Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5120401, 14 pages
https://doi.org/10.1155/2022/5120401

Research Article

Determining the Level of Service Scale of Public Transport System considering the Distribution of Service Quality

Min Wang[1], **Baohua Mao**[1], **Yanqiang Yang**[1], **Ruijia Shi**[2] **and Junsheng Huang**[1]

[1]*Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China*
[2]*Beijing Municipal Institute of City Planning & Design, Beijing 100045, China*

Correspondence should be addressed to Baohua Mao; bhmao@bjtu.edu.cn

Received 6 June 2021; Revised 10 September 2021; Accepted 28 September 2021; Published 2 March 2022

Academic Editor: Ren-Yong Guo

In China, many cities are building themselves the transit metropolis, and the reasonable evaluation of level of service (LOS) of public transport system (PTS) is one important aspect. However, to determine the overall LOS is hard because the distribution of service in PTS is not homogeneous with regard to time and space. To address this problem, this study proposes a general framework to determine the LOS scale of PTS based on the distribution of service quality. Under the framework, two classification methods are discussed. Method 1 uses two parameters, the mean and coefficient of variation to model the distribution, and Method 2 is an existing approach that only considers mean. Then the specific use of the framework is expounded for the service attribute of crowding, and Beijing subway line LOS is evaluated. The line LOS is divided into I–IV, whose threshold is expressed as a function of mean and coefficient of variation. The results show that 57.8% of the sample points are in the most crowding level IV in morning peak hours by Method 1, but 60.9% of sample points are in a comfortable level II by Method 2, and the former is more consistent with reality. In addition, it reveals which lines and time periods need to improve the service level. The research proves the feasibility of considering the service distribution to determine the overall LOS of PTS, and it is useful for capturing more detailed information of the system performance in time and space. This research can provide an approach for evaluating and helping to improve the overall service level of PTS for public transport authorities.

1. Introduction

With the development of social economy and residents' living standards, the public has an increasingly high demand for public transport travel services, and operators need to provide higher quality operation services to meet the needs of passengers. A reasonable evaluation of service quality is a prerequisite for targeted and efficient improvement of service quality. Among a variety of evaluation ways for public transit service quality, the method based on the concept of level of service (LOS) is widely used for its effectiveness in practice.

The LOS is usually assessed for some equipment in the service system or for the overall service system, such as subway platforms [1], station corridors [2], airport terminals [3], BRT system [4], and so on. In this paper, we focus on the public transport system (PTS). The service level of PTS can be measured by quantitative or qualitative indicators of each service attribute, and it is usually classified into several discrete levels (e.g. “A” to “F”) [5].

The overall indicators used to measure the system service level are often represented by mean or weighted results over a certain time and space. However, the variation features of the indicators with time and space are often ignored, which may affect the evaluation results. Taking a subway line with 3 sections as an example, when the train load factors are (60%, 60%, 60%) or (90%, 60%, 30%), respectively, the mean value of the two cases is equal, but evidently the overall service quality of the two cases is not the same. Therefore, it is necessary to consider the deviation of distribution in public transport services to appraise a convincing service level.

It is a tough work to evaluate the overall LOS of PTS, if spatiotemporal variation characteristics of service quality are taken into account. PTS usually operates for a long time in a day and covers a city, so the determination should be based on an integrated result of all service equipment performances of all time and space. This means that massive data need to be collected and handled (data are generated by every equipment at any time). In addition, the operators need to find reasonable indicators to reflect the overall distribution of service levels, rather than only a mean indicator.

To address this problem, this paper discusses the issue of determining PTS LOS considering the distribution of service quality and proposes a method framework as a solution. In the method framework, the coefficient of variation is introduced to reflect the distribution deviation of service measures, and PTS LOS is determined by using it together with the mean. To present the use of the method, the crowding attribute of urban rail transit system is taken as an example.

The main work of this paper is as follows:

(1) An implementable framework for determining PTS LOS is proposed, considering the service quality distribution model as a base.

(2) Rather than only using the mean indicator, the coefficient of variation is used simultaneously with the mean indicator to determine the threshold of service level.

(3) The method proposed is compared with an existing method. It is found that the results of this method are closer to reality, which can be used for service improvement guidance.

The remainder of this paper is organized as follows. Section 2 is the literature review. Section 3 proposes the method framework of determining PTS LOS scale. Section 4 and Section 5 implement the method in Beijing subway for the attribute of train crowding and analyze the determination results and the validity. Finally, conclusions and discussion are included in Section 6.

2. Literature Review

The generally accepted concept of LOS was put forward in *Highway Capacity Manual* (1965) for highway facilities. Continuous parameters are divided into discrete levels, such as A through F, to characterize the quality of service at a specific point, segment, or facility [6]. Then, it is used on pedestrians, walkways, and stairways based on capacity or volume factors [7]. In public transport, Botzow [8] earlier started to adopt levels of service A through F for each attribute.

At present, the LOS of PTS commonly used comes from *Transit Capacity and Quality of Service Manual* (TCQSM) issued by Transportation Research Board of the United States. The first and second editions of TCQSM use A–F letters as LOS expression. In the third edition of the manual [5], LOS letters have been eliminated to be more practically guiding. Some studies get the LOS in a continuous normalized score, and it is not divided into discrete grades [9]. In this paper, the form of discrete levels is used and not limited to six levels.

PTS covers a wide time space, with many facilities and multiple service attributes, so the evaluation of its service quality is hard and complex. From existing literatures, an evaluation system with multiple LOS measures (attributes) cannot define the benchmarking of systems [10]. The practicable scheme is to combine the service attributes into one indicator, that is, to reflect the system LOS on a single scale. A widely implemented method is the weighted average.

In the weighted average method, the system LOS is treated as the function of all service attributes levels [11–14]. The model is TPS LOS = $\sum w_i \mathrm{LOS}(X_i)$, where $\mathrm{LOS}(X_i)$ is the service level of attribute X_i; w_i is the weight, and $\sum w_i = 1$. This method is simple in form and easy to use. $\mathrm{LOS}(X_i)$ can be obtained either from user surveys or existing LOS criteria. Meanwhile, w_i should be consistent with users' perception, otherwise the result of system service level may be unreliable.

The key of the method is to allocate w_i to determine the relative importance of each attribute. User surveys combined the approaches of Analytical Hierarchy Process [9], Logit model [15], Probit model [16], or Importance-Performance Analysis [17], etc. to achieve this work. In this regard, de Oña and de Oña [18] had made a detailed review. Considering that users often provide a fuzzy statement, Ndoh and Ashford [11] and Murugesan and Moorthy [12] used the fuzzy set theory, which can deal with the uncertainty of using nature language or letters quantization. Correia et al. [13] proposed a method to reduce the investigation workload, which let users answer the LOS of each component and a system LOS during an investigation, and then regression was used to get the weight by least square fitting.

An average of different user evaluations might lead to a deviation from the actual situation. This is because the users are not homogeneous on the perception of service quality and may judge the importance of each attribute differently, resulting from socioeconomic and many other factors [19]. Some studies developed LOS criteria for different types of users. Huo et al. [4] divided bus passengers into calm and anxious groups and found that for anxious passengers, the threshold value of waiting time was lower and that the value of bus running speed was higher. Yang et al. [20] found that the criteria of level A and B of male are higher than that of female when the authors studied the LOS of corridors in the subway station. Differently, Eboli and Mazzulla [21] introduced heterogeneity into user judgments and established Heterogeneous Customer Satisfaction Index model, which emphasized the attributes of homogeneity to users.

However, the weighted average method does not pay much attention to service distribution in time and space. It is more applicable for specific time-space scenarios because the distribution characteristics are more likely to change smoothly in one time period or a limited space. However, the service quality of public transport will be significantly different in different periods and different locations. It is discussed below.

In reality, with time and regions changing, the transportation demands and operation schemes provided by public transport agencies are varied. Latest studies have proved evidence that public transport services are not evenly distributed. For example, the research of Nikel et al. [17] showed apparent variance in the perceived service quality of passengers from route types of core, standard, express, and local routes. The research of Eboli et al. [22] showed a spatial autocorrelation of the transit service quality at rail stations in Milan, and the authors said spatial variance should be taken into account for rail service research.

Therefore, it should not simply average the statistical results of different time periods and regions to get the overall transit service level. Similar to the above-mentioned problem, Xin et al. [10] evaluated transit LOS along travel corridors using TCQSM in 2005, and at the end of the article, the authors pointed out several issues that need to be studied: how to evaluate the service quality of corridors with frequent departures in peak hours but no service in other periods; the analysis results depend on the definition of activity centers (it can be understood that evaluation results are affected by space); and others.

