12-2017

Finding the 'faster' path in vehicle routing

Jing GUO

Yaoxin WU

Xuexi ZHANG

Le ZHANG

Wei CHEN

See next page for additional authors

Author
Jing GUO, Yaoxin WU, Xuexi ZHANG, Le ZHANG, Wei CHEN, Zhiguang CAO, and Hongliang GUO

IET Intelligent Transport Systems

Research Article

Finding the 'faster' path in vehicle routing

ISSN 1751-956X
Received on 4th November 2016
Revised 21st July 2017
Accepted on 13th August 2017
E-First on 17th October 2017
doi: 10.1049/iet-its.2016.0288
www.ietdl.org

Jing Guo[1], Yaoxin Wu[2], Xuexi Zhang[2], Le Zhang[3], Wei Chen[2], Zhiguang Cao[2], Lu Zhang[4], Hongliang Guo[5] ✉

[1]*Advanced Robotic Center (ARC), National University of Singapore, Singapore, Singapore*
[2]*School of Automation, Guangdong University of Technology, People's Republic of China*
[3]*Advanced Digital Sciences Center (ADSC), the Singapore-based research center of the University of Illinois at Urbana-Champaign (UIUC), Singapore*
[4]*Energy Research Institute @ NTU (ERIAN), Nanyang Technological University, Singapore*
[5]*School of Automation Engineering, University of Electronic Science and Technology of China, 2006 Xiyuan Avenue, Gaoxin West Zone, Chengdu, People's Republic of China*
✉ *E-mail: guohl1983@uestc.edu.cn*

Abstract: In this study, the authors improve the faster criterion in vehicle routing by extending the bi-delta distribution to the bi-normal distribution, which is a reasonable assumption for travel time on each road link. Based on this assumption, theoretical models are built for an arbitrary path and subsequently adopted to evaluate two candidate paths through probabilistic comparison. Experimental results demonstrate the bi-normal behaviour of link travel time in practice, and verify the faster criterion's superiority in determining the optimal path either on an artificial network with bi-normal distribution modelling link travel time or on a real road network with real traffic data. This study also validates that when the link number of one path is large, the probability density function of the whole path can be simplified by a normal distribution which approximates the sum of bi-normal distributions for each link.

1 Introduction

With the increase in vehicles and human populations as well as economic activities, the roads in urban areas, such as Guangzhou in China and Singapore, are likely to get ever busier, which inevitably lead to traffic jams. Traffic jam has always been an urgent and common problem, and it attracts broad attention from both industry and academic community because of its close relation to people's daily life. So travellers are more and more accustomed to get an optimal path from the navigator before their travels to avoid traffic congestion. Nowadays, it is not challenging for a navigation systems to find an optimal path in terms of the shortest distance [1, 2]. However, the shortest distance does not always guarantee a desirable travel time because there are many uncertain factors on road. This may prevent the driver from travelling on a real reliable route [3–5], since the traffic is always random, and the optimal criterion of the stochastic shortest path (SSP) is not unique. A different criterion needs different ways to explore the traffic data and results in different performance [6]. So it is crucial to analyse these criteria for the SSP problem, which aim to find an optimal path in stochastic traffic. Below are some commonly used criteria for the optimal path definition under a stochastic environment.

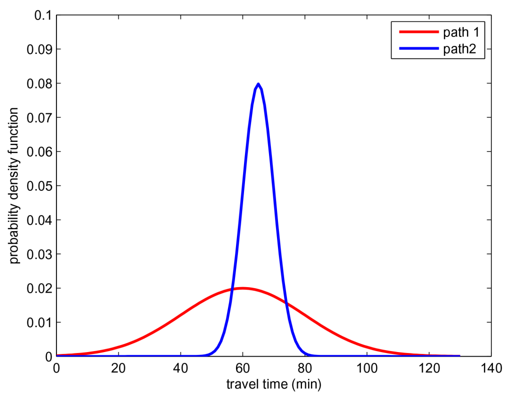

Fig. 1 *PDFs for two paths under a LET criterion*

Least expected travel time (LET). In a SSP problem, the LET is often used as one routing criterion. According to this criterion, the path is optimal only if it guarantees the minimum expected travel time. The travel time for each road link always follows a kind of distribution, so the mean of the travel time is usually fixed. In that case, finding the optimal path with LET is equal to solving one deterministic shortest path problem [7, 8], where the fixed weight for each road link is its expected travel time. Therefore, the general A^* or Dijkstra can be employed. However, a route with a minimum expected travel time might have a high variance due to the factors like weather and traffic signals, which may yield high risk of severe delay [9]. This can be intuitively illustrated by two probability distribution functions of travel time. As shown in Fig. 1, path 1 has a smaller mean and it is optimal according to LET. However, compared with path 2, path 1 has a higher variance, which means that the travellers on path 1 have less chance to traverse it using the expected travel time, namely, path 1 has a higher risk of a longer travel time. So, finding the SSP [10, 11] for a driver is much more desirable than merely returning a deterministic shortest path.

Mean-risk model. A mean-risk model is another widely used criterion in the SSP problem and on the basis of it, considerable work has been developed [12, 13]. In the mean-risk model, it seeks one path which minimises the weighted combination of the path's expected travel time and its standard deviation (i.e. minimise $E(x) + \gamma\,\mathrm{Var}(x)$, where x is the path's travel time distribution). This is a convex combination problem [14, 15], which can be converted into a deterministic shortest path problem with arc lengths being the linear combination of corresponding mean and variance. The mean-risk model does help solve the risk problem to some extent, but it still has one obvious limitation: the drivers may not understand the physical meaning of this model and they also do not know which γ is suitable.

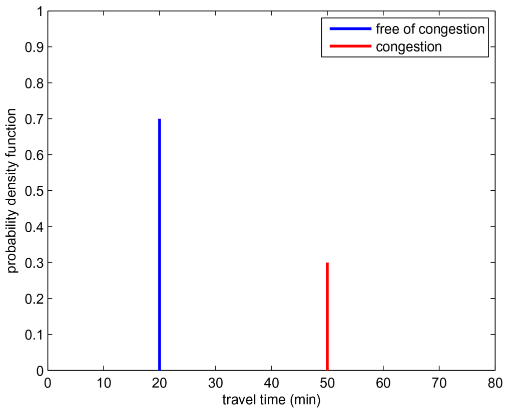

Fig. 2 *Illustrative example for the bi-delta distribution*

Travel time-based network performance model. In fact, the effectiveness of the routing strategy depends on its modelling of the real-world traffic data, which is generally stochastic, time-dependent and link-correlated. Numerous research works have managed to derive the realistic description of traffic network performance. Celikoglu [16] adopts conical delay function generated data to configure several neural networks for the approximation of flow rate-delay function, which is used to describe traffic dynamics. He also proposes an analytical node-based approach to solve the dynamic network loading problem, which integrates nodal rules with a link load computing process to simulate the acceleration behaviour of discrete vehicle packets [17]. Dell'Orco *et al.* [18] construct a quasi-anisotropic traffic flow model in which travel time prediction is based on a dynamic network loading process and the traffic assignment is implemented by bee colony optimisation. The research of a dynamic network loading model for travel time estimation has been developed in [19–21]. Additionally, Deniz *et al.* [22] conduct reliability evaluation of a macroscopic and a microscopic traffic model, respectively, analysing the distribution of reconstructed travel time by a random variable approach and comparing its statistical indicators. Celikoglu *et al.* [23] use dynamic classification and clustering technique to categorise flow patterns, evaluating the performance of a macroscopic flow model. Sun *et al.* [24] combine quadratic and constant basis function of time as piecewise truncated quadratic function to model speed trajectory, which can be used for travel time prediction. More research works on travel time estimation or network performance modelling can be found in [25, 26].

Stochastic routing over television networks. When the desirable network performance is well modelled, the real-time predicted travel data can be consulted in stochastic routing over a time-varying network. Laporte *et al.* [27] describe a routing problem with stochastic service and travel time as a chance constrained model and a recourse model, respectively, and solve them uniformly through branch and cut algorithm. Tas *et al.* [28] propose a vehicle routing model which minimises weighted sum of transportation cost and service cost and study on the trade-off between two types of costs and their impact on the optimal path. Miller-Hooks [29] considers updated arc traversal time en route and formulates recourse decisions as an adaptive LET hyper-path problem in a stochastic and time-varying (STV) network to determine optimal paths in different departure times. Cao *et al.* [30] formulate stochastic routing as a cardinality minimisation problem, approximating the probability of a path being the optimal by frequency technique. STV-based criteria can also be referenced in [31–33].

