Analyzing urban bus service reliability at the stop, route, and network levels

Xumei Chen, Lei Yu, Yushi Zhang, Jifu Guo

School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, PR China
Department of Transportation Studies, Texas Southern University, 3100 Cleburne Avenue, Houston, TX 77004, United States
Beijing Transportation Research Center, Beijing 100055, PR China

ABSTRACT

Improving the reliability of bus service has the potential to increase the attractiveness of public transit to current and prospective riders. An understanding of service reliability is necessary to develop strategies that help transit agencies provide better services. However, few studies have been conducted analyzing bus reliability in the metropolis of China. This paper presents an in-depth analysis of service reliability based on bus operational characteristics in Beijing. Three performance parameters, punctuality index based on routes (PIR), deviation index based on stops (DIS), and evenness index based on stops (EIS), are proposed for the evaluation of bus service reliability. Reliability involves routes, stops, punctuality, deviation, and evenness. The relationship among the three parameters is discussed using a numerical example. Subsequently, through a sampling survey of bus lines in Beijing, service reliability at the stop, route, and network levels are estimated. The effects of route length, headway, the distance from the stop to the origin terminal, and the use of exclusive bus lanes are also analyzed. The results indicate low service reliability for buses in Beijing and a high correlation between service reliability and route length, headway, distance from the stop to the origin terminal, and the provision of exclusive bus lanes.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Bus service, especially in busy urban areas, is facing the challenge to enhance its reliability. Studies have revealed that the preferences for public transportation in a standard commuting situation were enhanced not only by the belief that public transportation provided a shorter average travel time than traveling by car, but also by the belief that public transportation was at least very reliable (i.e., buses have an equal or lower variability in travel time compared to car) (Van Vugt et al., 1998). However, compared with subway or light rail, the operation of a bus is more likely to suffer from the effects of a range of factors, including traffic congestion, weather, passenger loading variations, and operating staff behavior. It is recognized that unreliability can seriously undermine the attractiveness of bus service and increase costs due to lost mileage and lower fleet utilization for operators (Lin et al., 2008). Hence, most transit agencies have monitored service reliability as one of the key performance measures for their bus operations (Benn and Barton-Ashman, 1995).

In China, the population density is so high that widespread car usage is very difficult to sustain. The preferential development of public transport systems is a vital national development strategy. During the period that various cities are experiencing rapid motorization, buses and, more generally at-grade public transportation remain the most important components of transit services in urban areas. However, while several cities assign dedicated lanes and transit signal
priorities for bus operations according to actual road conditions, in most cases, buses are running in the traffic flow where a high volume of non-motorized traffic mixes with motor vehicles. That means that the operation of buses is within the context of more congestion and disruption (Chen et al., 2008). Hence, accessing the transit service reliability/unreliability and identifying underlying causes in order to find countermeasures to provide consistent service is becoming increasingly important. But in practice, existing reliability assessments measure on-time performance of buses by means of statistically analyzing whether the whole running time along the routes is according to scheduled time (i.e., terminal on-time performance). Thus, transit agencies cannot precisely determine how buses run at the en route stops and may have a distorted view of the transit service. From the perspective of passengers, reliability at the stop level is more sensitive for their trips. To this end, this paper proposes a set of feasible reliability measures for buses regarding variability of journey time at routes and headway at stops. This information will equip transit agencies with the ability to identify potential service problems at the stop, route, or network level, which may ultimately lead to improved bus services.

In order to assess service reliability based on bus operational characteristics of China, we conducted an in-depth analysis of performance parameters for bus service reliability. Using the transit system in Beijing as an example, three performance parameters, punctuality index based on routes (PIR), deviation index based on stops (DIS), and evenness index based on stops (EIS), are proposed for the evaluation of bus service reliability. Reliability involves routes and stops, as well as punctuality, deviation and evenness. The relationship among these three parameters is discussed using a numerical example. Subsequently, through a sampling survey of bus lines in Beijing, service reliability at the stop, route, and network levels are estimated by employing the proposed three performance parameters. The effects of route length, headway, distance from the stop to the origin terminal, and exclusive bus lanes on service reliability are also analyzed. Several strategies for increasing urban bus transit reliability are suggested.

2. Existing studies

In recent years, some studies have been conducted to determine bus service reliability using measures of on-time performance, headway or headway adherence between buses, bus running times or running time variation, and excess waiting time. The factors influencing service reliability have also been widely analyzed.

The work of Stermen and Schofer (1976) was among the early studies on bus service reliability in the United States. This study was conducted to test a particular measure of reliability, the inverse of the standard deviation of point-to-point travel times, using data from bus services in the Chicago area. The selected measure was found to be a useful and easily collected indicator of service reliability. The reliability measured in this form was found to be significantly degraded by increasing the route length, intensity of intersection control, traffic volumes and, with less certainty, bus passenger loadings. Several strategies, such as decreasing the route length, intensity of intersection control and traffic volumes, were suggested to improve the urban bus transit reliability.

Turnquist (1978) proposed a model of estimating bus and passenger arrivals at a bus stop. The impacts on the expected waiting time of the service frequency and reliability for random and nonrandom arriving passengers were identified. The effects of the frequency and reliability on the proportion of the user population who plan their arrival times were also explored through a small empirical study.

Abkowitz and Engelstein (1983, 1984) studied factors affecting the running time on transit routes and methods for maintaining transit service regularity. These studies reported on estimating empirical models of transit mean running time and identifying operations-control actions to improve reliability. It was found that mean running time is strongly influenced by trip distance, people boarding and alighting, and signalized intersections. The proposed method for maintaining service regularity through improved scheduling and real-time control was found to be a reasonable solution. It was suggested that improvements of reliability may be possible through planning shorter routes or adopting investment strategies which emphasize improved control rather than modification of existing link characteristics.