In addition, there are also some other methods for determining PTS LOS. Wiley et al. [23] built a model called "Local Index of Transit Availability" to evaluate the service level in the space of transit system. Huo et al. [4] used the fuzzy c-means clustering method to determine BRT system LOS criteria in China for passenger arrival time, waiting time, and bus running speed. Liou et al. [24] used the dominance-based rough set method and established a set of "if-then" rules to predict airport service level, considering 24 attributes. Devasurendra et al. [25] said that a total generalized cost function is also a potential method. These methods have its own advantages in different terms, but they also do not consider the influence of space-time distribution of service on PTS LOS.

On the basis of the above analyzes, the structure of service quality of PTS considering the distribution in time and space is described, as shown in Figure 1.

In Figure 1, for one type of system components in the PTS, the spatial distribution of service quality is reflected by the service indicators of the facilities at various locations, and the temporal distribution of service quality is described by the service indicator changes over time. It should be noted that although the service attributes and system components have a many-to-many relationship in reality, like both vehicles and stations have the attributes of crowding and cleanliness, it is assumed that the service quality of each type of system component is described by an attribute that is the principal concern for passengers. For example, the service attributes of station stairs and platforms in metro system that passengers care most about are good lighting and clear visibility [26], and platform crowding [1], respectively, also the factors related to security and safety are highly important to users at transit stops and stations in some cities [27].

Thus, how to reasonably evaluate the overall LOS of the PTS with varying service quality? A feasible solution is proposed in this paper and that is to establish a comprehensive distribution model and take it into account the service level determination method. To achieve this purpose, the distribution model should at least be able to describe the mean and deviation of service quality of all components in the system. The method framework based on a distribution model is introduced below.

It is worth mentioning that although the spatiotemporal distribution characteristics of different attributes may be different, the methods are applicable, so it is reasonable to begin our research on a single attribute. Specifically, the passengers of regular bus systems are sensitive to schedule reliability at transit stops [28]; while the passengers of metro systems are more concerned with the crowding [1], especially in peak hours the congestion of stations and service trains seriously affects the service quality. In this paper, we will take the attribute of train crowding as an example for specific modeling after establishing the method framework.

3. Framework of Determining PTS LOS Scale for One Attribute

In this section, we introduce the method framework of determining PTS LOS for a kind of component. We consider the variations of the component service quality in time and space, and this method is suitable for evaluating the LOS in an operating time cycle (like one day).

The framework mainly includes four steps:

Step 1. Select the indicator to measure the component.

Step 2. Model PTS LOS parameters of a single sample point (the concept of sample point is defined below).

Step 3. Calculate the PTS LOS parameters of all sample points.

Step 4. Determine the LOS scale and analysis.

3.1. Step 1. Select the Indicator to Measure the Component.

In this framework, one attribute is corresponding to a kind of component (see Figure 1). In reality, a kind of component in the system, such as cars, platforms, ticket vending machines, and others, usually has a main indicator to measure its service quality. For example, the train load factor is usually chosen to measure the service quality of service trains because it can describe the crowding of trains intuitively, which is a major concern for passengers. Here, denote the indicator γ to measure the service quality of the component.

There is a lot of research on LOS standards of public transit system components, such as subway platforms [1] and station corridors [2, 20]. The LOS standard of the components may be a reference of the global LOS of PTS in the following steps. We can adopt the established component LOS standards. It should be noted that the standards of other regions cannot be completely copied because different countries or cities have different economic and social factors, which will lead to different applicable standards. For example, for India, the best LOS for "passenger comfort" is

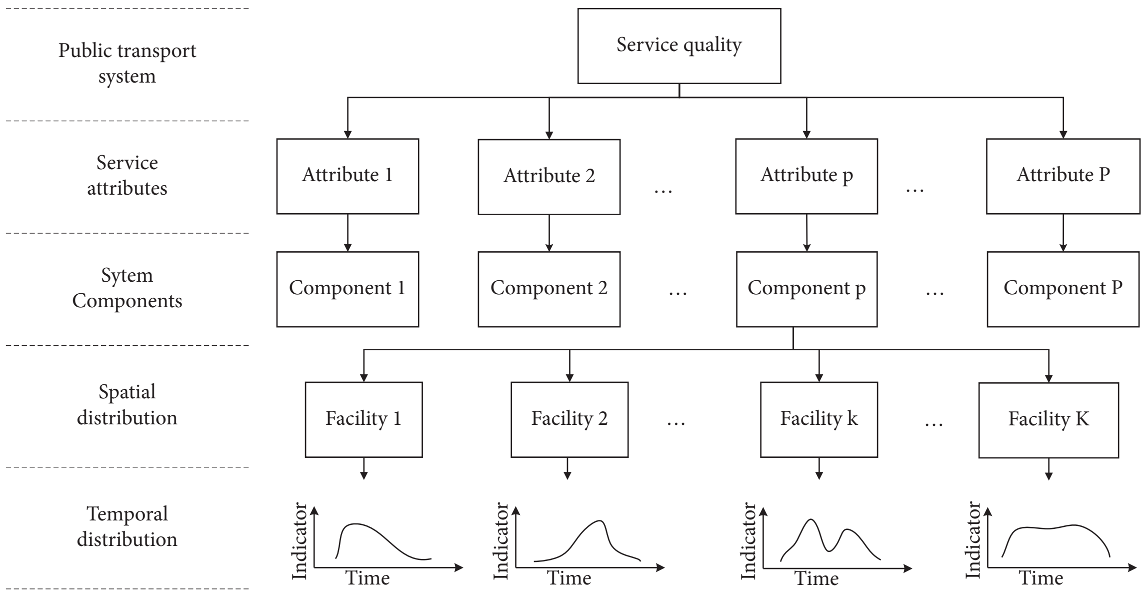

Figure 1: Structure of service quality of PTS considering the distribution in time and space.

≤ 1.5 persons per seat [29]; for the United States, it is 0–0.5; and for the city of Calcutta, it is ≤ 0.3 [30].

3.2. Step 2. Model PTS LOS Parameters of a Single Sample Point. In statistics, the mean value and coefficient of variation are often chosen to describe the data concentration trend. In the previous studies, these two parameters were already used to describe the distribution of subway transportation capacity utilization [31, 32], showing the applicability for public transport system. Therefore, in our model, these two parameters are applied simultaneously to model the distribution of service quality in the system.

With the change of time and location, we collect a lot of indicator values of the components. To reflect the characteristics of change over time and facilitate modeling, the whole operation time T (a periodic time, like a day) can be divided into n several periods, such as peak period and off-peak period, or division by 0.5 or 1 hour. $T = \{T_i \mid i = 1, 2, \ldots, n\}$. The characteristics are similar in each period but varied from time to time. In this paper, we call the calculation of parameters in a period a "sample point".

For the ith sample point, the model is established as

$$\begin{cases} A_i = f\left(\gamma_k^i\right) \\ C_i = g\left(\gamma_k^i\right) \end{cases}, \quad k = 1, 2, \ldots, K, \tag{1}$$

where f, g are the functions to determine the mean value A and the coefficient of variation C, respectively; γ_k^i is the indicator of the kth facility of the component in the system; and K is the total number of the facilities.

A sample point has the following characteristics:

(1) Representing the data in a period of the whole operation time

(2) Reflecting the results of the indicator values from all facilities of a kind of component in the system

(3) Consisting of two parameters: mean value and coefficient of variation

The aggregation of all sample points will reflect the spatiotemporal distribution of the service of PTS in an operational time cycle.

3.3. Step 3. Calculate the PTS LOS Parameters of All Sample Points. In this step, we apply the model in Step 2 to calculate the parameters of all sample points. The basis of this step is that the required data have been collected and processed.

3.4. Step 4. Determine the LOS Scale and Analysis. After getting all sample points, the goal of this step is to produce the thresholds of each level of service of PTS. Two specific methods are discussed: based on 1-dimensional coordinates and scale based on 2-dimensional coordinates.

3.4.1. Method 1. Based on 1-Dimensional Coordinates. When determining PTS LOS, the common approach is to calculate the mean or weighted mean value. For example, Li et al. used the average comfort level of all trains based on a timetable to represent the overall comfort level of the subway line in a day [33]. In this paper, we call the LOS scale based on the mean value A as scale based on 1-dimensional coordinates (Method 1).

In the form of mathematics, the PTS LOS scale determined by Method 1 can be expressed as

$$\text{Scale}\,(\text{PTS LOS}) = \Omega\,(A), \tag{2}$$

where Ω is the function to determine the thresholds of two successive levels. Take an example of dividing PTS LOS into

A–E levels, and the LOS scale by Method 1 is shown in Figure 2.

