

# The Prediction of the Peak Time of People Taking School Bus Based on Martingale Process

Yuxuan Zhou, Shujun Liu, Yufeng Gui\*

College of Science, Wuhan University of Technology, Wuhan, China

Email: \*guiyufeng@whut.edu.cn

**How to cite this paper:** Zhou, Y.X., Liu, S.J. and Gui, Y.F. (2017) The Prediction of the Peak Time of People Taking School Bus Based on Martingale Process. *Applied Mathematics*, 8, 630-636.

<https://doi.org/10.4236/am.2017.85049>

**Received:** March 24, 2017

**Accepted:** May 15, 2017

**Published:** May 18, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

As the main tool for students and professors to come and go to different districts in-campus, the school bus is of great importance. In order to improve its efficiency and convenient students and professors, a rational schedule of departure is badly needed. A model is established to predict the peak time of people taking school bus based on martingale process, and it is solved according to stopping time of martingale. Then it is applied to Wuhan University of Technology. A large amount of data is collected and the peak time for each day is predicted combined with the actual situation of the college. In doing so, suggestions are given for those who are in charge of the school buses.

## Keywords

School Bus, Martingale Process, Prediction, Peak Time

---

## 1. Introduction

There are more and more districts being constructed to provide students with comfortable environment to study and live in [1] [2]. In a meanwhile, it becomes further and further from one district to another, which makes it inconvenient for students and professors to travel in the campus [3]. As people need to come and go to different districts in a short period of time, the school bus has become the most important means of transportation [4]. Accordingly, the optimization of its schedule is badly needed to convenient those who take the bus [5].

According to literatures, problems of the optimization of distributions of the bus stations, dispatches and routes of the buses have already been discussed for many times [6]. And good algorithms are given by the predecessors, such as classic VRP (Vehicle Routing Problem) Algorithm [1] and complex SBRP Algorithm [2]. However, only few literatures that focus on predicting peak time of

people taking the buses can be founded. The main problem exists in school buses in some universities is that the number of buses is not big enough when it comes to peak time while it is too much when it comes to non-peak time due to the irrational dispatch. Thus, a martingale process based on collected data is established and it can predict peak time of the day according to stopping time theory [7].

## 2. The Peak Time Prediction Model

### 2.1. Definitions of Peak Time

The phenomenon of being crowded and short of buses in a time when there are a large amount of people arriving at the station continuously in adjacent period of time is of a high possibility to happen in-campus according to some literatures [3]. Thus, a peak time is defined as a period of time when there are more than  $a_i$  people arriving at the station continuously in adjacent  $L$  unit of time. Where  $i$  is the  $i$ th workday,  $a_i$  is a threshold that stands the number of people taking the school bus in adjacent  $L$  continuous unit of time on the  $i$ th workday, which is related to real situations.

### 2.2. The Establishment of the Model

In order to establish our model, several variables are assumed as follows (Table 1):

As it is inferred in 3.1 that a peak time is a period of time when there are more than  $a_i$  people arriving at the station continuously in adjacent  $L$  unit of time, it can be concluded that the total number of people taking the school bus in adjacent  $L$  unit of time should be  $La_i (i = 1, \dots, 5)$  when it comes to peak time. That is:

$$N_i = \inf \left\{ n_i : \sum_{j=n_i-L_i+1}^n X_i^j = La_i \right\} \tag{1}$$

Hence, the peak time prediction model on the  $i$ th workday can be expressed as:

$$E[N] = E \left[ \inf \left\{ n_i : \sum_{j=n_i-L_i+1}^n X_i^j = La_i \right\} \right] \tag{2}$$

### 2.3. Model Solving

The stopping time theory of martingale process is applied to solve the model.

#### 2.3.1. Establishment of Martingale Process

At first, a stochastic process  $L: Z_n^{i(1)}, \dots, Z_n^{i(l)}$  is defined, where

**Table 1.** Symbol description.

Variable	Definition
$X_n^i$	The number of people taking the bus on the $n$ th time unit of the $i$ th workday.
$N_i$	The exact time of a peak time.
$p_i$	$P(X_n^i = a_i)$

$$Z_n^{i(l)} = \{z_n^{i(l)}, n \geq l\}, l = 1, \dots, L; i = 1, \dots, 5.$$

In a meanwhile, it can be concluded that:

$$Z_n^{i(l)} = 1 - p^{-1} I \left( \sum_{j=n_l-L+1}^n X_j^n = La_i \right) \tag{3}$$

It is obvious that  $E[z_n^{i(l)}] = 0$ .

