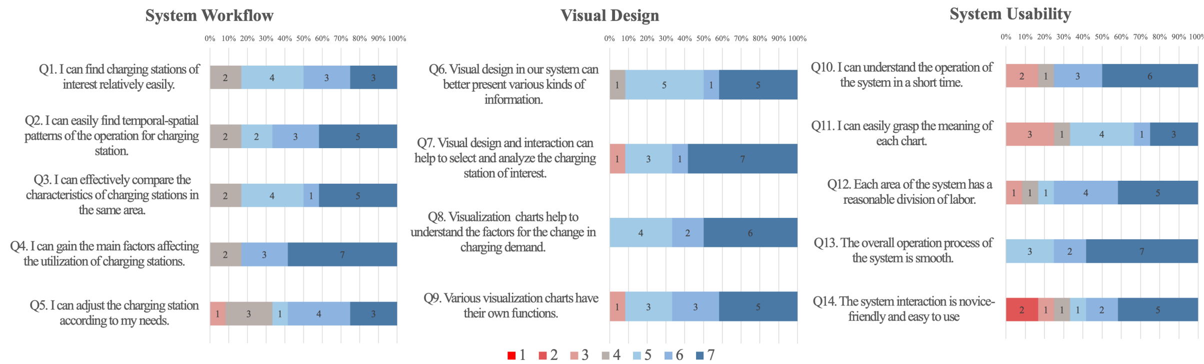


FIGURE 7. Result of user study. 1-7 represents "strongly disagree" to "strongly agree" for each statement.

the rationality of the workflow and test whether the visual design benefits the representation of information. In this task, participants were asked to find out the spatiotemporal patterns of charging stations and compare different charging stations. **Task II** was to examine whether *EVCSeer* could help users develop a new deployment strategy. Participants were asked to identify important factors affecting the utilization and then adjust parameters to improve the utilization.

Procedure. First, we introduced the overall design ideas, workflow and framework of the system to the participants. After that, the system was demonstrated in around 10-15 minutes with a detailed demonstration of the functional and visual design of each view. We then allowed them to pre-operate the system and ask us questions for another 5 minutes. When participants were ready, they were asked to perform the tasks described above using our system. Participants were also asked to think aloud about their ideas while performing all tasks. After completing all tasks, participants had to complete a questionnaire containing 7-point Likert scale questions, with scores ranging from 1-7, with increasing levels of approval.

Results. The result of the questionnaire can be found in Fig. 7. In general, participants rated *EVCSeer* well. The average score of the workflow, visual design and system usability was above 5.5. For the smoothness and rationality of the workflow, participants generally agreed that they could easily find some spatial and temporal patterns of different charging stations and compare the characteristics of various charging stations in the same area. In terms of visual design, the system organized the various information about the charging stations in a better way. To some extent, the different views helped users to quickly find the charging stations they were interested in. Some charts (e.g., the temporal view and spatial view) helped participants understand what influences the use of charging stations. Two participants felt that the map was a bit cluttered when showing multiple charging stations. They suggested that the system should quickly compare the differences between different charging stations in the same area. For the system usability, after a short training, participants generally understood the overall design idea of the system and the main functions of different modules. In general, they could easily grasp the meaning of each visual design. The overall operation of the system was also smooth enough. Some participants would like the system to be more customizable for charging stations and would also like more guidance for new users.

Discussion and Limitations

In this section, we discuss the high-level synthetics and limitations of our approach.

Influencing Factors. In our study, we make full use of the available data to explore the factors associated with charging station utilization. Considering the temporal factors, spatial geographic information (e.g., POI and traffic) and attributes of charging stations (e.g., the number of chargers, charging price, competition quantity), *EVCSeer* can predict the utilization well and is recognized by experts. However, due to the limitations of the platform, it is difficult for us to obtain the duration of land use and power consumption capacity that experts are interested in. Therefore, the relationship between more factors and utilization needs to be further explored.

Generalizability and Scalability. In our study, *EVCSeer* focuses on analyzing the influencing factors of charging station utilization and siting optimization. Although different factors need to be considered for charging stations for different vehicles (e.g. electric buses, logistics vehicles, sanitation vehicles), our analysis pipeline and strategy are generic and can be easily applied. It can also be applied to other application scenarios, such as the analysis of electricity consumption forecasting. *EVCSeer* can help managers predict electricity consumption peaks and analyze the influencing factors according to the characteristics of residential areas. As for scalability, there exist visual clutters in the charging clusters (Fig. 1A_1) when analyzing a large number of charging stations. Also, *EVCSeer* uses too many colors to encode different factors. In the future, we will adopt filtering techniques to filter charging stations and explore other encodings for influencing factors.

