This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2021.3070202, IEEE Transactions on Knowledge and Data Engineering



# Periodic Weather-Aware LSTM with Event Mechanism for Parking Behavior Prediction

Feng Zhang, Yani Liu, Ningxuan Feng, Cheng Yang, Jidong Zhai, Shuhao Zhang, Bingsheng He, Jiazao Lin, Xiao Zhang, Xiaoyong Du

**Abstract**—There are plenty of parking spaces in big cities, but we often find nowhere to park. For example, New York has 1.4 million cars and 4.4 million on-street parking spaces, but it is still not easy to find a parking place near our destination, especially during peak hours. The reason is the lack of prediction of parking behavior. If we could provide parking behavior in advance, we can ease this parking problem that affects human well-being. We observe that parking lots have periodic parking patterns, which is an important factor for parking behavior prediction. Unfortunately, existing work ignores such periodic parking patterns in parking behavior prediction, and thus incurs low accuracy. To solve this problem, we propose PewLSTM, a novel periodic weather-aware LSTM model that successfully predicts the parking behavior based on historical records, weather, environments, weekdays, and events. PewLSTM includes a periodic weather-aware LSTM prediction module and an event prediction module, for predicting parking behaviors in regular days and events. PewLSTM is extremely useful for drivers and parking lot owners to improve customer experience. For example, the probability of parking space that will be available soon can be provided even if the parking lot is full. Based on 910,477 real parking records in 904 days from 13 parking lots, PewLSTM yields 93.84% parking prediction accuracy, which is about 30% higher than the state-of-the-art parking behavior prediction method. We have also analyzed parking behaviors in events like holidays and COVID-19. PewLSTM can also handle parking behavior prediction in events and reaches 90.68% accuracy.

## 1 Introduction

We often find it hard to find a parking space in big cities. However, in building cities, designers usually reserve enough parking spaces. For example, New York has at least 4.4 million on-street parking spaces [1], but has only 1.4 million cars [2]. The reason why it is not easy to find a parking place, especially during peak hours, is the lack of prediction of parking behavior. In this paper, for a parking lot, we express parking behavior as the number of parking arrivals and departures in one hour, which can also be converted into the number of available parking spaces. Clearly, if we could provide the parking behavior in advance, we can ease such a parking problem that affects human well-being.

AI technology has transformed our everyday lives in many aspects, and now AI is shaping our parking behavior. More intelligent parking lots with AI management systems appear in our cities. Among the intelligent parking management systems, providing the parking behavior prediction is the most critical and beneficial function. First, accurate predictions can save us time and effort to find a proper parking space. For example, a motorist in New York spends an average of 107 hours per year on finding a parking space [3], while such time waste could be saved when we obtain the parking behavior for the expected parking lots in advance. Second, based on the precise parking behavior prediction, the parking lots can be utilized more efficiently, and parking space reservations can also be made based on the prediction. Third, based on the parking behavior prediction, the department of transportation is better able to understand the traffic conditions and dispatch resources such as traffic police. However, parking behavior is hard to predict, since it can be affected by many factors, such as weather, weekdays, and environments.

Many intelligent parking methods and systems have been developed in recent years, such as pricing parking [4], on-street parking [5], smart vehicle parking [6], and smart parking guidance [7]. However, none of these works involve parking behavior prediction. A real-time parking space availability system [8] has been developed, but this work only considers historical records. The state-of-the-art parking behavior prediction research [9] used regression methods regarding parking records and weather, but we find that their method only considers parking arrivals without departures, and ignores the periodic parking patterns and thus incurs low accuracy. Even worse, they can only predict parking behavior in one hour, while we are looking forward to a long-term prediction.

Building an accurate long-term parking behavior prediction system requires overcoming three challenges. The first challenge is to find an appropriate model that can understand the complicated parking-related behaviors by using historical records and other factors such as weather and time information. The second challenge is how to involve the complex periodic parking patterns, including the influence


- F. Zhang, Y. Liu, X. Zhang, and X. Du are with the Key Laboratory of Data Engineering and Knowledge Engineering (Renmin University of China), Ministry of Education, and with the School of Information, Renmin University of China, Beijing 100872, China.
- N. Feng and J. Zhai are with the Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China.
- C. Yang is with the School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, 100876, China.
- S. Zhang is with the Database Systems and Information Management Group (DIMA), Technische Universität Berlin, Berlin 10623, Germany.
- B. He is with the School of Computing, National University of Singapore, 119077, Singapore.
- J. Lin is with the Department of Information Management, Peking University, Beijing, 100871, China.

from a specific time in the past. For example, customers would like to drive to a shopping mall every Friday night for discounts, and such patterns should be integrated. The third challenge is how to predict parking behavior for a long time in the future with the current data. The fourth challenge is how to predict parking behavior in events.

We propose **PewLSTM**, a **pe**riodic **w**eather-aware **LSTM** model for long-term car parking behavior prediction. PewLSTM is extremely useful for drivers and parking lot owners to improve customer experience. For example, the probability of parking space that will be available soon can be provided even if the parking lot is full. Our preliminary work has been presented in [10], which supports only daily parking behavior prediction, and we extend it to support parking behavior prediction in events in this work. In contrast to [10], we develop novel technologies for PewLSTM to predict parking behaviors in events, because large-scale events, such as festival activities and COVID-19 [11], also have a great impact on parking behaviors. We analyze regular events such as new year, and unexpected events such as COVID-19 in this paper. When impactful event happens, PewLSTM switches to the new event prediction module for parking behavior prediction in events. Moreover, we add new optimizations on data preprocessing, accuracy tuning, and performance acceleration to improve both accuracy and performance, and conduct new experiments on PewLSTM to verify the performance benefits from our new optimizations with larger datasets.

Our system consists of four parts: an input layer with a classifier, a deep PewLSTM module, an event prediction module, and an output layer integrated with a linear layer. First, to depict the parking behavior for a parking lot, we use both the arrivals and departures, because these two indicators depict changes in parking vehicles. Second, to involve weather information, we add input gates for weather information into the LSTM block. The cell state and hidden state are decided by both previous states and weather information. Third, to utilize the periodic parking patterns from historical records, we add special gates into the LSTM block to input the hidden states from past specific time steps when certain conditions are met, such as the same time in previous weekdays. Fourth, to predict parking behavior in events, we build a polynomial model to analyze the events in history. For parking behavior prediction in events, we find the patterns in history, perform polynomial model fitting, and then model adaptation.

Our proposed periodic weather-aware LSTM with event mechanism, PewLSTM, has been integrated into a real parking system, ThsParking [12] (developed by Huaching Tech[1]), which is one of the top smart parking platforms in China. Based on the prediction, the parking system can launch a segmented pricing strategy to utilize parking spaces better and gain more profit.

We evaluate PewLSTM with real parking records. We collect 910,477 parking records from 13 parking lots. These records come from different environments, including hotels, shopping malls, and streets. Along with these parking records, we have collected related weather and date information. When predicting parking behavior in the next hour, PewLSTM achieves 93.84% arrival accuracy and 93.34% departure accuracy. For the next two hours, PewLSTM achieves 76.78% arrival accuracy and 73.45% departure accuracy. For the next three hours, its arrival accuracy is 67.54% and its departure accuracy is 67.48%. On average, PewLSTM achieves 30% higher accuracy than the state-of-the-art parking prediction method [9].

We further evaluate PewLSTM in supporting parking behavior prediction in events. We have analyzed the parking behaviors in regular events like holidays and unexpected events like coronavirus pandemic (COVID-19) [11]. We develop solutions to identify events based on historical records and perform event prediction for parking behaviors in the future. In our experiments, we measure the prediction accuracy at day granularity for the event one to three months ahead. For the one months ahead prediction, PewLSTM achieves 89.52% accuracy. For the two months ahead prediction, PewLSTM achieves 90.68% accuracy. For the three months ahead prediction, PewLSTM achieves 89.89% accuracy.

We summarize our contributions as follows.

- We exhibit our observations, insights, and rules about the periodic patterns from different parking records, types of parking lots, and weather factors.
- We propose PewLSTM, a periodic weather-aware LSTM with event mechanism, which incorporates the weather and periodic parking patterns for parking behavior prediction.
- We analyze the parking behaviors in events, including regular events like holidays and unexpected events like COVID-19, and provide our solution.
- We evaluate PewLSTM with real parking records and it achieves 93.84% accuracy, which is about 30% higher than the state-of-the-art method.

## 2 BACKGROUND

As far as we know, this work is the first to use the LSTM-based method to predict parking behavior by using records along with weather information.

### 2.1 Vehicle Parking

Many studies have been conducted on different aspects of vehicle parking, including parking demand prediction [13], smart parking guidance [7], [14], [15], parking space prediction [8], [16], [17], [18], [19], [20], and parking behavior prediction [9]. The regression-based parking behavior prediction [9] is the closest work to PewLSTM, which also considers weather conditions. They explored linear regression [21], ridge regression [22], Lasso regression [23], decision tree [24], and random forest [25] to predict the parking arrivals in one hour, and found that random forest achieves the highest accuracy.

However, the work [9] has three limitations. First, it provides the prediction for only parking arrivals, without departures. Departures are equally important as to arrivals, and they together define the parking behavior. Second, it does *not* provide the prediction results for more than one hour, which limits the applicability, since users may plan to book a parking space in an uncertain future. Third, their

1. http://www.huaching.com/

methods fail to capture the periodic patterns from long periods. For example, the parking space availability may have strong correlations with previous parking patterns, such as weekday and holiday patterns. Even worse, that work used records from only one parking lot for validation, which *cannot* represent all parking lots. Different types of parking lots exhibit various parking behaviors, and the locations of parking lots have a high influence on parking behaviors. For example, in holidays, the parking lots near a shopping mall are busy while the parking lots in an industrial area are idle.

### 2.2 Recurrent Neural Network

Recurrent neural network (RNN) [26] is a neural network that incorporates temporal sequence data as input. RNN performs recursion along the temporal sequence. Long Short-Term Memory (LSTM) [27] is a kind of RNN and its basic unit is a cell or called LSTM block. An LSTM block is usually composed of input gates, forget gates, and output gates, which can capture the complex non-linear relations among different factors. LSTM has been applied to different aspects of our daily life, such as financial market prediction [28], disease progression modeling [29], predictive phenotyping [30], text categorization [31], and sentiment analysis [32]. Particularly, LSTM conquers the limitations that the state-of-the-art parking behavior prediction method [9] failed to consider from long and short terms.

