

The Evaluation of Bus Operation Reliability at the Stops Based on the Monte Carlo Simulation

Lin Qiang^{1, a}, Zhai Lei^{2, b}, Liu Song^{3, c}

¹ School of Economics and Management Chang'an University Xi'an, 710064, China

² School of Economics and Management Chang'an University Xi'an, 710064, China

³ School of Economics and Management Chang'an University Xi'an, 710064, China

^aemail: linqiang6666@gmail.com, ^bemail: 303597966@qq.com, ^cemail: liusongll@qq.com

Keywords: Monte Carlo Simulation; Public Transport; Operation Reliability

Abstract. To enhance the bus operation stability is the key to improve bus service. Traditional evaluation of the bus operation stability always relies on the whole line, the stability of the bus operation at the stops can not be reflected. Besides, traditional evaluation of the bus operation stability drawing from the data which is collected in a period of short time can not reflect the affection from the randomness of the environment. This article discussed the bus operation reliability based on the stops, using the Monte Carlo simulation to deal with the probabilistic affection from the operation environment. Setting a bus line from a city in Zhejiang province as a example to illustrate how to analyse and calculate the bus operation reliability on the stops, in the end we get the proper evaluation indicators which is from simulation sample capacity in a certain extent error precision.

The Meaning of Bus Operation Reliability'S Punctuality at the Stops

The evaluation indicators of the bus operation reliability's punctuality [1] [2] [3] based on the whole line can reflect whether buses arrive at terminus punctually and the probability of it. But it only use the operation time between the origin station and terminus to analyse the operation reliability. Sometimes a number of drivers not only slower the speed of buses where the passengers is sufficient but also accelerate it where the passengers is not so much. Maybe in this situation, the evaluation of bus operation reliability on the stops based on the whole line is stable, but it is not meet with people's cognition.

The meaning of bus operation reliability's punctuality at the stops is that the probability of whether buses reach the stops on time according to the schedule, usually it's used to microcosmic evaluation of bus operation, and it can reflect the probability that whether buses reach the stops on time, avoid the shortcoming of the bus operation reliability evaluation based on the whole line.

The Evaluation Model of Bus Operation Reliability at the Stops

$$PIS_S = P\{T_{Arr} \in [T_{Sch} + \theta_1, T_{sch} + \theta_2]\} = P\{T_{Arr} - T_{Sch} \in [\theta_1 - \theta_2]\} \quad (1)$$

$T_{Arr, s}$ —The actual bus arriving time at the stop S

$T_{Sch, s}$ —The bus arriving schedule time at the stop S

θ_1, θ_2 The parameters of time

It is not suitable for no stop schedule bus lines. In theory the interval time of every bus is equal when it is equal at origin stop. In fact the interval time of every bus is affected by various environment factors, it is not always equal among the whole bus stops. So we can use the interval bus arriving time instead of the actual bus arriving time to calculate

$$PIS_S PIS_S = P\{H_s - H_0 \in [\theta_1, \theta_2]\} \quad (2)$$

H_s —The interval bus arriving time between one bus and the next one at stop S

H_0 —The interval time between one bus and the next one at origin station

θ_1, θ_2 —The parameters of time

For the parameters of time θ_1, θ_2 , usually use the linear function of H_0 to instead of them. When $\theta_1 = 0$ $\theta_2 = k \cdot H_0$, the expression of PIS as follows

$$PIS_S = P \left\{ \in \left[\frac{|H_S - H_0|}{H_0} \leq k \right] \right\} \quad (3)$$

It is easy to learn that $\frac{|H_S - H_0|}{H_0}$ is the bus operation deviation ratio at stop S, and it is used as the main evaluation indicator we can calculate one PIS_S for every stop from GPS data in one direction every day. But the data is affected by large quantities of random environmental factors, and it can't be used to illustrate the normal operation situation because of the sample capacity is not sufficient. Also it is not practical to gather a lot of data to calculate PIS_S owing to the high time cost and economic cost. Considering those factors we use Monte Carlo simulation to calculate PIS_S .

