Use Rate Prediction For Charging Stations

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

***Abstract*—This paper provides a time frame based prediction model for use rate of charging stations in Shanghai. The approach proposed in this paper takes both station's geographical information and working elements including price and type, into consideration, aimed to classify whether it's a high-use-rate station or a low-use-rate one during different time periods. Experimental results show that our method performs well on LR, Random Forest, SVM and XGBOOST.**

***Index Terms*—time frame, use rate, charging station, geographical information, POIs**

I. Introduction

An electric vehicle charging station, also called EV charging station is an element in an infrastructure that supplies electric energy for the recharging of electric vehicles, such as plug-in electric vehicles and plug-in hybrids. At home or work, some electric vehicles have onboard converters that can plug into a standard electrical outlet or a high-capacity appliance outlet.

However, in most cases, others require a charging station that provides electrical conversion, monitoring, or safety functionality. These stations are also needed when travelling, and many support faster charging at higher voltages and currents than are available from residential EVSEs. Public charging stations are typically on-street facilities provided by electric utility companies or located at retail shopping centers and operated by many private companies.

EVs in China is experiencing an overall growth in the past few years, which directly results in the massive construction and modification of charging stations and current road networks. However, the cost of construction of charging stations is considerable and are often very costly or even impracticable to reallocate. This raise the question of how to select the locations for building the charging stations.

In a typical view, a charging station no matter where it is located, its 'success' is often determined by the use rate of a station.

In order to help with station setting strategies, we explore a time frame based prediction model to classify stations into high use-rate and low-use-rate in various time frames, using features like georaphical information as well as other working elements.

We evaluate our method on RL, Random Forest, SVM and XGBOOST, which achieve relatively satisfactory results.

II. Problem Definition

It's a binary classification problem. The aim is to predict that one station is of high use rate or low use rate according to its georaphical information and working elements during different time periods.

III. Related Work

Hence, we studied the optimization problem of how to deploy charging stations. Existing works [1] mainly falls into the domain of bike-sharing. Current works of location selection are usually based on the flow prediction of a single station. Futhermore, they rely heavily on the historical data. We argue that the surrounding point of interests as well as distances to important POIs(e.g, metro stations, estates, etc.) play an important role in selecting the optimal location for stations.

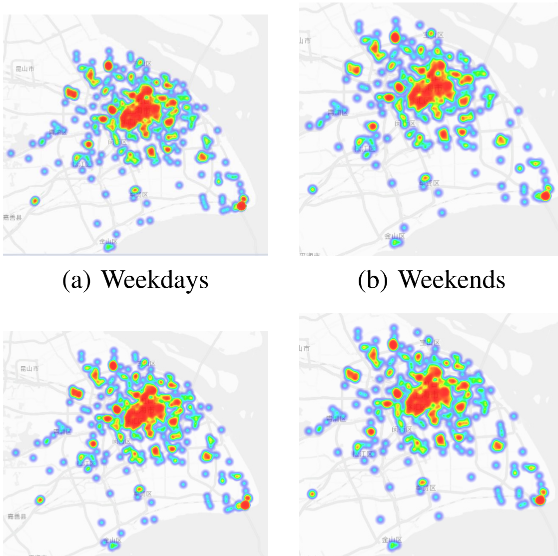
IV. Feature Description

A. Feature Extraction

From the data we gathered, we made several observations that benefits the features to be included in the model we present. We separate the data into different time frames in order to determine the overall difference among them, since charging network is a dynamic system. From Fig.1, it is easy to notice that in different time frames, the charging hotspots stays at almost the same locations, meaning that during different time periods, the use rate of a given charging station is determined by its spatiotemporal context.

B. Point Of Interest

To the prosperity of Shanghai, there are so many points of interest(e.g., shopping malls, schools. estates, companies, etc.) located in the city. Understanding the purpose of the trip by



(a) Weekdays (b) Weekends

(c) Mornings (d) Evenings

Fig. 1. Charging Hotspots in Shanghai in different time frames

each person who participates in charging system will help us to analyze the use rate prediction. So, we need to extract the POI around each existing charging station. There are a lot of POIs in Shanghai. The POIs are very close to each other. We set a radius around each station and then collect the POIs within the radius. Based on our experiment, choosing a radius as 300 meters is proper. In our work, we get 80 different types of POI totally. Some of the POIs are very similar to each other. Therefore, we group 80 POIs in further step. By grouping POIs, we get 10 groups at last. Table.I gives the groups and POIs in detail.