A typical LSTM model can be formulized as the following equations, where at the time step $t$, $h_t$ represents the hidden state, $x_t$ represents the input vector, $c_t$ represents the cell sate, and $f_t$, $i_t$, and $o_t$ are the forget gate, input gate, and output gate. $W$ and $b$ are related to weight matrices. Additionally, $\sigma$ indicates the sigmoid activation function, $\odot$ indicates element-wise multiplication, and $\phi$ indicates the *tanh* function.

$$
\begin{aligned}
\boldsymbol{f}_t &= \sigma(W_{fx}\boldsymbol{x}_t + W_{fh}\boldsymbol{h}_{t-1} + \boldsymbol{b}_f) \\
\boldsymbol{i}_t &= \sigma(W_{ix}\boldsymbol{x}_t + W_{ih}\boldsymbol{h}_{t-1} + \boldsymbol{b}_i) \\
\boldsymbol{g}_t &= \phi(W_{gx}\boldsymbol{x}_t + W_{gh}\boldsymbol{h}_{t-1} + \boldsymbol{b}_g) \\
\boldsymbol{o}_t &= \sigma(W_{ox}\boldsymbol{x}_t + W_{oh}\boldsymbol{h}_{t-1} + \boldsymbol{b}_o) \\
\boldsymbol{c}_t &= \boldsymbol{g}_t \odot \boldsymbol{i}_t + \boldsymbol{c}_{t-1} \odot \boldsymbol{f}_t \\
\boldsymbol{h}_t &= \phi(\boldsymbol{c}_t) \odot \boldsymbol{o}_t
\end{aligned}
\tag{1}
$$

## 3 Overview

### 3.1 Problem Definition

The problem we are trying to solve is how to predict the parking behavior for a given hour in the future by using the previous parking records and weather information.

**Parking Behavior Definition**. In this study, the parking behavior for a parking lot is defined by its arrivals and departures, as shown below. $\mathbb{N}$ represents the natural number domain.

- **Arrivals**. The sequence of the number of arrivals for the past $K$ time steps is $\boldsymbol{C} = \{c_{t-K}, \dots, c_{t-2}, c_{t-1}\} \in \mathbb{N}^K$, where $t$ represents the current time and $c_i$ represents the parking number of arrivals of the past $i$-th hour.
- **Departures**. The sequence of the number of departures for the past $K$ time steps is $\boldsymbol{D} = \{d_{t-K}, \dots, d_{t-2}, d_{t-1}\} \in \mathbb{N}^K$, where $t$ represents the current time and $d_i$ represents the parking number of departures of the past $i$-th hour.

**Event Prediction**. In this paper, the event prediction refers to the parking behavior prediction with event influence, including the arrivals and departures, in the entire event time frame.

**Significance**. Different from parking space prediction, parking behavior prediction contains more information for both users and owners of parking lots. For example, a parking lot is worth waiting if the number of departures $d_t$ is large, even if the parking lot is full (parking space prediction is useless in this case since the predicted results show no vacancy). When we obtain the number of arrivals and departures for a parking lot, we can further obtain its growth of parking space usage by subtracting the amount of departures from the amount of arrivals, which is very useful for smart parking systems.

### 3.2 System Overview

We show our system overview in Figure 1. Our system consists of four parts: an input layer with a classifier, a deep PewLSTM module, an event prediction module, and an output layer integrated with a linear layer. The classifier along the input layer provides the probability $p_{t_i}$ of an event at time $t$ by cosine similarity [33], presented in Equation 2, where $d_s$ and $d_i$ represent standard and current input. The deep PewLSTM module and event prediction module are the major components of our system. The deep PewLSTM layer consists of PewLSTM blocks, which predicts periodic weather-aware parking behaviors. The event prediction module predicts the parking behaviors in events. Both the results from LSTM module and event prediction module are fused as $out_i$ and then fed to the output layer.

$$
\begin{aligned}
\boldsymbol{p}_{t_i} &= \frac{1}{2}\left(1 - \frac{\boldsymbol{d}_s \cdot \boldsymbol{d}_i}{\|\boldsymbol{d}_s\|\|\boldsymbol{d}_i\|}\right) \\
\boldsymbol{p}_t &= (\boldsymbol{p}_{t_1}, \boldsymbol{p}_{t_2}, ..., \boldsymbol{p}_{t_n}) \\
\boldsymbol{out}_t &= \boldsymbol{h}_t \cdot (\boldsymbol{e} - \boldsymbol{p}_t) + \boldsymbol{v}_t \cdot \boldsymbol{p}_t \\
\boldsymbol{r}_t &= W_{ro}\boldsymbol{out}_t + \boldsymbol{b}_r
\end{aligned}
\tag{2}
$$

### 3.3 Workflow

The workflow of our system prediction is as follows. First, the parking records and weather data are input into the input layer, which preprocess the input data and output to both the deep PewLSTM layer and event prediction module. The deep PewLSTM layer outputs the generated data to the output layer. If the event prediction module identifies a special event, the module also output the event prediction results to the output layer. The output layer analyzes the input and generates the final output.

In the following sections, Section 4 shows how to predict periodic weather-aware parking behaviors, Section 5 shows the parking behavior prediction in events, Section 6 shows our optimizations.

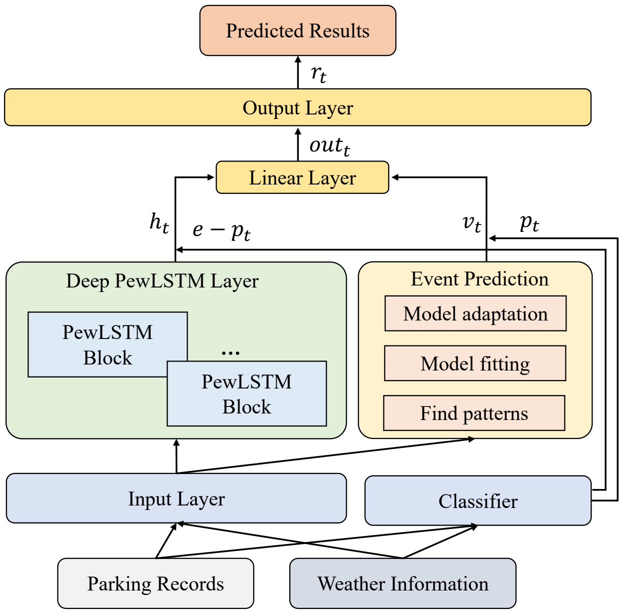



Fig. 1. System overview.

# 4 Periodic Weather-Aware Parking Behavior Prediction

## 4.1 Observations

In this part, we exhibit our parking behavior observations and analysis, which are also the motivation of this work.

**Data**. We collect both parking and weather data. For parking data, we collect 910,477 real parking records from 13 parking lots (detailed in Section 7.1). In our study, we reduce the analysis granularity to one hour, which is the same as [9]. In this section, we use roughly the parking counts of arriving cars in the parking lot within one hour as an indicator to reflect the parking behavior. For weather information, we collect the weather data, including temperature, humidity, wind speed, and precipitation.

**Observation 1: Periodic Parking Patterns**. Periodic parking patterns refer to that the arrivals and departures of a parking lot repeat in a cyclical and predictable manner. We find that most parking lots follow certain periodic parking patterns. For illustration, we show in Figure 2 the parking counts at hour granularity for parking lot P2 (see Table 1), which is close to a commercial street. Figure 2 shows that the parking lot follows a pattern that it is busy during the day (from 9am to 9pm). The reason is that the nearby shops open at 9am and close at 9pm. People go to these shops during the day and this activity lasts until the evening. Similar periodic parking patterns also exist at week and month granularities.

**Observation 2: Types of Parking Lots**. The types of parking lots also make the parking behaviors different from each other. For illustration, in Figure 3, we show the parking counts of two parking lots: the parking lot P4 located in an industrial park, and the parking lot P6 located near a market (see Table 1). We can see that these two parking lots show totally different behaviors. For explanation, people drive to work during working days, so P4 is busy only from Monday to Friday. In contrast, people tend to go to the market on weekends, so P6 is busy on Saturday and

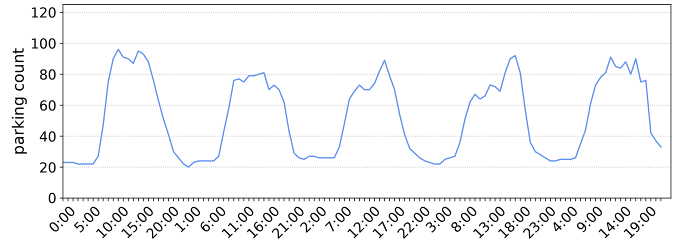
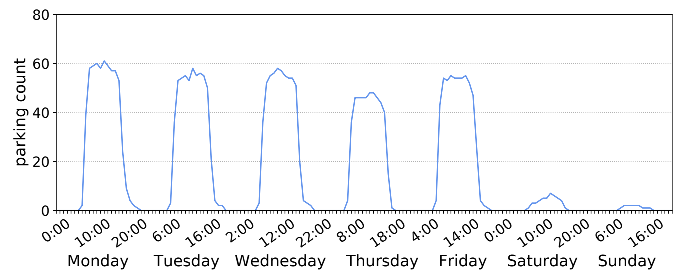
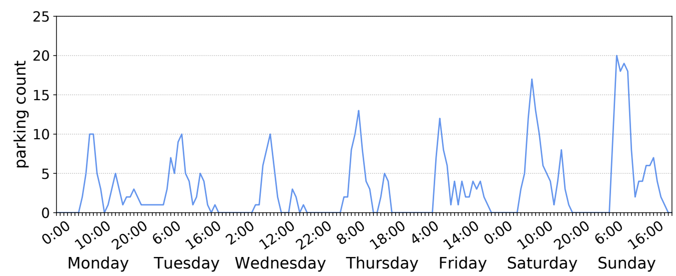



Fig. 2. Periodic parking patterns for P2 at hour granularity.