The Application of Monte Carlo Simulation in Bus Operation Reliability at the Stops

Monte Carlo simulation

Practical problems contain large quantities of random factors, simplify many assumption is needed in modeling to analyse problems. But it can't reflect the true situation that used to deal with the practical problems.

Monte Carlo simulation[4] [5] is a method which is different from the normal numerical computation method, it is a branch of math experiment. The basic thought of Monte Carlo is that when the solution of the problem is the probability of a certain event, we can use some mathematical test method to get the frequency of the event, the frequency is the solution of the problem.

To a certain distribution density function $p(x)$ which is from the random variable factor ξ , the mathematical expectation of its function $f(\xi)$ as follows:

$$E[f(\xi)] = \int_0^{\infty} f(x)p(x)dx \quad (4)$$

The observed value $\xi_1, \xi_2, \dots, \xi_n$ are drawing from random variable ξ obtained by the test, and then we can get the results $f(\xi_1), f(\xi_2), \dots, f(\xi_n)$, the arithmetic mean as follows:

$$\bar{f} = \frac{1}{n} \sum_{i=1}^n f(\xi_i) \quad (5)$$

According to the law of large numbers, for arbitrarily small positive number ε ,

$$\lim_{n \rightarrow \infty} P\{|\bar{f} - E[f(\xi)]| < \varepsilon\} = 1 \quad (6)$$

When n is large enough, we can use the arithmetic mean of $f(\xi_i)$ instead of the original problems' results.

The sample capacity of the Monte Carlo simulation

Assume the true average value of the multivariate random variable function is $\mu_z = E\{z\}$, the variance is $\sigma_z^2 = Var[z]$, it's statistical mean of test is \bar{z} . So the expected value and variance of \bar{z} are as follows:

$$E = (\bar{z}) = E \left\{ \frac{1}{N_s} \sum_{i=1}^{N_s} Z_i \right\} = \frac{1}{N_s} \sum_{i=1}^{N_s} E\{z_i\} = \frac{1}{N_s} N_s \mu_z = \mu_z \quad (7)$$

$$Var\{\bar{z}\} = Var \left\{ \frac{1}{N_s} \sum_{i=1}^{N_s} Z_i \right\} = \frac{1}{N_s^2} \sum_{i=1}^{N_s} Var\{z_i\} = \frac{1}{N_s} (N_s \sigma^2) = \frac{\sigma^2}{N_s} \quad (8)$$

According to the central limit theorem, if N_s is large enough, change normal distribution $\bar{z} \sim N(\mu_z, \sigma^2)$ into standard normal distribution as follows:

$$\frac{\bar{z} - \mu_z}{\sigma_z / \sqrt{N_s}} \sim N(0, 1) \quad (9)$$

At a given level of significance α , we can get:

$$-x_{\alpha/2} \leq \frac{\bar{z} - \mu_z}{\sigma_z / \sqrt{N_s}} \leq x_{\alpha/2} \quad (10)$$

When $x_{\alpha/2} > 0$

$$p \left\{ \left| \frac{\bar{z} - \mu_z}{\sigma_z / \sqrt{N_s}} \right| \leq x_{\alpha/2} \right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-x_{\alpha/2}}^{+x_{\alpha/2}} e^{-x^2/2} dx = 1 - \alpha \quad (11)$$

With the increase of the sample capacity N_s , also the accuracy of estimated value is increasing.

The relationship between \bar{z} and μ_z as follows:

$$|\bar{z} - \mu_z| \leq \frac{x_{\alpha/2} \sigma_z}{\sqrt{N_s}} \quad (12)$$

The relative error as follows

$$\Delta_z = \left| \frac{\bar{z} - \mu_z}{\mu_z} \right| \leq \frac{x_{\alpha/2} \sigma_z}{\sqrt{N_s} \mu_z} = \frac{x_{\alpha/2} \sigma_z}{\sqrt{N_s}} \delta \quad (13)$$

If set the given accuracy is Δ_z , the sample capacity for every stop as follows:

$$N_s \approx x_{\alpha/2}^2 \frac{\delta^2}{\Delta_z^2} \quad (14)$$

The calculation of bus operation reliability's punctuality on the stops

A bus line contains 40 stops in a city from Zhejiang province. To analyse the data from GPS on a working day, we can get the start time departure time and arrival time. Through these data we can calculate the PIS_s , so that we can learn the operation stability.

