

Research Article

A Capacity-Restraint Transit Assignment Model When a Predetermination Method Indicates the Invalidity of Time Independence

Haoyang Ding,^{1,2} Yu Bao,^{1,2} Sida Luo,^{1,2} Hanxia Shen,^{1,2} Wei Wang,^{1,2} and Man Long^{1,2}

¹Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210096, China

²Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Nanjing 210096, China

Correspondence should be addressed to Haoyang Ding; dinghy123@hotmail.com

Received 8 May 2015; Revised 25 July 2015; Accepted 28 July 2015

Academic Editor: Yuanchang Xie

Copyright © 2015 Haoyang Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The statistical independence of time of every two adjacent bus links plays a crucial role in deciding the feasibility of using many mathematical models to analyze urban transit networks. Traditional research generally ignores the time independence that acts as the ground of their models. Assumption is usually made that time independence of every two adjacent links is sound. This is, however, actually groundless and probably causes problematic conclusions reached by corresponding models. Many transit assignment models such as multinomial probit-based models lose their effects when the time independence is not valid. In this paper, a simple method to predetermine the time independence is proposed. Based on the predetermination method, a modified capacity-restraint transit assignment method aimed at engineering practice is put forward and tested through a small contrived network and a case study in Nanjing city, China, respectively. It is found that the slope of regression equation between the mean and standard deviation of normal distribution acts as the indicator of time independence at the same time. Besides, our modified assignment method performs better than the traditional one with more reasonable results while keeping the property of simplicity well.

1. Introduction

Urban transit network is becoming a hot issue especially in developing countries like China where transit priority has risen to become a national policy. Developed countries in Europe witness this trend as well. Even in the US where car traffic dominates, as a major focus of transit network, transit accessibility is widely researched (e.g., [1–3]). The essential part of transit network analysis, transit assignment, is paid attention to concerning its framework, algorithm, equilibrium solution, and so on. For example, the solution algorithms for the multicriteria multimodal shortest path problem (M-SPP), which is known as NP-hard, in urban transit network were proposed [4]. This was followed by a label-setting algorithm for finding optimal hyperpaths in large transit networks with realistic headway distributions

[5]. Among those models aimed at transit network, many of them have something to do with statistical independence, especially stochastic transit assignment models. A conditional logit model applied to traffic assignment and modal split was established which was able to cope with the independence of irrelevant alternatives (IIA) phenomenon in a very natural way [6]. Nested logit model that is widely believed to overcome the IIA characteristic was put forward as an extension of conditional logit model [7]. Cascetta et al. [8] introduced a commonality factor for overcoming the IIA problem in route choice, which was made good use of to estimate link flow variance and the variance of the path choice proportion in a proposed stochastic assignment model [9, 10]. Independence of route choice probability has been a focus since discrete choice model was brought in the area of traffic assignment.

The issue “independence” acts as one of the prerequisites that make a great difference to the validity of many models targeted at analyzing transit network. Traditional works, however, hardly paid enough attention to time independence of every two adjacent bus links. Substantially, this independence can never be neglected. MNP model, which is short for multinomial probit model, was first used in stochastic traffic assignment [11]. The model has no requirement on independence of irrelevant route choice but requires time independence of every two adjacent bus links. The requirement was applied to various logit models as well [12]. When probit-based stochastic user equilibrium (SUE) models were researched, most of them were implicitly based on the assumption of the correctness of time independence. Those researches varied from algorithms and dynamic pricing to sensitivity analysis [13–15]. Yet none of them concerned the validity of independence assumption, which might lead to inaccurate, even useless mathematical models when it came to their applications to real life. Meanwhile, plenty of research had something to do with independence test method in the statistics field. Entropy Theory is frequently made use of to test independence and the entropy was used as a measure of dependence, like serial dependence and spatial dependence, in these studies [16–18]. In addition, many other methods were proposed to conduct independence test. Brock et al. [19] presented an independence test method on the basis of an in-depth analysis on correlation dimension. Sugiyama and Suzuki [20] introduced least squares theory to nonparametric independence test. Those methods of test are subtle and theoretic. It is, however, rather hard to put them into practice for their complexity, especially to introduce them to time independence test in large urban transit networks.

This work is targeted at engineering practice of transit systems planning and management while two simple methods are proposed. To start with, we put forward a method to predetermine time independence of every two adjacent bus links. A characteristic quantity, α , the indicator of time independence, is derived to conduct independence predetermination, laying a foundation of the usage of many models. It is found that time independence is invalid when $\alpha > 0.5$, and data collected in metropolitan area of Nanjing city, China, is used to make a preliminary test on the practicability of our method. In the case $\alpha > 0.5$, transit network models on the basis of MNP lose their effects, especially those probit-based stochastic transit assignment models, so the number of alternative assignment models decreases. Meanwhile, capacity-restraint assignment method lacks the multipath characteristic [21] though widely used in engineering practice out of its simplicity and not bad precision. Therefore, a modified capacity-restraint transit assignment method is proposed following the predetermination method while taking good advantage of some predetermination results. The method maintains its effectiveness under the circumstance of invalidity of time independence. After introducing 95% quantile of normal distribution to the traditional capacity-restraint method to realize multipath transit assignment, it turns out that the results of the modified method are more sensible than the traditional one through a small contrived transit network.

2. Methodology

2.1. Time Independence Predetermination

2.1.1. Basic Assumptions

- (1) Most bus drivers are technical in the network.
- (2) The V/C (the ratio of road traffic volume to road capacity) can be somewhat large but no congestions occur when there exists a transit exclusive lane. Otherwise, V/C is medium or low.
- (3) There is only a little interference to transit operations caused by pedestrians and bicycles.
- (4) Transit priority control is advanced, so intersections have limited effect on transit operations.

These four basic assumptions are not impractical. Assumptions (1), (2), and (3) are valid in most cities of developed countries, while assumption (3) may be a bit questionable in some developing countries like China. These countries are known for mixed traffic flow that does make a big difference to transit operations in their big cities like Shanghai. Anyway, bicycles and pedestrians have little effect on public transit on the road with separation infrastructures. The effects are also quite limited in most of their medium and small cities. The most demanding assumption is (4) probably, which is easy to be satisfied in many European countries. It is, however, very difficult to guarantee an advanced transit priority control system. Anyhow, the priority control does exist in most cities of developed countries and many big cities of developing countries. In a word, it is likely that assumption (1) is the most accessible while assumption (4) is to the contrary.

As for time independence predetermination, the “time” here is to be defined strictly. The two sides of a bus line have a bus stop. When a bus begins to pull over at the upstream stop, the time is t_1 . Similarly, t_2 is defined. We call $|t_2 - t_1|$ “stop-stop time” that is simply denoted by SST, which is composed of the parking time at the upstream bus stop and travel time between two adjacent stops. If all basic assumptions are satisfied, we could infer that SST at the bus link conforms to normal distribution according to the data in Nanjing. This can be seen from the section of distribution test below. To analyze the whole transit network basically satisfying all the assumptions above, we assume that all the links in the network conform to normal distribution, which is the foundation of the proposed methodology. More insight concerning the normal distribution can be seen in Discussion.