Sunday. Moreover, we can analyze the parking lots from the two categories: 1) mandatory for living, and 2) entertainment. The parking lots that relate to mandatory for living usually exhibit strong periodicity and have relatively low correlation with weather conditions. In contrast, the parking lots that relate to entertainment are more likely influenced by other factors. P4 is located in a industrial park, which belongs to mandatory for living, while P6 is near a market, which relates to entertainment. We can see that P4 shows a more regular parking behavior waveform.




(a) P4: industrial park.



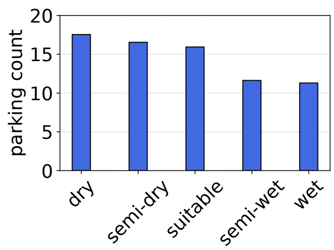

(b) P6: market.

Fig. 3. Different types of parking lots.

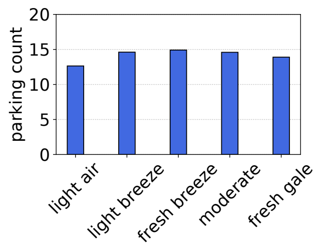
**Observation 3: Influence of Different Weather Factors on Parking Behavior**. The parking behavior exhibits different correlations with various weather factors. For illustration, we show in Figure 4 the average parking count of the parking lot P3 (see Table 1), which is close to a shopping mall. We show the relationship between the parking count and the relative humidity in Figure 4 (a), which implies that the parking behavior has a clear relationship with the relative humidity. There are two parks and a restaurant near this parking lot. Figure 4 (a) further implies that people are more willing to go to these places in dry weather instead of wet weather. We show the relationship between parking count and the wind speed in Figure 4 (b). We find that for this shopping mall, the wind speed shows less correlation with parking behavior.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2021.3070202, IEEE Transactions on Knowledge and Data Engineering






(a) Relative humidity.




(b) Wind speed.

Fig. 4. Weather influence on parking behavior for P3.

**Challenges**. We meet two challenges in applying LSTM in parking behavior prediction in this part. The first is how to involve periodic parking patterns since typical LSTM does *not* have the periodic design. The second is how to design the path from sources of parking patterns and weather into the various parts within the LSTM block.

### 4.2 Periodic Weather-Aware LSTM Model

In this part, we show our design of PewLSTM for parking behavior prediction and our solutions to the challenges mentioned in Section 4.1.

**Model Overview**. We show the model overview of PewLSTM for parking behavior prediction in Figure 5. To involve the periodic parking patterns, we add input gates for hidden states from previous LSTM functions at the same time at day, weekday, and month granularities. For example, for a time $t$, the current parking behavior may be influenced by the parking behavior at the same time from previous days (day granularity), so we add a gate for hidden states from the previous "$t-24 \cdot n$" time step, where $n$ is a positive integer. For example, people would like to park cars at a parking lot near a restaurant at noon, and the parking behavior at noon has a strong correlation with previous behaviors at noon in the past days. The parking behavior may be influenced by the same time from last week and last month, so we add gates for inputting the time pattern influence at the same time from the past LSTM states. As for model input, each LSTM function receives both weather information and records: for weather input, we design special gates for receiving weather information, including temperature, humidity, wind speed, and precipitation. For parking record input, the record includes the arrival and departure times, date, and weekday information. Furthermore, we also add an event prediction module to predict parking behaviors in events, which shall be discussed in Section 5.

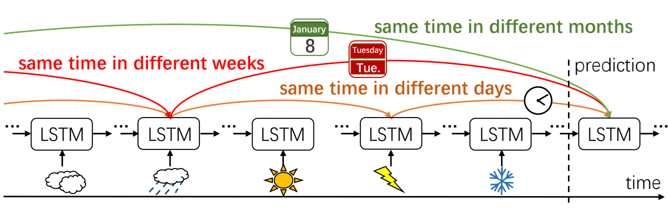



Fig. 5. Illustration of periodic weather-aware LSTM prediction.

**PewLSTM Block**. The LSTM function in PewLSTM is composed of PewLSTM blocks, as shown in Figure 6. Traditional LSTM block uses input gate $i_t$, output gate $o_t$, and forget gate $f_t$ to control dataflow in LSTM. We argue that for parking behavior prediction, we need to additionally involve 1) periodic pattern influence and 2) weather information. We consider these two factors as new gates involved in our PewLSTM model. We show our periodic weather-aware LSTM block in Figure 6. The main difference between our PewLSTM block and traditional LSTM block lies in the gates for previous periodic records $h_{day/week/mon}$ and weather input $weather_t$. First, for periodic time influence, the hidden states at the same time from previous steps are represented as $h_{day}$, $h_{week}$, and $h_{mon}$, and we design a special weight gate $\delta$ to integrate these input hidden states with $h_{t-1}$. The weight gate $\delta$ iterates continuously during training, adjusting the weight of each parameter in the model to achieve better results, as shown in Equation 3. Second, for weather influence, a weather-aware gating mechanism has been integrated into the PewLSTM block, as represented in Equation 4. We argue that the current weather information, along with the parking records and the previous parking patterns together decide the parking behaviors in the future. Based on this assumption, after a sigmoid function processing, the weather aspect vector $e_t$ is integrated to $f_t$, $i_t$, and $o_t$. Therefore, the cell state $c_t$ and the hidden state $h_t$ are influenced by the previous hidden states, cell states, the current weather information, and the periodic patterns.

$$\boldsymbol{h}_o = \delta(W_d \boldsymbol{h}_{day} + W_w \boldsymbol{h}_{week} + W_m \boldsymbol{h}_{mon} + W_{t-1} \boldsymbol{h}_{t-1}) \quad (3)$$

$$\boldsymbol{e}_t = \sigma(W_e weather_t + \boldsymbol{b}_f) \quad (4)$$

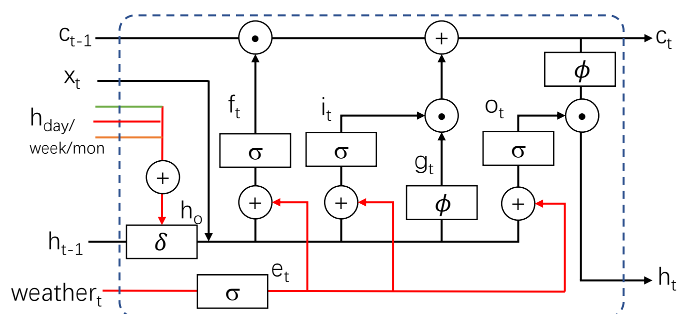



Fig. 6. PewLSTM block.

**Input Gates**. The input gate $i_t$ decides the amount of data that can be accumulated to the cell state $c_t$ from the previous state $c_{t-1}$. Different from the traditional input gates, the input gate in our model integrates the input vector $x_t$, the weighted hidden state $h_o$, and the weather input $e_t$, as shown in Equation 5.

$$\begin{aligned} \boldsymbol{i}_t &= \sigma(W_{ix}\boldsymbol{x}_t + W_{ih}\boldsymbol{h}_o + W_{fe}\boldsymbol{e}_t + \boldsymbol{b}_i) \\ \boldsymbol{g}_t &= \phi(W_{gx}\boldsymbol{x}_t + W_{gh}\boldsymbol{h}_o + \boldsymbol{b}_g) \end{aligned} \quad (5)$$

**Forget Gates**. The forget gate abandons unnecessary information so that only the useful information is retained in the cell state. Similar to the input gate, along with $h_{t-1}$ and the input vector $x_t$, the weather information and periodic hidden states from previous time steps also affect the retention from the $c_{t-1}$ state to the $c_t$ state, as shown in Equation 6.

$$\boldsymbol{f}_t = \sigma(W_{fx}\boldsymbol{x}_t + W_{fo}\boldsymbol{h}_o + W_{fe}\boldsymbol{e}_t + \boldsymbol{b}_f) \tag{6}$$

**Output Gates**. The output gate is used to control the data flow from the current input to the hidden state $h_t$, as shown in Equation 7. Similarly, the weighted hidden state $h_o$, input vector $x_t$, and weather information together decide the output $o_t$, which is different from the traditional output gate.

$$\boldsymbol{o}_t = \sigma(W_{ox}\boldsymbol{x}_t + W_{oh}\boldsymbol{h}_o + W_{fe}\boldsymbol{e}_t + \boldsymbol{b}_o) \tag{7}$$

**Cell States**. Traditional cell state is decided by previous state $c_{t-1}$, forget gate $f_t$, input gate $i_t$, and $g_t$. We need to involve the periodic and weather information into the cell state. However, $f_t$ and $i_t$ already preserve this information. Hence, the expression of the cell state $c_t$ keeps the same as in Equation 1, and so does to the hidden state $h_t$.

**Future Prediction**. To predict the parking behavior in a future time step $t+k$ ($k$ is a positive integer), we recursively predict the states in the near future, and input the predicted states into the model to predict the farther future states. Note that the periodic pattern influence also applies to the future prediction.

### 4.3 Model Details

Our PewLSTM model proposed in Section 4.2 needs further developments for different data formats. For example, the input data include different dimensions like *weekdays* and *months*, and numerical ranges like *temperature* and *wind speed*. Therefore, we add an input layer for data processing. To capture the complex relations among periodic time patterns, weather information, and parking records, we stack several PewLSTM blocks together for prediction. To output the predicted results, we add an output layer.

**Input Layer**. The input layer receives the input data for preprocessing. In our model, the processing granularity is one hour, so parking records are converted on an hourly granularity. Along with the parking records, the input weather data, including temperature $T$, humidity $H$, precipitation $P$, and wind speed $W$, are normalized to [0,1]. For example, as for temperature, the corresponding sequence for the past $K$ time steps is $\boldsymbol{T} = \{T_{t-K}, \cdots, T_{t-2}, T_{t-1}\}$, where $T_i$ represents the temperature of the past $i$-th hour.

**Deep PewLSTM Layers**. Different from traditional single-layer LSTM, we stack multiple PewLSTM layers together to form a deep neural network to capture the relations among parking records and other useful factors. Between two PewLSTM layers, the output of a PewLSTM block in one layer is the input to a PewLSTM block in the following layer.