Faster criterion. Except for the above criteria that minimises the expected travel time, some other optimality conditions of routing is derived from the perspective of probability. Fastenrath and Becker [34] and Sigal *et al.* [35] propose a faster criterion. Under this criterion, the SSP problem can be determined based on a probabilistic comparison, which is much more reliable and easier to understand. Specifically, the optimal path is the one having the highest probability of being faster than all alternatives. To find the optimal path under this faster criterion, Fastenrath and Becker [34] first assume that only two traffic situations exist for each road link: congestion and congestion-free. In other words, there are two peaks for the travel time of each road link. The reason to assume two peaks is that traffic jam usually happens during the rush hours or the traffic flow stays fluent. So the travel time for each road link follows a bi-model distribution which has been proven by analysing the traffic data detected from a road sensor network in one city of Germany [34]. When the path traffic situations are independent of each other, the probability of one path being faster than all the alternatives can be expressed as a multiplication, each multiplier of which means the probability that it is faster than each individual alternative. Finally, the path with highest probability would be the optimal path which is probabilistically faster than all alternatives. The basis of this criterion is achieving a probabilistic fastest path by probability comparisons between all path pairs, which is more desirable for travellers, rather than relying on mere comparison among the expected travel time. In other words, this faster criterion is much more robust and reliable [34, 35].

Our contributions. The faster criterion is promising because it involves a comprehensive probabilistic comparison. However, there are still some limitations in [34]. It is reasonable to assume a bi-model distribution for the travel time of each road link; however, it only employs the bi-delta distribution as the bi-model, which means that for one specific road link, there would be only two single values for the travel time. It is not reasonable because generally, as long as the travel time is around one comparatively high value, we will believe traffic jam happens. On the other hand, if it is around one comparatively small value, we will believe there is no traffic jam [36]. So, the travel time for each of the two traffic situations is one range instead of one fixed value. Therefore it is more convincing to assume that the travel time follows a bi-normal distribution rather than the bi-delta distribution [37] and accordingly, in this study, we extend the bi-delta distribution in [34] to the bi-normal distribution. More specifically, we develop original theoretical models and draw some crucial conclusions for the optimal paths computation.

2 Problem formulation for faster criterion

In this section, we first formulate the faster criterion for simple paths, each of which only consists of one road link, and then we extend it to complex paths, each of which consists of multiple road links.

2.1 Faster criterion for simple paths

As we have stated in Section 1, the most important part to determine the faster path is to compute the probability that one path is faster than an individual alternative. To make this objective more formal, one path (i.e. path 1) will be faster than another path (i.e. path 2) if it meets the following requirement

$$P(T_1 \le T_2) > P(T_1 > T_2), \tag{1}$$

where T_i is the travel time for path i, and $P(\cdot)$ is the probability function. More specifically, $P(T_1 \le T_2)$ is computed by the following equation:

$$P(T_1 \le T_2) = \int_0^{\infty} P(T_1 \le t) \cdot \rho_2(t)\,\mathrm{d}t, \tag{2}$$

where $\rho_i(\cdot)$ is the probability density function (PDF) for path i. At the same time, it is important to note that

$$P(T_1 > T_2) = 1 - P(T_1 \le T_2)\,. \tag{3}$$

To better illustrate the principle of this faster criterion, we suppose each path only consists of one road link. In addition, travel time for each road link is modelled by the bi-delta distribution [34], one example of which is shown in Fig. 2. With this assumption, only two traffic situations exist, i.e. congestion and congestion free, and only one value of travel time is assigned to each situation.

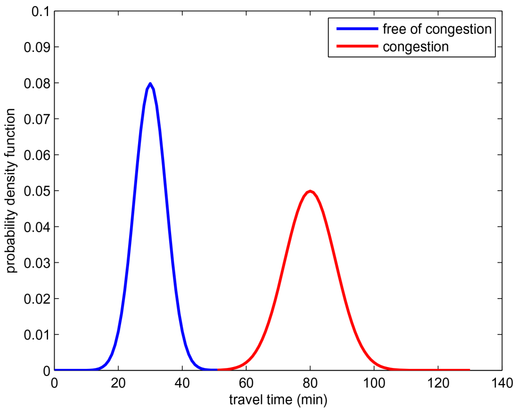

Fig. 3 *Illustrative example for the bi-normal PDF*

Moreover, each situation is also associated with one probability, which represents the chance that corresponding traffic situation will happen.

If we assume the travel time of path i ($i = 1, 2$) for congestion situation is $T_{i,\mathrm{c}}$, for free of congestion situation is $T_{i,\mathrm{f}}$, and $T_{i,\mathrm{f}}$ should be smaller than $T_{i,\mathrm{c}}$. At the same time, the probability for congestion situation of path 1 is p, and path 2 is q, and therefore, the probability for free of congestion situation would be $1 - p$ and $1 - q$, respectively. In this case, the PDF regarding the travel time for the two paths can be, respectively, expressed as follows:

$$\rho_1(t) = p \cdot \delta(t - T_{1,\mathrm{c}}) + (1 - p) \cdot \delta(t - T_{1,\mathrm{f}}), \tag{4}$$

$$\rho_2(t) = q \cdot \delta(t - T_{2,\mathrm{c}}) + (1 - q) \cdot \delta(t - T_{2,\mathrm{f}}), \tag{5}$$

where $\delta(\cdot)$ is the impulse function. Based on these two PDFs, we can further compute the probability that path 1 is faster than path 2 by integration, the result of which is stated as follows: (see (6)) where

$$\Theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 1 \end{cases} \tag{7}$$

To justify that the faster criterion is more reliable than others (i.e. LET), we show one example of finding a faster path with the following settings: $T_{1,\mathrm{f}} = 10$ min, $T_{2,\mathrm{f}} = 12$ min, $T_{1,\mathrm{c}} = 30$ min, $T_{2,\mathrm{c}} = 14$ min, $p = 0.4$ and $q = 0.6$. According to the LET criterion, the expected travel time for path 1 is 18 min, path 2 is 13.2 min, and the better path would be path 2. However, according to the faster criterion, we can get the corresponding probability as

$$\begin{aligned} P(T_1 \leq T_2) &= 0.6 \times 0.4 \times 1 + 0.6 \times 0.6 \times 1 \\ &\quad + 0.4 \times 0.4 \times 0 + 0.4 \times 0.6 \times 0 \\ &= 0.6 \end{aligned} \tag{8}$$

Therefore, path 1 is probabilistically faster and more reliable because it has a higher probability of being faster than path 2. Moreover, a probabilistic comparison is more robust than a pure expected travel time comparison.

2.2 Faster criterion for complex paths

In Section 2.1, we assume that each path consists of only one road link. However, it is more realistic that one path consists of multiple road links. To this end, we first assume that each path consists of two road links (for situation of more than two links, it is easy to be achieved based on the situation of two links). For the first road link of path 1, the travel time for its congestion situation is $T^1_{1,\mathrm{c}}$, the probability is p_1, and the travel time for the free of congestion situation is $T^1_{1,\mathrm{f}}$ and the probability is $1 - p_1$. Similarly, for the second road link, we have $T^2_{1,\mathrm{c}}$, p_2, $T^2_{1,\mathrm{f}}$ and $1 - p_2$. Thus, we can get the PDF regarding the travel time for each road link as follows:

$$\rho^1_1(t) = p_1 \cdot \delta(t - T^1_{1,\mathrm{c}}) + (1 - p_1) \cdot \delta(t - T^1_{1,\mathrm{f}}) \tag{9}$$

$$\rho^2_1(t) = p_2 \cdot \delta(t - T^2_{1,\mathrm{c}}) + (1 - p_2) \cdot \delta(t - T^2_{1,\mathrm{f}}) \tag{10}$$

Based on integration of (9) and (10), we can compute the PDF of entire path 1 as (see (11)) The PDF of path 2 can be obtained in the same way, and then we insert these two functions into (2) so that we can obtain the probability that path 1 is faster than path 2, which is computed according to (6). On the basis of this method, we can further extend to implement the probabilistic comparison for much more complex paths.

3 Extension from the bi-delta to bi-normal distribution

For the sake of computation simplification, the bi-delta distribution for the travel time of each road link is assumed in Section 2. However, the real travel time for each of the two traffic situations should be an arrangement rather than a single value. Therefore, it is more reasonable to assume the bi-normal distribution for each road link instead of the bi-delta distribution. This has been justified by analysing the traffic data detected by the road sensor network [38–40]. Accordingly, in this section, we will extend the bi-delta distribution to the bi-normal distribution for both simple path and complex path. We would like to remark that, we assume all the distributions we discussed in Sections 3 and 4 are independent of each other.