2.1.2. Mean and Standard Deviation Fitting. We fit the mean and std, which is short for standard deviation in this paper, of the normal distribution mentioned above. Bus links that satisfy all the basic assumptions to the greatest degree can be selected in a transit network. Accordingly, we could get several normal distributions: $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$, ..., $N(\mu_n, \sigma_n^2)$. μ and σ are estimated using the samples collected in transit survey. Note that μ and σ have the same dimension, so regression analysis is decided to be aimed at the std rather than variance. Furthermore, $\mu > 0$ and $\sigma > 0$, so instead of doing linear regression directly, we fit the μ and

σ in the Double Logarithmic Coordinate System and get the regression equation as follows:

$$\ln \sigma = \alpha \ln \mu + \theta. \quad (1)$$

The regression model is the fundamental of time independence predetermination and the modified capacity-restraint model. According to the data in Nanjing, when all chosen links meet the assumptions well, the goodness of fitting is perfect with high R -square which can be seen in Section 4. To analyze the whole transit network, we assume the regression equation has high goodness of fitting. If the regression model is not sufficiently significant, we may appropriately select other or more bus links that satisfy all basic assumptions well to avoid poor goodness of fitting. More analysis about the regression model can be seen in Section 5.

2.1.3. Indicator of Time Independence. After the regression analysis, we could predetermine time independence of every two adjacent bus links in a network. Time independence predetermination is to predetermine the independence of SST at every two adjacent links. Note that all bus links in the network conform to normal distribution and the goodness of fitting of regression equation is high on the basis of our analysis above. In this part, we put forward an indicator of independence whose value can simply predetermine time independence with practicability.

Considering normal distribution $N(x, y^2)$ with the mean x and std y , according to the previous section of fitting,

$$\ln y = \alpha \ln x + \theta \quad (x, y > 0). \quad (2)$$

α and θ are regression coefficients of the regression equation. The slope α is supposed to be greater than zero ($\alpha > 0$) because the std will increase when the mean increases. Another form of equation derived by deforming the regression equation above is

$$y = \mu x^\alpha, \quad \text{where } \ln \mu = \theta, \mu > 0. \quad (3)$$

Then the variance is given by the following expression:

$$\text{VAR} = y^2 = \mu^2 x^{2\alpha} = \lambda x^\beta \quad (4)$$

where $\lambda = \mu^2 > 0, \beta = 2\alpha$.

A transit route is given in Figure 1. Here A represents the origin and B represents the destination. The mean of a random variable D_{ab} , the difference between the moment when a bus stops at station A and that at station B, is set at fixed value T . S_k ($0 \leq k \leq n$) represents the intermediate bus stop, and t_r ($0 \leq r \leq n+1$) is the mean of SST at bus link. Obviously, T can be expressed as follows:

$$T = \sum_{k=1}^{n+1} t_k. \quad (5)$$

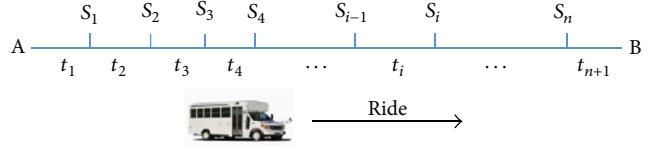


FIGURE 1: A transit route with its stops and SST.

Supposing that SST at every two adjacent links are mutually independent, then D_{ab} is normally distributed with variance σ_n^2 according to the linear superimposition principle of normal distribution. Considering there are n stops between station A and station B, we obtain

$$\sigma_n^2 = \sum_{k=1}^{n+1} \text{VAR}_k = \sum_{k=1}^{i-1} \lambda t_k^\beta + \lambda t_i^\beta + \sum_{k=i+1}^{n+1} \lambda t_k^\beta. \quad (6)$$

If there are $n+1$ stops between station A and station B, now we assume naturally the stop S_{n+1} is set up between the S_{i-1} and S_n . Thus, t_i is divided into two parts as t_j and t_{n+2} ($t_i = t_j + t_{n+2}$). Similarly, the variance is given as follows under this circumstance:

$$\sigma_{n+1}^2 = \sum_{k=1}^{n+2} \text{VAR}_k = \sum_{k=1}^{i-1} \lambda t_k^\beta + \lambda t_j^\beta + \lambda t_{n+2}^\beta + \sum_{k=i+1}^{n+1} \lambda t_k^\beta. \quad (7)$$

By subtracting σ_n^2 from σ_{n+1}^2 , the result is

$$\begin{aligned} \sigma_{n+1}^2 - \sigma_n^2 &= \lambda (t_j^\beta + t_{n+2}^\beta - t_i^\beta) \\ &= \lambda \left(t_j^\beta + t_{n+2}^\beta - (t_j + t_{n+2})^\beta \right), \quad \lambda > 0. \end{aligned} \quad (8)$$

Function $f(t_j, t_{n+2})$ is constructed where $t_j, t_{n+2} > 0$. Its partial deviation is obtained:

$$\begin{aligned} f(t_j, t_{n+2}) &= t_j^\beta + t_{n+2}^\beta - (t_j + t_{n+2})^\beta \\ \frac{\partial f}{\partial t_j} &= \beta \left(t_j^{\beta-1} - (t_j + t_{n+2})^{\beta-1} \right) \end{aligned} \quad (9)$$

$$\frac{\partial f}{\partial t_{n+2}} = \beta \left(t_{n+2}^{\beta-1} - (t_j + t_{n+2})^{\beta-1} \right).$$

The analytical results could be obtained. When $\beta > 1$ or $\beta < 0$ and $f(t_j, t_{n+2})$ is a decreasing function, then $f(t_j, t_{n+2}) < f(0, 0) = 0$, which is equivalent to $\sigma_{n+1}^2 < \sigma_n^2$; when $\beta = 1$, $f(t_j, t_{n+2}) = 0$, and $\sigma_{n+1}^2 = \sigma_n^2$; when $\beta = 0$, $f(t_j, t_{n+2}) = 1$, and $\sigma_{n+1}^2 > \sigma_n^2$; and when $0 < \beta < 1$ and $f(t_j, t_{n+2})$ is an increasing function, then $f(t_j, t_{n+2}) > f(0, 0) = 0$, which is equivalent to $\sigma_{n+1}^2 > \sigma_n^2$.

As is analyzed before, $\beta = 2\alpha$, and D_{ab} conforms to normal distribution. Note that $\alpha > 0$; the results are given as follows:

- (1) When $\alpha > 0.5$, the variance of D_{ab} decreases with the increasing of the number of intermediate bus stops.
- (2) When $\alpha = 0.5$, the variance of D_{ab} is constant with the changing of the number of intermediate bus stops.
- (3) When $0 < \alpha < 0.5$, the variance of D_{ab} increases with the increasing of the number of intermediate bus stops.

Under the premise of a fixed mean of D_{ab} , the variance should increase with the number of intermediate stops increasing. This is because the increasing of the number of bus stops results in starting and stopping more frequently, causing the uncertainty of travel time to become greater. The variance is a strong indicator of uncertainty. To sum up, all the analysis leads to a conclusion of predetermining time independence.

