Finding Most Reliable Path With Extended Shifted Lognormal Distribution

ZHENZHEN YANG[1,3], **ZIYOU GAO**[1], **HUIJUN SUN**[1], **FENG LIU**[2], **AND JIANDONG ZHAO**[1]

[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[2]Transportation Research Institute, Hasselt University, B-3590 Diepenbeek, Belgium
[3]Beijing PalmGo Infotech Co., Ltd., Beijing 100085, China

Corresponding author: Ziyou Gao (zygao@bjtu.edu.cn)

ABSTRACT Travel time uncertainty may cause late arrival and impose a high penalty on travelers. There is a growing interest in modeling travel time uncertainty to optimize the reliability of travel time at the path and network level. Real data analysis finds that the influence factors, including day-of-week, holidays, time-of-day, road grades, traffic states, and so on, often reduce the cumulative probability of travel time even in the same facility type (the same lane number and the same divided type). Thus, a novel aggregate approach is proposed to classify the travel time data based on these influence factors. The distribution with the new aggregate approach is defined as the extended shifted lognormal (ESLN) distribution. KS test indicates that the ESLN distribution can effectively describe travel time, and outperforms normal, lognormal, gamma, and beta distribution. Travel time correlations are calculated between new aggregate groups, which can effectively reduce the complexity compared with the link to link correlations. Finally, the ESLN distribution is used to find the most reliable path in a real-world large-scale network. The comparison results between ESLN distribution and shifted lognormal (SLN) distribution show the effectiveness and improvement of the proposed method in finding the most reliable path.

INDEX TERMS Travel time, most reliable path, extended shifted lognormal distribution.

I. INTRODUCTION

Advanced traveler information system has applied widely to alleviate traffic congestion. In congestion road network, link travel time is highly stochastic due to supply degradation and demand fluctuation [1], [2], [3], which may cause late arrival and impose a high penalty on travelers, such as missed flight and lost business [4], [5]. To ensure punctual arrival and avoid a penalty, travelers not only need the shortest path in a stochastic network but also pursue a reliable path within a given travel time budget. Consequently, quantifying, modeling, and optimizing travel time reliability is a growing interest and necessary research topic.

Quantitative measures for travel time reliability include statistic range methods [6], [7], [8], buffer time measures [3], [7], tardy-trip measures [7], and probabilistic measures [2], [4], [8], [9], [10], [11], [12]. In this study, probabilistic measures are adopted to capture travelers' different risk attitudes towards travel time uncertainty and maximize the on-time arrival probability within a given travel time budget. Travel time distribution is the foundation of probabilistic

measures, which has been proved as normal distribution [11], [13], [14], lognormal (LN) distribution [15], [16], shifted lognormal (SLN) distribution [2], Gamma distribution [17], [18], Gamma-Gamma distribution [19], Burr distribution [20], beta distribution [2], Weibull distribution [21], inverse Gaussian distribution [22], generalized extreme value distribution, and generalized Pareto distribution [23], etc. Among these distributions, normal distribution is convenient for analysis and calculation. However, the non-zero probability of negative travel time of normal distribution is unreasonable. LN distribution may cause unreasonable free-flow speed for its bounded below by zero. SLN distribution is proposed by Srinivasan *et al.* [2] to solve these shortcomings. However, SLN distribution only describes travel time distribution in the same facility type (the same lane number and the same divided type). Some influence factors, such as day-of-week, holidays, time-of-day, road grades, traffic states, etc., may cause the diversity of travel time and reduce the cumulative probability of travel time even in the same facility type. Thus, this paper proposes a novel aggregate approach of travel time

to improve the precision of travel time distribution and travel time reliability by isolating the influence factors.

Travel time correlation is an important parameter in finding the most reliable path. Ignoring correlation will result in substantial errors in estimating the benefits of projects that are expected to result in an improvement in reliability [24]. Factors affecting the correlation coefficient include spatial distance [25], traffic congestion [5], [27], speed limits, and turn intersection delays [28], etc. To solve the complexity of correlations between links, some techniques and approaches have been proposed, including nonparametric regression technique based on Bayesian natural cubic spines [27], regression model [25], Cholesky decomposition [5], Lagrangian relaxation approach [26]. Specifically, covariance matrix-based method requires a large amount of memory to store link to link correlation values, which is difficult to embed into standard shortest path algorithms directly [26]. To reduce the complexity of link travel time correlations efficiently, Srinivasan *et al.* [2] propose correlations between aggregate groups in the same facility type.

Some algorithms for finding the most reliable path have been explored in recent decades. The algorithms for solving the least expected travel time problem include the branch and bound solution procedure [29], the heuristic algorithm based on the K-shortest path algorithm [13], label-correcting algorithm [30], [31], [32], the multi-criteria A∗ algorithm based on the stochastic first-in-first-out property [4], [33], the multi-objective 0-1 optimization model and a tabu search algorithm [34]. Another category is the stochastic arrival on schedule and the reliable priori shortest path model. The optimal path has two definitions: the maximum arrival reliability within a given time budget [12], [35], [36] and the minimum travel time budget for a specified reliability [10]. For the non-linear and non-additive routing algorithm, the efficient methods include the non-dominance-based method [1], [11], the Lagrangian relaxation approach [26], [37], [38], [39], [40], and the simulation-based method [41], [42], [43]. The simulation-based method is expensive in computation and the maximum simulation number decides its precision. Srinivasan *et al.* [2] propose an algorithm to find the most reliable path with an approximated SLN distribution and a general correlation structure, which is computationally less expensive than traditional Monte-Carlo estimation techniques with an acceptable compromise on accuracy.

Considering the above literature review, the contributions of this paper to the growing body of knowledge in travel time reliability are as follows: A novel aggregate approach is proposed to classify travel time based on day-of-week, holidays, time-of-day, road grades, and traffic states. The distribution with the new aggregate approach is defined as the extended shifted lognormal (ESLN) distribution. ESLN distribution is used to find the most reliable path in a real-world large-scale network.

The remainder of this paper is structured as follows. Section II states the problem. Section III proposes a new aggregate approach of travel time, and uses the ESLN distribution to quantify travel time at link and path levels. In Section IV, ESLN distribution is used to find the most reliable path in a real-world large-scale network. Finally, the salient findings, conclusions, and directions for further research are outlined in Section V.

II. PROBLEM STATEMENT

SLN distribution has been proposed to describe travel time distribution in the same facility type (the same lane number and the same divided type), and indicates that SLN distribution outperforms normal distribution and LN distribution for almost all facility types [2]. However, day-of-week, holidays, time-of-day, road grades and traffic states, etc., may cause the diversity of travel time and reduce the cumulative probability of travel time. To explain the influence of these factors, real GPS data were collected by floating cars in Beijing from November 27, 2017, to January 1, 2018. Data contain the date, time, car ID, longitude, latitude, and speed. GPS data were associated with links with a map matching technique. Link travel time was calculated by the matched GPS data. As travel time increases with the increase in traffic congestion, we used congestion mileage to explain the influence factors. Fig. 1 (a) shows that congestion mileage varied greatly in different days of week from Monday to Sunday. Congestion mileage on holiday (e.g., New Year's Day) was also different from that of the other days. In addition, congestion mileage varied greatly in the peak hours and non-peak hours.

Road grades (RG) include six levels: expressway (RG=1), freeway (RG=2), arterial road and national road (RG=3), secondary road and provincial road (RG=4), branch road (RG=5), and others (RG=6). Fig. 1 (b) shows that the cumulative probability of all data was lower than that of RG =1-4, and was larger than that of RG=6. The cumulative probability of travel time per unit length varied greatly in different road grades. Moreover, each road grade had different lane numbers (see Fig. 1 (c)).

If we only aggerate travel time data with lane number and divided type, the cumulative probability will be higher or lower than the actual value. Thus, we need to aggregate travel time data by considering day-of-week, holidays, time-of-day, road grades, and traffic states to improve the accuracy of the cumulative probability of travel time and further improve the accuracy in finding the most reliable path.