Limitations. First, *EVCSeer* has only been evaluated by 4 experts and 12 users. Discussions with more domain experts are needed to better evaluate the usability and effectiveness of our system. Second, the datasets we used in *EVCSeer* were collected at different time. The varying time ranges of the datasets may affect the analysis result. Third, currently, we chose *XGBoost* as our final model because of its better performance. We realize that there is still room for improving the utilization prediction. For example, we did not explicitly consider the relationship between the factors in the model. In the future, we will work to obtain higher quality data and optimize our prediction model to better represent the relationship between factors and utilization rates.

Conclusion and Future Work

In this study, we present an interactive visual analytics system, *EVCSeer* to facilitate the evaluation of charging station operations and optimize charging station utilization. We identify a set of measurable influences (i.e., temporal variation in utilization, spatial geographic information, and attributes of charging stations) and employ representative predictive models to explore how these factors affect charging station utilization. In addition, elaborate visualizations are presented to explore spatiotemporal patterns of utilization data and support "what-if" analysis of charging station deployments. Two case studies, expert interviews, and a qualitative user study demonstrate the validity and usefulness of *EVCSeer*.

In the future, we plan to collect more data and consider more influencing factors, such as land factor and service area population, to improve system usability. Also, we want to improve the performance of the prediction model. In addition, we will conduct a long-term study with end users to evaluate the usability and effectiveness of *EVCSeer*.

ACKNOWLEDGMENT

This work is supported by grants from the National Natural Science Foundation of China (No. 62302531) and the Science and Technology Planning Project of Guangdong Province (No. 2023B1212060029).

REFERENCES

1. Lennart Adenaw and Sebastian Krapf. Placing bev charging infrastructure: Influencing factors, metrics, and their influence on observed charger utilization. *World Electric Vehicle Journal*, 13(4):56, 2022.
2. Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proc.SIGKDD*, pages 785–794, 2016.
3. Sanchari Deb, Kari Tammi, Karuna Kalita, and Pinakeswar Mahanta. Review of recent trends in charging infrastructure planning for electric vehicles. *Wiley Interdisciplinary Reviews: Energy and Environment*, 7(6):e306, 2018.
4. Guanpeng Dong, Jing Ma, Ran Wei, and Jonathan Haycox. Electric vehicle charging point placement optimisation by exploiting spatial statistics and maximal coverage location models. *Transportation Research Part D: Transport and Environment*, 67:77–88, 2019.
5. Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc.KDD*, volume 96, pages 226–231, 1996.
6. Mario Gellrich, Andreas Block, and Ninja Leikert-Böhm. Spatial and temporal patterns of electric vehicle charging station utilization: a nationwide case study of switzerland. *Environmental Research: Infrastructure and Sustainability*, 2022.
7. The White House. Fact sheet: Biden-harris administration catalyzes more than $700 million in private sector commitments to make ev charging more affordable and accessible, 2022.
8. Tim Jonas, Noah Daniels, and Gretchen Macht. Electric vehicle user behavior: An analysis of charging station utilization in canada. *Energies*, 16(4):1592, 2023.
9. Dmytro Karamshuk, Anastasios Noulas, Salvatore Scellato, Vincenzo Nicosia, and Cecilia Mascolo. Geo-spotting: mining online location-based services for optimal retail store placement. In *Proc.SIGKDD*, pages 793–801, 2013.
10. Mouna Kchaou-Boujelben. Charging station location problem: A comprehensive review on models and solution approaches. *Transportation Research Part C: Emerging Technologies*, 132:103376, 2021.
11. Quan Li, QQ Liu, CF Tang, ZW Li, SC Wei, XR Peng, MH Zheng, TJ Chen, and Qiang Yang. Warehouse vis: A visual analytics approach to facilitating warehouse location selection for business

