Developing Travel Reliability Inventory for the Chicago Region

Prepared by

Yu (Marco) Nie, Xing Wu, Qianfei Li
Department of Civil & Environmental Engineering
Northwestern University

John Dillenburg, Peter Nelson,
Department of Computer Science, University of Illinois, Chicago

October 11, 2011

Contents

List of Tables

List of Figures

ABSTRACT

Travel reliability is a critical performance dimension of transportation systems and services. This project aims to enable IDOT to document and monitor the reliability performance of its highway network. To this end, we develop a specialized Windows desktop application named Travel Reliability Inventory for Chicago (TRIC), available for download at NUTREND. TRIC integrates the reliable routing algorithm and the Gary-Chicago-Milwaukee (GCM) traffic data with an easy-to-use graphic user interface. With TRIC, the travel reliability inventory can be easily created and exported to ESRI shape files for post analyses. Two case studies are conducted to demonstrate (1) how to use TRIC to analyze travel reliability between a selected O-D pair; and (2) how to visualize and analyze travel reliability inventory generated by TRIC using a freely available GIS tool. The case study reveals interesting results about the relative performance of six major expressway segments (Kennedy Westbound, Eisenhower Westbound, Dan-Ryan Southbound, Stevenson Westbound, Lake Shore Dr North Northbound, Lake Shore Dr South Southbound) in the region during different time periods.

Chapter 1

Introduction

1.1 Background

Travel reliability is a critical performance dimension of transportation systems and services. It enables people and firms to make better use of available resources, including time, through effective personal and business activity scheduling. Shippers and freight carriers need predictable travel times to fulfill on-time deliveries and other commitments in order to remain competitive. The ability to arrive on-time with high reliability is imperative to emergency responders. However, urban transportation systems are affected by uncertainties of various sorts, which can be broadly classified as those affecting the supply of transportation (e.g., weather, accidents, natural and man-made disasters) and those associated with the demand for transportation (e.g., travel and activity behavior, special events). Taken individually or in combination, these factors could adversely affect and perturb the quality of transportation services. Travel behavior researchers have established that unanticipated long delays on highways typically produce much worse frustration among motorists than ”predictable” ones. The U.S. Federal Highway Administration (FHWA) estimates that 50-60% of congestion delays in most metropolitan areas are non-recurrent, and the percentage is even higher in smaller urban areas (FHWA 2000). To hedge against travel time fluctuations, travelers have to budget a sizable time buffer, much of which is often wasted. In 1982, a 20-minute free-flow trip requires on average an extra 12-minute buffer time if on-time arrival is important (FHWA 2005). In contrast, the same trip would require 60% more buffer time in 2003. The reliability issue is likely to get still worse in the years to come, in light of limited capacity addition in the face of continuing growth in demand. Currently, state and local agencies neither archive travel reliability data, nor have access to modeling tools that properly account for unreliability of travel times in the planning practice. Integrating travel reliability into transportation network analysis methods presents a pressing challenge that is of both theoretical and

practical importance. However, a prerequisite for the development and application of such tools is the availability of travel reliability data. The proposed research addresses precisely this prerequisite.

1.2 Objectives

The objective of the project is to develop necessary procedures and computer tools to systematically document travel reliability information for highway networks. To validate the techniques and provide necessary guidelines for its application, a case study for the Chicago area is conducted, which demonstrate how a travel time reliability inventory can be created for the area using the tool developed in the research.

1.2.1 What constitutes a travel reliability inventory?

The basic element of a travel reliability inventory is individual facility's reliability measures, which include the following three dimensions as recommended by Federal Highway Administration (FHWA): 95th percentile travel time, buffer index, and planning time index. Definitions of these reliability measures will be given in Section 3. The reliability performance may be evaluated for different types of facilities, such as individual road segments as defined in a travel planning model, or a specific route between an O-D pair. Sometimes, an O-D pair may be evaluated as a single facility by aggregating the reliability measures over the "dominating" routes that connect the O-D pair. The focus of this study is on road segment and route level, although the tools developed can be easily extended to produced O-D based reliability indexes.

Through an in-depth analysis of the traffic data, we identify time-of-day and day-of-week as two major factors that affect travel time distributions. Accordingly, we segment the original data into four time periods (off-peak, morning-peak, evening-peak and mid-of-day) and five day-of-week scenarios (weekend, weekday, Friday, Saturday and Sunday). For each combination, a different travel time distribution is generated from raw data and the corresponding travel reliability indexes are calculated.

1.2.2 The case study

Chicago is selected for the case study due to two reasons. First, as the third largest metropolitan area in the US, the city of Chicago and its surrounding areas are subject to significant congestion. According to the latest mobility report (Schrank & Lomax 2011), Chicago is the second most congested city in the US, next only to Washington DC, with an average per-commuter travel delay at 71 hours/year in 2010. The travel times in the Chicago area are also more unreliable than any other major metropolitan areas in the US.

An earlier TTI mobility report (2007) indicates that a traveler in Chicago area has to budget 2.07 times of the free flow travel time for an important trip (which requires 95% probability of on-time arrival), the highest index in this country. The second reason to prefer Chicago has to do with data availability. Chicago has archived a rich set of traffic data in both public and private sectors. In particular, the GCM (Gary-Chicago-Milwaukee corridor) traveler information system (www. gcmtravel.com) broadcasts real-time traffic data collected from loop detectors, toll transponders and other devices operated by IDOT, neighboring DOTs and toll roads.

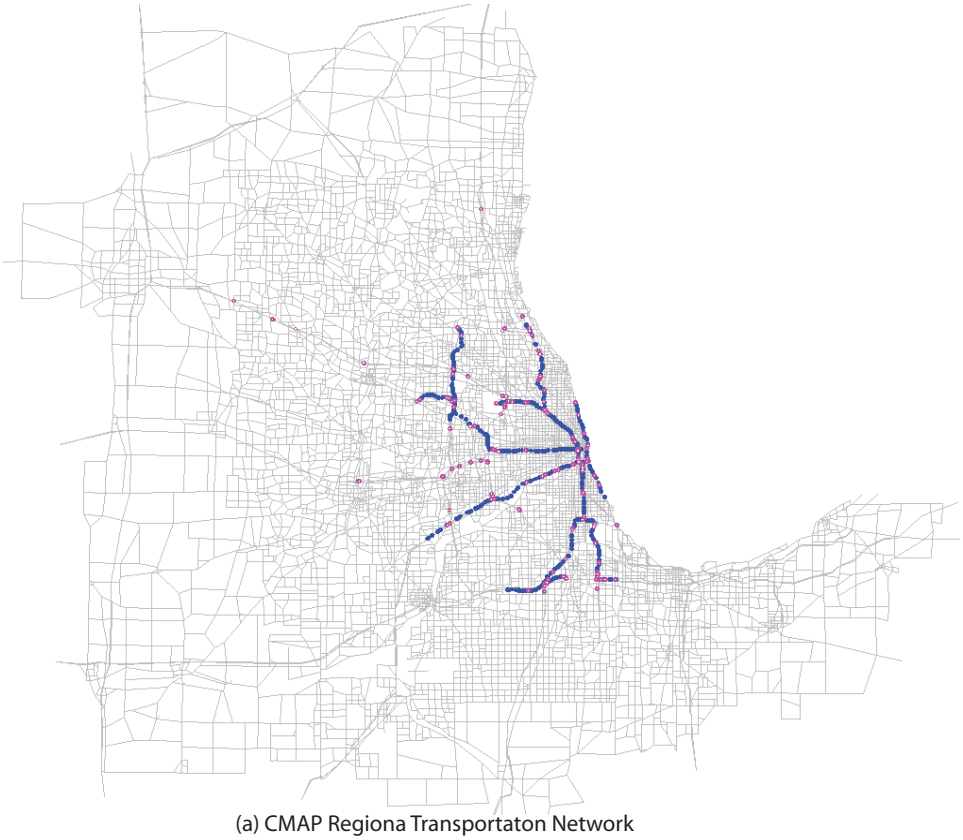
Because data on arterial and local streets are not available, this study will focus on freeways and expressways in the Chicago region.

1.3 Overview of research approach

In this project, travel reliability inventory are developed based on the regional transportation planning network built by Chicago Metropolitan Agency for Planning (CMAP) (see Figure 1.1-(a)). To reduce computational efforts, a subnetwork is first extracted from the CMAP planning network, which consists of most freeways and expressways in the region, as well as local streets necessary to maintain basic connectivity of the network. The subnetwork has in total 1970 links (the original CMAP network has over 40,000 links), as shown in Figure 1.1-(b). The main steps used to build the travel reliability inventory for this subnetwork are as follows:

- Step 1: Develop empirical travel time distributions on individual road segments in the study area from GCM traffic data and other data sources.
- Step 2: Compute the most reliable routes for selected corridors (Origin-destination pairs) in the region and the travel time distributions on these routes.
- Step 3: Calculate corresponding reliability measures for road segments and routes and archive the travel reliability inventory for further analysis.

Each of these steps involve challenges of its own. In the first step, the main issues is how to estimate travel time distributions on the roads that have no travel time observations (recall that these roads have to be included in the network to maintain connectivity). The second step requires not only computing travel time distribution on routes that connect the O-D pair, but also ranking the routes according to their travel time distributions. At the core of such computation is finding routes that are shortest to ensure a specified probability of on-time arrival. Solutions techniques and implementation details of this problem is one of the methodological focus in this report. The central question in the



(a) CMAP Regiona Transportaton Network

(b) The extracted Network for this project

Figure 1.1: Topology of CMAP and extracted expressway networks

last step is how to store the travel reliability inventory for easy future reference. The rest of this report details the efforts to address these challenges.

1.4 Organization

This report is organized as follows. Chapter 2 introduces the methodologies for finding most reliable route between an origin-destination pair, which is a key to evaluating route- and O-D based travel reliability. Chapter 3 discusses how to prepare input data necessary to creating travel reliability inventory. Chapter 4 provides a detailed description of TRIC, the software tool developed in this project, including the design concept, installation guidelines, and a user manual. In Chapter 5 we conduct two case studies to demonstrate how to use TRIC and the travel reliability inventory for travel reliability analysis.

Chapter 2

Methods for finding most reliable routes

In this chapter, we introduce the mathematical formulation and solution algorithm for the so-called *reliable a priori shortest path (RASP) problem*, which aims to find *a priori* paths that are shortest to ensure a specified probability of on-time arrival. Solving the RASP problem is at the core of the evaluation of route-based reliability indexes. In what follows, we first demonstrate the basic concept using an illustrative example. Then, a review of related work is provided. Finally, the formulation and algorithms are described.

2.1 Illustrative example

Simply speaking, the RASP problem attempts to rank routes based on a specific percentile value of their travel time distributions. Figure 2.1 shows the cumulative density functions (CDFs) of the random travel times of two routes. Let $F_i, i = 1, 2$ be the CDF of the two travel time distributions. That is, for a given probability α (e.g., 0.95), we have $F_i(X \leq t_i(\alpha)) = \alpha$, where $t_i(\alpha)$ is called the α percentile travel time of the route i. Figure 2.1 shows that for the given α, $t_1(\alpha) < t_2(\alpha)$. This implies route 1 is a better choice for this particular α because taking route 1 can save the traveler some "time budget" to ensure the desired probability of arriving on time (i.e. α).

In this project, we are interested in finding the route that gives the least 95 percentile travel time between a given origin-destination (O-D) pair. Essentially, the route (and its 95 percentile time) is employed as a rough estimation of the "best scenario" for anyone who has high reliability. Such a path can be found by enumerating all paths and comparing the distributions of their travel times, as shown in the above example. However, directly enumerating all paths is not practicable because it is difficult if not impossible for large networks. Specialized algorithms are thus needed to facilitate the search process.

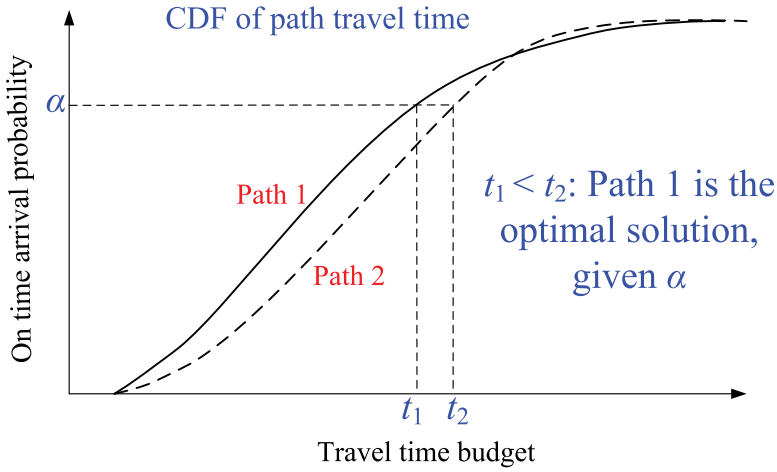


Figure 2.1: An illustrative example of the RASP problem

2.2 Literature review

Stochastic routing considers the circumstance where travelers have to choose the best from a set of routes with random travel times. The distribution of these random travel times are given and subject to no perception or measurement errors. In this context, the simplest behavioral assumption for route choice is that travelers would always minimize the expected travel time. This assumption leads to numerous variants of optimal path problems (Hall 1986, Fu & Rilett 1998, Fu 2001, Miller-Hooks & Mahmassani 2000, Miller-Hooks 2001, Waller & Ziliaskopoulos 2002, Fan, Kalaba & Moore 2005, Gao & Chabini 2006). However, it is easy to see that minimizing the expected travel time does not necessarily account for travel reliability.

Reliability-based routing has been studied extensively. Based on how travel reliability is defined, the existing studies may be categorized as follows.

On-time arrival probability Frank (1969) and Mirchandani (1976) define reliability as the probability of completing a trip within a given travel time budget, i.e. the probability of arriving on time. If this probability is given by the trip purpose, the routing problem becomes one that minimizes the travel time budget required for achieving that provability. Nie & Wu (2009*b*) show that the optimal paths defined above can be found from non-dominated paths under the first-order stochastic dominance (FSD) (Hadar & Russell 1969). Nie & Wu (2009*a*) incorporate the correlations between travel times on adjacent links into the above problem. Using FSD to compare paths with random travel times was also explored in Bard & Bennett (1991) and Miller-Hooks & Mahmassani (2003).

Effective travel time [1] In Hall (1983), travel reliability is modeled by effective travel time, which is defined as the sum of the mean travel time and a safety margin. Typically, the safety margin is the product of the standard deviation of travel time and a scalar called punctuality parameter (Sivakumar & Batta 1994, Sen, Pillai, Joshi & Rathi 2001, Uchida & Iida 1993, Lo, Luo & Siu 2006, Shao, Meng & Tam 2006). If all random travel times are normally distributed, minimizing effective travel time for a given punctuality parameter equals minimizing travel time budget for a corresponding on-time arrival probability, although this equivalence does not hold in general (Wu & Nie 2011*b*).

Expected utility Under the expected utility theory (von Neumann & Morgenstern 1967), travelers are assumed to choose paths that maximize the expected utility of random route travel times. Travel reliability can be incorporated in this framework by choosing a proper form for the utility function. For example, quadratic and exponential functions represent risk-averse and ruin-averse travelers, respectively, as shown in Wu & Nie (2011*b*). However, Bellman's principle of optimality can be used to find such optimal paths only when the utility is an affine or exponential function of the random route travel time (Loui 1983, Eiger, Mirchandani & Soroush 1985, Mirchandani & Soroush 1987). For other utility functions (e.g. Murthy & Sarkar 1996, Murthy & Sarkar 1998, Yin, Lam & Ieda 2004), the problem of finding optimal paths is generally intractable. When the utility function is quadratic and the distributions are uniquely determined by the first two moments, Loui (1983) showed that the optimal path problem is equivalent to a class of bicriteria shortest path problems, which can be formulated and solved using generalized dynamic programming (DP) (see, e.g., Carraway, Morin & Moskowitz 1990).

Connectivity Reliability can also be defined using the concept of connectivity (**?**, Chen, Bell, Wang & Bogenberger 2006, Kaparias, Bell, Chen & Bogenberger 2007). This approach models reliability as the probability that the travel time on a link is greater than a threshold. Accordingly, the reliability on a path is the product of link reliability when independence among the rand variables is assumed.

The routing model adopted in this project defines travel reliability as on-time arrival probability. Under this definition, the objective is to find routes that either maximize reliability for a given travel time budget, or alternatively, minimize the travel time budget for a desired reliability.

It seems natural for a traveler to express his / her perceived importance of a trip using the preferred likelihood of being punctual. Such a criterion is not only more intuitive, but

[1] Effective travel time and travel time budget are sometimes used interchangeably in the literature. In this paper, the term *travel time budget* always refers to the amount of time that a traveler allocates for a trip.

Table 2.1: Notation

s	the destination of routing
c_{ij}	travel times on link ij which is a random variable
$f_{ij}(\cdot)$	probability density function of c_{ij}
$\pi_{ij}(\cdot)$	cumulative distribution function (CDF) of c_{ij}
k^{rs}	path k that connects origin-destination path $r-s$.
K^{rs}	a set of all paths that connect r and s
$u_k^{rs}(b)$	the maximum probability of arriving at s through path k^{rs} on-time or earlier, departing from r with a time budget b.
$u^{rs}(b)$	the maximum probability of arriving at s through any path $k^{rs} \in K^{rs}$ on-time or earlier, departing from r with a time budget b.
Γ^{rs}	FSD-admissible paths between the OD pair rs
Ω^{rs}	FSD-optimal paths between the OD pair rs

also easier to measure and compare. In contrast, it is generally difficult for a traveler to determine their punctuality coefficient (in order to define the effective travel time) or to specify their utility function form. Moreover, it has been noted that travelers do tend to reserve a sizable buffer time for important trips in order to hedge against uncertainties (Hall 1983). According to an FHWA report (FHWA 2005), in 2003, US drivers reserve on average a 20-minute extra buffer time for a 20-minute trip in free flow conditions if arriving on time is important. A survey of students from Hong Kong University of Science and Technology (Lo et al. 2006) shows that all responders allocate a travel time budget above mean travel time for trips to examinations.

Previous studies (Frank 1969, Miller-Hooks 1997, Nie & Wu 2009*b*, Nie & Wu 2009*a*, Wu & Nie 2009) have provided mathematical formulations, solution algorithms and complexity analysis of the RASP problem. For the sack of completeness, the following section provides a detailed review of these results.

2.3 Problem formulation and a solution algorithm

Consider a directed network $G(\mathcal{N}, \mathcal{A}, \mathcal{P})$ consisting of a set of nodes $\mathcal{N}$ ($|\mathcal{N}| = n$), a set of links $\mathcal{A}$ ($|\mathcal{A}| = m$), and a probability distribution $\mathcal{P}$ describing the statistics of link travel times. Table 2.1 summarizes the notation to be used in what follows. This study does not consider the correlations among different c_{ij}, since these are difficult to establish from existing data.

The problem of determining reliable route guidance can be formulated as the *reliable a priori shortest path* (RASP) problem (Nie & Wu 2009*b*). Before presenting the formulation, let us first define the first-order stochastic dominance (FSD) and the associated admissibility.

Definition 2.1 (First-order stochastic dominance (FSD) $\succ_1$) *Path k^{rs} dominates path l^{rs} in the first order, denoted as $k^{rs} \succ_1 l^{rs}$, if the cumulative distribution function (CDF) of π_k^{rs} never lies below that of π_l^{rs} and the inequality holds strictly at least at one point.*

Definition 2.2 (FSD-admissible path) *A path l^{rs} is FSD-admissible if $\exists$ no path $k^{rs} \in K^{rs}$ such that $k^{rs} \succ_1 l^{rs}$.*

The RASP problem is equivalent to the problem of identifying all FSD-admissible paths between $(i, s), \forall i \neq s$. However, it is possible that an FSD-admissible path is not shortest for any on-time arrival probability. To clarify this point, we define FSD optimality in the following.

Definition 2.3 (FSD-optimal path) *A path k^{rs} is FSD-optimal if 1) it is FSD-admissible and 2) it provides the highest on-time arrival probability from node r to node s for some time budget b.*

We shall denote the set of FSD-admissible and FSD-optimal paths between the OD pair (r, s) with Γ^{rs} and Ω^{rs}, respectively. Note that Ω^{rs} is the subset of Γ^{rs} by definition. At any node $i \in \mathcal{N}$, define $u^{is}(b) \equiv \max\{u_k^{is}(b), \forall k^{is} \in \Omega^{is}, \forall b\}$. The function $u^{is}(\cdot)$ is called *Pareto frontier function* at node i, which constitutes optimal solutions of the RASP problem.

The RASP problem can be solved using the following label correcting algorithm (see Nie & Wu (2009*b*) for the proof of convergence and complexity analysis):

Algorithm FSD-LC

Step 0 Initialization. Let 0^{ss} be a dummy path from the destination to itself. Initialize the scan list $Q = \{0^{ss}\}$. set $\pi_0^{ss}(b) = 1, \forall b$.

Step 1 Select the first path from Q, denoted as l^{js}, and delete it from Q.

Step 2 For any predecessor node i of j, create a new path k^{is} by extending l^{js} along link ij.

step 2.1 Calculate the distribution of π_k^{is} from the distribution of π_l^{js} by convolution integral (Details are given below).

step 2.2 Compare the new path k^{is} with current *Pareto frontier*. If the frontier is dominated by k^{is}, update the frontier with the distribution of π_k^{is}, drop all existing FSD-admissible paths at node i, and set $\Gamma^{is} = \{k^{is}\}, \Omega^{is} = \{k^{is}\}$; otherwise, further compare the distribution of the new path to those of the existing FSD-admissible paths to check FSD admissibility. If any of the existing path dominates k_{is}, drop k_{is} and go back to Step 2; otherwise, delete all paths that are dominated by k_{is} from Γ^{is}, set $\Gamma^{is} \cup \{k_{is}\}$, and update $Q = Q \cup \{k^{is}\}$.

Step 3 If Q is empty, retrieve Ω^{is} from Γ^{is}, and evaluate the Pareto frontier function $u^{is}(\cdot), \forall i$, stop; otherwise go to Step 1.

2.4 Evaluation of convolution integral

2.4.1 Direct method and its shortcomings

If the random link travel time follows a continuous probability density function p_{ij}, the distribution of path travel time π_k^{rs} can be calculated recursively from the following convolution integral

$$u_k^{is}(b) = \int_0^b u_k^{js}(b-w) f_{ij}(w) dw, \ \forall b \in [0, T] \tag{2.1}$$

where T is the longest travel time of interest. Typically, the convolution integral has to be evaluated using numerical methods which involves discretization. The simplest discretization scheme divides $[0, T]$ evenly into L intervals of length ϕ. The corresponding probability mass function P_{ij} reads

$$P_{ij}(b) = \begin{cases} \int_b^{b+\phi} p_{ij}(w)dw & b = 0, \phi, ..., (L-1)\phi \\ \int_b^{\infty} p_{ij}(w)dw & b = L\phi \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

Accordingly, the evaluation of the convolution integral in Equation (2.1) is replaced with a finite sum as follows:

$$u_k^{is}(b) = \sum_0^b u_k^{js}(b-\phi) P_{ij}(\phi), \ \forall b = 0, \phi, \cdots, L\phi \tag{2.3}$$

While simple and relatively easy to implement, this direct method has several shortcomings.

1 Since the support of the probability function u_k^{is} is evenly discretized, the number of discrete support points will grow as the convolution is carried along the path. It is hard to bound the number of support points precisely a priori, because it requires estimating the shortest and longest possible travel time between a given O-D pair (denoted as T_0 and T_1 respectively). Recognizing that estimating T_0 and T_1 is a difficult exercise in its own right, several previous studies (e.g., Fan & Nie 2006, Nie & Wu 2009*b*) simply set the support of all distributions as $[0, \tilde{T}]$ (where $\tilde{T}$ is a predetermined large number), and divide it evenly. This approach is not only arbitrary but often turns out to be very inefficient. Alternatively, one can employ a dynamic data structure to manage the growth of the discrete support points. However, since the number of support points is not bounded, a dynamic data structure may lead to excessive memory and computational overheads.

2 The uniformly spaced discrete points are not effective in representing heterogenous concentrations of probability mass. They often over-represent the flat portions on the probability density function, and fail to capture sharp changes. As a result, it is difficult to achieve reasonable accuracy unless the resolution ϕ is sufficiently small.