III. QUANTIFICATION OF TRAVEL TIME DISTRIBUTION
A. NEW AGGREGATE APPROACH OF TRAVEL TIME
1) NEW AGGREGATE APPROACH
The transportation network is modeled as a directed graph $G(N, A)$, where $N = \{1, 2, \cdots, n\}$ represents the set of nodes and A denotes the set of directed arcs. Travel time varies greatly around upstream and downstream at some special road segments, such as traffic light, on/off ramp, toll station, service area, etc. Thus, these road segments are divided into upstream, middle position, and downstream (see Fig. 2).
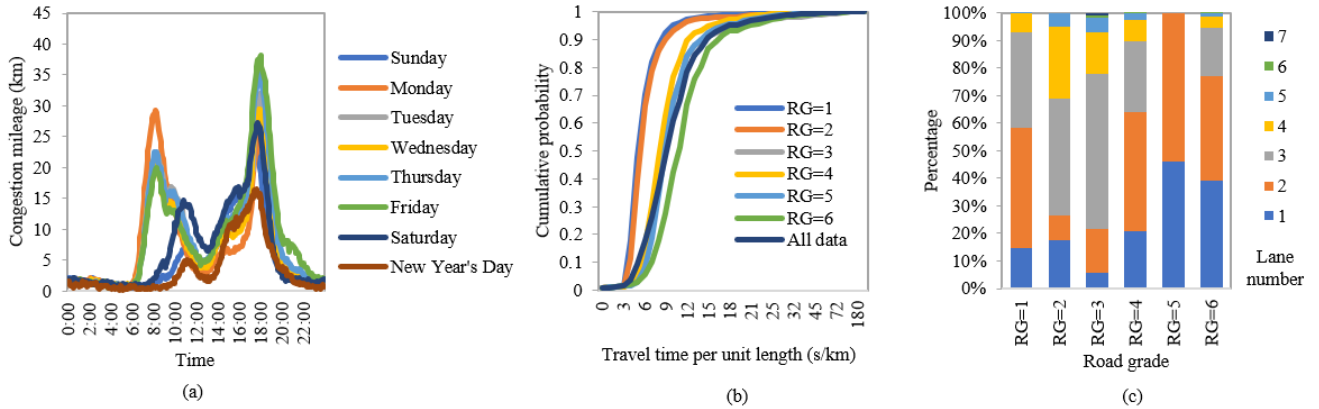
FIGURE 1. Problem statement: (a) the influence of day-of-week, holidays, time-of-day; (b) the influence of road grades (RG); (c) the percentage of lane number with different road grades.
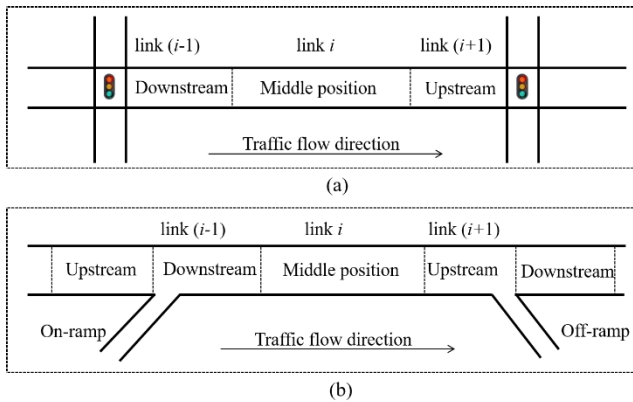

FIGURE 2. Schematic diagram of road segments: (a) signalized intersection; (b) ramp.

Data are first divided into eight groups: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, Holidays. Further, data on holidays are divided into Spring Festival, Qingming Festival, May Day, Dragon Boat Festival, National Day, Mid-Autumn Festival, New Year's Day.

Then, data are divided into morning peak period, evening peak period, non-peak period, and night period based on the statistical characteristics of traffic flow. For example, the periods on working days in Beijing are divided as follows: morning peak period (7:00-9:00), evening peak period (17:00-19:00), non-peak period (9:00-17:00, 19:00-22:00), and night period (22:00-7:00).

Finally, data are aggregated based on road grades and traffic states. Road grades are characterized by free-flow travel time. Traffic states include very smooth traffic, smooth traffic, light congestion, moderate congestion, and severe congestion. Traffic index is introduced to quantify traffic states. The detailed classification method based on free-flow travel time and traffic index is as follows.

Define γ_i as the free-flow travel time. Set e_i as the excess travel time. Travel time per unit length on link t_i is described by the following formula:

$$t_i = \gamma_i + e_i \tag{1}$$

Define β_i as the ratio between the excess travel time and the free-flow travel time:

$$\beta_i = \frac{e_i}{\gamma_i} \tag{2}$$

Set α as a conversion factor. The meaning of α is that travelers need to spend α times more travel time compared with free-flow travel time. Define I_i as the traffic index of link i. Thus,

$$I_i = \frac{\beta_i}{\alpha} = \frac{e_i}{\alpha \cdot \gamma_i} \tag{3}$$

$$e_i = \alpha I_i \gamma_i \tag{4}$$

t_i can be described as the function of γ_i and I_i:

$$t_i = \gamma_i + e_i = \gamma_i + \alpha I_i \gamma_i = \gamma_i(1 + \alpha I_i) \tag{5}$$

Thus, an aggregate approach of travel time is proposed based on γ_i and I_i. Travel time data with the same value of γ_i and I_i are aggregated into the same group. To reduce the complexity, γ_i and I_i are converted into discrete variables. Define $\overline{\gamma_i}$ and $\overline{I_i}$ as the discrete value of γ_i and I_i. Thus, travel time data with the same value of $\overline{\gamma_i}$ and $\overline{I_i}$ are aggregated into the same group (see Fig. 3(a)).

Free-flow travel time differs greatly among different road grades. Usually, free flow is measured by speed. For example, the maximum limited speed of the expressway in Beijing is 120 km/h or 100 km/h; but the maximum limited speed of the freeway is 80 km/h or 60 km/h. Thus, speed interval is used as the discrete interval of γ_i. Define $v_{f,i}$ (km/h) as the free-flow speed of link i, $v_{f,i} = \frac{L_i}{\gamma_i L_i / 3600} = \frac{3600}{\gamma_i}$, where L_i (km) is the length of link i. Set Δ (km/h) as the speed interval. The discretized free-flow travel time $\overline{\gamma_i}$ is calculated by:

$$\overline{\gamma_i} = \left\lceil \frac{3600}{\gamma_i \Delta} \right\rceil \tag{6}$$

where $\lceil \cdot \rceil$ is the function that rounds the element to the nearest integer in the direction of positive infinity. For example, if $\Delta = 20km/h$ and $\gamma_i = 40s/km$, then $\overline{\gamma_i} = \left\lceil \frac{3600}{40 \times 20} \right\rceil = 5$.

Define N_0 as the number of $\overline{I_i}$. For example, the default value of N_0 used by Beijing Transport Institute is 10. $\overline{I_i}$ is calculated by:

$$\overline{I_i} = \begin{cases} 1, & I_i = 0 \\ \lceil I_i \rceil, & 0 < I_i \leq N_0 - 1 \\ N_0, & N_0 - 1 < I_i \leq I_{max} \end{cases} \quad (7)$$

$\overline{I_i} = N_0$ covers the value of I_i from $N_0 - 1$ to I_{max}, where $I_{max} = \frac{e_{max}}{\alpha \gamma_{min}}$. If $e_{max} = 3600s/km$, $\gamma_{min} = 18s/km$, $\alpha = 0.15$, $N_0 = 10$, then $I_{max} = \frac{e_{max}}{\alpha \gamma_{min}} = 1333$, $9 < I_i \leq 1333$, which means $\overline{I_i} = 10$ covers I_i from 9 to 1333. The range is too wide. The relationship between t_i and I_i is non-linear (see Fig. 3(b)), which approximately follows the power law distribution. After logarithmic transform, the value of $\log 2(I_i)$ varies from 3.17 to 10.38 (see Fig. 3(c)). The value of $\lceil \log 2(I_i) \rceil$ varies from 4 to 11. Thus, logarithmic transform is introduced to further divide traffic index while $N_0 - 1 < I_i \leq I_{max}$ (see Fig. 3(a)). The logarithmic transform formation of I_i is defined as $\overline{I_{i,\log}}$:

$$\overline{I_{I,\log}} = \begin{cases} 0, & 0 \leq I_i \leq N_0 - 1 \\ \lceil \log 2(I_i) \rceil, & N_0 - 1 < I_i \leq I_{max} \end{cases} \quad (8)$$

$\overline{I_i}$, $\overline{I_{i,\log}}$, I_i, and their corresponding traffic states are shown in Table 1, where $N_0 = 10$. While $I_{max} = 1333$, $9 < I_i \leq 1333$, $\overline{I_{i,\log}} = \lceil \log 2(I_i) \rceil = \{4, 5, 6, 7, 8, 9, 10, 11\}$.

TABLE 1. The discretized traffic index with $N_0 = 10$.