We can find that Method 1 is similar to the determination of the LOS scale of a component of PTS, and the difference is that the research object is the whole system or one kind of facility. Thus, it can refer to the threshold of component LOS to some extent. However, this method does not reflect the service distribution in the system. The following method is an improvement.

3.4.2. Method 2. Based on 2-Dimensional Coordinates. According to the model in Step 2, each sample point has 2 dimensions of parameters (mean value A and coefficient of variation C). Thus, we propose a method based on 2-dimensional coordinates (Method 2), where taking A and C as the axes of the coordinate system.

In the form of mathematics, the PTS LOS scale determined by Method 2 can be expressed as

$$\text{Scale}\,(\text{PTS LOS}) = \Omega\,(A, C). \tag{3}$$

If A and C are in linear relationship, two possible PTS overall LOS scales in the 2-dimensional coordinates are shown in Figures 3(a) and 3(b). The dashed lines represent the thresholds of adjacent service levels. It could also be nonlinear relationship. To get accurate thresholds, one should combine user perceptions, which is the same as component LOS.

Besides, an alternative approach is proposed to get the scale of PTS LOS here. The component LOS can be referred in this step. This approach takes the LOS of the representative component as the sample point's LOS. The LOS of the representative component may be at the worst level or at the best level or others during a period. Based on the LOS of a large number of sample points, we can easily get the thresholds. This is a compromise without a large-scale user investigation, which is not the part we focus on in this paper. With the existing data collected by operators, this method can help to preliminarily analyze the characteristics of the scale based on 2-dimensional coordinate systems.

To illustrate the framework clearly, we consider the train crowding attribute in urban rail transit in the following example. Step 1 and Step 2 correspond to Sections 4.1 and 4.2, respectively. Step 3 and Step 4 correspond to Sections 5.2 and 5.3, respectively.

4. PTS LOS of train crowding attribute

There are mainly two reasons why the attribute of train crowding in urban rail lines is chosen.

First, crowding has been widely studied in the evaluation of the service quality of public transit, especially in-vehicle crowding [34, 35]. Therefore, it would be representative to focus on evaluating crowding level of trains.

Second, train is a vital kind of component in urban rail transit service system. The crowding indicator of service trains varies obviously with time and station (see Figure 4). As a result, the distribution of train service in the whole day and all sections is uneven. It may have to add several short turning trains to relieve congestion during peak hours [36]. It is meaningful to assess this situation by applying the method proposed in this paper.

4.1. Indicator Measuring the Component LOS. Load factor or density of standees per square meter is usually used to assess in-vehicle crowding level [37]. With the using of information equipment such as automatic fare collection system and automatic passenger counting system, these indicators can be collected accurately. In this paper, train load factor is chosen as the crowding measure.

In this paper, the component LOS scale is established (see Table 1). Train LOS is divided into 6 levels, and the crowding is increasing from A to F. The limit of train load factor is 140%. This scale refers to the standards of China, such as "Code for Design of Metro" (Chinese national standard GB50157-2013) and "Construction Standard of Urban Rail Transit Project" (Chinese industry standard JB104-2008).

4.2. Model PTS LOS Parameters of a Single Sample Point

4.2.1. Symbol Definition. We focus on an urban rail transit network with L lines. The whole operational time T is divided into n periods, $T = \{[t_i, t_{i+1}] \mid i = 1, 2, \ldots, n\}$.

Each line is considered as a transportation system, independent of the others. The set of stations of line l is denoted as $S_l = \{1, 2, \ldots, s, \ldots s_l\}$. Every line has two-way operation and a 0-1 variable f is defined as the running direction. During the ith period $[t_i, t_{i+1}]$, the set of trip services of line l between stations s and $s+1$ in f direction is denoted as $K^i_{l,f,s} = \{1, 2, \ldots, k, \ldots, k^i_{l,f,s}\}$, where $k^i_{l,f,s}$ is the number of trip services.

Note that one train may cover two periods, which will change the value of K. In this paper, trip services are counted according to space-time diagram, and at the ith period, K equals to the number of trains which depart at $[t_i, t_{i+1}]$. Take a line with four stations as an example, as shown in Figure 5. During $[t_1, t_2]$ for f direction, the trip services between stations 2 and 3 are train 1–train 4, so $k^1_{l,f,2}$ equals 4.

According to the framework in Step 2, a sample point for one period is calculated. In order to distinguish the two-way results, the sample point is calculated for each direction. Thus, for each line, we will get $2n$ sample points, and thus $2nL$ sample points will be calculated for the whole subway network.

4.2.2. Parameters of Sample Points. The calculation scope of each sample point includes all trip segments (between two consecutive stations) in a direction of a line. According to Step 2, each sample point has 2 parameters.

The first one is the mean value, denoted by A. This parameter represents the general loading level of all service trains. The calculation of A is

$$A^i_{l,f} = \frac{\sum_{s=1}^{s_l-1} \sum_{k=1}^{k^i_{l,f,s}} \gamma^i_{l,f,s,k}}{\sum_{s=1}^{s_l-1} k^i_{l,f,s}}, \tag{4}$$

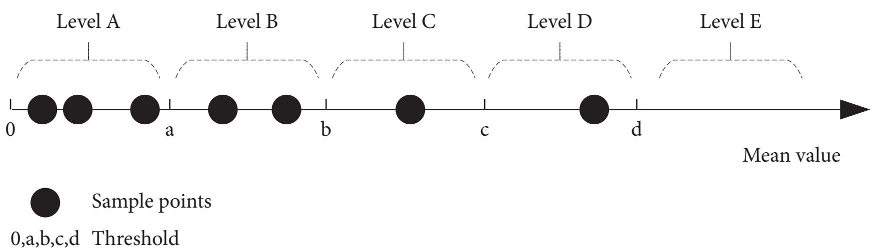

Figure 2: PTS overall LOS scale by Method 1.

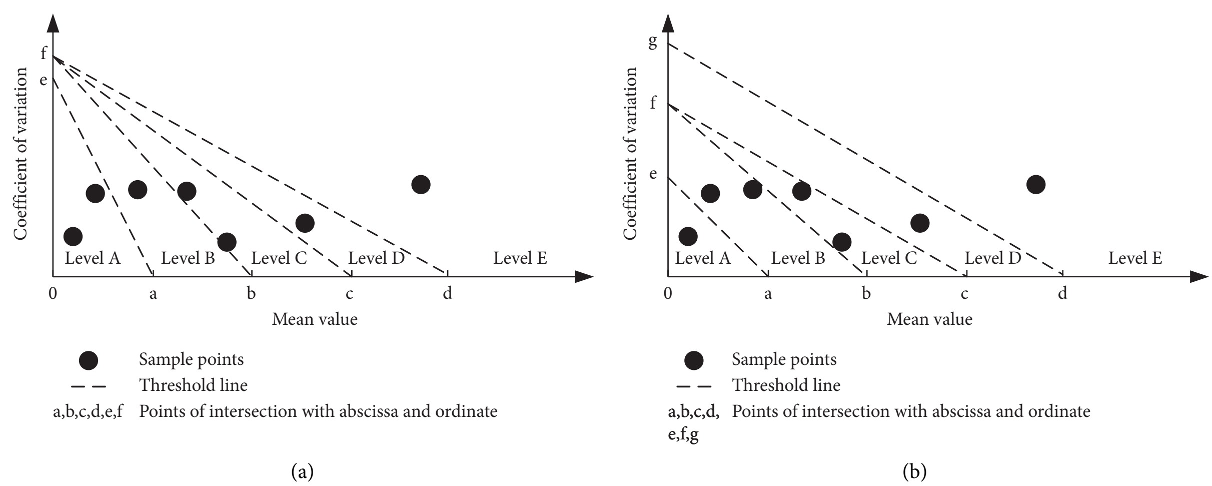

Figure 3: Possible overall LOS scales of PTS by Method 2.

