Path Finding Under Uncertainty

Anthony Chen
Zhaowang Ji

Path finding problems have many real-world applications in various fields, such as operations research, computer science, telecommunication, transportation, etc. In this paper, we examine three definitions of optimality for finding the optimal path under an uncertain environment. These three stochastic path finding models are formulated as the expected value model, dependent-chance model, and chance-constrained model using different criteria to hedge against the travel time uncertainty. A simulation-based genetic algorithm procedure is developed to solve these path finding models under uncertainties. Numerical results are also presented to demonstrate the features of these stochastic path finding models.

Introduction

The path finding problem under uncertainty is to find a feasible path between an origin and a destination such that the optimality condition defined in a stochastic context is optimized. In a deterministic environment, the path finding problem is usually defined as the shortest path (SP) problem in terms of distance, time, cost, or a combination of deterministic attributes (i.e., the value of the attribute is assumed known with certainty). However, in real life situations, the environment is often uncertain. For transportation, these uncertainties could arise from many sources, such as the variation of traffic demands and link capacities during the peak and non-peak hours, the variation of seasons and weather conditions, the unpredictable intersection delays, the mixed traffic flow by different vehicle types, etc. These uncertainties result in travel time variability in the network (i.e., environment is uncertain). Under such situation, the uncertainty of link travel times and their associated probability density functions should be explicitly considered when

Anthony Chen and Zhaowang Ji, Utah State University, UT, USA
determining the optimal path. The question is how should an optimal path be defined and identified when link travel times are defined by a probability distribution? The answer to this question is not obvious. One popular method for dealing with travel time uncertainty is to find the minimum expected shortest path [e.g., Fu and Rilett, 1998; Hall, 1986; Miller-Hooks and Mahmassani, 2000]. The advantage of this method is that efficient shortest path algorithms developed for the deterministic setting [Bellman, 1958; Dijkstra, 1959] can be readily adapted to identify the minimum expected travel time path in a stochastic setting. However, the path with the minimum expected travel time may not be reliable (i.e., high variability in travel time). Such a path is not optimal (or even risky) for travelers who are more concerned about the reliability of path travel time when finding paths in an uncertain environment. Recent empirical studies [Abdel-Aty et al., 1996; Kazimi et al., 2000; Lam, 2000; Lam and Small, 2001; Small et al., 1999] seem to support that travel time reliability is indeed an important criterion in route choice decisions. These studies concluded that travelers are interested in not only travel time saving but also reduction of travel time variability when making their route choice decisions under circumstances where they do not know with certainty about the outcome of their decisions. From this observation, it appears the notion of reliability should be incorporated into the path finding problem. Frank (1969) proposed defining the path that maximizes the probability of realizing a travel time less than a given travel time budget to be the optimal path. Sigal et al. (1980) suggested the optimal path should be the path that has the greatest probability of realizing the least cost. Both problems are concerned with maximizing the probability that the path travel time is less than or equal to some threshold value. The maximum probability model is equivalent to finding the most reliable path. It can also be interpreted as the definition of travel time reliability [Bell and Iida, 1997]. However, both studies did not provide a solution method to identify such optimal paths. Recently, Fan (2003) proposed the stochastic on-time arrival (SOTA) problem, which is the same as the maximum probability problem defined by Frank (1969) and Sigal et al. (1980), and provided a tractable numerical method that relies on Bellman’s Principle of Optimality and Picard’s method of successive approximation for solving the SOTA problem. Chen et al. (2004) provided an alternative definition of optimality that allows the travelers to specify a confidence level $\alpha$ for finding a reliable path with the minimum travel time budget required to meet the travel time reliability constraint. The $\alpha$ reliable path problem is formulated as a
chance constrained model and solved by a simulation-based genetic algorithm procedure. The advantage of this model is that it is able to identify a portfolio of paths with different levels of reliability to suit the traveler's risk preference towards travel time variability. To exert an increasing level of control of the risk of exceeding the minimum travel time budget, the traveler can increase the confidence level in the travel time reliability constraint. In this paper, we examine how the expected value model, the maximum probability (or most reliable) model, and the $\alpha$ reliable model deal with travel time variability when finding paths in an uncertain environment.

The paper is organized as follows: the following section presents three definitions of optimal path under uncertainty. Then, the mathematical formulations corresponding to these definitions are presented. After that, the simulation-based genetic algorithm procedure for solving these path finding models under uncertainty is described. Numerical experiments are conducted to illustrate the different path finding models under uncertainty, and conclusions and further research directions are addressed in the final section.