TABLE I
GROUPS OF POIS

Group	Points of interests
Food	chinese & foreign restaurant, snack bar, cake & dessert shop, cafe, bar
Hotels	star hotel, express hotel, apartment hotel
Shopping	shopping centers, department stores, supermarkets, convenience stores, home building materials, home appliances digital, shops, markets
Education	institutions of higher learning, middle schools, primary schools, kindergartens, adult education, parent-child education, special education schools, study agencies, research institutions, training institutions, libraries, science and technology museums
Cultural venue	press and publication, radio and television, art groups, art galleries, exhibition halls, cultural palaces
Medical	general hospitals, specialist hospitals, clinics, pharmacies, medical examination institutions, nursing homes, emergency centers, disease control centers
Car service	car sales, car repair, car beauty, auto parts, car rental, car inspection field
Transportation	airport, railway station, subway station, subway line, long-distance bus station, bus station, bus line, port, parking lot, refueling station, service area, toll station, bridge, charging station, roadside parking space
Estates	Office building, residential area, dormitory

C. Distance

In use rate prediction, we need to consider distance. People will not choose to park their electric cars for charging if the destination they planned to go is far away. In the system, a station with nearer distance to metro stations, financial centers and major functional buildings would easily be used more often. We select the nearest distance to the following to be the distance we considered as features: company, estate, hospital, metro station, shopping center and university.

D. Price

By futher digging into the data, we find that there is a 0.3 correlation between price for charging and the use rate. Since there are two types of charging ports: DC and AC. We would include the number of ports and the price of both types in a charging station as one of its feature.

E. Private or public

By observing the data we've collected, it can be seen that most of the charging stations are private charging stations, which means they are typically used by electric public buses and rent cars, or used by specific companies for their employees, accounting for over 70% of the total charging stations. Also, since the private are used by more regular users(e.g., buses, companies employees), its use rate are 5% higher compared to public ones. This alongside other observations will be taken into consideration.

V. TIME FRAME BASED MODEL

A. Time Frames

We separate the dataset into different time frames, including total time, weekday, weekend, daytime, evening time, morning_rush hours, evening_rush hours and travel_hours.Fig.2 shows the average use-rate of the time frames above.

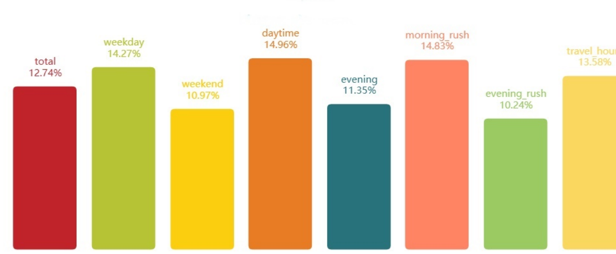


Fig. 2. Average use-rate of different time frames

B. Model Description

In each time frame, we add the geographical information as well as working elements into the prediction model, in order to predict whether it's a high use_rate station or a low use_rate one.

VI. EXPERIMENT

A. Datasets

In this paper, we gather the charging stations logs from all the existing charging station companies who provide their services in Shanghai, with a total of over 2,000,000 lines. The log has a length of one month, from 2018/10 to 2018/11, in which an hourly summarize of each charging station is recorded, showing whether it's occupied or not. After preprocessing, the training set has 80% of valid data, the test set has 20% left.

B. Implementation

With features extracted from our collected dataset, there are mainly three groups of them, including POIs, charing price and station types, that are added into the time frame based models. Fig.3 shows the general flow path of our work.