3.1 Mathematical model for the bi-normal distribution

The most important factor for a distribution is the PDF. For bi-normal PDF $\rho(t)$, it is usually expressed by a weighted sum of two single normal PDFs as follows:

$$\rho(t) = \phi \cdot \rho_\mathrm{f}(t) + (1 - \phi) \cdot \rho_\mathrm{c}(t), \tag{12}$$

where

$$\rho_\mathrm{f}(t) = \frac{1}{\sqrt{2\pi}\sigma_\mathrm{f}} \mathrm{e}^{-((t - \mu_\mathrm{f})^2/2\sigma_\mathrm{f}^2)}, \quad \rho_\mathrm{c}(t) = \frac{1}{\sqrt{2\pi}\sigma_\mathrm{c}} \mathrm{e}^{-((t - \mu_\mathrm{c})^2/2\sigma_\mathrm{c}^2)}.$$

In this formula, $\rho_\mathrm{f}(t)$ is the PDF for the free of congestion situation and $\rho_\mathrm{f}(c)$ is for congestion situation. ϕ represents the weight coefficient with $0 \leq \phi \leq 1$. For the parameters in two PDFs, $\mu_\mathrm{f} < \mu_\mathrm{c}$ is established obviously. Fig. 3 shows an example of bi-normal PDF with $\mu_\mathrm{f} = 30$, $\sigma_\mathrm{f} = 5$; $\mu_\mathrm{c} = 80$, $\sigma_\mathrm{c} = 8$, $\phi = 0.7$.

3.2 Faster criterion for simple paths

In this section, we assume that each path consists of only one road link and the travel time of each link follows a bi-normal distribution. We first suppose that there are two paths, and for paths 1 and 2, the PDFs are separately formulated as follows:

$$\begin{aligned} P(T_1 \leq T_2) = &(1 - p) \cdot (1 - q) \cdot \Theta(T_{2,\mathrm{f}} - T_{1,\mathrm{f}}) + (1 - p) \cdot q \cdot \Theta(T_{2,\mathrm{c}} - T_{1,\mathrm{f}}) \\ &+ p(1 - q) \cdot \Theta(T_{2,\mathrm{f}} - T_{1,\mathrm{c}}) + p \cdot q \cdot \Theta(T_{2,\mathrm{c}} - T_{1,\mathrm{c}}), \end{aligned} \tag{6}$$

$$\begin{aligned} \rho_1(T_1 = t) = &p_1 p_2 \delta(t - (T^1_{1,\mathrm{c}} - T^2_{1,\mathrm{c}})) + p_1(1 - p_2)\delta(t - (T^1_{1,\mathrm{c}} - T^2_{1,\mathrm{f}})) \\ &+ (1 - p_1)p_2\delta(t - (T^1_{1,\mathrm{f}} - T^2_{1,\mathrm{c}})) + (1 - p_1)(1 - p_2)\delta(t - (T^1_{1,\mathrm{f}} + T^2_{1,\mathrm{f}})) \end{aligned} \tag{11}$$

$$\rho_1(t) = \phi_1 \cdot \rho_{1,\mathrm{f}}(t) + (1-\phi_1)\cdot \rho_{1,\mathrm{c}}(t) \tag{13}$$

$$\rho_2(t) = \phi_2 \cdot \rho_{2,\mathrm{f}}(t) + (1-\phi_2)\cdot \rho_{2,c}(t) \tag{14}$$

The mean and standard deviation of $\rho_{1,\mathrm{f}}(t)$ are $\mu_{1,\mathrm{f}}$ and $\sigma_{1,\mathrm{f}}$, and $\rho_{1,\mathrm{c}}(t)$ are $\mu_{1,\mathrm{c}}$ and $\sigma_{1,\mathrm{c}}$, respectively. Similarly, $\mu_{2,\mathrm{f}}$, $\sigma_{2,\mathrm{f}}$ and $\mu_{2,\mathrm{c}}$, $\sigma_{2,\mathrm{c}}$ represent the mean and standard deviation of $\rho_{2,\mathrm{f}}(t)$ and $\rho_{2,\mathrm{c}}(t)$, respectively. We also assume that the travel time of path i is T_i. Then a double integral of $\rho_{1,\mathrm{f}}(t_1)\rho_{2,\mathrm{f}}(t_2)$ is calculated to determine the probability that path 1 is faster than path 2, which is expressed as

$$\begin{aligned}P(T_1 \le T_2) &= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{t_2}\rho_1(t_1)\,\mathrm{d}t_1\right]\cdot\rho_2(t_2)\,\mathrm{d}t_2\\ &= \phi_1\phi_2\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{f}}(t_1)\cdot\rho_{2,\mathrm{f}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\\ &\quad+\phi_1(1-\phi_2)\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{f}}(t_1)\cdot\rho_{2,\mathrm{c}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\\ &\quad+(1-\phi_1)\phi_2\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{c}}(t_1)\cdot\rho_{2,\mathrm{f}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\\ &\quad+(1-\phi_1)(1-\phi_2)\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{c}}(t_1)\cdot\rho_{2,\mathrm{c}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\end{aligned} \tag{15}$$

where the four parts in the sum can be individually represented as

$$P_{\mathrm{A}} = \phi_1\phi_2\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{f}}(t_1)\cdot\rho_{2,\mathrm{f}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \tag{16}$$

$$P_{\mathrm{B}} = \phi_1(1-\phi_2)\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{f}}(t_1)\cdot\rho_{2,\mathrm{c}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \tag{17}$$

$$P_{\mathrm{C}} = (1-\phi_1)\phi_2\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{c}}(t_1)\cdot\rho_{2,\mathrm{f}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \tag{18}$$

$$P_{\mathrm{D}} = (1-\phi_1)(1-\phi_2)\int_{-\infty}^{+\infty}\int_{-\infty}^{t_2}\rho_{1,\mathrm{c}}(t_1)\cdot\rho_{2,\mathrm{c}}(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \tag{19}$$

Each of the four parts can be further expressed as a more compact form. An example of P_{A} is shown here and we suppose variables $X_{1,\mathrm{f}}$ and $X_{2,\mathrm{f}}$ follow normal distributions with PDF $\rho_{1,\mathrm{f}}(t)$ and $\rho_{2,\mathrm{f}}(t)$, respectively. Looking into P_{A} in (16), we can reformulate it as

$$\begin{aligned}P_{\mathrm{A}} &= \phi_1\phi_2 P(X_{1,\mathrm{f}} \le X_{2,\mathrm{f}})\\ &= \phi_1\phi_2 P(X_{1,\mathrm{f}} - X_{2,\mathrm{f}} \le 0)\end{aligned} \tag{20}$$

Since $X_{1,\mathrm{f}}$ and $X_{2,\mathrm{f}}$ are two normal variables and it is reasonable to assume another normal variable $X_{\mathrm{A}} = X_{1,\mathrm{f}} - X_{2,\mathrm{f}}$, whose mean and standard deviation are $\mu_{\mathrm{A}} = \mu_{1,\mathrm{f}} - \mu_{2,\mathrm{f}}$ and $\sigma_{\mathrm{A}}^2 = \sigma_{1,\mathrm{f}}^2 + \sigma_{2,\mathrm{f}}^2$, respectively [41]. So we can obtain the compact form of P_{A} as

$$P_{\mathrm{A}} = \phi_1\phi_2\cdot\Psi_{\mu_{\mathrm{A}},\sigma_{\mathrm{A}}^2}(0), \tag{21}$$

where $\Psi_{\mu_{\mathrm{A}},\sigma_{\mathrm{A}}^2}(0)$ is the probability that normal variable X_{A} is not larger than 0, with mean μ_{A} and standard deviation σ_{A}. Similarly, we can obtain the compact form for P_{B}, P_{C} and P_{D}, and the simplified result of (15) is eventually expressed as

$$\begin{aligned}P(T_1 \le T_2) &= \phi_1\phi_2\Psi_{\mu_{\mathrm{A}},\sigma_{\mathrm{A}}^2}(0) + \phi_1(1-\phi_2)\Psi_{\mu_{\mathrm{B}},\sigma_{\mathrm{B}}^2}(0)\\ &\quad+(1-\phi_1)\phi_2\Psi_{\mu_{\mathrm{C}},\sigma_{\mathrm{C}}^2}(0) + (1-\phi_1)(1-\phi_2)\Psi_{\mu_{\mathrm{D}},\sigma_{\mathrm{D}}^2}(0)\end{aligned} \tag{22}$$

According to the value of $P(T_1 \le T_2)$ in (22), we can determine which path is faster based on probability. This criterion refers to the historical observations of travel time and the final decision avoids the uncertainty caused by travel time variation which is ill-considered in the LET criterion.