When $\alpha \geq 0.5$, if the independence of SST holds, a clear contradiction exists. The SST at every two adjacent bus links, therefore, is not mutually independent. Time independence is not valid in this case.

When $0 < \alpha < 0.5$, the SST at every two adjacent links is probably mutually independent, where a further demonstration is needed. In this case, we could not decide whether or not time independence is valid. This is what “predetermination” derives from.

Therefore, the parameter α is not only a regression coefficient of the equation expressed as $\ln y = \alpha \ln x + \theta$, but also an indicator of time independence. On the one hand, the larger α is, the higher time uncertainty is in urban transit network. Considering uncertainty avoidance theory, we hope α is not too large. On the other hand, the relative magnitude of the indicator and 0.5 predetermines whether SST in transit network is independent or not. This decides whether some stochastic models are valid to a large extent. Taking MNP-based assignment model as an example, it should not be applied if $\alpha \geq 0.5$. Anyway, we cannot determine the validity of this model under the condition $0 < \alpha < 0.5$.

2.2. A Modified Capacity-Restraint Transit Assignment Method. If time independence is predetermined to be not valid, many transit assignment methods such as MNP-based models should never be used. Under our basic assumptions, methods based on logit models are untenable due to the normal distribution for SST. As a result, many assignment methods prove ineffective. Traditional capacity-restraint transit assignment method is not affected, though it could not take time uncertainty into consideration. So this part proposes a modified capacity-restraint transit assignment method on the basis of normal distribution for SST and regression equation between the distribution mean and std. When the invalidity of time independence has been predetermined, unlike many assignment methods, the proposed one keeps its effectiveness. It remains useful even though time independence is proved to be solid.

Compared with traditional capacity-restraint method, the modified one considers time uncertainty by using 95% quantile of normal distribution while keeping the advantages of traditional method like the simplicity of calculation and so on. It is believed that traditional heuristic assignment methods are faced with the “common lines” problem, and many modifications have been made to overcome the obstacle [22]. These modifications can be used in our method as well if the “common lines” problem is supposed to be considered in a specific case. It is not difficult to generalize the proposed method using the modifications but requires a little future work. So, in this paper, the “common lines” problem is not much concerned including the contrived network we use below. Flow chart of the modified capacity-restraint transit assignment method is shown in Figure 2. When given a traditional transit network, we could regard the time at bus links as mean value of SST. Then, the previous regression equation is made use of to work out the std at all links. In this way, the normal distribution of every bus link is determined. As Figure 3 shows where A, B, C, D, and E represent bus stops and $N(\mu, \sigma^2)$ stands for normal distribution, the modified method is based on transit network with SST treated as random variable rather than deterministic variable, which is closer to reality.

The rationale of the modified method is similar to the traditional one, though the coefficient of assignment γ ($0 < \gamma < 1$) is introduced to the model. We assign 100γ percent of the OD, which is short for origin-destination, to the shortest path on the basis of the mean value of SST to obtain the passenger volume X_a . And $100(1 - \gamma)$ percent is assigned to the shortest path after each SST values their 95% quantile on the shortest route to obtain Y_a . The 95% quantile is chosen because it has a good sense in probability studies. With this method, the total travel time of shortest path with mean SST could be maximized by using 95% quantile of the SST of each bus link composing the route, which is aimed at achieving multishortest paths. It is suggested that γ should be specified as 0.75 [23]. So the final passenger volume Z_a could be calculated using the formula

$$Z_a = 0.75X_a + 0.25Y_a. \quad (10)$$

Note that the generalized cost of bus travel changes after loading one part of OD into transit network, but the std of SST is to remain the same unless some changes about the mean of SST take place. Then the remaining parts of the modified method are the same as that of the traditional one. To make this revised method more useful in engineering practice, efficiency should be paid special attention to. As a result, we recommend dividing the whole OD into three shares by 50%, 30%, and 20%, respectively.

3. Data Preparation

3.1. Bus Line Selection. Data has been collected through transit survey in metropolitan area of Nanjing, China. Considering all the basic assumptions of our methods and the traffic

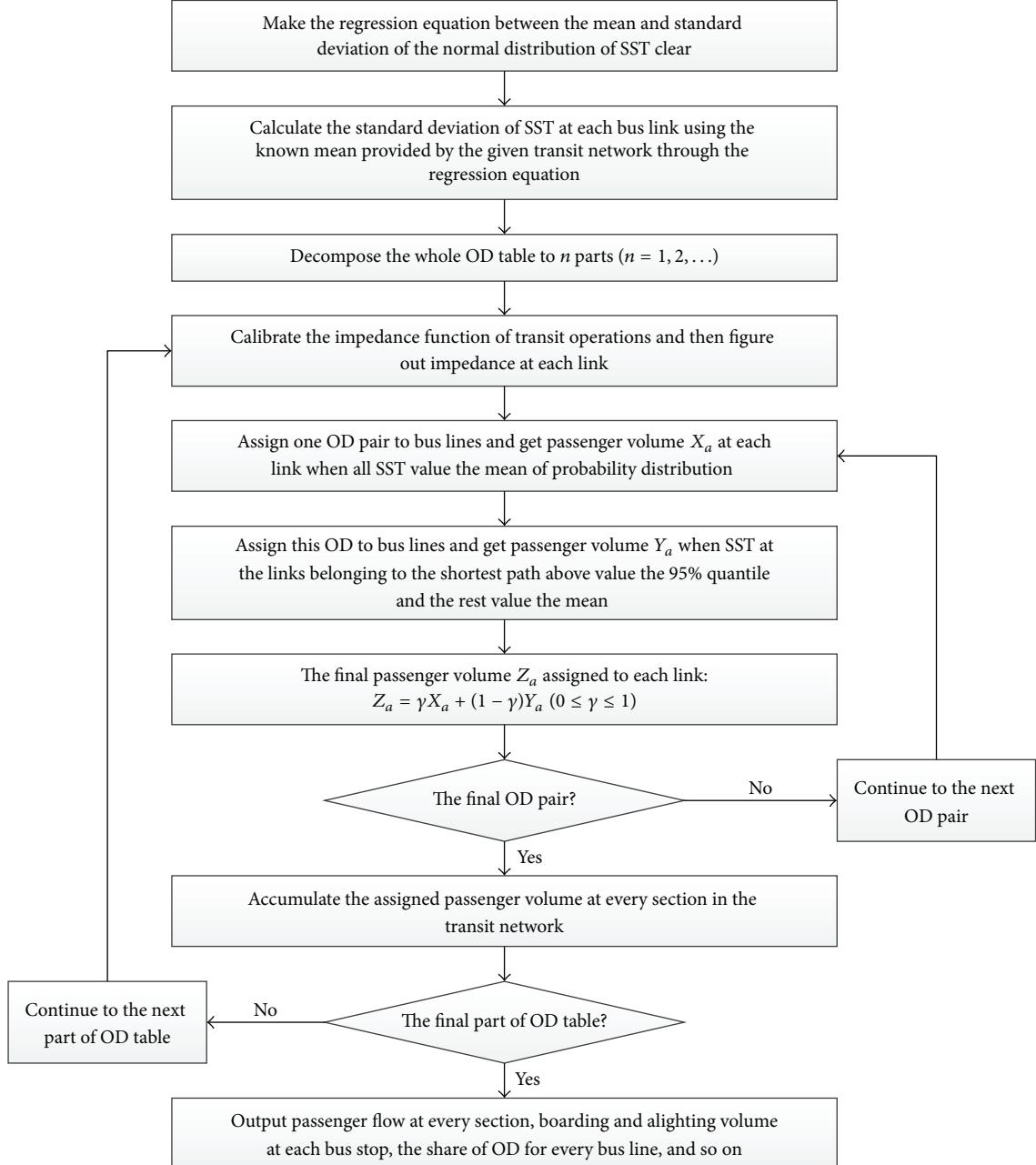


FIGURE 2: Flow chart of the modified capacity-restraint assignment method.

