Efficient Solution Algorithm for Finding Spatially-dependent Reliable Shortest Path in Road Networks

Bi Yu Chen[a,b*], William H. K. Lam[b] and Qingquan Li[c,a]

[a]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China; [b]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong; [c]Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen, China

Abstract

Travel times are generally stochastic and spatially correlated in congested road networks. However, very few existing route guidance systems (RGS) can provide reliable guidance services to aid travelers planning their trips with taking account explicitly travel time reliability constraint. This study aims to develop such a RGS with particular consideration of travelers' concern on travel time reliability in congested road networks with uncertainty. In this study, the spatially-dependent reliable shortest path problem (SD-RSPP) is formulated as a multi-criteria shortest path-finding problem in road networks with correlated link travel times. Three effective dominance conditions are established for links with different levels of travel time correlations. An efficient algorithm is proposed to solve SD-RSPP by adaptively using three established dominance conditions. The complexities of road networks in reality are also explicitly considered. To demonstrate the applicability of proposed algorithm, a comprehensive case study is carried out in Hong Kong. The results of case study show that the proposed solution algorithm is robust to take account of travelers' multiple routing criteria. Computational results demonstrate that the proposed solution algorithm can determine the reliable shortest path on real-time basis for large-scale road networks.

Keywords: Reliable shortest path; spatial correlation; travel time reliability; route guidance systems

1. Introduction

Shortest path problems have been intensively studied owing to their broad applications (Fu et al., 2006). In the literature, substantial attention has been given to the development of efficient shortest path algorithms for the route guidance system (RGS) (Huang et al., 2007; Yang and Zhou, 2014; Zhou et al., 2014; Li et al., 2015). The values of RGS have been widely recognized by transportation researchers and practitioners that RGS not only can help travelers to make better route choice decisions but also can improve overall network traffic conditions (Li et al., 2012).

---

[*] To cite this paper: Chen, B.Y., Lam, W.H.K. and Li, Q.Q., 2016, Efficient solution algorithm for finding spatially dependent reliable shortest path in road networks. Journal of Advanced Transportation, 50, pp. 1413-1431.

Most existing RGSs assume that travel times in the road network are deterministic. In reality, travel times are highly stochastic, because of demand fluctuations and interruptions caused by traffic control devices, accidents and road construction, etc. In addition, adjacent link travel times are strongly correlated (Chen et al., 2012). For example, traffic accidents on a link may also lead to serious travel delays on its upstream links. It has been well recognized in the literature that travel time uncertainties have significant impacts on individual's route choice behaviors (Lam and Small, 2001; Bates et al., 2001; Carrion and Levinson, 2012; Taylor, 2013). In face of travel time uncertainty, travelers tend to budget additional travel times to guarantee a high probability of on-time arrival (i.e., travel time reliability in the literature). However, most of the existing RGS do not allow users to incorporate the concerns of travel time reliability in their route choice decisions. Responding to this need, this study aims to develop such a RGS with particular consideration of travel time reliability.

The shortest path problems considering travel time reliability are known as reliable shortest path problem (RSPP). The RSPP have been studied extensively, with the majority of the literature focused on a priori path problems. In his seminal work, Frank (1969) introduced the most reliable path concept by maximizing travel time reliability for a pre-determined travel time budget. Chen and Ji (2005) proposed the α-reliable path concept with an objective of minimizing travel time budget required to satisfy a travel time reliability constraint. Chen et al. (2013a) noted that the travel time reliability constraint can be pre-determined by travelers based on their trip purposes, and thus the α-reliable path concept can better reflect travelers' decision-making process in the RGS applications.

Due to the non-linear feature of their objective functions, RSPP cannot be solved by classical shortest path algorithms (e.g., Dijkstra's algorithm). In the literature, several heuristic algorithms have been developed to solve RSPP. Chen and Ji (2005) developed a genetic algorithm for searching both α-reliable path and most reliable path. This genetic algorithm was further extended to consider travel time correlations (Ji et al., 2011). However, these genetic algorithms can be computationally expensive and may not obtain the optimal solution. For finding the most reliable path with travel time correlations, Srinivasan et al. (2014) developed an approximation algorithm based on K-shortest path approach. Nevertheless, the approximation accuracy depends on the number of generated K shortest paths.

Many research efforts also have been given to develop exact algorithms for solving RSPP. Existing exact algorithms can be roughly classified into two categories, including decomposition based approach and multi-criteria based approach. The decomposition based approach decomposes original RSPP into a set of subproblems that can be efficiently solved as standard shortest path algorithms (e.g., Dijkstra's algorithm). To find α-reliable path with travel time correlations, Xing and Zhou (2011) proposed a Lagrangian relaxation algorithm based on a sample-based formulation. Zeng et al. (2015) extended the Lagrangian relaxation algorithm by representing travel time correlations based on a Cholesky decomposition. Khani and Boyles (2015) developed a labelling algorithm by repeatedly solving a sub-problem of minimizing the sum of mean and variance of path travel time. Although these decomposition based algorithms can efficiently solve RSPP in finite iterations of standard shortest path searches, they may occasionally failed to find the optimal path in large-scale networks (Khani and Boyles, 2015).

As an alternative, multi-criteria based approach directly solves RSPP by a single multi-criteria shortest path search, and can guarantee to obtain the optimal path. Nevertheless, the multi-criteria shortest path algorithms generally have a non-deterministic polynomial complexity in the worst case, since the number of non-dominated paths grows exponentially with network size. To find both α-reliable path and most reliable path, Nie and Wu (2009a) proposed a multi-criteria label-correcting algorithm based on the first-order stochastic dominance (FSD) condition. To efficiently solve the α-reliable path problem, Chen et al. (2013a) established more effective Mean-Variance (M-V) and Mean-Budget (M-B) dominance conditions to reduce the number of generated FSD non-dominated paths. Using the M-B condition, a multi-criteria A* algorithm was developed for efficiently finding the α-reliable path in large-scale road networks. Chen et al. (2012) further extended the multi-criteria A* algorithm to consider travel time correlations amongst adjacent K links. However, the M-B dominance condition is not validate when link travel times are negatively correlated. Based on the M-V dominance condition, a spatially-dependent multi-criteria A* algorithm (called SD-RSP-A* algorithm) was developed to solve α-reliable path with travel time correlations.

Built on the our priori work (Chen et al., 2012, 2013a, 2013b, 2014, 2016a), this study moves one step further to the implementation of a real-world RGS in Hong Kong. By integrating with real-time travel time distribution information, a prototype RGS in Hong Kong, called Reliable Path Searching System (RPSS), is developed for providing reliable route guidance services to road users in face of travel time uncertainties. The contributions of this study are summarized as below.

Firstly, a solution algorithm (called ERSP-A* algorithm) is proposed for efficiently finding the α-reliable path in large-scale road networks with travel time correlations. In this study, the stricter M-B dominance condition is established for adjacent links with positive correlations. The proposed ERSP-A* algorithm adaptively uses M-B condition for adjacent links with positive correlations, and M-V condition for adjacent links with negative correlations. The ERSP-A* algorithm is further extended to consider complexities of road networks in reality, including turn restrictions (e.g. no-right-turn at intersections), origins or destinations at links instead of nodes. Travelers' multiple routing criteria toward travel time reliability, travel distance and toll charge are also explicitly considered.

Secondly, a comprehensive case study is carried out to demonstrate the applicability of the developed RPSS. The results of case study indicate that travel times between adjacent links are strongly correlated, and ignoring travel time correlations can significantly underestimate path travel time variation. The developed RPSS is robust to take account of travelers' multiple routing criteria, including travel time reliability, travel distance and toll charge. Computational experiments showed that the proposed ERSP-A* algorithm can find the α-reliable path in large-scale networks within satisfactory computational times, and has an obvious computational advantage over the SD-RSP-A* algorithm (Chen et al., 2012).

The remainder of this paper is organized as follows. The next section describes the proposed solution algorithm for solving RSPP in real road networks. Section 3 introduces the architecture of the developed RPSS. Section 4 reports the experimental and computational results of RPSS. Section 5 gives conclusions and future research directions.

2. Spatially-dependent reliable shortest path problem

2.1. Problem statement

Let $\mathrm{G} = (\mathrm{N}, \mathrm{A}, \Psi)$ denote a directed network. N is a set of nodes, A is a set of links, and Ψ is a set of movements. A link $a_{ij} \in \mathrm{A}$ denotes a direct road segment from nodes i to j. Travel time of link a_{ij} (denoted by T_{ij}) is a random variable and could be represented by normal, gamma, log-normal or Burr distributions (Kaparias et al., 2008; Susilawati et al., 2013, Chen et al., 2016). Let t_{ij} and σ_{ij} be the mean and standard deviation of link travel time T_{ij} respectively. A movement $\psi_{ijk} = a_{ij} \oplus a_{jk} \in \Psi$ represents an allowed movement at node j (passing through successive links a_{ij} and a_{jk}); while a movement $\psi_{ijk} \notin \Psi$ means restricted at node j (e.g., no U-turn). $\Psi_{ij} \in \Psi$ represents the set of movements starting from link a_{ij}.

Let r and s be origin and destination (O-D) pair, and $P^{rs} = \{p_1^{rs}, ..., p_n^{rs}\}$ be a set of paths between the O-D pair. The travel time of a path p_u^{rs} (denoted by T_u^{rs}) can be expressed as the summation of related link travel times along the path

$$T_u^{rs} = \sum_{a_{ij} \in A} T_{ij} x_{ij}^{rs} \tag{1}$$

where x_{ij}^{rs} be the decision variable representing link-path incidence relationship, $x_{ij}^{rs} = 1$ indicates that a_{ij} is on p_u^{rs}, and $x_{ij}^{rs} = 0$ otherwise. Obviously, the path travel time T_u^{rs} is a random variable whose distribution is the joint distribution of all links along the path. Its mean and standard deviation respectively are denoted by t_u^{rs} and σ_u^{rs}.

In this study, it is assumed that path travel time follows normal distributions. Using real-world traffic data, recent empirical studies based on field observations showed that the use of normal distributions appears to reflect observed path travel time distributions (Rakha et al., 2006). Chen et al. (2016a) found that the normal distribution can well approximate path travel time distribution. The normal distribution approximation can respectively achieve 98.3% and 94.9% of accuracy at 10^{th} and 90^{th} percentiles. It is also assumed that travel time correlations are restricted only to adjacent links, in a similar way to the work of Waller and Ziliaskopoulos (2002), and Nie and Wu

(2009b). To some extent, this limited spatial dependence can be interpreted as Tobler's First Law of Geography that 'all things are related, but nearby things are more related than distant things' (Tobler, 1970). Empirical studies based on field observations also found travel times amongst adjacent links to be strongly correlated (Gajewski and Rilett, 2003; Chan et al., 2009; Chen et al., 2012); the correlation is usually very low for links that are spatially distant, even on the same street (El Esawey and Sayed, 2011). This assumption can be relaxed to consider adjacent K links by using two-level hierarchical network model proposed by Chen et al. (2012).