3 The direct discrete convolution (2.3) has a quadratic complexity $O(L^2)$, where L is the upper bound for the number of discrete support points. When L is large (in thousands or larger), the direct method may be too slow to be practically useful. It is known that convolution based on Fast Fourier Transform (FFT) has a complexity of $O(L \log L)$, which is a significant improvement in term of computation speed. The FFT technique has so far not been adapted in solving reliable routing problems, to the best of our knowledge.

2.4.2 An adaptive discretization scheme

The above discussion suggests that a fixed number of discrete support points should be used to represent any distribution, even if the range of the support may grow with convolution. Accordingly, the length of the discrete intervals ϕ should change with the range of the support. In light of this observation, Wu & Nie (2009) proposed to divide the support such that each discrete interval has the same predetermined probability mass ϵ. The scheme discretizes the support $[T_0, T_1]$ to $L = [1/\epsilon]_-$ intervals, where $[a]_-$ denotes the largest integer smaller than or equal to a. Since the discrete interval is not of uniform length, the formula (2.3) is no longer applicable. Instead an alternative method was proposed to perform the convolution in (Wu & Nie 2009). Wu and Nie's method is not fully satisfactory for two reasons. First, it still poorly responds to the irregular concentration of probability mass. Note that an identical probability mass is required for each interval, which can produce unnecessarily short intervals where the concentration is high. Second, the corresponding convolution requires an extra sorting over a vector of L^2, which increases the complexity of the overall operation to $O(L^2 \log L)$. In the following, we first introduce a discretization scheme fully adaptive to the probability mass concentration, and then propose a convolution method.

Our adaptive discretization scheme consists of two steps. In the first, the range of support of a given distribution $[T_0, T_1]$ is divided into to a set of L uniform intervals $[b_0, b_1], [b_1, b_2], ...[b_{L-1}, b_L]$, where $b_0 = T_0, b_L = T_1, b_i - b_{i-1} = (T_1 - T_0)/L, \forall i = 1, L$. The probability mass for each interval $[b_{i-1}, b_i]$, denoted as p_i, is either directly obtained from an empirical distribution or computed using Equation (2.2). Consequently, the scheme represents any distribution with no more than L discrete support points, regardless of the actual range. The second step, called *consolidation*, merges consecutive intervals such that no interval has probability mass smaller than $1/L$ (with the exception of the last

interval). As shown below, the consolidation produces a set of effective support points $W = \{w_1, w_2, ..., w_{L_0}\}$ and a corresponding probability mass vector $Q = \{q_1, q_2, ..., q_{L_0}\}$, where L_0 is the number of effective support points.

Consolidation

Input: Vectors of initial discrete intervals and probability $\{b_0, b_1, ..., b_L\}, \{p_1, ..., p_L\}$

Output: Number of effective support points, L_0 and $\{w_1, ..., w_{L_0}\}, \{q_1, ..., q_{L_0}\}$

Step 0 Initialization. Set $l = 1, i = 1, \epsilon = 1/L$. Set $tp = 0, ts = 0$.

Step 1 if $i > L$; stop, set $L_0 = l$; otherwise, $tp = p_i, ts = 0.5(b_{i-1} + b_i)p_i$, go to Step 2.

Step 2 If $tp > \epsilon$, $q_l = tp, w_1 = ts/tp$, set $i = i + 1, l = l + 1$, go to Step 1; otherwise, go to Step 3.

Step 3 Set $i = i + 1, tp = tp + p_i; ts = ts + 0.5(b_{i-1} + b_i)p_i$, go to Step 2.

Remark: When several consecutive intervals are merged to meet the minimum probability mass requirement, the discrete support for the merged interval is obtained by taking the probability-weighted average of the original discrete support, which is assumed to locate at the center of the discrete interval (i.e., $0.5(b_{i-1} + b_i)$).

After the consolidation, the probability density function $f(t)$ can be approximated by

$$f(t) = \begin{cases} q_1/(w_2 - w_1) & t \in [1.5w_1 - 0.5w_2, 0.5(w_1 + w_2)] \\ q_{L_0}/(w_{L_0} - w_{L_0-1}) & t \in [0.5(w_{L_0} + w_{L_0-1}), 1.5w_{L_0} - 0.5w_{L_0-1}] \\ q_i/(0.5(w_{i+1} - w_{i-1})) & t \in [0.5(w_i + w_{i-1}), 0.5(w_i + w_{i+1})], i = 2, ..., L_0 - 1 \end{cases} \tag{2.4}$$

Thus, the cumulative probability function $F(t)$ can be obtained similarly from $f(t)$. Moreover, we note that the mean (ν) and variance (σ^2) of a consolidated distribution can be computed by

$$\nu = \sum_{i=1}^{L_0} q_i w_i, \sigma^2 = \sum_{i=1}^{L_0} q_i (w_i - \nu)^2 \tag{2.5}$$

We now illustrate the above discretization scheme using a numerical example. Consider the Gamma distribution whose probability density function is given by

$$f(x) = \frac{1}{\theta^\kappa \Gamma(\kappa)} (x - \mu)^{\kappa-1} e^{-(x-\mu)/\theta}; x \geq \mu, \theta, \kappa \geq 0 \tag{2.6}$$

where θ is the *scale parameter*; κ is the *shape parameter*; μ is the *location parameter*; and $\Gamma(\cdot)$ is the Gamma function. In this example, we set $\mu = 0, \kappa = 3, \theta = 2$. Figure 2.2-(a) shows the discrete probability mass function produced from a 100,000-point sample ($L = 100$),

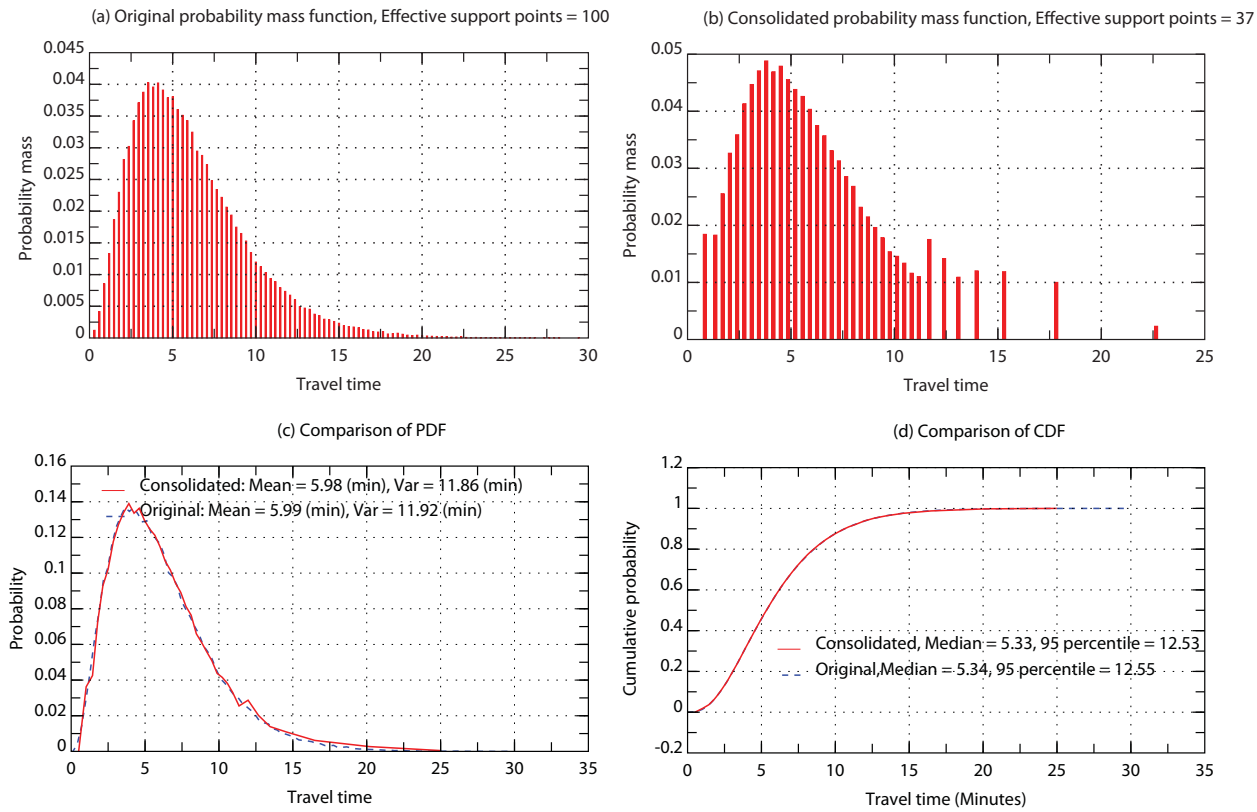


Figure 2.2: Comparison of original and consolidated distributions

whereas Figure 2.2-(b) demonstrates its consolidated counterpart. Note that, due to consolidation, (1) the probability mass in the figure is always larger than 0.01 except at the rightmost support point, and (2) the discrete support points are no longer uniformly spaced. Perhaps more important, the consolidated distribution only uses about one third as many discrete support points to satisfactorily represent the original distribution. As shown in Figure 2.2-(c) and 2.2-(d), the PDF and CDF of the consolidated distribution (produced based on (2.4)) is almost identical to that of the original distribution. In particular, note that the mean and variance closely replicate theoretical values ($\nu = 6, \sigma^2 = 12$). Thus, consolation seems to promise significant computational benefits in evaluating the convolution integral, without sacrificing numerical accuracy.

2.4.3 Convolution based on the adaptive discretization scheme

We proceed to show how the convolution can be performed using the adaptive discretization scheme. Consider two random variables x and y, which, after discretization, can be represented by a set of discrete support points, W_x and W_y, and associated prob-

ability mass vectors, Q_x and Q_y. Let the number of effective support points for x and y be D and E respectively. That is, we have

$$W_x = [w_1^x, ..., w_D^x], W_y = [w_1^y, ..., w_D^y], Q_x = [q_1^x, ..., q_E^x], Q_y = [q_1^y, ..., q_E^y]$$

The objective is to compute $z = x \oplus y$, where $\oplus$ denotes convolution integral. We first present a direct algorithm that has a complexity of $O(L^2)$, where $L \geq \max(D, E)$ is the maximum number of discrete support points as specified before.

Direct Convolution

Step 0 Set $b_0^z = 1.5(w_1^x + w_1^y) - 0.5(w_2^x + w_2^y), b_z^L = 1.5(w_D^x + w_E^y) - 0.5(w_{D-1}^x + w_{E-1}^y)$. Divide $[b_0^z, b_K^z]$ into L intervals of uniform length, and compute $\phi = (b_0^z - b_L^z)/L$. Initialize $p_l^z = 0, \forall l = 1, \cdots, L$.

Step 1 for $i = 1, 2, \cdots, D$,

for $j = 1, 2, \cdots, E$,

Calculate $ts = w_i^x + w_j^y$ and $tp = p_i^x p_j^y$. Define $l = \left[\frac{ts - b_0^z}{\phi}\right]_-$, and set $p_l^z = p_l^z + tp$

end for

end for

Step 2 Call Procedure Consolidate to get effective support points and the associated probability mass vector.

It is well-known that convolution of two integrable functions can be reduced to pointwise multiplication of their corresponding Fourier Transforms (see e.g., Bracewell 2000). That is, if $\hat{f}(s)$ denotes the Fourier Transform of function $f(x)$, i.e.,

$$\hat{f}(s) = \int_{-\infty}^{\infty} f(t) e^{-2\pi its} dt, \tag{2.7}$$

and $h(t) = f(t) \oplus g(t)$, then

$$\hat{h}(s) = \hat{f}(s) \cdot \hat{g}(s) \tag{2.8}$$

Note that i in Equation (2.7) is the imaginary unit, i.e., $i = \sqrt{-1}$. In the discrete case, the sequence of L complex numbers $[p_1, ..., p_L]$ can be transformed into the sequence of complex numbers $[\hat{p}_1, ..., \hat{p}_L]$ using the following formula

$$\hat{p}_k = \sum_{l=1}^{L} p_k \exp\left(\frac{-2\pi i}{L}(k-1)(l-1)\right), k = 1, ..., L \tag{2.9}$$

Moreover, the inverse discrete Fourier Transform reads

$$p_l = \sum_{k=1}^{L} \hat{p}_k \exp\left(\frac{2\pi i}{L}(k-1)(l-1)\right), l = 1, ..., L \tag{2.10}$$

Since the discrete Fourier Transform can be computed in time $O(L \log L)$ using the Fast Fourier Transform (FFT) algorithm (e.g., Cormen, Leiserson, Rivest & Stein 2001), the convolution based on discrete Fourier Transform has a theoretical complexity of $O(L \log L)$ (note that it involves two FFTs, one forward, one backward, and a point-wise multiplication). However, to apply the FFT technique in convolution requires that the discrete intervals of both distributions have the same length, which is inconsistent with the adaptive discretization scheme adopted in this study. Wu & Nie (2011*a*) closely examined the computational issues associated with FFT-based convolution methods. Becasue the results are inconclusive, the implementations of these methods are not used in this project.

Chapter 3

Data preparation

This chapter discusses issues associated with input data necessary to creating travel reliability inventory. The focus is on the construction of link travel time distributions from loop detector data.

The GCM database are used as the primary source for traffic data on expressways (including both freeways and toll roads). The GCM data come from two main sources: loop detectors and electronic toll transponders (I-PASS), which cover freeways and toll roads respectively. Figure 3.1 shows the topology of the expressway subnetwork in which the colored links represent those covered by either or both of the data sources. As revealed in the figure, many links in this network do not have data. For one thing, not all expressway road segments have data because there is only about one loop detector per mile. Some links may fall into the "gaps" between two loop detectors. More importantly, to maintain connectivity, many local streets are added when building the expressway network. No data are available on these streets at all. In what follows, we first describe procedures used to construct travel time distributions on links that have data, and then present methods to estimate distributions on those that do not. Note that similar results have been reported in (Nie, Wu, Dillenburg & Nelson 2010) and they are presented here only for the convenience of the reader. Finally, we show how to compute the reliability indexes from a given travel time distribution.

3.1 Links with data

For links covered by loop detector(s), the recorded speed in a 5-minute interval is used to estimate link travel time for the corresponding interval, i.e.,

$$\tau_a^d(t) = l_a / v_a^d(t) \tag{3.1}$$

where $\tau_a^d(t)$ and $v_a^d(t)$ are travel time and speed on link a recorded by detector d for time interval t, and l_a is the link length. If for an interval, a link contains more than one

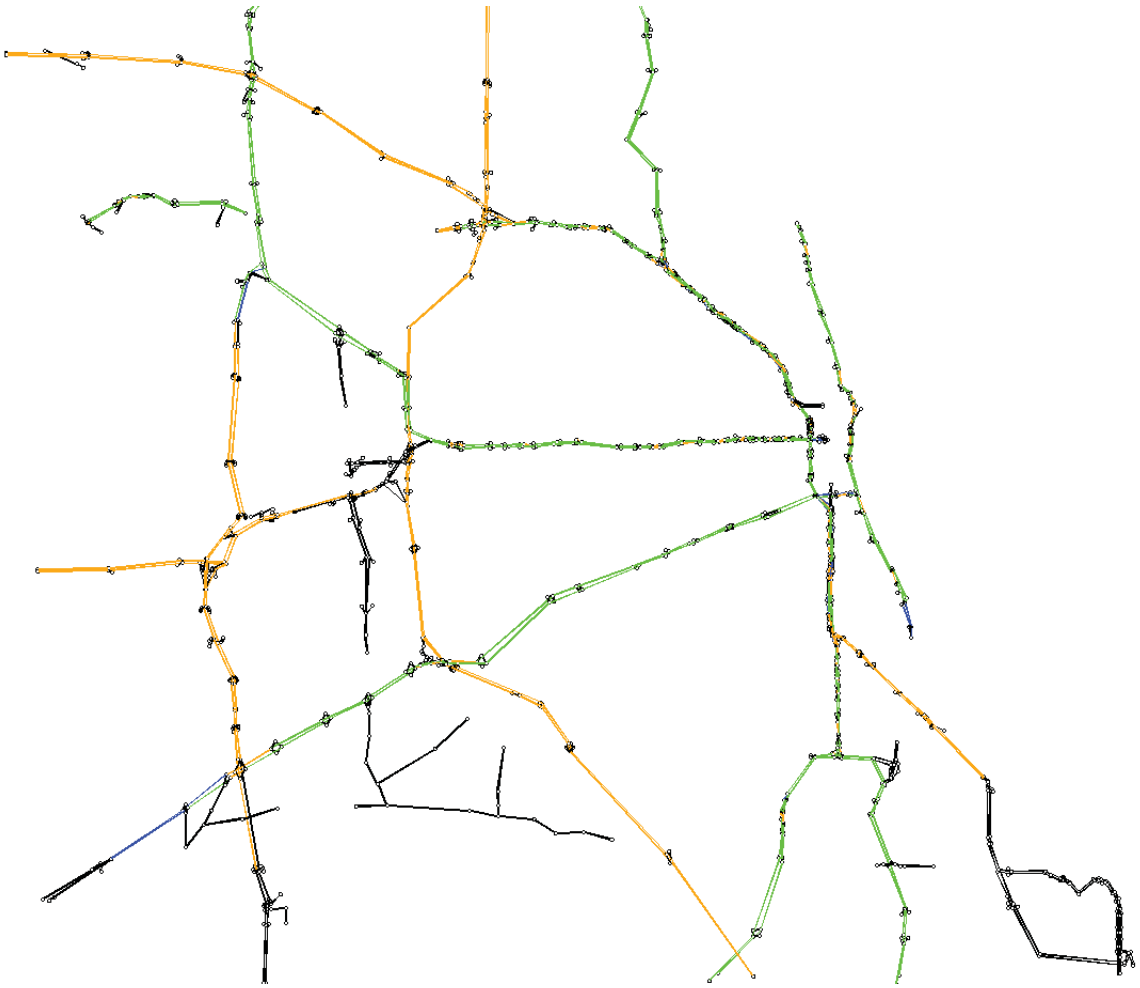


Figure 3.1: Data coverage of the expressway network (colored links are covered by either loop detectors or traffic report from IPASS data. Blue: covered by loop detectors only; Orange: covered by IPASS only; Green: covered by both sources)

recorded travel time, the arithmetic average of calculated travel time values is taken as the nominal link travel time, that is,

$$\tau_a^d(t) = \frac{\sum_{d \in D(a)} l_a / v_a^d(t)}{|D(a)|} \tag{3.2}$$

where $D(a)$ is the set of loop detectors associated with link a at a given time. As for I-PASS detectors, we need to estimate link travel times on covered links based on the recorded path travel times for a given time period. This is a difficult exercise for two reasons. First, how the travel delays (if there is any) experienced on a path may be *spatially* distributed is unknown. Second, an I-PASS record tagged by one time interval might contribute to link travel times at other time intervals [1]. It is hard to solve either problem unless further information is available, such as supplementary loop detector data. For simplicity, we assume that path travel times are distributed to links proportional to their lengths, that is

$$\tau_a^i(t) = \frac{l_a}{\sum_{a \in k^{rs}} l_a} c_k^{rs}(t) \tag{3.3}$$

where k^{rs} denotes the shortest path connecting nodes r (the starting node of the link associated with the origin toll booth) and s (the ending node of the link associated with the destination toll booth), and $c_k^{rs}(t)$ is the recorded travel time on the path for time interval t. While this simplification would certainly introduce errors, we note that the magnitude of errors may be alleviated when multiple I-PASS records are available for the same stretch of toll roads. Equation (3.3) also implies that the travel time on a path at one time interval contributes to its covered links for the same interval. This shortcoming is not as serious as it sounds, since eventually the travel time data will be aggregated on a period of a couple of hours. That is to say, as long as the misplaced link travel times do not go into a wrong period (which is certainly possible but is much less likely), they will not seriously distort the aggregated distributions. To summarize, the travel time on link a at time t is given by

$$\tau_a(t) = \begin{cases} \tau_a^d(t) & \text{if loop data are available} \\ \tau_a^i(t) & \text{if I-PASS data are available} \end{cases} \tag{3.4}$$

Once link travel times are obtained, the empirical distributions can be constructed using the following procedure.

Construct Empirical Distribution

[1] Note that the time interval that identifers an I-PASS record must be tied to either the entry (origin) or exit (designation), since the travel times between most I-PASS toll collection plazas are longer than 5 minutes.

Step 1 Find $L_a = \min\{\tau_a(t), \forall t \in \Lambda\}, U_a = \min\{10 l_a / v_a^0, \max\{\tau_a(t), \forall t\}\}$, where Λ is a set of valid time intervals, and v_a^0 is free flow speed (or speed limit) on link a. Note that the maximum possible travel time on a link is bounded by 10 times free flow travel time.

Step 2 Divide $[L_a, U_a]$ into M intervals, and let $\delta_a = (U_a - L_a) / M$. Find the set $D_m = \{\tau_a(t) | \forall t \in \Lambda, (m-1)\delta a \leq \tau_a(t) < m\delta\}, \forall m = 1, ..., M$

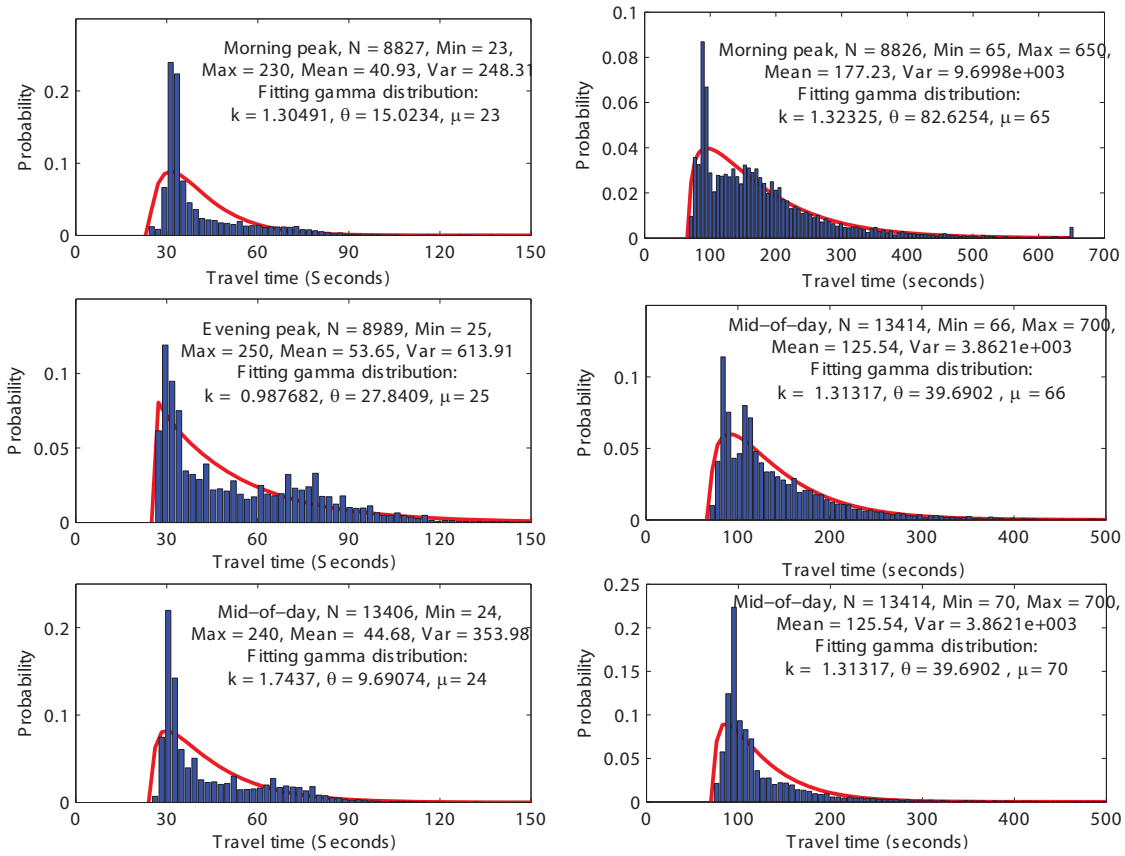
Step 3 Obtain the probability mass for each interval m using $P_m = \frac{|D_m|}{|\Lambda|}$.