environment in real world, the bus lines are selected according to the following points:

- (a) All the bus lines pass the metropolitan area of Nanjing, which are widely distributed in the city center.
- (b) Bicycles and pedestrians make a little difference to bus operations, especially where the side median exists.
- (c) Buses have priority on signalized intersections.
- (d) Survey had better be conducted on a sunny day to avoid small probability events.

Taking all these factors into consideration, 7 groups of adjacent bus stops were randomly picked out, and corresponding bus lines were selected in the metropolitan area of Nanjing. Note that there may exist several choices of bus lines when a pair of bus stops is determined. So it is rationally assumed that there exists no difference between these lines concerning their operations between the pair of bus stops. To avoid waiting for a long time at bus stops, generally bus lines with too low frequency were never selected in this survey.

3.2. The Survey. The transit survey was conducted from 4:00 p.m. to 7:00 p.m. on weekdays to collect data at peak hours.

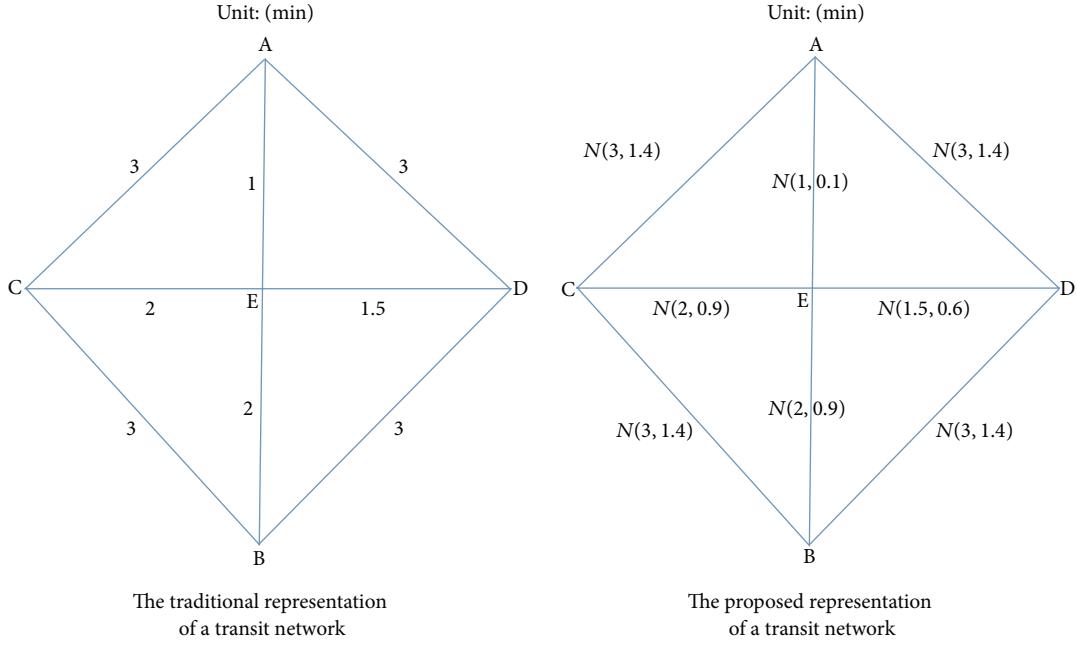


FIGURE 3: Comparison of the proposed and traditional abstraction of networks.

The investigators must record the time when their targeted buses stop for the first time at the bus stop. Then they should go on their buses and record the time when the buses pull over at the next stop. In theory, investigators should repeat the process for 16 times in single direction. In this way, however, it is hard for the survey to get finished in 3 hours. So investigators were permitted to repeat the process for 8 times in dual directions if the bus frequency was rather low. The error is typically rather limited if the bus stop on the other side of road is very close to the stop on this side in terms of the distance along the road.

3.3. Preliminary Data Analysis. From the recorded time when buses pull over at stops, SST samples are figured out and displayed in Table 1. The quality of our data is evaluated by the seven frequency histograms in Figure 4. According to these figures, it could be found qualitatively that the SST at bus links might conform to normal distribution, which is consistent with our conjecture in the former section. In addition, the correlation coefficient between the logarithm of sample mean and std is 0.996. This means the logarithm of the two basic statistics is highly correlated indicating a significant linear relationship. Therefore, the data acquired from this survey is preliminarily considered reliable. It can be made use of in the following analysis of the proposed methods.

4. Results

4.1. Transit Network in Metropolitan Area of Nanjing

4.1.1. Distribution Test of SST. Although it has been assumed that SST at bus links conforms to normal distribution, distribution test could be made to give a preliminary proof. We

TABLE 1: Result of the SST survey in metropolitan area of Nanjing.

Number	201	100	792	405	126	160	16
1	159	32	42	61	55	35	49
2	127	37	49	52	56	36	46
3	166	32	42	59	54	37	51
4	158	28	46	56	65	43	50
5	122	31	41	67	67	41	58
6	134	33	51	60	56	35	60
7	138	35	41	62	59	35	47
8	148	33	45	53	51	36	52
9	166	29	42	58	52	39	51
10	140	31	47	70	52	37	53
11	155	35	42	71	62	35	51
12	138	35	40	67	55	40	57
13	151	30	45	58	48	36	52
14	144	33	49	64	55	37	62
15	138	36	45	57	54	35	48
16	132	33	48	57	52	38	51
Mean	144.8	32.7	44.7	60.8	55.8	37.2	52.4
std	13.4	2.5	3.4	5.7	5.1	2.4	4.6

conduct Jarque-Bera test on the basis of the data in Nanjing using MATLAB. According to the results displayed in Table 2, h values 0 under the level of significance of 0.05. Besides, it is seen that the scatters are basically uniformly distributed around the line that indicates a normal distribution from one of the test graphs (Figure 5). In summary, this assumption of normal distribution is tenable.