Bates (1986) conducted a survey to determine basic practices and attitudes concerning the on-time performance in bus transit operations. Based on responses from 146 bus operators, it appeared that there was a wide variation in the definition of on-time performance, although the standard definition of no more than 1 min early and no more than 5 min late would include the definitions employed by most systems. The determination of on-time performance appears to be largely informal practice with little statistical basis. There was almost universal agreement that on-time performance is an important aspect of transit operations and there was strong support for research in this area.

Strathman and Hopper (1993) presented an empirical assessment of factors affecting the on-time performance of the fixed route bus system in Portland, Oregon. A multinomial logit model relating early, late, and on-time bus arrivals to route, schedule, driver, and operating characteristics was developed and estimated. The model results showed that the probability of on-time failures increased during PM peak periods, with longer headways and higher levels of passenger activity, and as buses progress further along their routes. Part-time drivers were also more likely to fall behind the schedule. With few exceptions, schedule changes and operations-control actions could mitigate these effects. Later, Strathman et al. (1999) conducted a baseline analysis of service reliability on selected routes, focusing on running times, headways, and on-time performance when Tri-Met, the transit provider in Portland, Oregon, was implementing a new computer aided bus dispatching system that used a satellite-based global positioning system (GPS) to track vehicle locations. Reliability was found to vary according to the route characteristics, direction, and time of day.

Strathman and Hopper (1993) presented an empirical assessment of factors affecting the on-time performance of the fixed route bus system in Portland, Oregon. A multinomial logit model relating early, late, and on-time bus arrivals to route, schedule, driver, and operating characteristics was developed and estimated. The model results showed that the probability of on-time failures increased during PM peak periods, with longer headways and higher levels of passenger activity, and as buses progress further along their routes. Part-time drivers were also more likely to fall behind the schedule. With few exceptions, schedule changes and operations-control actions could mitigate these effects. Later, Strathman et al. (1999) conducted a baseline analysis of service reliability on selected routes, focusing on running times, headways, and on-time performance when Tri-Met, the transit provider in Portland, Oregon, was implementing a new computer aided bus dispatching system that used a satellite-based global positioning system (GPS) to track vehicle locations. Reliability was found to vary according to the route characteristics, direction, and time of day.
Nakanishi (1997) described NYCT's new bus performance indicators, which were established in 1994; suggested further analyses; and recommended an additional indicator that may be added to the program. The suggestions were not intended to detract from the agency's successful implementation of the indicators, but were meant to further enhance the quality and usefulness of the indicators.

Yin et al. (2004) developed a generic simulation-based approach to assess transit service reliability, taking into account the interaction between the network performance and passengers’ route choice behavior. Three types of the reliability, system-wide travel time reliability, schedule reliability, and direct boarding waiting-time reliability, were defined from perspectives of the community or transit administration, the operator, and passengers. A Monte Carlo simulation approach with a stochastic user equilibrium transit assignment model embedded in it was proposed to quantify these three reliability measures of the transit service. A simple transit network with a Bus Rapid Transit (BRT) corridor was analyzed as a case study in which the impacts of BRT components on the transit service reliability were evaluated preliminarily.

An exhaustive list of transit reliability measure examples was shown in the Transit Capacity and Quality of Service Manual (TCQSM) (Kittelson et al., 2003). On-time performance and headway adherence, the most widely used reliability measures in the transit industry, were especially discussed. The on-time performance LOS (i.e., Level of Service) and headway adherence LOS were provided in this manual. Later, Camus et al. (2005) discussed the advantages and limitations of the TCQSM methods for the estimation of the reliability LOS, based on automated vehicle location (AVL) data collected in Trieste, Italy. A new performance measure called weighted delay index was proposed, allowing the amount of delay, the effect of early departures on users, and a flexible tolerance to be considered. This allowed the limitations of the TCQSM to be overcome.

Lu and Ismutulla (2006) set up a model that contained the transferring via three public transport routes with different running time reliabilities. The model was applied to simulate the impacts of the departure time reliability of public transport services on the arrival lateness. A simulation method for the analysis and evaluation of running time reliability of transport systems based on Matlab was introduced. Furthermore, according to the simulation method, a case study was carried out.

The studies identified in the literature review have offered insights into bus service reliability and the factors contributing to the bus service variability. They suggested that the on-time/punctuality performance and headway evenness are primary focuses in the practice of bus reliability analysis. However, most of the studies reviewed primarily relied upon analyzing one or several transit routes and the performance parameters’ lack of necessary correlation. Little effort has been made to explore bus reliability assessment on a large scale network and investigate the relationship between different performance parameters for bus reliability. This work has expanded upon previous research on transit service reliability by analyzing the relationship between performance parameters of on-time performance and headway evenness. The methodology of the reliability assessment with short frequency due to large passenger demand on a large scale bus network has also been explored. Performance parameters developed in this research attempt to more closely address transit service reliability from the perspective of passengers. Passenger flow based weight is proposed to aggregate reliability from the stop level to route level.