Under above two assumptions, the path travel time follows multivariate normal distributions. Its mean and standard deviation can be calculated as

$$t_u^{rs} = \sum_{a_{ij} \in A} t_{ij} x_{ij}^{rs} \tag{2}$$

$$\sigma_u^{rs} = \sqrt{\sum_{a_{ij} \in A} \left(\sigma_{ij}^2 x_{ij}^{rs} + 2\operatorname{cov}(a_{ij}, a_{jk}) x_{ij}^{rs} x_{jk}^{rs} \right)} \tag{3}$$

where $\operatorname{cov}(a_{ij}, a_{jk})$ is the covariance between adjacent links a_{ij} and a_{jk}. Let $\Phi_{T_u^{rs}}^{-1}(\alpha)$ be the inverse of the cumulative distribution function (CDF) of T_u^{rs}. It can be expressed as

$$\Phi_{T_u^{rs}}^{-1}(\alpha) = t_u^{rs} + z_\alpha \sigma_u^{rs} \tag{4}$$

where z_α is the inverse CDF of standard normal distribution at α confidence level. This confidence level $\alpha \in (0,1)$ also called travel time reliability in the literature. It represents the probability of travelers arriving at the destination within the travel time budget $\Phi_{T_u^{rs}}^{-1}(\alpha)$. The travel time reliability α can be pre-determined based on travelers' trip purposes (Chen et al., 2012). $\alpha > 0.5$, $\alpha = 0.5$ and $\alpha < 0.5$ respectively represent individual's risk-averse, risk-neutral and risk-seeking attitudes.

According to the α-reliable path concept, the spatially-dependent reliable shortest path problem (SD-RSPP) can be formulated as following minimization problem:

$$\min_{x_{ij}^{rs}} \quad \Phi_{T_u^{rs}}^{-1}(\alpha) = t_u^{rs} + z_\alpha \sigma_u^{rs} \tag{5}$$

Subject to

$$\sum_{a_{ij} \in A} x_{ij}^{rs} - \sum_{a_{ki} \in A} x_{jk}^{rs} = \begin{cases} 1 & \forall\, i = r \\ 0, & \forall\, i \neq r;\, i \neq s \\ -1 & \forall\, i = s \end{cases} \tag{6}$$

$$x_{ij}^{rs} \in \{0, 1\}, \quad \forall a_{ij} \in A \tag{7}$$

$$\psi_{ijk} \in \Psi, \quad \forall \psi_{ijk} \in p_u^{rs} \tag{8}$$

Eq. (5) is the objective function that individuals want to minimize. Eq. (6) ensures that the path is feasible. Eq. (7) indicates that link-path incidence variables should be binary. Eq. (8) is concern with feasibility of all movements along the path.

The above formulation of SD-RSPP can be regarded as a generalization of traditional shortest path problem. The optimal solution of SD-RSPP depends on pre-determined travel time reliability α. For risk-neutral scenarios ($\alpha = 0.5$), $z_{\alpha} = 0$ and thus SD-RSPP becomes the traditional shortest path problem to find the least mean travel time path. For risk-averse scenarios ($\alpha > 0.5$), travelers make route choice decisions by simultaneously minimizing t_u^{rs} and σ_u^{rs}. For risk-seeking scenarios ($\alpha < 0.5$), travelers tend to choose the optimal path with smaller t_u^{rs} but larger σ_u^{rs}.

2.2. Stochastic dominance conditions

The SD-RSPP can be formulated and solved as a multi-criteria shortest path problem. The multi-criteria shortest path algorithm typically builds upon dominance conditions for determining a set of Pareto-optimal (or non-dominated) paths. To effectively identify dominated paths, Chen et al. (2012) proposed a mean-variance (M-V) dominance condition by considering travel time correlations amongst adjacent K links.

Let $p_u^{r,ij} = p_u^{ri} \oplus a_{ij} \in P^{rj}$ denote a path going through link a_{ij}. When spatial correlations between adjacent links (i.e., K=1) are considered, the M-V dominance can be defined as below:

Definition 1. Given a travel time reliability α and two paths $p_u^{r,ij} \neq p_v^{r,ij} \in P^{rj}$, $p_u^{r,ij}$ dominates $p_v^{r,ij}$ (denoted by $p_u^{r,ij} \succ p_v^{r,ij}$), if and only if $\Phi^{-1}_{T_u^{rw}}(\alpha) < \Phi^{-1}_{T_v^{rw}}(\alpha)$ holds for $p_u^{rw} = p_u^{r,ij} \oplus p^{jw}$ and $p_v^{rw} = p_v^{r,ij} \oplus p^{jw}, \forall p^{jw} \in P^{jw}\ \forall w \in N$.

Proposition 1. (Mean-Variance dominance) Given a travel time reliability α and two paths $p_u^{r,ij} \neq p_v^{r,ij} \in P^{rj}$, $p_u^{r,ij} \succ p_v^{r,ij}$ if $p_u^{r,ij}$ and $p_v^{r,ij}$ satisfy either

(i) $t_u^{r,ij} \leq t_v^{r,ij}$ and $z_{\alpha}\sigma_u^{r,ij} < z_{\alpha}\sigma_v^{r,ij}$ or

(ii) $t_u^{r,ij} < t_v^{r,ij}$ and $z_{\alpha}\sigma_u^{r,ij} \leq z_{\alpha}\sigma_v^{r,ij}$.

Proof. See Proposition 3.2 in Chen et al. (2012). □

Let $p_u^{r,ijk} = p_u^{ri} \oplus a_{ij} \oplus a_{jk}$ denote a path going through the movement $\psi_{ijk} = a_{ij} \oplus a_{jk}$. If $\text{cov}(a_{ij}, a_{jk}) \geq 0$ holds, a stricter Mean-Budget (M-B) dominance condition is introduced in this study as follows:

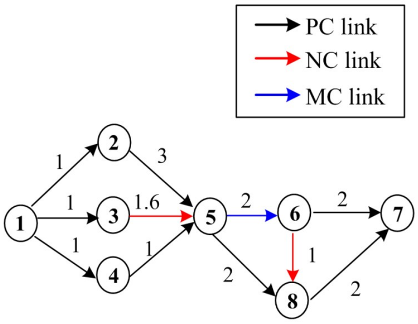
Proposition 2. (Mean-Budget dominance**)** Given a travel time reliability α and two paths $p_u^{r,ijk} \neq p_v^{r,ijk} \in P^{rk}$, $p_u^{r,ijk} \succ p_v^{r,ijk}$ if $p_u^{r,ij}$ and $p_v^{r,ij}$ satisfy $t_u^{r,ij} \leq t_v^{r,ij}$, $\Phi^{-1}_{T_u^{r,ij}}(\alpha) < \Phi^{-1}_{T_v^{r,ij}}(\alpha)$ and $\text{cov}(a_{ij}, a_{jk}) \geq 0$.

Proof. See Appendix 1.□

Figure 1 illustrates these two dominance conditions using a simple network. The mean travel times are given on the links. In the variance-covariance matrix, elements along the diagonal are the variance of link travel times; while off-diagonal elements are the covariance between two adjacent links. As the matrix is symmetric, only a lower triangular matrix is shown in the figure. There are three paths (i.e., $p_1^{1,56} = a_{12} \oplus a_{25} \oplus a_{56}$, $p_2^{1,56} = a_{13} \oplus a_{35} \oplus a_{56}$ and $p_3^{1,56} = a_{14} \oplus a_{45} \oplus a_{56}$) from Node 1 to Node 6 going through the same link a_{56}. According to Proposition 1, path $p_1^{1,56}$ is M-V dominated by path $p_2^{1,56}$ (since $t_2^{1,56} = 4.6 < t_1^{1,56} = 6$ and $\sigma_2^{1,56} = \sqrt{4} < \sigma_1^{1,56} = \sqrt{12}$) and thus the path $p_1^{1,56}$ can be eliminated at link a_{56}. Because $t_2^{1,56} = 4.6 > t_3^{1,56} = 4$ and $\sigma_2^{1,56} = \sqrt{4} < \sigma_3^{1,56} = \sqrt{6}$ hold, $p_2^{1,56}$ and $p_3^{1,56}$ are two M-V non-dominated paths and both of them should be maintained at link a_{56}.

When these two M-V non-dominated paths (i.e., $p_2^{1,56}$ and $p_3^{1,56}$) are extended to link a_{67}, the extended paths $p_2^{1,567} = p_2^{1,56} \oplus a_{67}$ and $p_3^{1,567} = p_3^{1,56} \oplus a_{67}$ are still M-V non-dominated paths at link a_{67} (since $t_3^{1,567} = 6 < t_2^{1,567} = 6.6$ and $\sigma_3^{1,567} = \sqrt{8} > \sigma_2^{1,567} = \sqrt{6}$ holds). However, according to Proposition 2, $p_2^{1,567}$ can be identified and eliminated as M-B dominated path at the link a_{67} (since $t_3^{1,56} = 4 < t_2^{1,56} = 4.6$, $\Phi^{-1}_{T_3^{1,56}}(0.9) = 7.14 < \Phi^{-1}_{T_2^{1,56}}(0.9) = 7.16$ and $\mathrm{cov}(a_{56}, a_{67}) = 0.5 \geq 0$). Thus, the M-B dominance can help identifying more dominated paths, which may not be identified by the M-V dominance.