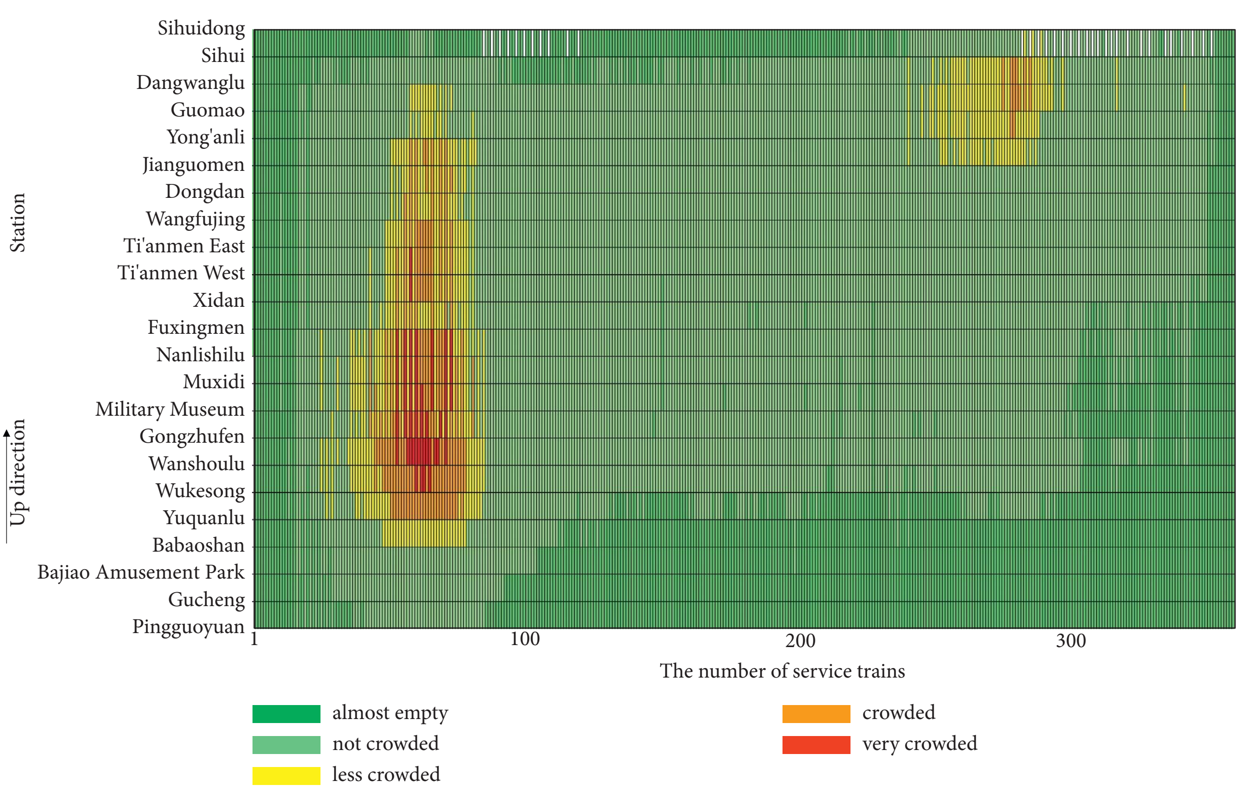

Figure 4: Crowding of trains in Beijing subway line 1 in up direction.

Table 1: The scale of train LOS.

Train LOS	Train load factor (%)	Description
A	[0, 17.5]	Almost empty
B	(17.5, 60]	Not crowded
C	(60, 80]	Less crowded
D	(80, 100]	Crowded
E	(100, 120]	Very crowded
F	(120, 140]	Overcrowded

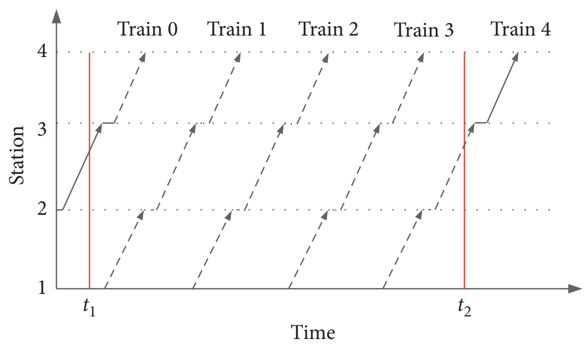

Figure 5: Calculation of K based on space-time diagram.

where $A^i_{l,f}$ donates the mean value of all trip services in the ith period in the direction f of line l; $\gamma^i_{l,f,s,k}$ is the train load factor of the kth train between station s and $s+1$; and $k^i_{l,f,s}$ is the number of trip services between station s and $s+1$.

The second variable is the coefficient of variation, denoted by C. This parameter represents the distribution deviation of the loading level of all trip segments during one period from the mean. The calculation of C is

$$C^i_{l,f} = \sqrt{\frac{\sum_{s=1}^{s_l-1}\sum_{k=1}^{k^i_{l,f,s}}\left(\gamma^i_{l,fs,k} - E\right)^2}{\sum_{s=1}^{s_l-1} k^i_{l,f,s}}} \times \frac{1}{E}, \qquad (5)$$

where $C^i_{l,f}$ denotes the coefficient of variation of all trip services in the ith period in the direction f of line l. And E is the expectation and $E = A^i_{l,f}$.

5. Case Study

A case of Beijing subway network is carried out to determine the line LOS scale, as shown in Figure 6. The network consists of 8 lines. Lines 1, 5, and 13 run through the urban center and extend to suburban; lines Batong, Changping and Yizhuang are suburban lines; and lines 2 and 10 are two loops. Operating time is from 5:00 to 24:00.

5.1. Basic Data. The basic data obtained include three parts: ridership of each trip segment during 5:00–24:00 of a weekday; timetables of each line in space-time diagram form; and the capacity of trains allocated to each line. The operation time is divided into 38 periods by 0.5 hours. Thus, each sample point represents a result during 0.5 hours. The load factor of a train at each trip segment can be easily obtained by dividing the number of passengers on train by train capacity.

5.2. Results of All Sample Points. According to equations (4) and (5), the mean value and coefficient of variation of each sample point are calculated. Since there is no train or passenger in some lines during the beginning or ending periods of the day, the coefficient of variation cannot be calculated. There are totally 589 valid sample points after removing 19 invalid samples. The results of parameters of all sample points are shown in Figure 7.

In Figure 7, each sample point is an aggregation of train crowding distribution at up or down direction during half an hour. Take the sample point α (0.71, 0.57) as an example. The sample point α represents train crowding distribution result of line 13 at up direction during 7:30–8:00 am, and specifically the space-time diagram of these trains is shown in Figure 8. Therefore, all sample points together will cover the network in Figure 6 and the whole operation time periods. On the basis of Figure 7, the line LOS scale can be determined using the two methods established in Step 4.

5.3. Determining Line LOS Scale. To obtain the line system LOS scale based on the 2-dimensional coordinates (Method 2), we take the representative component LOS as the line LOS of the sample point, as mentioned in Step 4. The line LOS of each sample point is represented by the most crowded level of the trains. The result is shown in Figure 9(a).

To compare with the line LOS scale based on 1-dimensional coordinate (Method 1, only uses the mean value), 2-dimensional coordinate system for Method 1 is also used. The result is shown in Figure 9(b). In Figure 9(b), the line LOS of each sample point is consistent with the level corresponding to the average load factor of all trains on the line according to Table 1.

Each level from A to F represents a distribution scenario of train crowding on the line. For example, in Figure 9(a), level D means that the maximum load factor of the service trains does not exceed 1.0; while in Figure 9(b), level D means that the average load factor of the service trains does not exceed 1.0. The difference is caused by the different models of the two methods.

We can see that there is a certain relationship between mean and coefficient of variation of each level in Figure 9. The comparison of the difference between the two results will help us to determine the LOS scale.

First, the number of sample points of each service level is different. In Figure 9(a), the number of sample points with service level A/B/C//E/F is 53/264/142/73/28/29, respectively, while in Figure 9(b), the number is 134/434/14/7/0/0, respectively, and 96% samples are in level A or B, without crowding results of levels E and F. The difference between the results of the two methods is quite apparent.