Then a martingale process can be established:

$$Y_n^i = \sum_{j=L}^n z_j^{i(L)} + \sum_{l=1}^{L-1} z_n^{i(l)}, n \geq L, i = 1, \dots, 5 \tag{4}$$

### 2.3.2. Prove of $Y_n^i$ Being a Martingale Process

At first, it is known that:

$$E[z_{n+1}^{i(l)} | X_1^i, \dots, X_n^i] = z_n^{i(l-1)}, l = 1, \dots, L, n \geq l-1, i = 1, \dots, 5 \tag{5}$$

According to literature [3],  $z_n^{i(0)} = 0$ . Thus, regulate that  $z_n^{i(0)} = 0$ .

Then it can be proved that  $Y_n^i (i = 1, \dots, 5)$  is a martingale process, that is to prove

$$E[Y_{n+1}^i | X_1^i, \dots, X_n^i] = Y_n^i, n \geq L, i = 1, \dots, 5 \tag{6}$$

However,

$$\begin{aligned} & E[Y_{n+1}^i | X_1^i, \dots, X_n^i] \\ &= E \left[ \sum_{j=L}^{n+1} z_j^{i(L)} + \sum_{l=1}^{L-1} z_{n+1}^{i(l)} | X_1^i, \dots, X_n^i \right] \\ &= E \left[ \sum_{j=L}^n z_j^{i(L)} | X_1^i, \dots, X_n^i \right] + E[z_{n+1}^{i(L)} | X_1^i, \dots, X_n^i] + E \left[ \sum_{l=1}^{L-1} z_{n+1}^{i(l)} | X_1^i, \dots, X_n^i \right] \\ &= E \left[ \sum_{j=L}^n z_j^{i(L)} | X_1^i, \dots, X_n^i \right] + E \left[ \sum_{l=1}^L z_{n+1}^{i(l)} | X_1^i, \dots, X_n^i \right] \end{aligned}$$

According to (5), it is clear that:

$$E \left[ \sum_{j=L}^n z_j^{i(L)} | X_1^i, \dots, X_n^i \right] = \sum_{j=L}^n z_j^{i(L)} \tag{7}$$

while it is known that:

$$E[z_{n+1}^{i(l)} | X_1^i, \dots, X_n^i] = z_n^{i(l-1)} \tag{8}$$

Hence, it can be concluded that:

$$E[Y_{n+1}^i | X_1^i, \dots, X_n^i] = \sum_{j=L}^n z_j^{i(L)} + \sum_{l=1}^L z_n^{i(l-1)} \tag{9}$$

As  $z_n^{i(0)} = 0$ , it is clear that:

$$\begin{aligned} E[Y_{n+1}^i | X_1^i, \dots, X_n^i] &= \sum_{j=L}^n z_j^{i(L)} + \sum_{l=1}^L z_n^{i(l-1)} = \sum_{j=L}^n z_j^{i(L)} + \sum_{l=2}^L z_n^{i(l-1)} \\ &= \sum_{j=L}^n z_j^{i(L)} + \sum_{l=2}^{L-1} z_n^{i(l)} = Y_n^i \end{aligned}$$

From what have been discussed above,  $Y_n^i (i = 1, \dots, 5)$  is proved to be a martingale process and  $E[Y_L^i] = E\left[\sum_{l=1}^L z_L^{(l)}\right] = \sum_{l=1}^L E[z_L^{(l)}] = 0$ .

**2.3.3.  $N_i$  is a Stopping Time of  $\{X_n^i, n \geq L, i = 1, \dots, 5\}$  and Suits What Stopping Time Theory Require**

The demonstration can be founded in literature [3].

**2.3.4. Prove  $E[N_i] = \frac{p_i^{-L} - 1}{1 - p_i}, i = 1, \dots, 5$**

At first, proof of  $Y_{N_i}^i = N_i + \frac{1 - p_i^{-L}}{1 - p_i}$  is needed.