It is noted that link travel time distribution may be affected by various factors, such as time-of-day and day-of-week. Conceivably, the most reliable route for a given O-D pair may be different for rush hour and off-peak periods. To address this issue, the travel time data are disaggregated according to two key factors: time-of-day and day-of-week. Specifically, each day is divided into four periods, namely, morning peak period (6 am - 10 am), mid-of-day period (10 am - 3 pm), evening peak period (3 pm - 8 pm) and off-peak period (8 pm - 6 am). Days in a week are first grouped into weekends and weekdays. In addition, Friday, Saturday and Sunday form individual groups because they have somewhat different travel patterns. For each of the two factors, an additional group is added to address the case of no-segmentation. For instance, the segmentation for time-of-day contains 5 instead of 4 groups: morning peak, mid-of-day, evening peak, off-peak and whole-day (no segmentation for time-of-day). Therefore, in total, there are $5 \times 6 = 30$ possible combinations. Accordingly, we generate 30 different distributions for each of the covered links. Figure 3.2 plots the histograms of travel time distributions on two links (Links 19185 and 3087) for three periods (morning peak, evening peak and middle-of-day) of a weekday using the CMAP data.

3.2 Links without data

The travel time distributions on these links have to be estimated indirectly. The estimation process involves two main steps: select an appropriate functional form, and estimate its parameters.

Road travel times are known to roughly follow a Gamma distribution (e.g., Polus 1979). To verify this result, we fit the distributions reported in Figure 3.2 using Gamma distributions. The best fitted gamma distributions are obtained using MATLAB's `fitdist` function, and the resulting PDFs are added to Figure 3.2 for comparison. As shown in the figure, the freeway travel time data can be satisfactorily fit using the Gamma distribution. Thus, the Gamma distribution is adopted to describe the travel time distribution on arterial and local streets in this study. Having selected the functional form (cf. Equation



(a) Link 19185 (b) Link 8097

Figure 3.2: Original travel time distributions and the fitting gamma distribution curves for two expressway links in the network (aggregated over all weekdays in Spring) N: number of observations. κ, θ, μ are defined as in Equation (2.6).

(2.6)), we proceed to show how to estimate the three parameters required in a Gamma distribution. We first note that the mean and variance of a Gamma distribution are $\kappa\theta$ and $\kappa\theta^2$, respectively. Thus, if we know mean (denoted as ν), variance (denoted as σ^2) and μ, then κ and θ can be obtained by

$$\theta = \frac{\sigma^2}{\nu - \mu}, \kappa = (\frac{\nu - \mu}{\sigma})^2 \tag{3.5}$$

The CMAP travel demand model (CMAP 2006) is used to estimate congested travel times on arterial streets. Note that the planning model is designed to capture average traffic conditions in the network for the designated time period on a typical weekday. The CMAP travel demand model represents a classical four-step process of *trip generation*, *trip distribution*, *mode choice*, and *traffic assignment*, with considerable modifications used to enhance the distribution and mode choice procedures. The original CMAP model divides a day into eight periods: off-peak (8 PM - 6 AM), pre-morning-peak (6-7 AM), morning-peak (7 - 9 AM), post-morning-peak (9-10 AM), mid-of-day (10 AM - 2 PM), pre-evening-peak (2 - 4 PM), evening-peak (4 - 6 PM), and post-evening-peak (6 - 8 PM). Note that in GCM data peak periods combine the pre and post periods defined in the CMAP model. For simplification, the assignment results for the peak periods (morning and evening) in the CMAP model are used to represent those from pre to post peak periods. Specifically, each link obtains from the CMAP model a mean travel time for each of the four predetermined GCM periods: morning peak, mid-of-day, evening peak and off-peak.

We postulate that the mean and variance of travel times on a link are related to its free flow travel time τ^0 and the level of congestion $\rho = \tau - \tau^0$, where τ is travel time from traffic assignment (note that the subscript a is suppressed for simplicity). This relationship may be estimated from freeway data using statistical models. The simplest linear regression model reads

$$\nu = a_1\tau^0 + b_1\rho + c_1 \tag{3.6}$$

$$\sigma = a_2\tau^0 + b_2\rho + c_2 \tag{3.7}$$

where a_1, b_1, c_1, a_2, b_2 and c_2 are coefficients to be estimated from linear regression. For all 765 links covered by GCM data, ν and σ can be obtained from the empirical distribution and ρ is known from the CMAP travel demand model. Thus, a linear regression can be performed to determine the coefficients, which in turn are employed to estimate ν and σ for arterial streets. We note that a linear model is needed for each of the four time-of-day periods.

A similar linear model can be constructed to estimate the location parameter μ, which delineates the smallest possible travel time on a link. We note that μ is likely to be smaller

Table 3.1: Results of the linear regression for the mean-variance model (Equations 3.6-3.7) and the location model (Equation 3.8)

time-of-day	Variance Model				Mean Model				Location Model		
periods	a_1	b_1	c_1	R^2	a_2	b_2	c_2	R^2	a	b	R^2
AM PEAK	0.309	0.870	0.580	0.444	1.127	0.546	-2.056	0.910	0.843	-4.106	0.958
PM PEAK	0.368	0.685	2.967	0.400	1.143	0.563	0.336	0.872	0.860	-3.533	0.964
MIDDAY	0.283	1.076	2.040	0.346	1.100	0.630	-1.145	0.889	0.857	-3.608	0.956
OFF PEAK	0.178	0	-1.031	0.516	1.043	0	-5.854	0.907	0.831	-5.257	0.937

than τ_0 because motorists may drive well beyond the speed limit or the nominal "free-flow travel speed". Moreover, it seems reasonable to assume that the level of congestion does not affect μ. Thus, the linear model used to estimate μ is given by

$$\mu = a\tau^0 + b \tag{3.8}$$

The linear regression results for each of the four time-of-day periods are given in Table 3.1. As shown, the mean model and the location model fit the data rather well (high R^2). Note that the coefficient of the congestion index is near zero in both mean and variance models for the off-peak period because congestion is negligible in that period. The fitness of the variance model is not impressive. Various forms of non-linearity into the model were tried, such as using the variance (σ^2) instead of the standard deviation on the left hand side of Equation (3.7), or considering ρ^2 on the right hand side. However, these transformations made little difference in R^2 value. Apparently, the travel time variances are affected by other factors not included in the simple linear model. While a more in-depth investigation of the variance model is clearly needed, we believe that the result is good enough for a proof-of-concept study.

For illustration purpose, Figure 3.3 plots the estimated travel time distributions on a road segment of Michigan Ave. in Downtown Chicago The results show that the street is slightly more congested during the mid-of-day and evening peak periods.

3.3 Computation of reliability indexes

In this project, the basic element of the travel reliability inventory consists of three indexes: 95-percentile travel time, buffer index (B-index) and planning index (P-index). For each link a, these indexes are denoted as T^a_{95}, I^a_B and I^a_P, respectively [2]. Let $\bar{\tau}_a$ be the

[2]These indexes can be similarly defined for a route r, provided that the distribution of travel time on the route is known.

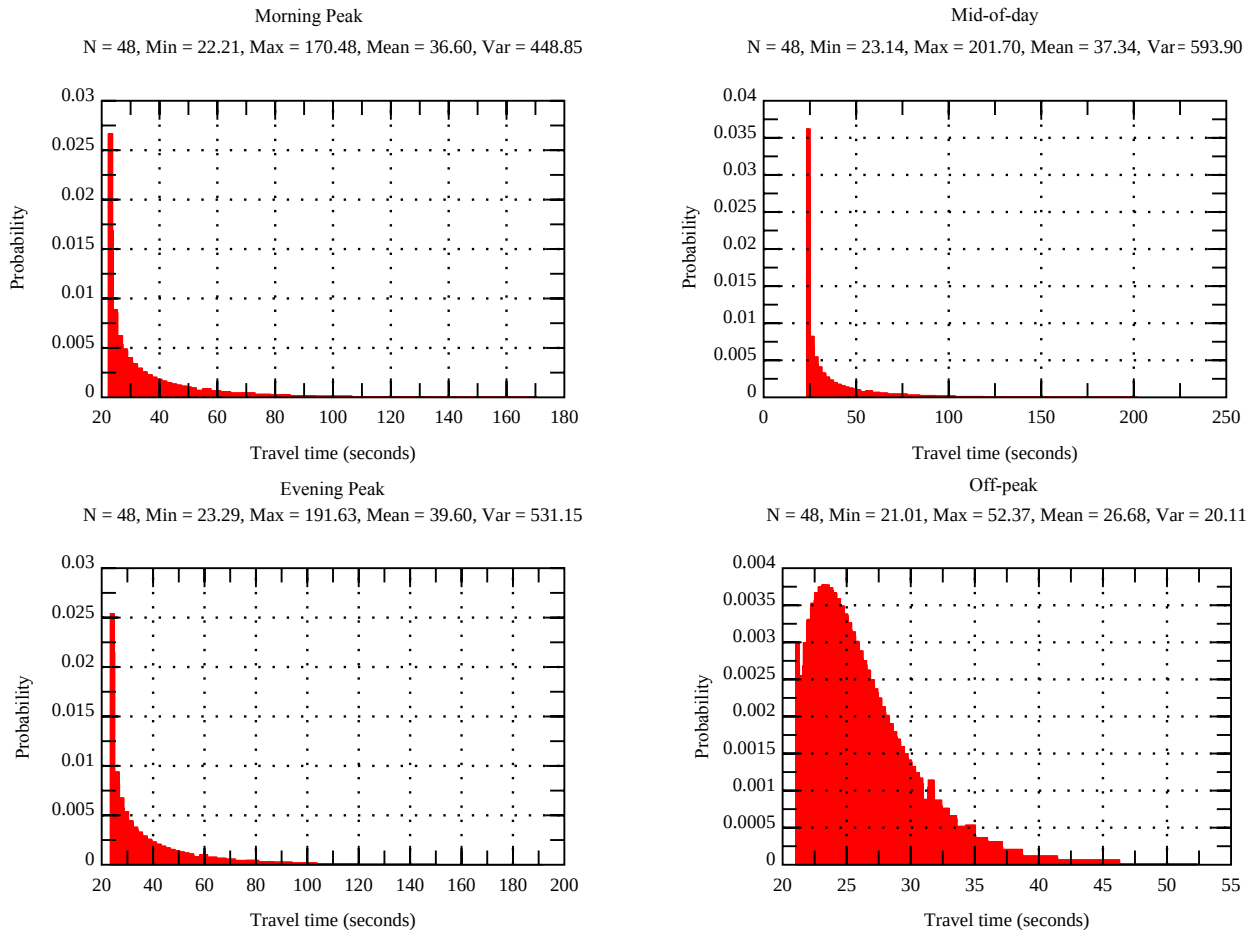


Figure 3.3: Estimated travel time distributions for different time periods on a road segment in Downtown Chicago (length = 0.26 mile, free flow speed = 30 mph)

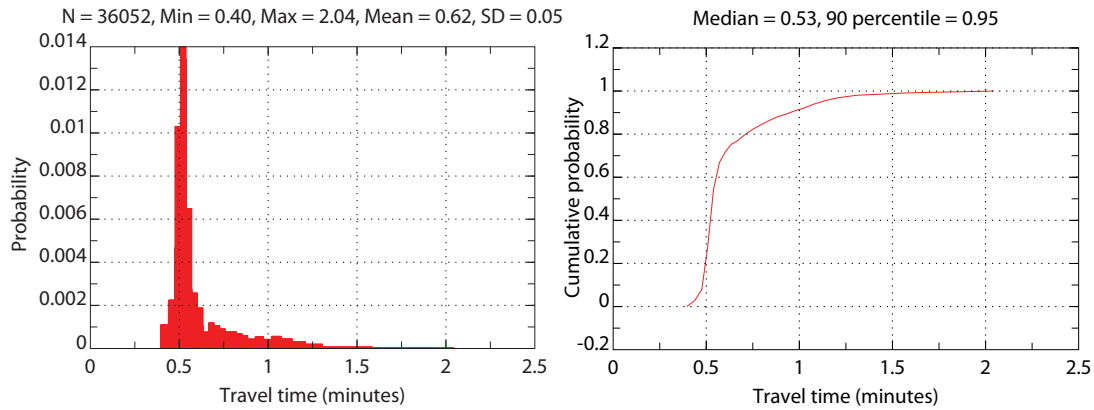


Figure 3.4: Travel time distribution on link 19815 built over the morning peak period of all weekdays

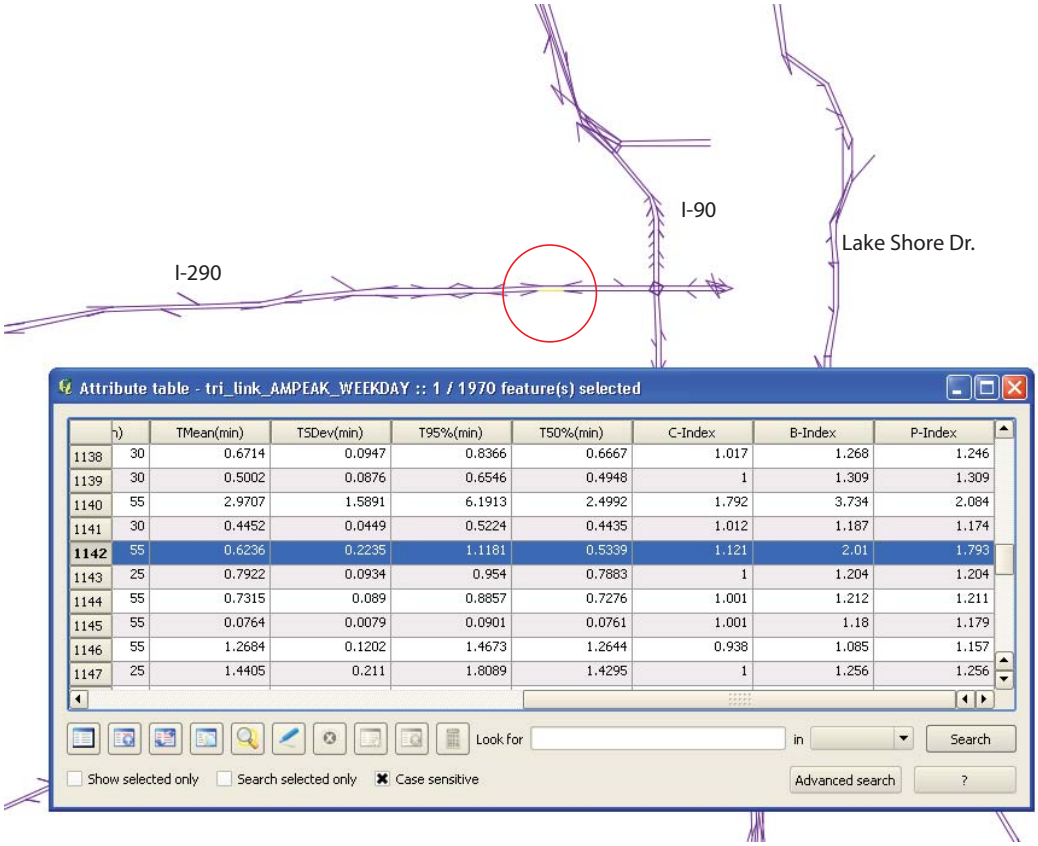
average travel time on link a, τ_a^0 be the free-flow travel time (computed from the speed limit), and F_a be the cumulative density function of the travel time on link a. The above reliability indexes can be calculated as follows:

$$T_{95}^a = F_a^{-1}(0.95) \tag{3.9}$$

$$I_P^a = \frac{T_{95}^a}{\bar{\tau}_a} \tag{3.10}$$

$$I_B^a = \frac{T_{95}^a}{\tau_a^0} \tag{3.11}$$

For an illustration, Figure 3.4 shows the travel time distribution on link 19815 over the morning peak period of all weekdays. In Figure 3.5, the location of the link and the values of the reliability indexes (computed based on the above formula), are presented.



	n)	TMean(min)	TSDev(min)	T95%(min)	T50%(min)	C-Index	B-Index	P-Index
1138	30	0.6714	0.0947	0.8366	0.6667	1.017	1.268	1.246
1139	30	0.5002	0.0876	0.6546	0.4948	1	1.309	1.309
1140	55	2.9707	1.5891	6.1913	2.4992	1.792	3.734	2.084
1141	30	0.4452	0.0449	0.5224	0.4435	1.012	1.187	1.174
1142	55	0.6236	0.2235	1.1181	0.5339	1.121	2.01	1.793
1143	25	0.7922	0.0934	0.954	0.7883	1	1.204	1.204
1144	55	0.7315	0.089	0.8857	0.7276	1.001	1.212	1.211
1145	55	0.0764	0.0079	0.0901	0.0761	1.001	1.18	1.179
1146	55	1.2684	0.1202	1.4673	1.2644	0.938	1.085	1.157
1147	25	1.4405	0.211	1.8089	1.4295	1	1.256	1.256

Figure 3.5: Location of Link 19185 and its reliability indexes

Chapter 4

Introduction to TRIC

The RASP algorithm described in Chapter 2 has been previously implemented on the top of Toolkit of Network Modeling (TNM), a C++ class library for solving various transportation network problems (Nie 2006). This project develops a specialized Windows desktop application named Travel Reliability Inventory for Chicago (TRIC) using MFC (Microsoft Foundation Class Library) and MYSQL (a free database management tool). TRIC integrates the RASP algorithm and the GCM database with an easy-to-use graphic user interface (GUI). With TRIC, the travel reliability inventory can be easily created and exported to shape files for post analyses. In the following, we first introduce the overall design and data management strategies adopted by TRIC, and then present a detailed user manual.

4.1 Design and data management

TRIC is integrated with processed traffic data required to create travel reliability inventory. That is, network and travel time data will be automatically loaded when TRIC is installed. This project does not develop any interface between the internal data files and potential external sources, because such efforts are beyond the scope of the proposed work. However, it is certainly possible to develop such interfaces if needed (for instance, when the similar analysis has to be performed for other regions in the state). Such an interface may be developed to create internal TRIC files from network topology and link travel time distributions (for different time periods and day of week).

A main design issue in TRIC is how to manage the generated travel reliability inventory. For several reasons we propose to store the inventory using the ESRI shape files. First, the shape file format provide built-in GIS support to accommodate the geographical information in the inventory (note that links and routes referred in the inventory have to be identified on a map, eventually). Second, storing the inventory in the shape

file format allows the end user to access it using GIS tools (free GIS tools are abundant nowadays, such as Quantum GIS, MapWindowsGIS, Udig). Finally, the shape files can be easily converted to other formats, notably, tables in database. For instance, shape files can be easily loaded into postgresql (another popular free database management tool), retaining all geometrical information in the generated tables.

The installation folder of TRIC contains two data-related folders. The folder named `data` stores all the inputs, including network files and traffic data files. The folder named `tric` stores the travel reliability inventory (TRI). For each combination of time period and day-of-week (in total, there are 30 combinations), two files are needed to store TRI, one for all links, the other for all O-D pairs. For instance, for the morning peak period of all weekdays, the two TRI file set are named as `tri_link_AMPEAK_WEEKDAY` and `tri_path_AMPEAK_WEEKDAY`, for link-based and path-based TRI, respectively. Note that by the convention of the shape file format, each TRI file set consists of four different files, identified by the suffix. Specifically, *.shp file is the main shape file, *.shx is the index file, *.dbf is the database file associated with the shapes, and finally *.prf defines the projection information.

4.2 Installation

You first need to download the setup file at the NUTREND web site (http://translab.civil.northwestern.edu/nutrend/). Once you open the home page, click "software" tab on the top of the page, as shown in Figure 4.1. If you click on the "Travel Reliability Inventory for Chicago" entry, a dialog will be prompted that allows you to specify a local folder to save the installation file (downloading may take a while depends on your internet speed, because the data files are included). To install TRIC on your Windows machine, simply double click on the installation file (tricsetup.msi) to launch the setup program. You should see a welcome dialog as shown in Figure 4.2. At this point, you only need to accept all the default settings by clicking on the "Next" button until you reach the end of the installation. The setup program will place a TRIC icon on your Windows desktop. Double click on the icon will launch the application.

4.3 User manual

This section provides a detailed introduction to the graphic user interface of TRIC.

4.3.1 Get started

To start, click "File → Start" on the Menu bar at the upper left corner of the window to load the built-in Chicago Expressway Network, as shown in Figure 4.3. As explained

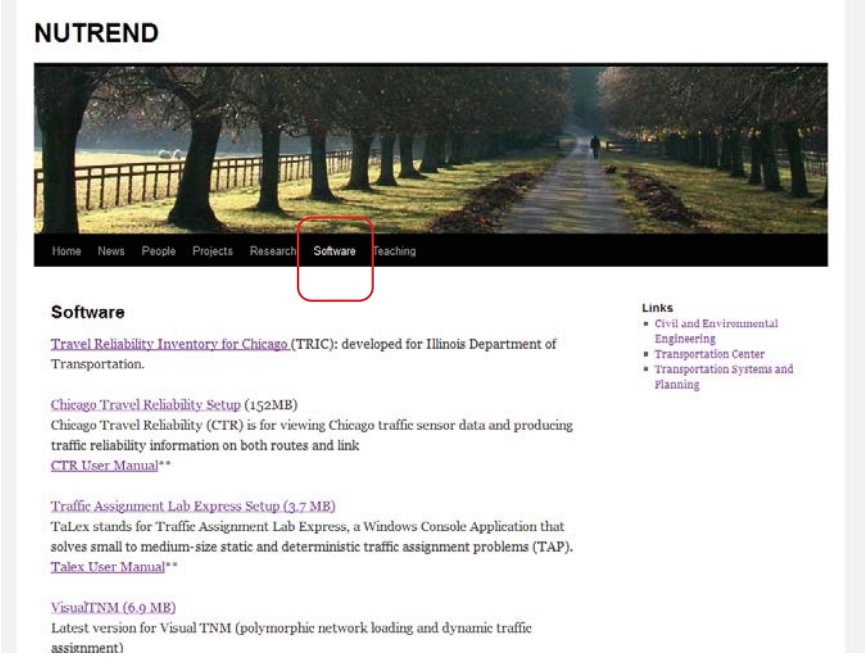


Figure 4.1: Software page at NUTREND

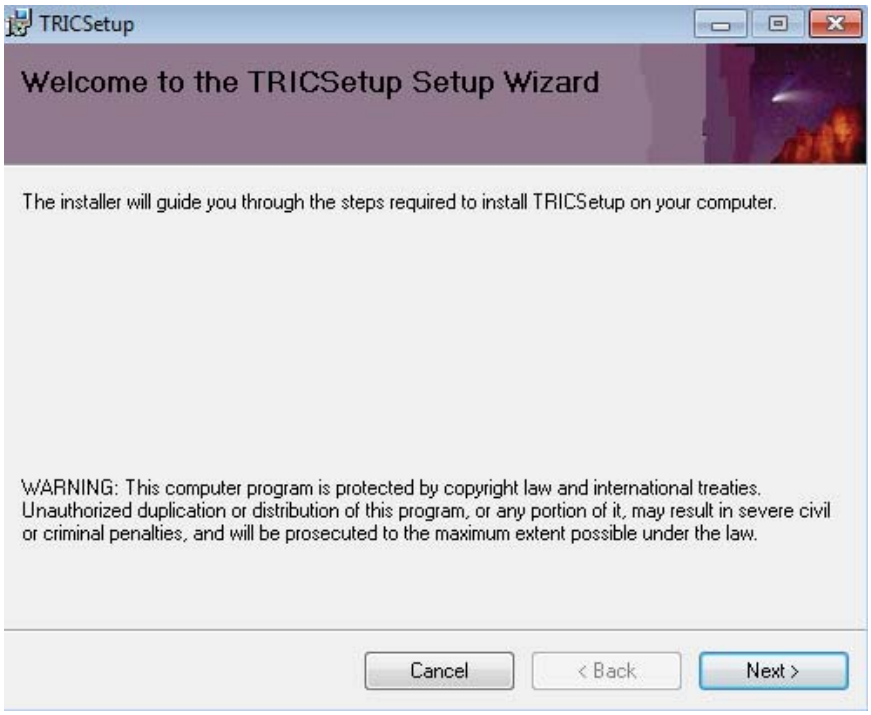


Figure 4.2: Welcome dialog of the TRIC setup program

earlier, this network is trimmed from the CMAP planning network.