As summarized in Chen et al. (2013a), for risk-averse travelers ($\alpha > 0.5$), M-V dominated paths satisfying $\Phi^{-1}_{T_u^{rj}}(\beta) < \Phi^{-1}_{T_v^{rj}}(\beta)\ \forall \beta \in [0.5,\ 1)$ can be discarded during the path finding process. Using M-B dominance, some M-V non-dominated paths satisfying $\Phi^{-1}_{T_u^{rj}}(\beta) < \Phi^{-1}_{T_v^{rj}}(\beta)\ \forall \beta \in [0.5,\ \alpha)$ can be further eliminated, so as to reduce the number of generated non-dominated paths, during the path search process. Similarly, for risk-seeking travelers ($\alpha < 0.5$), M-B dominance can eliminate dominated paths satisfying $\Phi^{-1}_{T_u^{rj}}(\beta) < \Phi^{-1}_{T_v^{rj}}(\beta)\ \forall \beta \in (\alpha,\ 0.5]$; while M-V dominance can only eliminate dominated paths satisfying $\Phi^{-1}_{T_u^{rj}}(\beta) < \Phi^{-1}_{T_v^{rj}}(\beta)\ \forall \beta \in (0,\ 0.5]$. Therefore, the established M-B dominance condition can reduce the number of non-dominated paths generated, evaluated, and stored when searching for the α-reliable path, so as to improve the efficiency of the search process.



	a_{12}	a_{13}	a_{14}	a_{25}	a_{35}	a_{45}	a_{56}	a_{58}	a_{67}	a_{68}	a_{87}
a_{12}	1										
a_{13}	0	1									
a_{14}	0	0	1								
a_{25}	1	0	0	1							
a_{35}	0	1	0	0	1						
a_{45}	0	0	1	0	0	1					
a_{56}	0	0	0	1	**-0.5**	0.5	1				
a_{58}	0	0	0	1	**-0.5**	0.5	0	1			
a_{67}	0	0	0	0	0	0	0.5	0	1		
a_{68}	0	0	0	0	0	0	**-1**	0	0	1	
a_{87}	0	0	0	0	0	0	0	1	0	**-1**	1

Figure 1. An illustration of dominance conditions

As described above, the path $p_2^{1,56}$ in Figure 1 (satisfying $t_3^{1,56} = 4 < t_2^{1,56} = 4.6$, $\Phi_{T_3^{1,56}}^{-1}(0.9) = 7.14 < \Phi_{T_2^{1,56}}^{-1}(0.9) = 7.16$) can be regarded as a potential M-B dominated path at link a_{56}. When a_{56} has positive correlation with its successor link a_{67}, this potential M-B dominated path is not necessary to be extended to the successor link a_{67}, since $p_2^{1,567} = p_2^{1,56} \oplus a_{67}$ is M-B dominated path. If a link has positive correlations with its all successor links, then following Lemma holds for identifying the M-B dominated path:

Lemma 1. Given a travel time reliability α and two paths $p_u^{r,ij} \neq p_v^{r,ij} \in P^{rj}$, $p_u^{r,ij} \phi\, p_v^{r,ij}$ if $p_u^{r,ij}$ and $p_v^{r,ij}$ satisfy $t_u^{rj} \le t_v^{rj}$, $\Phi_{rj,u}^{-1}(\alpha) < \Phi_{rj,v}^{-1}(\alpha)$ and $\mathrm{cov}(a_{ij}, a_{jk}) \ge 0 \;\; \forall \psi_{ijk} \in \Psi_{ij}$.

Proof. According to Proposition 2, $p_u^{r,ij} \oplus a_{jk}$ dominates $p_v^{r,ij} \oplus a_{jk}$ for any successor link a_{jk} ($\forall \psi_{ijk} \in \Psi_{ij}$). According to Definition 1, we have $\Phi_{T_u^{rw}}^{-1}(\alpha) < \Phi_{T_v^{rw}}^{-1}(\alpha)$ for paths $p_u^{r,ij} \oplus a_{jk} \oplus p^{kw}$ and $p_v^{r,ij} \oplus a_{jk} \oplus p^{kw}$, $\forall \psi_{ijk} \in \Psi_{ij}$, $\forall p^{kw}$. Thus, we have $\Phi_{T_u^{rw}}^{-1}(\alpha) < \Phi_{T_v^{rw}}^{-1}(\alpha)$ for $p_u^{r,ij} \oplus p^{jw}$ and $p_v^{r,ij} \oplus p^{jw}$, $\forall p^{jw}$. Therefore, $p_u^{r,ij}$ dominates $p_v^{r,ij}$ according to Definition 1. □

Lemma 1 is illustrated using the same example shown in Figure 1. There are three paths (i.e., $p_1^{1,58} = a_{12} \oplus a_{25} \oplus a_{58}$, $p_2^{1,58} = a_{13} \oplus a_{35} \oplus a_{58}$ and $p_3^{1,58} = a_{14} \oplus a_{45} \oplus a_{58}$) from Node 1 to Node 8 going through the same link a_{58}. As the link a_{58} has positive correlations with its all successor links, both $p_1^{1,58}$ and $p_2^{1,58}$ can be identified as M-B dominated paths and thus eliminated at link a_{58} (since $t_3^{1,58} = 4 < t_2^{1,58} = 4.6 < t_1^{1,58} = 6$ and $\Phi_{T_3^{1,58}}^{-1}(0.9) = 7.14 < \Phi_{T_2^{1,58}}^{-1}(0.9) = 7.16 < \Phi_{T_1^{1,58}}^{-1}(0.9) = 9.39$). Compared with Proposition 2, Lemma 1 can help identify M-B dominated paths in the early stage of path searching process (i.e., at link a_{58} instead of link a_{87}), so as to improve the efficiency of the search process.

According to the above three dominance conditions, network links can be classified into three categories: PC (positive covariance) links, NC (negative covariance) links and MC (mix covariance) links as following

- PC link: A link a_{ij} is said to be the PC link if $\text{cov}(a_{ij}, a_{jk}) \geq 0$ holds for its any successor link a_{jk} ($\forall \psi_{ijk} \in \Psi_{ij}$).
- NC link: A link a_{ij} is said to be the NC link if $\text{cov}(a_{ij}, a_{jk}) < 0$ holds for its any successor link a_{jk} ($\forall \psi_{ijk} \in \Psi_{ij}$).
- MC link: A link a_{ij} is said to be the MC link if $\text{cov}(a_{ij}, a_{jk})$ is either positive or negative for its successor links.

Different dominance conditions can be applied for these three types of links. For the NC link (e.g., a_{35} or a_{68} in Figure 1), the M-B dominance is not valid; and all M-V non-dominated paths have to be maintained at the link according to Proposition 1. For MC link (e.g., a_{12}), all M-V non-dominated paths also have to be maintained at the link, but the potential M-B dominated paths (e.g., $p_2^{1,56}$) are no need to be extended to their successor links with positive correlations. For the PC link (e.g., a_{58}), only M-B non-dominated paths (e.g., $p_3^{1,58}$) need to be maintained at the link, while M-B dominated paths (e.g., $p_2^{1,58}$) can be eliminated according to Lemma 1.

2.3. Solution algorithm

Using above established dominance conditions, a link-based multi-criteria A* algorithm (called ERSP-A* algorithm) is proposed to solve SD-RSPP. Similar to multi-criteria A* algorithm (Chen et al., 2013a), the ERSP-A* algorithm utilizes a heuristic value function $F(p_u^{r,ij}) = \Phi^{-1}_{T_u^{r,ij}}(\alpha) + h(j)$ for labelling path $p_u^{r,ij}$, where $h(j)$ is estimated cost from node j to the destination, and $h(s) = 0$ at the destination. By using this $F(p_u^{r,ij})$ function instead of $\Phi^{-1}_{T_u^{r,ij}}(\alpha)$, a higher priority can be assigned to links closer to the destination, so as to reduce the number of examined links and speed up the search process.

In the proposed ERSP-A* algorithm, each link a_{ij} maintains a set of non-dominated paths $P^{r,ij} = \{p_1^{r,ij}, \ldots, p_m^{r,ij}\}$ which are sorted in ascending order by mean travel time $t_u^{r,ij}$. Non-dominated paths from all links are maintained in a scan eligible set, denoted by $SE = \{p_u^{r,ij}, \ldots, p_v^{r,kw}\}$. As a label-setting technique is adopted, the non-dominated paths in SE are ordered by increasing value of $F(p_u^{r,ij})$. At each iteration, path $p_u^{r,ij}$ with minimum $F(p_u^{r,ij})$ (i.e., at the top of SE) is selected for path extension. Only for each allowed movement $\psi_{ijk} \in \psi_{ij}$, a temporary path (denoted by $p_u^{r,ijk} := p_u^{r,ij} \oplus a_{jk}$) is constructed by extending $p_u^{r,ij}$ to the successor link a_{jk}. All restricted turns $\forall \psi_{ijk} \notin \psi_{ij}$ are excluded in the path searching process. At successor link a_{jk}, the dominance relationship between the newly generated path $p_u^{r,ijk}$ and the existing non-dominated path set $P^{r,jk}$ is then determined. If $p_u^{r,ijk}$ is a non-dominated path at link a_{jk}, it is inserted into $P^{r,jk}$ and SE. The newly generated path $p_u^{r,ijk}$ may also dominate a set of paths in $P^{r,jk}$ (denoted by $P_D^{r,jk}$). All dominated paths in $P_D^{r,jk}$ are eliminated from $P^{r,jk}$ and SE. The algorithm continues this process until SE becomes empty or the destination is reached.

To determine the dominance relationship between the newly generated path $p_u^{r,ijk}$ and the existing non-dominated path set $P^{r,jk}$, the proposed ERSP-A* algorithm adaptively uses three dominance conditions according to the type of link a_{jk}. If a_{jk} is a NC link, then all M-V non-dominated paths are kept at the link for further path extensions. If a_{jk} is a PC link, then all M-B dominated paths can be eliminated at the link without further consideration. These two scenarios are implemented in *CheckDominance* procedure. If a_{jk} is a MC link, then all M-V non-dominated paths have also to be kept at the link. However, in this scenario, the potential M-B dominated paths (denoted by $MBD = true$) can be determined using *CheckMBDominance* procedure. There is no need for these potential M-B dominated paths to extend to the successor links a_{jk} with $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$. The detailed steps of the ERSP-A* algorithm are given as below.

Algorithm: ERSP-A*
Inputs: O-D pair (r, s) and travel time reliability α
Returns: the reliable shortest path
Step 1. Initialization:
  For each link a_{rj} emanating from origin node r
    Create a path $p_u^{r,rj}$ at link a_{rj} and calculate $h(j)$ and $F(p_u^{r,rj})$.
    Set $P^{r,rj} := \{p_u^{r,rj}\}$ and $SE := SE \cup \{p_u^{r,rj}\}$.
  End for
Step 2. Path selection:
  If $SE = \phi$, then Stop; otherwise, continue.
  Select $p_u^{r,ij}$ at the top of SE and set $SE := SE \setminus \{p_u^{r,ij}\}$.
  If $j = s$, then Stop; otherwise continue.
Step 3. Path extension:
  If link a_{ij} is MC link, Call procedure $MBD :=$ CheckMBDominance($p_u^{r,ij}$).
  For every movement $\psi_{ijk} \in \Psi_{ij}$
    If link a_{ij} is a MC link
      If $MBD = true$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$, then scan next movement; otherwise, continue.
    End if
    Generate a new path $p_u^{r,ijk} := p_u^{r,ij} \oplus a_{jk}$ and calculate $h(k)$ and $F(p_u^{r,ijk})$.

Call procedure $P_D^{r,jk} := CheckDominance(p_u^{r,ijk}, P^{r,jk})$.

If $p_u^{r,ijk}$ is a non-dominated path, then set $SE := SE \cup \{p_u^{r,ijk}\}$ and $SE := SE \setminus P_D^{r,jk}$.