3. Performance parameters

Transit service reliability has been defined in a variety of ways. From one point of view, it is defined as the ability of a transit system to adhere to a schedule or maintain regular headways and a consistent travel time. In other words, reliability can be defined as the on-time/punctuality performance and headway evenness. In China, only some long distance routes or routes with a long headway are designed with a schedule at each stop, while most of other routes do not have such schedule. Moreover, due to the high passenger volume, the scheduled frequencies are generally high even during off-peak hour in Beijing. Therefore, in Beijing, more attention should be paid to the headway measures. From another point of view, transit service reliability can be defined as route-based reliability and stop-based reliability. Lower-level measures (e.g., stop level) are also applicable at higher levels (i.e., the route or network level) (Kittelson et al., 2003). Route-based reliability reflects the reliability performance of a bus operation at the route level, which can be used to assess the reliability of a specific route or a whole transit network. Conversely, stop-based reliability focuses on reliability at the stop level, which can be employed to assess the reliability of a specific stop, route, or transit network. Considering both points of views, three types of bus reliability measures are proposed and discussed (Zhang, 2008). These are punctuality index based on routes (PIR), deviation index based on stops (DIS), and evenness index based on stops (EIS). The specific definitions of each reliability measure used in this paper are given in Sections 3.1–3.3. It should be noted that only fixed bus lines were considered in this study.

3.1. Punctuality index based on routes

PIR is defined as the probability that a bus can arrive at the terminals in a given time period. In regard to a Route L, its PIR is formulated in the following equation:

$$\text{PIR}_L = P(T_{\text{Arr},L} \in [T_{\text{Sch},L} + \delta_1, T_{\text{Sch},L} + \delta_2]) = P(T_{\text{Arr},L} - T_{\text{Sch},L} \in [\delta_1, \delta_2])$$

(1)

where $T_{\text{Arr},L}$ is the actual bus arrival time at terminals of Route L, $T_{\text{Sch},L}$ the scheduled bus arrival time at terminals of Route L, and $\delta_1$, $\delta_2$ are the time period factors used to determine on time.
In general, the actual departure time is considered consistent with the scheduled departure time in practices. Hence, $PIR$ can be further expressed in terms of running times. It is the possibility that a bus can run within scheduled running times between terminals, as shown in the following equation:

$$PIR_i = P\{t_{run} \in [t_{Sch} + \delta'_1, t_{Sch} + \delta'_2]\} = P\{t_{run} - t_{Sch} \in [\delta'_1, \delta'_2]\} \quad (2)$$

where $t_{run}$ is the actual running times of Route $L$ during one bus trip, $t_{Sch}$ the scheduled running times of Route $L$ during one bus trip, and $\delta'_1, \delta'_2$ are the time period thresholds used to determine on time.

In fact, $PIR$ is the possibility that the average differences between actual and scheduled running times fall in the span of $[\delta'_1, \delta'_2]$ in Eq. (2). The span of $[\delta'_1, \delta'_2]$ represents a reliable or on-time span. If the actual arrival time at the terminal is kept consistent with the scheduled arrival time at terminal, $\delta'_1 = \delta_1$ and $\delta'_2 = \delta_2$. If average differences between actual and scheduled running times fall in the span of $[\delta'_1, \delta'_2]$, the bus trip is on time. $\delta'_1$ and $\delta'_2$ are either constants or functions of scheduled running times $t_{Sch}$, which will make some difference about the reliability of the results. $PIR$ is another form of on-time performance measures, which are widely used in the transit industry. It can be used to analyze at route and network levels. In the network level assessment of bus reliability, weighted boardings should be considered.

### 3.2. Deviation index based on stops

An on-time performance indicates the level of success of the bus service remaining on the published schedule. In general, PIR can only reflect the reliability performance of routes to some extent. Under some circumstances, bus drivers in China tend to extend dwelling time at a stop to carry more passengers when they begin running along a route and attempt to catch up with the schedule due to late running on the rest of a route. In addition, transit agencies face a difficult challenge since bus operations in China are affected by other factors such as the characteristics of mixed traffic flow, high bus passenger flow during peak hours, congested roadway environment, and insufficient transit preferential treatment. Some delays caused by the transportation system cannot be controlled through strategic changes in services. Therefore, the headway deviation at the stop level is inevitable. $PIR$ is not detailed enough to describe the service variability at the stop level. For a transit service with short headways and riders arriving more randomly in relation to the schedule, reliability is better reflected by a transit agency’s ability to maintain headways and minimize a typical passenger’s waiting time. Thus, the deviation index based on stops ($DIS$) is designed to capture the operational characteristics at the stop level.

In a given time period, a bus is dispatched with a consistent headway at a terminal in normal operations. In an ideal situation, the headways between successive buses at each stop should remain consistent with each other. But the actual headway will deviate from the scheduled one in a real-world running. $DIS$ is defined as the possibility that a bus will adhere the headway between successive buses at each stop within a given time period. In regard to a stop $S$, this can be expressed mathematically as shown in the following equation:

$$DIS_S = P\{H_i - H_0 \in [\theta_1, \theta_2]\} \quad (3)$$

where $H_i$ is the headway between successive buses at stop $S$ for the same route, $H_0$ the headway between successive buses at terminals for the same route, and $\theta_1, \theta_2$ are the time period thresholds used to determine headway regularity.

In Eq. (3), $H_0$ is a reference to headway. $\theta_1$ and $\theta_2$ are also either constants or functions of $H_0$. Generally, a functional expression for $\theta_1$ and $\theta_2$ are recommended. When $\theta_1$ and $\theta_2$ are the linear function of $H_0$, and $\theta_1 = 0$ as well as $\theta_2 = k \times H_0$, $DIS$ for stop $S$ is given in the following equation:

$$DIS_S = P\left\{\frac{|H_i - H_0|}{H_0} \leq k\right\} \quad (4)$$

In Eq. (4), $\frac{|H_i - H_0|}{H_0}$ is the headway deviation ratio and $k$ is a positive number. $DIS_S$ is a stop level performance parameter for the bus reliability, which can be used to assess stop level, route level, and network level reliability. Weight based on boardings will be proposed when estimating the route level reliability $DIS_L$ and network level reliability $DIS_N$ in this paper, which represents the importance of different routes. This can be expressed mathematically as shown in Eqs. (5) and (6).