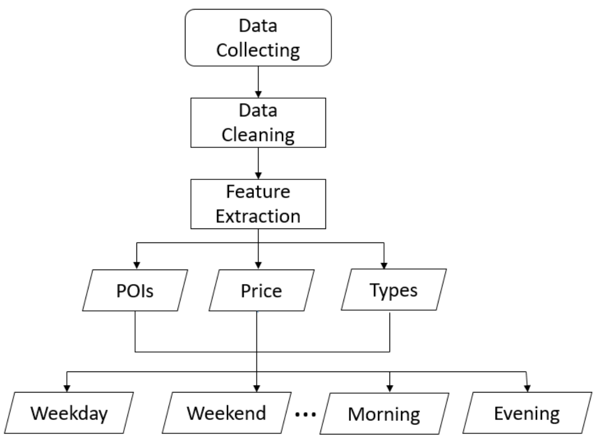


Fig. 3. Implementation of the work

C. Results

We run our dataset on LR, Random Forest, SVM and XGBOOST respectively, along with the features mentioned above added into time frame based models. Table.II shows the prediction accuracy of different models. We can see that they can all perform well based on our settings, and XGBOOST achieves the most favourable result.

TABLE II
EVALUATION RESULTS ON DIFFERENT MODELS

Model	Accuracy
LR	60.9%
SVMt	68.9%
Random Forest	70.1%
XGBOOST	73.6%

VII. CONCLUSION AND FUTURE WORK

In this paper, we propose the time frame based prediction model for use rate of charging stations. We study on some important features like station's surrouding POIs, charing price and station type. In experiments, we separate our dataset into different time frames and add the features into them, which obtains a relatively favourable result, indicating that the use rate of charing station is highly influenced by its geographical information and working elements. Furthermore, it also has a greate impact on location choosing problem.

There is still a lot of work to be continued in the future. First, we only consider three mainly types of features that might affect station use rate, many other important features also need to be extracted and incluede. Furthermore, we might explore a properly modified model to execute those features and gain a more satisfactory result.

REFERENCES

[1] Jie Bao, Tianfu He, Sijie Ruan, Yanhua Li, and Yu Zheng. Planning bike lanes based on sharing-bikes' trajectories. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*, pages 1377–1386, 2017.

[2] Minh X. Hoang, Yu Zheng, and Ambuj K. Singh. FCCF: forecasting citywide crowd flows based on big data. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS 2016, Burlingame, California, USA, October 31 - November 3, 2016*, pages 6:1–6:10, 2016.

[3] Yaguang Li, Kun Fu, Zheng Wang, Cyrus Shahabi, Jieping Ye, and Yan Liu. Multi-task representation learning for travel time estimation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 1695–1704, 2018.

[4] Yexin Li, Yu Zheng, and Qiang Yang. Dynamic bike reposition: A spatio-temporal reinforcement learning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 1724–1733, 2018.

[5] Binbing Liao, Jingqing Zhang, Chao Wu, Douglas McIlwraith, Tong Chen, Shengwen Yang, Yike Guo, and Fei Wu. Deep sequence learning with auxiliary information for traffic prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 537–546, 2018.

[6] Xinyue Liu, Xiangnan Kong, and Yanhua Li. Collective traffic prediction with partially observed traffic history using location-based social media. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016, Indianapolis, IN, USA, October 24-28, 2016*, pages 2179–2184, 2016.

[7] Zhaoyang Liu, Yanyan Shen, and Yanmin Zhu. Where will dockless shared bikes be stacked?: - parking hotspots detection in a new city. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 566–575, 2018.

[8] Bilong Shen, Xiaodan Liang, Yufeng Ouyang, Miaofeng Liu, Weimin Zheng, and Kathleen M. Carley. Stepdeep: A novel spatial-temporal mobility event prediction framework based on deep neural network. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 724–733, 2018.

[9] Zidong Yang, Ji Hu, Yuanchao Shu, Peng Cheng, Jiming Chen, and Thomas Moscibroda. Mobility modeling and prediction in bike-sharing systems. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys 2016, Singapore, June 26-30, 2016*, pages 165–178, 2016.

[10] Xiao Zhou, Anastasios Noulas, Cecilia Mascolo, and Zhongxiang Zhao. Discovering latent patterns of urban cultural interactions in wechat for modern city planning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pages 1069–1078, 2018.