International Journal of the Physical Sciences Vol. 6(20), pp. 4848-4853, 23 September, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.925
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Optimal route choice in stochastic time-varying transportation networks considering on-time arrival probability

Siliang Wang* and Minghui Wang

School of Computer Science, Sichuan University, Chengdu 610065, China.

Accepted 5 July, 2011

This paper studies the problem of finding a priori optimal paths to guarantee a maximum likelihood of arriving on-time in a stochastic time-varying transportation network. The reliable path can help travelers better plan their trip by measuring the risk of late under uncertain conditions. We first identify a set of mathematical relationships between the on-time arrival probability, mean and variance of the dynamic and stochastic link travel times on the networks. The arriving time of each node can be computed by using central limit theorem for its independent link travel time. We show that the time varying problem is decomposable with respect to arrival times and therefore can be solved as easily as its static counterpart. An efficient algorithm is also proposed in the paper. Numerical results are provided using typical transportation networks.

Key words: Stochastic time-varying transportation networks, on-time arrival probability, solution algorithm, route choice, mathematical model.

INTRODUCTION

In stochastic, time-varying networks (STV networks), travel times are modeled as random variables with time-varying distributions, which provide a better modeling tool in transportation applications (Gao et al., 2006; Nielsen, 2004). Finding an optimal routing strategy under uncertainty is an important question in diverse fields of science and engineering, especially in emergency and disaster management. Various route criteria are used in the studies in order to cope with the uncertainty. Those criteria include expectation travel time, reliability, value at risk and utility functions. Most studies have focused on shortest path problems (Bellman, 1958; Fu et al., 1998; Current et al., 1993; Hall, 1986). Miller-Hooks and Mahmassani (2000) discuss time-dependent least expected time (LET) path problems without waiting policy. Fan and Nie (2006) study the LET path problems considering correlation between random link travel times. However, the LET path could not describe the risk measurement. The reason is that the lower expected travel time may have higher variance. Instead of minimizing the expectation, Frank (1969) considered the optimal path that maximizes the probability of realizing a travel time less than a predefined value. The on-time arrival probability measure, on the other hand, considers the percentage of trips that are completed within a reasonable buffered travel time. In a study by Fan and Kalaba (2005), a path finding algorithm was proposed to minimize the probability of arriving at the destination later than a specified arrival time. Recently, Nie and Wu (2009) developed solution algorithms with first-order stochastic dominance rules for the on-time arrival probability. However, all these formulations require enumerating paths and evaluating multiple integrals, and thus cannot be easily implemented for large-scale problems. It is known that it exists as many as paths between an origin and destination in a network of nodes. The result makes it unrealistic to enumerate all possible paths in large time-varying networks. As a result, how to decrease the computation complexity is the main challenge in finding optimal route choice in STV networks under arrival probability constraints (Miller Hooks et al., 1998).

The major goal of this paper is to propose an efficient method in finding paths in STV networks on maximizing on-time arrival probability. A novel approximation method

*Corresponding author. E-mail: hhren1980@126.com.

is proposed in reducing the computation complexity. The method presents a way in defining and computing the distribution of arriving time at each node. An expected function is deduced to reflect different risk of running late on basis of the arriving time distribution. Through the Taylor series, the conversion from dynamic to static condition is discussed in the network. The final optimal paths search algorithm proves its feasibility and efficiency. The rest of the paper is organized as follows. First, the descriptions of the problem are presented. Next, a mathematical programming model is presented and the novel approach is defined and its efficiency is analyzed. Then, an efficient algorithm is proposed based on the approach and finally conclusion of the paper.

NETWORK DESCRIPTION AND PROBLEM DEFINITION

The optimal route choice considering on-time probability problem is formulated as a mathematical programming in STV networks. Let $G(N,A)$ be a directed network, N is the set of nodes, $|N|=n$, and A is the set of arcs, $|A|=m$. The set of successor nodes is given by $\Gamma^{+1}(i)=\{j\,|\,(i,j)\in A\}$ and the set of predecessor nodes is given by $\Gamma^{-1}(i)=\{j\,|\,(j,i)\in A\}$. The time horizon $T=[0,\infty)d$ is divided into finite time intervals within each day. The set of time intervals is $\{t_0+l\theta\}$, and $l=0,1,2,...,L$, θ is the smallest increment of time.