$$DIS_L = \sum_{i=1}^{N} \frac{q_{Si}}{Q_L} \cdot DIS_{Si} \quad (5)$$

where $q_{Si}$ is the boardings at the ith stop of Route $L$ in a given time period, $Q_L$ the total sum of boardings along Route $L$ in a given time period, and $N$ the total number of stops for Route $L$.

$$DIS_N = \sum_{i=1}^{M} \frac{q_{Si}}{Q_N} \cdot DIS_{Li} \quad (6)$$

where $q_{Li}$ is the boardings of the ith route in a given time period, $Q_N$ the total sum of boardings for a transit network in a given time period, and $M$ the total number of routes for a transit network.
3.3. Evenness index based on stops

From a passenger’s perspective, whether buses are running regularly is more important than whether they are actually running on schedule under the conditions of short frequency. DIS is a comparatively detailed performance parameter to describe reliability compared with PIR. However, it is still a deviation ratio performance. Although it is a useful statistical measure of the dispersion of headway points in a headway series around the mean, the coefficient of variation (CV), which represents the ratio of the standard deviation to the mean of headway, is more preferred than the relative deviation ratio. Because the standard deviation in CV squares its deviations, it tends to give more weight to larger deviations and less weight to smaller deviations compared to the deviation ratio. Moreover, CV better measures the degree of variations from one headway series to another, even if the means are drastically different from each other. Thus, evenness index based on stops (EIS) is employed to capture the pattern of consistency or evenness of the headway between vehicles in a way that allows the consideration of characteristics of routes with heavy demands in China (i.e., short headways and random arrivals).

\[ B_i = \sqrt{\frac{\sum_{j=1}^{m} (H_j - H_0)^2}{H_0^2}} \]

where \( B_i \) is the CV of the headway for stop \( i \) in a given time period, \( H_j \) the actual headway between the \( j \)th bus and the \((j - 1)\)th bus in a given time period, and \( m \) the total number of buses dispatched along a route in a given time period.

Eq. (7) shows that the lower \( B_i \) is, the more even and reliable the bus running at the stop, and vice versa. Especially, for \( B_i = 0 \), \( H_0 \) is equal to \( H_0 \), which means an ideal situation in which the bus is running with a totally even headway, providing the most reliable services. The inclusion of \( B_i \) allows comparison across routes or sections of a route, time periods and alternative reliability indicators. It should be noted that in a given time period, \( H_0 \) is a constant. Therefore, before the assessment of reliability with EIS, time periods with different \( H_0 \) should be divided and \( B_i \) in different time periods should be calculated, respectively. Because \( B_i \) is a value within the range of \([0, +\infty)\), while the reliability performance parameter should be a value in the range of \([0, 1]\), a kind of mapping is designed to transform \( B_i \) to a parameter in the range of \([0, 1]\). The transformed parameter is named fluctuation index based on stops (FIS), which reflects the unreliability at each stop. So EIS can be expressed as follows:

\[ EIS = 1 - FIS \]

Similar to DIS, EIS can be used to assess bus reliability at a stop, route, or network level. Boardings based weight is also introduced for both route and network level assessments of bus reliability, as shown in Eqs. (9)–(12).

\[ FIS = \sum_{i=1}^{N} \frac{q_{li}}{Q_L} \cdot FIS_{si} \]

\[ EIS = 1 - FIS \]

\[ FIS = \sum_{i=1}^{M} \frac{q_{li}}{Q_N} \cdot FIS_{li} \]

\[ EIS = 1 - FIS \]

where \( FIS \) is the fluctuation index based on stops for Route \( L \), \( FIS \) the fluctuation index based on stops for a transit network.

3.4. Relationship among PIR, DIS, and EIS

Actually, PIR is used to measure terminal on-time performance which is being used in most cities in China including Beijing. In Beijing, DIS and EIS will be introduced to Beijing Municipal Committee of Communications to help them supervise the bus service quality in the near future. Compared with PIR, DIS and EIS are microscopic performance parameters to bus reliability assessment, as shown in Table 1. These three performance parameters allow a realistic representation of bus

<table>
<thead>
<tr>
<th>Performance parameter</th>
<th>Bus reliability assessment method</th>
<th>Level</th>
<th>Required data</th>
<th>Assessment objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIR</td>
<td>Punctuality assessment based on routes</td>
<td>Route based</td>
<td>Running time between terminals</td>
<td>Route and network level assessment</td>
</tr>
<tr>
<td>DIS</td>
<td>Deviation assessment based on stops</td>
<td>Stop based</td>
<td>Headway deviation at each stop; boardings at each stop</td>
<td>Stop, route and network level assessment</td>
</tr>
<tr>
<td>EIS</td>
<td>Evenness assessment based on stops</td>
<td>Stop based</td>
<td>CV of headway at each stop; boardings at each stop</td>
<td>Stop, route and network level assessment</td>
</tr>
</tbody>
</table>
operations from different levels. Transit companies pay more attention to PIR and transit customers are more interested in DIS and EIS. In essence, PIR are punctuality based performance parameters, whereas EIS and DIS is the deviation or evenness based parameter. They present a progressive pattern from macro-scale to micro-scale, as shown in the following equation:

\[
P_{\text{IR}} = \text{Considering deviation ratio of headway based reliability at each stop} \quad \text{DIS = Considering CV of headway based reliability at each stop} \quad \text{EIS}
\]

(13)

The following section analyzes the relationship between PIR, DIS, and EIS based on a numerical example. There are five stops serving a given Route L in one direction. During the peak hour, its headway is 6 min and the scheduled running time between two terminals is 20 min. In Table 2, actual arrival times from 7:00 to 7:59 are given. Except for Terminal 5, boardings at each stop are equal with each other. It is assumed that the stop spacing is an equal value.