Second, the distribution patterns of samples with the same service level are different. The results obtained by Method 2 show that the distribution shape of each level is approximately like triangular, and the higher the coefficient

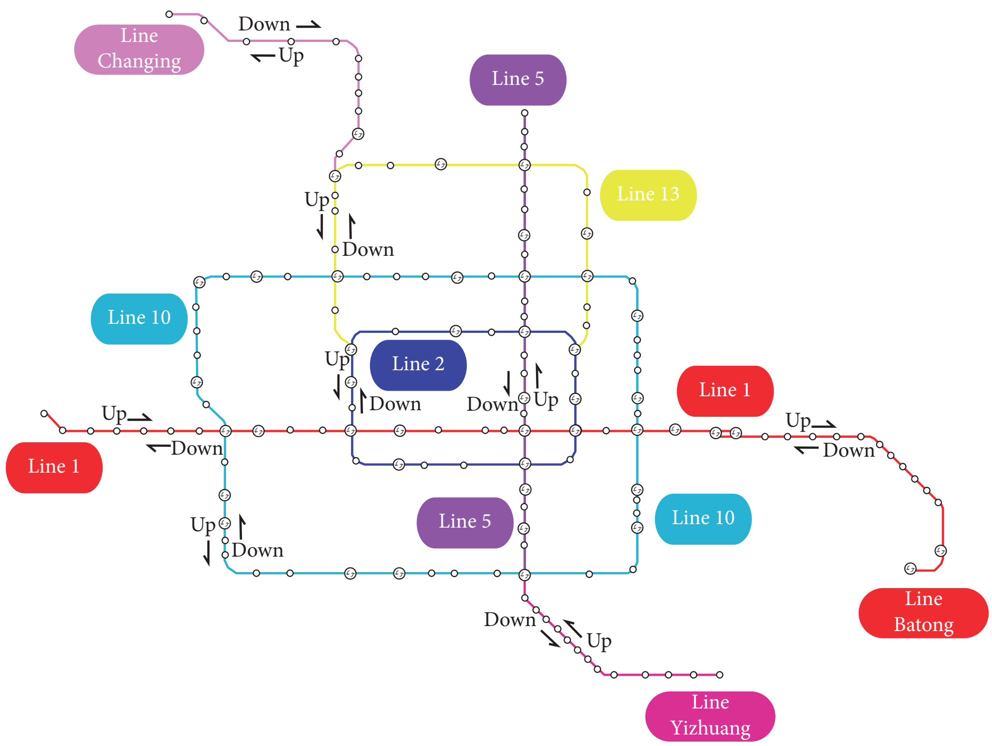

Figure 6: 8 lines of Beijing subway network.

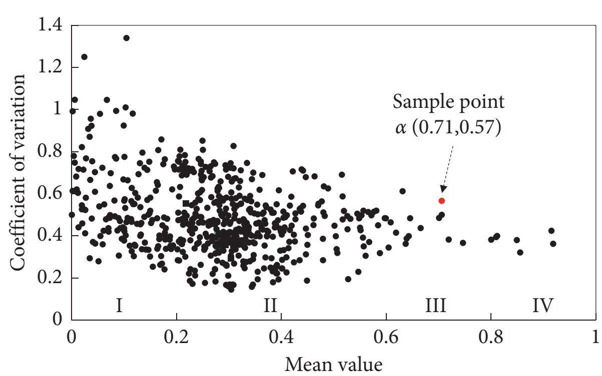

Figure 7: Results of parameters of all sample points.

of variation is, the lower the service level is. In reality, it is explained that a high coefficient of variation means the load is unbalanced; in other words, some trains are extremely crowded compared to the others. While the results obtained by Method 1 depend only on the mean value.

Then, we try to get the thresholds between adjacent levels. According to the distribution patterns of the sample points, the thresholds of the two methods are shown in Figures 10(a) and 10(b) dashed lines. In Figure 10(b), the overall service level of the subway line system is divided and renamed as level I – IV, with level I being the best and level IV the worst. In Figure 10(a), it is also divided into 4 levels to facilitate the comparison between the two methods.

In Figure 10(a), the solution of the dashed lines is to set the intersection points of the horizontal axis to be 0.175, 0.6, and 0.8, which are the same as Method 1, and then search the intercept of the vertical axis, to make the proportion of the sample points falling in the range the largest. The intercept of the vertical axis is 1.28. At this moment, 96% sample points of level A fall in zone I, 91% sample points of levels B and C fall in zone II, 62% sample points of level D fall in zone III, and 72% sample points of levels E and F fall in zone IV.

In Figure 10(a), the functions of the three dashed lines using parameters A (mean value) and C (coefficient of variation) are ① $C+7.31\,A=1.28$ (between levels I and II); ② $C+2.13\,A=1.28$ (between levels II and III); and ③ $C+1.6\,A=1.28$ (between levels III and IV). Therefore, the threshold of each level can be obtained according to the functions.

In summary, the results of the two methods are compared in terms of threshold, the number of sample points in each service level (NS_LOS), and the number of sample points in different time periods (NS_TP), as shown in Table 2.

The threshold of Method 2 is determined by two parameters A and C, but in Method 1 they are only decided by A. This is an important difference when taking the distribution of service quality into consideration to determine the LOS scale of PTS. In utilization, the overall LOS of the line can be evaluated by comparing the thresholds with the parameters of A and C. For example, when (A, C) is (0.4, 0.6), the line LOS is III by Method 2, meaning that there are trains in LOS D and it starts to be crowded for passengers; but the line LOS is II by Method 1, meaning that trains are in LOS B and it is overall comfortable for passengers. This is a phenomenon of practical significance, and for a detailed analysis, see 5.4.

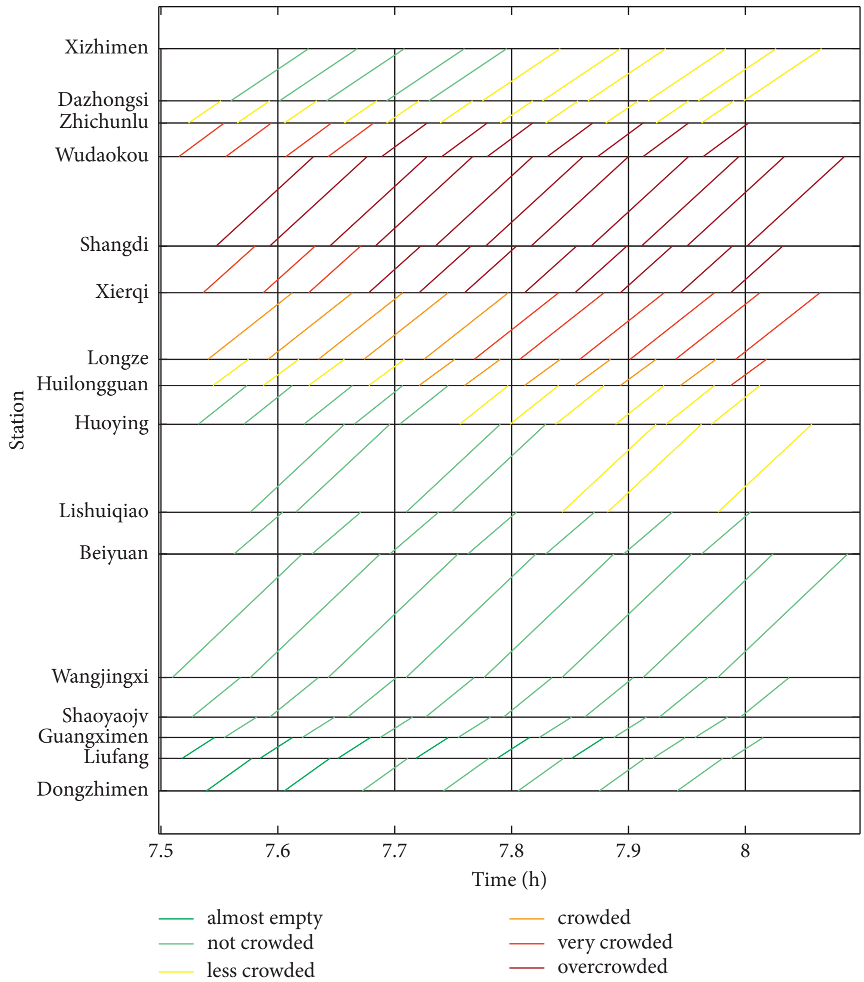

FIGURE 8: Crowding distribution reflected by sample point α

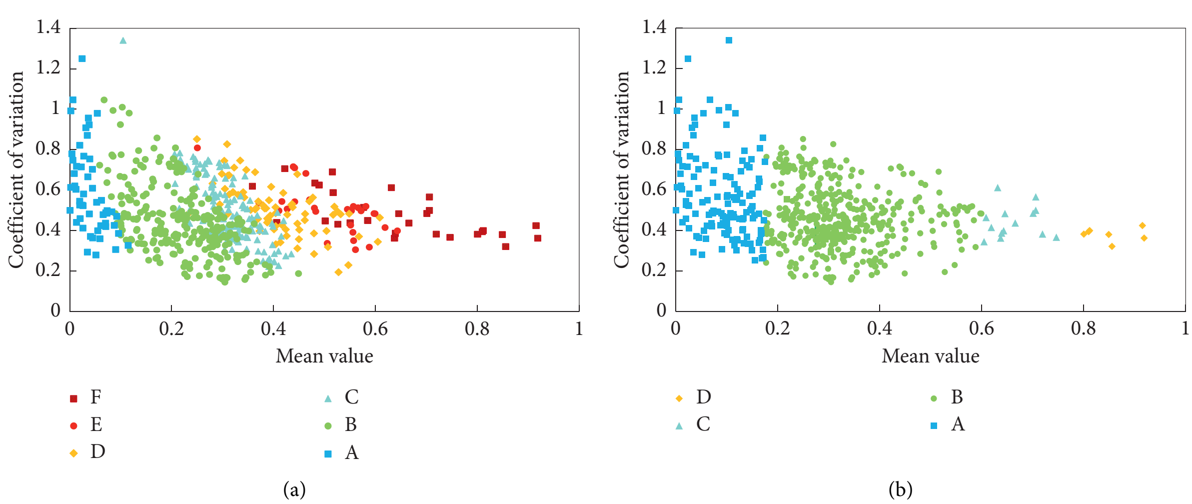

FIGURE 9: Initial results of line LOS of all sample points. (a) Based on Method 2 (b) Based on Method 1.