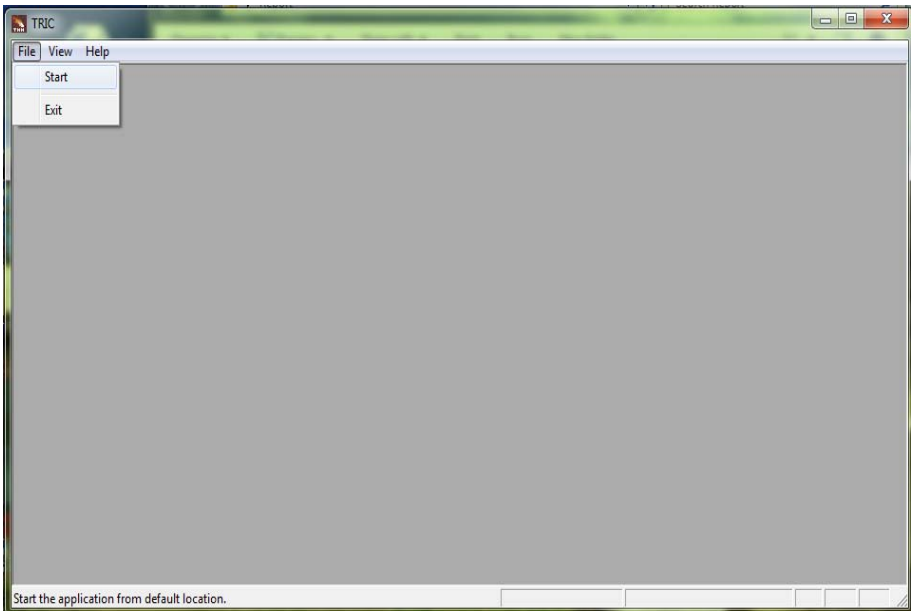


Figure 4.3: Start TRIC

After the map is loaded, the main interface of TRIC consists of the following components, as illustrated in Figure 4.4.

Title bar: The title bar displays the name of the software and map information.

Menu bar: The menu bar offers access to TRIC's main functions.

Tool bar: The tool bar provides shortcuts to frequent functions.

Map window: The map window displays the network.

Input panel: The input panel allows the user to provide necessary information for reliable path search.

Output panel: The output panel provides details of the reliable paths.

Reliability panel: The reliability panel allows the user to adjust on-time arrival probability.

Status bar: The status bar provides information about the loaded map.

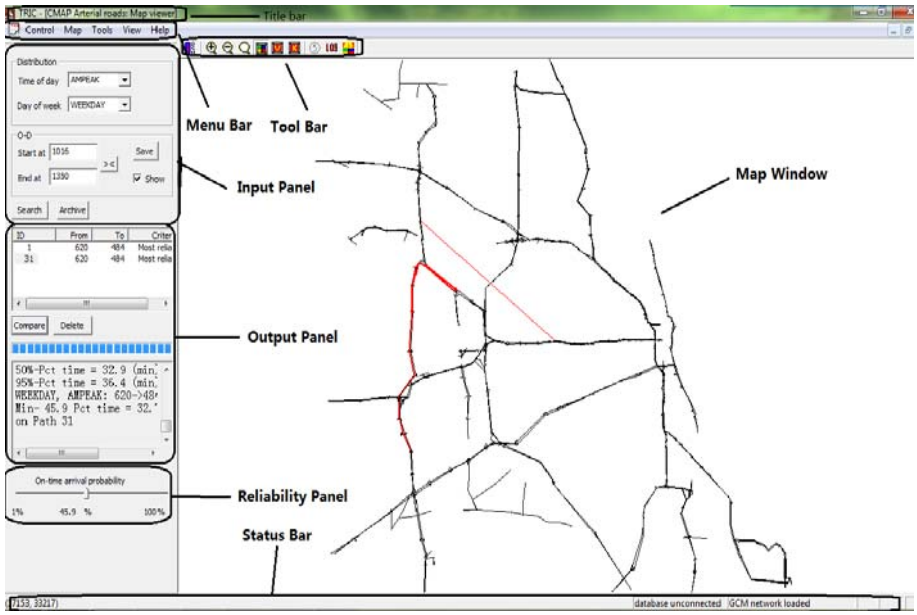


Figure 4.4: TRIC screen

4.3.2 Menu bar

The menu bar organizes the main functions into four categories: control, map, view, and window. In the following, each of these categories are described.

1. **Control:** This menu offers basic control functions (Figure 4.5).

   - **Database:** this function gives users the option of connecting to the database server located at NU-TREND. To connect to database, click "Control→Database" on the menu bar, a pop-up window is prompted out, as shown in Figure 4.6. In the pop-up window, select the server "NUTranslab" and click on "Connect" (Accept the default user name and password). Once the server is connected, the status bar would show "gcm at 129.105.69.220". To return to the main TRIC window, click "OK".
   - **Close:** To close the current map.
   - **Exit:** Exit TRIC.

2. **Map:** This category offers basic map navigation and setting functions, as shown in Figure 4.7. Note that most zoom functions can also be accessed through a context menu which can be activated by right clicking on anywhere in the map area.

   - **Zoom in/out:** Zoom in / out the current view.

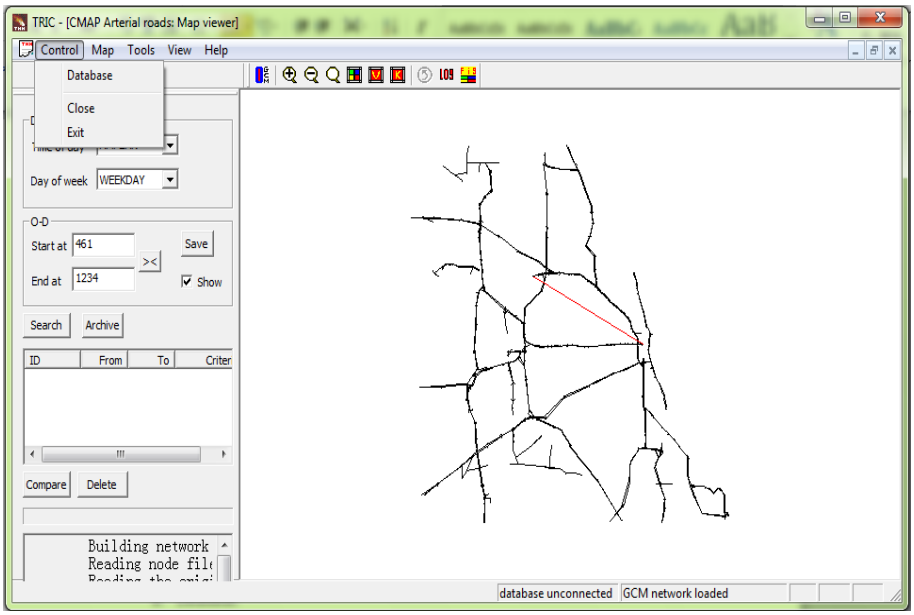


Figure 4.5: Control menu on the menu bar

- **Zoom full:** Zoom to the full size of the map.
- **Setting:** to change display options. *It is also accessible by double click on the map window.* Once click on ”Settings”, ”set display properties” window is popped up, as shown in Figure 4.8.
  - **Map size:** the default setting is ”Fit current network”. It can be changed by adjusting the width and length of the network.
  - **Text display:** to enable / disable the display of node / link name, and the text size of the display.
  - **Macroscopic:** This panel provides options to change the node size. Yet, max speed, max delay and max density are not subject to change at current stage. They are set as their default value.
  - **Microscopic:** This panel provides options to change lane width, as well as display of direction arrows and loop detectors.

3. **Tools:** This category provides main TRIC functions (see Figure 4.9).

- **Replay:** This allows the user to visualize the change of traffic data during a selected day. To use this function, the user has to connect to the MYSQL database. To enable ”replay”, switch to ”speed map” first (View→Maps→Speed). Then specify the time and data type, and click ”play” (see Figure 4.10). Traffic speed is denoted by different colors as a fraction of free flow travel time. The

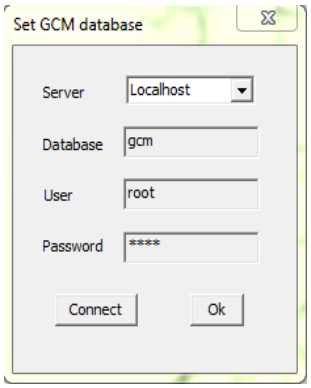


Figure 4.6: TRIC database

change of color on the map indicates the change of traffic speed. Users can control the replay by dragging the slide bar, or using F step (forward) button and B step (backward) button.

- **Log:** opens a log window which prints warnings or error massages.
- **Figure template:** The figure template command opens a PGL graph editor. All plots produced by CTR are in PGL format and can be saved as a PGL file. The saved PGL files can be loaded and edited using this function and exported to other formats such as jpeg, eps, etc. Once click on ”Figure template”, a figure template layout window is popped up to create figure template. The figure below shows a layout of 2 by 2 figures, as shown in Figure 4.11(a).
- **Archive all links:** This is the main function for creating link-based travel reliability inventory. It will save all link-based TRI to the folder ”TRI”. The user can export all link statistics in a single run (it will generate 30 shape file sets, each corresponding to a unique combinations of “Time of day” and “Day of week”). However, the user can also choose to write one shape file (corresponding only to the current distribution setting) at a time. The user can also decide whether or not the existing shape files should be overwritten or the existing shapes in the current file should be kept. As shown in Figure 4.11(b), when *archive all links*, two options are provided: ”Overwrite existing files” and ”All distributions”. The first option creates new files and overwrites all previously archived links. (Overwrite will greatly accelerate the process because it avoids checking the existence of particular shapes.) The second option determines if the operation is conducted for all combination or just the current combination.
- **Archive all O-D:** This is the main function for creating path-based travel reli-

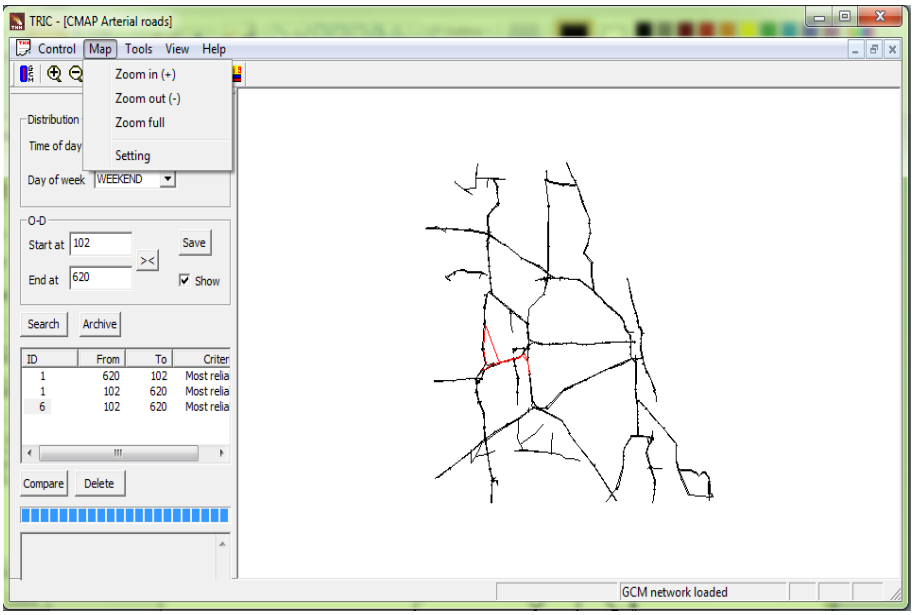


Figure 4.7: Map menu on the menu bar

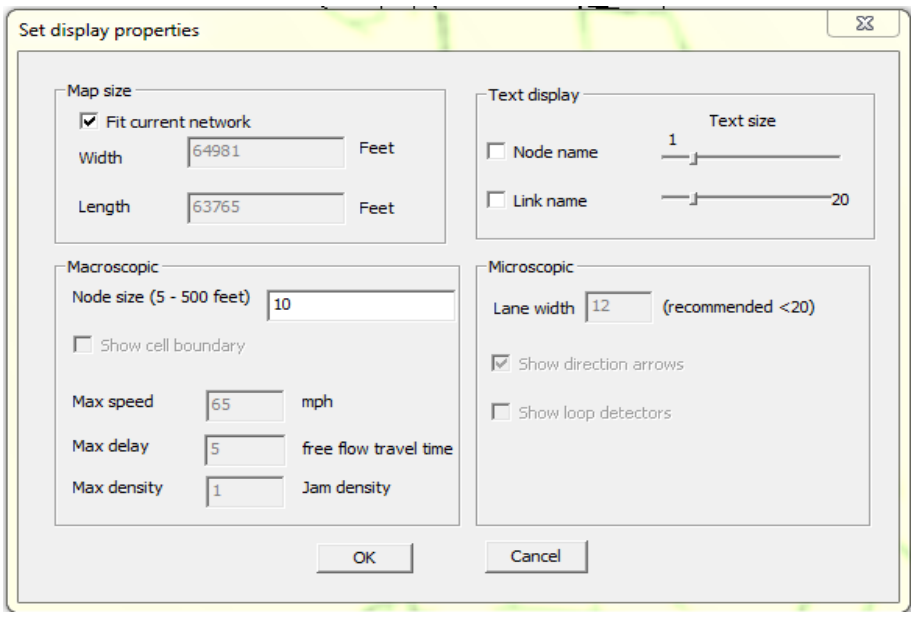


Figure 4.8: Set display properties

ability inventory. It writes all reliable path statistics for all O-D pairs stored in the *O-D pair list* into the corresponding shape files (it will generate 30 shape file sets, each corresponding to a unique combinations of "Time of day" and "Day of week") in a single run. Similarly, the user can choose whether or not

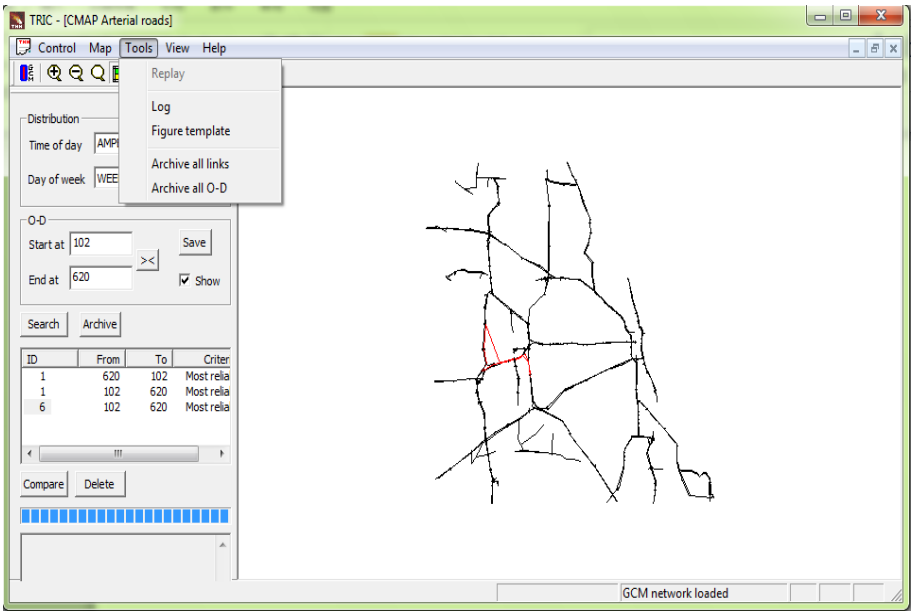


Figure 4.9: Tools menu on the menu bar

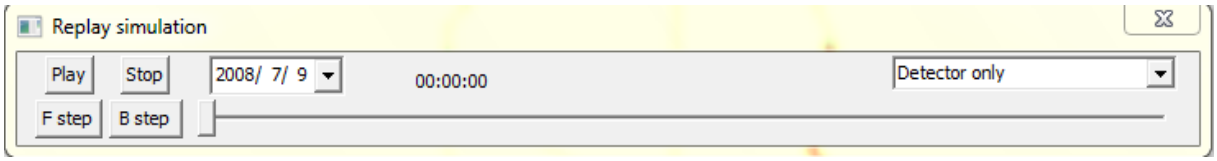
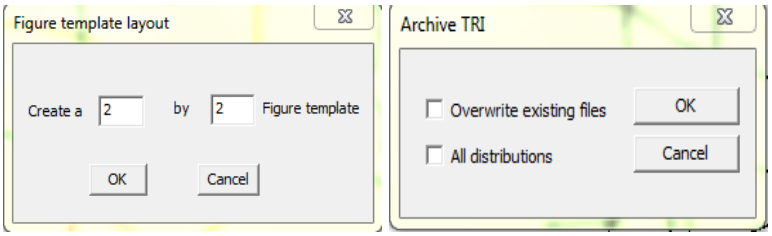


Figure 4.10: Replay function to show visualized traffic speed



(a) Figure template layout (b) Archive all links / OD

Figure 4.11: Figure template layout and the panel for archiving all links / OD

the TRI for all distributions will be generated in a single run, as well as the existing TRI files should be overwritten.

4. **View:** This category provides functions related to the display of link / OD information and other objects, as shown in Figure 4.12.

   - **Map:** The map option provides three display modes.

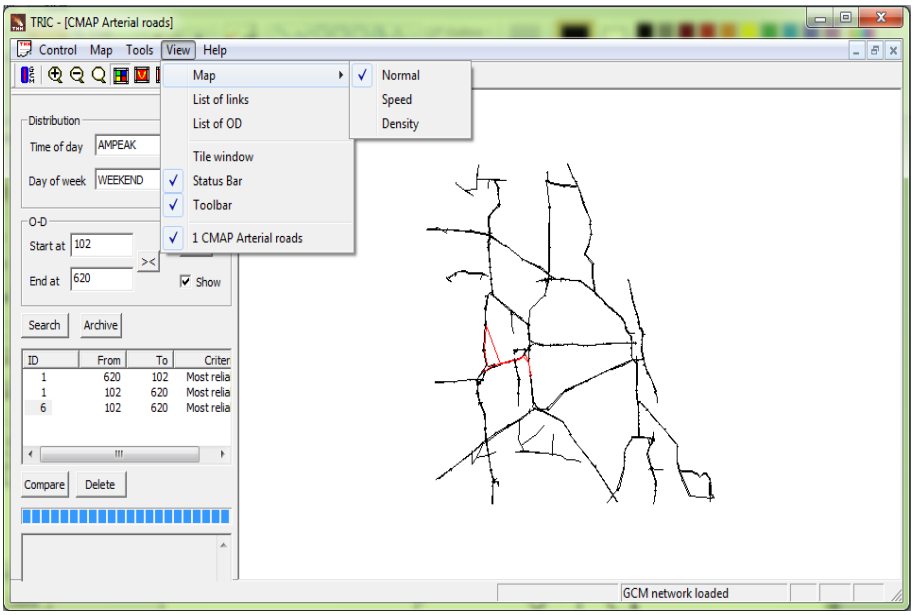


Figure 4.12: View menu on the menu bar

- **Normal:** The normal map.
- **Speed:** This map provides speed information on the map. The color bar on the right hand side links the speed value to color (in which the speed is measured as a percentage of free flow speed), see Figure 4.13.

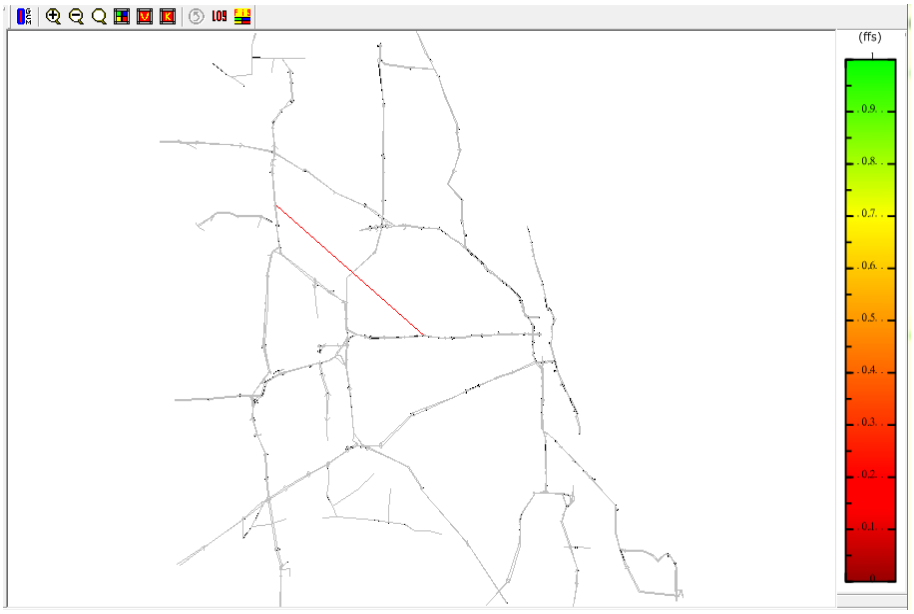


Figure 4.13: Speed map with the colored legend

– **Density:** This map provides a colored scheme for data coverage. As shown in Figure 4.14, Blue indicates that data are collected through loop detector; Orange indicates that data are collected through traffic report; and the green line indicates that data are collected through both sources.