It is assumed that \( \delta_1 \) and \( \delta_2 \) for PIR are designated as positive and negative 10% percent of the scheduled running time. Then PIR, DIS, and EIS will be calculated in this numerical example.

1. **PIR:**

\[ [\delta_1, \delta_2] = (0 \pm 10\%) \times t_{\text{Sch}}, \text{ then } \delta_2 = -\delta_1 = 0.1 \times t_{\text{Sch}} = 2. \]

Herein, a bus that has an actual running time within the range from 18 to 22 min is reliable. That is to say, Buses 4, 5, 6, 7, 9, and 10 provide reliable services. For Route L from 7:00 to 7:59, \( P_{\text{IR}} = P\{t_{\text{run}} - t_{\text{Sch}} \leq 2\} = 6/10 = 0.6 \).

2. **DIS:**

Suppose the length of Route L is \( R \). \( [\delta_1, \delta_2] \) represents time period thresholds used to determine on time when the bus runs a distance of \( R \). Then, when the bus runs a distance of \( x \) ( \( x \) is the distance to the origin terminal), the time period thresholds used to determine on time are expressed as \( [\delta_1, \delta_2] \). Similarly, for \( j \)th stop ( \( j \geq 2 \)), the time period thresholds used to determine on time can be expressed as \( [\sum_{i=2}^{j-1} \frac{t_i \delta_1}{R}, \sum_{i=2}^{j-1} \frac{t_i \delta_2}{R}] \). Herein, under reliable bus services, the actual running time between two consecutive buses should fall within \( \left[ \max \left(0, H_0 - \sum_{i=2}^{j-1} \frac{t_i \delta_1}{R}\right), \min \left(H_0 + \sum_{i=2}^{j-1} \frac{t_i \delta_2}{R}, 2H_0\right) \right] \) when the two consecutive buses arrive at stop \( j \). Hence, the Eq. (14) can be used to determine \( \text{DIS}_N \) under a condition that on-time thresholds is \( [\delta_1, \delta_2] \).

\[
\text{DIS}_N = P\left\{H_i \leq \left[ \max \left(0, H_0 - \sum_{i=2}^{j-1} \frac{t_i \delta_1}{R}\right), \min \left(H_0 + \sum_{i=2}^{j-1} \frac{t_i \delta_2}{R}, 2H_0\right) \right] \right\}
\]

(14)

If \( \sum_{i=2}^{j-1} \frac{t_i \delta_1}{R} \) and \( \sum_{i=2}^{j-1} \frac{t_i \delta_2}{R} \) are decimal, its value should be rounded up to an integer. With Eq. (14), connection between PIR and DIS are established.

Because the actual arrival time at the terminal is equal to the scheduled one at the terminal, \( \delta_1 = \delta_1 = -2 \) and \( \delta_2 = \delta_2 = 2. \)

According to Eq. (14), for Terminal 1, Stop 2, Stop 3, Stop 4, and Terminal 5, the reliable thresholds for \( H \) are \( [6, 6], [5, 7], [5, 7], [4, 8] \) and \( [4, 8] \), respectively. According to these thresholds, there are, respectively, 9, 4, 3, 5, and 4 arrivals for Terminal 1, Stop 2, Stop 3, Stop 4, and Terminal 5, which are reliable. Then \( \text{DIS} \) at Terminal 1, Stop 2, Stop 3, Stop 4, and Terminal 5 are \( 1, 0.444, 0.333, 0.556, \) and \( 0.444 \). Boardings based weight is considered as shown in Eq. (5), hence for Route L, \( \text{DIS}_L = 0.583 \).

3. **EIS:**

The network level reliability \( \text{DIS}_N \) can be calculated using \( \text{DIS}_L \) and boardings based weight. Among all \( N \) stops in a transit network, the possibility of reliability is \( \text{DIS}_N \); that is, a total of \( N \times \text{DIS}_N \) stops are reliable and the remaining stops are unreliable. Then the mapping is conducted to transform \( B_t \) to the fluctuation index based on stops (FIS3), which is in the range of \( [0, 1] \), as presented in the following two steps:
Step 1: all $B_i$ in the transit network are sorted in ascending order as $B_{1,1}$, $B_{1,2}$, $B_{1,3}$, ..., $B_{k,N}$.

Step 2: we design $[0, 1]$ mapping. For a reliable stop, a linear mapping is designed to transform $B_i$ to $FIS_i$, while $FIS_i$ is designated as 1 for the unreliable stop, as presented in the following equation:

$$FIS_i = \begin{cases} 
\frac{B_{i,n}}{B_{i,N \cdot PIS_N}}, & n < N \cdot PIS_N \\
1, & n \geq N \cdot PIS_N 
\end{cases} \quad (15)$$

Then $EIS$ can be calculated according to Eqs. (9)–(12). It should be noted that such mapping makes $DIS$ and $EIS$ applicable in a framework with similar metrics.

According to Eq. (15), $FIS_i$ for Terminal 1, Stop 2, Stop 3, Stop 4, and Terminal 5 are 0, 0.669, 1, 1, and 1. Correspondingly, $EIS_i$ for Terminal 1, Stop 2, Stop 3, Stop 4, and Terminal 5 are 1, 0.331, 0, 0, and 0. Considering the boardings based weight, route level $FIS_i$ and $EIS_i$ are 0.667 and 0.333, respectively.

In summary, these three performance parameters are analyzed under a framework with similar metrics. $PIR_i = 0.6$, $DIS_i = 0.583$, and $EIS_i = 0.333$ under such a framework. From this numerical example, the progressive pattern from a macro-scale to a micro-scale of $PIR$, $DIS$, and $EIS$ are proved again. Compared with $PIR$ and $DIS$, $EIS$ is more sensitive to any unevenness of the headway.

