A Novel Travel Adviser Based on Improved Back-propagation Neural Network

Min Yang  
Department of Electronic Engineering  
Tsinghua University  
Beijing, China  
m-yang13@mails.tsinghua.edu.cn

Xuedan Zhang  
Department of Electronic Engineering  
Tsinghua University  
Beijing, China  
zhangxuedan@sz.tsinghua.edu.cn

Abstract — Nowadays, public bicycle system has emerged as a good solution to terminal traffic. Compared with other public transportation systems (like bus or taxi), public bicycle system is clean, cheap, flexible and convenient. In particular, public bicycle doesn’t need to follow fixed schedule. However, this flexibility of public bicycle system also brings in a problem: we don’t know whether there exist available bikes when we get to a certain bike station. This paper proposes a novel method to predict the number of available bikes after a given period of time so as to optimize users’ travel choice. In our paper, we use improved back-propagation neural network as our prediction algorithm and a novel prediction model considering the impact of surrounding stations. Experimental results show that our novel method can properly handle this non-linear problem.

Keywords - public bicycle system; back-propagation neural network; terminal traffic; time pattern

I. INTRODUCTION

With the rapid expansion of urban population and extreme lack of fossil energy, building a green and efficient public transportation system has nowadays became a big challenge. Generally speaking, there exist 3 public transportation systems: bus system, taxi system and subway system. Bus system and subway will undoubtedly be the first choice in terms of cleanness. However, neither subway nor bus system can completely handle the problem of terminal traffic: people who choose public transportation systems: bus system, taxi system and subway system. Bus system and subway system will undoubtedly be the first choice in terms of cleanness. However, neither subway nor bus system can completely handle the problem of terminal traffic: people who choose these two means of transportation still need to walk to their final destinations after they get off at corresponding stations. In terms of convenience, taxi is relatively a better choice. But taxi will cost the users more and pollute the air because of the use of gas.

Public bicycle system has drawn great attention for its advantages on cost, environmental protection, user experience and convenience. In general, public bicycle system has following strong points: (1) cleanliness. Public bicycle system can effectively reduce urban air pollution, improve urban air quality, save fossil energy, serve as an exercise mode, enhance people’s physical fitness and promote the city’s image; (2) flexibility. Public bicycle system doesn’t need to follow fixed schedule. So the users can start their journey whenever they want, which consequently promote users’ travel efficiency greatly; (3) convenience. Users can ride a bike to their final destinations. This means public bicycle system can properly handle the problem of terminal traffic; (4) low cost. Public bicycle system has a relatively lower cost and saves road resources. It can effectively relieve the lack of parking space and traffic jam.

As discussed above, public bicycle system doesn’t need to follow fixed schedule. This flexibility brings high efficiency as well as uncertainty [1]. We don’t know whether there are available bikes or bike stands when we get to a certain bike station. Fig. 1 describes this situation. Suppose the user’s name is Jack. There are many nearby bike stations to choose. But he doesn’t know whether there are bikes left when he gets to the corresponding station. So to us the biggest challenge is to accurately predict the number of available bikes or bike stands after a given period of time.

Many talented scholars have already done meaningful research in this field. Froehlich et al. [2] use Bayesian Networks for short-term and long-term (5-minute and 2-hour-ahead) predictions of bike availability in Barcelona’s Bicing system. Their prediction model incorporates the time of the day, currently available bikes and prediction window. Experimental results show remarkable improvements when compared with last view and historic trend algorithm. Kaltenbrunner et al. [3] use another quiet different algorithm. Considering the time sequence of available bikes, they use Autoregressive Moving Average (ARMA) time series model. Experimental results show that the number of input surrounding stations has an optimum. Based on the instability of available bikes’ number, Yoon et al. [4] use Autoregressive Integrated Moving Average (ARIMA) time series model to promote prediction accuracy. Chen et al. [1] not only consider internal, but also the impact of external factors (such as weather) on the prediction results. By applying Generalized Additive Model (GAM) [5], they can predict the number of available bikes as well as provide estimates of the waiting time distribution when there are not available bikes. Labadi et al. [14] develop an original discrete event approach for modelling and performance evaluation of public bicycle system by using Petri nets with time, inhibitor arcs and variable arc weights.
In this paper, we mainly address the problem mentioned above: choosing the fittest bike station at the very beginning. In other words, we provide a novel method to predicting the number of available bikes after a given period of time. In our paper, we use a kind of improved back-propagation neural network as our prediction algorithm and provide a novel prediction model. After that, we discuss factors that impact the performance of our prediction. At last, we will say something about our future work.

II. DATASET AND PARAMETER DEFINITION

A. Bicing System

Bicing is the name of a public bicycle system in Barcelona established on March 22, 2007. As a third-generation public bicycle system, it is similar to the Vélo'v service in Lyon or SmartDC system in Washington, DC. Its purpose is to handle the problem of terminal traffic and support a new means of public transportation. In general, there are about 421 stations in this system and the government of Barcelona also maintain a website from which we can get the real-time data (includes the number of available bikes and bike slots) of all stations. In fact, the data used in our paper was crawled from this website.