3.3 Faster criterion for complex paths

Similar to Section 2.2, we assume that each path consists of two road links and the travel time of each link follows a bi-normal distribution. For the first link, the PDF is expressed as follows:

$$\rho_1(t) = \phi_1\cdot\rho_{1,\mathrm{f}}(t) + (1-\phi_1)\cdot\rho_{1,\mathrm{c}}(t) \tag{23}$$

where

$$\rho_{1,\mathrm{f}}(t) = \frac{1}{\sqrt{2\pi}\sigma_{1,\mathrm{f}}}\mathrm{e}^{-((t-\mu_{1,\mathrm{f}})^2/2\sigma_{1,\mathrm{f}}^2)},\quad \rho_{1,\mathrm{c}}(t) = \frac{1}{\sqrt{2\pi}\sigma_{1,\mathrm{c}}}\mathrm{e}^{-((t-\mu_{1,\mathrm{c}})^2/2\sigma_{1,\mathrm{c}}^2)} \tag{24}$$

and for the second road link, the PDF can be expressed in the same way as

$$\rho_2(t) = \phi_2\cdot\rho_{2,\mathrm{f}}(t) + (1-\phi_2)\cdot\rho_{2,\mathrm{c}}(t), \tag{25}$$

where

$$\rho_{2,\mathrm{f}}(t) = \frac{1}{\sqrt{2\pi}\sigma_{2,\mathrm{f}}}\mathrm{e}^{-((t-\mu_{2,\mathrm{f}})^2/2\sigma_{2,\mathrm{f}}^2)},\quad \rho_{2,\mathrm{c}}(t) = \frac{1}{\sqrt{2\pi}\sigma_{2,\mathrm{c}}}\mathrm{e}^{-((t-\mu_{2,\mathrm{c}})^2/2\sigma_{2,\mathrm{c}}^2)} \tag{26}$$

Consequently, we can obtain the PDF for the whole path as

$$\begin{aligned}\rho(t) &= \int_{-\infty}^{t}\rho_1(\tau)\cdot\rho_2(t-\tau)\mathrm{d}\tau\\ &= \rho_1(t)*\rho_2(t)\\ &= \phi_1\phi_2\rho_{1,\mathrm{f}}(t)*\rho_{2,\mathrm{f}}(t) + \phi_1(1-\phi_2)\rho_{1,\mathrm{f}}(t)*\rho_{2,\mathrm{c}}(t)\\ &\quad+(1-\phi_1)\phi_2\rho_{1,\mathrm{c}}(t)*\rho_{2,\mathrm{f}}(t) + (1-\phi_1)(1-\phi_2)\rho_{1,\mathrm{c}}(t)*\rho_{2,\mathrm{c}}(t),\end{aligned} \tag{27}$$

where $*$ is the convolution operator. Since each $\rho_{i,\{\mathrm{f},\mathrm{c}\}}(t)$ in (27) is a normal PDF, and the convolution of them is another normal PDF according to the following fact [42]:

$$\begin{aligned}g_1(t) &= \rho_{1,\mathrm{f}}(t)*\rho_{2,\mathrm{f}}(t)\\ &= \frac{1}{\sqrt{2\pi(\sigma_{1,\mathrm{f}}^2+\sigma_{2,\mathrm{f}}^2)}}\mathrm{e}^{-((t-(\mu_{1,\mathrm{f}}+\mu_{2,\mathrm{f}}))^2/2(\sigma_{1,\mathrm{f}}^2+\sigma_{2,\mathrm{f}}^2))}\end{aligned} \tag{28}$$

So all convolutions between two PDFs can be replaced by one normal PDF and the travel time regarding the whole path can be further expressed as

$$\begin{aligned}\rho(t) &= \phi_1\phi_2 g_1(t) + \phi_1(1-\phi_2)g_2(t)\\ &\quad+(1-\phi_1)\phi_2 g_3(t) + (1-\phi_1)(1-\phi_2)g_4(t),\end{aligned} \tag{29}$$

where $g_i(t)$ represents a new normal PDF and its mean and standard deviation could be easily derived based on (28). Similarly, we can also compute the PDF for any path with N (e.g. $N > 2$) road links. After obtaining the PDF for each whole path, we can use (15) and (22) to determine the optimal path according to the faster criterion.

4 Using the normal distribution to approximate the bi-normal distribution

In this section, we prove and conclude that the normal distribution can be used to approximate the bi-normal distribution so that in

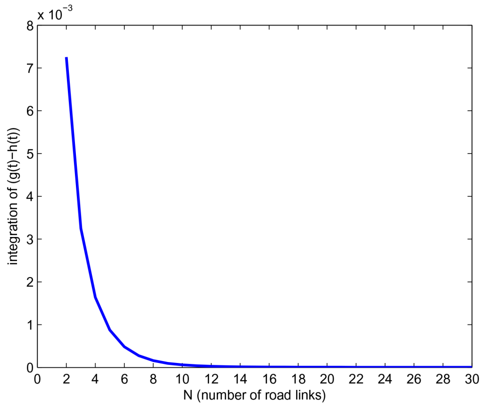

Fig. 4 *Integration of* $(g(t) - h(t))^2$ *with different N*

general, the computation would be greatly simplified and more efficient with travel time following a normal distribution.

4.1 PDF of the whole path by individual normal distributions

First we still assume one bi-normal PDF $\rho_i(t)$ for road link i, which is described in the same way as (23). Then for validating the proposed argument, we approximate the bi-normal distribution by a normal distribution with the mean and standard deviation achieved as follows:

$$\mu_i = \int_{-\infty}^{+\infty} t \cdot \rho_i(t)\,\mathrm{d}t = \phi_i\mu_{i,\mathrm{f}} + (1 - \phi_i)\mu_{i,\mathrm{c}} \tag{30}$$

(see (31)) We suppose that there are N road links for the whole path, each of which follows a bi-normal distribution with the same parameters. Then the parameters of the approximated normal distribution can be computed by (30) and (31). Consequently, the PDF of the whole path by the approximated normal distribution can be expressed as

$$h(t) = \frac{1}{\sqrt{2\pi}\sigma_h}\mathrm{e}^{-((t - \mu_h)^2/2\sigma_h^2)} \tag{32}$$

where

$$\mu_h = N[\phi\mu_\mathrm{f} + (1 - \phi)\mu_\mathrm{c}] \tag{33}$$

$$\sigma_h^2 = N[\phi(1 - \phi)(\mu_\mathrm{c} - \mu_\mathrm{f})^2 + \phi\sigma_\mathrm{f}^2 + (1 - \phi)\sigma_\mathrm{c}^2] \tag{34}$$

4.2 PDF of the whole path by individual bi-normal distributions

In contrast, if we directly use the bi-normal distribution, then the accurate PDF of the same path can be described as

$$\begin{aligned} g(t) &= \rho(t)^{N^*} \quad (* \text{ means convolution}) \\ &= \sum_{i=0}^{N}\binom{N}{i}\phi^i(1 - \phi)^{N-i}\rho_\mathrm{f}(t)^{i^*} * \rho_\mathrm{c}(t)^{(N-i)^*} \end{aligned} \tag{35}$$

Since convolutions of several normal PDF is another normal PDF, we can get that

$$\rho_\mathrm{c}^{i^*} * \rho_\mathrm{f}^{(N-i)^*} = \rho_{i,N-i}(t) \tag{36}$$

where $\rho_{i,N-i}(t)$ is a new normal PDF with

$$\mu_{i,N-i} = \mathrm{i}\mu_\mathrm{f} + (N - i)\mu_\mathrm{c} \tag{37}$$

$$\sigma_{i,N-i}^2 = \mathrm{i}\sigma_\mathrm{f}^2 + (N - i)\sigma_\mathrm{c}^2 \tag{38}$$

so $g(t)$ can be simplified as

$$g(t) = \sum_{i=0}^{N}\binom{N}{i}\phi^i(1 - \phi)^{N-i}\rho_{i,N-i}(t) \tag{39}$$

4.3 Comparison of the approximated and accurate PDFs

The integrated square difference between the approximated PDF $h(t)$ and the accurate PDF $g(t)$ can be expressed as

$$\begin{aligned} &\int_{-\infty}^{+\infty}(g(t) - h(t))^2\,\mathrm{d}t \\ &= \int_{-\infty}^{+\infty}\left[\sum_{i=0}^{N}\binom{N}{i}\phi^i(1 - \phi)^{N-i}\rho_{i,N-i}(t)\right]^2\mathrm{d}t + \int_{-\infty}^{+\infty}h(t)^2\,\mathrm{d}t \\ &-2\int_{-\infty}^{+\infty}\sum_{i=0}^{N}\binom{N}{i}\phi^i(1 - \phi)^{N-i}\rho_{i,N-i}(t)h(t)\,\mathrm{d}t, \end{aligned} \tag{40}$$

where we note that the production of two normal PDFs is another weighted normal PDF [43] because (see (41)) So we can simplify the (40) as follows:

$$\begin{aligned} &\int_{-\infty}^{+\infty}(g(t) - h(t))^2\,\mathrm{d}t \\ &= \sum_{i=0}^{N}\left[\binom{N}{i}\phi^i(1 - \phi)^{N-i}\right]^2 C_{i,i} - 2\sum_{i=0}^{N}\binom{N}{i}\phi^i(1 - \phi)^{N-i}C_{i,h} \\ &+2\sum_{i=0}^{N-1}\sum_{j=i+1}^{N}\binom{N}{i}\binom{N}{j}\phi^i(1 - \phi)^{N-i}\phi^j(1 - \phi)^{N-j}C_{i,j} + C_{h,h} \end{aligned} \tag{42}$$