The travel time $C_{ij}(t)$ along the link ij is assumed to be continuous stochastic process, and t is the departure time at node i. Each travel time is independent and follows a random distribution with a probability density function varying over time. We denote the probability density function for travel time on link ij by $f_{ij}(C_{ij}(t))$. For simplicity, no waiting is permitted at node.

For a given source s and destination d, we have a deadline in time t_d, and we would link to find a path which maximizes the probability that we reach the destination within time t_d. Thus, the optimal path considering on-time arrival probability is to solve the equation follows:

$$\max_{\pi} \Pr\left(\sum_{ij\in A} C_{ij} \le t_d\right) \tag{1}$$

For any path π, $\sum_{ij\in A} C_{ij} = T_d$, and each link travel time is independent, then, $T_d \sim N(\mu,\sigma^2)$ according to the central limit theorem. Therefore, the on-time probability can be computed by:

$$\Pr\left(\sum_{ij\in A} C_{ij} \le t_d\right) = \Pr\left(\frac{\sum_{ij\in A} C_{ij} - \mu}{\sigma} \le \frac{t_d-\mu}{\sigma}\right) = \phi\left(\frac{t_d-\mu}{\sigma}\right) \tag{2}$$

where $\phi(\cdot)$ is the cumulative distribution function of the standard normal random variable $N(0,1)$. Since $\phi(\cdot)$ is monotone increasing, the optimal path problem is equivalent to finding the path which maximizes its argument:

$$\max\left[\frac{t_d-\mu}{\sigma}\right] \tag{3}$$

In Equation 3, t_d is a constant. Let $t_d=\mu+\lambda\sigma$, then, the Equation 3 is equal to the equation that follows:

$$\min[\mu+\lambda\sigma] \tag{4}$$

where λ denotes the measurement of on-time arrival probability. Consequently, the optimal route choice considering on-time probability is equal to find the value of the Equation 4 in the STV network.

MATHEMATICAL MODEL

Definition 1

The link travel times are modeled as the sum of free flow travel time and random delay. That is, $C_{ij}(t)=ff_{ij}+D_{ij}(t)$, where ff_{ij} is the travel time when the link ij is in state of free flow and $D_{ij}(t)$ is the random delay time on link ij when departing at t.

Definition 2

According to Equation 4, the shortest path considering risk measurement defined here is $\min[mean(p(t))+B\cdot\sqrt{\mathrm{var}(p(t))}]$, where $mean(p(t))$ is mean value of path travel time when departing at t and $\mathrm{var}(p(t))$ is the variance of path travel time as departing at t. B is the coefficient relating to the risk measurement.

Definition 3

The indicator variable δ_{ij} is defined as follows:

$$\delta_{ij} = \begin{cases} 1 & if\ arc\ ij\ is\ selected; \\ 0 & otherwise \end{cases} \tag{5}$$

Through the aforementioned definitions, the mathematical model of SP problem is formulated as follows:

$$\min_{p \in P^{rs}} [e(t,p) + \beta\sqrt{v(t,p)}] \tag{6}$$

The objective function is subject to the following equations:

$$e(t,p) = t + \sum_{i \in N}\sum_{ij \in A} E[C_{ij}(t_i)]\delta_{ij}, \quad \forall t_i \in T, p \in P^{rs} \tag{7}$$

$$v(t,p) = \sum_{i \in N}\sum_{ij \in A} Var[C_{ij}(t_i)]\delta_{ij}, \quad \forall t_i \in T, p \in P^{rs} \tag{8}$$

$$\sum_{j \in \Gamma^{-1}(i)} \delta_{ij} - \sum_{j \in \Gamma^{+1}(i)} \delta_{ji} = \begin{cases} 1 & if\ i = s, \\ -1 & if\ i = d, \\ 0 & otherwise. \end{cases} \tag{9}$$

$$\delta_{ij} \in \{0,1\}, \forall i, j \in N \quad (i,j) \in A, t = t_s, \beta \in (-\infty, +\infty) \tag{10}$$

where $E[\cdot]$ denotes expected value and $Var[\cdot]$ is the variance of the random variable.