B. Dataset

We got the data used in this paper by crawling the website of Bicing system. The website offers real-time data of 421 bicycle stations. Our program collects station ID, available bikes and available bike slots of each of the 421 bicycle stations in Barcelona every 5 minutes. We separate the collected data into 2 parts. The first part is the data we collected from June 1st to July 17th, and we use it as our training set; the second part is the data we collected from September 1st to September 30th, and we use it as test set. Some of the data we collected is useless (because of the disconnection of network) and it may do harm to our prediction accuracy. So we reject the useless data and set its value to the previous value to get better prediction result.

C. Parameter Definition

In this part, I will give the definition of some parameters that will be used in our paper.

Normalized available bikes (NAB):

\[ NAB = \frac{\alpha}{\alpha + \beta} \]  
(1)

In this equation, \( \alpha \) stands for the number of available bikes at time \( t \), \( \beta \) stands for the number of available bike stands at time \( t \). Since the sum of available bikes and available bike slots is not a constant (some bikes or bike slots may be damaged, so we cannot use them), NAB can effectively reflect the percentage of available bikes.

Normalized activity score (NAS):

\[ NAS = \frac{\gamma - \alpha}{\alpha + \beta} \]  
(2)

In this equation, \( \gamma \) stands for the number of available bikes at time \( t-1 \). NAS can effectively indicate how active a station is at a given time \( t \). In this paper, we use the parameter to show the activeness of a station.

D. Time pattern analysis

Before we start our prediction, we must first discuss the time pattern of bike stations to optimize our prediction result [13].

In our paper, we use the two parameters defined in the last part (NAB and NAS) to indicate the time pattern of bike stations. Fig. 2 compares weekday NAB with weekend NAB over 24 hours at Station47; Fig. 3 compares weekday NAS with weekend NAS over 24 hours at Station47; According to these two pictures, we can see that both NAB and NAS differ greatly on weekdays and weekends. So we must discuss these two situations separately. In the following part of this article we only discuss weekday situation and we can handle weekend situation with the same method.
III. PREDICTION METHODOLOGY

A. Back-propagation Neural Network and Genetic Algorithm

Back-propagation (BP) neural network [6] is a kind of multi-layer feed-forward neural network. There are connections between neurons of adjacent layers while the errors of the output layer provide feedback to the input layer. To be exactly, BP neural network consists of 3 layers: the input layer, the middle layer (often called “the hidden layer”) and the output layer (we can see the structure clearly in Fig. 4). Inputs propagate from input layer to output layer while errors propagate in the opposite direction. It is a common method of training artificial neural networks used in conjunction with an optimization method such as gradient descent. The method calculates the gradient of a loss function with respect to all the weights in the network. The gradient is fed to the optimization method which in turn uses it to update the weights, in an attempt to minimize the loss function.

Genetic algorithm (GA) [7] is a search heuristic that mimics the process of natural selection. This heuristic is often used to figure out optimal solutions to optimization and search problems. In fact, Genetic algorithm belongs to evolutionary algorithm (EA), which generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In other words, when we want to solve optimization problems, we don’t need to think about exactly what we should do. On the contrary, we just need to figure out the fitness function and apply it to GA. In this way, we can easily get the final result without excessive deduction.

In this paper, we propose a novel prediction model which considers the effect of surrounding bicycle stations (Fig. 6). In contrast to ordinary prediction model (Fig. 5), the bicycle number of surrounding stations can be explicitly factored into prediction model, resulting in significant gains in terms of prediction accuracy.

In the ordinary model, we don’t consider the impact of surrounding stations on the prediction result. This model has 3 inputs: (1) time point. It represents time of the day and has 24 possible values corresponding to each of the hours of the day; (2) available bikes. It is the value of NAB at time t and is discretized into five categories, where a value of one corresponds to 0~20%, two corresponds to 20%~40%, etc.; (3) prediction window. The size of the prediction window, with six possible values corresponding to 10, 20, 30, 60, 90 and 120 minutes. This model also has an output named delta, which corresponds to the change of NAB from time t to time t+PW. So the final prediction result is made by adding the value of delta to current NAB. In our new prediction model, we consider the impact of surrounding stations (in this paper we only consider the impact of their NABs). Compared with the ordinary model, the new model has more inputs. They are the NABs of the chosen surrounding stations at time t. Fig. 7 and Fig. 8 shows the performance curves of ordinary and new prediction model (use improved BP as prediction algorithm). Table 1 compares their performances in detail. As shown in table 1, our new model obtains about 38.2% improvement in performance result.
According to our experimental results, we know that the number of surrounding stations considered in our prediction model also has a great influence on the final prediction result. As shown in Fig. 9, when the number of surrounding stations considered in our model is too big or too small, the prediction accuracy will descend. In this case, the best number of surrounding stations is 10. In other words, when we determine our prediction model, we must first calculate the fittest number of considered surrounding stations.

3) the selection strategy of surrounding stations

In the last section, we have discussed the optimal number of considered surrounding stations. And in that part, we choose these stations by distance. In other words, we choose the nearest surrounding stations. And in this part, we will choose surrounding stations that share the same time pattern (we use NAB) with our target station.