Definitions of Optimal Path Under Uncertainty

In an uncertain environment, the definition of optimal path is not obvious. In this section, we provide three possible definitions of optimality for finding path under uncertainty. It should be mentioned that these three definitions are not exhaustive. There are many others that also deserve examination (e.g., mean-variance model [Sen et al., 2001]; variance-constrained shortest path [Sivakumar and Batta, 1994]; robust shortest path [Yu and Yang, 1998]). In this paper, we are primarily interested in the minimum expected travel time path, the most reliable path, and the $\alpha$ reliable path definitions.

(i) Minimum expected travel time path. A path $x$ is called the minimum expected travel time path from origin $r$ to destination $s$ if

$$\min \{E[t(x, \xi)]\} \leq \min \{E[t(x', \xi)]\}$$

for any path $x'$ from origin $r$ to destination $s$, where $\xi$ is a vector of random link travel times, $t(x, \xi)$ is the path travel time from origin $r$ to destination $s$ (a random variable), and $E[\cdot]$ is the expectation operator. The minimum expected travel time path problem is one of the popular methods for dealing with travel time uncertainty. However, travelers are
not always concerned with minimizing the expected travel time. A path may have the minimum expected travel time, yet a large variability in travel time. Travelers sometimes have to consider the reliability of arriving on time.

(ii) Most reliable path. A path \( x \) is called the most reliable path from origin \( r \) to destination \( s \) if

\[
\Pr\{t(x, \xi) \leq \bar{T}\} \geq \Pr\{t(x', \xi) \leq \bar{T}\}
\]

for any path \( x' \) from origin \( r \) to destination \( s \), where \( \bar{T} \) is a predetermined travel budget (or threshold value) specified by the traveler, and \( \Pr\{\} \) denotes the probability of the event in \( \{} \). The most reliable path problem is concerned with arriving on time for a given travel time budget (or threshold). The maximum probability problem proposed by Frank (1969) and studied by Sigal et al. (1980), the stochastic on time arrival problem defined by Fan (2003), and the most reliable path definition above are essentially the same.

(iii) \( \alpha \) reliable path. A path \( x \) is called the \( \alpha \) reliable path from origin \( r \) to destination \( s \) if

\[
\min \left\{ \bar{T} \mid \Pr(t(x, \xi) \leq \bar{T}) \geq \alpha \right\} \leq \min \left\{ \bar{T} \mid \Pr(t(x', \xi) \leq \bar{T}) \geq \alpha \right\}
\]

for any path \( x' \) from origin \( r \) to destination \( s \), where \( \bar{T} \) is the travel time budget (a variable in the \( \alpha \) reliable path definition) and \( \alpha \) is a predetermined confidence level specified by the traveler. The \( \alpha \) reliable path problem is meaningful for travelers who are concerned with arriving at the destination within a certain level of confidence while minimizing the travel time budget required satisfying the travel time reliability constraint.

Mathematical Formulations

Let \( G(N, A) \) denotes a stochastic network, where \( N \) is the set of nodes and \( A \) is the set of links. Also, we denote \( O(n) \) as the set of outbound links emanating from node \( n \) and \( I(n) \) as the set of inbound links feeding into node \( n \). \( \xi_{ij} \) is the random travel time on link \((i, j)\), which is an element of the random link travel time vector \( \xi = \left( \xi_1, \xi_{ij}, \xi \right) \). \( x_{ij} \) is the decision variable, where \( x_{ij} = 1 \) means that link \((i, j)\) is in the path, and 0 otherwise, and \( x = \left( \xi, x_{ij}, \xi \right) \) denotes the
corresponding vector with element $x_{ij}$. In this section, we provide mathematical formulations for the three definitions of optimal path presented in Section 2. The minimum expected travel time path can simply be formulated as an expected value model (EMV). For the latter two definitions which explicitly consider the notion of reliability, they are formulated as an uncertain program [Liu, 1999]. Specifically, the most reliable path is formulated as a dependent chance model (DCM) and the $\alpha$ reliable path is formulated as a chance constrained model (CCM).