$\overline{I_i}$	$\overline{I_{i,\log}}$	I_i	Traffic state
1	0	$0 \leq I_i \leq 1$	Very smooth traffic
2	0	$1 < I_i \leq 2$	Very smooth traffic
3	0	$2 < I_i \leq 3$	Smooth traffic
4	0	$3 < I_i \leq 4$	Smooth traffic
5	0	$4 < I_i \leq 5$	Light congestion
6	0	$5 < I_i \leq 6$	Light congestion
7	0	$6 < I_i \leq 7$	Moderate congestion
8	0	$7 < I_i \leq 8$	Moderate congestion
9	0	$8 < I_i \leq 9$	Severe congestion
10	$\lceil \log 2(I_i) \rceil$	$9 < I_i < I_{max}$	Severe congestion

To identify the aggregate groups distinctly, group ID is defined as the function of $\overline{\gamma_i}, \overline{I_i}$, and $\overline{I_{i,\log}}$. For example, groupID $= 100000\overline{\gamma_i} + 1000\overline{I_i} + \overline{I_{i,\log}}$. This expression can make each parameter in a fixed position. Travel time data with the same group ID are aggregated into the same group. Define N_1 as the the maximum group number of γ_i, $N_1 = \left\lceil \frac{3600}{\Delta} \left(\frac{1}{\gamma_{min}} - \frac{1}{\gamma_{max}} \right) \right\rceil$. Define N_2 as the maximum group number of I_i. When $0 \leq I_i \leq N_0 - 1$, $N_2 = (N_0 - 1)$; when $N_0 - 1 < I_i \leq I_{max}$, $N_2 = \lceil \log 2(I_{max}) \rceil$; thus, for all I_i, $N_2 = N_0 - 1 + \lceil \log 2(I_{max}) \rceil$. The maximum group number of all aggregate groups $N_{max} = N_1 \cdot N_2$. For example, if $e_i \in [0, 3600]$, $\gamma_{min} = 18$, $\gamma_{max} = 120$, $\Delta = 20$, $N_0 = 10, \alpha = 0.15$, then $N_{max} = \left\lceil \frac{3600}{20} \times \left(\frac{1}{18} - \frac{1}{120} \right) \right\rceil \times \left\{ 10 - 1 + \left\lceil \log 2 \left(\frac{3600}{0.15 \times 18} \right) \right\rceil \right\} = 165$.

The aggregate approach of travel time based on free-flow travel time and traffic index is illustrated in Algorithm 1.

Algorithm 1 Travel Time Data Aggregation Based on Free-Flow Travel Time and Traffic Index

---

for (i=1; i≤m; i++) //m is the total number of links.
{
   $\gamma_i = 3600/v_{f,i}$;
   if $((t_i - \gamma_i) \geq 0)$
   {
      $\overline{\gamma_i} = \left\lceil \frac{3600}{\gamma_i \Delta} \right\rceil$;
      $I_i = \frac{e_i}{\alpha \cdot \gamma_i} = \frac{t_i - \gamma_i}{\alpha \cdot \gamma_i}$;
      // Calculate $\overline{I_i}$:
      if $(0 < I_i \leq N_0 - 1)$
      {
         $\overline{I_i} = \lceil I_i \rceil$;
      }
      else if $(I_i > N_0 - 1)$
      {
         $\overline{I_i} = N_0$;
      }
      else
      {
         $\overline{I_i} = 1$;
      }
      //Calculate $\overline{I_{i,\log}}$:
      if $(I_i > N_0 - 1)$
      {
         $\overline{I_{i,\log}} = \lceil \log 2(I_i) \rceil$;
      }
      else
      {
         $\overline{I_{i,\log}} = 0$;
      }
      // Calculate group ID:
      groupID $= 100000\overline{\gamma_i} + 1000\overline{I_i} + \overline{I_{i,\log}}$; // This expression
      can make each parameter in a fixed position. Travel time
      data with the same group ID are aggregated into the
      same group.
   }
}

---

2) DISTRIBUTION OF AGGREGATE GROUPS

Aggregated results with real data and Kolmogorov-Smirnov (KS) test are given as follows. Real data were collected by floating cars in Beijing in the morning peak period, non-peak period, and evening peak period on Tuesday, October 6, 2015. In the morning peak period, 2189 samples were aggregated into 29 groups. In the non-peak period,
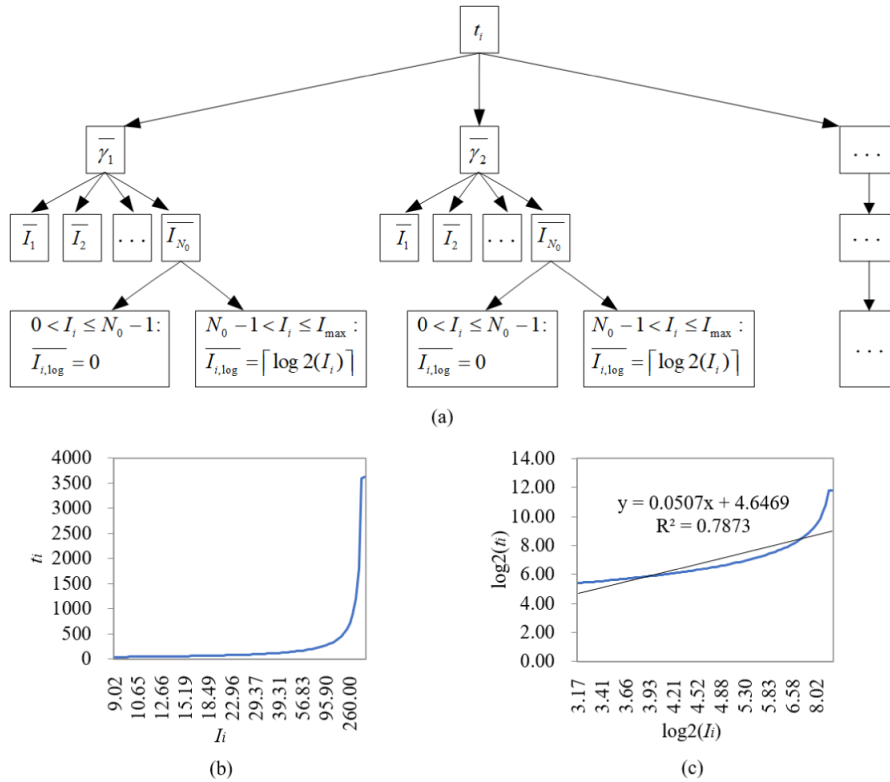
FIGURE 3. Aggregate approach: (a) the decision tree; (b) the relationship between t_i and I_i; (c) the relationship between $\log 2(t_i)$ and $\log 2(I_i)$.

2375 samples were aggregated into 37 groups. In the evening peak period, 2380 samples were aggregated into 37 groups.

Tested distribution included normal, LN, ESLN, gamma, beta distribution. If the statistic value is less than the critical value (CV) of the KS test, the distribution passes the KS test. The detailed results of the KS test are shown in Fig. 4. Most of the statistic values of normal, LN, ESLN, and gamma distribution were less than CV. In the morning peak period, the pass rates of ESLN, normal, LN, gamma, and beta distribution were 93%, 69%, 69%, 69%, and 7%, respectively (see Table 2). In the non-peak period, the pass rates of ESLN, normal, LN, gamma, and beta distribution were 92%, 59%, 62%, 59%, and 11%, respectively. In the evening peak period, the pass rates of ESLN, normal, LN, gamma, and beta distribution were 97%, 70%, 73%, 70%, and 8%, respectively. For all periods, ESLN distribution had the highest accuracy.

Average travel time per unit length in the new aggregate group is modeled as ESLN distribution, $t_{GID} \sim ESLN(\mu_{GID}, \sigma_{GID}, \gamma_{GID})$, where GID is the group ID. Parameters are the same across all links in the same aggregate group. However, for link i, t_i should satisfy $t_i \geq \gamma_i \geq \gamma_{min}$. If free-flow travel time in the same aggregate group is set as the same value, for example, $\gamma_{GID} = \gamma_{min}$, the maximum pass rates of the KS test in the morning peak period, non-peak period, evening peak period were 72%, 62%, and 73%,

TABLE 2. The accuracy of KS test.

Period	Result	Sample size	Group size	Beta	Gamma	LN	Normal	ESLN
Morning Peak	Pass	2189	29	2	20	20	20	27
	Reject			27	9	9	9	2
	Pass rate (%)			7	69	69	69	93
Non-peak	Pass	2375	37	4	22	23	22	34
	Reject			33	15	14	15	3
	Pass rate (%)			11	59	62	59	92
Evening peak	Pass	2380	37	3	26	27	26	36
	Reject			34	11	10	11	1
	Pass rate (%)			8	70	73	70	97

Note that data used in Table II and Fig. 4 are aggregated with the new proposed aggregate method. SLN cannot be compared in Table II and Fig. 4, because the aggregate method of SLN distribution is based on the facility type (the same lane number and the same divided type).

respectively, which were lower than 93%, 92%, and 97%. Thus, the parameter of free-flow travel time is not estimated in the same aggregate group.