districts. In *Computer Graphics Forum*, volume 39, pages 483–495. Wiley Online Library, 2020.
12. Yuhong Li, Yu Zheng, Shenggong Ji, Wenjun Wang, Leong Hou U, and Zhiguo Gong. Location selection for ambulance stations: a data-driven approach. In *Proc.SIGSPATIAL*, pages 1–4, 2015.
13. Dongyu Liu, Di Weng, Yuhong Li, Jie Bao, Yu Zheng, Huamin Qu, and Yingcai Wu. Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):1–10, 2016.
14. Mustafa Lokhandwala and Hua Cai. Siting charging stations for electric vehicle adoption in shared autonomous fleets. *Transportation Research Part D: Transport and Environment*, 80:102231, 2020.
15. Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.
16. Simone Maase, Xiomara Dilrosun, Martijn Kooi, and Robert Van den Hoed. Performance of electric vehicle charging infrastructure: Development of an assessment platform based on charging data. *World Electric Vehicle Journal*, 9(2):25, 2018.
17. Marc-Olivier Metais, O Jouini, Yannick Perez, Jaâfar Berrada, and Emilia Suomalainen. Too much or not enough? planning electric vehicle charging infrastructure: A review of modeling options. *Renewable and Sustainable Energy Reviews*, 153:111719, 2022.
18. Easwaran Narassimhan and Caley Johnson. The role of demand-side incentives and charging infrastructure on plug-in electric vehicle adoption: analysis of us states. *Environmental Research Letters*, 13(7):074032, 2018.
19. Michael Nicholas. Estimating electric vehicle charging infrastructure costs across major us metropolitan areas. *Int. Counc. Clean Transp*, 14(11), 2019.
20. Tim Unterluggauer, Jeppe Rich, Peter Bach Andersen, and Seyedmostafa Hashemi. Electric vehicle charging infrastructure planning for integrated transportation and power distribution networks: A review. *ETransportation*, 12:100163, 2022.
21. Di Weng, Heming Zhu, Jie Bao, Yu Zheng, and Yingcai Wu. Homefinder revisited: Finding ideal homes with reachability-centric multi-criteria decision making. In *Proc.CHI*, pages 1–12, 2018.
22. Rick Wolbertus, Maarten Kroesen, Robert Van Den Hoed, and Caspar Chorus. Fully charged: An empirical study into the factors that influence connection times at ev-charging stations. *Energy Policy*, 123:1–7, 2018.
23. Zuzhao Ye, Ran Wei, and Nanpeng Yu. Short-term forecasting for utilization rates of electric vehicle charging stations. In *Proc.ISC2*, pages 1–7. IEEE, 2021.
24. Zhiyan Yi, Xiaoyue Cathy Liu, and Ran Wei. Electric vehicle demand estimation and charging station allocation using urban informatics. *Transportation Research Part D: Transport and Environment*, 106:103264, 2022.
25. Chenyang Zhang, Xiyuan Wang, Chuyi Zhao, Yijing Ren, Tianyu Zhang, Zhenhui Peng, Xiaomeng Fan, and Quan Li. Promotionlens: Inspecting promotion strategies of online e-commerce via visual analytics. *arXiv preprint arXiv:2208.01404*, 2022.
26. Yang Zhao, Zhenpo Wang, Zuo-Jun Max Shen, and Fengchun Sun. Data-driven framework for large-scale prediction of charging energy in electric vehicles. *Applied Energy*, 282:116175, 2021.

Yutian Zhang is a master student in School of Intelligent Systems Engineering at Sun Yat-sen University (SYSU). His research interests include interpretable machine learning, visual analytics and transportation big data. Zhang received his B.S. degree in transportation engineering from SYSU. Contact him at zhangyt85@mail2.sysu.edu.cn.

Shuxian Gu is a master student in School of Intelligent Systems Engineering at Sun Yat-sen University (SYSU). Her research interests include intelligent transportation, visualization, and urban visual analytics. Gu received her B.S. degree in transportation engineering from SYSU. Contact her at gushx3@mail2.sysu.edu.cn.

Quan Li is currently a tenure-track assistant professor at School of Information Science and Technology, ShanghaiTech University. His research interests include visual analytics, interpretable machine learning, and human-computer interaction. Li received his Ph.D. in Computer Science from the Hong Kong University of Science and Technology. Contact him at https://faculty.sist.shanghaitech.edu.cn/liquan/.

Haipeng Zeng is currently an associate professor in School of Intelligent Systems Engineering at Sun Yat-sen University (SYSU). His research interests include data visualization, visual analytics and intelligent transportation. Zeng received his Ph.D. in Computer Science from the Hong Kong University of Science and Technology. Contact him at http://www.zenghp.org/.