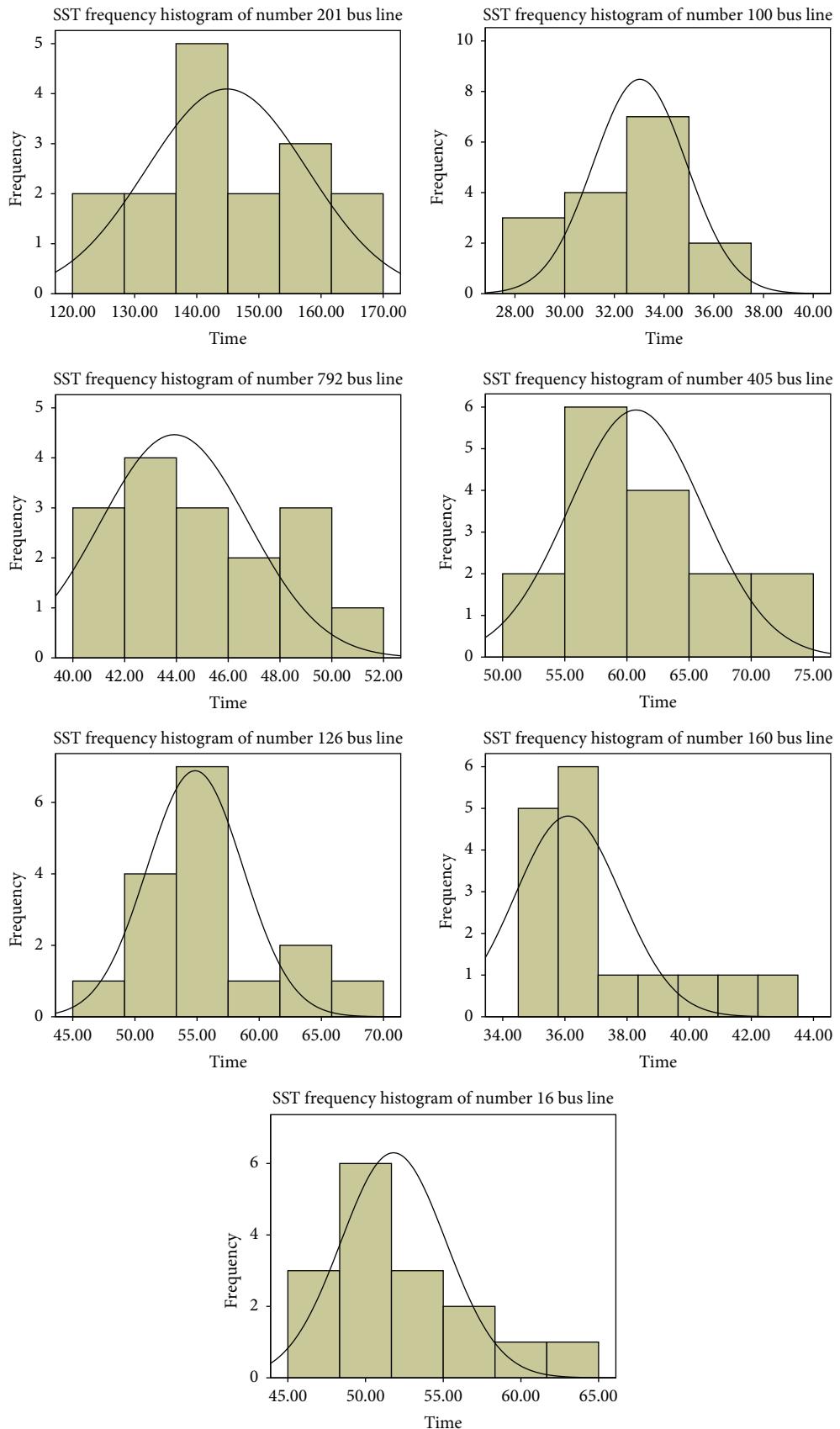


FIGURE 4: Frequency histogram of SST of each bus line.

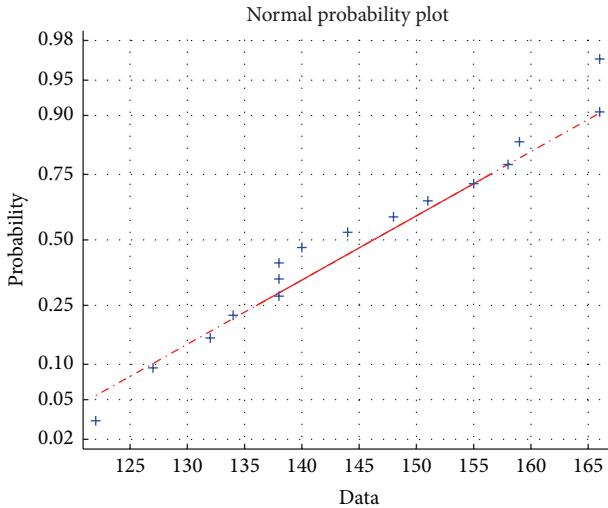


FIGURE 5: Fitting curve of the SST of number 201 bus line.

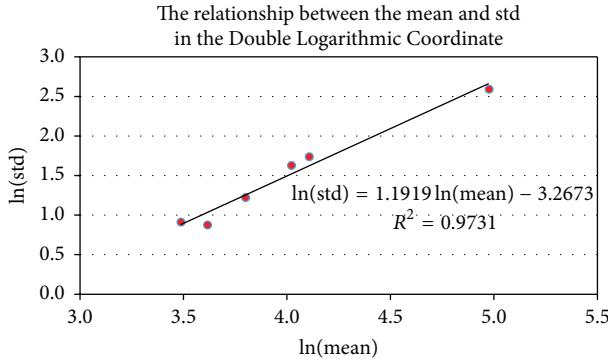


FIGURE 6: Regression analysis between the mean and std of normal distributions.

4.1.2. Regression Analysis between the Mean and Standard Deviation. Data from 6 lines is used for fitting the relationship between the mean and std, and the rest is for examining the fitting result simply. According to Section 2.1.2, the regression analysis is conducted after taking logarithm with the help of MATLAB followed by an evaluation of its results. The coefficients of regression and some statistics are listed in Table 3 while the regression equation is shown in Figure 6.

Judging from the results, the coefficient of determination is close to 1 and the probability of significance values 0.0003 that is smaller than 0.05 under the level of significance of 0.05. Furthermore, the R -square is 0.9731 indicating that the regression relationship appears reliable. To determine the validity of the equation, the rest of the bus lines are used for checking through comparing the calculation from the regression equation with the observed data. For the std of SST of line number 16, the result by the regression equation is 4.2665 while data shows the value 4.6025. The relative error is 7.3%, which is small enough to prove the validity of the regression analysis. To sum up, the regression equation is significant. In addition, it has been assumed that the equation

TABLE 2: Result of the distribution test of SST.

Bus lines	h	CV	Mean	Variance
201	0	3.4140	144.75	180.47
100	0	3.4140	32.69	6.36
792	0	3.4140	44.69	11.56
405	0	3.4140	60.75	32.47
126	0	3.4140	55.81	26.30
160	0	3.4140	37.19	5.90
16	0	3.4140	52.38	21.18

Note: h = result of the test; if $h = 0$, X conforms to normal distribution; if $h = 1$, X does not conform to normal distribution; CV = the critical value whether to reject the null hypothesis or not.