For each aggregate group, the mean (μ_{GID}) and the standard deviation (σ_{GID}) in different periods are estimated using
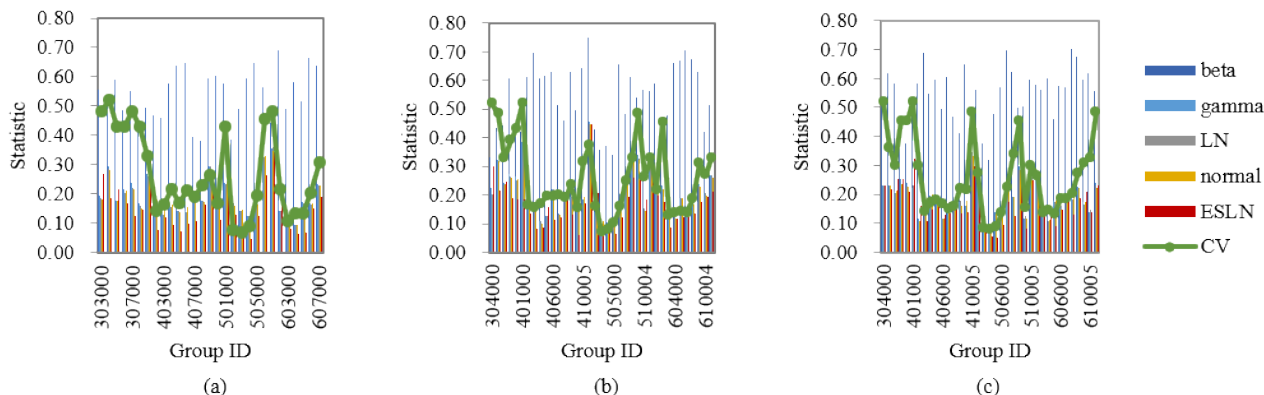
FIGURE 4. KS statistic results: (a) morning peak; (b)non-peak; (c) evening peak. Note that "CV" is the critical value of KS test.
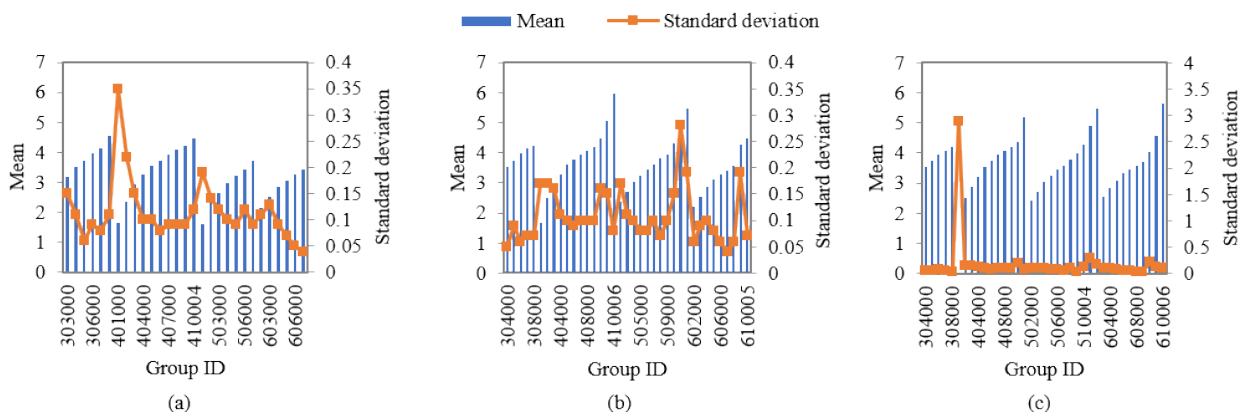

FIGURE 5. The mean and the standard deviation of aggregate groups with ESLN distribution in different periods: (a) morning peak; (b) non-peak; (c) evening peak.
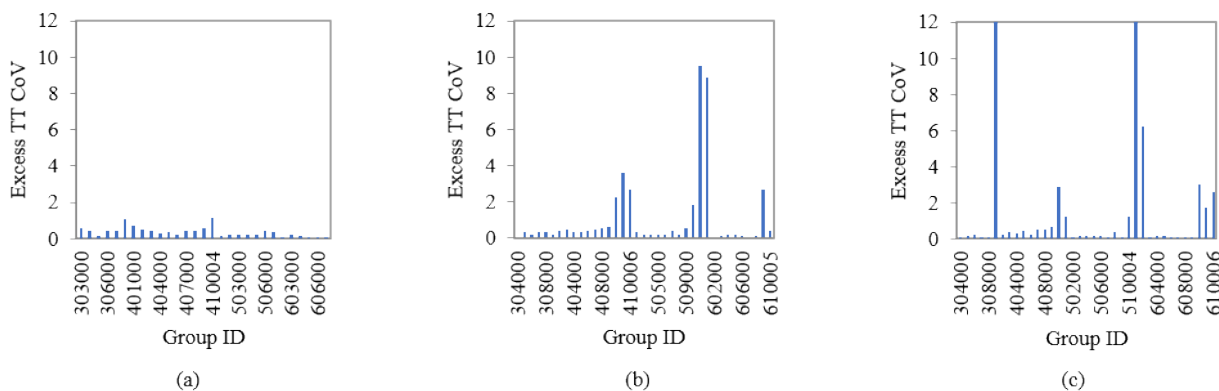

FIGURE 6. The Excess TT CoV in different periods: (a) morning peak; (b) non-peak; (c) evening peak.

maximum likelihood estimation technique. Fig. 5 shows the estimated results.

ANOVA tests are used to confirm the estimated results. Fig. 6 shows the coefficient of variation (CoV) of excess travel time (TT) in different periods. In the morning peak period, all Excess TT CoV were smaller than 1.2. In the non-peak and evening peak periods, the Excess TT CoV of most groups were smaller than 1, except for some groups whose $\bar{I}_i$ was 10. The main reason for this phenomenon is that $\bar{I}_i = 10$ covers $9 < I_i \le 1333$, which causes the Excess TT CoV of $\bar{I}_i = 10$ are generally higher than that of $\bar{I}_i = 1$-9.

3) CORRELATIONS BETWEEN AGGREGATE GROUPS

The correlations between aggregate groups can efficiently reduce the complexity compared with the link to link correlations. Maximum likelihood estimation technique is used to obtain the covariance matrix between aggregate groups. Link within groups has the same group ID, thus, link travel time correlations within groups are the correlations between groups with the same group ID.

TABLE 3. p-values of the hypothesis of no correlation.

	Morning peak	Evening peak	Non-peak
p-value<0.05	12%	15%	18%
p-value≥0.05	88%	85%	82%

The hypothesis of no correlation is tested by p-value. p-value is the probability of getting a correlation as large as the observed value by random chance, where the true correlation is zero. If p-value <0.05, the correlation is significant. Table 3 shows the percentage of p-values in different periods. For p-value <0.05, the proportions in the morning peak period, evening peak period, and non-peak period were 12%, 15%, and 18%, respectively. In the non-peak period, more groups were correlative compared with morning peak period and evening peak period. This phenomenon shows that heavy congestion can reduce the correlation, which coincides with the conclusion of Gajewski and Rilett [27] and further verifies the effectiveness of the proposed method.

B. LINK TRAVEL TIME MODEL

As ESLN distribution is chosen to represent link travel time per unit length in each aggregate group, μ_i, σ_i, and γ_i are defined as the parameters of ESLN distribution. t_i is the travel time per unit length on link i, $t_i \sim ESLN(\mu_i, \sigma_i, \gamma_i)$. t_i is described as the following structure:

$$t_i = \gamma_i + \exp(\mu_i + \sigma_i z_i) \tag{9}$$

where $e_i = \exp(\mu_i + \sigma_i z_i)$ is the excess travel time. Specifically, μ_i and σ_i are the mean and the standard deviation of the excess travel time. z_i is a standard normal random variable, $z_i \sim N(0, 1)$. Therefore, e_i is a random variable with lognormal distribution. The mean and the variance of t_i are calculated by the following equations:

$$E[t_i] = T_i = \gamma_i + \exp(\mu_i + 0.5\sigma_i^2) \tag{10}$$

$$Var[t_i] = V_i = \exp\left(2\mu_i + \sigma_i^2\right)\left[\exp\left(\sigma_i^2\right) - 1\right] \tag{11}$$

The mean excess travel time is calculated by the following equation:

$$M_i = E[t_i] - \gamma_i \tag{12}$$

Due to the use of aggregate approach, the average travel time per unit length is modeled as following an ESLN distribution. This assumption is relaxed in the subsequent sections while calculating the most reliable path, where the parameters can vary even across the links in the same aggregate group.

C. PATH TRAVEL TIME DISTRIBUTION

For t_i is the link travel time per unit length, the total travel time of link i should consider the link length, which is calculated by the following equation:

$$t_i l_i = [\gamma_i + \exp(\mu_i + \sigma_i z_i)] l_i = \gamma_i l_i$$
$$+ \exp(\mu_i + \sigma_i z_i) l_i = \gamma_i l_i + \exp(\mu_i + \ln l_i + \sigma_i z_i) \tag{13}$$

Thus, the total link travel time $t_i l_i$ can be distributed as $t_i l_i \sim ESLN(\mu_i + \ln l_i, \sigma_i, \gamma_i l_i)$, where l_i is the length of link i.