We can only get one PIS_s from data at a stop everyday. The results as follows.

Table 1 Test Data

Stop	1	2	3	4	5	6	7	8	9	10
PIS	0.57	0.53	0.59	0.65	0.65	0.63	0.70	0.72	0.74	0.72
Stop	11	12	13	14	15	16	17	18	19	20
PIS	0.70	0.67	0.70	0.70	0.70	0.70	0.68	0.68	0.66	0.59
Stop	21	22	23	24	25	26	27	28	29	30
PIS	0.61	0.61	0.59	0.59	0.59	0.59	0.57	0.57	0.57	0.57
Stop	31	32	33	34	35	36	37	38	39	40
PIS	0.54	0.54	0.52	0.50	0.45	0.50	0.50	0.50	0.48	0.50

In order to decrease the error caused by the short term data survey, we use Monte Carlo simulation to analyse data. Generally speaking the error is proper under the 3%, looking up the standard normal distribution table, $x_{\alpha/2}$ is 2.107. Through sample trial calculation we get δ is 0.1346, so the N_s is 89.

Use EXCEL produce 40 columns random numbers between 0 and 1 expressed by $s_1, s_2, s_3 \dots s_n$, and every column has 89 numbers. These 40 columns random numbers represent 40 bus stops. Set x_i as bus arriving time at the stop i . We can get $PIS_s \geq 0.45$, so set the critical value is 0.45. The relationship between s_i and x_i as follows:

$$x_i = \begin{cases} 1, & s_i \in [0.45, 1) \\ 0, & s_i \in (0, 0.45) \end{cases} \quad i = 1, 2, 3 \dots, n \quad (15)$$

n —Total number of bus stops

Through this formula, we get 3560 samples in total. These are the time state at the bus stop. The statistics result's as follows:

Table 2 Simulation Data

Stop	1	2	3	4	5	6	7	8	9	10
PIS simulate value	0.43	0.45	0.63	0.59	0.63	0.57	0.55	0.55	0.54	0.50
Stop	11	12	13	14	15	16	17	18	19	20
PIS simulate value	0.60	0.61	0.56	0.46	0.43	0.50	0.57	0.53	0.55	0.64
Stop	21	22	23	24	25	26	27	28	29	30
PIS simulate value	0.49	0.47	0.53	0.54	0.58	0.57	0.59	0.56	0.62	0.54
Stop	31	32	33	34	35	36	37	38	39	40
PIS simulate value	0.57	0.59	0.63	0.48	0.57	0.52	0.59	0.56	0.60	0.50

Conclusion

a: Through Monte Carlo simulation ,the error is under 3% between survey value and simulative value about bus operation reliability, and the reliability is higher than normal analysis methods.

b: The sophisticated environment factors in bus operation are considered in the simulation procedure, but the traditional analysis can't contain such influencing factors.

c: Use PIS_5 to evaluate bus operation reliability can recognize delay unloading and driving fast between the origin station and the terminus. The result is objective and true.

References

- [1]Zhang,Y.S..Research and evaluation of public transportation operation reliability in metropolis. Graduate Thesis, Beijing Jiaotong University, China, 2008
- [2] Abkowitz,M.,Engelstein,I.Factors affecting running time on transit routes. Transportation Research Part A 17A(2),107-113,1983
- [3] Zhang Man, Li Wen-quan. Factors affecting headway regularity on bus routes [J]. Journal of Southeast University. 2013:29 (1): 99-102
- [4] Bi Zhong-wei. Random simulation times of engineering reliability [J]. HYDRO-SCIENCE AND ENGINEERING.2005 (1) 44-46
- [5] Fan Hai-yan, Yang Xiao-guang. Application of Monte-Carlo method in the reliability calculation of bus travel time [J]. University of Shanghai for Science and Technology. 2006:28 (1):59-62