Based on (41) and (42), we can derive that

$$C_{i,i} = \frac{1}{\sqrt{2\pi}\sqrt{2\sigma_i^2}} \tag{43}$$

$$\begin{aligned} \sigma_i^2 &= \int_{-\infty}^{+\infty}(t - \mu_i)^2 \cdot \rho_i(t)\,\mathrm{d}t \\ &= \int_{-\infty}^{+\infty}[t - (\phi_i\mu_{i,\mathrm{f}} + (1 - \phi_i)\mu_{i,\mathrm{c}})]^2[\phi_i \cdot \rho_{i,\mathrm{f}}(t) + (1 - \phi_i) \cdot \rho_{i,\mathrm{c}}(t)]\,\mathrm{d}t \\ &= \phi_i(1 - \phi_i)(\mu_{i,\mathrm{f}} - \mu_{i,\mathrm{c}})^2 + \phi_i\sigma_{i,\mathrm{f}}^2 + (1 - \phi_i)\sigma_{i,\mathrm{c}}^2 \end{aligned} \tag{31}$$

$$\begin{aligned} \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-((t - \mu_1)^2/2\sigma_1^2)}\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-((t - \mu_2)^2/2\sigma_2^2)} &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}}\mathrm{e}^{-((\mu_1 - \mu_2)^2/2(\sigma_1^2 + \sigma_1^2))} \\ &\times \frac{1}{\sqrt{2\pi}(\sigma_1\sigma_2/\sqrt{\sigma_1^2 + \sigma_2^2})}\mathrm{e}^{-((t - ((\sigma_2^2\mu_1 + \sigma_1^2\mu_2)/(\sigma_1^2 + \sigma_2^2)))^2/2(\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)))} \end{aligned} \tag{41}$$

where $\sigma_i^2 = i\sigma_f^2 + (N - i)\sigma_c^2$

$$C_{h,h} = \frac{1}{\sqrt{2\pi}\sqrt{2\sigma_h^2}} \tag{44}$$

where $\sigma_h^2 = N[\phi(1 - \phi)(\mu_c - \mu_f)^2 + \phi\sigma_f^2 + (1 - \phi)\sigma_c^2]$.

$$C_{i,j} = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_i^2 + \sigma_j^2}} e^{-((\mu_i - \mu_j)^2/2(\sigma_i^2 + \sigma_j^2))} \tag{45}$$

where $\mu_i = i\mu_f + (N - i)\mu_c$, $\mu_j = j\mu_f + (N - j)\mu_c$, $\sigma_i^2 = i\sigma_f^2 + (N - i)\sigma_c^2$ and $\sigma_j^2 = j\sigma_f^2 + (N - j)\sigma_c^2$

$$C_{i,h} = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_i^2 + \sigma_h^2}} e^{-((\mu_i - \mu_h)^2/2(\sigma_i^2 + \sigma_h^2))} \tag{46}$$

where $\mu_i = i\mu_f + (N - i)\mu_c$, $\mu_h = N[\phi\mu_f + (1 - \phi)\mu_c]$, $\sigma_i^2 = i\sigma_f^2 + (N - i)\sigma_c^2$ and $\sigma_h^2 = N[\phi(1 - \phi)(\mu_f - \mu_c)^2 + \phi\sigma_f^2 + (1 - \phi)\sigma_c^2]$.

To visualise the result, we plot the integration of the square difference against the road link number N in Fig. 4 from which we observe that as N increases, the difference between the accurate and approximate PDF of the travel time for the whole path becomes smaller.With large N, e.g. ten or larger, it is clear that the normal distribution can well approximate the bi-normal distribution.

This result is promising since the normal distribution is much easier for transformation and computation, and many conclusions have been drawn regarding the normal distribution of travel time in vehicle routing algorithms.

5 Experimental results and analysis

In this section, we provide experimental results and analysis to justify that our faster criterion is superior to the LET. The whole experiment is implemented in MATLAB 2014a, and divided into two parts containing tests on an artificial road network and a real road network, respectively. The artificial road network is represented by a directed graph, which is shown in Fig. 5. The real traffic data are derived from an area of Beijing in which some link travel times will be displayed with histogram to show their bi-normal behaviours. In addition, we must emphasise that the routing problem here only involves two candidate paths, for convenience to compare the travel time distribution on paths and analyse the comparison result. In the real-world scenario, our criterion can be used when we extend the probability comparison to all candidate paths between the origin and destination.

5.1 Experiments on the artificial road network with artificial data

5.1.1 Faster road in one-to-one case: We first compare the faster criterion with LET in one-to-one case, which means two simple paths both with one link are considered to determine the better one. A varied value of ϕ and variance is assigned to show results under various road conditions.

The experimental setting is stated as follows:

- We randomly generate the u_f, σ_f, u_c and σ_c for each bi-normal distribution.
- We randomly generate 100 traffic data (travel time) for each link according to the bi-normal distribution.
- We randomly choose O–D pairs, satisfying D is reachable from O and both roads consist of one link.
- According to these data, we compute the optimal path P_f with the highest probability of being fastest (i.e. the path with the most times being fastest), and P_l with the LET.
- The first four steps are repeated for 200 times and the number of P_f and P_l being different is recorded.
- The values of $\Phi_i (i = 1, 2)$ is adjusted from 0.0 to 1.0 for each bi-normal distribution, and the previous steps are repeated to obtain the corresponding difference numbers.

The left subfigure in Fig. 6 shows a one-to-one case with low variance. Table 1 displays difference numbers with varied values of Φ_1 and Φ_2 which represent probability of congestion in paths 1 and 2, respectively. Due to the low variance, the bi-normal distribution is close to the bi-delta distribution, and the result can be similarly analysed by an example in Section 2.1. The difference is that sample size is assumed to be sufficient in a previous example but is merely 100 in the test which is more common in practical applications. As we can see, when $\Phi_1 \leq 0.4$ and $\Phi_1 \geq 0.8$, difference number is equal to or close to zero, which means

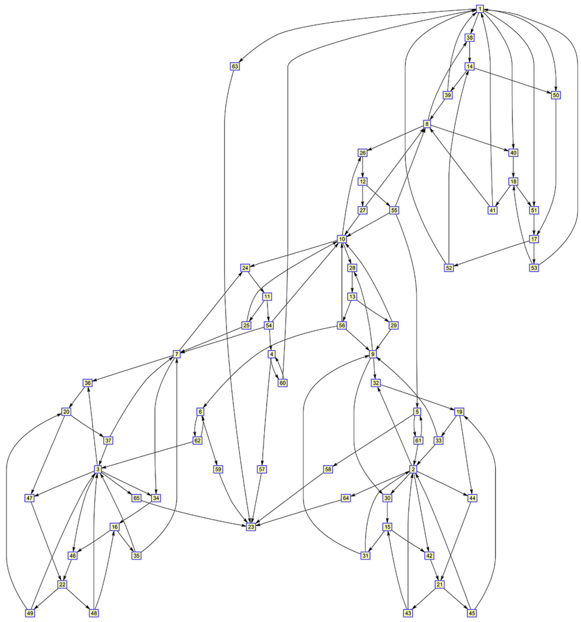

Fig. 5 *65-node, 123-arc artificial network*

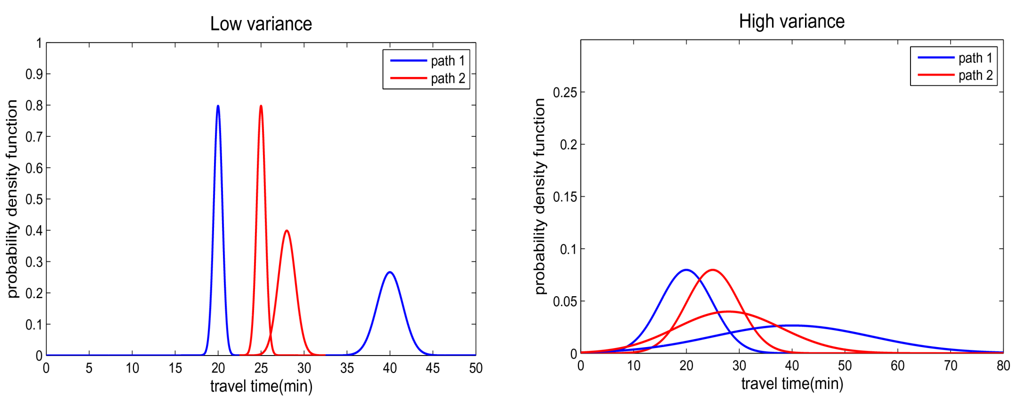

Fig. 6 *Bi-normal distributions in two one-to-one cases*

Table 1 Difference number with low variance

Φ_2	Probability for congestion situation of path 1 Φ_1 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	0	0	0	7	107	101	1	0	0
0.2	0	0	0	5	105	150	8	0	0
0.4	0	0	0	6	101	179	44	0	0
0.6	0	0	0	3	104	192	82	2	0
0.8	0	0	0	6	104	194	126	4	0
1.0	0	0	0	6	113	196	180	18	0