TABLE 3: Result of the regression analysis.

b	1.1919	-3.2673
bint	[0.9169, 1.4669]	[-4.3756, -2.1590]
stats	0.9731	144.8071

Note: b = the least-square estimated value of the two regression coefficients; bint = the 95% confidence interval of the regression coefficients; stats = the coefficient of determination, F statistics of variance, the probability of significance of variance, and the estimate of variance, respectively.

between the mean and std of normal distribution of SST has high goodness of fitting. The results of the regression analysis can well support this assumption.

4.1.3. Time Independence of Every Two Adjacent Bus Links. On the basis of our regression equation (11), the indicator of independence α is 1.1919 and another parameter θ values -3.2673 where the unit of time is the second:

$$\ln(\text{std}) = 1.1919 \ln(\text{mean}) - 3.2673. \quad (11)$$

α is larger than 0, which suits our common sense. More importantly, α is larger than 0.5. First, it means that the independence of SST at every two adjacent bus links is invalid in metropolitan area of Nanjing. Consequently, models like MNP-based assignment can hardly be used to estimate passenger flow of transit system in the city center. It is worth mentioning that most logit-based models are groundless to be adopted in this city area because SST conforms to normal distribution, let alone the invalidity of time independence. It is these widely used models that lose their effects. Second, the time uncertainty of transit operations is never low because α is significantly larger than zero. These may account for the reasons why many models adopted by traffic engineers appeared ineffective on the transportation network in Nanjing, which indicates the significance of the proposed indicator to a great extent. More analysis could be made through comparison with that of other cities.

4.2. A Contrived Transit Network. A contrived transit network is established to compare our proposed assignment method with the traditional one. Impact from other traffic modes on transit operations is not considered in this example. There are three transit lines in Figure 7. The mean of SST

TABLE 4: Cost update of transit routes.

OD pair	Transit route	Our cost during loading 50%/30%/20% of OD (95% quantile SST if necessary)
Origin 1 to destination 3	1 → 2 → 3	4.9 (5.52)/8.1 (9.216)/10.02 (11.4336)
	1 → 5 → 6 → 3	28.5/32.7/34.5
	1 → 9 → 10 → 6 → 3	11.7/13.1/14.66
Origin 1 to destination 6	1 → 3 → 6	12.9/17.7/20.58
	1 → 5 → 6	8.7 (9.63)/10.8/ 11.7 (13.095)
	1 → 9 → 10 → 6	8.9/10.3 (11.757)/11.86
Origin 1 to destination 12	1 → 2 → 3 → 4 → 12	10.9 (11.55)/17.1 (19.611)/20.82
	1 → 9 → 10 → 6 → 3 → 4 → 12	27.9/34.5/39.54
	1 → 5 → 6 → 7 → 12	17.7/19.8/20.7 (23.49)
	1 → 9 → 10 → 6 → 7 → 12	23.7/25.8/28.14

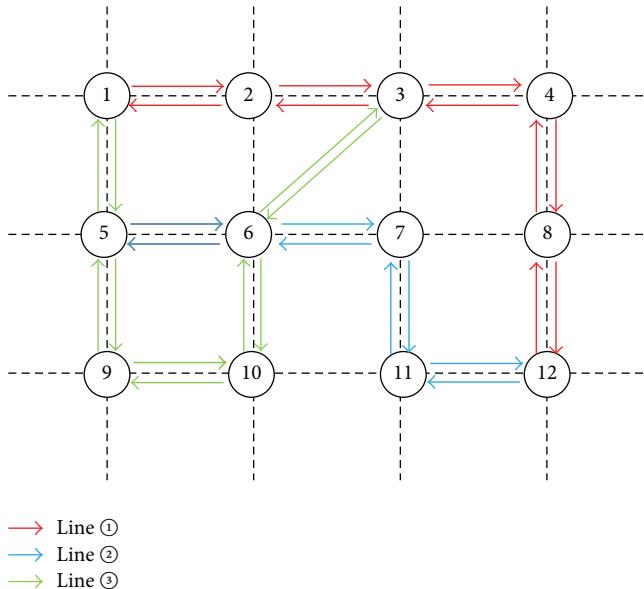


FIGURE 7: A contrived transit network.

between two adjacent stops is 2 min while it is 2.8 min between stop 3 and stop 6. Their fees are all 0.9 units. The frequencies and vehicle capacities of these bus lines are deterministic, and it is worked out that the rated capacity and maximum capacity of all the three lines are 500 passengers per hour and 800 passengers per hour, respectively [24]. OD from stops 1 to 3, 1 to 6, and 1 to 12 is 300, 400, and 500 passengers per hour, respectively.

Generalized cost function is commonly used in transportation studies [25]. Referring to the document mentioned above, the shortest path defined in this example is on the basis of generalized cost while passenger congestion constitutes one part of the cost. Meanwhile, we consult the form of generalized cost function used in [26]. Combining what

is referred altogether, the cost of a route is determined as follows:

$$\begin{aligned} \text{GC} &= (\eta \sum T \cdot k + F) \cdot q \\ V &\leq MC, \end{aligned} \quad (12)$$

where GC is generalized cost; T is mean SST; F indicates fees; η is value of travel time; k indicates discomfort coefficient; q is transfer penalty; V is passenger volume; MC is maximum capacity.

As is recommended in [26], it is decided that η is 1 unit/min and k is defined to be $V/C+1$, where V/C is the ratio of passenger volume to the rated capacity. And if i denotes transfer times, q equals $1.5i$, while it values 1 unless a transfer occurs. Traditional capacity-restraint method in this case is under the condition that OD is divided into 3 parts: 50%, 30%, and 20%, which is the same as that of our modified method. And the coefficient of assignment γ values 75% as recommended. If some routes have the same cost, passenger flow will be assigned to the route with the fewest transfers while they share the equivalent flow if their transfer times are the same. All routes that are feasible to be assigned are listed in Table 4 which provides the cost update during the assignment. These exclude the routes with too much cost (e.g., route 1 → 3 → 6 → 7 → 12 with the cost of 46.5 units) and reverse routes (e.g., route 1 → 5 → 9 → 5 → 6). Table 5 presents how our assignment method performs in each OD share, and results of the assigned passenger flow are shown in Figure 8.