End for

Goto Step 2.

Procedure: *CheckMBDominance*

Inputs: A path $p_u^{r,ij}$ at link a_{ij}

Returns: MBD indicating whether or not $p_u^{r,ij}$ is a potential M-B dominated path

Set $v := 1$.

While $v \le |P^{r,ij}|$ and $t_u^{r,ij} \ge t_v^{r,ij}$ ($|P^{r,ij}|$ is the number of paths in $P^{r,ij}$)

If $\Phi_{T_u^{r,ij}}^{-1}(\alpha) > \Phi_{T_v^{r,ij}}^{-1}(\alpha)$, then return $MBD := true$.

Set $v := v+1$.

End while

Return $MBD := false$.

Procedure: *CheckDominance*

Inputs: A path $p_u^{r,ijk}$ and a set of non-dominated path $P^{r,jk}$

Returns: $P_D^{r,jk}$ storing the set of paths dominated by $p_u^{r,ijk}$, and update $P^{r,jk}$

Step 1: Initialization

If link a_{jk} is a PC link, then $\beta := \alpha$.

If link a_{jk} is a NC link or a MC link, then

If $\alpha > 0.5$, then $\beta := 0.999$.

If $\alpha = 0.5$, then $\beta := 0.5$.

If $\alpha < 0.5$, then $\beta := 0.001$.

End If

Set $P_D^{r,jk} := \varphi$ and $v := 1$.

Step 2: Dominance relationship determination

While $v \le |P^{r,jk}|$ and $t_u^{r,ijk} \ge t_v^{r,jk}$ ($|P^{r,jk}|$ is the number of paths in $P^{r,jk}$)

If $\Phi_{T_u^{r,ijk}}^{-1}(\beta) > \Phi_{T_v^{r,jk}}^{-1}(\beta)$, then return $P_D^{r,jk}$.

Set $v := v+1$.

End while

Insert $p_u^{r,ijk}$ into $P^{r,jk}$ at v^{th} position and set $v := v+1$ (by default $|P^{r,jk}| := |P^{r,jk}|+1$).

While $v \le |P^{r,jk}|$ and $\Phi_{T_u^{r,ijk}}^{-1}(\beta) < \Phi_{T_v^{r,jk}}^{-1}(\beta)$

Set $P^{r,jk} := P^{r,jk} \setminus \{p_v^{r,jk}\}$ and $P_D^{r,jk} := P_D^{r,jk} \cup \{p_v^{r,jk}\}$.

Set $v := v+1$.

End while

Return $P_D^{r,jk}$.

If the following inequality is satisfied, the heuristic function is admissible

$$F(p_u^{r,ijk}) = \Phi_{T_u^{r,ijk}}^{-1}(\alpha) + h(k) \geq F(p_u^{r,ij}) = \Phi_{T_u^{r,ij}}^{-1}(\alpha) + h(j) \tag{9}$$

Eq. (9) means that the $F(p_u^{r,ij})$ value monotonically increase with path extension. The Euclidean distance function $h(j) = e_{js} / v_{\max}$ is a commonly used admissible $h(j)$; where e_{js} is Euclidean distance from node j to the destination, and $v_{\max}$ is the design speed in road network. It can be proved that the ERSP-A* algorithm can obtain the optimal solution for SD-RSPP as follows.

Proposition 3. If the used heuristic function is admissible, the proposed ERSP-A* algorithm can obtain the optimal solution when the destination was reached.
Proof. See Appendix 1. □

The complexity of the proposed ERSP-A* algorithm is analysed. In the worst case, the algorithm generates $|A||P|$ non-dominated paths in SE, where $|A|$ is link number in the network and $|P|$ is the maximum number of generated non-dominated paths at a link. With the implementation of SE using F-heap data structure (Fredman and Tarjan, 1987), the path selection step requires $O(|A||P| Log(|A||P|))$. As both *CheckMBDominance* and *CheckDominance* procedures require $O(|P|)$, the path extension step requires $O((|A| + |\Psi|)|P|^2)$, where $|\Psi|$ is the number of allowed movements in the network. Therefore, the complexity of ERSP-A* algorithm is $O((|A| + |\Psi|)|P|^2 + |A||P| Log(|A||P|))$. Theoretically, $|P|$ grows exponentially with the network size and the ERSP-A* algorithm has a non-polynomial complexity. In practice, $|P|$ is much smaller than its maximum size, especially for transportation networks (Nie and Wu, 2009a; Chen et al., 2013a).

To provide route guidance services in real RGS, several complexities of road networks should be explicitly considered. For instance, origins or destinations at links instead of nodes. Travelers' multiple routing criteria toward travel distance and toll charge also should be formulated. These issues are addressed in this study and given in Appendix 2.

3. System architecture of reliable path searching system

Using the proposed ERSP-A* algorithm, this section presents the architecture of the developed Reliable Path Searching System (RPSS). Figure 2 illustrates its concept framework. As shown in the figure, RPSS comprises of four key components: Service and Application Provider (SAP), Content and Data Provider (CDP), Website Portal (WP) and Clients.

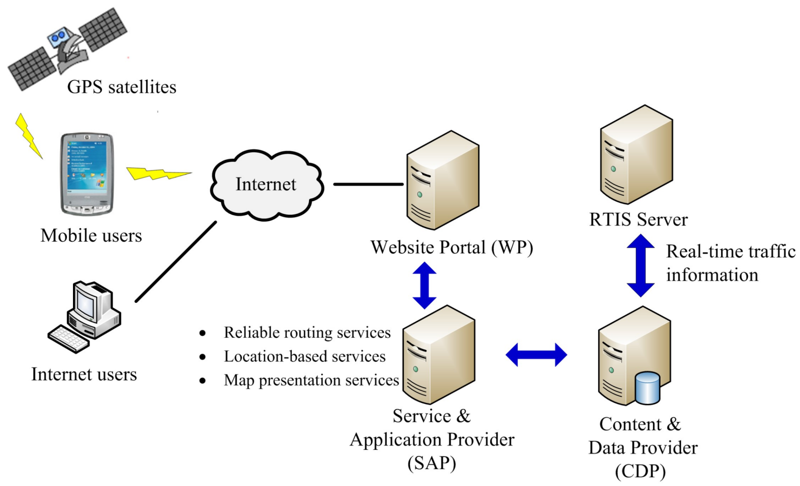


Figure 2. System architecture of RPSS prototype

The SAP acts as an application server for proceeding and answering core service requests. The core services implemented in the RPSS include location-based services (LBS), reliable routing services, and map presentation services. The LBS provides users with capabilities to find the nearest or a specific point of interesting (POI) (e.g. "Where is The Hong Kong Polytechnic University (PolyU)?", "Where is the nearest hotel to PolyU?"). As formulated in Section 2, the reliable routing services can determine the optimal route between a specified O-D pair, satisfying travelers' multiple routing criteria toward travel time reliability, travel distance and toll charge. The O-D pair can be determined through a search of LBS, or acquired by GPS (global positioning system) or user inputs. The map presentation services provide a way to render outputs of the LBS and the reliable routing services on geographical maps.

In this study, the SAP is implemented using a Web Service technique, which is an open industrial standard to allow interoperate across programming languages, platforms and operating systems. The Web Service offers protocols and mechanisms for rapidly publishing, discovering, and invocating services via internet. With the capabilities of Web Service, the SAP can provide a flexible and loosely-integrated platform not only for the implementation of routing guidance but also for the development of other applications (e.g. logistic applications).

The CDP maintains the needed data in RPSS, including GIS (geographical information system) maps and traffic information. The GIS maps, used for LBS and map presentation services, consist of a base map, a road network map and a set of POI maps (e.g. hotels, restaurants and banks, etc.). The real-time and historical traffic information from Real-time Traveler Information System (RTIS) (Tam and Lam, 2008) are used for reliable routing services. Historical traffic information in the same date of the week are collected from RTIS to generate off-line hourly travel time distributions. For each date of the week, 24 datasets of hourly travel time distributions are maintained and used as long-term predictions of traffic conditions. The routing services can adaptively choose the corresponding hourly travel time distributions according to the users' preferred arrival time (or departure time). To provide real-time routing services in the current system hour, real-time travel time distributions from RTIS are updated once every five minutes from RTIS.

The WP provides interfaces for users to access the services through the internet or mobile phones. When users request a service, the WP will pass the request to the SAP, and then transfer the SAP processing results to the user. Figure 3 illustrates a typical user interface of internet users to access the reliable routing services. As shown in Figure 3, the O-D pair can be set by entering their addresses or the names of POI. The on-time arrival probability (travel time reliability constraint α), the VOT and VOD parameters can also be inputted in this interface. By entering the preferred arrival time, the reliable routing services can determine the latest departure time and the associated route to users.

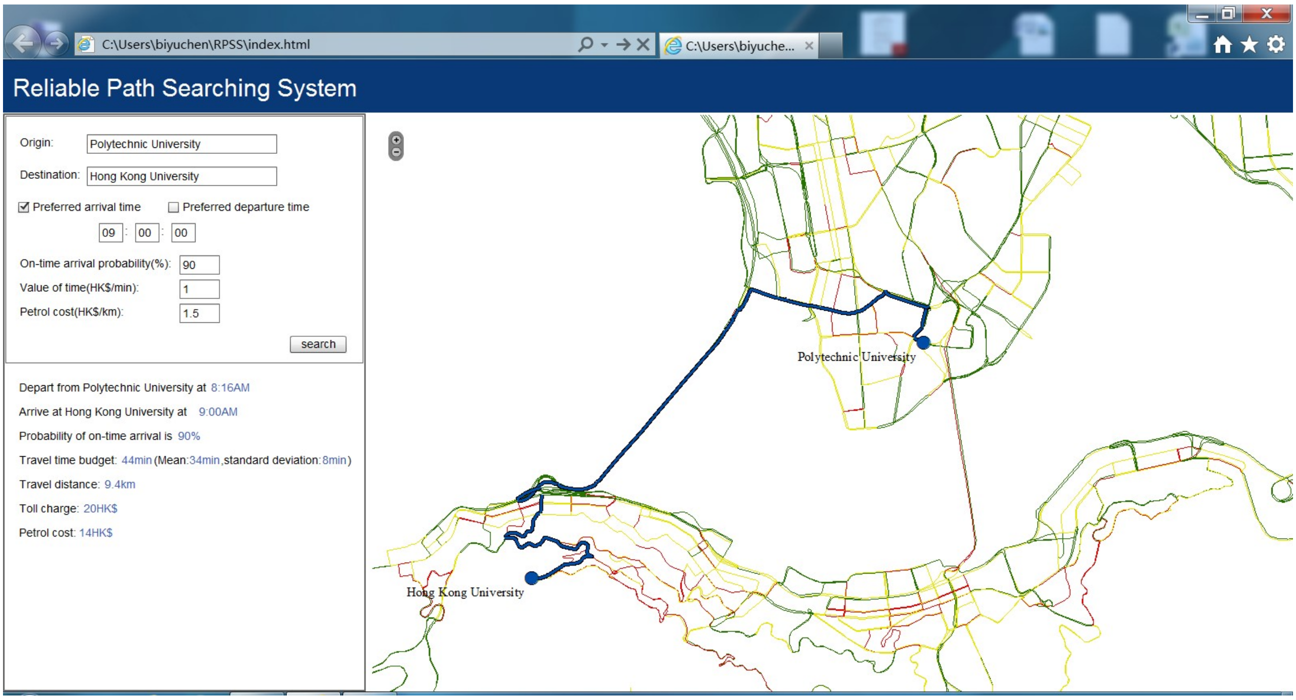


Figure 3. User interface for reliable routing service

4. Case Study

4.1. Data description

This section reported the numerical results of developed RPSS prototype. As can be seen in Figure 4, the Hong Kong RTIS network comprises 1,367 nodes, 3,655 links and 11,849 movements. The travel time information (including mean and variance-covariance matrix of link travel times) were collected from RTIS during a typical morning peak hour period (08:00-09:00) on 23 Sep 2010 (Thursday).