Table 2 Difference number with high variance

Φ_2	Probability for congestion situation of path 1 Φ_1 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	0	0	8	41	92	88	20	2	0
0.2	0	1	8	38	96	90	28	1	0
0.4	0	1	2	24	81	126	49	5	0
0.6	0	0	1	16	85	142	82	13	0
0.8	0	0	0	11	66	143	112	22	0
1.0	0	0	0	7	60	134	148	44	0

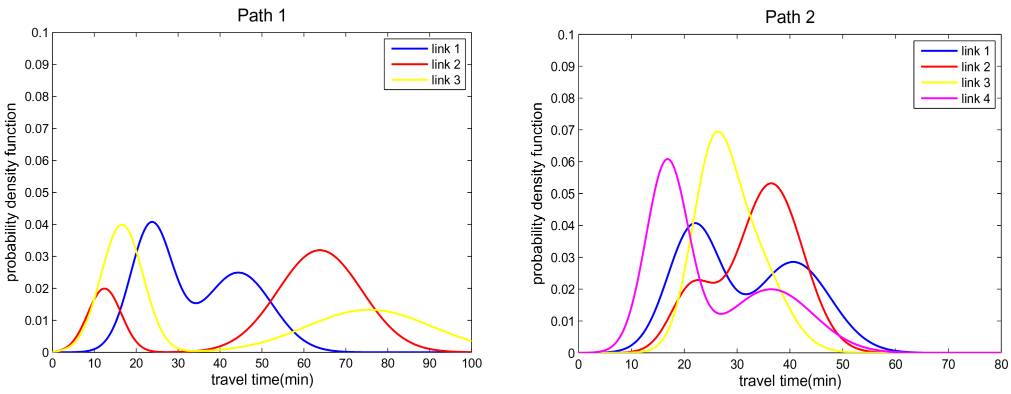

Fig. 7 *Bi-normal distributions in one 3 links-to-4 links case*

optimal paths determined by two criteria always keep the same. In these situations, Φ_1 is small or large enough to guarantee a path has large probability of less travel time than the other and meanwhile has less mean of travel time, or the opposite. Obviously, difference number appears large when $0.5 \le \Phi_1 \le 0.7$, and it is affected by both Φ_1 and Φ_2. When $\Phi_1 = 0.6$ and 0.7, the proposed criterion mostly regards path 1 as the optimal during 200 comparisons, but the changed Φ_2 lead to the reduction of the mean of travel times in path 2, so that the LET criterion gradually inclines to path 2 as the optimal and therefore difference number rises rapidly. The case of $\Phi_1 = 0.5$ can be analysed in the same way, where LET generally takes path 2 as the optimal, but the proposed criterion assigns almost the same chance (i.e. 50%) to two paths to be the optimal one.

To validate the impact of variance on difference number, the experiment is repeated with variance which is set ten times as much as the value used above. The bi-normal distributions are displayed as the right subfigure in Fig. 6 and Table 2 contains the test result. Evidently, Table 2 shares the similar pattern with Table 1 when $\Phi_1 \ge 0.6$, but its result distributes more widely along the values of Φ_1 considering the much more uncertainty caused by high variance. However, we find that the difference number reduces when $\Phi_1 \le 0.5$. In these situations, the probability calculated by (15) has changed because of the variance. So the faster criterion gradually selects path 2 as the optimal path when Φ_2 turns larger, which has the same trend as LET. In a word, we can conclude that in one-to-one cases, there are always some values of Φ_1 and Φ_2 corresponding to certain road conditions, which cause different choices of optimal roads determined by two criteria. It is wise for travellers to adopt the faster criterion in order to avoid the risk of a long travel time induced easily by LET.

5.1.2 Faster road in many-to-many cases: Based on Section 5.1.1, we extend the simple one-to-one case to many-to-many cases, taking more multiple-link roads into consideration. However, it is intractable to implement the test by the same method in simple cases, since the number of parameters significantly increases with the link number. We can no longer display the variation trend of the difference number with $\Phi_i (i = 1, 2, \ldots, N)$ in detail due to the excessive combinations of $\Phi_i (i = 1, 2, \ldots, N)$. To better reveal the distribution of difference number in each case, we randomly generate 200 groups of Φ_i and for each group, comparison between the faster criterion and LET is repeated 100 times to determine the difference number. The procedure is performed with respect to changing variance and different cases are verified. Fig. 7 shows one 3 links-to-4 links case, where we notice the bi-normal PDFs for each link differs from the ones in Fig. 6, because they have been computed by (12)

Table 3 shows the experimental results in four many-to-many cases. The difference number is counted recorded according to seven intervals to reveal its distribution. We can easily find that a similar pattern is shared by different cases, no matter how many links both roads contain. In addition, more conclusions can be made: (i) in each case, the most of the difference numbers is always 0, which means the most combinations of Φ_i (i.e. quantities of congestion situations) generate the same decision about optimal road by the faster criterion and LET; (ii) the distribution of difference number is remarkably reduced, as the interval rises and there are barely occurrences of difference number being larger than 40; (iii) with the increase of variance in bi-normal PDFs, difference number is generally increased and more difference numbers concentrate on 1–40, which means the optimal roads determined by two criteria become more distinct, and the result is more explicit

Table 3 The distribution of difference number in different many-to-many cases

Different cases with varied variance	Difference number intervals						
	0	1–10	11–20	21–30	31–40	41–50	51–100
3 links-to-4 links							
low variance	115	33	20	9	7	7	9
medium variance	102	71	27	28	5	1	0
high variance	95	48	27	20	10	0	0
5 links-to-6 links							
low variance	131	37	11	16	4	1	0
medium variance	122	36	25	17	0	0	0
high variance	89	56	37	18	2	0	0
7 links-to-7 links							
low variance	148	25	12	13	2	0	0
medium variance	134	35	16	15	0	0	0
high variance	111	46	25	18	0	0	0
8 links-to-8 links							
low variance	114	53	20	13	0	0	0
medium variance	99	50	26	24	1	0	0
high variance	64	72	37	23	4	0	0