4. Evaluation, results, and analysis

4.1. The surveyed transit network by sampling technique

The bus transit network in Beijing is one of the largest and most comprehensive in China. The whole system is managed by two city public transit companies, Beijing Public Transport Holdings, Ltd. and Beijing Xianglong Bus CO, Ltd. The bus transit system in Beijing is characterized by its comparatively short headway and fixed schedules for most of the bus lines. In 2007, more than 4.2 billion passengers were carried on nearly 600 different routes. Buses have amounted to nearly 26,000. Annual vehicle travel kilometers by buses in Beijing are now higher than at any time since 2001 (1.67 billion kilometers in 2007). Considering that Beijing is plagued with increasingly serious transport problems, the Beijing city government is attempting to increase the share of total transit trips by the provision of more reliable and convenient bus services.

By analyzing the bus operational characteristics of Beijing’s transit system, several principles with high priority were used for sampling the bus network were determined. The sample selection is based on route location and layout, road classification along the route, the proportion of routes with different fare structures the proportion of routes with vs. without air conditioning, and the proportion of routes from different public transit companies.

(1) The selected routes should be within the Fifth Ring Road of Beijing, which are distributed as uniformly as possible without overlaps.

(2) The major routes along existing ring road expressways and arterials should be encompassed. A small proportion of routes along the local roads are also considered.

(3) The proportions of selected routes with flat fares vs. routes with distance-based fares, and routes with air conditioning vs. routes without air conditioning are equal to the real proportions.

(4) The sampled routes of each sub-company of the two city public transit companies should have approximately similar sampling proportions.

According to these four principles and surveyors used in this study, the sampling proportion for bus routes equaled 5% of total bus routes. Thirty bus routes, six of which were circular lines, were selected to be surveyed, as shown in Fig. 1. In these 30 bus routes, one of them is the Bus Rapid Transit (BRT) Line 1, along which BRT vehicle is running on exclusive bus lanes with barriers and the transit signal priority is provided. In the following section, Bus Rapid Transit (BRT) Line 1 is only used as a contrast analysis because it has different preferential treatments to provide better reliability. The sampling proportion for the scheduled bus for the 30 bus routes amounted to 10%; the bus that ran during the peak hour was close to 50%. By analyzing headways by direction and time of day (i.e., morning peak hours, off-peak hours at noon, evening peak hours) for the 30 sampled bus routes, the authors found that 75.3% of the headways are less than 10 min and 94% of the headways are less than 15 min. 4% of the headways are equal to 15 min. Only 2% of headways are more than 15 min but less than 20 min. A total of 396 surveyors were arranged to go aboard the sampled bus on each day from November 20 to 22, 2006 to manually record the bus arrival time at each stop, boarding numbers at each stop, departure time at each stop, stop time at intersections, stop time on roadway segment, and vehicle stops. Summarizing the data from the survey enabled the generation of a reliable sample ($N = 70,723$) at stop level to be used in a statistical analysis of different performance parameters on bus reliability, while controlling for the variations introduced by the differences in patterns.

The transit on-board surveys of the 30 lines were conducted on November 20–22, 2006 (from Tuesday to Thursday), which covered PM peak, AM peak, and off-peak times. During this period of time, no major weather issues were present (i.e., snow storms) that might have had an effect on travel time and headway adherence. Weekend and holiday services are not included in the analysis. The stop to board by surveyor and the number of surveyors along a route was determined according to the headway, route length, and location of a terminal or depot.
4.2. Service reliability analysis at different levels

This section presents information about service reliability at stop, route, and network levels based on the case study of the bus transit system in Beijing. It should be noted that $d_0^1$ and $d_0^2$ for PIR were designated as positive and negative 10% percent of the scheduled running time.

4.2.1. Stop level

Bus stops are major components of both bus routes and a bus network. Bus reliability assessments at route or network level are based on stop-level analysis to a great extent. Stop-level reliability can directly affect the waiting time of passengers. Passengers value the waiting time for urban transit almost twice as much as in-vehicle time (Oort and Nes, 2003). Therefore, bus reliability at the stop level is from a passenger’s point of view, which can be used to identify a decline or improvement in the reliability from a passenger’s perspective.

In this paper, DIS and EIS were used to measure stop-level bus reliability. The statistical results are shown in Table 3. From the analysis in Table 3, we concluded that the mean, standard deviation, maximum, minimum, as well as 50% percentile of DIS are less than those of EIS.

4.2.2. Route level

PIR, DIS, and EIS can be employed to the route-level analysis on bus reliability assessment. The results are illustrated in Table 4 and Fig. 2.
As shown in Table 4 and Fig. 2, for most of the sampled routes, \( \text{PIR} > \text{PIS} > \text{EIS} \), with only a few exceptions, such as Routes 45, 125, and 452. The reason for these exceptions was that the bus running time along these routes was unreliable due to the impacts of traffic flows on the roadway. However, successive buses ran at comparatively even headways. The PIR of seven
4.3.1. The effects of route length on service reliability

Adjustments and long-term planning to improve reliability can be provided to bus operators. The long headway results in the reduction of reliability. The significant decline in performance happens in the up to 30 km length range actually. This proved that a route that exceeds 30 km yielded a more unreliable service. It should be noted that the fluctuation may be caused by the difference of roadway conditions for different routes with different lengths. The lengthening of headway resulted in the decrease of \( \text{EIS} \). A relatively large headway deviation may occur for the route with a long headway if other operational conditions are similar. Moreover, the same headway deviation at a stop will cause a bigger impact on the reliability of the whole route.