**Output Layer**. Because the final output from the model is the predicted parking numbers of arrivals and departures, we add a feed-forward neural network following the deep PewLSTM layers for post-processing. Similar processes have been conducted in previous research, such as [34].

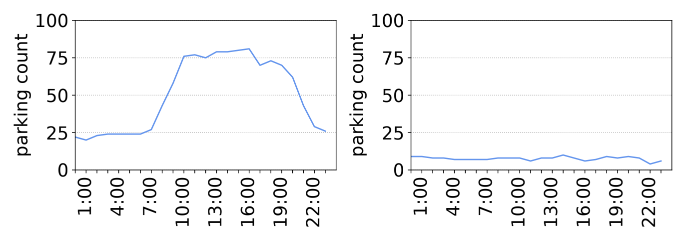
## 5 Parking Behavior Prediction with Event Influence

Large-scale events, such as festival activities, also have a great impact on parking behaviors. When an impactful event happens, PewLSTM switches to the event prediction module for parking behavior event prediction. In this section, we analyze the influence on parking behavior from regular social events and unexpected events, and then provide our solution.

### 5.1 Regular Event

Regular events refer to the events that appear periodically according to certain rules, such as holidays. We use the Chinese New Year as an example to show the influence of regular events on parking.

**Hour Granularity**. Figure 7 shows the parking behavior of parking lot P2, which is close to a commercial street in Ningbo, in non-holiday and holiday for 24 hours. Figure 7 (a) shows its parking behavior on March 8th 2018. This parking pattern is common for the other non-holidays. The parking lot turns to be busy from the morning till the night. The maximum parking count reaches 81. In contrast, Figure 7 (b) shows the parking behavior on February 16th 2018, which is during the Chinese New Year. The utilization rate of the parking lot is 8.76% on average, and the parking lot turns to be idle in almost all the time. The reason for such differences between non-holiday and holiday is that people turn to go home from big cities during the new year. Accordingly, the parking lot becomes idle.




(a) Non-holiday. (b) New Year.

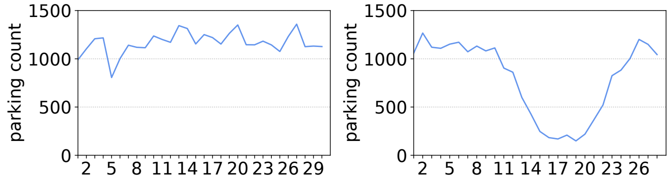
Fig. 7. P2: Influence from Chinese New Year at hour granularity.

**Day Granularity**. Figure 8 shows the parking behavior difference of parking lot P2 in non-holiday and holiday at day granularity. Figure 8 (a) shows the non-holiday parking from April 1st 2019 to April 30th 2019, while Figure 8 (b) shows the holiday parking from February 1st 2018 to February 28th 2018, in which the new year holidays start from February 10th 2018 to February 26th 2018. From Figure 8, we can see that the parking lot turns to be empty at the beginning of the new year holiday, and returns to normal close to the end of the New Year holiday. Such phenomena confirm our conjecture: large-scale population movements will occur during long holidays such as the New Year, and some parking lots will be vacant while others will be the opposite.

**Small Cities**. Note that in small cities, the parking behavior could become the opposite. People flow out of big cities during the New Year and flow into their hometowns. After the new year, a large number of people return to the big cities from their hometowns and return to work.

### 5.2 Unexpected Event

The parking behavior could also be influenced by unexpected events. In this section, we analyze the influence of coronavirus pandemic (COVID-19) [11] on parking.




(a) April. (b) February.

Fig. 8. P2: Influence from Chinese New Year at day granularity.

**Short-Term Impact**. We use the parking lot P2 to show the short-term impact of COVID-19 on parking behavior in Figure 9. The Chinese pulmonologist Zhong Nanshan warned that COVID-19 can be transmitted from person to person in an interview with CCTV on January 20th, 2020 [35], and we can see that this warning had a great impact on parking behavior: within 4 days, the number of parking at the peak dropped by 76.53%. The reasons are as follows. First, the government encourages everyone to stay at home and not go out. Second, many non-essential entertainment venues are closed and employees stay at home and wait for notification. Third, the epidemic news has brought fear, worries, and anxiety to the public [36]. Accordingly, the parking lot became vacant and underutilized.

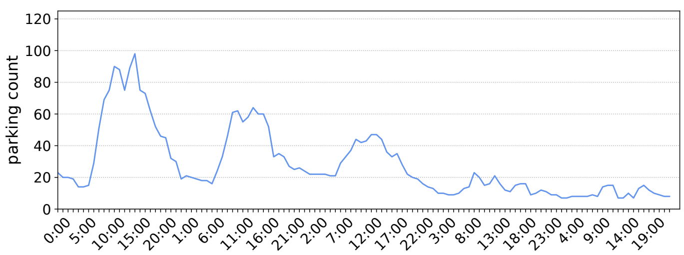
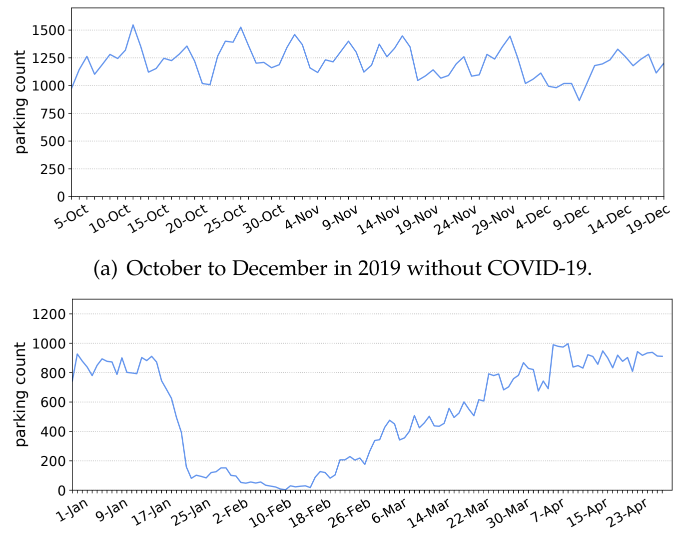



Fig. 9. P2: Short-term COVID-19 influence from January 21 to January 25, 2020.

**Long-Term Recovery**. With effective measures in various aspects and effective control of the COVID-19 epidemic, people gradually regained their confidence, and the parking lot gradually resumed use. We show such a long-term recovery in Figure 10. Figure 10 (a) shows the normal parking behavior from October to December, 2019, while Figure 10 (b) shows the parking behavior with COVID-19 influence from January to April, 2020. The utilization of the parking lot drops within 20 days from 66.57% to 8.00%. On February 8th, the parking lot utilization rate dropped to freezing point, and this state lasted for 10 days. Fortunately, on February 18th, the government sent a positive signal: COVID-19 pandemic has been efficiently controlled. Next, such information has promoted the normal operation of society and the restoration of normal use of parking lots. Till March 25th 2020, the parking lot has resumed 74.39% of its previous usage.

**Insights**. Based on these observations, we have the following insights. First, the parking behavior has a strong correlation with the news. For example, the news reports the COVID-19 control situation. Second, no fixed pattern in parking behavior exists in unexpected events, since the parking behavior greatly depends on the current situation.




(a) October to December in 2019 without COVID-19.

(b) January to April in 2020 with COVID-19.

Fig. 10. P2: Accumulated parking count of long-term COVID-19 influence.

Third, many factors need to be involved to predict the parking behaviors in unexpected events, such as the time when the epidemic was under control.

### 5.3 Event Prediction Solution

We show our solution to such an event influence on parking behavior prediction in this part.

**Event Identification**. Before predicting the parking behavior during events, we need to identify the events that affect the parking behavior. We category the events into long-term events that may last several days, and short-term events that last only a few hours.

1) *Long-term events*. To identify the long-term events that potentially affect parking behavior, we use the LSTM model to find all possible events in the training dataset. In detail, we first mark some events in one parking lot: we map the records in events to a higher weight while remaining the other records to a low weight. Second, we train our model on the manually labeled dataset. Third, we use our trained model to identify events.

2) *Short-term events*. Because parking behavior exhibits obvious periodic patterns and the crests and troughs are relatively fixed, we can detect the short-term events by comparing current waveform to a generalized average waveform. In detail, we first generate the average parking counts in each hour of one day to form a baseline vector. Second, we use the cosine similarity [33] to calculate the similarities between different waveforms and the baseline. Third, we mark the identified short-term events. Using the above method can identify both long-term and short-term events. However, most of the identified short-term events relate to unpredictable human factors such as shopping mall promotion, which is hard to be controlled, so we only consider long-term events in PewLSTM. Long-term events include both regular and unexpected long-term events.

**Event Prediction**. Based on the event identification, we perform the event prediction, which includes three steps:

finding pattern, polynomial model fitting, and model adaptation.

1) *Finding the pattern*: We first need to find the parking pattern of the event. The event identification provides the length of the event, and we record the trend of parking behaviors during the interval. We also record the number of parking spaces for further parking behavior prediction.

2) *Polynomial model fitting*: After we collect the data to depict the parking pattern during the event, we then use a high-order polynomial model to fit the parking data change pattern. Higher-order models usually fit better, but it may also bring about over-fitting problems. The experience value is 15 in our experience.

3) *Model adaptation*: Our model can be applied to three situations. The first is to predict parking behaviors for the same parking lot in the future. The second is to predict parking behaviors cross different events. Because events may share similarities, we need to adjust the length to fit different events. The third is to predict parking behaviors cross different parking lots, which needs further adjustment to proportionally put or reduce the capacity of the parking lot. Note that only the same type of parking lot can make cross-parking behavior predictions.

After these steps, we can conduct event prediction on upcoming events for parking behaviors.

# 6 Optimization

In this section, we describe our optimizations on data preprocessing and accuracy tuning.

## 6.1 Data Preprocessing

The data preprocessing is important to our method, which includes data selection of parking records, weather information, and parking environment, as well as the data filtering and correlation analysis.

**Parking Records**. We use real parking records from 13 parking lots. The format of the original parking records includes the start time, end time, and the parking space ID. In our design, we reduce the analysis granularity to one hour. We roughly use the number of cars parked in the parking lot at the hour granularity as an indicator to reflect the parking behavior.