The objective function describes the shortest path considering on risk measurement. The risk considered here bases on evaluation of deviating from the schedule arriving time. Constraint (7) denotes the expected value of path travel time of path p when departing at time t . Constraint (8) denotes the variance of path travel time of path p when departing at time t . Constraint (9) denotes the relationship between path and its corresponding links. It is easily deduced that the complexity of the solution space is in connection with start time t (also refers to distribution of arrival time at the node) and links. The complexity of solution space is $\Theta(nL\prod_{\forall ij \in A}|S_{ij}|)$, and $|S_{ij}|$ is the number of possible link travel time. L is the number of time intervals, which would become larger when considering a long time periods (Loui, 1983; Guan et al., 2011). The reason is that the method of evaluation of the path travel time. Thus, better of the description of path travel time also makes more efficient of the computation method. The paper presents a novel method of evaluation of the path travel time in order to provide an efficient algorithm.

EVALUATION OF ARRIVING TIME OF THE NODE

Let the random variable Y_i denote the arrival time at node i , then Y_i is equal to the departure time at node i or the time link a is entered. The departure at the origin node s, Y_s is assumed to be deterministic and known a priori. Let random variable Z_a be the travel time on link a under a given instance. The path travel time is equal to the sum of the travel time of all the links along the path. This path travel time can be obtained by calculating the arrival time at each node along the path using a recursive formula until the destination node is reached. The relationship equation is as follows:

$$Y_j = Y_i + Z_a \tag{11}$$

The relationship between the mean arrival times at node i and the mean arrival time at the downstream node j is as follows:

$$E[Y_j] = E[Y_i] + E[Z_a] \tag{12}$$

The variance of the arrival time at a downstream node j may be derived using Equation 12 which is based on the following formula:

$$Var[Y_j] = Var[Y_i] + Var[Z_a] + 2COV(Y_i, Z_a) \tag{13}$$

The symbol $COV(\cdot)$ denotes covariance of random variables. $Var[\cdot]$ denotes variance of random variables. Assume the probability that Y_i falls into the nth time interval at node i is as follows:

$$p_i^n = \varphi(z_n) - \varphi(z_{n-1}) \tag{14}$$

where $\varphi(z_n)$ denotes the cumulative distribution of a standard normal random variable, The reason is that the sum of independent variables satisfy the central limit theorem. The time of entering the node i is as follows:

$$z_n = (t_0 + l\theta - E[Y_i]) / \sqrt{\mathrm{var}[Y_i]} \tag{15}$$

Thus, Equation 8 is reformulated as the following:

$$E[Y_j] = E[Y_i] + ff_{ij} + \sum_{n=1}^{L} p_i^n \rho_{ij}^n \tag{16}$$

where ff_{ij} is the travel time of link as in free flow state. ρ_{ij}^n is the mean delay time on link ij . Similarly, Equation 9 is reformulated as the following:

$$\begin{aligned} Var[Y_j] = & Var[Y_i] + Var[\mu(Y_i)] \\ & + E[v(Y_i)] + 2E[Y_i\mu(Y_i)] \\ & - 2E[Y_i]E[\mu(Y_i)] \end{aligned} \tag{17}$$

where $\mu(Y_i)$ is expected value of link travel time and $\mu(Y_i) = E[C_{ij}(t_i)]$. Likewise, $v(Y_i)$ is variance of link travel time and $v(Y_i) = Var[C_{ij}(t_i)]$. On the down sides of the aforementioned equation are five parts. Each can be computed by simultaneous Equation (14 to 17). The results of the computation are as follow:

$$E[v(Y_i)] = \sum_{n=1}^{L} p_i^n \sigma_{ij}^{2n} \tag{18}$$

$$\begin{aligned} & Var[\mu(Y_i)] = \\ & \sum_{n=1}^{L} p_i^n [ff_{ij} + \rho_{ij}^n]^2 - [ff_{ij} + \sum_{n=1}^{L} p_i^n \rho_{ij}^n]^2 \end{aligned} \tag{19}$$