5. Discussions

As for the independence predetermination, it is hoped that the indicator of time independence α is small when an urban transit network has low time uncertainty. Meanwhile, chances are that time independence is valid when α is small; to be specific, $0 < \alpha < 0.5$. Therefore, the most ideal situation is $0 < \alpha < 0.5$ when the transit network has low time

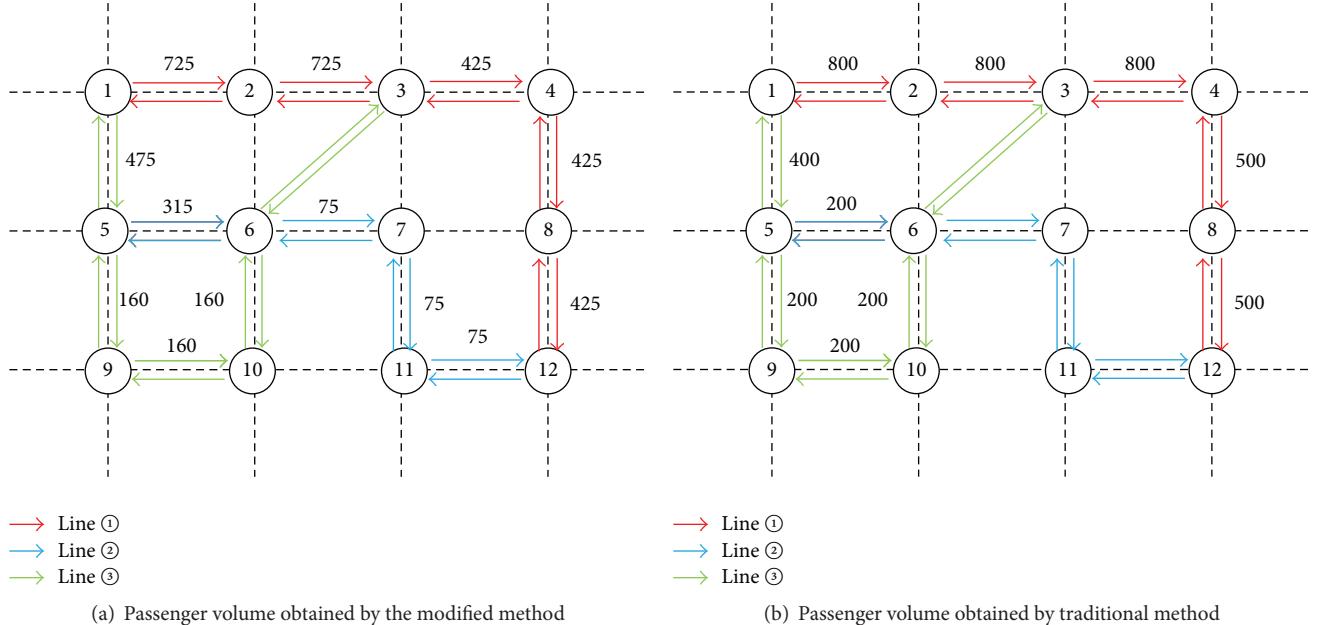


FIGURE 8: Results of transit assignment.

TABLE 5: Process of passenger flow assignment.

Transit route	50% OD share	30% OD share	20% OD share
1 → 2 → 3	400	240	85
1 → 5 → 6	150	30	60
1 → 9 → 10 → 6	50	90	20
1 → 2 → 3 → 4 → 12	250	150	25
1 → 5 → 6 → 7 → 12	0	0	75

Note: routes in Table 4 that are assigned zero flow during the process are not listed in this table.

uncertainty and network models requiring time independence, like MNP-based assignment models, can probably be used. With regard to the modified capacity-restraint method, it has proved to be more accurate than the traditional one. Heuristic assignment algorithms like capacity-restraint method are more common in engineering practice, so the modified method with the same time complexity as that of the traditional method is highly prospective. Comparing our assignment methods with equilibrium ones, it is valueless to take many SUE models into account when time independence is not valid. But the assertion is hard to be made that the modified heuristic model will be more accurate than SUE models when $0 < \alpha < 0.5$. Anyhow, the two methods we put forward have an advantage over others on simplicity. They can be adopted in engineering practice in any urban transit network with basic assumptions satisfied by and large, which is not a tough requirement.

The results of metropolitan area of Nanjing indicate the key point before modeling urban transit network. It is the

basic assumptions that researchers should lay great emphasis on. Instead of asserting that all model assumptions could be satisfied with qualitative analysis, some quantitative analysis is also indispensable. The proposed indicator of time independence along with some statistics of probabilistic tests exactly plays the role. Some uncomplicated transit survey mentioned in this paper could be conducted to figure out these key indicators before selecting reasonable models.

Some basic assumptions could be strengthened with the development of urban transportation. For assumption (2), big cities in developing countries are promoting the construction of transit special lane network. This strategy is aimed at ensuring “Transit Priority” by providing successive transit right of way. For assumption (3), separation of different types of traffic flow is being advocated all over the world considering efficiency and safety issue. Transit operations are becoming increasingly steady. For assumption (4), with the devotion to transit priority at intersections, actuated signal control systems have a brilliant future. Anyway, the basic assumptions are aimed at guaranteeing a normal distribution of SST to the maximum extent. As an adequate alternative of assumption (4), bus delays at intersections must be very small.

Although it is assumed that SST conforms to normal distribution when bus links satisfy all the four basic assumptions, we could conduct distribution test to confirm the normal distribution assumption just like the case study in Nanjing, China. Only in a few cases may such SST at bus links not conform to normal distribution though there is no exception in Nanjing. When the assumption of normal distribution does not pass the test, the modified assignment method will lose its effect, and the method to predetermine time independence cannot be used in the whole urban transit network. However, we can focus on the subnetworks that SST

at all bus links are tested to conform to normal distribution, so the predetermination method is able to be adopted in these subnetworks. In this way, the results of predetermination could also act as part of evidence to decide the validity of some models.

Another assumption is made in which the regression equation between the mean and std of normal distribution in the Double Logarithmic Coordinate is significant. Nevertheless, this cannot be guaranteed even though all bus links satisfy those basic assumptions. To obtain high goodness of fitting, samples could probably be optimized by selecting locations. This measure is quite flexible, which can make successive adjustments in the whole network, and there is a high chance that a significant regression equation exists characterizing a transit network that satisfies the basic assumptions. Even if poor goodness of fitting remains unsettled, which is not likely to happen, our proposed assignment method will work all the same if all std of SST at bus links could be worked out through transit survey. Then, it is evident the transit network cannot be too large. But in this case, anyhow, our method to predetermine time independence will become invalid.

6. Conclusions

A simple method has been proposed to predetermine time independence of every two adjacent bus links, which can pre-judge the validity of many stochastic network models for public transit. Then a modified capacity-restraint transit assignment method is put forward and aimed at engineering practice when such independence is predetermined to be invalid.

Some basic assumptions have been made to elicit two direct assumptions to the proposed methods: First, SST at bus links conform to normal distribution. Second, the regression equation between the mean and standard deviation has high goodness of fitting. These two assumptions are tenable according to the results in Nanjing. Then the indicator of time independence α is proposed to show the predetermination result. When $\alpha \geq 0.5$, SST at every two adjacent bus links interact with each other. When $0 < \alpha < 0.5$, time independence has potential validity that still needs proving. Meanwhile, α acts as the slope of the regression equation, which can reflect the time uncertainty in urban transit network. The modified assignment method is put forward, which appears effective when predetermination indicates the validity of time independence. We introduce the coefficient of assignment γ and 95% quantile of probabilistic distribution on the basis of the traditional capacity-restraint method. The traditional transit network with fixed SST at bus links is transformed into the network with normal distributed SST, which supports the modified method.