**Expected Value Model**

The expected value model is perhaps the most commonly used method for finding the least expected travel time paths. The main idea is to optimize the expected value of a linear (or additive) objective function subject to path feasibility constraints.

$$\text{Min } E\left[ \sum_{(i,j) \in A} \xi_{ij} x_{ij} \right]$$

Subject to

$$\sum_{j \in O(n)} x_{nj} - \sum_{i \in I(n)} x_{in} = \begin{cases} 1 & \text{if } n = r \\ 0 & \text{if } n \neq r; i \neq s, \forall n \in N \\ -1 & \text{if } n = s \end{cases}$$

$$x_{ij} = (0, 1), \forall (i, j) \in A.$$  \hspace{1cm} (6)

In the expected value path finding problem, the objective function in equation (4) is to minimize the expected path travel time. Equation (5) ensures the links in the path are feasible, connecting from origin $r$ to destination $s$. Equation (6) ensures the decision variables are binary. In the expected value model, it only considers the average value of network travel times, travel time variability is totally ignored. Under this model, travelers would be indifferent between two paths that have equal expected path travel time but with different path travel time variability. The path identified in this model can be risky since it may select a path with higher travel time variability. Such a path is suboptimal for travelers who are concerned with the reliability of path travel time.
Dependent Chance Model

The most reliable path problem is to find the path that satisfies the travel time budget with the highest probability. This problem can be formulated as a dependent chance model (DCM) as follows.

\[
\text{Max } \Pr(\sum_{(i,j) \in A} \xi_{ij} x_{ij} \leq \overline{T}) \tag{7}
\]

Subject to

\[
\sum_{j \in O(n)} x_{nj} - \sum_{i \in I(n)} x_{in} = \begin{cases} 
1 & \text{if } n = r \\
0 & \text{if } n \neq r; i \neq s, \forall n \in N \\
-1 & \text{if } n = s 
\end{cases} \tag{8}
\]

\[
x_{ij} = (0, 1), \quad \forall (i, j) \in A, \tag{9}
\]

where \(\overline{T}\) is the travel time budget between origin \(r\) and destination \(s\) given by the traveler. In DCM, \(\overline{T}\) is a predefined threshold and the objective in equation (7) is to find a path that maximizes the probability of arriving on time or earlier than the given travel time budget. Equivalently, the objective is to find a path that maximizes the travel time reliability defined by the probability for a given travel time budget. Equations (8) and (9) are the same as (5) and (6) which ensure the links in a path are feasible and the decision variables are binary.

Chance Constrained Model

The \(\alpha\) reliable path problem can be formulated as a chance constrained model (CCM) as follows.

\[
\text{Min } \overline{T} \tag{10}
\]

Subject to

\[
\Pr(\sum_{(i,j) \in A} \xi_{ij} x_{ij} \leq \overline{T}) \geq \alpha, \tag{11}
\]

\[
\sum_{j \in O(n)} x_{nj} - \sum_{i \in I(n)} x_{in} = \begin{cases} 
1 & \text{if } n = r \\
0 & \text{if } n \neq r; i \neq s, \forall n \in N \\
-1 & \text{if } n = s 
\end{cases} \tag{12}
\]

\[
x_{ij} = (0, 1), \quad \forall (i, j) \in A, \tag{13}
\]

where \(\overline{T}\) represents the travel time budget (a variable in CCM) that travelers want to minimize in equation (10). Equation (11) is the probabilistic constraint that ensures the probability of path travel time
less than the budget is greater than or equal to a confidence level \( \alpha \) (a predetermined value in CCM). Equations (12) and (13) are the same as (5) and (6) which ensure the links in a path are feasible and the decision variables are binary.

**Simulation-Based Genetic Algorithm**

Stochastic integer optimization problems are generally difficult to solve by traditional calculus-based methods. To solve the path finding models under uncertainty described in above section, we develop a simulation-based genetic algorithm procedure that integrates stochastic simulation and genetic algorithm.

**Stochastic Simulation**

Stochastic simulation is an important tool for performing sampling experiments on the models of stochastic systems [Liu, 1999]. It is based on sampling random variables from probability distributions to compute the uncertain functions. Let \( \xi = (\xi_i, \xi_j, \xi_k) \) be the random link travel time vector defined on the probability space \((\Omega, \Theta, \Pr)\), where \( \Omega \) is a set of all outcomes of a random experiment (a non-empty set), \( \Theta \) is called a \( \sigma \)-algebra, and \( \Pr \) is referred to as a probability measure. For each \( \omega \in \Omega \), \( \xi(\omega) \) is a realization of the random link travel time vector \( \xi \). In order to compute the uncertain functions used in the path finding models under uncertainty, we assume that a set of paths between an origin and a destination has been determined from the genetic algorithm procedure to be described later. The three uncertain functions to be computed are (i) expected value function, (ii) probability function, and (iii) chance constrained function.