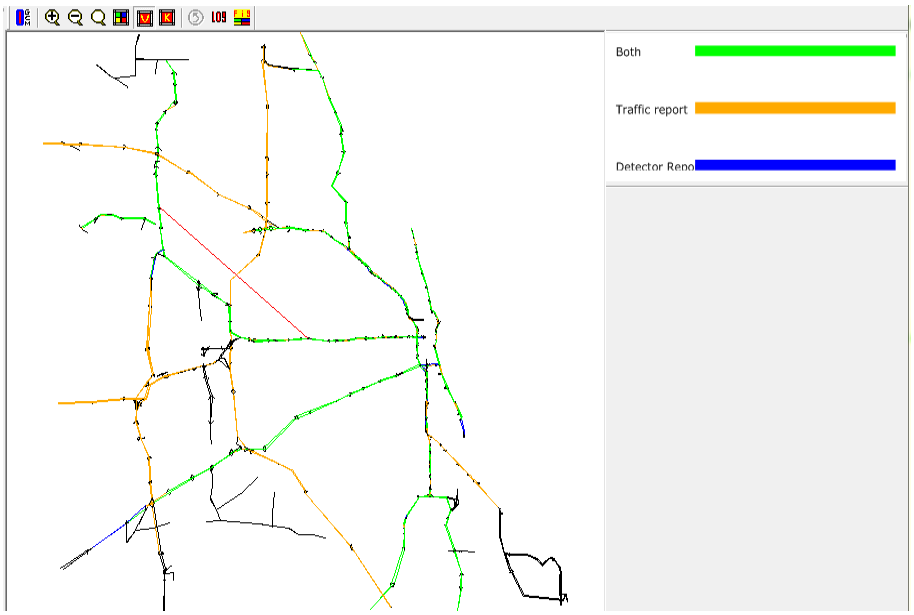


Figure 4.14: Density map

- **List of Links:** This function activates a link data table, as shown in Figure 4.15. The total number of links is 1970 in the network. The link table contains link ID, Type, Start node, End node, Length, Capacity, Free flow speed, RoadName, MODES, Stamp ID and detection type. Link list and map view are linked together inherently. That is, when you click a link on the link view, the map view will automatically zoom into the link and highlight it. This makes it easier to locate a link on a map according to its ID (and other properties) or vice versa. Make sure to tile the window when you need the dynamic linkage between link view and map view.
- **List of OD:** This function opens a text file that stores a list of archived O-D pairs. These are the O-D pairs whose reliable path are generated and archived into the TRI, when the “Archive All O-D” function is called. Unless you know what you are doing, it is not recommended to directly modify this file. To add a new O-D pair, use the input panel to specify an O-D pair and then click ”Save” to add it into the list.
- **Tile window:** This is used to tile windows vertically when map view and link

ID	Type	Start node	End no...	Length (ft)	Capacity	Free flow speed	RoadName	MODES	Stamp ID	Detection
1	GCMLK	1	1392	0.280000	1000.000000	30.000000	I-294	ASHThmlb	18	NONE
2	GCMLK	1	1103	0.260000	1000.000000	30.000000		ASHThmlb	20	NONE
3	GCMLK	2	474	0.120000	1000.000000	30.000000		ASHThmlb	191	NONE
4	GCMLK	3	1056	0.890000	12000.000000	55.000000	I-294	ASHThmlb	204	IPASS
5	GCMLK	4	263	0.250000	4000.000000	55.000000	unknown	ASHThmlb	205	IPASS
6	GCMLK	5	238	0.080000	4000.000000	55.000000	LAKE COOK	ASHThmlb	206	NONE
7	GCMLK	6	1340	0.570000	2000.000000	25.000000	unknown	ASHThmlb	207	NONE
8	GCMLK	7	934	0.150000	1000.000000	30.000000		ASHThmlb	386	NONE
9	GCMLK	8	988	0.450000	6000.000000	50.000000		ASHThmlb	391	IPASS
10	GCMLK	11	448	0.150000	1000.000000	30.000000	IL-83	ASHThmlb	629	NONE
11	GCMLK	11	974	0.360000	4800.000000	55.000000	IL-83	ASHThmlb	630	NONE
12	GCMLK	11	275	0.150000	1000.000000	30.000000		ASHThmlb	632	NONE
13	GCMLK	12	122	0.160000	1000.000000	30.000000	EISENHOWER	ASHThmlb	663	NONE
14	GCMLK	13	1360	0.150000	1000.000000	30.000000	OLD ORCHARD	ASHThmlb	703	NONE
15	GCMLK	14	795	0.250000	1000.000000	30.000000	41A	ASHThmlb	709	NONE
16	GCMLK	15	870	0.150000	6000.000000	55.000000	EAST-WEST	ASHThmlb	753	NONE
17	GCMLK	16	1200	0.750000	5400.000000	55.000000		ASHThmlb	757	NONE
18	GCMLK	16	742	0.160000	1000.000000	30.000000	YORK	ASHThmlb	758	NONE
19	GCMLK	17	1232	1.250000	6000.000000	55.000000		ASHThmlb	800	DET+IPA
20	GCMLK	19	803	0.440000	1000.000000	30.000000	I-55	ASHThmlb	815	NONE
21	GCMLK	19	1079	1.100000	6000.000000	55.000000	I-55	ASHThmlb	816	DET+IPA
22	GCMLK	20	635	0.390000	6000.000000	55.000000	I-190	ASHThmlb	840	IPASS
23	GCMLK	20	773	0.200000	2000.000000	30.000000		ASHThmlb	841	NONE
24	GCMLK	21	1234	0.540000	6000.000000	55.000000	I-190	ASHThmlb	842	IPASS
25	GCMLK	23	105	0.200000	2000.000000	25.000000		ASHThmlb	905	NONE
26	GCMLK	24	231	0.280000	1000.000000	30.000000	271B	ASHThmlb	1019	NONE
27	GCMLK	24	1358	2.020000	6000.000000	55.000000	I-55	ASHThmlb	1020	DET+IPA
28	GCMLK	24	1096	0.430000	1000.000000	30.000000		ASHThmlb	1021	NONE
29	GCMLK	26	753	0.040000	4000.000000	30.000000		ASHThmlb	1155	NONE
30	GCMLK	27	26	0.060000	2000.000000	30.000000		ASHThmlb	1157	NONE
31	GCMLK	28	716	0.060000	2000.000000	30.000000	I-355	ASHThmlb	1160	NONE
32	GCMLK	29	1053	1.460000	6000.000000	55.000000	EAST-WEST	ASHThmlb	1166	IPASS
33	GCMLK	30	31	0.300000	2000.000000	30.000000		ASHThmlb	1177	NONE
34	GCMLK	31	92	0.050000	4000.000000	27.500000		ASHThmlb	1179	NONE
35	GCMLK	32	986	0.310000	2000.000000	30.000000	ARLINGTON HEIG...	ASHThmlb	1186	NONE
36	GCMLK	33	810	0.540000	1000.000000	45.000000	unknown	ASHThmlb	1207	NONE
37	GCMLK	33	390	0.470000	4000.000000	55.000000	ELGIN OHARE	ASHThmlb	1208	IPASS
38	GCMLK	34	668	0.420000	1600.000000	30.000000	I-55	ASHThmlb	1237	NONE
39	GCMLK	34	188	3.860000	6000.000000	60.000000	I-55	ASHThmlb	1238	DETECT

Figure 4.15: Link table

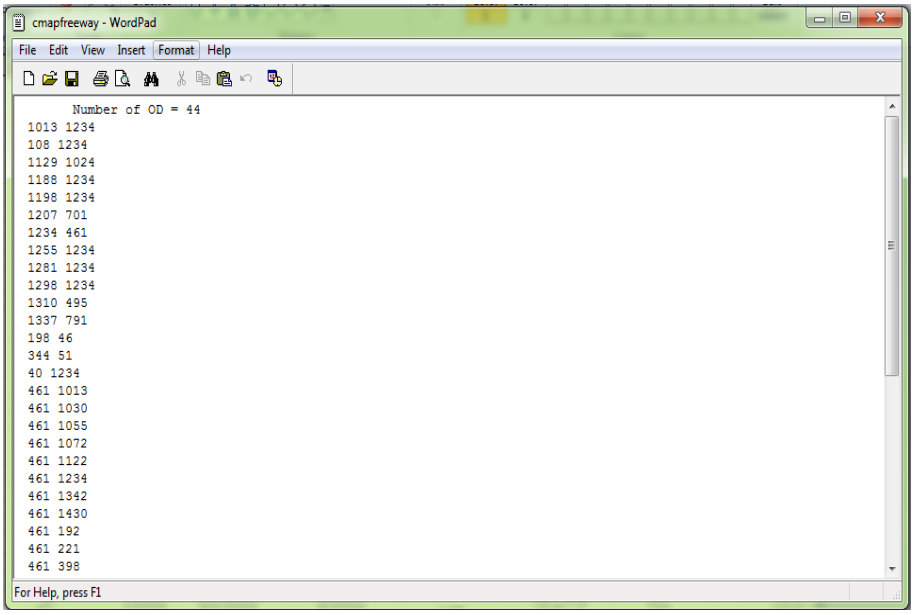


Figure 4.16: Table of O-D pairs

list window are open simultaneously.

- **Status bar/toolbar:** These functions allow the user to display / hide status bar / toolbar.

5. **Help**: The help menu offers help information regarding TRIC.

4.3.3 Tool bar

The tool bar provides shortcuts to commands described in the menu bar section, as shown in Figure 4.17.

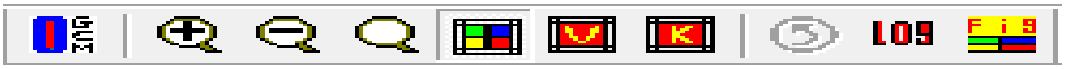


Figure 4.17: Tool bar

 Press this button to open the database window.

 Adjusts the size of the map.

 Displays different kinds of maps.

 Press this button to replay.

 Press this button to open the log window.

 Press this button to open the figure template window.

4.3.4 Map window

Map window displays the map of expressway for the Chicago metropolitan network, as shown in Figure 4.18.

- Double click the map to open ”display properties” window, as shown in Figure 4.19.
- Right click on the map to open a context menu which allows the user to zoom the map or to select origin point (from here) or destination point (to here), as shown in Figure 4.20. Once the origin and destination are specified, a straight line is plotted in the map to connect these two points. This identify these two points as the ”current” O-D pair. If this is deemed unnecessary, the display can be shut off by uncheck the “show” button in the input panel.

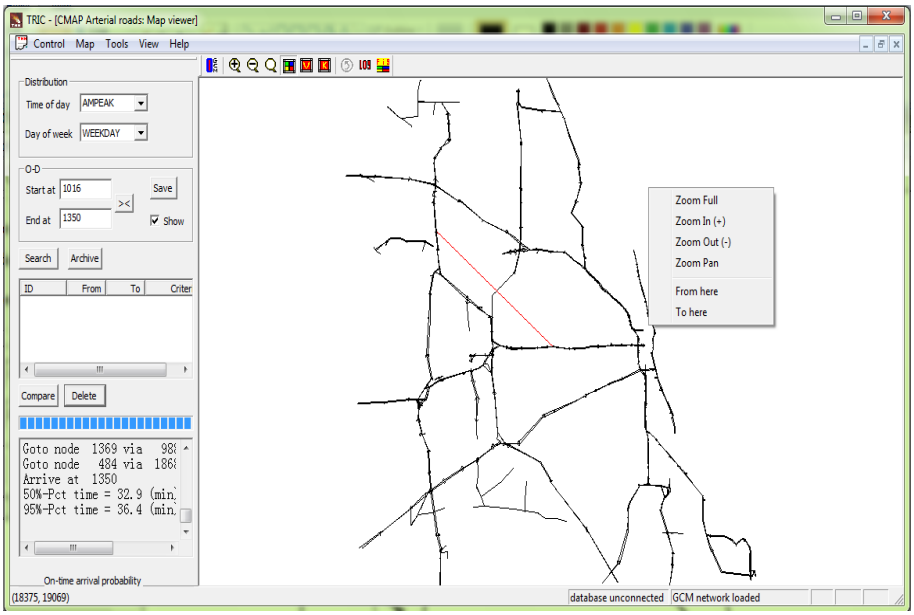


Figure 4.18: Map window

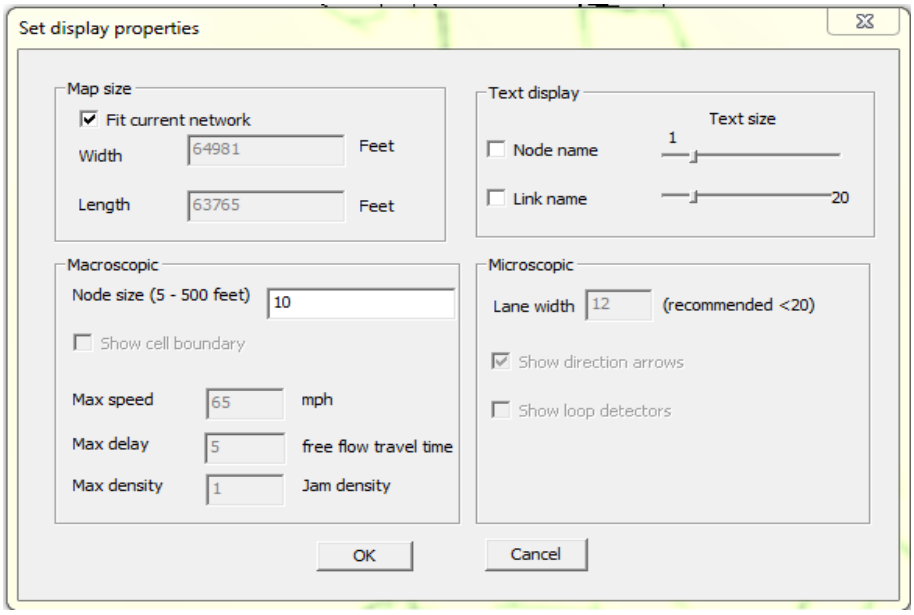


Figure 4.19: Map settings and properties

- Right click on a link to open a context menu that offers two functions: archive and inquiry (see Figure 4.21). Click ”Archive” to save the link (corresponding to current specified distribution). Or click ”Inquiry” to open a travel time distribution window. Travel time distribution window provides information about the CDF and

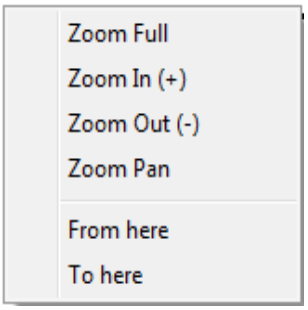


Figure 4.20: Context menu for the map

PDF of the link travel time, as shown in Figure 4.22. To display travel time distribution, specify ”traffic quantity” first by choosing IPASS data or detector data. Then specify parameters related to time (season, day of week and time of day). Users can also change resolution and lower / upper bound of the distribution. Once all parameters are set up, click ”Distribution” button to display travel time distribution (as shown in Figure 4.22).

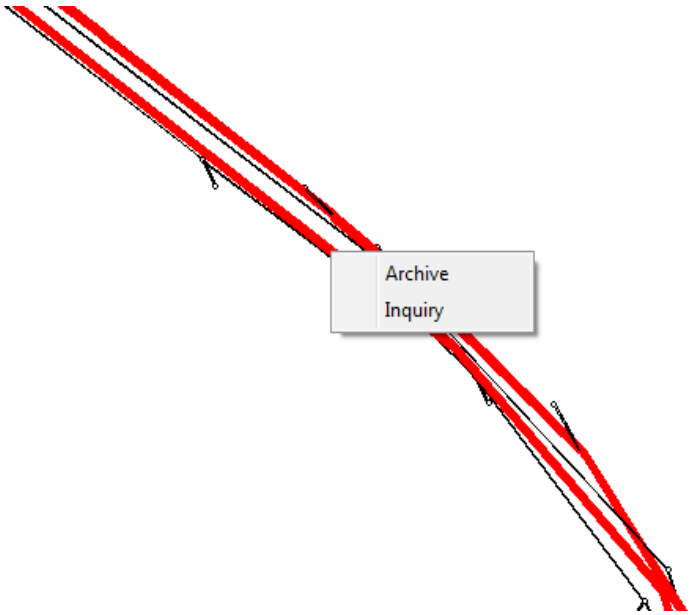


Figure 4.21: Context menu for selected links

4.3.5 Input panel

Input panel provides users options to enter different parameters. Users can choose both distribution of link travel time and O-D pairs, as shown in Figure 4.23(a).

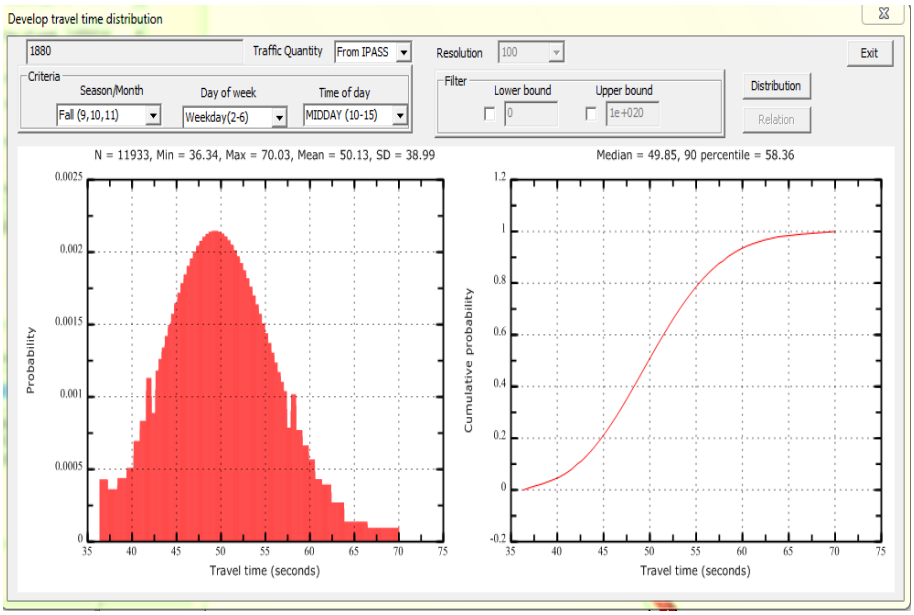
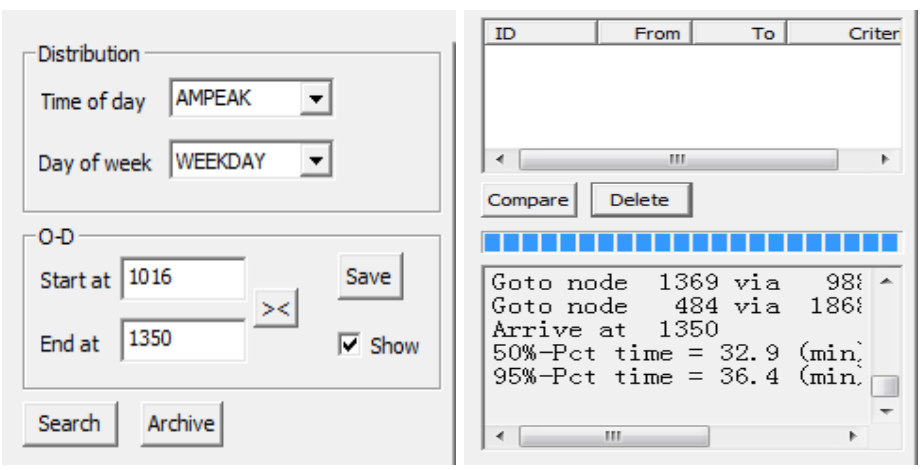


Figure 4.22: Travel time distribution window for a specified path



(a) Input panel (b) Output panel

Figure 4.23: Input and output panel

1. **Distribution panel:** Change link travel time distributions based on the combination of “time of day” and ”day of week”. Whenever the choice for ”time of day” or ”day of week” is changed, the distribution information will be reloaded.

   - ”Time of day” provides 5 options from the drop-down list.
     - AMPEAK: morning peak hour (6-10 am)
     - PMPEAK: afternoon peak hour (4-8pm)

  - MIDDAY: the middle of the day (10am-4pm)
  - OFFPEAK: non-peak hour (other time)
  - NA: all day
- ”Day of week” provides 6 options from the drop-down list.
  - WEEKDAY
  - WEEKEND
  - 0: Sunday
  - 6: Saturday
  - 5: Friday
  - NA: All week

2. **O-D panel:** allows the users to choose different O-D pairs. Users can either enter the O-D pair on this panel or right click the map to set up O-D pairs (select ”From here” and ”To here” respectively). Users can click the ”show” box to display/hide the red line which indicates the origin and destination of a route.

   to switch origin and destination nodes.

   to save O-D pairs to list of O-D pairs. This list will be used when an O-D archive is performed.
   Once distribution and O-D pairs are set up, click ”Search” or ”Archive” to find or save the reliable path between the O-D pair for the specified distribution. All reliable paths for the selected O-D pair will be listed. The search result is reported in the output panel.

4.3.6 Output panel

Output panel provides information about reliable paths based on user specified input, as shown in Figure 4.23(b).

- **Display the CDF and PDF of a reliable path:** double click the path.

  At the top of PDF plot, minimum, maximum, mean and standard deviation are calculated (see Figure 4.24).
  At the top of CDF plot, 50% percentile, 95% percentile, B-Index (buffering index, the ratio of 95 percentile travel time to free flow travel time) and P-Index (planning index, the ratio of 95 percentile travel time to mean travel time) are calculated, as shown in Figure 4.24.

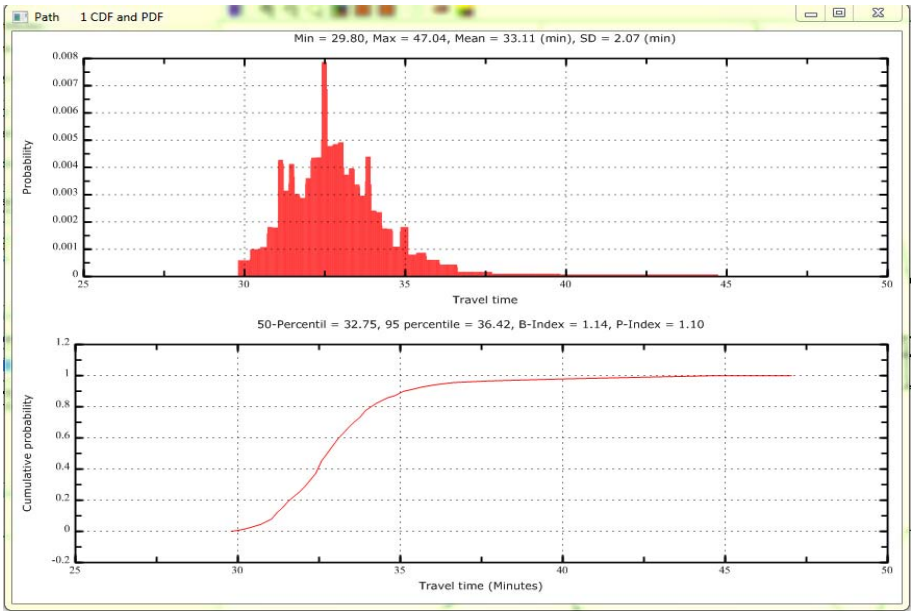


Figure 4.24: Display PDF and CDF of reliable path

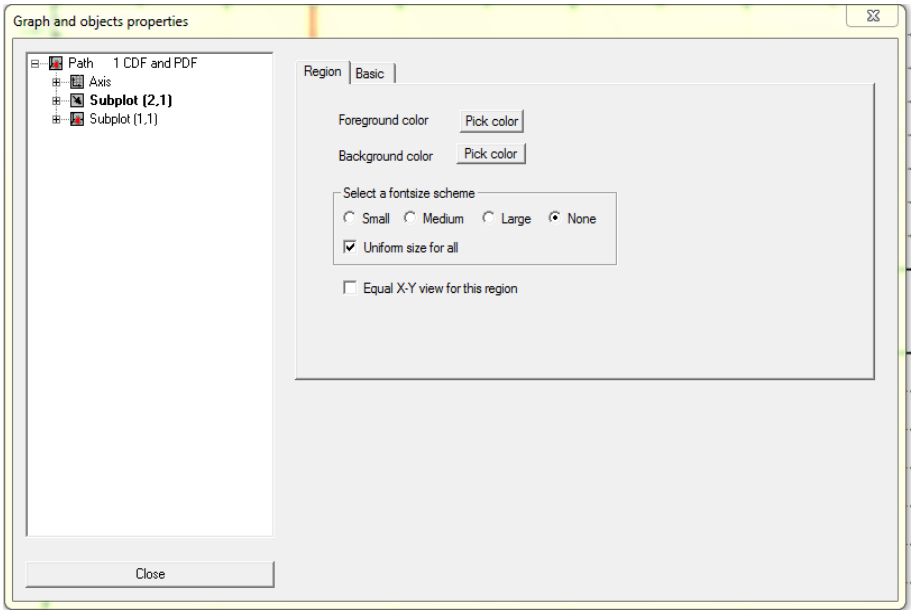


Figure 4.25: Graph properties

- **Adjust the properties of the plots:** double click the plot.

  Compare to compare multiple paths. Select multiple paths with ”Shift” or ”Ctrl”, the click ”compare” to display the comparison graph of multiple paths, as shown

in Figure 4.26.

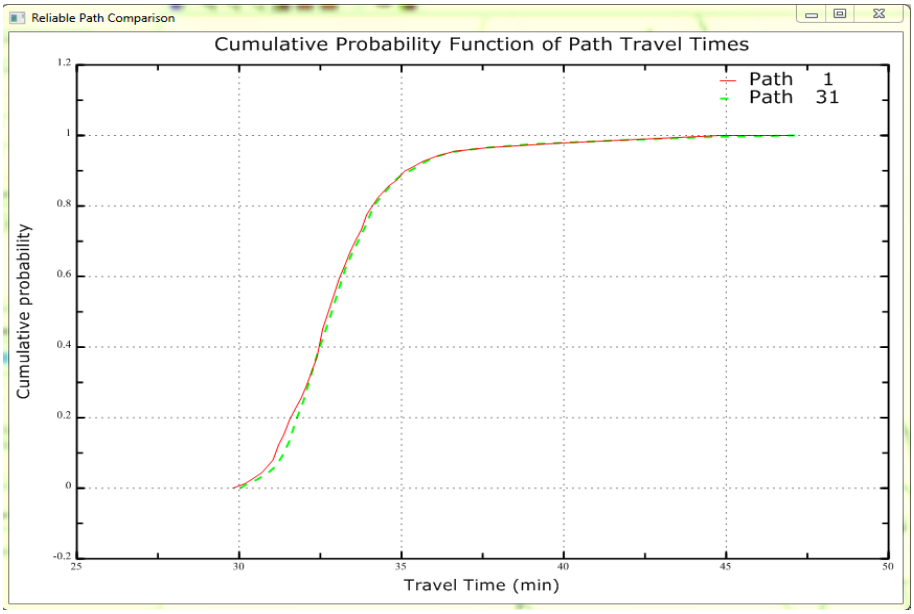
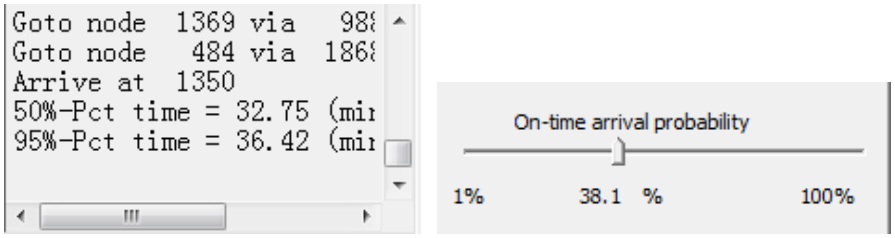


Figure 4.26: Multiple paths comparison

Delete to delete paths from the list. Select paths from the list and click ”Delete” to remove the paths from the list.