As link travel time follows an ESLN distribution, the path travel time distribution does not have a closed form CDF. Several approaches have been proposed to compute the complementary distribution function of the sum of lognormal random variables, such as Fenton-Wilkinson's approach [44], Schwartz and Yeh's approach [45], and cumulants matching approach [46]. Abu-Dayya and Beaulieu [47] has proved that Fenton-Wilkinson's approach is the best one. Moreover, Fenton-Wilkinson's approach has been used extensively in signal processing applications [47], [48], [49] and path travel time distribution [2]. Thus, Fenton-Wilkinson's approach is introduced to approximate the distribution of path travel time t_P by ESLN distribution with parameters μ_P, σ_P, and γ_P.

The mean value of excess travel time is calculated by the following formula:

$$\text{Exp}\left(\mu_P + 0.5\sigma_P^2\right) = \sum_{i \in P} \exp\left(\mu_i + \ln l_i + 0.5\sigma_i^2\right) \tag{14}$$

The variance of travel time is calculated by the following formula:

$$\text{Exp}\left(2\mu_P + \sigma_P^2\right)\left[\exp\left(\sigma_P^2\right) - 1\right]$$
$$= \sum_{i \in P} \exp\left[2(\mu_i + \ln l_i) + \sigma_i^2\right]\left[\exp\left(\sigma_P^2\right) - 1\right]$$
$$+ \sum_{i \in P, j \in P} \rho_{i,j}\left\{\exp\left[2(\mu_i + \ln l_i) + \sigma_i^2\right]\left[\exp\left(\sigma_P^2\right) - 1\right]\right\}^{0.5}$$
$$\times \left\{\exp\left[2(\mu_j + \ln l_j) + \sigma_j^2\right]\left[\exp\left(\sigma_P^2\right) - 1\right]\right\}^{0.5} \tag{15}$$

where $i, j \in A$ represent links and $\rho_{i,j}$ denotes the correlation coefficient between link travel times.

The shift parameter of travel time is calculated by the following formula:

$$\gamma_P = \sum_{i \in P} \gamma_i l_i \tag{16}$$

Thus, the travel time reliability of path P is computed using the CDF of the approximate path travel time distribution $t_P \sim ESLN(\mu_P, \sigma_P, \gamma_P)$.

The total travel time t_P can be expressed as:

$$t_P = \gamma_P + \exp(\mu_P + z\sigma_P) \tag{17}$$

The expected value and the variance of travel time are calculated by:

$$T_p = E[t_p] = \gamma_p + \exp(\mu_p + 0.5\sigma_p^2) \tag{18}$$

$$V_P = Var[t_P] = \exp\left(2\mu_P + \sigma_P^2\right)\left[\exp\left(\sigma_P^2\right) - 1\right] \tag{19}$$

where $z \sim N(0, 1)$ represents a normal random variable. e_P is defined as a random variable denoting excess travel time.

$$e_P = t_P - \gamma_P = \exp(\mu_P + z\sigma_P) \quad (20)$$

The expected value of excess travel time of path P is calculated by:

$$M_P = \text{E}[e_P] = \exp\left(\mu_P + 0.5\sigma_P^2\right) \quad (21)$$

From (21), μ_P can be written as:

$$\mu_P = \text{In}(M_P) - 0.5\sigma_P^2 \quad (22)$$

From (19) and (21), σ_P can be written as:

$$\sigma_P = \left[\ln\left(1 + \frac{V_P}{M_P^2}\right)\right]^{0.5} \quad (23)$$

D. MOST RELIABLE PATH MODEL WITH ESLN DISTRIBUTION

The objective of finding the most reliable path problem is to determine the path on a network with maximum reliability between a specific origin-destination (OD) pair (s, t) for a pre-specified travel time threshold T_0, where $s \in N$, $t \in N$.

The path travel time reliability R_P is defined as the probability that path travel time t_P is less than an associated threshold of travel time T_0:

$$R_P = P[t_P < T_0] \quad (24)$$

The objective function is to maximize the path travel time reliability R_P:

$$\text{Max } R_p(x) = \Phi\left(\frac{\text{In}(T_0 - \gamma_P) - \mu_P}{\sigma_P}\right) \quad (25)$$

$$st. \sum_{(h,k)\epsilon A} x_{hk} - \sum_{(k,h)\epsilon A} x_{kh} = \begin{cases} 1, & h = s \\ -1, & h = t \\ 0, & h \in N - \{s,t\} \end{cases} \quad (26)$$

where γ_P is the free-flow travel time of path P, obtained from (16); μ_P and σ_P are the mean and the standard deviation of the excess travel time of path P, obtained from (22) and (23). (26) denotes flow conservation constraints. Each link (h, k) has a tail node h and a head node k. $x_{hk} \in \{0, 1\}$: $x_{hk} = 1$ means that link (h, k) is on the path P; $x_{hk} = 0$ means that link (h, k) is not on the path P.

The most reliable path with ESLN distribution and general correlation structure is computed by the convergence algorithm of reliability bounds [2]. A set of $K_0 + K_1$ least expected travel time paths are computed. The first K_0 paths are added to the path set P_{K_0}, whose excess travel time (M_P) are lower than the threshold $M_0 = \max\left(\sqrt{V_{min}}, \frac{T_0 - \gamma_{min}}{e}\right)$ where e is the Euler's number. The next K_1 paths are added to the path set P_{K_1}. From the last K_1 paths, a sequence of K_1 largest and progressively decreasing upper bounds on path reliability are obtained. The largest reliability among the paths in the set of $K_0 + K_1$ paths forms a lower bound on the optimal reliability objective. Yen's K-shortest path algorithm is

implemented to determine the least expected excess travel time paths.

The reliability of path P can be reformulated as follows:

$$
\begin{aligned}
R_P &= \Phi\left(\frac{\ln(T_0 - \gamma_P) - \mu_P}{\sigma_P}\right) \\
&= \Phi\left(\frac{\ln(T_0 - \gamma_P) - \left(\ln(M_P) - 0.5\sigma_P^2\right)}{\sigma_P}\right) \\
&= \Phi\left(\frac{\ln\left(\frac{T_0 - \gamma_P}{M_P}\right)}{\sigma_P} + 0.5\sigma_P\right) \\
&= \Phi\left(\frac{\ln\left(\frac{T_0 - \gamma_P}{M_P}\right)}{\ln\left(1 + \frac{V_P}{M_P^2}\right)^{0.5}} + 0.5\sigma_P\right) \quad (27)
\end{aligned}
$$

The upper bound of R_P can be constructed by replacing γ_P, V_P, and σ_P with γ_{min}, V_{min}, and σ_{max}, respectively. γ_{min} is the shift parameter of the path with the minimum free-flow travel time. γ_{min} is determined by a standard shortest path algorithm with link costs as the free-flow travel time, $\gamma_{min} = \min_P \gamma_P$. V_{min} represents the minimum variance of travel time on any path for the given OD pair. σ_{max} represents the largest standard deviation of underlying normal distribution for any link on the network, $\sigma_{max} = \max_{i \in A} \sigma_i$.

Define R_{PUB} as the upper bound on the reliability of path P, then

$$R_P \leq \Phi\left(\frac{\ln\left(\frac{T_0 - \gamma_{min}}{M_P}\right)}{\ln\left(1 + \frac{V_{min}}{M_P^2}\right)^{0.5}} + 0.5\sigma_{max}\right) = R_{PUB} \quad (28)$$

The algorithm for computing the most reliable path based on ESLN distribution between an origin node and a destination node is shown in Algorithm 2.

IV. PERFORMANCE OF FINDING MOST RELIABLE PATH WITH ESLN DISTRIBUTION

In this section, a case study is conducted on the real-world large-scale network of Beijing to estimate the performance of finding the most reliable path with ESLN distribution. Road segments were divided at the positions where traffic was easily disturbed, including intersection, traffic light, pedestrian crossing, on/off ramp, toll station, service area, and gas station, etc. In addition, road segments around traffic light, on/off ramp, toll station, and service area, were divided into upstream, middle position, and downstream. After dividing, the road network of Beijing consisted of 80291 nodes (N) and 160533 directed links (A). Each link contained link ID, tail node ID, head node ID, name, length, direction, road grade, city code, width, lane number, etc. Real data were collected by floating cars in Beijing from September 16, 2015, to November 30, 2015, seventy-six days in total. The program was written in C++ and tested on Windows 10 (64-bit) system

Algorithm 2 Computing the Most Reliable Path

Initially, the lower bound on path reliability $R_{KLB} = 0$.