**Weather Information**. The weather condition also affects the parking behavior. For instance, as discussed in Section 4.1, we find that people tend to drive in windy weather. In detail, we collect the following weather information from the local weather bureau where the parking lot is located: temperature, humidity, wind speed, and precipitation. These data need to be normalized.

**Parking Environment**. The surroundings of a parking lot play an important role in the parking behaviors of the parking lot. As discussed in Section 4.1, parking lots with different surroundings may have totally different parking behaviors. We categorize the parking environments into seven types: hotel, commercial street, shopping mall, industrial park, hotel, market, and community. Accordingly, we train each type of parking lots separately.

**Data Filtering**. In our study, we mainly meet two data issues: noisy data and missing attributes. We perform data filtering to solve these two issues. For noisy data in a dataset, we perform outlier analysis and remove the outlier data. For missing attributes of a record, if more than 20% of attribute values are missing, we abandon that record.

**Correlation Analysis**. We perform correlation analysis for the involved attributes. We only keep the attributes that have a low correlation with each other to reduce the model complexity.

## 6.2 Accuracy Tuning

Many configurations could affect the accuracy of PewLSTM, and we discuss the accuracy tuning for these configurations in this section.

**Hidden Size**. Hidden size refers to the number of output features of the hidden state $h_o$ in LSTM blocks. We find that the hidden size of LSTM blocks affects the parking behavior prediction accuracy. A larger hidden size represents parking behavior from more dimensions but may contain more redundant information. We have explored a wide range of hidden size and find that PewLSTM achieves the highest accuracy when the hidden size is 64.

**Training Epochs**. In PewLSTM, an epoch represents the training process for the input dataset. In our tuning, we find that the accuracy of the model increases with the number of epochs. However, when the number of training epochs reaches 300, the accuracy reaches its maximum. Hence, we set the number of training epochs to 300.

**Number of LSTM Layers**. As stated in Section 4.3, we stack PewLSTM blocks together to form deep PewLSTM layers to capture the complicated non-linear relations in parking behavior. To decide the number of PewLSTM layers, we use a small training set for validation and choose the number of layers that can provide the best results. In PewLSTM, we use two LSTM layers.

**Output Layer**. Although we can configure the LSTM block to output the prediction results, we find that add another transformation could further improve the accuracy. After trying different strategies, such as linear transformation and softmax function, we find that the linear transformation achieves the optimal results.

# 7 Experiments

In this section, we evaluate PewLSTM on 910,477 records from 13 parking lots and compare it with the state-of-the-art parking behavior prediction method.

## 7.1 Experimental Setup

**Methodology**. We evaluate the prediction accuracy of five methods for the future three hours. The "PewLSTM" method is our periodic weather-aware LSTM with event mechanism based on all influencing factors. The "simple LSTM" method is an LSTM implementation based on only parking records without other factors. We use the regression method of random forest from [9], denoted as "regression", as our baseline. Because the regression method only supports the prediction in one hour, we extend it to support prediction for two and three hours. We use the baseline implementation with the datasets described below for comparison. Additionally, we show our preliminary

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2021.3070202, IEEE Transactions on Knowledge and Data Engineering



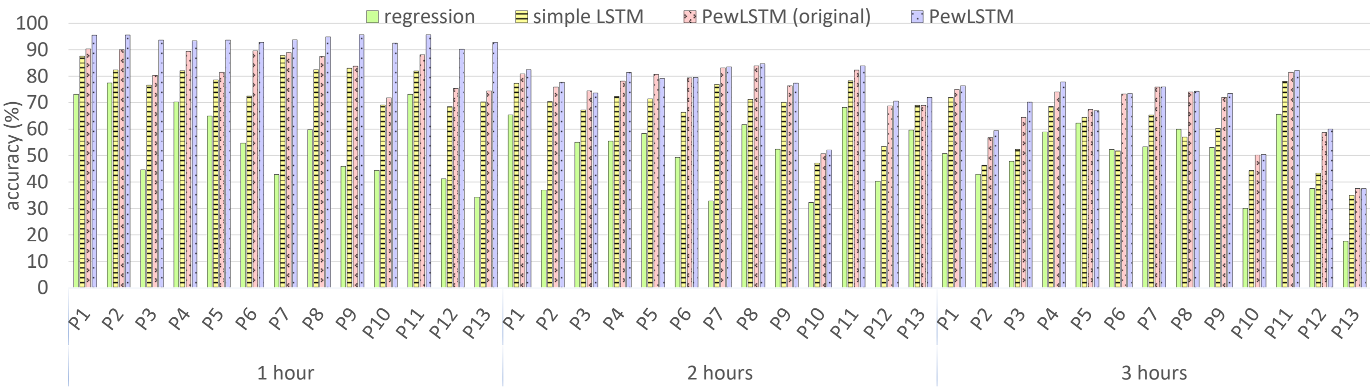



Fig. 11. Arrival prediction accuracy.

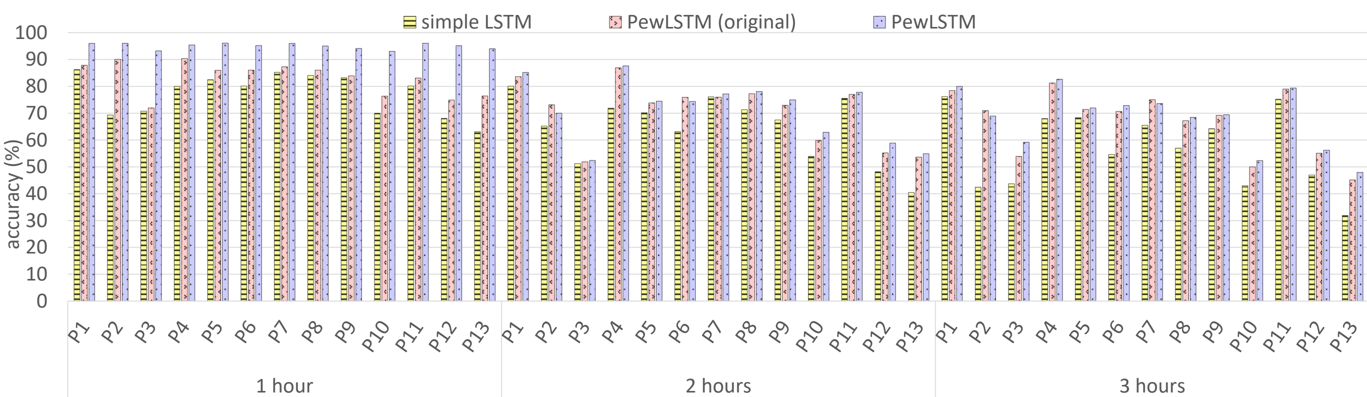



Fig. 12. Departure prediction accuracy.

version without the optimizations in Section 6, denoted as "PewLSTM(original)", which is also the initial version of PewLSTM presented in [10]. The accuracy is defined in Equation 8.

$$accuracy = 1 - \frac{|count_{observed} - count_{predicted}|}{count_{observed}} \quad (8)$$

**Datasets**. In this paper, our parking datasets are composed of parking records from 13 parking lots in China, including shopping malls, hotels, communities, and so on[2]. We list the total number of parking spaces, parking records, surroundings, and location for each parking lot in Table 1. The dataset spans 30 months from October 20th, 2017 to April 7th, 2020, which consists of 910,477 parking records in 904 days. For each record, we obtain the parking information including the arrival time, the departure time, date, parking space, and the price. Please note that the parking records with a duration of fewer than five minutes are regarded as noise data. The related weather dataset includes the hourly weather data from four districts where the 13 parking lots are built. For each hour, we collect the related weather information including temperature, wind speed, precipitation, and humidity.

**Platform**. Our experiments are conducted on a server equipped with an Intel i7-7700K CPU and an NVIDIA GTX 1080 GPU. Its memory capacity is 32 GB, and the operating system is Ubuntu 16.04.4 LTS. Our PewLSTM is implemented by PyTorch 1.2.0.

2. https://github.com/NingxuanFeng/PewLSTM

### 7.2 Prediction Accuracy Evaluation

We show the arrival prediction accuracy in Figure 11, and the departure prediction accuracy in Figure 12. We have the following findings.

First, our periodic weather-aware LSTM with event mechanism achieves the highest prediction accuracy in most cases. For arrivals, PewLSTM achieves an accuracy of 93.84% in one hour, 76.78% in two hours, and 67.54% in three hours. In contrast, the *regression* method achieves 57.8% accuracy in one hour, 51.2% in two hours, and 50.0% in three hours. For departures, PewLSTM achieves an accuracy of 93.34% in one hour, 73.45% in two hours, and 67.48% in three hours. The "PewLSTM(original)" achieves a moderate accuracy that is higher than the accuracy of simple LSTM, which implies that periodic patterns and weather information are useful for improving prediction accuracy.

Second, with the growth of the prediction time, the prediction accuracy decreases. In our method, to predict parking behavior at a far time point, we recursively input the predicted results in a near time point into the model. We use the predicted results as training data, which inevitably introduces uncertainties and noise factors. However, our

TABLE 1
Parking lot information.

| Parking Lot | Space# | Start Time | End Time | Record# | Surrounding | Location (district/city/province) |
|---|---|---|---|---|---|---|
| P1 | 23 | 2017/10/16 | 2019/10/6 | 22,955 | hotel | Haishu, Ningbo, Zhejiang |
| P2 | 87 | 2017/10/20 | 2020/4/7 | 280,687 | commercial street | Haishu, Ningbo, Zhejiang |
| P3 | 9 | 2018/6/29 | 2020/4/7 | 38,509 | shopping mall | Yinzhou, Ningbo, Zhejiang |
| P4 | 62 | 2019/3/12 | 2020/4/7 | 13,476 | industrial park | Yinzhou, Ningbo, Zhejiang |
| P5 | 46 | 2017/11/6 | 2019/9/11 | 29,131 | hotel | Yinzhou, Ningbo, Zhejiang |
| P6 | 16 | 2018/4/27 | 2019/8/24 | 32,009 | market | Yuelu, Changsha, Hunan |
| P7 | 31 | 2018/5/9 | 2019/8/22 | 49,708 | market | Yuelu, Changsha, Hunan |
| P8 | 27 | 2018/5/29 | 2020/4/7 | 63,502 | shopping mall | Yuelu, Changsha, Hunan |
| P9 | 49 | 2018/6/22 | 2020/4/7 | 55,631 | shopping mall | Yuhua, Changsha, Hunan |
| P10 | 65 | 2018/12/25 | 2020/4/7 | 38,581 | community | Yuhua, Changsha, Hunan |
| P11 | 11 | 2018/12/19 | 2020/4/7 | 10,156 | commercial street | Yuhang, Hangzhou, Zhejiang |
| P12 | 152 | 2018/5/11 | 2020/4/7 | 140,034 | commercial street | Yuhang, Hangzhou, Zhejiang |
| P13 | 114 | 2018/6/27 | 2020/4/7 | 136,098 | shopping mall | Yuhang, Hangzhou, Zhejiang |

PewLSTM still maintains high accuracy in terms of long-term predictions: for every additional hour of prediction, the accuracy drops by only 13.1% for arrivals and 12.9% for departures on average.