From Table 2, it is found that the result of Method 2 is more reasonable. There are totally 64 sample points in the morning peak periods, 37 of which by Method 2 (see item NS_TP) are distributed in level IV (trains are very crowded or overcrowded), and the ratio is 57.8%. According to Method 1, 39 sample points are in level (trains are nor or less

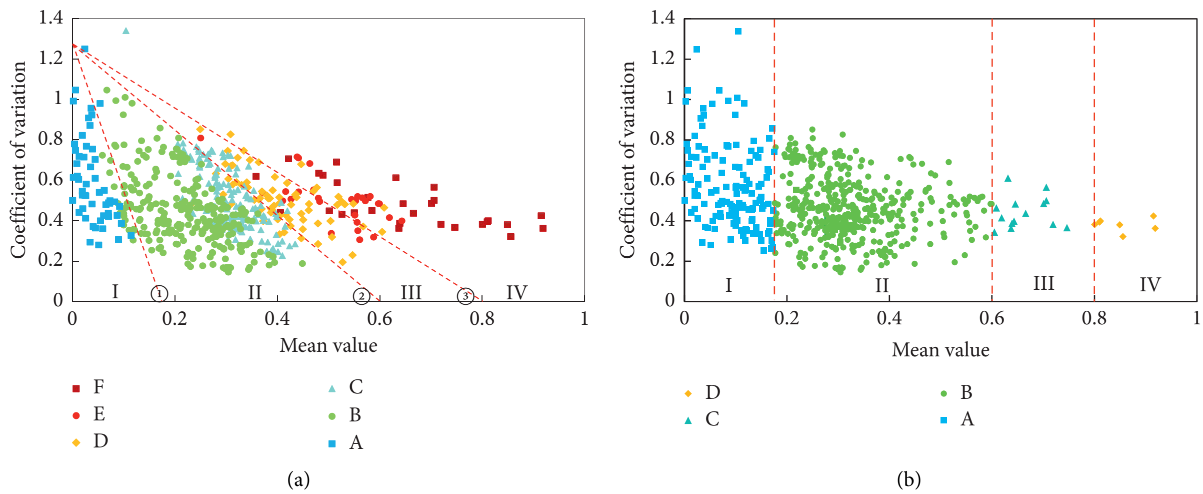

FIGURE 10: Thresholds of line system LOS. (a) Based on Method 2. (b) Based on Method 1.

TABLE 2: Comparison of Method 1 and Method 2.

Method	Item	I	II	III	IV
Method 2	Threshold	$0 \le C < -7.31A + 1.28$; and $0 \le A < 0.175$	$-7.31A + 1.28 \le C < -2.13A + 1.28$; and $0 \le A < 0.6$	$-2.13A + 1.28 \le C < -1.6A + 1.28$; and $0 \le A < 0.8$	$-1.6A + 1.28 \le C$; and $0 \le A$
	NS_LOS	51/17/0/0/0/0	1/247/124/16/0/0	0/0/17/45/12/4	1/0/1/12/16/25
	NS_TP	32/4/6/26	11/10/20/347	0/13/24/41	2/37/14/2
Method 1	Threshold	$0 \le A < 0.175$	$0.175 \le A < 0.6$	$0.6 \le A < 0.8$	$0.8 \le A$
	NS_LOS	134/0/0/0/0/0	0/434/0/0/0/0	0/0/14/0/0/0	0/0/0/7/0/0
	NS_TP	45/8/9/72	0/39/51/344	0/10/4/0	0/7/0/0

Note: (1) NT_LOS indicates the number of sample points whose worst LOS of trains is A/B/C/D/E/F, respectively. (2) NS_TP indicates the number of sample points in the beginning and ending periods (5 : 00–6 : 00, 23 : 00–24 : 00)/morning peak (7 : 00–9 : 00)/evening peak (17 : 30–19 : 30)/flat peak (other periods), respectively.

crowded), with a ratio of 60.9%. A similar situation occurs in the evening peak periods. In fact, during the peak periods in Beijing subway, the overall service level is low and the train crowding is generally very high because the transportation capacity cannot fully meet the demand. An evidence is that even more than 80 stations have implemented passenger flow control in Beijing subway because of overcrowded stations and trains in peak hours.

5.4. Sample Points That Operators Should Focus on. By comparing the results from Method 1 with those from Method 2, some sample points appear in different levels. It is reasonable for operators to pay attention to the sample points in level IV and optimize the train plan. Specifically, it is discussed in three situations.

First, for the sample points in level IV by Method 2 but in a better level than IV by Method 1, the difference of load factor among the trains on the line is large, and the utilization of line transportation capacity is uneven. Some trains with high load reduce the overall service level of line. This implies that operators should focus on improving the evenness of capacity utilization and reducing the high load factor.

Second, for the sample points in level IV by both two methods, the load factor of most trains is too high, and the overall transportation capacity needs to be improved.

In addition, for the samples in better level than IV by Method 2, the train plan can meet the needs of passengers, and the service is relatively well. Thus, it does not need to give priority to improvement for these sample points and no more discussion in this paper.

Take the sample point α (0.71, 0.57) as an example as before. The line LOS is IV according to Method 2 and III according to Method 1. This period belongs to the peak periods, and the headway is 5.5 minutes per shift. The spatial distribution of all trains LOS is shown in Figure 11.

It can be seen that the trains in LOS F (overcrowded, train load factor is greater than 120%) are concentrated in sections 11–13, and the maximum load factor is 129%–139%, which is close to the limit of train capacity. For passengers, this crowding level is unbearable. However, for sections 1–8 and 14–15, the train LOS are no worse than level C (train load factor does not exceed 80%). The utilization of transport capacity of the line is very uneven. Even if some trains are added at sections 8–15, the problem is still serious. At this time, it may be necessary to continue to increase transport capacity if possible, or implement passenger flow

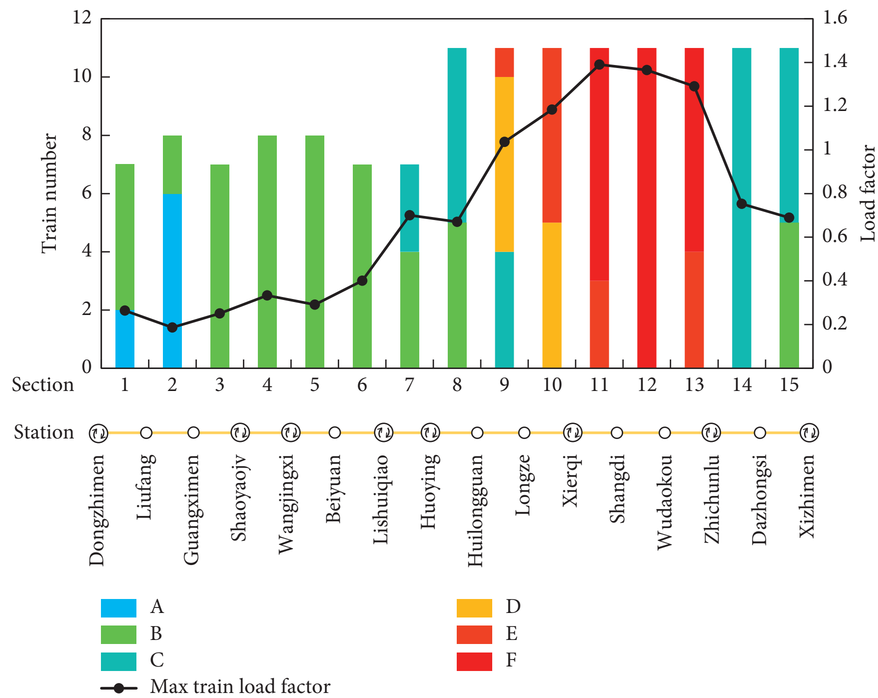

Figure 11: Spatial distribution of train LOS of line 13 in up direction during 7 : 30–8 : 00.