(i) **Expected Value Function**
The objective function of the expected value model is

\[
U_1 : (x) \rightarrow E \left[ \sum_{(i,j) \in A} \xi_{ij} x_{ij} \right] \quad (14)
\]

To compute it, we use the following stochastic simulation procedure:

Step 1. Set \( U_1(x) = 0 \).

Step 2. Generate \( \omega \) from \( \Omega \) according to the probability measure \( \Pr \).
Step 3. For each $\xi(\omega)$, calculate $t(x, \xi)$ for all paths in the path set.

Step 4. $U_1(x) \leftarrow U_1(x) + t(x, \xi)$.

Step 5. Repeat the second to fourth steps for $N$ times, where $N$ is a sufficiently large number.

Step 6. Return $U_1(x)/N$.

(ii) Probability Function

The objective function of the dependent chance model is

$$U_2: (x) \rightarrow \left\{ \Pr\left( \sum_{(i,j) \in A} \xi_{ij} x_{ij} \leq T \right) \right\}$$

The steps of the stochastic simulation procedure are:

Step 1. Set $N' = 0$.

Step 2. Generate $\omega$ from $\Omega$ according to the probability measure $Pr$.

Step 3. For each $\xi(\omega)$, calculate $t(x, \xi)$ for all paths in the path set.

Step 4. If $t(x, \xi) \leq T$, then $N' \leftarrow N' + 1$.

Step 5. Repeat the second to fourth steps for $N$ times, where $N$ is a sufficiently large number.


(iii) Chance Constrained Function

The objective function of the chance constrained model is

$$U_3: (x) \rightarrow \left\{ T \mid \Pr\left( \sum_{(i,j) \in A} \xi_{ij} x_{ij} \leq T \right) \geq \alpha \right\}$$

The steps of the stochastic simulation procedure are:

Step 1. Generate $\omega_1, \omega_2, \omega_3, \omega_N$ from $\Omega$ according to the probability measure $Pr$, where $N$ is a sufficiently large number.

Step 2. For each $\xi(\omega_k)$, calculate $t(x, \xi)$ for all paths in the path set and denote the path travel time by $t_k(x, \xi)$ for $k = 1, 2, \ldots, N$.

Step 3. Set $N'$ as the integer part of $\alpha N$.

Step 4. Return the $N'th$ least element in $\{t_1(x, \xi), t_2(x, \xi), \ldots, t_N(x, \xi)\}$. 
Genetic Algorithm Procedure

Genetic Algorithm (GA) is an intelligent stochastic search method for optimizing complex problems. After Holland (1975) first proposed the algorithm and with decades of development, GA has been widely applied in many fields because of its globality, parallelism, and robustness features (for detailed knowledge about GA, readers can refer to Goldberg (1989) and Gen and Cheng (2000). Compared to numerous GA applications in other fields, research on path finding using GA is relatively scarce. One reason may be the difficulty of representing a path as a chromosome (since the number of nodes on each path is not fixed, the length of a chromosome is variable). Another problem is maintaining the feasibility of a chromosome under GA’s crossover and mutation operations. Network topology has to be considered in the GA’s operators in order to maintain feasibility of the chromosomes. Gen and Cheng (2000) proposed a priority orientated chromosome representation. Gomez-Sanchez (2001) provided another variable length chromosome presentation by using a dynamic adjacent matrix. In this study, we adopt the variable length chromosome representation and develop appropriate GA operators to ensure feasibility for the path finding problems under uncertainty. For this section, we briefly describe the major components of the GA procedure and refer the readers to Ji and Chen (2004) for the details of implementation of the GA procedure.

Chromosome Representation

One important aspect of applying GA to any problem is the representation of the decision variables in the genetic fashion or, as a chromosome. For the path finding problem, the chromosome is coded as a node sequence. Each gene in the chromosome represents a node index, and the gene sequence is arranged according to the node sequence that a path would go through to get from the origin to the destination. Under this coding scheme, the length of each chromosome can be different. By adopting a variable length chromosome representation, the GA procedure uses less computer memory and is more efficient in manipulating the chromosomes to obtain better solutions.
Reproduction Operator

Reproduction is a selection process that selects chromosomes from a population pool based on their fitness for mating. The fitness of the chromosome implies the number of times each chromosome will be in the mating pool. The most commonly used selection schemes are roulette wheel and tournament selection. In this study, the tournament selection and the half replacement strategy are adopted. After evaluating the fitness of all chromosomes in the population pool, they are ranked in an ascending order based on the fitness values. Chromosomes in the top half are eligible to be selected as parent chromosomes, and the bottom half will be replaced by the new offspring chromosomes generated from the crossover and mutation operators.