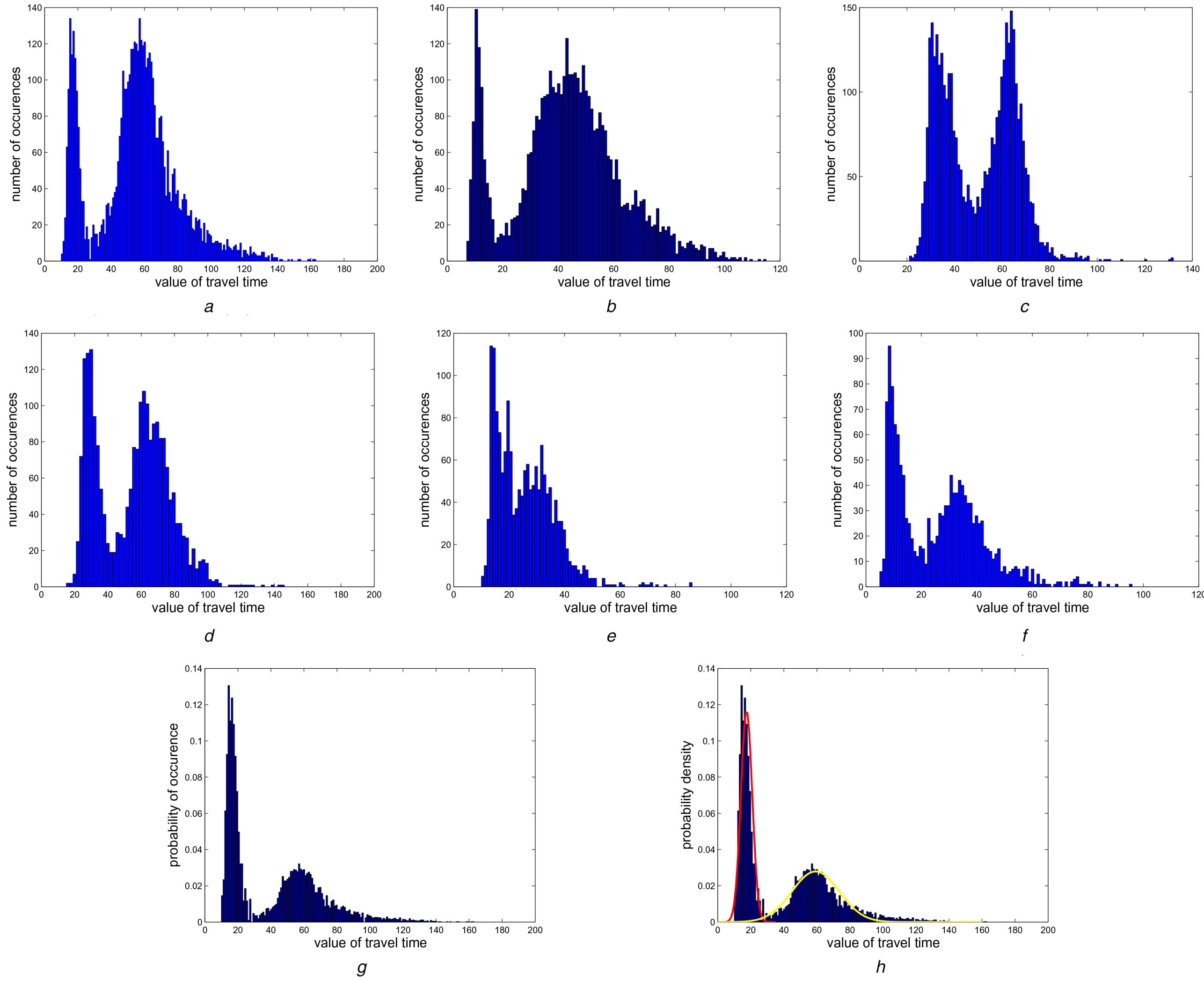

Fig. 8 *Histograms of travel times on some road links in Beijing*
(a) 5257 samples, Link ID:192086, ***(b)*** 4397 samples, Link ID:192087, ***(c)*** 3899 samples, Link ID:29450, ***(d)*** 2297 samples, Link ID:183085, ***(e)*** 1657 samples, Link ID:32174, ***(f)*** 1424 samples, Link ID:95029, ***(g)*** PMF of travel time on Link 192086, ***(h)*** Fitting PDF for travel time on link 192086

when comparison is carried out between roads with more links. In summary, there always exist some road conditions leading to clear distinction between the faster criterion and LET like10–40 differences in 100 comparisons, which is notable in practical applications. As the faster criterion is based on the historical observations of travel time data, it bears less risk of long travel time than LET. Therefore, the faster criterion is more reasonable and worth adopting.

5.2 Experiments on the real road network with real traffic data

Fig. 9 *Directed graph extracted from the real road network in Beijing*

Table 4 Difference number of optimal path by two criteria with real traffic data

Different cases in a real-world scenario	Sample size			
	100	200	500	1000
1 links-to-2 links	18.28	12.19	3.94	0.79
2 links-to-2 links	26.08	20.00	10.62	3.59
	49.20	47.79	59.70	71.68
	11.98	4.85	0.49	0.005
2 links-to-3 links	5.35	0.87	0	0
2 links-to-4 links	29.80	27.72	23.89	17.54
3 links-to-3 links	20.37	14.43	5.63	1.24
3 links-to-5 links	37.35	32.16	22.00	12.75
5 links-to-5 links	1.82	0.17	0	0
5 links-to-6 links	39.29	46.11	62.92	78.98

To better apply our theoretical results to the real road network and further conduct the contrast experiment of faster criterion and LET, firstly, we demonstrate bi-normal behaviours of real travel times on real road links. Herein, we use sampled trajectory data in the dataset given in [44]. The sampled trajectory data contains trajectories generated by over 30,000 taxicabs during one day in Beijing, which is recorded with four types of information, i.e. road segments ID, user ID, time slots, and travel time. Indeed, most link travel times in practice bear the bi-normal distribution, and we extract travel times on various links and display their histograms in Fig. 8 which readers can refer to.

There are six histograms shown in Fig. 8, and each of them approximates a diverse discrete bi-normal pattern, which reveal the bi-normal behaviour of link travel time in practice. Furthermore, as can be seen, the number of samples in each histogram is distinct, and it is clear that in Figs. 8*a–f*, the larger number of samples generates a smoother appearance. These changing cases of roughness rely on the character of the histogram, which is said that the statistical noise (i.e. roughness) of histogram is inversely proportional to the square root of sample quantity. As a consequence, we can deduce that with the number of samples increasing, the intact bi-normal curves will be obtained. In fact, we can convert the histogram into probability mass function (PMF), i.e. discrete PDF, and perfectly fit the PMF with two specified normal PDFs. As an example, Figs. 8*g* and *h* display the PMF and fitting normal PDFs of the histogram in Fig. 8*a*, in which the PMF of link travel time accords with two given normal PDFs obviously. Thus, the hidden bi-normal distribution of link travel time in a real-world scenario can be confirmed.

5.2.2 Contrast experiment in a real-world scenario: On the basis of bi-normal behaviour of link travel time, we further implement the contrast experiment of two criteria on the real road network with real traffic data, considering that more stochastic factors exist under real traffic conditions. To this end, we adopt the dataset in [44], which contains the information of road connection, to construct a directed graph corresponding to the real road network in Beijing. Fig. 9 displays a directed graph associated with part of the road network, in which the components are weakly connected and the paths can be derived. Using the directed graph, we find out various cases in the real-world scenario, and carry on the contrast experiment. The experimental result is given out in Table 4. It needs to be emphasised that, we do not set difference number intervals as we did in Section 5.1.2, since the probability of congestion is hidden in real travel times on each link (i.e. the parameters Φ_1 and Φ_2 are fixed). Instead, we show the average difference of optimal path by two criteria in 100 tests, when different number of samples are used. As can be seen, Table 4 contains various cases and each case corresponds to a contrast of two paths sharing the same origin and destination. It is clear that the difference of two criteria in the decision of the optimal path is fairly common, no matter how many links the two alternative paths have. In fact, the difference number is not correlated with link number, but the whole travel time of two paths. In some cases (e.g. 2 links-to-3 links and 5 links-to-5 links in Table 4), the whole travel times on two paths differ greatly, i.e. a path is always faster than the other, so there is almost no difference between two criteria. For other cases, the travel time distributions of two paths overlap, and there is chance that two criteria achieve different optimal paths. Even in the same case, like 2 links-to-2 links in Table 4, the difference number holds different patterns on each two paths due to various travel time distributions on them. Furthermore, the result shows that the difference of the optimal path by two criteria exists commonly, no matter how many samples we use in two criteria. When more samples are used, the optimal path determined by LET will tend to be fixed, which depends on the expected value of travel time on two paths. However, what matters in routing is not only the LET, but also the travel time variation. The decision by LET neglects the stochasticity of travel time in a real-world scenario. In contrast, the faster criterion focuses on probability comparison of the whole travel time on alternative paths, and reflects the travel time variation by samples from real travel times on each path, so it is more reliable than LET. Furthermore, the effect will be more evident when sample size rises, because the travel time distribution can be better approximated. As a consequence, the superiority of the faster criterion is also validated on the real road network with real traffic data and it is more reasonable and reliable than LET.

6 Conclusion and future works

This study aims at finding an optimal path under a faster criterion by exploring the traffic data, the essential part of which is performing a probabilistic comparison between two candidate paths. Regarding this comparison, a bi-model distribution of travel time for each road link is usually assumed. This study improves the faster criterion by extending the bi-delta distribution to the bi-normal distribution and deriving some theoretical models for both a simple path and a complex path. In addition, the result in Fig. 4 shows that as the road link number increases, it can generate a similar result if we approximate the bi-normal distribution of each road link by one normal distribution. The experimental result shows that, when the bi-normal distribution is significant on the road link, it is reasonable and beneficial to use the faster criterion to evaluate the paths. Moreover, via the experiment on a real road network with real traffic date, the criterion's superiority is also validated in practical terms, after the demonstration of the bi-normal distribution hidden in real link travel time is provided.

In the future, we would like to extend our work in the following aspects: first, more attention will be paid on how to directly determine the optimal path among all the alternatives instead of pair comparisons; second, for the result in Fig. 4, each road link is assumed following the same bi-normal distribution for the sake of computation simplification, we will improve it with random parameters; third, we will further implement experiments to compare with other routing criteria such as SSP methods.