(a) Log panel (b) Reliability panel

Figure 4.27: Log panel and reliability panel

- **Log panel**: Log panel is below the path list, as shown in Figure 5.6(a). It provides summary of the path. It directs the path choice at each node and summarized the 50% percentile travel time and 95% percentile travel time.

4.3.7 Reliability panel

Reliability panel adjusts user’s perception of on-time arrival probability, as shown in Figure 4.27(b). Move the scroll bar to adjust on-time arrival probability. Travel time

corresponding to that on-time arrival probability is displayed in the log panel after adjustment. When there are more than one path for a given O-D pair (and if they are all selected), the minimum percentile travel time will be displayed, and the path that gives the optimal value will be highlighted.

Chapter 5

Case Study

In this chapter, we conduct two case studies using TRIC on the Chicago expressway network. Section 5.1 performs a travel reliability analysis for a selected O-D pair using TRIC. Section 2 shows how to make use of travel reliability inventory generated by TRIC with the help of a freely available GIS tool.

5.1 Travel reliability analysis

In this study, *Node 908* is selected as the origin, and *Node 132* is selected as the destination, see Figure 5.1. *Node 902* is close to Clinton Blue Line Station of Chicago Transit Authority, whereas *Node 132* is close to Montrose Avenue where Kennedy Expressway merges with the southern end of Edens Expressway. We choose to analyze this particular O-D pair because it connects Downtown Chicago and O'Hare International airport and it has two major alternatives (I-90 and I-90 Expressway). For the demonstration purpose, we only consider three distributions, namely: (a) **WEEKDAYS AMPEAK**, (b) **WEEKDAY MIDDAY** and (c) **WEEKEND AMPEAK**. We note that traffic condition are quite different in these time periods.

5.1.1 Case 1: WEEKDAYS-AMPEAK

We first set up the distribution for the morning rush hour of weekdays by selecting the two parameters under "Distribution" as "AM PEAK" and "WeekDay", respectively. Then we can either choose O-D pair from the map by right clicking the nodes on the map or entering the node IDs directly into TRIC. See Figure 5.2(a) for an illustration. Once all input parameters are specified, click "Search" to find the reliable paths (Figure 5.2(a)). In our case, two reliable paths, 1 and 31, are found and added into a list view, as shown in Figure 5.2(b). A close examination reveals that path 1 predominately uses I-90 Expressway and path 31 predominately use I-90.

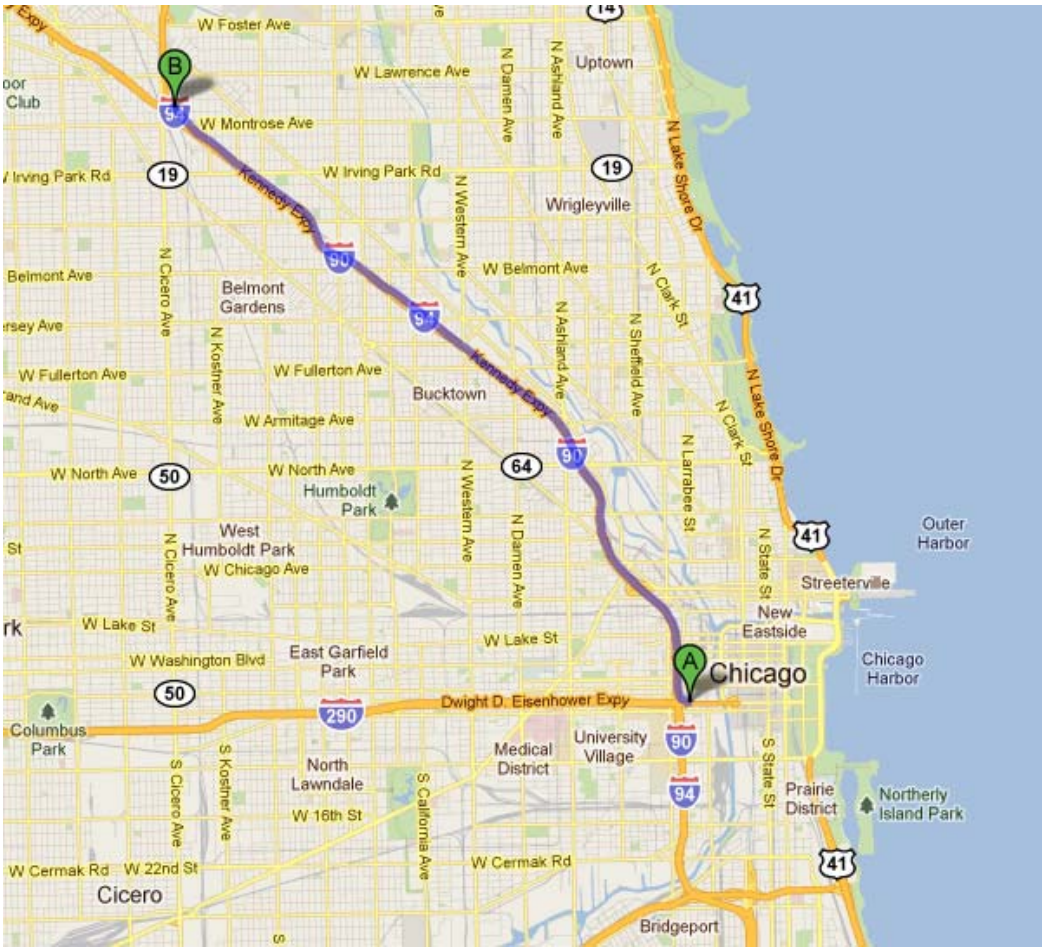
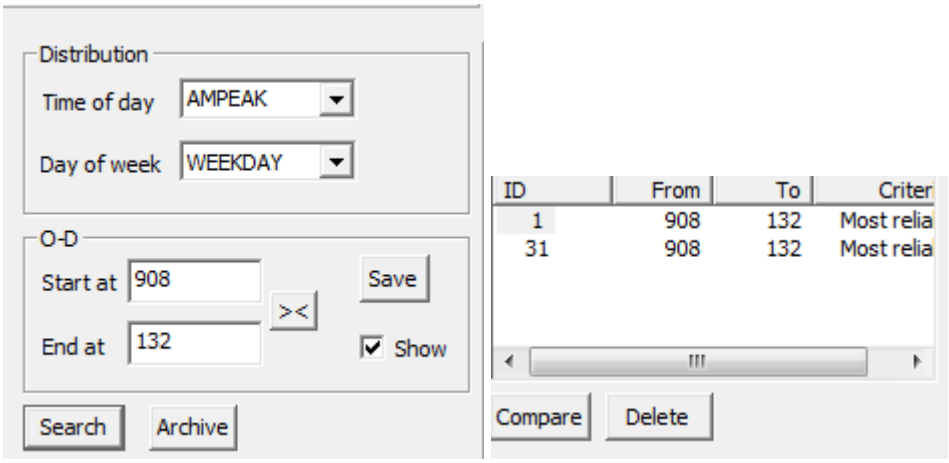


Figure 5.1: Snapshot between Node 908 (A) and Node 132 (B)

Display detailed path information: Click on either path to display the detailed path information in log panel (see Figure 5.2(b)). At the end of the window display the 50 and 95-percentile travel times of the selected path. In this case, the 50% percentile time = 12.92 (min). This means if one wants to guarantee a 50% probability of arriving on time, 12.92 minutes should be reserved for the trip. If a higher (95%) on-time arrival probability is desired, 18.17 minutes should be reserved for the trip. The result is consistent with the intuition that a higher arrival probability requires a larger time budget.

Display CDF and PDF: To display the cumulative density function (CDF) and the probability density function (PDF) of the travel time on the selected path, simply double click on the path in the list view. The CDF and PDF of Path 1 is shown in Figure 5.4. The PDF plot reports that the minimum travel time on path 1 is 8.05 minutes, the maximum travel time is 23.41 minutes, the mean travel time is 13.26 minutes, and standard deviation is 2.67 minutes. The CDF plot reports that 50 percentile time and 95



(a) Input parameters in Case 1 (b) Reliable shortest paths in Case 1

Figure 5.2: Input parameters and the identified reliable shortest paths in Case 1

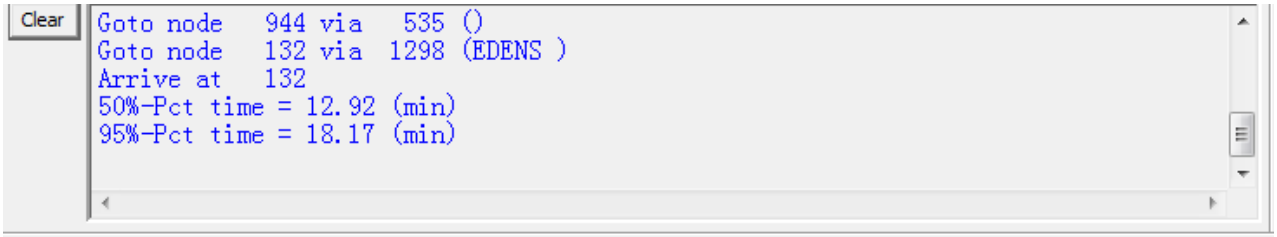


Figure 5.3: Detailed path information as shown in the log panel

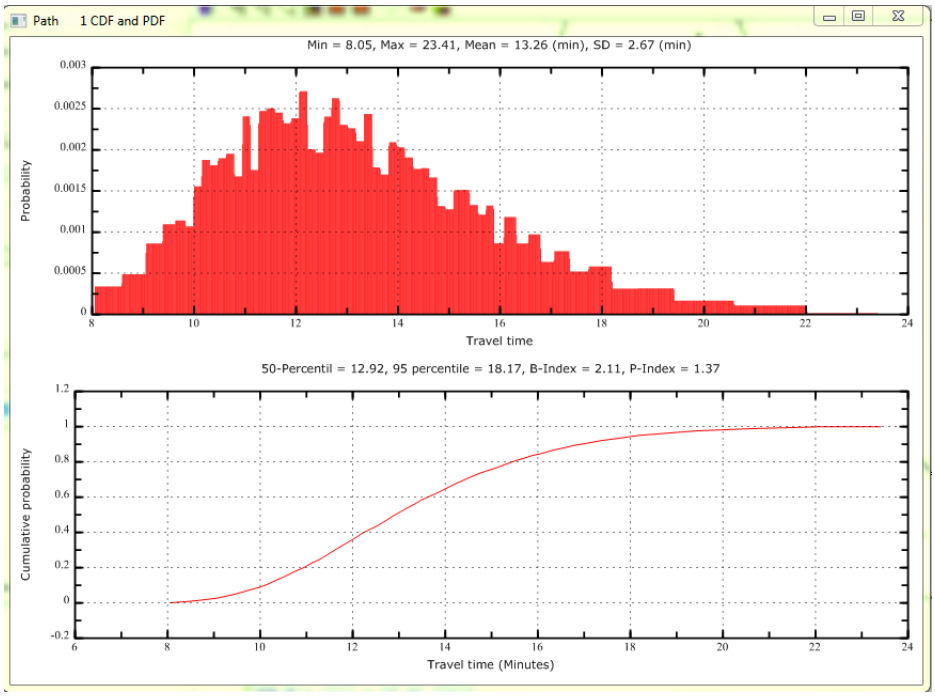


Figure 5.4: PDF and CDF of the first path in Case 1

percentile time are 12.92 minutes and 18.17 minutes respectively. Moreover, B-index and P-index are 2.11 and 1.37, respectively.

Compare different paths: We can compare the CDF of different paths shown in the list view. To do this, simply select multiple paths from the list view (hold ctrl key) and click ”compare” button. In our example, two paths are selected, Path 1 (red line) and Path 31 (green line), as shown in Figure 5.5. The two CDFs intersect with each other around point (12.5, 0.42). This indicates that if on-time arrival probability is less than 0.42, Path 1 (or I-90 Expressway) is better than Path 31 in that using path 1 requires less travel time budget for the same on-time arrival probability. However, Path 31 is a better alternative when $\alpha > 0.42$. Because we are interested in paths that are best at high reliability (i.e. $\alpha = 0.95$) in this study, Path 31 (or I-90) will be selected to represent this O-D pair in the travel reliability inventory.

Adjust on-time arrival probability: we can use the reliable panel to adjust on-time arrival probability. Once on-time arrival probability is adjusted, travel time budget will be re-calculated and shown in the log panel. In this example, we calculate different on-time arrival probability and summarize the result below in Table 5.1.

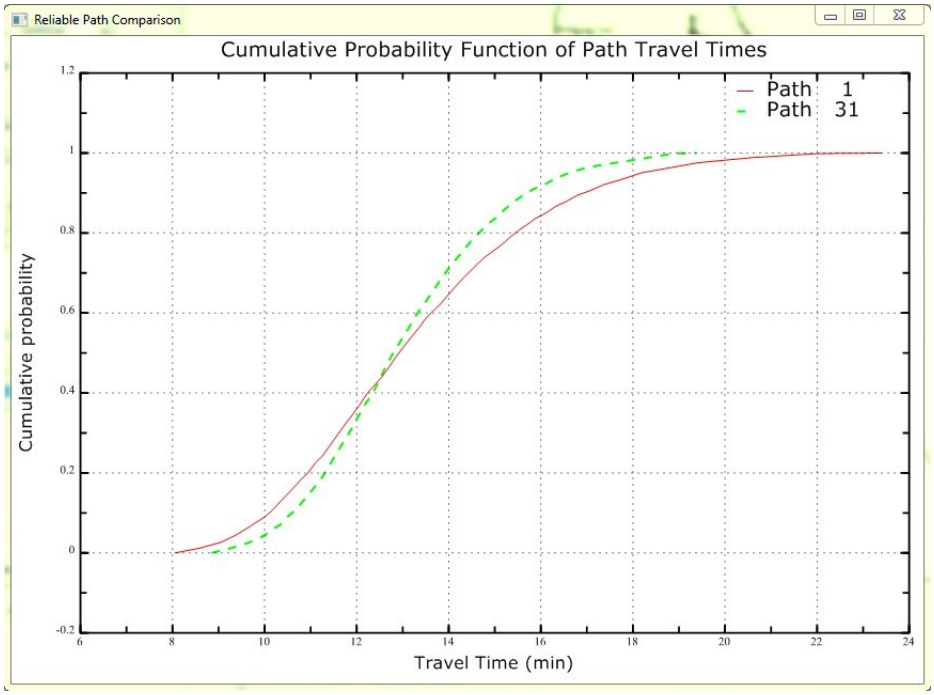
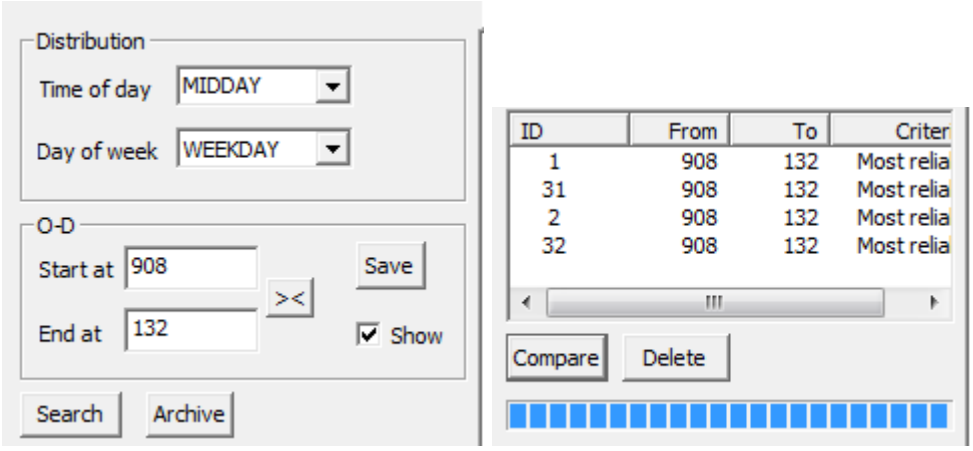


Figure 5.5: Path comparison in Case 1

Table 5.1: Travel time budgets for different on-time arrival probabilities

On-time arrival probability	Travel time (min)
38.1%	12.1
50%	12.9
78.1%	15.2
96 %	18.2



(a) Input parameters in Case 2 (b) Reliable shortest paths in Case 2

Figure 5.6: Input parameters and the identified reliable shortest paths in Case 2

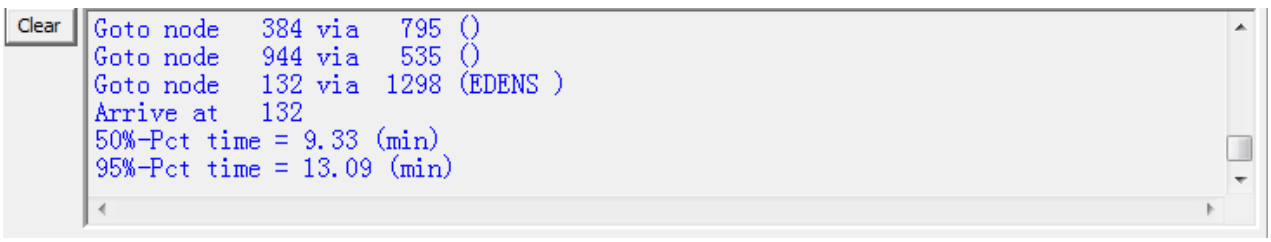


Figure 5.7: Log panel in case 2

5.1.2 Case 2: WEEKDAYS-MIDDAY

The inputs for this case is identical the Case 1 except the time of day parameter is changed from AMPEAK to MIDDAY, as shown in Figure 5.6(a). After clicking ”Search”, two reliable paths are found for this case, Path 2 and Path 32. A close examination reveals that paths 2 and 32 predominately uses I-90 Expressway and I-90, respectively. Note that these paths are added to the list view, following the paths from the first case, as shown in Figure 5.6(b). This makes it easy to compare paths obtained from different scenarios.

Display detailed information of the path: This can obtained by clicking on the path in the list view. As shown in Figure 5.7, 50 and 95 percentile travel times on Path 2 are 9.33 and 13.09 minutes, respectively.

Display CDF and PDF: The CDF and PDF of Path 2 are shown in Figure 5.8. Note that the minimum travel time is 7.23 minutes, the maximum travel time is 17.04 minutes, the mean travel time is 9.69 minutes, and standard deviation is 1.75 minutes. Also, the B-index is 1.52 and P-index is 1.35.

Compare different paths: We can similarly compare different paths generated for the

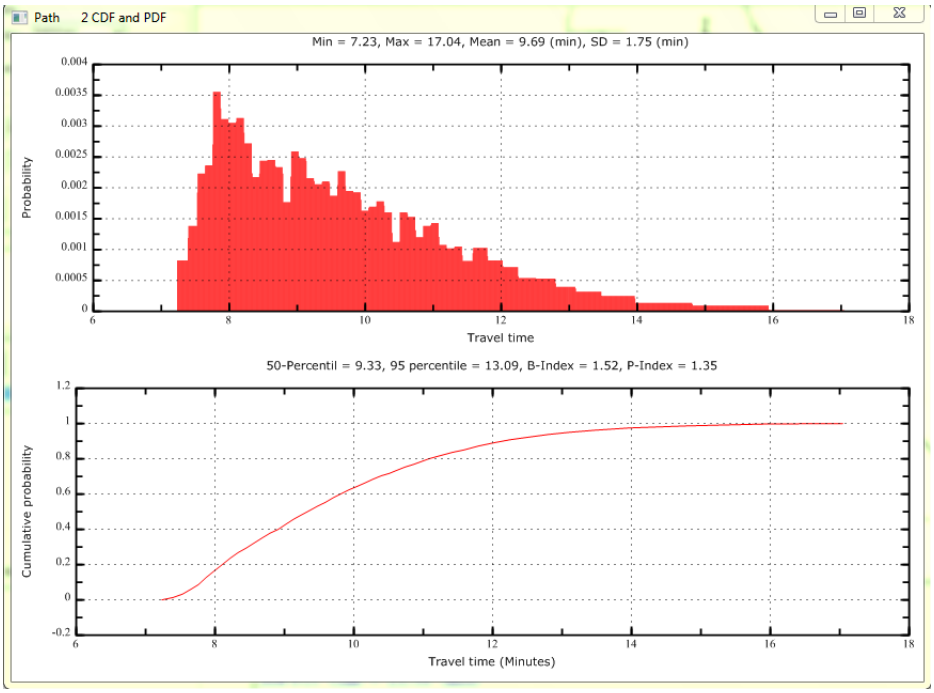


Figure 5.8: CDF and PDF in case 2

Table 5.2: Comparing multiple paths

Path	Path 1(AMPEAK)	Path 31(AMPEAK)	Path 2(MIDDAY)	Path 32(MIDDAY)
Min	8.05	8.85	7.23	7.55
Max	23.41	19.38	17.04	14.55
Mean	13.26	13.02	9.69	9.76
50 % percentile	12.92	12.79	9.33	9.52
95 % percentile	18.17	16.61	13.09	12.37

same time period, e.g. Path 2 and Path 32. The comparison yields a similar conclusion as in the case 1, that is, Path 2 (I-90 Expressway) performs better for lower reliability and Path 32 (I-90) is the winner when high reliability is desired. Moreover, we can compare the paths from different time period, e.g. Path 1 and Path 2. Figure 5.1.2 performs such a comparison. As shown in the plot, the CDF of Path 2 (green line) lies completely above that of Path 1 (red line), which indicates that for all on-time probability, path 2 requires less travel time budget. This observation is expected because the road is typically more congested in the morning rush hour.

Comparing multiple paths: So far, we find four reliable paths for two different time periods. These paths are compared in Table 5.2. We can see that the performance of Path 2 and Path 32 are consistently better than their counterparts. As mentioned before, this

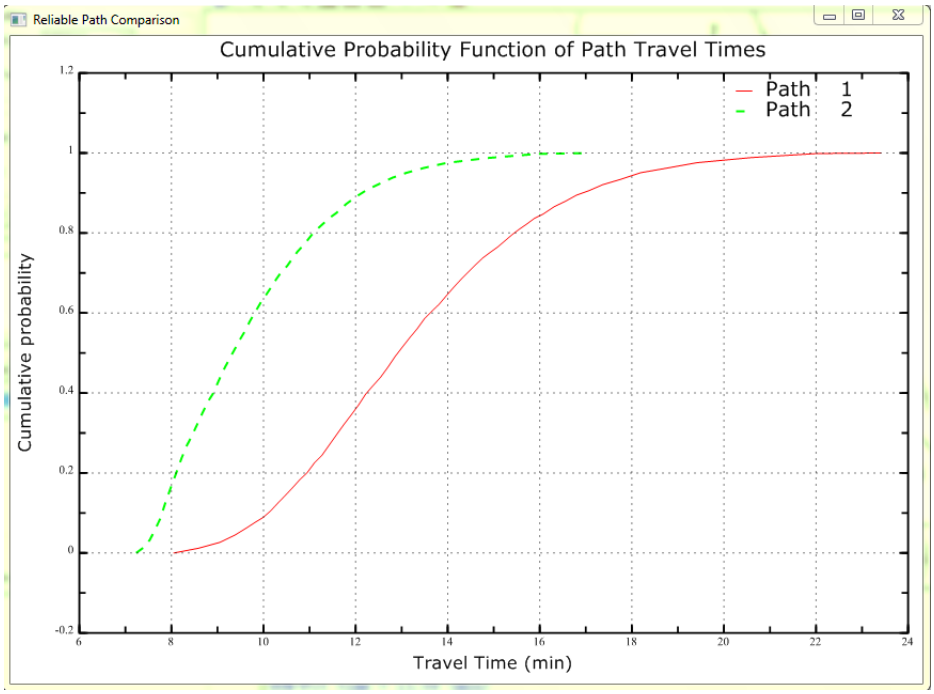


Figure 5.9: Path comparison in case 2

is because the traffic congestion is worse in the morning rush hour. More importantly, the choice between I-90 Expressway (paths 1 and 2) and I-90 (Path 31 and 32) depends on on-time arrival probability. For low on-time arrival probability,I-90 Expressway is preferred, and for high on-time arrival probability, I-90 is preferred.