The upper bound on path reliability $R_{KUB} = 1$. Set two path sets P_{K_1} and P_{K_0} to be empty. Set a pre-specified tolerance ε and a pre-specified limit $\widehat{K}$. Set $k = 1$.

do

{

  Determine the kth shortest path P_k with link excess
    means as the costs;

  Compute the travel time distribution parameters
    $\left(\mu_{P_k}, \sigma_{P_k}, \gamma_{P_k}\right)$ using (14), (15), and (16);

  Compute the path reliability using the ESLN CDF

with

    the given reliability threshold T_0;

if $\left(M_{P_k} > \max\left(\sqrt{V}_{min}, \frac{T_0 - \gamma_{min}}{e}\right)\right)$

{

  Compute $R_{P_k UB}$ (the path reliability upper bound

of

   path P_k) using (28);

  Set $R_{KUB} = R_{P_k UB}$;

  Add the path to the path set P_{K_1};

}

else

{

  Add the path to the path set P_{K_0};

}

  Identify the path with the highest reliability in the path
   sets;

  $R_{KLB} = \max_{P \in P_{K_1} \cup P_{K_0}} R_P$;

  $k = k + 1$;

} while $\left(\frac{R_{KUB} - R_{KLB}}{R_{KLB}} > \varepsilon \text{ or } k \leq \widehat{K}\right)$

with Inter (R) Core (TM) i7-8550U processor, 1.80 GHz and 2.00 GHz CPU, and 8 GB RAM.

For link i, the mean and the standard deviation of the excess travel time per unit length μ_i and σ_i were calculated within the divided periods. To avoid the unreasonable value of travel time, the minimum travel time of each link (corresponding to the maximum value of speed, which may be beyond the reasonable range) was removed. Thus, the free-flow travel time γ_i was the second minimum value. Group ID was calculated with μ_i, σ_i, and γ_i. Travel time correlations were calculated between aggregate groups.

The origin was the Institute of Transportation System Science and Engineering, Beijing Jiaotong University (TSSE-BJTU). The destination was Terminal 3 of Beijing Capital International Airport (T3-BCIA). Fig. 7 shows the K-shortest paths with K=3 based on ESLN distribution. Fig. 8 shows the reliabilities of three paths with $T_0 =1200$ to 5000. While $T_0 =1200$ to 2402, the most reliable path was Path 3; while $T_0 >2402$, the most reliable path was Path 1. While $T_0 >2982$ (49.7 minutes), the maximum reliability was more than 0.95, which coincided with the actual running time and verified that path travel time should consider link length.
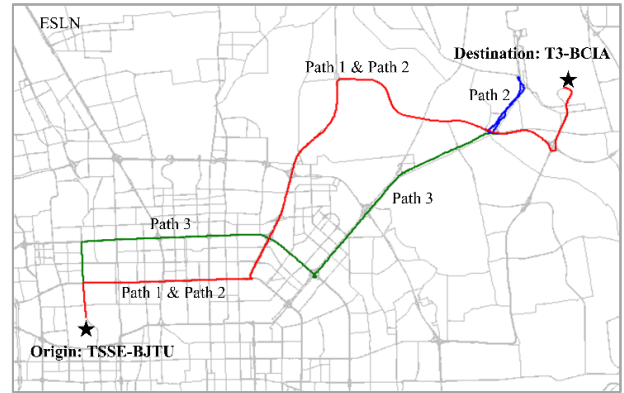
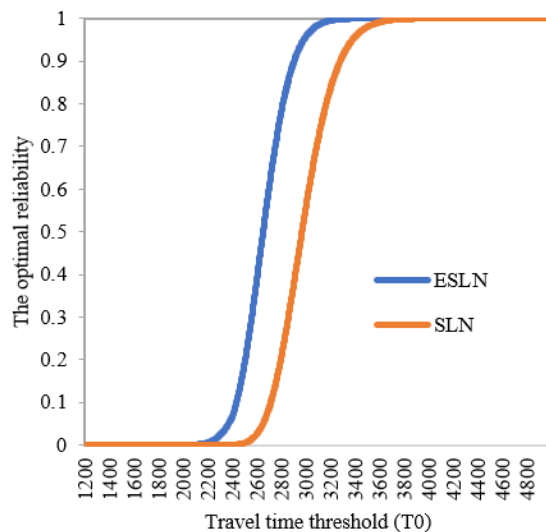**FIGURE 7.** The K- shortest paths with K=3 from TSSE-BJTU to T3-BCIA based on ESLN distribution.
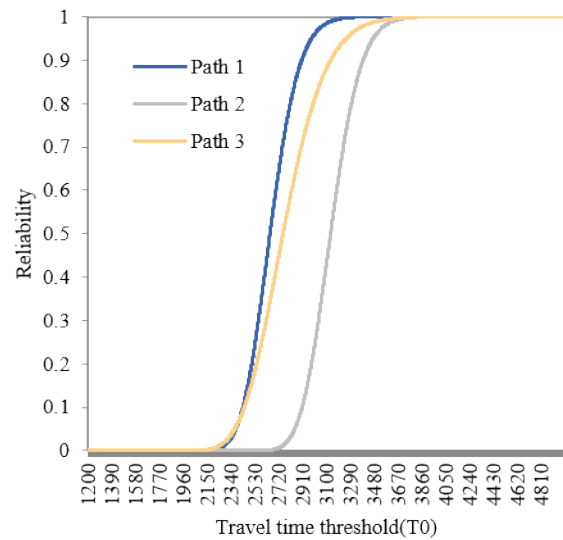
FIGURE 8. The reliabilities of three paths with T_0 =1200 to 5000 based on ESLN distribution.

To compare the difference between ESLN distribution and SLN distribution in finding the most reliable path, the corresponding parameters μ, σ, and γ within the same facility type (the same lane number and the same divided type) were calculated with real data collected in Beijing (see Table 4). Fig. 9 shows the K- shortest paths with K=3 from TSSE-BJTU to T3-BCIA based on SLN distribution whose parameters were aggregated within the same facility type. The reliabilities of three paths with T_0 =1200 to 5000 based on SLN distribution are shown in Fig. 10. For all travel time threshold, the most reliable path was Path 1.
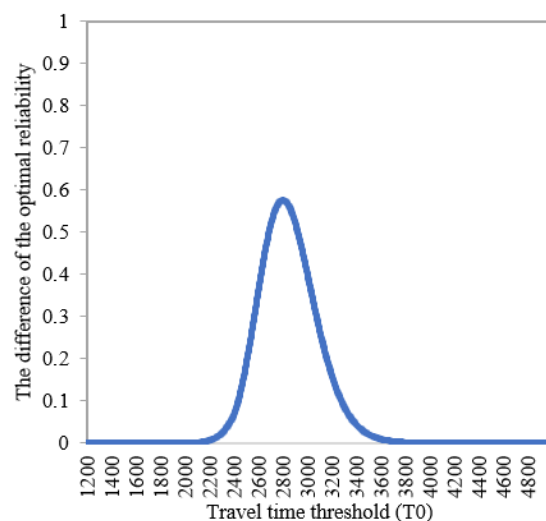
The comparison results between ESLN distribution and SLN distribution are shown in Fig. 11, where the optimal reliability of ESLN distribution was larger than that of SLN distribution. This phenomenon is caused by the different aggregate approaches of travel time. The aggregate approach of SLN distribution [2] is based on the same facility type (the same lane number and the same divided type), where the parameters are the same across all links in the same facility

TABLE 4. Parameters of SLN distribution within the same facility type.

Facility type	Sample size	γ(s/km)	μ(s/km)	σ(s/km)
1L-2W-D	39787179	50.0000	3.55144	0.935004
1L-2W-UD	38049718	60.0000	3.82166	0.917194
2L-2W-D	46367011	46.7532	3.64476	0.894093
2L-2W-UD	8027189	65.4545	3.68447	0.902696
3L-2W-D	32950974	41.8605	3.50524	0.942594
3L-2W-UD	2757672	61.0169	3.68293	0.936812
4L-2W-D	10083342	46.1538	3.28074	1.059950
4L-2W-UD	902806	63.1579	3.65142	0.935460
5L-2W-D	2015894	48.6486	3.27006	1.072690
5L-2W-UD	164086	65.4545	3.59622	0.944231
6L-2W-D	229435	55.3846	3.2914	0.968954
6L-2W-UD	32823	64.2857	3.58883	0.933673
7L-2W-D	49190	55.3846	3.06811	0.909617
7L-2W-UD	12251	85.7143	3.67127	0.811223

FIGURE 9. The K- shortest paths with K=3 from TSSE-BJTU to T3-BCIA based on SLN distribution.