Third, the arrival prediction accuracy is higher than the departure prediction accuracy. The reason is that departure situations are more complex than arrival situations. The departure is also affected by the purpose or event from the driver side during parking, which is hard to capture. However, even in the departure prediction, our prediction accuracy is only 1.3% lower than that in arrival prediction.

### 7.3 Qualitative Analysis

We use the root-mean-square deviation (RMSE) to measure the deviation between the predicted results and the observed results. Low RMSE values indicate higher accuracy and stability. In our evaluation, RMSE is defined in Equation 9. Assume we perform $N$ predictions. For the $i$-th prediction, $predicted_i$ represents the predicted parking results while $observed_i$ represents the observed parking results.

$$RMSE_{count} = \sqrt{\frac{\sum_{i=0}^{N-1}(predicted_i - observed_i)^2}{N}} \quad (9)$$

We show the RMSE results of P1 in Table 2 for illustration. The other parking lots show similar RMSE results. We have the following observations. First, our PewLSTM achieves the lowest RMSE results for both arrivals and departures, which implies that periodic patterns and weather information are useful for reducing deviation and the optimizations in Section 6 are effective. Second, as the prediction time becomes longer, the RMSE result increases. However, even so, PewLSTM can still maintain a small deviation. Third, the arrival RMSE results are lower than the departure RMSE results, which is due to the complexity in departure behaviors. Additionally, the regression-based method only predicts arrivals and it has high RMSE results that are more than 1.5, so we do not show it in Table 2.

**LSTM Configuration**. We analyze the influence of PewLSTM configuration over the model accuracy, and we mainly adjust the number of PewLSTM layers. The accuracy results of different layers are shown in Figure 13, which shows

TABLE 2
RMSE results of P1 for different methods.

| | Arrival | | | Departure | | |
|---|---|---|---|---|---|---|
| Method | 1hour | 2hours | 3hours | 1hour | 2hours | 3hours |
| simple LSTM | 0.99 | 1.28 | 1.36 | 1.09 | 1.28 | 1.45 |
| PewLSTM(original) | 0.92 | 0.99 | 1.08 | 0.99 | 0.87 | 1.19 |
| PewLSTM | 0.90 | 0.89 | 0.96 | 0.94 | 0.88 | 1.02 |

that when the number of layers is two, our model achieves the highest accuracy; when the number of layers further increases, the accuracy decreases. The optimal number of layers relates to both the input data and the problem. We mainly obtain it empirically based on the training set.

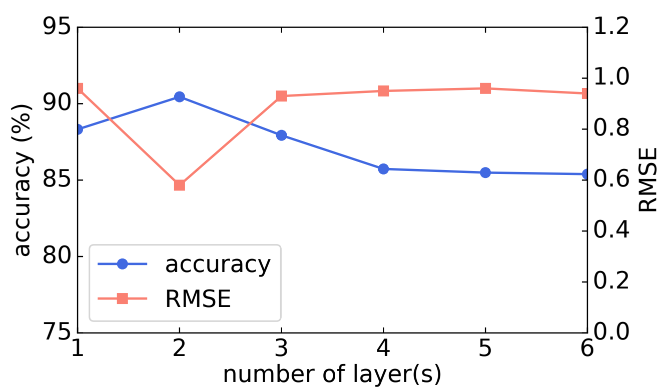



Fig. 13. Accuracy under different layers.

### 7.4 Event Analysis

In this part, we evaluate the accuracy of the event prediction of PewLSTM. Because the parking behavior with the hourly granularity is greatly affected by environmental factors, we mainly use the day as the granularity to predict parking behaviors.

**Selected Events**. We use the event identification method in Section 5.3 to detect the training set and select three regular events: National Day, Spring Festival, and May Day, which are the three longest holidays in China. Parking behaviors in these events exhibit distinct waveforms compared to the other periods. Note that we do not consider the unexpected events, though we have identified them. The reasons are as follows. First, the unexpected events usually happen by different reasons, which could be hard to reproduce. Second, only similar events can be related to each other for prediction, and the unexpected events are hard to use. Third, the records of most parking lots are only about one year. Therefore, we only use the three regular events for validation.

**Event Validation Method**. We use the parking lots, P2, P5, P6, P7, P8, P12, and P13 for evaluation, because these parking lots exhibit clear event patterns, and have enough records with a relatively long time range. In detail, the validation involves three types of model adaptation: 1) same event prediction in one parking lot, like P2, which has enough records across two years; 2) cross event prediction in one parking lot, since no enough records exist but different events show similar parking patterns; 3) cross parking lot prediction, for the parking lots that are similar to each other. Furthermore, the input is from two sources. The first one is the pre-trained fitting model for expressing patterns. The second one is the base data ahead of the event for model adjustment; in our experiments, we conduct parking behavior prediction within one month, two months, and three months.

**Event Prediction Results**. We show our event prediction results in Figure 15. On average, our PewLSTM achieves 72.91%, 77.19%, and 78.90% prediction accuracy for three months, two months, and one month respectively. For the Spring Festival, the average prediction accuracy for Spring Festival is 64.42%. For Figure 15 (c), we perform the same event prediction in one parking lot, in which we use the parking records in Spring Festival of 2018, along with the month data ahead, to predict the parking behavior in Spring Festival of 2019. For Figure 15 (d), we use the records from P2 in the 2018 Spring Festival to predict the P12 parking behavior of the Spring Festival of 2019. For National Day and May Day prediction, the average prediction accuracies are 79.60% and 80.09% respectively. Figures 15 (a,g) are from same event prediction in one parking lot, Figures 15 (e) is from cross event prediction in one parking lot, and Figures 15 (b,f,h,i) are from cross parking prediction. The unexpected events, such as COVID-19, could be rare and lack information. However, similar events sometimes happens. For example, in 2002, SARS broke out in China [37], in which traffic restrictions in major cities are similar to those during COVID-19. If we can collect sufficient data and build models, we can prepare for the next similar events in the future. When we have no similar records, cross-event prediction can be used by expert experience. For example, Chinese pulmonologist Nanshan Zhong predicted on February 27th that COVID-19 could be controlled by on the Chinese mainland by the end of April [38], so we can use a cross-event prediction such as New Year model with length set from February 28th to April 30th. The predicted parking behavior is shown in Figure 14, which is similar to the real situation.

**Event Analysis**. We analyze the event prediction results in Table 3, and we have the following observations and insights. First, in most cases, the closer to the event, the better prediction accuracy PewLSTM can achieve. The reason is that the input data is closer to the event data. Second, the

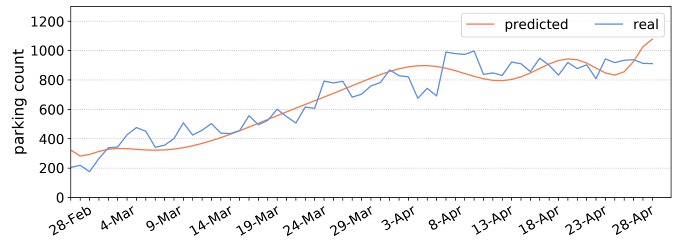



Fig. 14. P2: Parking behavior prediction in COVID-19.

event RMSE results are higher than the hourly RMSE results of Section 7.3, because of the longer forecast time range and more uncertainty factors. Third, even in such complicated situations where the prediction time range is long and the training data are limited, PewLSTM can still achieve significant prediction results, which proves the effectiveness of our method. Note that an unexpected event can also be predicted if we could obtain its event type and have similar historical event records.

TABLE 3
Accuracy and RMSE for event prediction.