Crossover Operator

Crossover is a means of exchanging genetic materials between two parent chromosomes such that two new offspring chromosomes, containing genetic materials from the parent chromosomes, can be generated. Crossover occurs with a constant probability, which implicitly indicates the expected number of chromosomes in the mating pool undergoing crossover. There are many crossover schemes in the literature (e.g., single-point crossover, multiple-point crossover, uniform crossover, etc.) In the path finding problem, the single-crossover scheme is adopted. However, we only perform the crossover operation when there are at least one common node (origin and destination nodes are excluded) between the two parent chromosomes. The crossover operator must also ensure the paths are feasible and cycle-free after performing the operation.

Mutation Operator

Mutation alters the value of genetic units for the purpose of introducing new genetic structures to the new offspring. All new offspring are subjected to the mutation operator with a predefined mutation rate. Mutation allows the GA to explore new regions of the solution space and helps to prevent convergence to a sub-optimal solution. In the path finding problem, the mutation operation works as follow. A chromosome is randomly selected to perform the mutation operation. Among the nodes in the selected chromosome, a node with
multiple emanating links (again excluding the origin and the destination) is randomly selected as a mutation node to generate a new subpath from the mutation node to the destination. The mutation operator is designed to increase the diversity of the solutions in the population set.

A Simulation-Based Genetic Algorithm Procedure

In this section, we summarize the major steps of the simulation-based genetic algorithm procedure for solving the path finding models under uncertainty. For more details about the design of the procedure, readers are referred to Ji and Chen (2004).

Step 0. Define input parameters: population size, crossover and mutation rates, maximum number of generations, and maximum number of simulations.

Step 1. Generate an initial path set and initialize the generation index.

Step 2. Evaluate the fitness of all chromosomes in the population pool using the stochastic simulation procedure described above.

Step 3. Check whether the predefined maximum generation number is reached or not. If yes, go to step 6; otherwise, go to step 4.

Step 4. Rank the chromosomes based on their fitness values and use the tournament selection scheme to select parent chromosomes for reproduction.

Step 5. Update the chromosomes using the crossover and mutation operators, increment the generation index, and go to step 2.

Step 6. Report the best chromosome as the optimal path.

Numerical Results

To demonstrate the feasibility of the three stochastic path finding models, a small network depicted in Figure 1 is used. In this network, there are four types of links. For simplicity, travel times on all four types of links are assumed to follow a normal distribution with similar means but different variances (note that this is not a requirement for the simulation-based GA procedure). Type D has the highest variability in travel times; type B has the lowest variability in travel times; and type A and type C are in between Type D and Type B. The link travel time distributions are given in Table 1.

There are 6 possible paths connecting between node 1 and node 9. The node sequence of path 1 is 1-2-3-6-9 which contains four type A links; path 2 is 1-4-7-8-9 which contains four type B links; path 3 is 1-4-
5-8-9 which has two type B links and two type D links; path 4 is 1-2-5-6-9 which has two type A and two type C links; path 5 is 1-4-5-6-9 which contains one type A, one type B, one type C and one type D links; and path 6 is 1-2-5-8-9 which has the same link type combination as path 5. Thus, path 5 and path 6 have the exactly the same path travel time distribution albeit they are different paths.

![Test Network Diagram]

**Figure 1. Test Network**

<table>
<thead>
<tr>
<th>Link Type</th>
<th>Link Travel Time Distribution (mean, variance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Normal (200, 400)</td>
</tr>
<tr>
<td>B</td>
<td>Normal (207, 20)</td>
</tr>
<tr>
<td>C</td>
<td>Normal (202.5, 30)</td>
</tr>
<tr>
<td>D</td>
<td>Normal (198, 1500)</td>
</tr>
</tbody>
</table>
The following parameters are set for GA implementation.
- Population size is 20 chromosomes
- The maximum number of generations is 10
- Maximum number of simulations is 1000.
- Crossover rate is 0.6
- Mutation rate is 0.6

Summary of Descriptive Statistics

After running the simulation-based genetic algorithm procedure, the summary of path travel time statistics is provided in Table 2. These descriptive statistics include the mean, standard deviation (STD), range, minimum, maximum, 5th percentile, 15th percentile, 50th percentile, 85th percentile, and 95th percentile. For each statistical measure, the minimum value is highlighted.