Fig. 10 shows the fitted curve of considering 10 surrounding stations that have approximate time pattern. Compared with Fig. 8, we can see that when we consider stations that have approximate time pattern, we can get better prediction accuracy.

C. Improved BP Neural Network

In this paper, we choose a kind of improved BP neural network as our prediction algorithm. As we know, the performance of BP neural network is determined by 3 factors: the structure of neural network (in this case we mean the number of hidden layers and their corresponding
neurons), the initial weights and biases and its training method. As to the number of hidden layers and the number of neurons each layer has, we often follow empirical equations listed below:

\[ m = \sqrt{n + l + \alpha} \]  
\[ m = \log_2 n \]  
\[ m = \sqrt{nl} \]

In the above 3 equations, \( m \), \( l \), \( n \) stand for the number of neurons in hidden layers, input layer and output layer, respectively. \( \alpha \) stands for a constant ranging from zero to one. And we often choose gradient descent method as training algorithm. However, this training algorithm doesn’t guarantee global optimization solution. In fact, we often get local optima because of random initial weights and biases. To solve this problem, we use GA to get the optimum initial weights and biases.

The improved algorithm based on genetic algorithm mainly fix following steps:

1) Use the fittest empirical equation to determine the structure of BP neural network (often we only need one hidden layer);
2) Choose gradient descent method as training algorithm;
3) Determine the fitness function. As to the fitness function, the fittest chromosome has the biggest value, in our program, we use square of the difference between output computed with initial weights and biases and true output as our fittest function;
4) Initiation. The population size depends on the nature of the problem, but typically contains several hundreds or thousands of possible solutions. Often, the initial population is generated randomly, allowing the entire range of possible solutions (the search space).
5) Selection. Each generation we choose chromosomes that have bigger fitness;
6) Mutation and crossover. Use mutation and crossover to get the next generation;
7) Repeat step(5)~step(6) until the number of generation get its maximal value;
8) Use the initial weights and biases for the training of BP neural network;

Fig. 11 indicate the training procedures of improved BP neural network. From this figure we can see that the neural network gets its target after 15 iterations. Compared with the original BP neural network (as displayed in Fig. 12), the improved one has a faster convergence velocity. Fig. 13 and Fig. 14 show the performance curves of improved BP neural network and original BP neural network (they share the same structure and training method). And table 2 compared their performance clearly. From this table we can see that the improved BP neural network obtains 24.3% improvement in performance. In other words, the improved algorithm has better generalization ability.

### TABLE II. PREDICTION ERROR OF ORDINARY AND IMPROVED BP NEURAL NETWORK

<table>
<thead>
<tr>
<th>The style of neural network</th>
<th>The number of neurons in input layer</th>
<th>The number of neurons in hidden layer</th>
<th>Prediction errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard BP</td>
<td>13</td>
<td>25</td>
<td>1.52(PW=30min)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.08(PW=120min)</td>
</tr>
<tr>
<td>Improved BP</td>
<td>13</td>
<td>25</td>
<td>1.15(PW=30min)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.32(PW=120min)</td>
</tr>
</tbody>
</table>

### D. Factors that impact prediction result

In this part, we use the novel model and improved BP neural network to illustrate factors that impact prediction result [13].

1) Prediction Window

As shown in Fig. 14, the prediction value is about 0.71 times of real value. That is to say, the error is about 0.29 times of real value. As we know, real value (real
delta) will increase with the increase of prediction window. So we can conclude that the error will increase as a consequence. On the other hand, real value (real data) will decrease with the decrease of prediction window, then the error will decrease consequently. In a word, bigger prediction window will contribute to bigger error.

2) Time Point of Prediction

We can see from Fig. 2 that NAS gets its first peak on 9:00 a.m. and first valley on 5:00 a.m. This tells us that the bicycle station is active around 9:00 a.m. and relatively inactive around 5:00 a.m. To explore the impact of time point we can make our prediction on 9:00 a.m. and 5:00 a.m. and compare their prediction results. Fig. 15 and Fig. 16 show the prediction results in these two situations. From these two pictures, we know that prediction result will be more accurate when the station is relatively not so active.

![Figure 15. The fitted curve on 5:00 a.m.](image)

![Figure 16. The fitted curve on 9:00 a.m.](image)

IV. CONCLUSION AND FUTURE WORK

Public bicycle system nowadays has emerged as a good solution to terminal traffic [12]. Compared with other public transportation systems (like bus or taxi), public bicycle system is clean, cheap, flexible and convenient. However, the flexibility of public bicycle also brings in a problem: we don’t know whether there exist available bikes when we get to a certain bike station. In this paper we have handled this problem. We use improved back-propagation neural network as our prediction algorithm and a novel prediction model considering the impact of surrounding stations to predict the number of available bikes at the given station after a given period of time. Prediction results show that our novel method can get a relatively better prediction result than previous prediction model with ordinary back-propagation neural network.

As we know, the distribution of bike stations is also a critical problem in public bicycle system [9–11]. In the future, we will mainly do some research on the optimal distribution of bike stations to promote global efficiency.

REFERENCES