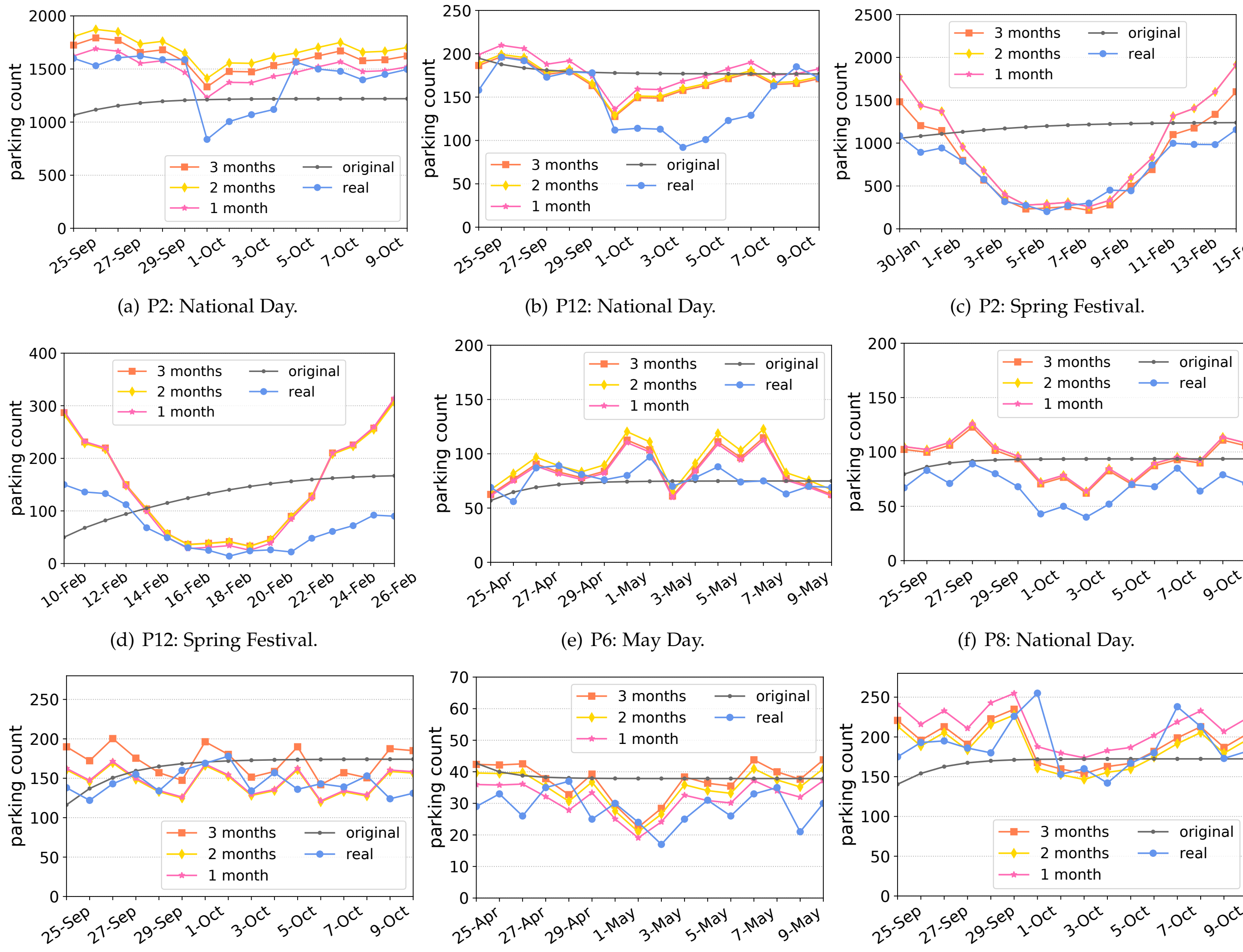
| | Accuracy (%) | | | RMSE | | |
|---|---|---|---|---|---|---|
| ParkingLot/Event | 3mon | 2mon | 1mon | 3mon | 2mon | 1mon |
| P2: National Day | 82.07 | 82.50 | 89.52 | 262.24 | 257.05 | 158.50 |
| P12: National Day | 75.76 | 86.91 | 87.28 | 55.33 | 29.32 | 28.50 |
| P2: Spring Festival | 53.60 | 67.00 | 83.76 | 518.13 | 354.41 | 122.28 |
| P12: Spring Festival | 61.72 | 62.49 | 63.25 | 27.93 | 27.11 | 26.38 |
| P6: May Day | 81.61 | 78.12 | 83.32 | 17.87 | 26.62 | 15.95 |
| P8: National Day | 60.99 | 57.65 | 57.50 | 25.52 | 28.65 | 28.75 |
| P13: National Day | 78.99 | 86.30 | 86.39 | 36.12 | 22.30 | 22.51 |
| P5: May Day | 71.58 | 83.06 | 82.87 | 46.03 | 33.68 | 33.97 |
| P7: National Day | 89.89 | 90.68 | 81.58 | 30.29 | 30.34 | 38.13 |

## 7.5 Performance Comparison

We show the performance results of our previous model of PewLSTM (original) and our new model of PewLSTM (original+extra) on both CPU and GPU in Figure 16. We have the following observations. First, our new model does not bring significant performance overhead compared to the original model, which means that our method greatly improves the prediction accuracy, as discussed in Section 7.2, with acceptable overhead. Second, the GPU acceleration is more effective on the training process than on the prediction process, because the training involves huge computation. Third, on the prediction process, we do not need to use GPU since it saves less than 0.15 second, which is marginal.

## 7.6 Deployment

Our periodic weather-aware LSTM with event mechanism has been integrated into a real smart parking system, ThsParking [12], which is designed to provide users with convenient and fast parking. The system mainly provides services of online payment, smart self-parking by phones, and parking space reservation. Our method can be used to prediction the earliest time for an empty spot in a parking, which is important especially when a parking lot is full. The earliest time for an empty spot in a parking lot only relates




(a) P2: National Day. (b) P12: National Day. (c) P2: Spring Festival.

(d) P12: Spring Festival. (e) P6: May Day. (f) P8: National Day.

(g) P13: National Day. (h) P5: May Day. (i) P7: National Day.

Fig. 15. Parking behavior prediction in events. “3 months”, “2 months”, and “1 month” indicate utilizing PewLSTM to predict the parking behavior for the event based on the current data three months, two months, and one month ahead. “original” indicates the PewLSTM without event mechanism, which is the state-of-the-art method.

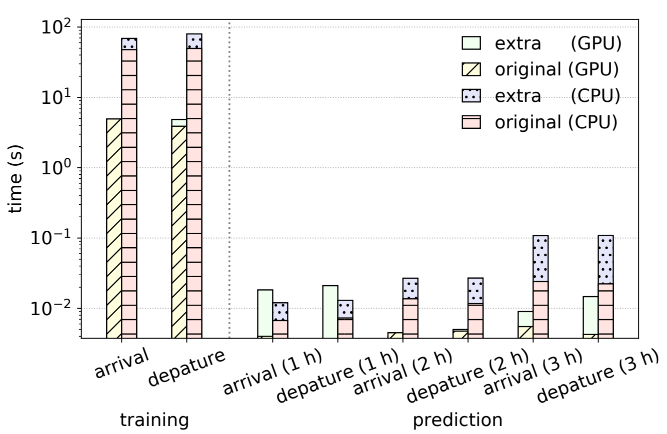



Fig. 16. Illustration for performance comparison between PewLSTM (original) and PewLSTM. The “original” indicates the execution time of PewLSTM (original) and the sum of “original” and “extra” indicates the execution time of PewLSTM.

to the departures: when a car leaves, there will be an empty spot. Because we prediction the parking behavior at hour granularity, the earliest time for an empty spot should be the reciprocal of the predicted number of departures per hour, assuming the vehicle departure time is evenly distributed.

## 8 Related Work

In recent years, there have been many studies about different aspects of parking prediction, LSTM application, and event analysis, and we show these related work as follows.

**Parking Prediction**. Parking relates to our daily life and has been a hot research topic for many years. Accurate parking prediction can help car owners to choose parking lots. Caicedo *et al.* [8] proposed a method to perform parking space prediction in intelligent parking reservations. Chen [17] studied parking occupancy and patterns, and Zheng *et al.* [18] analyzed the sensor-enabled car parks. Hossinger *et al.* [16] presented a real-time model for parking space occupancy prediction. Vlahogianni *et al.* [19] developed a framework for smart real-time parking availability prediction. We explored the use of weather information and regression models in parking prediction, which has been presented in [9], [10]. Moreover, high-quality smart parking guidance systems also appear, which are convenient for users [7], [14], [15], [39], [40].

**LSTM**. As a variant of RNN [26], LSTM [27] has been widely used in knowledge and data engineering. For instance, Yu *et al.* [41] applied reinforcement learning with Tree-structured LSTM to join order selection in data management systems. Wang *et al.* [28] developed CLVSA, which applied LSTM to predict the trends in financial markets. Zhang *et al.* [29] proposed ATTAIN, which is LSTM networks to model disease progression. Fu *et al.* [30] studied how to apply LSTM to predictive phenotyping. Huang *et al.* [31] applied LSTM text categorization. Xing *et al.* [32] proposed an aspect-aware LSTM for sentiment analysis based on aspects. Different from these work, we integrate the periodicity and weather information into the LSTM model, and applied it to the parking behavior prediction.

**Event Analysis**. With event analysis, we can understand the influence of events, and respond better to emergencies. Cortez *et al.* [42] presented a novel architecture for event prediction based on LSTM. Yuen *et al.* [43] developed a model to identify unusual events in videos with reference to a large collections of videos. Antunes *et al.* [44] developed a Bayesian predictive method to give event prediction for optimal alarm systems. Jin *et al.* [45] studied the event stream dissemination in large-scale online social network systems. Du *et al.* [46], [47] developed an event recommendation framework to help people with upcoming social events. Aa *et al.* [48] studied the conformance checking based on the analysis of observed events. In this work, we model the event influence on parking behaviors and enable PewLSTM to predict parking behaviors in events.

## 9 Conclusion

In this paper, we argue that periodic patterns shall be involved in parking behavior prediction. We have shown that parking behaviors exhibit strong periodic patterns, and such parking patterns shall be retained in classic LSTM models. To address this problem, we proposed a novel periodic weather-aware LSTM with event mechanisms for parking behavior prediction, which considers the periodic patterns from parking records and weather data. We also support event prediction for PewLSTM. We exhibited our model, data collection, and data processing in detail. We evaluate our proposed model with 13 parking lots in different cities, and experiments show that our model achieves 93.84% prediction accuracy, which is about 30% higher than the state-of-the-art prediction method.

## References


[1] Noah Kazis, "New York Has 81,875 Metered Parking Spaces, And Millions of Free Ones," https://www.statista.com/statistics/183505/number-of-vehicles-in-the-united-states-since-1990/, 2011.

[2] NYCEDC, "New Yorkers and Their Cars," https://edc.nyc/article/new-yorkers-and-their-cars, 2018.

[3] E. Brandt, "New Study Shows How Much Time and Money We Waste on Parking," https://www.thedrive.com/sheetmetal/12389/new-study-shows-how-much-time-and-money-we-waste-on-parking, 2017.

[4] G. Pierce and D. Shoup, "Getting the prices right: an evaluation of pricing parking by demand in San Francisco," *Journal of the American Planning Association*, 2013.

[5] C. Roman, R. Liao, P. Ball, S. Ou, and M. de Heaver, "Detecting on-street parking spaces in smart cities: performance evaluation of fixed and mobile sensing systems," *IEEE Transactions on Intelligent Transportation Systems*, 2018.

[6] A. Shahzad, J.-y. Choi, N. Xiong, Y.-G. Kim, and M. Lee, "Centralized connectivity for multiwireless edge computing and cellular platform: a smart vehicle parking system," *Wireless Communications and Mobile Computing*, 2018.