7 References

[1] Isaac, G., Lange, T.: 'Split routing as a part of the urban navigation'. 19th ITS World Congress, 2012, pp. 1–8

[2] Li, S., Meng, M.-H., Chen, W., *et al.*: 'Sp-nn: a novel neural network approach for path planning'. IEEE Int. Conf. on Robotics and Biomimetics, 2007 (ROBIO 2007), 2007, pp. 1355–1360

[3] Le, M.K., Bhaskar, A., Chung, E.: 'Public transport travel-time variability definitions and monitoring', *J. Transp. Eng.*, 2015, **141**, pp. 04014068-1–04014068-9

[4] Montgromery, F., May, A.: 'Factors affecting travel times on urban radial routes', *Traffic Eng. Control*, 1987, **28**, (9), pp. 452–458

[5] Fan, Y., Nie, Y.: 'Optimal routing for maximizing the travel time reliability', *Netw. Spatial Econ.*, 2006, **6**, (3–4), pp. 333–344

[6] Lim, S., Balakrishnan, H., Gifford, D., *et al.*: 'Stochastic motion planning and applications to traffic'. Algorithmic Foundation of Robotics VIII, 2009, pp. 483–500

[7] Miller-Hooks, E.D., Mahmassani, H.S.: 'Least expected time paths in stochastic, time-varying transportation networks', *Transp. Sci.*, 2000, **34**, (2), pp. 198–215

[8] Dean, B.C.: 'Algorithms for minimum-cost paths in time-dependent networks with waiting policies', *Networks*, 2004, **44**, (1), pp. 41–46

[9] Mazloumi, E., Currie, G., Rose, G.: 'Using GPS data to gain insight into public transport travel time variability', *J. Transp. Eng.*, 2010, **136**, (7), pp. 623–631

[10] Cao, Z., Guo, H., Zhang, J., *et al.*: 'Multiagent-based route guidance for increasing the chance of arrival on time'. Proc. 30th AAAI Conf. on Artificial Intelligence, 2016

[11] Fan, Y., Kalaba, R., Moore, J.II: 'Arriving on time', *J. Optim. Theory Appl.*, 2005, **127**, (3), pp. 497–513

[12] Nie, Y.M., Wu, X., Homem-de Mello, T.: 'Optimal path problems with second-order stochastic dominance constraints', *Netw. Spatial Econ.*, 2012, **12**, (4), pp. 561–587

[13] Boyles, S.D., Waller, S.T.: 'A mean-variance model for the minimum cost flow problem with stochastic arc costs', *Networks*, 2010, **56**, (3), pp. 215–227

[14] Nikolova, E.: 'Approximation algorithms for reliable stochastic combinatorial optimization'. Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 2010, pp. 338–351

[15] Hutson, K.R., Shier, D.R.: 'Extended dominance and a stochastic shortest path problem', *Comput. Oper. Res.*, 2009, **36**, (2), pp. 584–596

[16] Celikoglu, H.B.: 'A dynamic network loading model for traffic dynamics modeling', *IEEE Trans. Intell. Transp. Syst.*, 2007, **8**, (4), pp. 575–583

[17] Celikoglu, H.B., Gedizlioglu, E., Dell'Orco, M.: 'A node-based modeling approach for the continuous dynamic network loading problem', *IEEE Trans. Intell. Transp. Syst.*, 2009, **10**, (1), pp. 165–174

[18] DellOrco, M., Marinelli, M., Silgu, M.A.: 'Bee colony optimization for innovative travel time estimation, based on a mesoscopic traffic assignment model', *Transp. Res. C, Emerg. Technol.*, 2016, **66**, (1), pp. 48–60

[19] Celikoglu, H.B.: 'A dynamic network loading process with explicit delay modelling', *Transp. Res. Emerg. Technol.*, 2007, **15**, (5), pp. 279–299

[20] Celikoglu, H.B.: 'Reconstructing freeway travel times with a simplified network flow model alternating the adopted fundamental diagram', *Eur. J. Oper. Res.*, 2013, **228**, (2), pp. 457–466

[21] Celikoglu, H.B.: 'Flow-based freeway travel-time estimation: a comparative evaluation within dynamic path loading', *IEEE Trans. Intell. Transp. Syst.*, 2013, **14**, (2), pp. 772–781

[22] Deniz, O., Aksoy, G., Celikoglu, H.B.: 'Analyzing freeway travel times within a case study: reliability of route traversal times'. Proc. 16th Int. IEEE Conf. on Intelligent Transportation Systems, 2013, pp. 195–202

[23] Celikoglu, H.B., Silgu, M.A.: 'Extension of traffic flow pattern dynamic classification by a macroscopic model using multivariate clustering', *Transp. Sci.*, 2016, **50**, (3), pp. 966–981

[24] Sun, L., Yang, J., Mahmassani, H.: 'Travel time estimation based on piecewise truncated quadratic speed trajectory', *Transp. Res. A, Policy Pract.*, 2008, **42**, (1), pp. 173–186

[25] Mori, U., Mendiburu, A., lvarez, M., *et al.*: 'A review of travel time estimation and forecasting for advanced traveller information systems', *Transportmetrica A, Transp. Sci.*, 2015, **11**, (2), pp. 119–157

[26] Celikoglu, H.B.: 'Dynamic classification of traffic flow patterns simulated by a switching multimode discrete cell transmission model', *IEEE Trans. Intell. Transp. Syst.*, 2014, **15**, (6), pp. 2539–2550

[27] Laporte, G., Louveaux, F., Mercure, H.: 'The vehicle routing problem with stochastic travel times', *Transp. Sci.*, 1992, **26**, (3), pp. 161–170

[28] Ta, D., Dellaert, N., Van Woensel, T., *et al.*: 'Vehicle routing problem with stochastic travel times including soft time windows and service costs', *Comput. Oper. Res.*, 2013, **40**, (1), pp. 214–224

[29] Miller Hooks, E.: 'Adaptive least expected time paths in stochastic, time varying transportation and data networks', *Networks*, 2015, **37**, (1), pp. 35–52

[30] Cao, Z., Guo, H., Zhang, J., *et al.*: 'Finding the shortest path in stochastic vehicle routing: a cardinality minimization approach', *IEEE Trans. Intell. Transp. Syst.*, 2016, **17**, (6), pp. 1688–1702

[31] Chien, S.I., Kuchipudi, C.M.: 'Dynamic travel time prediction with real-time and historic data', *J. Transp. Eng.*, 2003, **129**, (6), pp. 608–616

[32] Dong, W., Hai, L.V., Nazarathy, Y., *et al.*: 'Shortest paths in stochastic time-dependent networks with link travel time correlation', *Transp. Res. Rec. J. Transp. Res. Board*, 2013, **2338**, (1), pp. 58–66

[33] Cao, Z., Guo, H., Zhang, J., *et al.*: 'Maximizing the probability of arriving on time: a practical q-learning method'. Proc. 31th AAAI Conf. on Artificial Intelligence, 2017, pp. 4481–4487

[34] Fastenrath, U., Becker, M.: 'Process for selection of a route for a dynamic navigation on private transport'. German Patent: EP 2071287A2, 2008

[35] Sigal, C.E., Pritsker, A.A.B., Solberg, J.J.: 'The stochastic shortest route problem', *Oper. Res.*, 1980, **28**, (5), pp. 1122–1129

[36] Davidson, K.: 'A flow travel time relationship for use in transportation planning'. Proc. 3rd Australian Road Research Board (ARRB) Conf., Sydney, 1966, vol. **3**

[37] Allsop, R.E.: 'Some possibilities for using traffic control to influence trip distribution and route choice'. Proc. Transportation and Traffic Theory, 1974, vol. **6**

[38] Wenjie, C., Liqiang, G., Zhilei, C., *et al.*: 'An intelligent guiding and controlling system for transportation network based on wireless sensor network technology'. Proc. 5th Int. IEEE Conf. on Computer and Information Technology, 2005, pp. 810–814

[39] Mogridge, M., Fry, S.: 'Variability of car journey times on a particular route in central London', *Traffic Eng. Control*, 1984, **25**, (HS-038 005), pp. 510–511

[40] Rakha, H., El-Shawarby, I., Arafeh, M., *et al.*: 'Estimating path travel-time reliability'. Proc. Int. IEEE Conf. on Intelligent Transportation Systems, 2006, pp. 236–241

[41] Bennett, G.: 'Probability inequalities for the sum of independent random variables', *J. Am. Stat. Assoc.*, 1962, **57**, (297), pp. 33–45

[42] Bondesson, L.: '*Generalized gamma convolutions*' (Wiley Online Library, 1997)

[43] Goodman, N.: 'Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)', *Ann. Math. Stat.*, 1963, **34**, (1), pp. 152–177

[44] Wang, Y., Zheng, Y., Xue, Y.: 'Travel time estimation of a path using sparse trajectories'. Proc. ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, 2014, pp. 25–34