4.3.2. The effects of headway on service reliability

Collecting information about any underlying factors causing bus unreliability, particularly those that are controllable by transit providers, will help bus operators improve the situation. There has been considerable research on the underlying causes of unreliable services. The route length is found to be the most important variable affecting bus reliability, followed by service frequency or headway and provision of bus lanes (Sterman and Schofer, 1976; Abkowitz and Engelstein, 1983, 1984; Strathman et al., 1999). In order to better identify the threshold for route length, the distance from a stop to the origin terminal was further identified as one of the key factors that impact service reliability in this paper. Thus, route length, headway, distance from a stop to origin terminal, and exclusive bus lanes were selected to estimate their effects on service reliability in this research. These four factors are controllable by transit providers. Therefore, decision making on short-term adjustments and long-term planning to improve reliability can be provided to bus operators.

4.2.3. Network level

\( \text{PIR} \), \( \text{DIS} \), and \( \text{EIS} \) were all used to assess bus reliability at the network level, as presented in Table 5.

This result shows the relationship among \( \text{PIR} \), \( \text{DIS} \), and \( \text{EIS} \). These three performance parameters assess reliability from the macroscopic to the microscopic scale. Moreover, the results indicate that there was low service reliability for the bus transit network in Beijing in the existing situation. According to TCQSM, the reliability LOS for the bus network in Beijing is estimated “F” when positive and negative 10% of the scheduled running time was adopted as the on-time thresholds. The results may change when different on-time thresholds are selected. Table 6 presents \( \text{PIR} \) under different on-time thresholds. The results in Table 6 indicate that \( \text{PIR} \) was improved when on-time thresholds were increased. When ±15 of scheduled running time was adopted, LOS “E” could be determined. However, a low reliability LOS for the bus transit network in Beijing is shown in general.

Table 5
Reliabilities at network level.

<table>
<thead>
<tr>
<th>PIR</th>
<th>DIS</th>
<th>EIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.681</td>
<td>0.499</td>
<td>0.311</td>
</tr>
</tbody>
</table>

Table 6
PIR based on different LOS thresholds.

<table>
<thead>
<tr>
<th>On-time thresholds</th>
<th>±5% of scheduled running time</th>
<th>±10% of scheduled running time</th>
<th>±15% of scheduled running time</th>
<th>±20% of scheduled running time</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIR 0.401</td>
<td>0.510</td>
<td>0.681</td>
<td>0.761</td>
<td>0.814</td>
</tr>
</tbody>
</table>

4.3. Estimation of effects of different factors on service reliability

Collecting information about any underlying factors causing bus unreliability, particularly those that are controllable by transit providers, will help bus operators improve the situation. There has been considerable research on the underlying causes of unreliable services. The route length is found to be the most important variable affecting bus reliability, followed by service frequency or headway and provision of bus lanes (Sterman and Schofer, 1976; Abkowitz and Engelstein, 1983, 1984; Strathman et al., 1999). In order to better identify the threshold for route length, the distance from a stop to the origin terminal was further identified as one of the key factors that impact service reliability in this paper. Thus, route length, headway, distance from a stop to origin terminal, and exclusive bus lanes were selected to estimate their effects on service reliability in this research. These four factors are controllable by transit providers. Therefore, decision making on short-term adjustments and long-term planning to improve reliability can be provided to bus operators.

4.3.1. The effects of route length on service reliability

Fig. 3 illustrates the effects of route length on the three different performance parameters on reliability (i.e., \( \text{PIR} \), \( \text{DIS} \), and \( \text{EIS} \)). All of \( \text{PIR} \), \( \text{DIS} \), and \( \text{EIS} \) decreased with the increase of route length although a fluctuation is present, which means that a long route results in the reduction of reliability. The significant decline in performance happens in the up to 30 km length range actually. This proved that a route that exceeds 30 km yielded a more unreliable service. It should be noted that the fluctuation may be caused by the difference of roadway conditions for different routes with different lengths.

4.3.2. The effects of headway on service reliability

Bus service reliability under different headways is exhibited in Fig. 4. The relationships between the headway and \( \text{PIR/DIS/EIS} \) are different because different curves are presented. The headway is shown to have no obvious relationship with \( \text{PIR} \) because \( \text{PIR} \) is the running time-based performance parameter. As headway-based performance parameters, \( \text{DIS} \) and \( \text{EIS} \) are strongly related to headway. With lengthening headway, \( \text{DIS} \) increased because its thresholds increased with a longer headway. The lengthening of headway resulted in the decrease of \( \text{EIS} \). A relatively large headway deviation may occur for the route with a long headway if other operational conditions are similar. Moreover, the same headway deviation at a stop will
create more significant effects to routes with short headways than to routes with long headways. Thus, EIS-based reliability has worsened.

4.3.3. The effects of distance between a stop and the origin terminal on service reliability

Fig. 5 provides DIS and EIS that were obtained by varying the distance from the stop to the origin terminal. In general, service reliability declined with the increase of the distance from the stop to the origin terminal. This was caused by the fact that unreliability is cumulated due to the effect of the headway deviation at the previous location. However, when 30 km were exceeded, DIS began to fluctuate remarkably, which means that more random factors that impact reliability may exist when the distance from a stop to the origin terminal is longer than 30 km. Correspondingly, a distance between a stop and the origin terminal that exceeds 30 km yielded very low EIS. These results further show that the route length should be kept to a limited span in transit planning practices.