$$E[Y_i\mu(Y_i)] = \sum_{n=1}^{L} (ff_{ij} + \rho_{ij}^n) P_i^n \tag{20}$$

$$E[\mu(Y_i)] = ff_{ij} + \sum_{n=1}^{L} p_i^n \rho_{ij}^n \tag{21}$$

As a result, the value of mean and variance of the random Y_j are decided by several parameters, such as the following:

$$E[Y_j] = E[Y_i] + f(ff_{ij}, p_i^n, \rho_{ij}^n) \tag{22}$$

$$Var[Y_j] = Var[Y_i] + u(ff_{ij}, p_i^n, \rho_{ij}^n) \tag{23}$$

Where $E[Y_s] = t$, $Var[Y_s] = 0$, σ_{ij} is variance of delay time. The expected value and variance of arriving time at each node can be computed by the Equations 22 and 23.

Proposition 1

Given a start time t_1 , the shortest path problem considering on risk measurement can be reformulated as deterministic optimization problem.

Proof: consider the definition of link travel time $C_{ij}(t_i)$, $C_{ij}(t) = ff_{ij} + D_{ij}(t_i)$, moreover, waiting at the node is not allowed and $Y_j = Y_i + Z_a$, thus, $Y_j(t + t_a(t)) = Y_i(t) + Z_a(t)$. Because of the object function, $\min_{i \in N}[nE[Y_i] + m\sqrt{Var[Y_i]}]$, and the object function can be formulated as:

$$\begin{aligned} & \min[n' \cdot E[Y_1(t_1)] + l \cdot \sum_{a \in A, i \in N} \mu_{Z_a}(f(\mu_i), ff_a)) \\ & + m' \cdot Var[Y_1(t_1)] + r \cdot \sum_{a \in A, i \in N} v_{Z_a}(f(\mu_i), \sigma_a)] \end{aligned}$$

Therefore, the proposition is proved.

SOLUTION ALGORITHM

According to Proposition 1, we devise the solution algorithm like Dijkstra algorithm. The pseudo code of the algorithm is as follows.

Algorithm 1

Compute shortest path considering on-time arrival probability in STV networks.

Step 1: Initializations

Bool s[maxnum];
// Judging whether the node is visited
for(int i=1; i<=n; ++i)// n is the max number of the nodes
{
dist[i] = c[s][i];
//Distance between two nodes, and dist[i]=nE[Y_i]+mVar[Y_i],c[i][j]=E[Z_a].
s[i] = 0;
// Initialization the set.
if(dist[i] == maxint)
prev[i] = 0;// previous node on basis of current node i.
else
prev[i] = s;
}

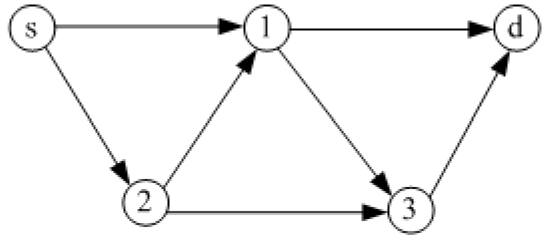

Figure 1. A simple STV networks.

Table 1. Inputs of parameters for links in path.

Parameter	Link 1 (s-1)	Link 2 (s-2)	Link 3 (1-d)	Link 4 (1-3)	Link 5 (2-1)	Link 6 (2-3)	Link 7 (3-d)
Free flow time	2	1	2	3	0.5	3	2
Delay time probability distribution	exp(2)	0≤t<10, exp(1); 10≤t, exp(2)	0≤t<12, exp(2); 12≤t<20, exp(3); 20≤t, exp(3)	0≤t<14, exp(2); 14≤t<24, exp(3); 24≤t, exp(1)	0≤t<10, exp(6); 10≤t<16, exp(5); 16≤t, exp(2)	0≤t<11, exp(1); 11≤t<22, exp(2); 22≤t, exp(3)	0≤t<13, exp(2); 13≤t<20, exp(3); 20≤t, exp(2)

```
dist[s] = 0;
s[s] = 1;
```

Step 2: Main loop

```
// Find minimum value of objective function
for(int i=2; i<=n; ++i)
{
int tmp = maxint;
int u = s;
for (int j=1; j<=n; ++j)
if((!s[j]) && dist[j]<tmp)
{
u = j;
//the minimum distance in the adjacency table
tmp = dist[j];
}
s[u] = 1;
// Update the distance
for(int j=1; j<=n; ++j)
if((!s[j]) && c[u][j]<maxint)
{
int newdist = dist[u]+ c[u][j];
if(newdist < dist[j])
{
dist[j] = newdist;
prev[j] = u;
}
}
}
```

Step 3: Output the shortest path.