Figure 4 shows the travel time distributions of all links in the RTIS network. Figure 4 (a) illustrates the mean link travel times. Links shown in red represent congested links (< 15 km/h); yellow represents slightly congested links (15–30 km/h); and green represents uncongested links (> 30 km/h). It was found that 19.8% of links in the RTIS network were congested. Most of congested links were located in Hong Kong Island and some of them in Kowloon Urban Area.

Figure 4(b) shows travel time variation in the RTIS network. In this study, link travel time variation was measured by coefficient of variation (CV), which is the ratio of the standard deviation to the mean as

$$cv_{ij} = \sigma_{ij} / t_{ij} \tag{10}$$

As can be seen from the figure, travel times in the RTIS road network were highly stochastic, with the average CV value was equal to 0.35. Therefore, the travel time uncertainty should be necessarily considered in RGS.

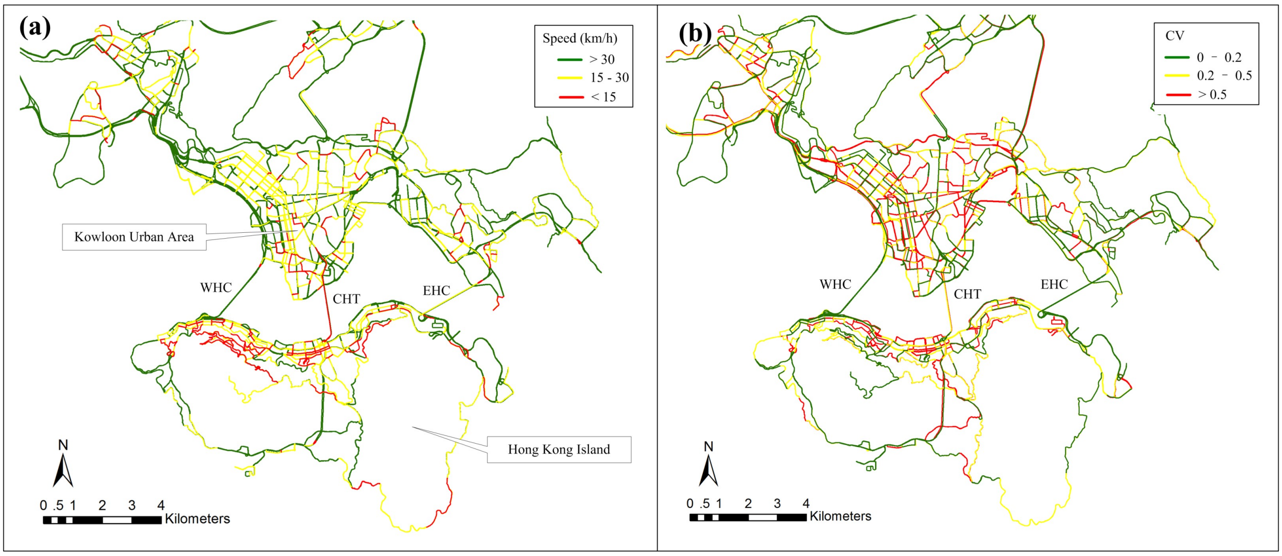


Figure 4. Travel time distributions in RTIS network (a) mean speeds (b) travel time variation

To normalize link travel time covariances in the network, a correlation coefficient for each pair of adjacent links was calculated as

$$\rho_{ijk} = \mathrm{cov}(a_{ij}, a_{jk}) / \sigma_{ij}\sigma_{jk} \tag{11}$$

The ρ_{ijk} value ranges between -1 and $+1$. $\rho_{ijk} = +1$ is the scenario of perfect positive correlation, while $\rho_{ijk} = -1$ is the scenario of perfect negative correlation. Figure 5 gives the correlation coefficients of all adjacent RTIS links. As shown in the figure, a considerable number of links were strongly correlated and the average ρ_{ijk} was 0.29. Therefore, the travel time correlations amongst adjacent links should not be ignored in reliable shortest path finding algorithms.

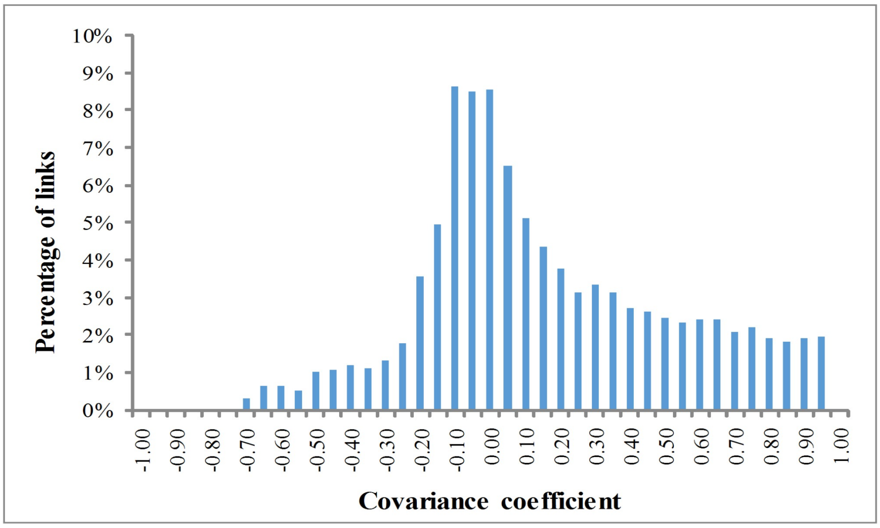


Figure 5. Histogram of correlation coefficient between adjacent links

It can also be seen from Figure 5 that travel times between adjacent links were correlated either positively or negatively. Amongst them, 70.3% of covariance values (i.e., $\text{cov}(a_{ij}, a_{jk})$) were positive. This was positive evidence supporting the use of M-B dominance to eliminate dominated paths. The proposition for PC, NC and MC links were 34.2%, 0.7%, and 65.1% respectively. This low proposition of NC links is good for developing efficient ERSP-A* algorithm by using the M-B dominance.

4.2. Computational performance

The computational performance of the proposed ERSP-A* algorithm is examined in this section. The ERSP-A* algorithm was coded in Visual C# programming language. The F-heap data structure (Fredman and Tarjan, 1987) was employed in the algorithm as the priority queue (i.e., SE). The link-based adjacent list data structure (Gutierrez and Medaglia, 2008) was adopted for loading network data into the memory. The used $h(j)$ was the Euclidean distance function. The SD-RSP-A* algorithm proposed by Chen et al. (2012) was also implemented for comparison study. All experiments were conducted on a MacBook Air laptop with a four-core Intel i7 CPU running at 2.0 GHz (only one core was used) and the Windows 7 operating system.

The computational tests were conducted on four networks (see Table 1). For Hong Kong RTIS network, traffic information were collected as described in Section 4.1. For other three networks, the mean and variance-covariance matrix of link travel times were randomly generated. Firstly, the link speeds were uniform distributed from [10 km/hour, 100 km/hour] to generate the mean travel time by calculating the link length ratio to link speed. Secondly, the link travel times SD was generated by randomly selecting the CV value (see Eq. (13)) from a uniform distribution in the range of [0.1, 1]. Thirdly, the correlation coefficient of each adjacent link pair was uniformly distributed from [-1, 1] to calculate link travel times covariance according to Eq. (14). Finally, the propositions of three types of links were checked. Based on the link propositions in the RTIS network, the minimum propositions of PC, NC and MC links were set as 30%, 0.5% and 50% respectively.

Table 1. Basic characteristics of testing networks

Road networks				Grid networks			
Network	Nodes	Links	Movements	Network	Nodes	Links	Movements
Hong Kong RTIS	1,367	3,655	11,849	G1 (40*50)	2,000	7,820	30,748
Chicago Regional	12,982	39,018	119,326	G2 (50*100)	5,000	19,700	77,908

Table 2 reports computational performance of the ERSP-A* and SD-RSP-A* algorithms. Reported results were the average of 100 runs using randomly generated O-D pairs for every network. In these tests, $\alpha = 0.8$ was set. It should be noted that these two algorithms can obtain identical reliable shortest path for every O-D pair.

As can be seen from the table, the ERSP-A* algorithm performed substantially better than the existing SD-RSP-A* algorithm for all networks. For example, the ERSP-A* algorithm was 2.87 (36.4/12.7) times as fast as the SD-RSP-A* algorithm to determine reliable shortest paths in the RTIS network. This is because the SD-RSP-A* algorithm using only M-V dominance generates a large amount of unnecessary paths. The SD-RSP-A* algorithm generated 4,678 M-V non-dominated paths. However, 46.9% of these M-V non-dominated paths can be eliminated as M-B dominated paths when compared with the results of the proposed ERSP-A* algorithm (2,484). These results clearly showed that the use of M-B dominance at PC and MC links can significantly reduce the number of generated non-dominated paths in the search process so as to improve the computational efficiency. The computational advantage of the proposed ERSP-A* algorithm was also evidenced in other three networks. The proposed ERSP-A* algorithm was 1.66, 3.08 and 2.32 times as fast as the SD-RSP-A* algorithm in the G1, G2 and Chicago Regional networks respectively.