4.3.4. The effects of exclusive bus lanes on service reliability

Among the sampling routes, route segments with vs. without an exclusive bus lane were divided to analyze their bus reliabilities. The results are presented in Table 7. Bus service reliability on a route segment with an exclusive bus lane was generally better than reliability on a route segment without an exclusive bus lane. These results reflect the provision of an
exclusive bus lane plays a significant role in improving service reliability. With regard to time of day, the improvement of service reliability during both morning and evening peak hours was better than reliability during off-peak hours. Such results can be explained by the fact that exclusive bus lanes in Beijing only operate during peak hours (i.e., 7:00–9:00, 16:00–19:00), except for the exclusive bus lane on Chang’an Street. During off-peak hours, most buses are mixed with the general traffic flow.

In order to further estimate the effect of bus preferential treatments on bus service reliability, the Bus Rapid Transit (BRT) Line 1 was analyzed specifically because a large portion of the line was designated by exclusive bus lanes with barriers and the transit signal priority was provided at intersections. As shown in Table 8, its service reliability was higher than other route segments with exclusive bus lanes, and especially higher than other route segments without exclusive bus lanes.

### 5. Conclusions

Bus service reliability is one of the key factors in the acceptability of bus services because “arriving when planned” is one of the most important desires of transit riders. The assessment of transit system reliability and efficiency has attracted considerable attention in the literature because the vast array of useful and important policy results can be extracted from such an analysis. In this paper, we use three performance parameters on bus service reliability, employing sampling survey tech-

### Table 7

Reliabilities for route network with vs. without bus lanes.

<table>
<thead>
<tr>
<th></th>
<th>DIS</th>
<th></th>
<th>EIS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without exclusive bus lane</td>
<td>With exclusive bus lane</td>
<td>%</td>
<td>Without exclusive bus lane</td>
</tr>
<tr>
<td>Morning peak hour</td>
<td>0.412</td>
<td>0.461</td>
<td>11.78</td>
<td>0.283</td>
</tr>
<tr>
<td>Off-peak hour</td>
<td>0.533</td>
<td>0.567</td>
<td>6.29</td>
<td>0.406</td>
</tr>
<tr>
<td>Evening peak hour</td>
<td>0.385</td>
<td>0.453</td>
<td>17.59</td>
<td>0.226</td>
</tr>
</tbody>
</table>

### Table 8

Reliability of BRT line with preferential treatments.

<table>
<thead>
<tr>
<th></th>
<th>DIS</th>
<th></th>
<th>EIS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRT line 1</td>
<td>Without exclusive bus lane</td>
<td>%</td>
<td>BRT line 1</td>
</tr>
<tr>
<td>Morning peak hour</td>
<td>0.571</td>
<td>0.412</td>
<td>38.52</td>
<td>0.530</td>
</tr>
<tr>
<td>Off-peak hour</td>
<td>0.599</td>
<td>0.533</td>
<td>12.26</td>
<td>0.572</td>
</tr>
<tr>
<td>Evening peak hour</td>
<td>0.506</td>
<td>0.385</td>
<td>31.17</td>
<td>0.381</td>
</tr>
</tbody>
</table>
niques to estimate service reliability at the stop, route, and network levels. We considered the effects of route length, headway, distance from a stop to the origin terminal, and the use of exclusive bus lanes on service reliability.

Using the bus transit system in Beijing as a case study and considering the operational characteristics of China, the research shows that $\text{PIR} > \text{DIS} > \text{EIS}$ can be used to effectively assess bus service reliability at the stop, route, and network levels. These three performance parameters allow the reliability estimation from the macro-scale to the micro-scale under a framework with similar metrics. In general, it was found that $\text{PIR} > \text{DIS} > \text{EIS}$. In terms of Beijing, the results imply that its transit system is operated at a relatively low reliability level.

In addition, the effects of four factors on service reliability were analyzed. First, with the increase of the route length, bus services were found to be unreliable for all performance parameters of $\text{PIR}$, $\text{DIS}$, and $\text{EIS}$. Second, an obvious relationship between headway and $\text{PIR}$ was not discovered. However, $\text{DIS}$ merely increased and $\text{EIS}$ declined as the headway was gradually increased. Third, changing the distance between a bus stop and the origin terminal in the range from 0 to 30 km caused bus service reliability to worsen because $\text{DIS}$ and $\text{EIS}$ were decreased with a slight fluctuation. Further, bus service reliability was uneven as the distance from the stop to the origin terminal exceeded 30 km. In this case, $\text{DIS}$ began to fluctuate remarkably and $\text{EIS}$ tended to be very low. Fourth, the provision of an exclusive bus lane can effectively enhance bus service quality with improved reliability. As a result, it is recommended that existing long bus lines be split into two or more separate sections. A length less than 30 km for a bus route is suitable to improve service reliability. This recommendation is adopted by the public transit companies at present. Meanwhile, bus preferential treatments, such as providing more exclusive bus lanes, extending bus lane hours, and implementing transit signal priority on more routes, are recommended for Beijing.

The analysis presented here is limited to a set of data actually collected by surveyors over a 3-day test period. Although limited, this analysis shows a potential of extracting a great amount of information about bus operations if a large and comprehensive data set including AVL and IC card data can be made available in Beijing in the future. The study attempts to develop a methodology for analyzing other routes facing similar problems in China. Future studies should consider using more extensive data including the AVL and IC card data to test the effects of additional contributing factors on bus service reliability, thereby offering more practical strategies to monitor and improve the quality of bus service.

Acknowledgements

The authors acknowledge that this paper was prepared based on the National High Technology Research and Development Program of China (863 Program) (2008AA11Z202), the Project of Foundation of Beijing Jiaotong University (2007XM021), and the Project (T0710620 and CZ200704) funded by Beijing Transportation Research Center. The authors would also like to thank Shawn Turner from Texas Transportation Institute for supporting their efforts for this paper. The MOE Key Laboratory for Transportation Complex Systems Theory and Technology of Beijing Jiaotong University also supports the research in this paper.

References