In Algorithm 1, the complexity of the computation is $\Theta(n^2)$.

NUMERICAL EXAMPLES AND ANALYSIS

Consider a simple STV networks as Figure 1, and it is very common in transportation application. The node s is source node and d is destination node. The inputs of parameters for links in path are in Table 1. The research time is within one day, and assumes the link travel time follow exponential distribution. According to the computation with algorithm 1, the results of SP problem in STV networks in Figure 1 considering the risk measurement is in Table 2. The results prove its efficiency of the algorithm.

CONCLUSION

This paper studies the problem of finding a priori optimal paths to guarantee a maximum likelihood of arriving on-time in a stochastic, time-varying transportation network. The reliable path can help travelers better plan their trip by measuring the risk of late under uncertain conditions.

The paper first identifies a set of mathematical relationships between the on-time arrival probability and mean and variance of the dynamic and stochastic link travel times on the networks. The arriving time of each node can

Table 2 . Results of route choice in the STV network.

Start time (t_s)	On-time arrival probability (λ)	Shortest path
1	0.3	(s-1-d)
2	0.8	(s-2-3-d)
3	0.9	(s-2-1-d)
4	0.1	(s-1-3-d)
5	0.6	(s-1-3-d)
6	0.7	(s-2-1-d)
7	0.8	(s-2-1-d)
8	0.7	(s-1-3-d)
9	0.4	(s-1-3-d),(s-2-1-3-d)
10	0.6	(s-1-3-d)

be computed by using central limit theorem for its independent link travel time. We show that the time varying problem is decomposable with respect to arrival times and therefore can be solved as easily as its static counterpart.

The paper only considers single constraint (travel time) in the mathematical model. The future work is considering multiple constraints in real transportation networks.

REFERENCES

Gao S, Chabini I (2006). Optimal Routing Policy Problems in Stochastic Time-Dependent Networks. Transp. Res., Part B. 2(40): 93-122.

Nielsen LR (2004). Route Choice in Stochastic Time-Dependent Networks. PhD thesis, Department of Operations Research, University of Aarhus, pp. 13-25.

Bellman R (1958). On a routing problem. Q. Appl. Math., 1(16): 87-90.

Liping Fu, Rilett LR (1998). Expected shortest paths in dynamic and stochastic traffic networks, Transp. Res.-B, 7(32): 499-516.

Current J, Marsh M (1993). Multi objective transportation network design and routing problems: Taxonomy and annotation. Eur. J. Oper. Res., 65: 4-19.

Hall RW (1986). The fastest path through a network with random time-dependent travel times. Transp. Sci., 3(20): 182-188.

Miller-Hooks ED, Mahmassani HS (2000). Least expected time paths in stochastic, time-varying transportation networks. Transp. Sci., 34(2): 198-215.

Fan Y, Nie Y (2006). Optimal Routing for Maximizing the Travel Time Reliability. Netw. Spatial Econ., 6(3): 333-344.

Frank H (1969). Shortest Paths in Probabilistic Graphs. Oper. Res., 2(17): 583-599.

Fan YR, Kalaba JM (2005). Shortest Paths in Stochastic Networks with Correlated Link Costs. Comput. Math. Appl., 3(49): 1549-1564.

Nie Y, Wu X (2009). The Shortest Path Problem Considering On-Time Arrival Probability. Trans. Res. Part B, 6(43): 597-613.

Miller-Hooks ED, Mahmassani HS (1998). Optimal routing of hazardous materials in stochastic, time-varying transportation networks. Trans. Res. Records, 1(1645): 143-151.

Loui RP (1983). Optimal paths in graphs with stochastic or multidimensional weights. Comm. ACM, 9(26): 670-676.

Guan G, Zhangxin C, Qun W, Anmin H (2011). Estimation of multipath time delays based on bi-sparse constraint. Int. J. Phys. Sci., 6(3): 402-406.