5.1.3 Case 3: WEEKEND-AMPEAK

For the morning period of weekends, only one reliable path (path 6) is found, which predominately uses I-90. On this path, the 50% percentile travel time is 6.63 minutes and 95% percentile travel time is 7.26 minutes, as shown in Figure 5.10. These values are far lower than the previous two cases, likely because the early morning period of weekends is not subject to congestion at all.

CDF and PDF: From the CDF and PDF plots, we see that the minimum travel time on Path 6 is 6.29 minutes, the maximum travel time is 8.35 minutes, the mean travel time is 6.70 minutes, and standard deviation is 0.29 minutes. The B-index is 0.91 and P-index is 1.08. It is interesting to note that the B-Index is smaller than 1. This indicates that the average travel time during this period is smaller than the free flow travel time calculated from the speed limit. Thus, an average drive will drive at a speed higher than the speed limit if possible.

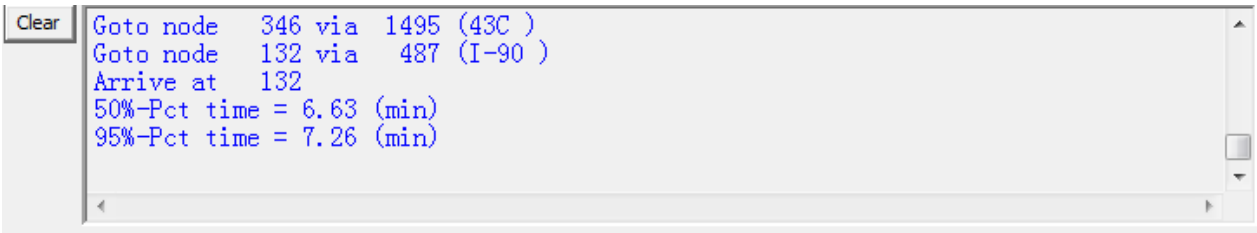


Figure 5.10: Path choice in case 3

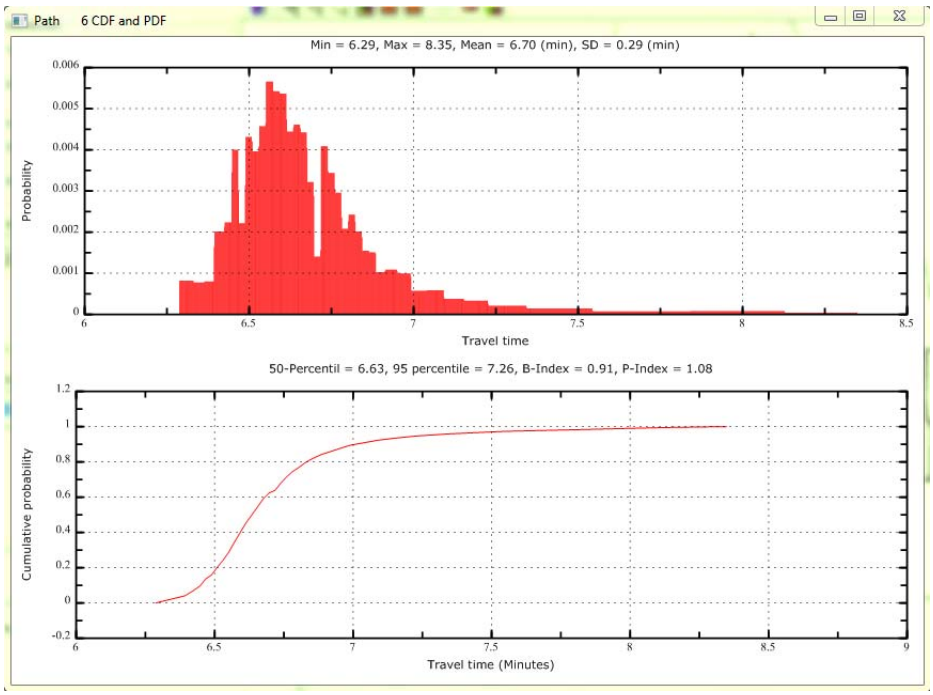


Figure 5.11: CDF and PDF of Path 6

Compare different paths: Finally, we compare the CDF of Path 6 (WEEKEND, AMPEAK) with that of Path 1 (WEEDDAYS, AMPEAK) in Figure 5.12. It is clear from the plot that travel time budget for any reliability is far less on Path 6 than on Path 1.

5.2 Integration TRI with GIS

We now show how to visualize travel reliability inventory using a GIS tool. In this study, a freely available GIS tool named **Quantum GIS** is used. However, other GIS software packages will work equally well. By default, the input and the trave reliability inventory (TRI) files are stored at `installation folder\Data` and `installation folder\TRI`, respectively. Included in the input folder are two shape file sets named

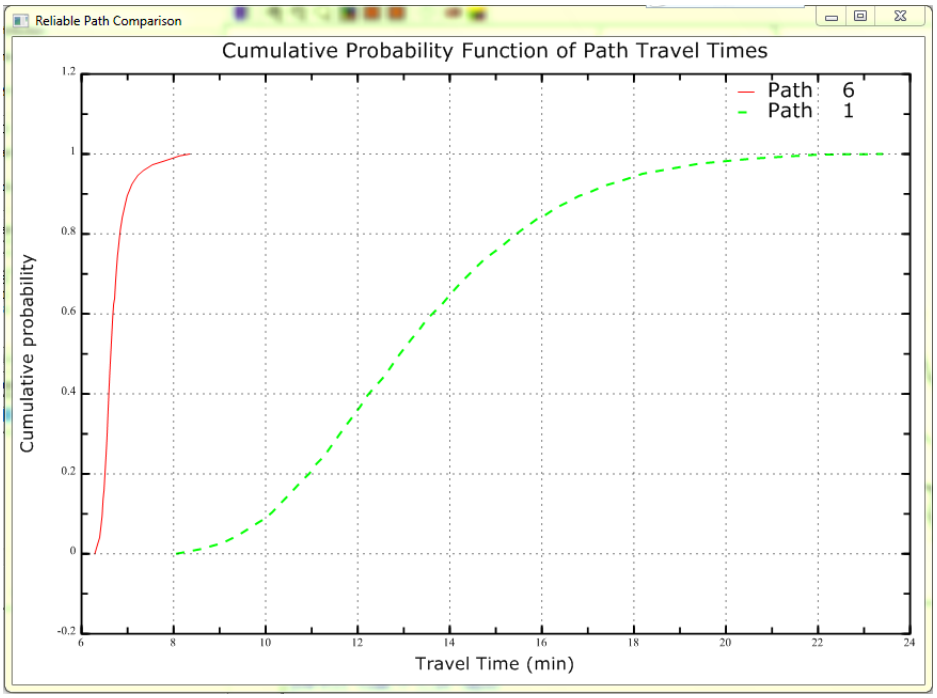


Figure 5.12: Path comparison in case 3

`CMAP_ntwk` and `CMAP_nodes`. These shape files, provided by CMAP, contain the GIS information of the Chicago Metropolitan network, and can be used to overlay the TRI data if needed. Because all TRI files are in shape file format, they can be easily imported into Quantam GIS as vector layers. We do not provide a detailed instruction on how to use Quantum GIS because (1) the end users may opt to a different GIS tool, and (2) it is beyond the scope of our project. The reader is referred to Quantum GIS Web Site for more details.

We first examine the link-based TRI for the three cases considered in the previous section, namely, AMPEAK-WEEKDAY, MIDDAY-WEEKDAY and AMPEAK-WEEKENDS. By default, the link-based and path-based TRI corresponding to a time period (`thisPeriod`) and a day of week (`thisDay`) is stored in shape file sets named `tri_link_thisPeriod_thisDay` and `tri_path_thisPeriod_thisDay`, respectively. For instance, the link-based TRI for the morning peak of weekdays is stored in `tri_link_AMPEAK_WEEKDAY` (with four suffixes: .prj, .shp., .shx and .dbf).

We import each shape file set into Quantum GIS and then color the links based on the calculated B-index and P-index. Recall that B-index is the ratio of 95 percentile travel time to free flow travel time, while P-index is the ratio of 95 percent travel time to the mean travel time. Both indices measures the travel reliability associated with a facility (a path or a link). Specifically, a higher B- and P-index means the travelers who use the facility

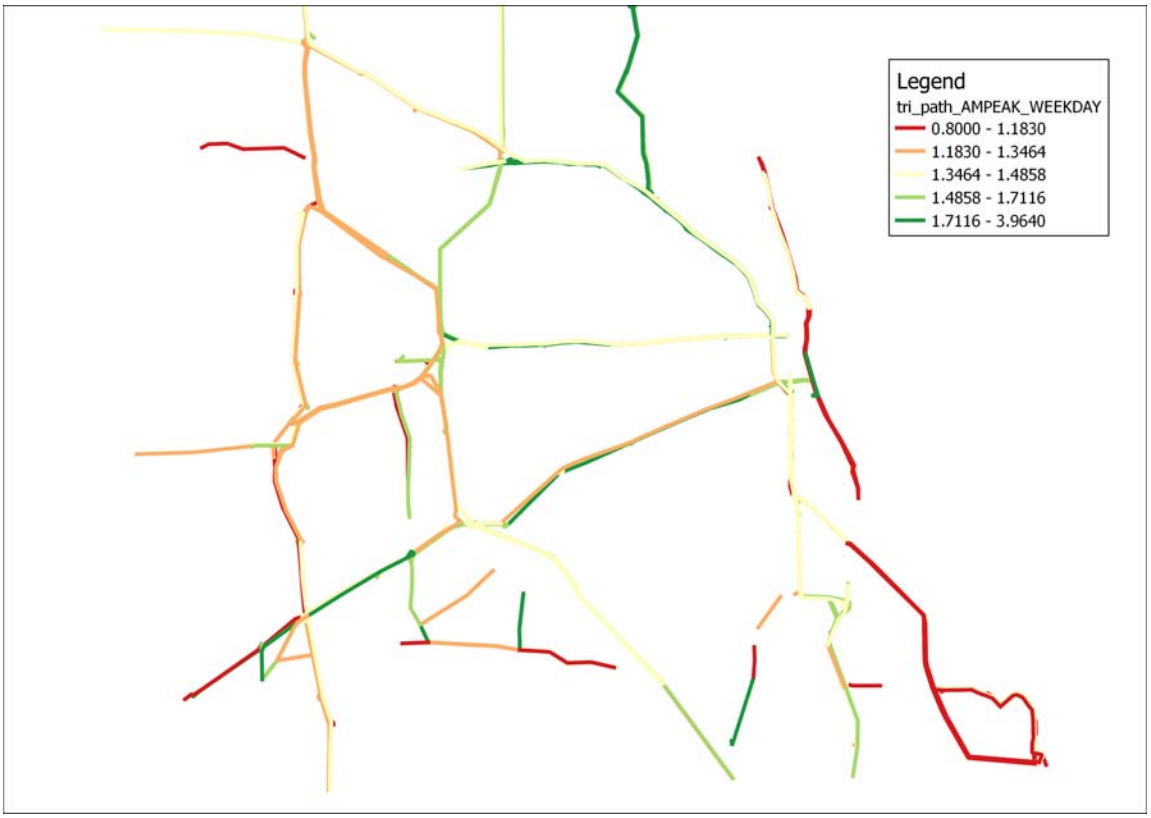


Figure 5.13: Link B-Index for AMPEAK-Weekday

will have to reserver more time for travel in order to achieve high reliability. The link B-index plots for the three time periods are shown in Figures 5.13 - 5.15. First, note that these indexes are generally the highest during the morning peak period of weekdays. Of all the major expressway, the eastbound of Edens and Kennedy seems to have the highest B-index during the morning period, whereas the Stevenson and Dan-Ryan have the highest B-index in the middle of day. For the morning of weekends, most roads have relatively low B-index (near or lower than 1), as expected. The most unreliable portion in this case are the Dan-Ryan and the segments that connect Eisenhower and Lake Shore Drive.

Figures 5.16 - 5.18 report the link P-index plots for the three time periods. Interestingly, Edens and Kennedy, which have the highest B-index in the weekday mornings, have relatively low P-index for the same period. This indicates that these roads are generally more congested but the conditions are relatively stable. The south bound of Lake shore Drive North, on the other hand, have the highest P-index in the weekday mornings, indicating high volatility in traffic conditions. For the mid-of-day, Eisenhower and Stevenson have the highest P-index of all major expressways. Stevenson also has the

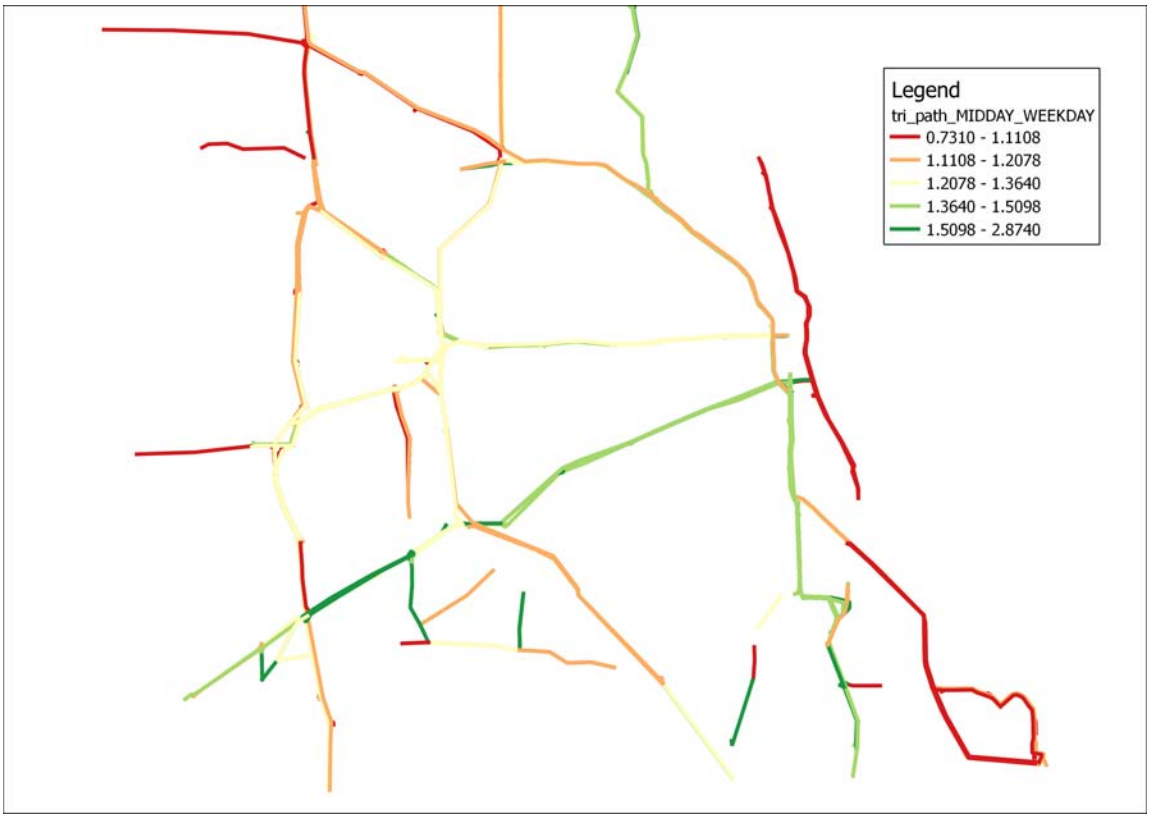


Figure 5.14: Link B-Index for MDIDAY-Weekday

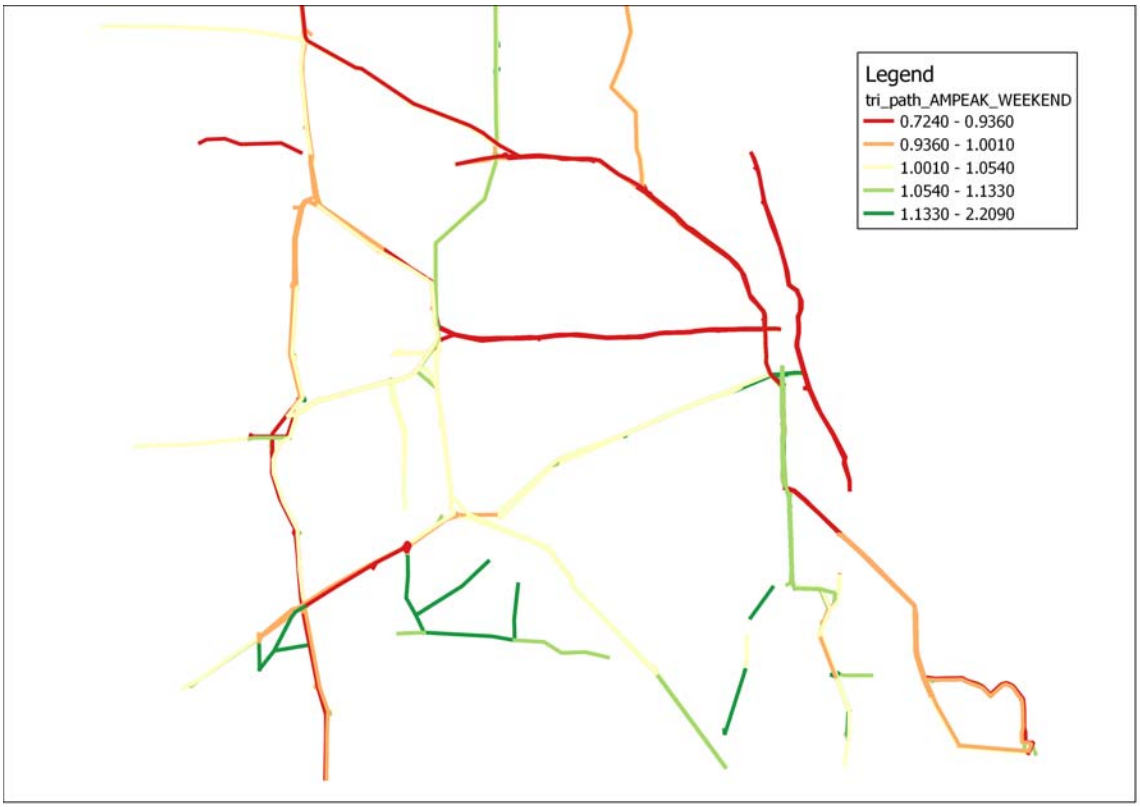


Figure 5.15: Link B-Index for AMPEAK-Weekend

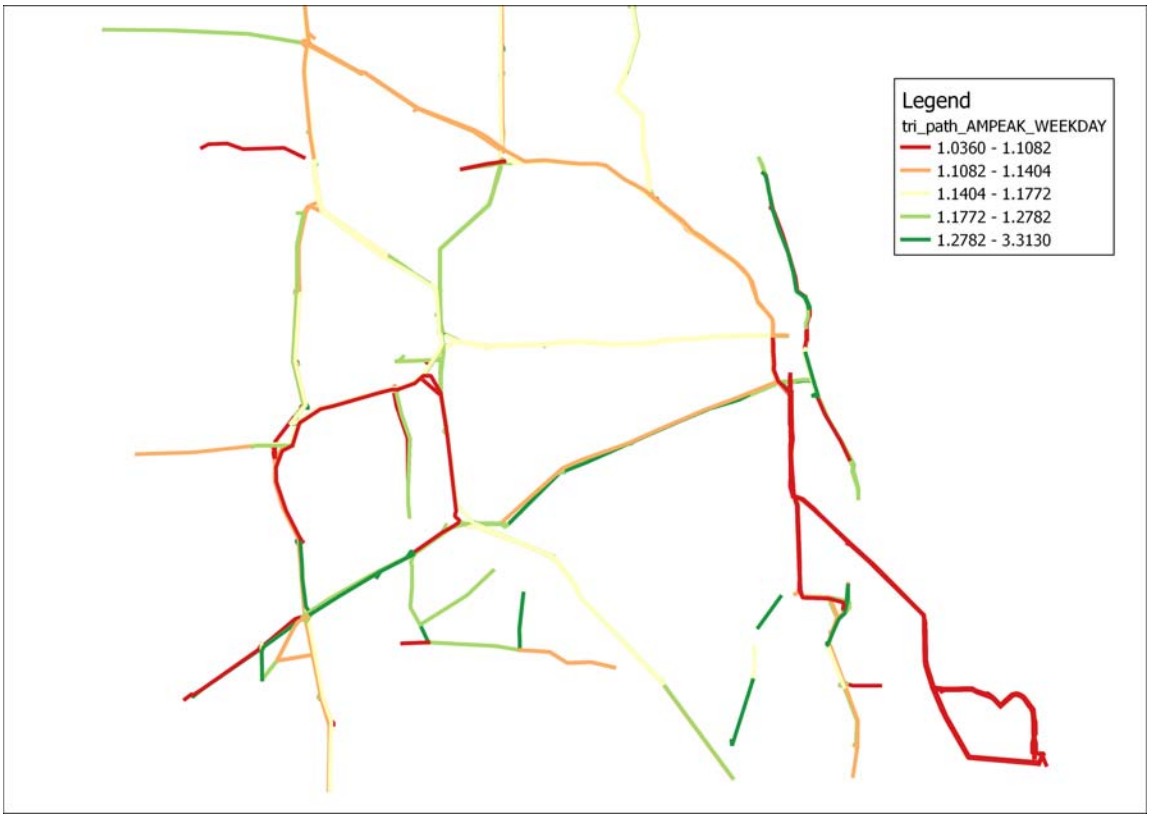


Figure 5.16: Link P-Index for AMPEAK-Weekday

highest P-index for the weekend mornings.

We proceed to examine the TRI of six paths, which correspond to six segments of main expressways in the Chicago area: namely, Kennedy westbound, Eisenhower westbound, Stevenson westbound, Dan-Ryan (southbound), Lake Shore Dr. North (northbound), and Lake shore Dr. South (southbound).

Inspecting these plot we have the following findings: (1) for weekday morning period, Kennedy westbound has the highest B-index among all major expressways that lead away from Downtown Chicago, followed by Eisenhower westbound, and Stevenson westbound; (2) for weekday mid-of-day period, Stevenson has the highest B-index, followed by Eisenhower; (3) for weekday morning period and mid-of-day period, Eisenhower westbound and Stevenson westbound have the highest P-index. In general, these observations are consistent with those made from link-based plots.

It is clear that the above analysis can be extended to other routes or O-D pairs. The focus of this report is to show the methodology. We leave comprehensive numerical analyses to TRIC'S end users.