FIGURE 10. The reliabilities of three paths with T_0 =1200 to 5000 based on SLN distribution.

type. In practice, travel time in the same facility type may vary in many cases, such as day-of-week, holidays, time-of-day, road grades, and traffic states, et al. The aggregate approach

FIGURE 11. Comparison results between ESLN distribution and SLN distribution: (a) the optimal reliability; (b) the optimal reliability difference between ESLN distribution and SLN distribution.
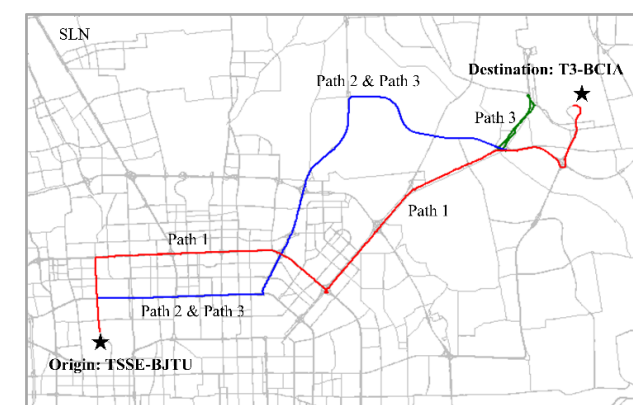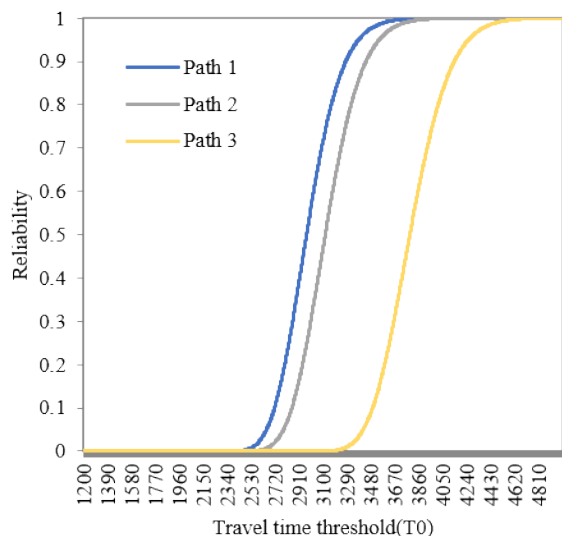
of SLN distribution does not distinguish these influence factors, which lead to a decline in the cumulative probability of travel time. The aggregate approach of ESLN distribution proposed in this study first divides travel time data into eight groups: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, Holidays. Further, travel time data on holidays are aggregated with different holidays. Then, based on the statistical characteristics of traffic flow, travel time data are divided into morning peak period, evening peak period, non-peak period, and night period. Finally, travel time data are aggregated based on road grades and traffic states. New proposed aggregate approach can separate different influence factors from each other. With these improvements, the optimal reliability of ESLN distribution exceeds that of SLN distribution.

V. CONCLUSIONS

The main research contents and results of this study in finding the most reliable path are as follows:

(1) Real data analysis found that day-of-week, holidays, time-of-day, road grades, and traffic states could cause the diversity of travel time and further reduced the cumulative probability of travel time even in the same facility type (the same lane number and the same divided type). A new aggregate approach was proposed to classify travel time based on the influence factors. The distribution with new aggregate approach was defined as the extended shifted lognormal (ESLN) distribution. KS test proved that ESLN distribution could effectively describe travel time distribution.

(2) Travel time correlations were calculated between new aggregate groups, which could reduce the complexity compared with the link to link correlations. Moreover, statistical results in different periods validated that heavy congestion could reduce the correlations.

(3) For travel time distribution was verified based on travel time per unit length, path travel time model considering link length was proposed.

(4) ESLN distribution was used to find the most reliable path in the real-world large-scale network (with 80291 nodes and 160533 links). The comparison results between ESLN distribution and SLN distribution indicated the improvement of the proposed method in finding the most reliable path.

Potential directions for future research include: propose a method for finding the optimal cutting points of daytime, explore a more reasonable method for determining the free-flow travel time. The computational time is proportional to the K value of K-shortest path algorithm and the distance between an origin node and a destination node. Thus, we will further study the acceleration algorithm for improving the computational efficiency. Furthermore, the algorithms for determining the departure time with a given on-time arrival probability based on ESLN distribution are intended to be proposed.

REFERENCES

[1] B. Y. Chen, W. H. K. Lam, A. Sumalee, Q. Li, and Z.-C. Li, "Vulnerability analysis for large-scale and congested road networks with demand uncertainty," *Transp. Res. A, Policy Pract.*, vol. 46, no. 3, pp. 501–516, Mar. 2012.

[2] K. K. Srinivasan, A. A. Prakash, and R. Seshadri, "Finding most reliable paths on networks with correlated and shifted log–normal travel times," *Transp. Res. B, Methodol.*, vol. 48, no. 66, pp. 110–128, Aug. 2014.

[3] A. T. Hojati, L. Ferreira, S. Washington, P. Charles, and A. Shobeirinejad, "Reprint of: Modelling the impact of traffic incidents on travel time reliability," *Transp. Res. C, Emer.*, vol. 70, pp. 86–97, Sep. 2016.

[4] C. Bi-Yu, W. H. K. Lam, Q. Li, A. Sumalee, and K. Yan, "Shortest path finding problem in stochastic time-dependent road networks with stochastic first-in-first-out property," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1907–1917, Dec. 2013.

[5] W. L. Zeng, T. Miwab, Y. Wakita, and T. Morikawa, "Application of Lagrangian relaxation approach to α-reliable path finding in stochastic networks with correlated link travel times," *Transp. Res. C, Emerg. Technol.*, vol. 56, pp. 309–334, Jul. 2015.

[6] J. Bates, J. Polak, P. Jones, and A. Cook, "The valuation of reliability for personal travel," *Transp. Res. E, Logistics Transp. Rev.*, vol. 37, no. 2, pp. 191–229, Apr. 2010.

[7] T. Lomax, D. Schrank, S. Turner, and R. Margiotta, "Selecting travel reliability measures," Texas Transp. Inst., Cambridge Systematics, Inc., Tech. Rep., Medford, MA, USA, May 2003.

[8] J. W. C. Van Lint, H. J. Van Zuylen, and H. Tu, "Travel time unreliability on freeways: Why measures based on variance tell only half the story," *Transp. Res. A, Policy Pract.*, vol. 42, no. 1, pp. 258–277, Jan. 2008.

[9] M. G. H. Bell, "Measuring network reliability: A game theoretic approach," *J. Adv. Trans.*, vol. 33, no. 2, pp. 135–146, Sep. 1999.

[10] A. Chen and Z. W. Ji, "Path finding under uncertainty," *J. Adv. Trans.*, vol. 39, no. 1, pp. 19–37, Jun. 2005.

[11] B. Y. Chen, W. H. K. Lam, A. Sumalee, Q. Li, H. Shao, and Z. Fang, "Finding reliable shortest paths in road networks under uncertainty," *Netw. Spatial Econ.*, vol. 13, no. 2, pp. 123–148, May 2012.

[12] H. Frank, "Shortest paths in probabilistic graphs," *Oper. Res.*, vol. 17, no. 4, pp. 583–599, Aug. 1969.

[13] L. Fu and L. R. Rilett, "Expected shortest paths in dynamic and stochastic traffic networks," *Transp. Res. B, Methodol.*, vol. 32, no. 7, pp. 499–516, Sep. 1998.

[14] R. Seshadri and K. K. Srinivasan, "Algorithm for determining most reliable travel time path on network with normally distributed and correlated link travel times," *Trans. Res. Rec.*, vol. 2196, no. 1, pp. 83–92, Jan. 2010.

[15] I. Kaparias, M. G. Bell, and H. Belzner, "A new measure of travel time reliability for in-vehicle navigation systems," *J. Intell. Trans. Syst.*, vol. 12, no. 4, pp. 202–211, Nov. 2008.

[16] Y. Hollander and R. Liu, "Estimation of the distribution of travel times by repeated simulation," *Transp. Res. C, Emer.*, vol. 16, no. 2, pp. 212–231, Apr. 2008.

[17] A. Polus, "A study of travel time and reliability on arterial routes," *Transportation*, vol. 8, no. 2, pp. 141–151, 1979.

[18] Y. M. Nie, X. Wu, J. F. Dillenburg, and P. C. Nelson, "Reliable route guidance: A case study from Chicago," *Transp. Res. A, Policy Pract.*, vol. 46, no. 2, pp. 403–419, Feb. 2012.

[19] J. Kim and H. S. Mahmassani, "A finite mixture model of vehicle-to-vehicle and day-to-day variability of traffic network travel times," *Transp. Res. C, Emerg. Technol.*, vol. 46, pp. 83–97, Sep. 2014.

[20] M. A. P. Taylor and Susilawati, "Modelling travel time reliability with the burr distribution," *Proc. Social Behav. Sci.*, vol. 54, pp. 75–83, Oct. 2012.