<table>
<thead>
<tr>
<th></th>
<th>Path 1 (sec.)</th>
<th>Path 2 (sec.)</th>
<th>Path 3 (sec.)</th>
<th>Path 4 (sec.)</th>
<th>Paths 5 &amp; 6 (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td><strong>799.13</strong></td>
<td>828.14</td>
<td>807.97</td>
<td>804.93</td>
<td>806.15</td>
</tr>
<tr>
<td>Median</td>
<td><strong>800.18</strong></td>
<td>828.20</td>
<td>808.02</td>
<td>803.49</td>
<td>805.95</td>
</tr>
<tr>
<td>STD</td>
<td>40.37</td>
<td><strong>9.11</strong></td>
<td>55.45</td>
<td>30.08</td>
<td>45.00</td>
</tr>
<tr>
<td>Range</td>
<td>255.96</td>
<td><strong>58.79</strong></td>
<td>349.91</td>
<td>198.86</td>
<td>285.06</td>
</tr>
<tr>
<td>Minimum</td>
<td>673.59</td>
<td>802.10</td>
<td><strong>642.56</strong></td>
<td>716.18</td>
<td>669.93</td>
</tr>
<tr>
<td>Maximum</td>
<td>929.55</td>
<td><strong>860.89</strong></td>
<td>992.46</td>
<td>915.04</td>
<td>954.99</td>
</tr>
<tr>
<td>5th percentile</td>
<td>729.24</td>
<td>813.37</td>
<td><strong>712.75</strong></td>
<td>758.06</td>
<td>733.16</td>
</tr>
<tr>
<td>15th percentile</td>
<td>756.88</td>
<td>818.65</td>
<td><strong>750.96</strong></td>
<td>773.78</td>
<td>760.67</td>
</tr>
<tr>
<td>50th percentile</td>
<td><strong>800.13</strong></td>
<td>828.18</td>
<td>807.97</td>
<td>803.44</td>
<td>805.76</td>
</tr>
<tr>
<td>85th percentile</td>
<td>840.62</td>
<td>837.55</td>
<td>864.22</td>
<td><strong>836.17</strong></td>
<td>853.08</td>
</tr>
<tr>
<td>95th percentile</td>
<td>864.82</td>
<td><strong>842.64</strong></td>
<td>898.22</td>
<td>854.86</td>
<td><strong>880.58</strong></td>
</tr>
</tbody>
</table>

From the above table, we can observe the following:
- Path 1 has the minimum expected path travel time;
- Paths 2 has the smallest standard deviation of path travel time and the smallest range of path travel time;
- Path 3 has the smallest minimum path travel time (i.e., minimizing the best outcome);
- Paths 2 has the smallest maximum path travel time (i.e., minimizing the worst outcome);
Depending on the percentile value, different paths would be selected.

**Most Reliable Path Finding in DCM**

In this section, we examine the effect of different travel time budgets on the path travel time reliability (i.e., the probability of path travel time less than a given travel time budget) of the most reliable path finding model. Table 3 provides the most reliable path results for different \( T \) values, ranging from 730 to 880 seconds at an increment of 25 seconds. Under each given travel time budget, the most reliable path with the maximum probability value of arrival on time is highlighted. From the table, we can observe that with an increase of the travel time budget, path travel time reliability is also increased. This is because travelers are more likely to arrive at the destination on time if a larger travel time budget is allowed. Another observation is that with different travel time budgets, different most reliable paths can be found. When the travel time budget is set between 730 to 755 seconds, path 3 the most reliable path; it changes to path 1 when the travel time budget is set between 780 to 805 seconds; path 4 for a travel time budget of 830 seconds; and paths 2 for a travel time budget greater than 855 and less than 880 seconds. This reflects that DCM is able to capture travelers' risk preferences towards travel time variability for different travel time budgets. It should be noted that when \( T = 805 \), which is close to the minimum expected path travel time (see table 2), the optimal path is path 1 which is also the optimal path under EVM. The corresponding travel time reliability is 54.6\% which is close to the 50th percentile. When the travel time budget is specified below the expected value, DCM tends to select larger travel time variability paths (path 3); and when the travel time budget is specified above the expected value, DCM tends to select smaller travel time variability paths (path 4 and path 2).

