Data-driven prediction of fine-grained EV charging behaviors in public charging stations

Fuli Qiao
Shanghai Jiao Tong University
qiaoflii@sjtu.edu.cn

Shan Lin
Stony Brook University
Shan.X.Lin@stonybrook.edu

ABSTRACT

With the rapid growth of electrical vehicle public charging stations, accurate predictions of local charging demand enable many prospective applications. In this paper, we explore a data-driven approach to predict future charging demand, and build predictive models to characterize behaviors of both registered long-term users and unregistered short-term users. With a real-world dataset of 28053 records over 798 days at multiple locations, evaluation results demonstrate that our model with XGBoost outperforms existing solutions, reducing the prediction error up to 40.8% at the finest time granularity (15-minute interval).

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**.

KEYWORDS

EV charging prediction, user behaviors, machine learning

ACM Reference Format:
Fuli Qiao and Shan Lin. 2021. Data-driven prediction of fine-grained EV charging behaviors in public charging stations. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Prediction of charging demand is very important for both electric vehicle (EV) drivers and charging station owners to manage and optimize their operations[4]. However, a main challenge is how to understand the aggregated charging behaviors of both registered users and unregistered users in the system[2]. The registered users tend to use the system for longer terms; whereas the unregistered users typically are short-term users who use charging stations occasionally. To address this issue, we design a data-driven approach to predict usage patterns of EV charging piles at public stations. Our prediction design includes one predictive model for registered users and another one for unregistered users. Working on a real-world charging record dataset collected in Caltech[3], we apply supervised learning based algorithms, specifically XGBoost, Support Vector Regression (SVR), and Gradient Boost Decision Tree (GBDT), to predict sequences of future availability.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 EV CHARGING DEMAND PREDICTION MODELS

Goal: Obtain a function f that inputs a sequence of the occupied number of charging piles in a previous continuous time period $OP_{TG,M} = \{o_1, o_2, ..., o_m\}$ and outputs a sequence of the available number of charging piles in a future continuous time period $AP_{TG,N} = \{a_1, a_2, ..., a_n\}$. Here, TG presents time granularity. The element in $OP_{TG,M}$ and $AP_{TG,N}$ respectively mean the occupied and available number of charging piles at the time divided by TG. Besides, M and N present the number of TG.

$$f(OP_{TG,M}) = AP_{TG,N} \tag{1}$$

Considering different charging patterns of registered users R and unregistered users U, the improved function is as follows.

$$f(OP_{TG,M}) = f_{re}(OP^R_{TG,M}) + f_{unre}(OP^U_{TG,M}) \tag{2}$$

Method: As Figure 1 shown, for registered/unregistered users, use XGBoost to train two predictive models separately and then combine middle prediction results to obtain final prediction results.

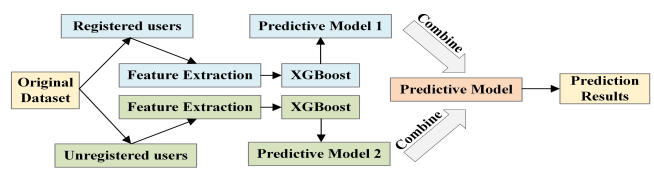

Figure 1: Proposed predictive model

Data processing: Calculate the availability (unavailability: 0, availability: 1) of each charging pile to obtain total occupied number in a certain period. Figure 2 shows data samples of unregistered/registered users from 20:00 to 24:00 of a day. In order to predict fine-grained, we divide the time into 15, 30, 45, 60, 80 minutes.

Prediction process: To predict $AP_{TG,S-M} = \{a_{m+1}, a_{m+2}, ..., a_s\}$, use $OP_{TG,M} = \{o_1, o_2 ..., o_m\}$. First, use $\{o_1, o_2, ..., o_m\}$ to fit a function and get $\{a_{m+1}\}$. Then, use $\{o_2, o_3, ..., a_{m+1}\}$ and get $\{a_{m+2}\}$. Conduct this multi-step iterative process until obtaining $\{a_s\}$.

Prediction results: As Figure 1 shown, for registered users, use predictive model 1 to get prediction results:

$$AP^R_{TG,N} = \{R_1, R_2, ..., R_N\} \tag{3}$$

Use predictive model 2 to obtain prediction results for unregistered users:

$$AP^U_{TG,N} = \{U_1, U_2, ..., U_N\} \tag{4}$$

Finally, combine these two prediction results to get the available number of EV charging piles in a future time period:

$$AP_{TG} = AP^R_{TG,N} + AP^U_{TG,N} = \{R_1 + U_1, ..., R_N + U_N\} \tag{5}$$