[21] E. B. Emam and H. Ai-Deek, "Using real-life dual-loop detector data to develop new methodology for estimating freeway travel time reliability," *Trans. Res. Rec.*, vol. 1959, no. 1, pp. 140–150, Jan. 2006.

[22] Z. H. Zhang, Q. He, J. Z. Gou, and X. L. Li, "Performance measure for reliable travel time of emergency vehicles," *Transp. Res. C, Emer.*, vol. 65, pp. 97–110, Apr. 2016.

[23] F. Lei, Y. Wang, G. Lu, and J. Sun, "A travel time reliability model of urban expressways with varying levels of service," *Transp. Res. C, Emer.*, vol. 48, pp. 453–467, Nov. 2014.

[24] A. Nicholson, "Travel time reliability benefits: Allowing for correlation," *Res. Trans. Econ.*, vol. 49, pp. 14–21, Jun. 2015.

[25] P. Rachtan, H. Huang, and S. Gao, "Spatiotemporal link speed correlations: Empirical study," *Trans. Res. Rec.*, vol. 2390, no. 1, pp. 34–43, Jan. 2013.

[26] T. Xing and X. S. Zhou, "Finding the most reliable path with and without link travel time correlation: A Lagrangian substitution based approach," *Transp. Res. B, Methodol.*, vol. 45, no. 10, pp. 1660–1679, Dec. 2011.

[27] B. J. Gajewski and L. R. Rilett, "Estimating link travel time correlation: An application of Bayesian smoothing splines," *J. Trans. Stat.*, vol. 7, nos. 2–3, pp. 53–70, May 2004.

[28] E. Jenelius and H. N. Koutsopoulos, "Travel time estimation for urban road networks using low frequency probe vehicle data," *Transp. Res. B, Methodol.*, vol. 53, pp. 64–81, Jul. 2013.

[29] R. A. Sivakumar and R. Batta, "The variance-constrained shortest path problem," *Trans. Sci.*, vol. 28, no. 4, pp. 309–316, Nov. 1994.

[30] E. D. Miller-Hooks and H. S. Mahmassani, "Least expected time paths in stochastic, time-varying transportation networks," *Trans. Sci.*, vol. 34, no. 2, pp. 198–215, May 2000.

[31] E. D. Miller-Hooks, "Adaptive least-expected time paths in stochastic, time-varying transportation and data networks," *Netw., Int. J.*, vol. 37, no. 1, pp. 35–52, Jan. 2001.

[32] E. D. Miller-Hooks and H. S. Mahmassani, "Path comparisons for a priori and time-adaptive decisions in stochastic, time-varying networks," *Eur. J. Oper. Res.*, vol. 146, no. 1, pp. 67–82, Apr. 2003.

[33] B. Y. Chen, W. H. K. Lam, A. Sumalee, Q. Q. Li, and M. L. Tam, "Reliable shortest path problems in stochastic time-dependent networks," *J. Intell. Trans. Syst.*, vol. 18, no. 2, pp. 177–189, Apr. 2014.

[34] L. Wang, Z. Gao, and L. Yang, "*A priori* least expected time paths in fuzzy, time-variant transportation networks," *Eng. Optim.*, vol. 48, no. 2, pp. 272–298, Feb. 2016.

[35] Y. Y. Fan, R. E. Kalaba, and J. Moore, II, "Shortest paths in stochastic networks with correlated link costs," *Comput. Math. Appl.*, vol. 49, no. 9, pp. 1549–1564, May 2005.

[36] Y. Nie and Y. Fan, "Arriving-on-time problem: Discrete algorithm that ensures convergence," *Trans. Res. Rec.*, vol. 1964, no. 1, pp. 193–200, Jan. 2006.

[37] T. Xing and X. Zhou, "Reformulation and solution algorithms for absolute and percentile robust shortest path problems," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 943–954, Jun. 2013.

[38] L. X. Yang and X. S. Zhou, "Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem," *Transp. Res. B, Methodol.*, vol. 59, pp. 22–44, Jan. 2014.

[39] L. Wang, L. Yang, and Z. Gao, "The constrained shortest path problem with stochastic correlated link travel times," *Eur. J. Oper. Res.*, vol. 255, pp. 43–57, Nov. 2016.

[40] L. Wang, L. Yang, Z. Gao, S. Li, and X. Zhou, "Evacuation planning for disaster responses: A stochastic programming framework," *Transp. Res. C, Emer.*, vol. 69, pp. 150–172, Aug. 2016.

[41] Z. Ji, Y. S. Kim, and A. Chen, "Multi-objective α-reliable path finding in stochastic networks with correlated link costs: A simulation-based multi-objective genetic algorithm approach (SMOGA)," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 1515–1528, Mar. 2011.

[42] A. Zockaie, Y. M. Nie, X. Wu, and H. S. Mahmassani, "Impacts of correlations on reliable shortest path finding: A simulation-based study," *Trans. Res. Rec.*, vol. 2334, no. 2, pp. 1–9, Jan. 2013.

[43] A. A. Prakash and K. Srinivasan, "Sample-based algorithm to determine minimum robust cost path with correlated link travel times," *Transp. Res. Rec.*, vol. 2467, no. 1, pp. 110–119, Jan. 2014.

[44] L. Fenton, "The sum of lognormal probability distribution in scatter transmission systems," *IRE Trans. Syst.*, vol. 8, no. 1, pp. 57–67, Mar. 1960.

[45] S. C. Schwartz and Y. S. Yeh, "On the distribution function and moments of power sums with lognormal components," *Bell Syst. Tech. J.*, vol. 61, no. 7, pp. 1441–1462, Sep. 1982.

[46] D. Schleher, "Generalized Gram–Charlier series with application to the sum of lognormal variates," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 2, pp. 275–280, Mar. 1977.

[47] A. A. Abu-Dayya and N. C. Beaulieu, "Outage probabilities in the presence of correlated lognormal interferers," *IEEE Trans. Veh. Technol.*, vol. 43, no. 1, pp. 164–173, Feb. 1994.

[48] F. Graziosi, L. Fuciarelli, and E. Santucci, "Second order statistics of the SIR for cellular mobile networks in the presence of correlated co-channel interferers," in *Proc. IEEE VTS 53rd Vehic. Tech. Conf.*, Rhodes, Greece, May 2001, pp. 2499–2503.

[49] N. C. Beaulieu, A. A. Abu-Dayya, and P. J. McLane, "Estimating the distribution of a sum of independent lognormal random variables," *IEEE Trans. Commun.*, vol. 43, no. 12, p. 2869, Dec. 1995.

ZIYOU GAO received the Ph.D. degree in operations research and control theory from the Institute of Applied Mathematics, Academia Sinica, Beijing, China, in 1994. He is currently a Professor with the School of Traffic and Transportation, Beijing Jiaotong University. His research interests include the modeling of transportation network systems, network design problems, traffic flow theory, and dynamic traffic assignment.

HUIJUN SUN received the B.S. degree in industrial engineering and the M.S. degree in management science and engineering from the Shandong University of Science and Technology in 1997 and 2000, respectively, and the Ph.D. degree in traffic and transportation planning and management from Beijing Jiaotong University, Beijing, China, in 2003. From 2003 to 2011, she was a Lecturer and an Associate Professor with the School of Traffic and Transportation, Beijing Jiaotong University. Since 2011, she has been a Professor with the School of Traffic and Transportation Engineering, Beijing Jiaotong University. Her research interests include travel behavior analyses, travel demand management, congestion pricing, and traffic policy.

FENG LIU received the M.S. degree in statistics from the Catholic University of Leuven, Belgium, in 2002, and the Ph.D. degree in computational biology from Hasselt University in 2007. She is currently an Associate Professor with the Travel Behavior Research Division, Transportation Research Institute (IMOB), Hasselt University, Belgium. Her current research focuses on travel demand modeling and transport network analysis using big data.

ZHENZHEN YANG received the B.S. degree from the School of Civil Engineering and Architecture, Hubei University of Technology, in 2009, and the M.S. degree from the School of Management, University of Shanghai for Science and Technology, in 2012. She is currently pursuing the Ph.D. degree with the School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China. Her research interests include intelligent transportation systems, traffic big data analysis, and machine learning and its application.

JIANDONG ZHAO received the B.S. and M.S. degrees from Xi'an Jiaotong University, China, in 1997 and 2000, respectively, and the Ph.D. degree from Tsinghua University in 2004. He is currently an Associate Professor with the School of Traffic and Transportation, Beijing Jiaotong University, China. His research interests include traffic forecast, traffic control and safety, and traffic emergency management. He is a member of the Key Laboratory of Urban Transportation Complex Systems Theory and Technology, Ministry of Education, Beijing Jiaotong University.

● ● ●