**α Reliable Path Finding in CCM**

In this section, we examine the effect of different confidence levels on the minimum travel time budget of the α reliable path finding model. Table 4 provides the α reliable path results for different α values, ranging from 0.1 to 1.0 at an increment of 0.1. From Table 4, we can clearly observe that with an increase of the confidence level, the minimum path travel time budget increases. This is because if travelers want to be more certain in arriving at the destination, a higher
travel time budget is required. Another observation is that under different $\alpha$ values, different $\alpha$-reliable paths can be found. When $\alpha$ is specified between 0.1 and 0.2, the optimal path becomes path 3; when $\alpha$ is specified between 0.3 and 0.6, path 1 is optimal; path 4 is optimal when $\alpha$ is specified between 0.7 and 0.8 and path 2 is the optimal path when $\alpha$ is greater or equal to 0.9. This indicates that CCM can accommodate travelers' risk preferences towards travel time variability.

**Table 3. Most Reliable Path Results for Different Travel Time Budgets in DCM**

<table>
<thead>
<tr>
<th>$\bar{T}$ Value (sec.)</th>
<th>Path 1</th>
<th>Path 2</th>
<th>Path 3</th>
<th>Path 4</th>
<th>Path 5 &amp; 6</th>
<th>Optimal Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>730</td>
<td>5.0%</td>
<td>0.0%</td>
<td><strong>8.4%</strong></td>
<td>0.5%</td>
<td>4.3%</td>
<td>Path 3</td>
</tr>
<tr>
<td>755</td>
<td>13.8%</td>
<td>0.0%</td>
<td><strong>16.4%</strong></td>
<td>3.9%</td>
<td><strong>12.3%</strong></td>
<td>Path 3</td>
</tr>
<tr>
<td>780</td>
<td>31.6%</td>
<td>0.0%</td>
<td>30.4%</td>
<td>20.4%</td>
<td>28.1%</td>
<td>Path 1</td>
</tr>
<tr>
<td>805</td>
<td>54.6%</td>
<td>0.4%</td>
<td>47.4%</td>
<td>51.9%</td>
<td>49.1%</td>
<td>Path 1</td>
</tr>
<tr>
<td>830</td>
<td>77.7%</td>
<td>56.4%</td>
<td>64.3%</td>
<td><strong>80.0%</strong></td>
<td>71.3%</td>
<td>Path 4</td>
</tr>
<tr>
<td>855</td>
<td>91.3%</td>
<td>99.5%</td>
<td>79.9%</td>
<td>95.0%</td>
<td>85.7%</td>
<td>Path 2</td>
</tr>
<tr>
<td>880</td>
<td>97.9%</td>
<td><strong>100.0%</strong></td>
<td>91.2%</td>
<td>98.8%</td>
<td><strong>94.7%</strong></td>
<td>Path 2</td>
</tr>
</tbody>
</table>