Table 2. Computational performance of two reliable shortest algorithms

Network	SD-RSP-A* algorithm		ERSP-A* algorithm	
	Computational times (milliseconds)	Number of non-dominated paths	Computational times (milliseconds)	Number of non-dominated paths
Hong Kong RTIS	**36.4**	4,678	**12.7**	2,484
G1	43.1	4,026	25.9	2,912
G2	469.8	16,263	152.3	9,419
Chicago Regional	764.7	28,974	329.6	18,043

Figure 6 shows the computational performance of two testing algorithms in the RTIS network under various values of travel time reliability (i.e., α parameter). As shown in the figure, the proposed ERSP-A* algorithm performs significantly better than the SD-RSP-A* algorithm for risk-averse scenarios (i.e., $\alpha \in [0.5, 0.99]$). The computational advantage of the ERSP-A* algorithm was more obvious when α value approaches to 0.5. For example, when $\alpha = 0.95$, the ERSP-A* algorithm was 1.61 (29.8/18.5) times as fast as the SD-RSP-A* algorithm. This computational advantage was improved to 4.29 times when $\alpha = 0.5$. This is because the M-B dominance condition used by the ERSP-A* algorithm becomes more effective to eliminate dominated paths (according to $\Phi^{-1}_{rij,u}(\beta) < \Phi^{-1}_{rij,v}(\beta)$ for any confidence level $\beta \in [0.5, \alpha]$ (Chen et al., 2013a)). In the case of $\alpha = 0.5$, only one least mean travel time path is maintained at each PC link. Thus, the ERSP-A* algorithm runs fastest (i.e., 10.1 milliseconds) in this scenario by generating the least number of non-dominated paths (i.e., 2,180 paths). Conversely, the SD-RSP-A* algorithm uses the M-V dominance condition (according to $\Phi^{-1}_{rij,u}(\beta) < \Phi^{-1}_{rij,v}(\beta)$ for any confidence level $\beta \in [0.5,1)$ (Chen et al., 2013a)) cannot benefit from the decrease of input α value.

It can be seen from the figure that the ERSP-A* algorithm also has a significant computational advantage over the SD-RSP-A* algorithm in risk-seeking scenarios (i.e., $\alpha \in [0.1, 0.5]$). For example, when $\alpha = 0.4$, the ESP-A* algorithm consumed 12.5 milliseconds which was 7.02 times (87.8/12.5) as fast as the SD-RSP-A* algorithm. Therefore, the proposed ERSP-A* algorithm can efficiently find the α-reliable path in the RTIS network under various travel time reliability constraints (within 30 milliseconds). This result was evident that the developed reliable routing services can be applicable for the online routing applications in the large-scale road network of Hong Kong.

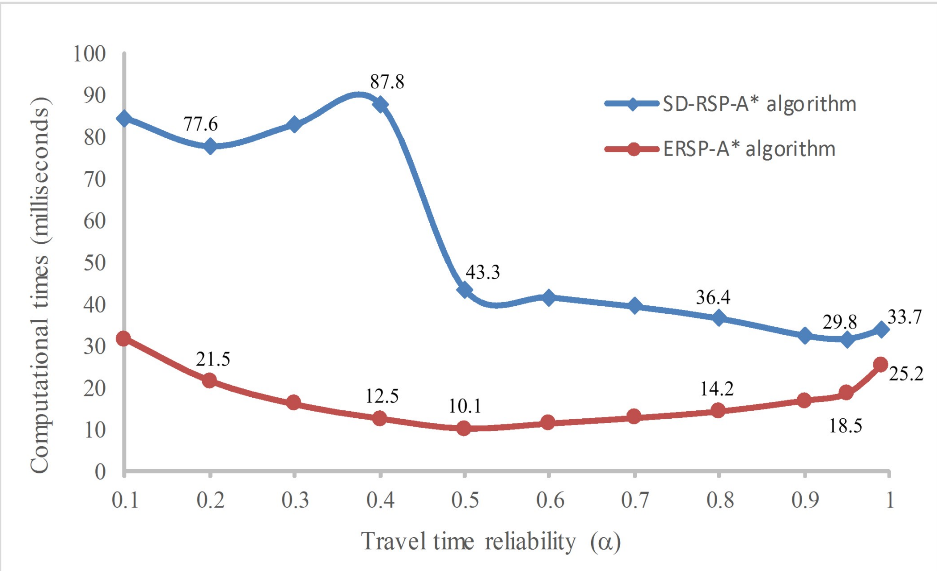


Figure 6. Computational performance of algorithms under various travel time reliability values

4.3. Numerical example

In Hong Kong, the crossing harbor traffic is one of the most important travel patterns. As illustrated in Figure 7, there are three cross-harbor tunnels in Hong Kong with significantly different characteristics. Western Harbor Crossing (WHC), with the most expensive toll charge (HK$ 50) amongst these three tunnels, has the least mean and variance of travel times. Conversely, Cross Harbor Tunnel (CHT) with cheapest toll charge (HK$ 20) has the largest mean and variance of travel times. Eastern Harbor Crossing (EHC) charges HK$ 25; and its mean and variance of travel time are slightly larger than that of WHC.

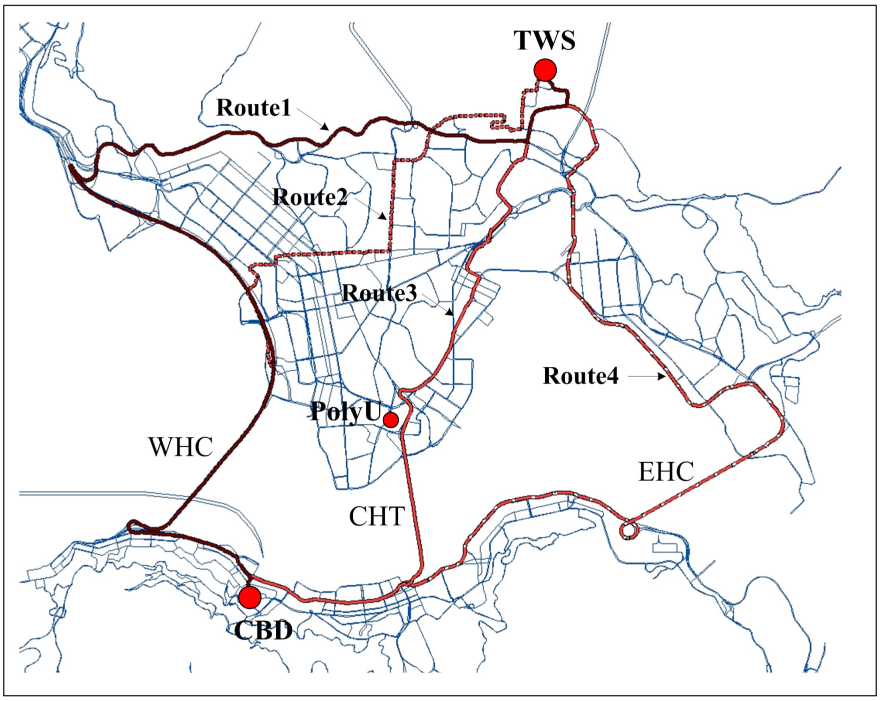


Figure 7. Results of reliable route searching from TWS to CBD

As shown in Figure 7, we tested the reliable routing services using a crossing harbor journey from Tsz Wan Shan (TWS) to the Central Business District (CBD). The travelers' preferred arrival time was set as 9AM. The reliable path finding results are reported in Figure 7 and Table 3.

Table 3. Reliable route searching results from TWS to CBD

Travel time reliability α (%)	*VOT* (HK$ / min)	*VOD* (HK$ / km)	Route	Departure Time	Toll charge + petrol cost (HK$)	Mean travel time (min)	Std. dev. of travel time (min)
10	1	1.5	Route3	**8:50:13**	38.9	37.56	21.69
30	1	1.5	Route3	**8:33:48**	38.9	37.56	21.69
50	1	1.5	Route3	**8:22:26**	38.9	37.56	21.69
70	1	1.5	Route3	**8:11:04**	38.9	37.56	21.69
90	1	1.5	Route4	**8:14:53**	53.3	31.41	10.70
99	1	1.5	Route4	**8:03:45**	53.3	31.41	10.70
99	**10**	1.5	Route1	**8:10:07**	76.6	29.02	8.99
99	10	**7.0**	Route2	**8:08:09**	149.6	32.31	8.42

Four routes were obtain by using various combinations of α, VOT and VOD values. When the VOT was small (i.e., HK$ 1 / min), risk-neutral travelers (i.e., $\alpha = 0.5$) preferred to use Route3, which was the cheapest route (HK$ 38.9) passing through CHT. In this scenario, travelers would depart from the origin at 8:22:26AM only based on mean travel time (i.e., 37.5 min). With the increase of α parameter, risk-averse travelers tended to budget additional time to achieve a high travel time reliability. For example, when $\alpha = 0.7$, risk-averse travelers using the same route would depart at 8:11:04AM, which was 11.4 min earlier than risk-neutral travelers. In addition to departure time adjustment, risk-averse travelers may change their route choice to achieve a higher travel time reliability. For example, when $\alpha = 0.9$, travelers would switch to Route 4 by using EHC, because it was faster and more reliable than Route 3. By using the more reliable route, travelers can depart late from their origin at 8:14:53 AM while achieve a high on-time arrival probability of 90%. Another observation in Table 1 was that high-income travelers were more willing to travel through reliable but expensive routes for achieving a high travel time reliability. For example, when *VOT* = HK$ 10 / min, travelers would prefer to use Route 1 (passing through WHC) to obtain the on-time arrival probability of 99%. In addition, we can observe that travelers tended to use a shorter route (Route2) when *VOD* became larger (*VOD* = HK$ 7 / km is the cost by taxi in Hong Kong). These results demonstrated that RPSS was robust to take account of travelers' various routing criteria toward travel time reliability, travel distance and toll charge.

To examine effects of travel time correlation, the results of SD-RSPP were compared to those ignoring the travel time correlations amongst adjacent links. Let σ_u^{rs} be the actual standard deviation of travel time of path p_u^{rs}; and $\bar{\sigma}_u^{rs}$ be the standard deviation of travel time of p_u^{rs} when ignoring the link travel time correlations. The percentage difference between σ_u^{rs} and $\bar{\sigma}_u^{rs}$ can be calculated as $\lambda = (\bar{\sigma}_u^{rs} - \sigma_u^{rs}) / \sigma_u^{rs}$. In this study, the λ values for 333 O-D pairs from all zone centres to CBD were calculated. Figure 8 shows the percentage cumulative frequency distributions of these λ values. It can be observed that 94.9% of path travel time standard deviations were underestimated when the link travel time correlations were ignored. At 50th percentile (median value), half of O-D pairs underestimated standard deviations of path travel times by 11.6%.