Table 3: Sample points that operators should pay attention to.

Sample point (A, C)	Line LOS		Line and period			Very crowded section	
	Meth-od 1	Meth-od 2	Line	Direction	Period	Section index	Percentage (%)
(0.92, 0.36)	IV	IV	Line Changping	Down	07 : 00-07 : 30	1–5	83.3
(0.92, 0.42)	IV	IV	Line Changping	Down	07 : 30-08 : 00	1–4	66.7
(0.86, 0.32)	IV	IV	Line Batong	Down	07 : 30-08 : 00	2–5	33.3
(0.85, 0.38)	IV	IV	Line 10	Up	08 : 00-08 : 30	4–8, 11, 12, 14–20, 24–29, 31	48.9
(0.81, 0.40)	IV	IV	Line Batong	Down	08 : 00-08 : 30	2–5	33.3
(0.81, 0.39)	IV	IV	Line Changping	Down	08 : 00-08 : 30	1–3	50.0
(0.80, 0.38)	IV	IV	Line 10	Up	07 : 30-08 : 00	4–8, 11–20, 24, 26, 28	40.0
(0.75, 0.37)	III	IV	Line 1	Up	08 : 00-08 : 30	6–11, 18	31.8
(0.72, 0.38)	III	IV	Line 10	Down	08 : 00-08 : 30	4, 6, 18–22, 37–42, 44	31.1
(0.71, 0.57)	III	IV	Line 13	Up	07 : 30-08 : 00	9–13	33.3
(0.71, 0.50)	III	IV	Line 5	Down	07 : 30-08 : 00	14–20	31.8
(0.70, 0.49)	III	IV	Line 5	Down	08 : 00-08 : 30	14–20	31.8
(0.67, 0.44)	III	IV	Line 1	Up	07 : 30-08 : 00	5–11	31.8
(0.65, 0.48)	III	IV	Line Batong	Up	18 : 30-19 : 00	2–4	25.0
(0.64, 0.39)	III	IV	Line 10	Down	07 : 30-08 : 00	4, 6, 21, 22, 37, 39–42, 44, 45	26.7

Table 3: Continued.

Sample point (A, C)	Line LOS		Line and period			Very crowded section	
	Meth-od 1	Meth-od 2	Line	Direction	Period	Section index	Percentage (%)
(0.64, 0.36)	III	IV	Line 10	Down	18 : 00-18 : 30	4, 6, 13–17, 24	17.8
(0.63, 0.61)	III	IV	Line 13	Up	08 : 00-08 : 30	9–13	33.3
(0.58, 0.45)	II	IV	Line 5	Up	18 : 00-18 : 30	13, 15, 16	13.6
(0.58, 0.52)	II	IV	Line 13	Down	07 : 30-08 : 00	3–5, 7	26.7
0.55, 0.44)	II	IV	Line 13	Down	08 : 00-08 : 30	4	6.7
(0.52, 0.59)	II	IV	Line 13	Down	18 : 00-18 : 30	10–12	20.0
(0.52, 0.69)	II	IV	Line Yizhuang	Up	07 : 30-08 : 00	10–13	23.1
(0.49, 0.63)	II	IV	Line 13	Up	07 : 00-07 : 30	11–13	20.0
(0.48, 0.63)	II	IV	Line Yizhuang	Up	07 : 00-07 : 30	13	7.7
(0.42, 0.71)	II	IV	Line 13	Down	18 : 30-19 : 00	11–12	13.3

Note: Section index is the sequence number of each line's section in up direction.

control measures in Huilongguan, Longze, and Xierqi stations to reduce the train load factor.

According to the same analysis as above, the sample points that operators should pay attention to are separated. These sample points are in level IV by Method 2. Due to space limitation, only some sample points with trains in LOS F are listed here (see Table 3). If there are trains in LOS E or F in a section, we call it a very crowded section. The service information of these sample points is also listed in Table 3.

It can be seen that all of the sample points listed in Table 3 are distributed in the morning and evening peak hours, and mostly in the morning peak hours. In specific, in the evening peak hour of 18 : 00–19 : 00, operators only need to pay attention to line 5 in up direction, line 10 in down direction, line 13 in down direction, and line Batong in up direction.

Furthermore, we get the indexes of crowded and very crowded sections, and the ratio of the number of such sections to the total number of sections of the line, as shown in the last two columns of Table 3. For the sample points in levels IV, III, and II by Method 1, the percentages are 33.3–83.3%, 17.8–33.3%, and 6.7–26.7%, respectively. It means that the percentage of crowded and very crowded sections is changed in a relatively wide range. It is caused by the alternative approach in Step 4. But the present results can already show the relevant conclusions above and can be accepted.

For the lines whose sections are mostly crowded and very crowded, such as line Changping in down direction during 7 : 00–8 : 00 and the percentage exceeds 66.7%, we suggest that the transport capacity of the whole line should be improved. The measures such as increasing departure frequency or train capacity can be adopted. On the other hand, for the lines with a certain percentage of crowded sections, such like line 1 in up direction during 7 : 30–8 : 30 and the percentage is 31.8%, we suggest that the strengthening of the transportation capacity for the very crowded sections should be considered. The specific optimization measures need to consider the actual constraints such as the minimum departure frequency, length of the platform, train fleet size, and other factors.

6. Conclusion

In this paper, we study the problem of determining the level of service (LOS) scale of public transport system considering the distribution of service quality. A method framework is proposed to solve it for a single service attribute. Taking the attribute of crowding in urban rail transit system as an example, the distribution model consisting of coefficient of variation and mean is concretely carried out. In the end, the method is applied to Beijing subway. The main conclusion is stated below.

The influence of coefficient of variation and mean on the subway line LOS is obtained. Coefficient of variation describes the distribution deviation of train service quality in the system. A high value of coefficient of variation means that the line loading is unbalanced, and the crowdedness of some trains is relatively high, which does not favor the overall LOS of the line system. The larger the mean is, the more unfavorable the line LOS is, which is the same as the result from the existing method. At the same service level, the coefficient of variation is negatively correlated with the mean indicator.

The LOS scale threshold of the line system is obtained, which is expressed by the parameters of mean (denoted as A) and coefficient of variation (denoted as C). The LOS of Beijing subway lines in this paper are divided into level I, II,

Table 4: Subway line LOS.

Line LOS	Threshold	Description
I	$0 \le C < -7.31A + 1.28$; and $0 \le A < 0.175$	Seats of all trains are not fully utilized; passengers feel very comfortable.
II	$-7.31A + 1.28 \le C < -2.13A + 1.28$; and $0 \le A < 0.6$	Seats of all trains are fully utilized and some passengers have to stand; passengers feel comfortable.
III	$-2.13A + 1.28 \le C < -1.6A + 1.28$; and $0 \le A < 0.8$	The overall capacity of all trains is utilized well; some trains start to be crowded.
IV	$-1.6A + 1.28 \le C$; and $0 \le A$	There are crowded and overcrowded trains; passengers feel the worst.

III, and IV. Based on the standard of train LOS, the capacity utilization and congestion of each line LOS can be described, as shown in Table 4.

The determination result is better by using the two parameters, mean and coefficient of variation, than that of only using mean. A convincing evidence is that, in the early peak-hour period, 57.8% of the line service is in the worst service level IV, while this percentage is 60.9% in level II by the present method. The result of the proposed method based on the distribution is more consistent with the reality.

We suggest that operators should pay more attention to the lines with a service level of IV. In particular, when the evaluation results using the proposed method are inconsistent with that of using the existing method, it indicates that the service quality of trains is unevenly distributed, and optimization measures should be considered. This implies that the research in this paper can provide an effective evaluation tool for judging and improving the service quality of public transport system.

The validity of the proposed method framework has been confirmed by the case study results from the train crowding models of Beijing subway network. These results show that considering service distribution to determine the overall service level of public transport system is very useful and necessary for capturing more detailed information of the system performance in time and space and targeted improving system service level.

At last, this paper mainly focuses on the method of how to determine the LOS scale of public transport system considering service quality distribution and does not consider user perceptions. In the future, it should pay attention to the following questions: what is the user's overall feeling when he or she passes through multiple facilities with different service qualities? And what is the heterogeneity of the user's overall feeling? Do the public transport systems in other cities follow the same rules?