[7] J. Fang, A. Ma, H. Fan, M. Cai, and S. Song, "Research on smart parking guidance and parking recommendation algorithm," in *8th IEEE International Conference on Software Engineering and Service Science*, 2017.

[8] F. Caicedo, C. Blazquez, and P. Miranda, "Prediction of parking space availability in real time," *Expert Systems with Applications*, 2012.

[9] N. Feng, F. Zhang, J. Lin, J. Zhai, and X. Du, "Statistical analysis and prediction of parking behavior," in *IFIP International Conference on Network and Parallel Computing*, 2019.

[10] F. Zhang, N. Feng, Y. Liu, C. Yang, J. Zhai, S. Zhang, B. He, J. Lin, and X. Du, "PewLSTM: Periodic LSTM with Weather-Aware Gating Mechanism for Parking Behavior Prediction," in *IJCAI*, 2020.

[11] P. Mehta, D. F. McAuley, M. Brown, E. Sanchez, R. S. Tattersall, J. J. Manson, H. A. S. Collaboration *et al.*, "COVID-19: consider cytokine storm syndromes and immunosuppression," *Lancet (London, England)*, 2020.

[12] ThsParking, "ThsParking Homepage," http://www.thsparking.com, 2020.

[13] A. R. Alho and J. d. A. e. Silva, "Freight-trip generation model: predicting urban freight weekly parking demand from retail establishment characteristics," *Transportation Research Record*, 2014.

[14] S.-e. Yoo, P. K. Chong, T. Kim, J. Kang, D. Kim, C. Shin, K. Sung, and B. Jang, "PGS: Parking Guidance System based on wireless sensor network," in *International Symposium on Wireless Pervasive Computing*, 2008.

[15] M. Idris, E. Tamil, N. Noor, Z. Razak, and K. Fong, "Parking guidance system utilizing wireless sensor network and ultrasonic sensor," *Information Technology Journal*, 2009.

[16] R. Hössinger, P. Widhalm, M. Ulm, K. Heimbuchner, E. Wolf, R. Apel, and T. Uhlmann, "Development of a real-time model of the occupancy of short-term parking zones," *International Journal of Intelligent Transportation Systems Research*, 2014.

[17] X. Chen, "Parking occupancy prediction and pattern analysis," *Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep. CS229-2014*, 2014.

[18] Y. Zheng, S. Rajasegarar, and C. Leckie, "Parking availability prediction for sensor-enabled car parks in smart cities," in *IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2015.

[19] E. I. Vlahogianni, K. Kepaptsoglou, V. Tsetsos, and M. G. Karlaftis, "A real-time parking prediction system for smart cities," *Journal of Intelligent Transportation Systems*, 2016.

[20] C. Badii, P. Nesi, and I. Paoli, "Predicting available parking slots on critical and regular services by exploiting a range of open data," *IEEE Access*, 2018.

[21] G. A. Seber and A. J. Lee, *Linear regression analysis*. John Wiley & Sons, 2012.

[22] S. Le Cessie and J. C. Van Houwelingen, "Ridge estimators in logistic regression," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 1992.

[23] C. Hans, "Bayesian lasso regression," *Biometrika*, 2009.

[24] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE transactions on systems, man, and cybernetics*, 1991.

[25] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomForest," *R news*, 2002.

[26] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, 1986.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, 1997.

[28] J. Wang, T. Sun, B. Liu, Y. Cao, and H. Zhu, "CLVSA: A convolutional LSTM based variational sequence-to-sequence model with attention for predicting trends of financial markets," in *IJCAI*, 2019.

[29] Y. Zhang, X. Yang, J. Ivy, and M. Chi, "ATTAIN: attention-based time-aware LSTM networks for disease progression modeling," in *IJCAI*, 2019.

[30] T. Fu, T. N. Hoang, C. Xiao, and J. Sun, "DDL: Deep dictionary learning for predictive phenotyping," in *IJCAI*, 2019.

[31] T. Huang, G. Shen, and Z. Deng, "Leap-LSTM: Enhancing Long Short-Term Memory for Text Categorization," in *IJCAI*, 2019.

[32] B. Xing, L. Liao, D. Song, J. Wang, F. Zhang, Z. Wang, and H. Huang, "Earlier attention? aspect-aware LSTM for aspect-based sentiment analysis," in *IJCAI*, 2019.

[33] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Computer Vision – ACCV 2010*, 2011.

[34] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, 2017.

[35] Erin Schumaker, "Human-to-human transmission of new coronavirus reported in China," https://abcnews.go.com/Health/human-human-transmission-coronavirus-reported-china/story?id=68403105, 2020.

[36] D. K. Ahorsu, C.-Y. Lin, V. Imani, M. Saffari, M. D. Griffiths, and A. H. Pakpour, "The fear of COVID-19 scale: development and initial validation," *International journal of mental health and addiction*, 2020.

[37] Wikipedia, "Top Chinese expert: Epidemic to be under control by late April," https://en.wikipedia.org/wiki/2002\%E2\%80\%932004_SARS_outbreak, 2020.

[38] China Daily & Xinhua, "Top Chinese expert: Epidemic to be under control by late April," https://www.chinadailyhk.com/article/122555, 2020.

[39] A. Deshpande, S. Nath, P. B. Gibbons, and S. Seshan, "IRIS: Internet-scale Resource-Intensive Sensor Services," in *SIGMOD*, 2003.

[40] D. Teodorović and P. Lučić, "Intelligent parking systems," *European Journal of Operational Research*, 2006.

[41] X. Yu, G. Li, C. Chai, and N. Tang, "Reinforcement learning with tree-lstm for join order selection," in *ICDE*, 2020.

[42] B. Cortez, B. Carrera, Y.-J. Kim, and J.-Y. Jung, "An architecture for emergency event prediction using LSTM recurrent neural networks," *Expert Systems with Applications*, 2018.

[43] J. Yuen and A. Torralba, "A data-driven approach for event prediction," in *European Conference on Computer Vision*, 2010.

[44] M. Antunes, M. A. Turkman, and K. Turkman, "A Bayesian approach to event prediction," *Journal of Time Series Analysis*, 2003.

[45] H. Jin, C. Lin, H. Chen, and J. Liu, "QuickPoint: Efficiently Identifying Densest Sub-Graphs in Online Social Networks for Event Stream Dissemination," *TKDE*, 2020.

[46] Y. Du, X. Meng, Y. Zhang, and P. Lv, "GERF: A group event recommendation framework based on learning-to-rank," *TKDE*, 2020.

[47] Y. Du, X. Meng, and Y. Zhang, "CVTM: A content-venue-aware topic model for group event recommendation," *TKDE*, 2020.

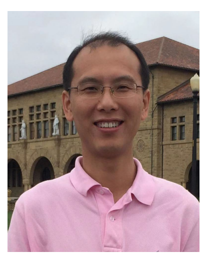
[48] H. van der Aa, H. Leopold, and H. A. Reijers, "Efficient process conformance checking on the basis of uncertain event-to-activity mappings," *TKDE*, 2020.


**Feng Zhang** received the bachelor degree from Xidian University in 2012, and the PhD degree in computer science from Tsinghua University in 2017. He is an assistant professor with the Key Laboratory of Data Engineering and Knowledge Engineer (MOE), Renmin University of China. His major research interests include big data management systems and parallel and distributed systems.

**Yani Liu** joined the Key Laboratory of Data Engineering and Knowledge Engineer (MOE), Renmin University of China in 2019. She is now a fourth-year undergraduate student at School of Information, Renmin University of China. Her major research interests include high performance computing, machine learning, and parallel and distributed systems.

**Ningxuan Feng** recived his B.S. degree in computer science from Renmin University of China in 2020. He is currently completing a PH.D in computer science at Tsinghua University. His major interests include neural network and AI security.

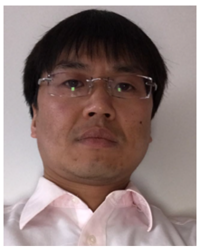
**Cheng Yang** received his B.E. degree and Ph.D. degree from Tsinghua University in 2014 and 2019, respectively. He is an assistant professor in Beijing University of Posts and Telecommunications. His research interests include natural language processing and network representation learning.



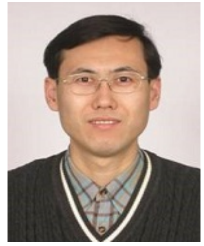
**Jidong Zhai** received the BS degree in computer science from University of Electronic Science and Technology of China in 2003, and PhD degree in computer science from Tsinghua University in 2010. He is an associate professor in Department of Computer Science and Technology, Tsinghua University. His research interests include performance evaluation for high performance computers, performance analysis and modeling of parallel applications.

**Shuhao Zhang** received the bachelor degree in computer engineering from Nanyang Technological University in 2014, and the PhD degree in computer science in National University of Singapore in 2019. He is an incoming Assistant Professor at the Singapore University of Technology and Design. His research interests include high performance computing, stream processing systems, and database systems.

**Bingsheng He** received the bachelor degree in computer science from Shanghai Jiao Tong University (1999-2003), and the PhD degree in computer science in Hong Kong University of Science and Technology (2003-2008). He is an Associate Professor in School of Computing, National University of Singapore. His research interests are high performance computing, distributed and parallel systems, and database systems.



**Jiazao Lin** received the Ph.D. degree in computing mathematics from Lanzhou University, China, in 2011. He is currently a Researcher with the Department of Information Management, Peking University, Beijing, China. His current research interests include the Internet of things, artificial intelligence, and data analysis.



**Xiao Zhang** received the master's degree from Renmin University and the PhD degree from the Institute of Computing Technology, Chinese Academy of Science, in 1998 and 2001, respectively, both in computer science and technology. He is a professor at the School of Information, Renmin University of China. His research interests include database architecture and big data management systems.

**Xiaoyong Du** obtained the B.S. degree from Hangzhou University, Zhengjiang, China, in 1983, the M.E. degree from Renmin University of China, Beijing, China, in 1988, and the Ph.D. degree from Nagoya Institute of Technology, Nagoya, Japan, in 1997. He is currently a professor with the School of Information, Renmin University of China. His current research interests include databases and intelligent information retrieval.