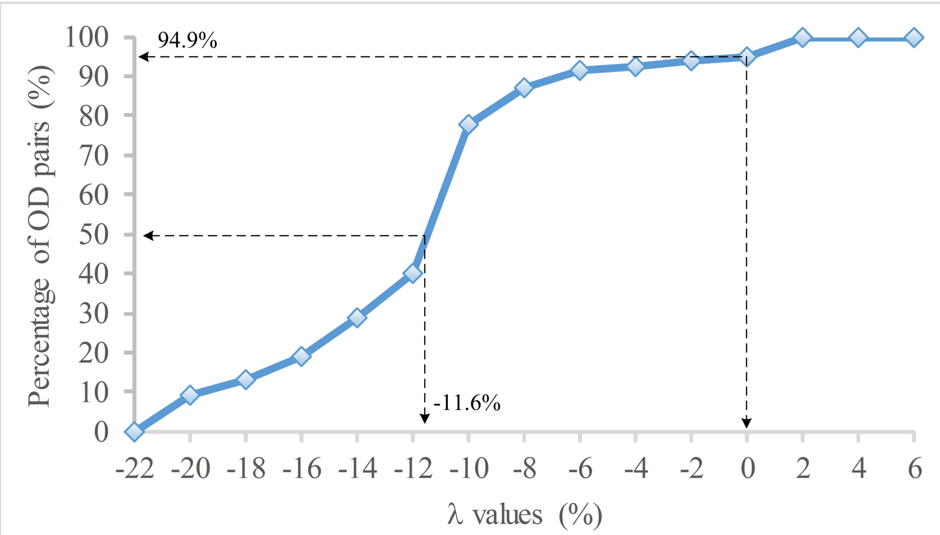


Figure 8. Cumulative distributions of λ values

5. Conclusions

This study proposed an efficient algorithm for solving spatially-dependent reliable shortest path problem (SD-RSPP) in the context of route guidance systems (RGS). In this study, SD-RSPP was formulated as a multi-criteria shortest path problem. The M-V dominance condition was adopted for NC links with negative travel time correlations. Stricter M-B dominance conditions were established for PC links and MC links to reduce the number of generated M-V non-dominated paths during the path searching process. A new multi-criteria A* algorithm (called ERSPA* algorithm) was proposed to solve efficiently the SD-RSPP by adaptively using M-B dominance and M-V dominance conditions. Based on the ERSP-A* algorithm, a reliable path search system (RPSS), was developed to provide reliable RGS to road users in Hong Kong.

A comprehensive case study was carried out in this study to demonstrate the applicability of the developed RPSS in reality. The results of case study showed that travel times are highly stochastic and spatial correlated in congested road networks. Ignoring travel time correlations can significantly underestimate path travel time variation. The case study results suggested that the developed RPSS was robust to take account of users' multiple routing criteria toward travel time reliability, travel distance and toll charge. Computational experiments showed that the proposed ERSP-A* algorithm can find the α-reliable path in large-scale networks within satisfactory computational times (within 30 milliseconds), and has an obvious computational advantage over the SD-RSP-A* algorithm (Chen et al., 2012).

Several directions for future research are worth for further investigations. Firstly, the developed RPSS can provide reliable routing services only to road users. The extension of the developed RPSS to multi-modal transportation networks for providing reliable routing service to transit users is an interesting topic for future work, particularly for Asian cities with higher transit patronage (Fu et al., 2014a, 2014b). Secondly, the RPSS was developed to find a priori α-reliable path for travelers in face of travel time uncertainties. In the literature, substantial attentions also have been given to adaptive en-route guidance applications for finding a set of routing strategies when the en-route traffic information are updated during journeys (Fan and Nie, 2006; Bell, 2009; Ma et al., 2013, 2015). How to extend the proposed ERSP-A* algorithm to the adaptive en-route guidance applications is another interesting topic of further studies. Finally, the A* technique was adopted in this study to speed up the reliable shortest path finding performance. In the literature, several advanced heuristic techniques recently have been developed for improving the efficiency of standard shortest path algorithms in deterministic networks, such as contraction hierarchy (Geisberger et al., 2012) and transit node routing techniques (Bast et al., 2007). How to incorporate these advanced heuristic techniques for reliable shortest path finding in stochastic networks needs further investigations.

Acknowledgements
The work described in this paper was jointly supported by research grants from the National Natural Science Foundation of China (Nos. 41571149 and 41231171), the Research Grants Council of the Hong Kong Special Administrative Region (Project Nos. PolyU 5242/12E and 152074/14E), the Research Committee of the Hong Kong Polytechnic University (Project No. 87PV), and the

Research Institute of Sustainable Urban Development of the Hong Kong Polytechnic University (Project No. 1-ZVEW).

Appendix 1. Proofs

Proposition 2. (Mean-Budget dominance**)** Given a travel time reliability α and two paths $p_u^{r,ijk} \neq p_v^{r,ijk} \in P^{rk}$, $p_u^{r,ijk} \succ p_v^{r,ijk}$ if $p_u^{r,ij}$ and $p_v^{r,ij}$ satisfy $t_u^{r,ij} < t_v^{r,ij}$, $\Phi^{-1}_{T_u^{r,ij}}(\alpha) \leq \Phi^{-1}_{T_v^{r,ij}}(\alpha)$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$.

Proof. According to Proposition 1, we have $p_u^{r,ijk} \succ p_v^{r,ijk}$ if $z_\alpha \sigma_u^{r,ij} \leq z_\alpha \sigma_v^{r,ij}$ and $t_u^{r,ij} < t_v^{r,ij}$ are satisfied. Then, we prove that $p_u^{r,ijk} \succ p_v^{r,ijk}$ holds when $z_\alpha \sigma_u^{r,ij} > z_\alpha \sigma_v^{r,ij}$, $t_u^{r,ij} < t_v^{r,ij}$, $\Phi^{-1}_{T_u^{r,ij}}(\alpha) \leq \Phi^{-1}_{T_v^{r,ij}}(\alpha)$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$ as following. According to Definition 1, $p_u^{r,ijk} \succ p_v^{r,ijk}$ is equivalent to $\pi_{uv}^{r,ij}(a_{jk} \oplus p^{kw}) = \Phi^{-1}_{T_u^{rw}}(\alpha) - \Phi^{-1}_{T_v^{rw}}(\alpha) < 0$, $\forall p^{kw} \in P^{kw}$. Because $\pi_{uv}^{r,ij}(p^{jw}) = (t_u^{r,ij} - t_v^{r,ij}) + z_\alpha \left(\sqrt{(\sigma_u^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2} - \sqrt{(\sigma_v^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2} \right)$, its gradient can be formulated as

$$\frac{\partial \pi_{uv}^{r,ij}(p^{jw})}{\partial\left(2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2\right)} = \frac{z_\alpha}{2} \frac{\sqrt{(\sigma_v^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2} - \sqrt{(\sigma_u^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2}}{\sqrt{(\sigma_u^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2} * \sqrt{(\sigma_v^{r,ij})^2 + 2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2}}$$

As $z_\alpha \sigma_u^{r,ij} > z_\alpha \sigma_v^{r,ij}$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$, we have $\frac{\partial \pi_{uv}^{r,ij}(p^{jw})}{\partial\left(2\,\mathrm{cov}(a_{ij}, a_{jk}) + (\sigma^{jw})^2\right)} < 0$ and thus $\pi_{uv}^{r,ij}(p^{jw})$ is a monatomic decreasing function. Since $\Phi^{-1}_{T_u^{r,ij}}(\alpha) \leq \Phi^{-1}_{T_v^{r,ij}}(\alpha)$, we have $\pi_{uv}^{r,ij}(p^{jw}) < 0$ for $\forall p^{kw} \in P^{kw}$. Therefore, $p_u^{r,ijk} \succ p_v^{r,ijk}$ holds when $z_\alpha \sigma_u^{r,ij} > z_\alpha \sigma_v^{r,ij}$, $t_u^{r,ij} < t_v^{r,ij}$, $\Phi^{-1}_{T_u^{r,ij}}(\alpha) \leq \Phi^{-1}_{T_v^{r,ij}}(\alpha)$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$ are satisfied. Since $p_u^{r,ijk} \succ p_v^{r,ijk}$ holds when $z_\alpha \sigma_u^{r,ij} \leq z_\alpha \sigma_v^{r,ij}$ and $t_u^{r,ij} < t_v^{r,ij}$ are satisfied, we thus have $p_u^{r,ijk} \succ p_v^{r,ijk}$ when $t_u^{r,ij} < t_v^{r,ij}$, $\Phi^{-1}_{T_u^{r,ij}}(\alpha) \leq \Phi^{-1}_{T_v^{r,ij}}(\alpha)$ and $\mathrm{cov}(a_{ij}, a_{jk}) \geq 0$ are satisfied. □

Proposition 3. If the used heuristic function is admissible, the proposed ERSP-A* algorithm can obtain the optimal solution when the destination was reached.

Proof. Let P^{rs} be the path set containing all possible non-dominated paths between the O-D pair, $\bar{P}^{rs} \in P^{rs}$ be the set of generated paths maintained in SE, and $\tilde{P}^{rs} = P^{rs} - \bar{P}^{rs}$ be set of paths that have not been generated by the algorithm. As path $p_u^{r,is}$ was selected from SE, we have $F(p_u^{r,is}) \leq F(p_v^{r,is})$ for $\forall p_v^{r,is} \in \bar{P}^{rs}$. According to Eq. (9), when the heuristic function is admissible, its value is monotonically increasing with path extension. Because all paths in $\tilde{P}^{rs}$ are extended from existing paths in SE, we have $F(p_u^{r,is}) \leq F(p_v^{r,is})$ for $\forall p_v^{r,is} \in \tilde{P}^{rs}$. Therefore, the selected path $p_u^{r,is}$ is the reliable shortest path with the minimum travel time budget in P^{rs}. □

Appendix 2. Consideration of real road network complexities in route guidance systems

To model the concern of travel cost in individual's route choice decision in reality, toll charge and petrol cost are incorporated into the objective function of SD-RSPP. Let d_{ij} and τ_{ij} be the link length and toll charge of link a_{ij}. Without loss of generality, τ_{ij} and d_{ij} are assumed to be deterministic. Let τ_i^{rs} and d_i^{rs} be toll charge and travel distance of a path p_i^{rs} between the O-D pair. They can be expressed as

$$\tau_i^{rs} = \sum_{a_{ij} \in A} \tau_{ij} x_{ij}^{rs} \tag{12}$$

$$d_i^{rs} = \sum_{a_{ij} \in A} d_{ij} x_{ij}^{rs} \tag{13}$$

Incorporating toll charge and petrol cost into Eq. (5), the objective function becomes

$$\min_{x_{ij}^{rs}} \quad f_u^{rs}(\alpha) = t_u^{rs} + (\tau_i^{rs} + d_i^{rs} VOD) / VOT + z_\alpha \sigma_u^{rs} \tag{14}$$

where VOD is for converting travel distance d_i^{rs} into the petrol cost; and VOT is the value of time for converting the toll charge and the petrol cost into the time unit. These two parameters can be determined based on vehicle type and road user type. The developed ERSP-A* algorithm can be easily modified for solving this problem by replacing mean travel time t_u^{rs} with the generalized travel cost $g_u^{rs} = t_u^{rs} + (\tau_i^{rs} + d_i^{rs} VOD) / VOT$.