According to (4), it can be concluded that:

$$Y_{N_i}^i = \sum_{n=L}^{N_i} z_n^{(L)} + \sum_{l=1}^{L-1} z_{N_i}^{(l)} \tag{10}$$

1) For  $z_n^{(L)}, n = L, L + 1, \dots, N_i$

As  $\sum_{j=n-L+1}^n X_j^i = La_i$  means there are continuously  $L$  units of  $a_i$ ,  $N_i$  is the first time of a day that have continuously  $L$  units of  $a_i$  on the  $n$ th day of a week.

Therefore, only when  $n = N_i$ , it can concluded that  $\sum_{j=n-L+1}^n X_j^i = La_i$ .

As for  $z_n^{(L)}, n = L, L + 1, \dots, N_i - 1$ , as  $\sum_{j=n-L+1}^n X_j^i \neq La_i$ , it is clear that:

$$z_n^{(L)} = \begin{cases} 1 - p_i^{-L} & n = N_i \\ 1 & L \leq n \leq N_i - 1 \end{cases} \tag{11}$$

which means:

$$\sum_{n=L}^{N_i} z_n^{(L)} = N_i - L + 1 - p_i^{-L} \tag{12}$$

2) For  $z_{N_i}^{(l)}, l = 1, \dots, L - 1$

According to definition of  $z_{N_i}^{(l)}, l = 1, \dots, L - 1$  and  $N_i$ , it is obvious that  $z_{N_i}^{(l)} = 1 - p_i^{-l}, l = 1, \dots, L - 1, i = 1, \dots, 5$ . Thus, it is clear that:

$$\sum_{l=1}^{L-1} z_{N_i}^{(l)} = \sum_{l=1}^{L-1} (1 - p_i^{-l}) = L - 1 - \frac{p_i^{-1}(1 - p_i^{-(L-1)})}{1 - p_i^{-1}} = L - 1 + \frac{1 + p_i^{1-L}}{1 - p_i} \tag{13}$$

According to what have been discussed above, it is known that:

$$Y_{N_i}^i = N_i - L + 1 - p_i^{-L} + L - 1 + \frac{1 - p_i^{1-L}}{1 - p_i} = N_i + \frac{1 - p_i^{-L}}{1 - p_i} \tag{14}$$

As  $N_i$  is a stopping time of  $\{X_n^i, n \geq L, i = 1, \dots, 5\}$  and suits what stopping time theory require, it is obvious that:

$$E[Y_{N_i}^i] = E[Y_L^i] \tag{15}$$

That is:

$$E[N_i] = \frac{p_i^{-L} - 1}{1 - p_i}, i = 1, \dots, 5 \quad (16)$$

### 3. Empirical Analysis

#### 3.1. Taking Wuhan University of Technology for Example

Taking Wuhan University of Technology for example, it has 4 districts in Ma Fangshan district, which are Nan Hu district, Jian Hu district, Dong Yuan district and Xi Yuan district, and it also contains an accommodation area for students outside of the campus named Shengsheng. It can be clearly seen from the distribution of the school that the need for students to take the school bus is very urgent as there are 5 districts. Therefore, a good dispatch of the school bus is of a great importance.

To collect data, the number of people those take the school bus in every 3 minutes, from Monday to Friday, 7:30 a.m. to 9:00 a.m. is recorded. By using martingale process to predict peak time for each day, the schedule of the departure time can be optimized.

The data was collected in Nan Hu district, as it is the biggest district of the 5 ones and holds the largest population. There are 14 school buses in Wuhan University of Technology. According to our research, the average traffic for each day can be obtained. After studying on the data, it appears to relate to the total number of courses of the whole school, which makes sense.

#### 3.2. The Determine of $a_i$

Analyzing the data,  $a_i$  is related to the real situations. Therefore, the values of  $a_i$  combine with the real situations are determined.

As the number of courses is of a large amount in Monday, the traffic should be largest of the weak. Thus, a peak time for Monday is a period of time when there are more than 9 people arriving at the station continuously in adjacent  $L$  unit of time.