We take good advantage of the data acquired in Nanjing to conduct a case study. From the result of distribution test, all the seven bus links conform to normal distribution with a good significance level of 0.05. Then regression analysis is made in the Double Logarithmic Coordinate with a large R^2 -square of 0.9731 reflecting its significance. The indicator of time independence or the regression coefficient is 1.1919 that is larger than 0.5. As a consequence, time independence is

not valid in Nanjing, so many network models could not be used in the city. Researchers have to turn to other methods. Finally, the proposed capacity-restraint assignment method is preliminarily tested by a simple network. Compared with traditional capacity-restraint method, the modified method inherits the property of simplicity and has a better performance with more sensible results.

Although the findings in this paper are some of engineering practices, this work is limited by some issues. Firstly, it is very difficult to completely satisfy basic assumption (4) nowadays. Even if we turn to searching for signalized intersections with much small bus delays, it may have some influence on the normal distribution assumption of SST at bus links. Besides, the sample size in this study is somewhat limited for the huge workload of data acquisition though the proposed methodology is of theoretical basis. There are great numbers of two adjacent links in an urban transit network, which is not that suitable to acquire SST data manually.

Further research could focus on how to loosen the proposed assumptions. Even if normal distribution is invalid, others like Negative Binomial Distribution and others may replace it. Different distributions can be introduced to a network simultaneously, and accordingly the methods of predetermination and assignment might exist as well, which is probably complicated to be put into practice yet highly prospective. In addition, existing approaches could be introduced to have our assignment model well consider the “common lines” problem as with the development of previous heuristic capacity-restraint models. Related details are worth exploring. Moreover, the validity of the modified assignment method could be verified using real data of transit passenger flow before getting adopted by traffic engineers. Last but not least, work could be continued in the case $0 < \alpha < 0.5$, and it is expected that further conclusions concerning the time independence will be drawn in future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is funded by the National Key Basic Research Program of China (no. 2012CB725402), the Fundamental Research Funds for the Central Universities (no. KYLX_0171), and the Scientific Research Foundation of the Graduate School of the Southeast University.

References

- [1] S. Chandra, M. E. Bari, P. C. Devarasetty, and S. Vadali, “Accessibility evaluations of feeder transit services,” *Transportation Research Part A: Policy and Practice*, vol. 52, pp. 47–63, 2013.
- [2] K. Sivakumaran, Y. Li, M. Cassidy, and S. Madanat, “Access and the choice of transit technology,” *Transportation Research Part A: Policy and Practice*, vol. 59, pp. 204–221, 2014.

- [3] S. Farber, M. Z. Morang, and M. J. Widener, "Temporal variability in transit-based accessibility to supermarkets," *Applied Geography*, vol. 53, pp. 149–159, 2014.
- [4] L. Z. Liu, J. H. Yang, H. B. Mu, X. J. Li, and F. Wu, "Exact algorithms for multi-criteria multi-modal shortest path with transfer delaying and arriving time-window in urban transit network," *Applied Mathematical Modelling*, vol. 38, no. 9-10, pp. 2613–2629, 2014.
- [5] Q. F. Li, P. Chen, and Y. Nie, "Finding optimal hyperpaths in large transit networks with realistic headway distributions," *European Journal of Operational Research*, vol. 240, no. 1, pp. 98–108, 2015.
- [6] D. Heidemann, "Conditional logit model and the independence of irrelevant alternatives phenomenon," in *Proceedings of the Conference on Traffic and Transportation Studies (ICTTS '98)*, pp. 255–264, July 1998.
- [7] H. C. Williams, "On the formation of travel demand models and economic evaluation measures of user benefit," *Environment and Planning A*, vol. 9, no. 3, pp. 285–344, 1977.
- [8] E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta, "A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks," in *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pp. 697–711, Lyon, France, July 1996.
- [9] K. S. Chan and W. H. K. Lam, "A stochastic traffic assignment model for estimating the variance of the path choice proportion," in *Proceedings of the 1st Conference of Hong Kong Society for Transportation Studies*, pp. 31–44, 1996.
- [10] K. S. Chan and W. H. K. Lam, "Estimation of link flow variance by stochastic traffic assignment," in *Proceedings of the Transportation Systems Conference*, pp. 1241–1246, Chania, Greece, June 1997.
- [11] C. F. Daganzo and Y. Sheffi, "On stochastic models of traffic assignment," *Transportation Science*, vol. 11, no. 3, pp. 253–274, 1977.
- [12] E. T. Kenneth, *Discrete Choice Methods with Simulation*, University of California, Berkeley and National Economic Research Associates, 2003.
- [13] Q. Meng and Z. Liu, "Mathematical models and computational algorithms for probit-based asymmetric stochastic user equilibrium problem with elastic demand," *Transportmetrica*, vol. 8, no. 4, pp. 261–290, 2012.
- [14] K. Zhang, H. S. Mahmassani, and C.-C. Lu, "Dynamic pricing, heterogeneous users and perception error: probit-based bicriterion dynamic stochastic user equilibrium assignment," *Transportation Research Part C: Emerging Technologies*, vol. 27, pp. 189–204, 2013.
- [15] S. D. Clark and D. P. Watling, "Sensitivity analysis of the probit-based stochastic user equilibrium assignment model," *Transportation Research Part B: Methodological*, vol. 36, no. 7, pp. 617–635, 2002.
- [16] A. Dionísio, R. Menezes, and D. A. Mendes, "Entropy-based independence test," *Nonlinear Dynamics*, vol. 44, no. 1–4, pp. 351–357, 2006.
- [17] M. Matilla-García and M. R. Marín, "A non-parametric independence test using permutation entropy," *Journal of Econometrics*, vol. 144, no. 1, pp. 139–155, 2008.
- [18] F. López, M. Matilla-García, J. Mur, and M. R. Marín, "A non-parametric spatial independence test using symbolic entropy," *Regional Science and Urban Economics*, vol. 40, no. 2–3, pp. 106–115, 2010.
- [19] W. A. Broock, J. A. Scheinkman, W. D. Dechert, and B. LeBaron, "A test for independence based on the correlation dimension," *Econometric Reviews*, vol. 15, no. 3, pp. 197–235, 1996.
- [20] M. Sugiyama and T. Suzuki, "Least-squares independence test," *IEICE Transactions on Information and Systems*, vol. E94-D, no. 6, pp. 1333–1336, 2011.
- [21] W. Wang and X. W. Chen, *Traffic Planning*, Southeast University, Nanjing, China, 2007.
- [22] Y. Liu, J. Bunker, and L. Ferreira, "Transit user's route-choice modelling in transit assignment: a review," *Transport Reviews*, vol. 30, no. 6, pp. 753–769, 2010.
- [23] H. J. Huang, *Urban Transportation Network Equilibrium Analysis: Theory and Practice*, Beihang University, Beijing, China, 1994.
- [24] W. Wang and X. W. Chen, *Urban Transit Systems Planning and Management Techniques*, Southeast University, Nanjing, China, 2002.
- [25] L. Zhang and J. F. Jia, "Study on generalized cost of passenger trip and product split of passenger transport," *Journal of Beijing Jiaotong University*, vol. 35, no. 3, pp. 68–71, 2011.
- [26] L. Q. Yang, W. Huang, and N. Zhang, "Transfer preferential benefit of public transport based on generalized cost," *Journal of Transport Information and Safety*, vol. 27, no. 3, pp. 20–23, 2009.