In classical shortest path problem, any journey always starts from an origin node and ends at a destination node. Nevertheless, in reality, the origins or destinations of travel in road network should not only be located at nodes but also along the road links. Given a point of interest (POI) denoted by a_{ij} in the road network, its location is represented by (a_{ij}, θ) by using linear reference technique (Miller and Shaw, 2001; Chen et al., 2016b). $\theta \in [0,1]$ indicates the POI's relative position on link a_{ij}, and $\theta = 0$ and $\theta = 1$ represent the beginning and end of the link respectively. Let $\hat{a}_{uk}$ be a part of link a_{ij}, and ($\hat{\tau}_{uk}, \hat{d}_{uk}, \hat{T}_{uk}$) be the toll charge, length and travel time of the partial link $\hat{a}_{uk}$. It is assumed that $\hat{\tau}_{uk}$ is equal to τ_{ij}, and $\hat{T}_{uk}$ is proportional to the link length $\hat{T}_{uk} = T_{ij} * \hat{d}_{uk} / d_{ij}$.

To address SD-RSPP for the situation of origin and/or destination at links, temporary nodes and links are added in the road network. When origin is at link a_{ij}, a temporal node o and partial link $\hat{a}_{oj}$ should be added into the road network. When destination is at link a_{ij}, a temporal node d and partial link $\hat{a}_{id}$ should be added into the road network. After the reliable shortest path finding using the ERSP-A* algorithm, all added temporal nodes and links are removed. In this way, origin and/or destination of any journey can be at links.

References

Bast, H., Funke, S., Sanders, P. and Schultes, D., 2007, Fast routing in road networks with transit nodes. *Science*, 316, pp. 566-566.

Bates, J., Polak, J., Jones, P. and Cook, A., 2001, The valuation of reliability for personal travel. *Transportation Research Part E*, 37, pp. 191-229.

Bell, M.G.H., 2009, Hyperstar: a multi-path astar algorithm for risk averse vehicle navigation. *Transportation Research Part B*, 43, pp. 97-107.

Carrion, C. and Levinson, D., 2012, Value of travel time reliability: A review of current evidence. *Transportation Research Part A*, 46, pp. 720–741.

Chan, K.S., Lam, W.H.K. and Tam, M.L., 2009. Real-time estimation of arterial travel times with spatial travel time covariance relationships. *Transportation Research Record*, 2121, 102-109.

Chen, A. and Ji, Z.W., 2005, Path finding under uncertainty. *Journal of Advanced Transportation*, 39, pp. 19-37.

Chen, B.Y., Lam, W.H.K., Sumalee, A. and Li, Z.L., 2012, Reliable shortest path finding in stochastic networks with spatial correlated link travel times. *International Journal of Geographical Information Science*, 26, pp. 365-386.

Chen, B.Y., Lam, W.H.K., Sumalee, A., Li, Q.Q., Shao, H. and Fang, Z.X., 2013a, Finding reliable shortest paths in road networks under uncertainty. *Networks & Spatial Economics*, 13, pp. 123-148.

Chen, B.Y., Lam, W.H.K., Li, Q.Q., Sumalee, A. and Yan, K., 2013b, Shortest path finding problem in stochastic time-dependent road networks with stochastic first-in-first-out property. *IEEE Transactions on Intelligent Transportation Systems*, 14, pp. 1907-1917.

Chen, B.Y., Lam, W.H.K., Sumalee, A., Li, Q.Q. and Tam, M.L., 2014, Reliable shortest path problems in stochastic time-dependent networks. *Journal of Intelligent Transportation Systems*, 18, pp. 177-189.

Chen, B.Y., Shi, C., Zhang, J., Lam, W.H.K., Li, Q.Q. and Xiang, S., 2016a, Most reliable path-finding algorithm for maximizing on-time arrival probability. *Transportmetrica B*, DOI:10.1080/21680566.2016.1169953.

Chen, B.Y., Yuan, H., Li, Q.Q., Shaw, S.-L., Lam, W.H.K. and Chen, X., 2016b, Spatiotemporal data model for network time geographic analysis in the era of big data. *International Journal of Geographical Information Science*, 30, pp. 1041-1071.

El Esawey, M. and Sayed, T., 2011, Travel time estimation in urban networks using limited probes data. *Canadian Journal of Civil Engineering*, 38, pp. 305-318.

Fan, Y.Y. and Nie, Y., 2006, Optimal routing for maximizing the travel time reliability. *Networks & Spatial Economics*, 6, pp. 333-344.

Frank, H., 1969. Shortest paths in probabilistic graphs. *Operations Research*, 17, pp.583-599.

Fredman, M.L. and Tarjan, R.E., 1987. Fibonacci heaps and their uses in improved network optimization algorithms. *Journal of the Association of Computing Machinery*, 34, pp. 596-615.

Fu, L., Sun, D. and Rilett, L.R., 2006, Heuristic shortest path algorithms for transportation applications: State of the art. *Computers & Operations Research*, 33, pp. 3324-3343.

Fu, X., Lam, W.H.K. and Chen, B.Y., 2014a, A reliability-based traffic assignment model for multi-modal transport network under demand uncertainty. *Journal of Advanced Transportation*, 48, pp. 66-85.

Fu, X., Lam, W.H.K. and Meng, Q., 2014b, Modelling impacts of adverse weather conditions on activity-travel pattern scheduling in multi-modal transit networks. *Transportmetrica B*, 2, pp. 151-167.

Gajewski, B.J. and Rilett, L.R., 2003. Estimating link travel time correlation: An application of bayesian smoothing splines. *Journal of Transportation and Statistics*, 7 (2-3), pp. 53-70.

Geisberger, R., Sanders, P., Schultes, D. and Vetter, C., 2012, Exact routing in large road networks using contraction hierarchies. *Transportation Science*, 46, pp. 388-404.

Gutierrez, E. and Medaglia, A.L., 2008, Labeling algorithm for the shortest path problem with turn prohibitions with application to large-scale road networks. *Annals of Operations Research*, 157, pp. 169-182.

Huang, B., Wu, Q. and Zhan, F.B., 2007. A shortest path algorithm with novel heuristics for dynamic transportation networks. *International Journal of Geographical Information Science*, 21 (6), 625-644.

Ji, Z.W., Kim, Y.S. and Chen, A., 2011, Multi-objective alpha-reliable path finding in stochastic networks with correlated link costs: A simulation-based multi-objective genetic algorithm approach (SMOGA). *Expert Systems with Applications*, 38, pp. 1515-1528.

Kaparias, I., Bell, M.G.H. and Belzner, H., 2008, A new measure of travel time reliability for in-vehicle navigation systems. *Journal of Intelligent Transportation Systems*, 12, pp. 202-211.

Khani, A. and Boyles, S.D., 2015, An exact algorithm for the mean-standard deviation shortest path problem. *Transportation Research Part B*, 81, pp. 252-266.

Lam, T.C. and Small, K.A., 2001. The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E*, 37 (2-3), 231-251.

Li, Q.Q., Chen, B.Y., Wang, Y. and Lam, W.H.K., 2015, A hybrid link-node approach for finding shortest paths in road networks with turn restrictions. *Transactions in GIS*, 19, pp. 915-929.

Li, Z.C., Huang, H.J. and Lam, W.H.K., 2012, Modelling heterogeneous drivers' response to route guidance and parking information system in stochastic and time-dependent networks. *Transportmetrica*, 8, pp. 105-129.

Ma, J., Fukuda, D. and Schmöcker, J.-D., 2013, Faster hyperpath generating algorithms for vehicle navigation. *Transportmetrica A*, 9, pp. 925-948.

Ma, J. and Fukuda, D., 2015, A hyperpath-based network generalized extreme-value model for route choice under uncertainties. *Transportation Research Part C*, 59, pp. 19-31.

Miller, H.J. and Shaw, S.L., 2001. Geographic information systems for transportation: principles and applications. New York: Oxford University Press.

Nie, Y. and Wu, X., 2009a. Shortest path problem considering on-time arrival probability. *Transportation Research Part B*, 43 (6), 597-613.

Nie, Y. and Wu, X., 2009b, Reliable a priori shortest path problem with limited spatial and temporal dependencies. In: *18th International Symposium on Transportation and Traffic Theory 2009: Golden Jubilee*, W.H.K. Lam, S.C. Wong and H.K. Lo (Eds.), pp. 169-195.

Rakha, H., El-Shawarby, I., Arafeh, M. and Dion, F., 2006, Estimating path travel-time reliability. In: *2006 IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, pp. 236-241.

Srinivasan, K.K., Prakash, A.A. and Seshadri, R., 2014, Finding most reliable paths on networks with correlated and shifted log-normal travel times. *Transportation Research Part B*, 66, pp. 110-128.

Susilawati, S., Taylor, M.A.P. and Somenahalli, S., 2013, Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, 47, pp. 720-736.

Tam, M.L. and Lam, W.H.K., 2008. Using automatic vehicle identification data for travel time estimation in Hong Kong. *Transportmetrica*, 4 (3), 179-194.

Taylor, M.A.P., 2013, Travel through time: the story of research on travel time reliability. *Transportmetrica B*, 1, pp. 174-194.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit Region. *Economic Geography*, 46, 234–240.

Waller, S.T. and Ziliaskopoulos, A.K., 2002, On the online shortest path problem with limited arc cost dependencies. *Networks*, 40, pp. 216-227.

Wu, X. and Nie, Y., 2011, Modeling heterogeneous risk-taking behavior in route choice: A stochastic dominance approach. *Transportation Research Part A*, 45, pp. 896-915.

Xing, T. and Zhou, X., 2013, Reformulation and solution algorithms for absolute and percentile robust shortest path problems. *IEEE Transactions on Intelligent Transportation Systems*, 14, pp. 943 - 954.

Yang, L. and Zhou, X., 2014, Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem. *Transportation Research Part B*, 59, pp. 22-44.

Zeng, W., Miwa, T., Wakita, Y. and Morikawa, T., 2015, Application of Lagrangian relaxation approach to a-reliable path finding in stochastic networks with correlated link travel times. *Transportation Research Part C*, 56, pp. 309-334.

Zhou, Y., Wang, W., He, D. and Wang, Z., 2014, A fewest-turn-and-shortest path algorithm based on breadth-first search. *Geo-spatial Information Science*, 17, pp. 201-207.