As the number of courses is not very crowded in Tuesday and Wednesday, a peak time for Tuesday and Wednesday is a period of time when there are more than 7 people arriving at the station continuously in adjacent  $L$  unit of time.

As most professors have their meetings on Thursday, the number of courses is not very large. Thus, a peak time for Thursday is a period of time when there are more than 5 people arriving at the station continuously in adjacent  $L$  unit of time.

As there are usually have practical courses on Friday, traffic is a little larger than Thursday. Thus, a peak time for Friday is a period of time when there are more than 6 people arriving at the station continuously in adjacent  $L$  unit of time.

That is **Table 2**.

#### 3.3. The Determine of $L$

The value of  $a_i$  has been determined by the feature of the data in section 3.2

while the value of  $L$  is not. According to what have been observed in each school bus station in Wuhan University of Technology, it will significantly be crowded when there are  $a_i$  person arriving at the station to wait for the bus continuously in adjacent 3 unit of time. Thus  $L$  is determined as:

$$L = 3. \tag{17}$$

### 3.4. Results and Conclusions

Applying the model in section 2, the peak time for each day is predicted as shown in the **Table 3**.

It can be seen from the table that the peak time for Monday is around 12:05 and 17:25.  $T$  the peak time for Tuesday is around 10:33, the peak time for Wednesday is around 14:33, the peak time for Thursday is around 17:25 and the peak time for Friday is around 11:23. According to our research, it is corresponded to the real situation.

As for the occurrence of 2 peak times on Monday, according to the collected data, there are usually more people taking the bus on Monday in Wuhan University of Technology, thus there are 2 peak times that 9 person arrives at the station continuously in adjacent 3 unit of time in the original data. In this way, there are 2 peak times predicted by the model.

### 4. Conclusions and Suggestions

A peak time prediction model based on martingale process is established and solved with the stopping time theory. Taking Wuhan University of Technology for example, the peak time for each day is predicted by applying the model.

The suggestion is to reduce the time for departure to increase the trips when it comes to peak times. In this way, the efficiency of the school bus will increase and the cost will decrease at the same time. It will also provide a lot of convenience for students and professors. Besides, it is also suggested that students should prevent these time to take the school bus as far as possible.

### Acknowledgements

The paper is financially supported by Students innovation and entrepreneurship

**Table 2.** The determine of  $a_i$ .

$i$	1	2	3	4	5
$a_i$	9	7	7	5	6

**Table 3.** The prediction of the peak time for each day.

	Monday	Tuesday	Wednesday	Thursday	Friday
$p_i$	1/5	2/15	7/30	1/5	1/6
$N_i$	155	20.9	101	155	39
Peak time	12:05 and 17:25	10:33	14:33	17:25	11:23

training program, Wuhan University of Technology, China (No.166814007).

## References

- [1] Song, Y.Z. (2013) Application of Heuristic Algorithm Based on Simulated Annealing Algorithm in VRP. Master's Thesis, Huazhong Normal University, Wuhan.
- [2] School of Economics and Management, Henan University of Science and Technology (2013) Research on the Model and Algorithm of School Bus Routing Problem. *Journal of Henan University (Natural Science Edition)*, **43**, 682-691.
- [3] Ma, M. (2011) The Martingale Prediction of the Peak of Infectious Diseases. *Practice and Understanding of Mathematics*, **41**, 103-107.
- [4] Bao, L. and Li, D. (2006) School Bus Weekend Peak Traffic Volume Forecast and Special Date Adjustment. *Road Traffic and Safety*, **6**, 29-32.
- [5] Zhou, Q. (2013) Research on Optimization of School Bus Dispatching. *Journal of Hebei University of Technology*, **23**, 32-35.
- [6] Liu, Q.S. (2013) Meta-Heuristic Algorithm Design and GIS Integration for Bus Routing Problem. Master's Degree Thesis, Henan University, Kaifeng.
- [7] Liu, S.S. (2014) The Cloud Is Not Be Able to Rain. The Shijiazhuang Air Quality in the First "Severe" in June. *Yanzhao Metropolis Net*. 2014-06-17(3).



**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact [am@scirp.org](mailto:am@scirp.org)