**Table 4. $\alpha$-Reliable Path Results for Different Confidence Levels ($\alpha$) in CCM**

<table>
<thead>
<tr>
<th>$\alpha$ Value (sec.)</th>
<th>Path 1 (sec.)</th>
<th>Path 2 (sec.)</th>
<th>Path 3 (sec.)</th>
<th>Path 4 (sec.)</th>
<th>Path 5 &amp; 6 (sec.)</th>
<th>Optimal Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>748.60</td>
<td>816.34</td>
<td><strong>735.23</strong></td>
<td>768.57</td>
<td>748.47</td>
<td>Path 3</td>
</tr>
<tr>
<td>0.2</td>
<td>764.58</td>
<td>819.75</td>
<td><strong>763.42</strong></td>
<td>779.53</td>
<td>767.90</td>
<td>Path 3</td>
</tr>
<tr>
<td>0.3</td>
<td><strong>777.64</strong></td>
<td>823.19</td>
<td>779.58</td>
<td>788.08</td>
<td>782.31</td>
<td>Path 1</td>
</tr>
<tr>
<td>0.4</td>
<td><strong>789.07</strong></td>
<td>825.79</td>
<td>793.81</td>
<td>797.19</td>
<td>794.71</td>
<td>Path 1</td>
</tr>
<tr>
<td>0.5</td>
<td><strong>800.13</strong></td>
<td>828.18</td>
<td>807.97</td>
<td>803.44</td>
<td>805.76</td>
<td>Path 1</td>
</tr>
<tr>
<td>0.6</td>
<td><strong>809.72</strong></td>
<td>830.71</td>
<td>823.78</td>
<td>811.21</td>
<td>816.88</td>
<td>Path 1</td>
</tr>
<tr>
<td>0.7</td>
<td>820.15</td>
<td>833.14</td>
<td>837.26</td>
<td><strong>820.05</strong></td>
<td>827.71</td>
<td>Path 4</td>
</tr>
<tr>
<td>0.8</td>
<td>833.31</td>
<td>835.83</td>
<td>855.12</td>
<td><strong>829.92</strong></td>
<td>843.57</td>
<td>Path 4</td>
</tr>
<tr>
<td>0.9</td>
<td>851.29</td>
<td><strong>839.56</strong></td>
<td>876.96</td>
<td>843.32</td>
<td>865.05</td>
<td>Path 2</td>
</tr>
<tr>
<td>1.0</td>
<td>929.55</td>
<td><strong>860.89</strong></td>
<td>992.46</td>
<td>915.04</td>
<td>954.99</td>
<td>Path 2</td>
</tr>
</tbody>
</table>

for different confidence levels. In addition, we can observe when $\alpha=0.5$, the difference between the minimum travel time budget and the expected travel time (from Table 2) is relatively small. This is because when $\alpha$ is set at 0.5, the travel time budget value is essentially the median value of
path travel time. Since network travel times in this numerical example are assumed to follow a normal distribution, the median value is close to the expected value of path travel time. However, the expected value and the median value may not be the same in general.

Figure 2 graphically displays the cumulative distribution function (CDF) of these six paths. From the figure, we can clearly observe that when $730 \leq \bar{T} \leq 769$ (in DCM) or $0.0 \leq \alpha \leq 0.23$ (in CCM), path 3 will be selected as the optimal path; when $769 < \bar{T} < 820$ (in DCM) or $0.23 < \alpha < 0.70$ (in CCM), path 1 will be selected; when $820 \leq \bar{T} \leq 838$ (in DCM) or $0.70 \leq \alpha \leq 0.86$ (in CCM), path 4 will be selected; when $\bar{T} > 838$ (in DCM) or $\alpha > 0.86$ (in CCM), paths 2 will be selected. These observations from the figure are consistent with the results presented in the above tables.

Figure 2. CDF of Path Travel Time
Conclusions and Future Research

In this study, three definitions of optimality for finding path under uncertainty were examined. The minimum expected travel time path definition is perhaps the most widely adopted because the expected value model can be efficiently solved by variants of the Dijkstra's algorithm. However, the path with the minimum expected travel time is not optimal for travelers who are more concerned with on time arrival (i.e., travel time reliability). The most reliable path and the $\alpha$ reliable path explicitly considered the notion of reliability in their definitions. However, the two definitions handle the reliability requirement in different ways. The most reliable path, formulated as a dependent chance model, identifies the optimal path by maximizing the probability of on time arrival for a given travel time budget, while the $\alpha$ reliable path, formulated a chance constrained model, determines the optimal path with the least travel time budget by constraining it to satisfy the travel time reliability requirement for a given confidence level. By adjusting the input parameters (travel time budget in DCM or $\alpha$ value in CCM), both models could find a portfolio of paths to suit the travelers' risk preferences towards travel time uncertainty in a stochastic network. These two reliable path finding models should be useful for both passenger and freight transportation. For future research, we plan to explore other definitions of optimality (e.g., mean-variance model) for finding paths under an uncertain environment. Particularly, we are interested in examining the mathematical properties of the two reliable path finding models and their relationship. On the computational side, we have tested the simulation-based genetic algorithm procedure on large networks with more than 2000 nodes and 7000 links, but the performance is not great. We plan to enhance the efficiency of the procedure to make it practical for real-life applications. In addition, we plan to explore designing more effective algorithms that can guarantee certain optimality for some special cases of the reliable path finding models using the insights obtained from the mathematical properties derived from the mathematical programming formulations for the path finding problem under uncertainty.

Acknowledgment

This research was supported by the NSF CAREER Grant: CMS-0134161.
References


Hall, R. 1986. The fastest path through a network with random time-dependent travel times, Transportation Science 20: 182-188.