About the classification on service levels in the following study, other approaches based on data mining like clustering analysis are worth trying for it is practical for public transport operators. Mapping the classification on the physical network can help to conveniently observe the distribution, just like the research of [38] on emission factor.

Data Availability

The data used to support the findings of this study are available on request through the third author at the e-mail address 18114052@bjtu.edu.cn.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the Fundamental Research Funds for the Central Universities (2020YJS101) and the National Natural Science Foundation of China (71971021). The authors deeply appreciate the support.

References

[1] S. Cao, Z. Yuan, Y. Li, P. Hu, and L. Johnson, "Study on service-level classification of platforms in Beijing Urban rail transit," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2112, no. 1, pp. 127–135, 2009.

[2] S. Cao, Z. Yuan, C. Zhang, and L. Zhao, "LOS classification for urban rail transit passages based on passenger perceptions," *Journal of Transportation Systems Engineering and Information Technology*, vol. 9, no. 2, pp. 99–104, 2009.

[3] A. R. Correia, S. C. Wirasinghe, and A. G. de Barros, "Overall level of service measures for airport passenger terminals," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 2, pp. 330–346, 2008.

[4] Y. Huo, J. Liu, J. Zhang, X. Li, and Y. Guo, "Development of fuzzy level of service criteria for bus rapid transit considering user heterogeneities in China," *Journal of Advanced Transportation*, vol. 2020, 2020.

[5] Transportation Research Board, *Transit Capacity and Quality of Service Manual*, Transportation Research Board, Washington, DC, USA, 3rd edition, 2013.

[6] R. P. Roess, M. A. Vandehey, and W. Kittelson, "Level of service: 2010 and beyond," *Transportation Research Record*, no. 2173, pp. 20–27, 2010.

[7] J. J. Fruin, "Designing for pedestrians: a level- of- service concept," *Highway Research Record*, vol. 335, pp. 1–15, 1971.

[8] H. Botzow, "Level-of-service concept for evaluating public transport," *Transportation Research Record*, vol. 519, pp. 73–84, 1974.

[9] T. Tallam, H. Yallabandi, and C. N. Kumar, "Determination of level-of-service for public transport: a case study for hyderabad metro," *Advances in Civil Engineering Select*, vol. 2021, pp. 439–451, 2019.

[10] Y. Xin, L. Fu, and F. F. Saccomanno, "Assessing transit level of service along travel corridors case study using the transit capacity and quality of service manual," *Transportation Research Record*, no. 1927, pp. 259–267, 2005.

[11] N. N. Ndoh and N. J. Ashford, “Evaluation of transportation level of service using fuzzy sets,” *Transportation Research Record*, no. 1461, pp. 31–37, 1994.

[12] R. Murugesan and N. V. R. Moorthy, “Level of public transport service evaluation: a fuzzy set approach,” *Journal of Advanced Transportation*, vol. 32, no. 2, pp. 216–240, 1998.

[13] A. R. Correia, S. C. Wirasinghe, and A. G. de Barros, “A global index for level of service evaluation at airport passenger terminals,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 4, pp. 607–620, 2008.

[14] D. Sun and S. Guan, “Measuring vulnerability of urban metro network from line operation perspective,” *Transportation Research Part A: Policy and Practice*, vol. 94, no. 800, pp. 348–359, 2016.

[15] L. Eboli and G. Mazzulla, “How to capture the passengers’ point of view on a transit service through rating and choice options,” *Transport Reviews*, vol. 30, no. 4, pp. 435–450, 2010.

[16] R. C. P. Wong, W. Y. Szeto, L. Yang, Y. C. Li, and S. C. Wong, “Elderly users’ level of satisfaction with public transport services in a high-density and transit-oriented city,” *Journal of Transport & Health*, vol. 7, pp. 209–217, 2017.

[17] C. Nikel, G. Eldeeb, and M. Mohamed, “Perceived quality of bus transit services: a route-level analysis,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 2, pp. 79–91, 2020.

[18] J. de Oña and R. de Oña, “Quality of service in public transport based on customer satisfaction surveys: a review and assessment of methodological approaches,” *Transp. Sci.*vol. 49, no. 3, pp. 605–622, 2014.

[19] M. Bordagaray, L. dell’Olio, A. Ibeas, and P. Cecín, “Modelling user perception of bus transit quality considering user and service heterogeneity,” *Transportmetrica: Transportation Science*, vol. 10, no. 8, pp. 705–721, 2014.

[20] Y. Yang, H. Li, X. Jiang, X. Zheng, X. Xu, and Z. Qingyu, “LOS of pedestrian perception for corridor in subway station considering the reliability and validity,” *J. Transp. Syst. Eng. Inf. Technol.*vol. 16, no. 2, pp. 212–218, 2016.

[21] L. Eboli and G. Mazzulla, “A new customer satisfaction index for evaluating transit service quality,” *Journal of Public Transportation*, vol. 12, no. 3, pp. 21–37, 2009.

[22] L. Eboli, C. Forciniti, and G. Mazzulla, “Spatial variation of the perceived transit service quality at rail stations,” *Transportation Research Part A: Policy and Practice*, vol. 114, pp. 67–83, 2018.

[23] K. Wiley, H. Maoh, and P. Kanaroglou, “Exploring and modeling the level of service of urban public transit: the case of the Greater Toronto and Hamilton Area, Canada,” *Transportation Letters*, vol. 3, no. 2, pp. 77–89, 2011.

[24] J. J. H. Liou, C.-H. Tang, W.-C. Yeh, and C.-Y. Tsai, “A decision rules approach for improvement of airport service quality,” *Expert Systems with Applications*, vol. 38, no. 11, pp. 13723–13730, 2011.

[25] K. W. Devasurendra, S. C. Wirasinghe, and L. Kattan, “A critical review of transit level of service measures and an overview of a proposed new approach,” in *Proceedings of theInternational Conference on Transportation and Development*, pp. 1–12, Seattle, WA, USA, May 2020.

[26] J. Y. S. Lee and W. H. K. Lam, “Levels of service for stairway in Hong Kong underground stations,” *Journal of Transportation Engineering*, vol. 129, no. 2, pp. 196–202, 2003.

[27] H. Iseki and B. Taylor, “Style versus service? An analysis of user perceptions of transit stops and stations,” *Journal of Public Transportation*, vol. 13, no. 3, pp. 23–48, 2010.

[28] L. A. Bowman and M. A. Turnquist, “Service frequency, schedule reliability and passenger wait times at transit stops,” *Transportation Research Part A: General*, vol. 15, no. 6, pp. 465–471, 1981.

[29] S. Das and D. Pandit, “Importance of user perception in evaluating level of service for bus transit for a developing country like India: a review,” *Transport Reviews*, vol. 33, no. 4, pp. 402–420, 2013.

[30] S. Das and D. Pandit, “Determination of level-of-service scale values for quantitative bus transit service attributes based on user perception,” *Transportmetrica: Transportation Science*, vol. 11, no. 1, pp. 1–21, 2015.

[31] R. Shi, *Calculation and Optimization for the Conveying Capacity Utlization of Urban Rail Transit*, Beijing Jiaotong University, Beijing, China, 2017.

[32] M. Wang, B. Mao, Y. Yang, R. Shi, and Y. Wang, “Line loading evaluation method for urban rail transit considering train loading evenness,” *Journal of Transportation Systems Engineering and Information*, vol. 21, no. 2, pp. 98–104, 2021.

[33] S. Li, R. Xu, and Z. Jiang, “Evaluation method for matching degree between train diagram capacity and passenger demand for urban rail transit,” *Journal China Railway Science*, vol. 38, no. 3, pp. 137–144, 2017.

[34] Z. Li and D. Hensher, “Crowding in public transport: a review of objective and subjective measures,” *Journal of Public Transportation*, vol. 16, no. 2, pp. 107–134, 2013.

[35] M. Börjesson and I. Rubensson, “Satisfaction with crowding and other attributes in public transport,” *Transport Policy*, vol. 79, pp. 213–222, 2019.

[36] X. Ding, S. Guan, D. J. Sun, and L. Jia, “Short turning pattern for relieving metro congestion during peak hours: the substance coherence of Shanghai, China,,” *European Transport Research Review*, vol. 10, no. 2, 2018.

[37] A. Tirachini, L. Sun, A. Erath, and A. Chakirov, “Valuation of sitting and standing in metro trains using revealed preferences,” *Transport Policy*, vol. 47, pp. 94–104, 2016.

[38] D. Sun, K. Zhang, and S. Shen, “Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data,” *Transportation Research Part D: Transport and Environment*, vol. 62, no. 800, pp. 699–714, 2018.