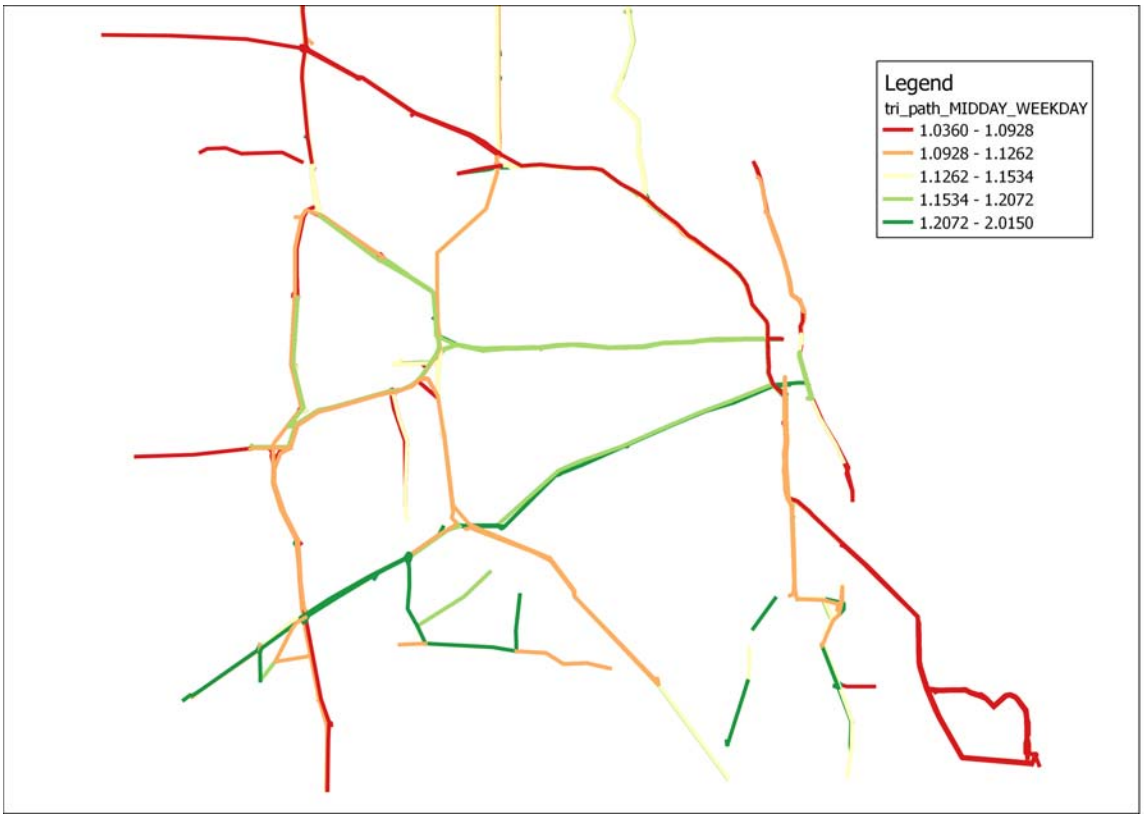


Figure 5.17: Link P-Index for MIDDAY-Weekday

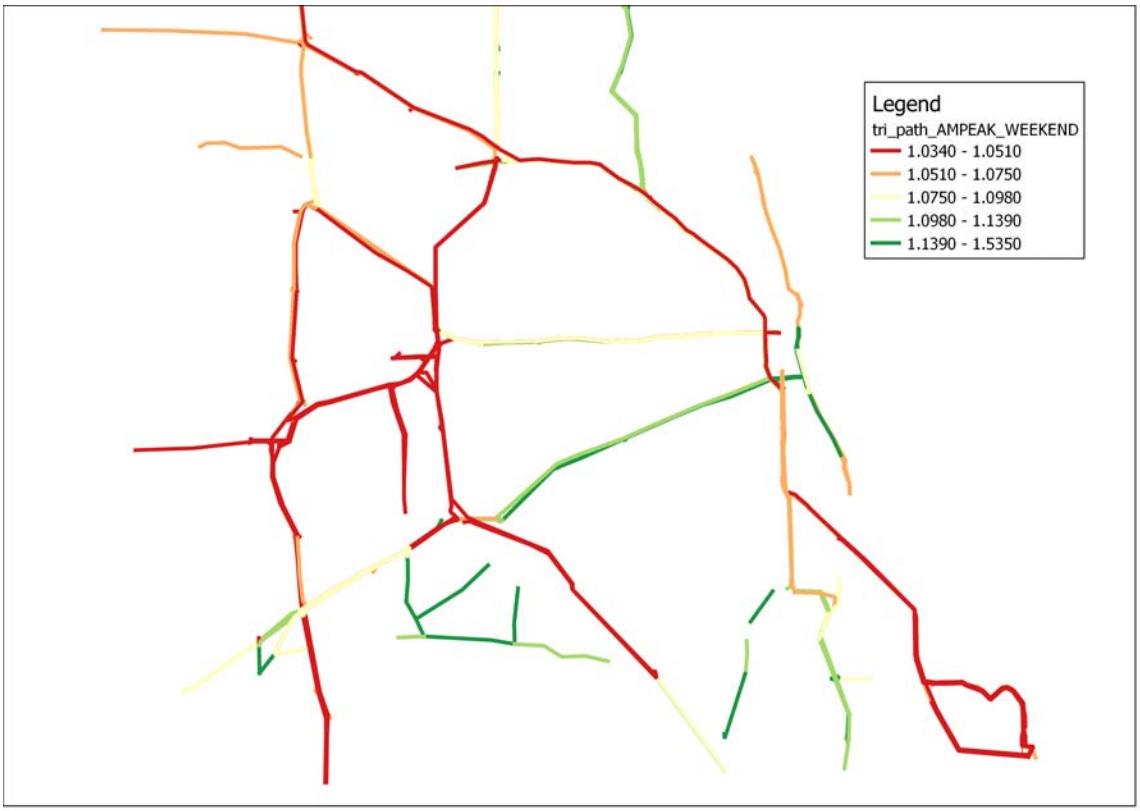


Figure 5.18: Link P-Index for AMPEAK-Weekend

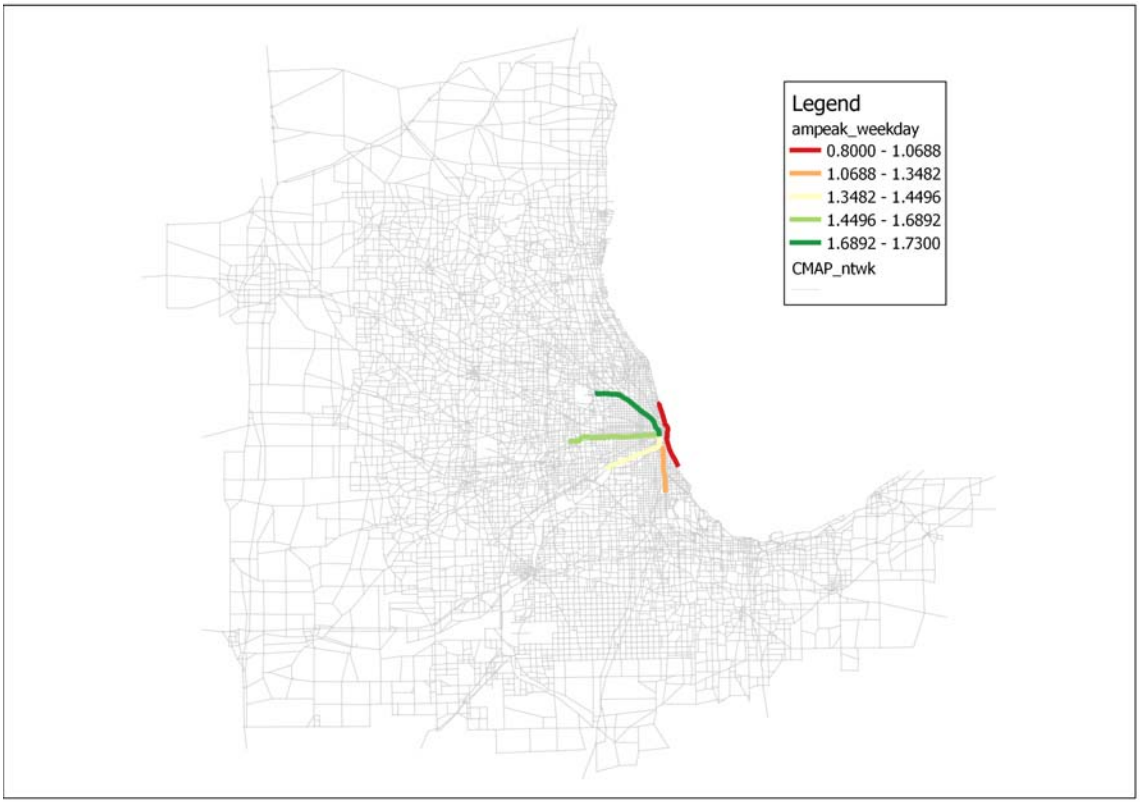


Figure 5.19: B-Index for AMPEAK-Weekday on six major routes

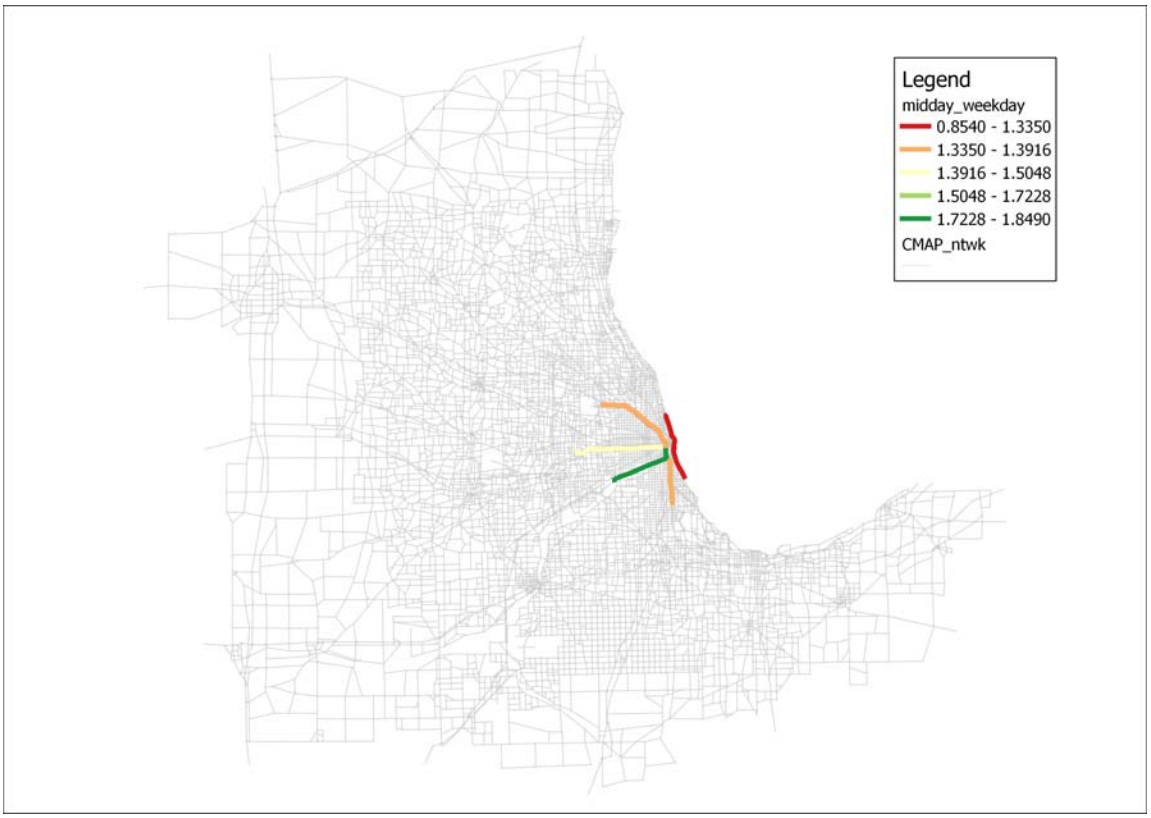


Figure 5.20: B-Index for MIDDAY-Weekday on six major routes

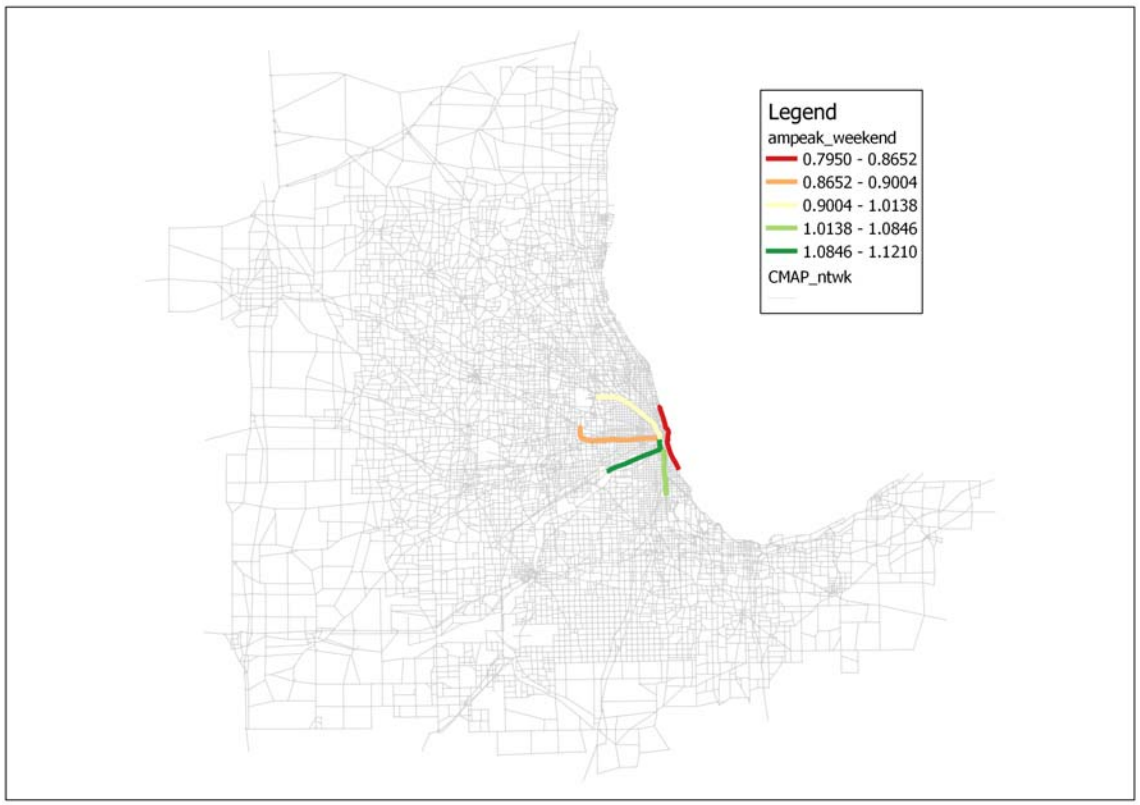


Figure 5.21: B-Index for AMPEAK-Weekend on six major routes

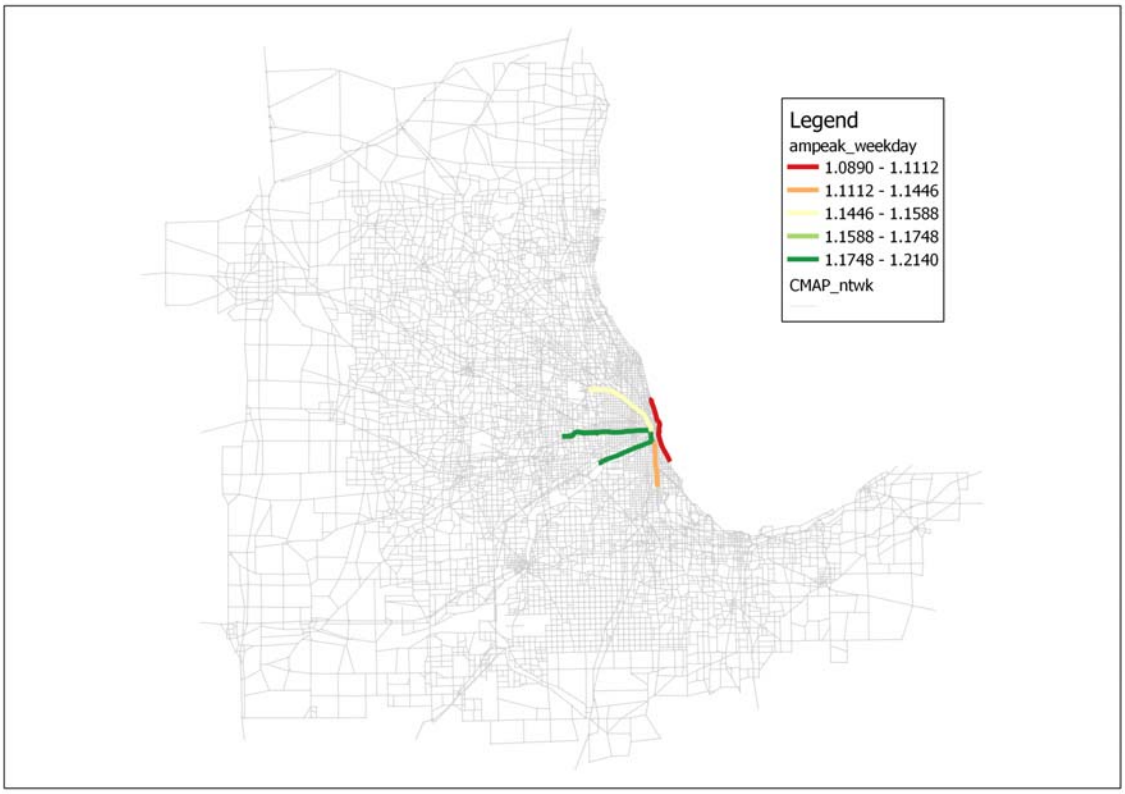


Figure 5.22: P-Index for AMPEAK-Weekday on six major routes

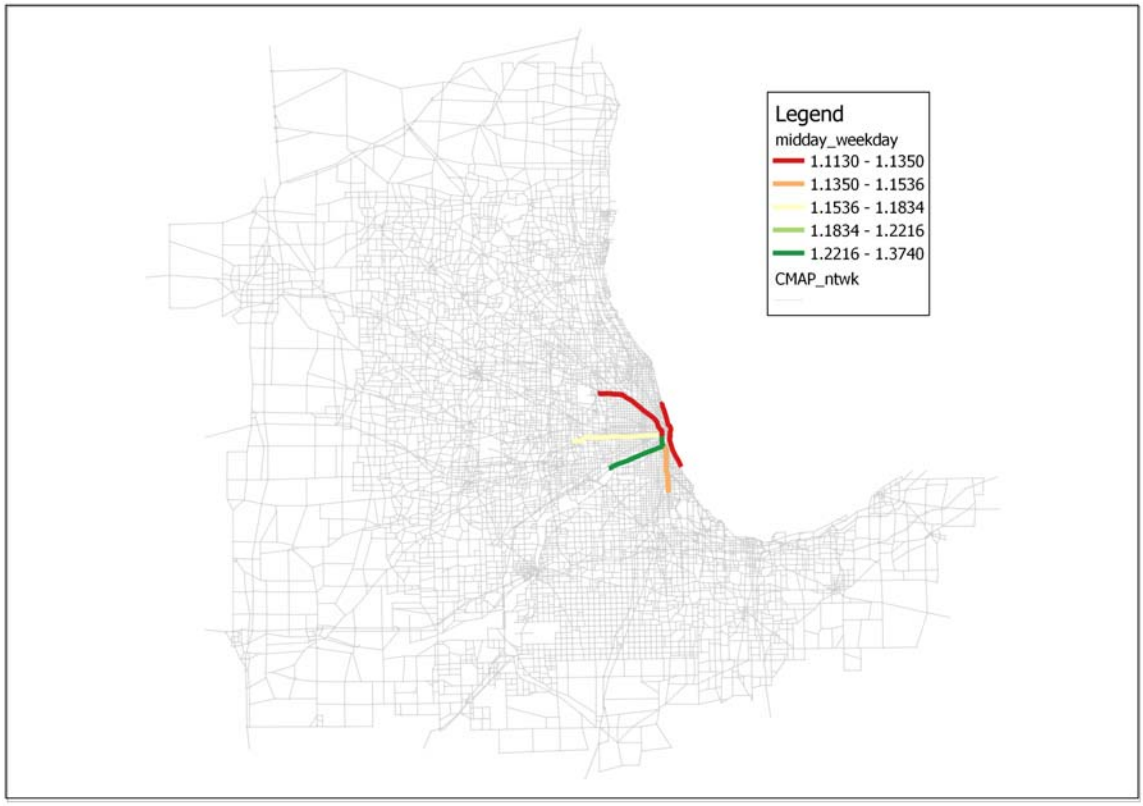


Figure 5.23: P-Index for MIDDAY-Weekend on six major routes

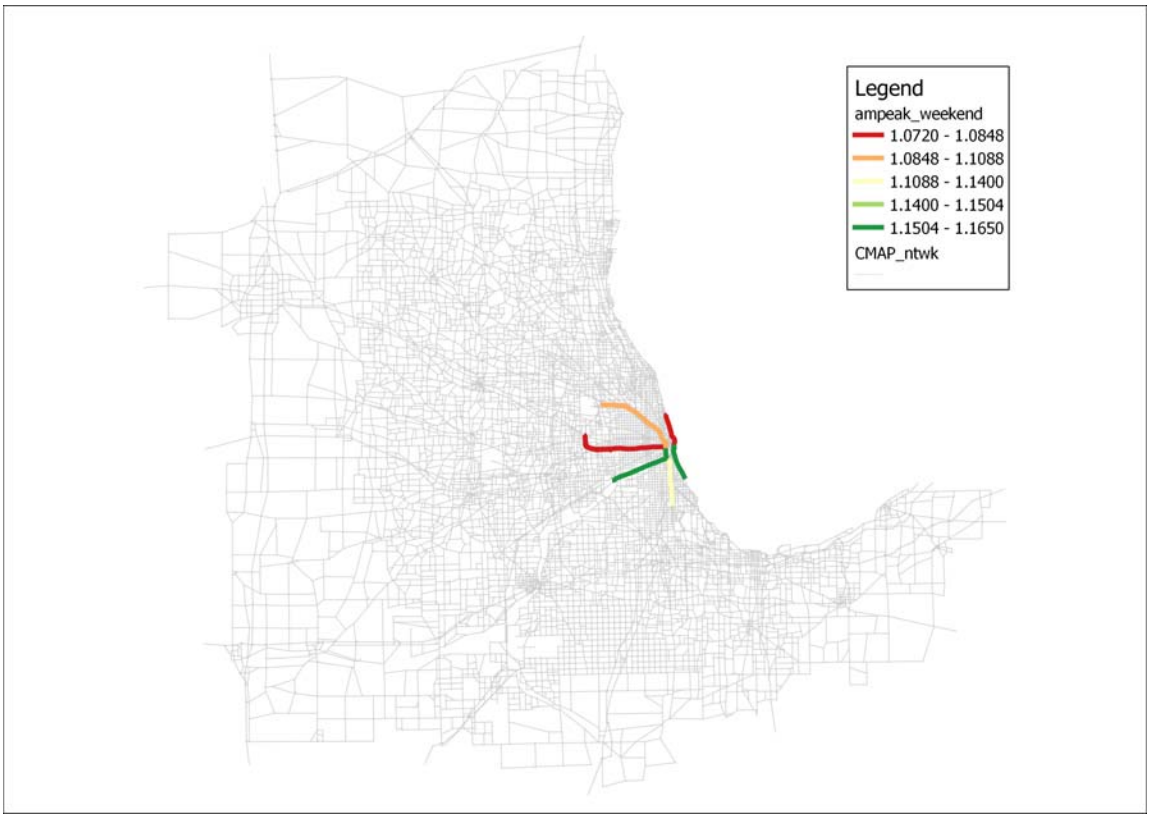


Figure 5.24: P-Index for AMPEAK-Weekend on six major routes

Bibliography

Bard, J. & Bennett, J. (1991), 'Arc reduction and path preference in stochastic acyclic networks', *Management Science* **37(2)**, 198–215.

Bracewell, R. (2000), *The Fourier Transform and Its Applications*, 3rd edn, McGraw-Hill, Boston.

Carraway, R. L., Morin, T. L. & Moskowitz, H. (1990), 'Generalized dynamic programming for multicriteria optimization', *European Journal of Operational Research* **44**(1), 95–104.

Chen, Y., Bell, M. G. H., Wang, D. & Bogenberger, K. (2006), 'Risk-averse time-dependent route guidance by constrained dynamic a* search in decentralized system architecture', *Transportation Research Record* **1944**, 51–57.

CMAP (2006), Travel demand modeling for the conformity process in northeastern illinois, available at http://www.cmap.illinois.gov /uploaded-files/publications/other_publications /pm25_conformity_analysis_b.pdf, last accessed on 6/10/2009, Chicago Metropolitan Agency for Planning.

Cormen, T., Leiserson, C., Rivest, R. & Stein, C. (2001), *Introduction to Algorithms*, MIT Press, Cambridge, MA.

Eiger, A., Mirchandani, P. B. & Soroush, H. (1985), 'Path preferences and optimal paths in probabilistic networks', *Transportation Science* **19**(1), 75–84.

Fan, Y., Kalaba, R. & Moore, J. (2005), 'Shortest paths in stochastic networks with correlated link costs', *Computers and Mathematics with Applications* **49**(9-10), 1549–1564.

Fan, Y. & Nie, Y. (2006), 'Optimal routing for maximizing the travel time reliability', *Networks and Spatial Economics* **3**(6), 333–344.

FHWA (2005), Traffic congestion and reliability: Trends and advanced strategies for congestion mitigatio, Technical report, Federal Highway Administration.

Frank, H. (1969), 'Shortest paths in probabilistic graphs', *Operations Research* **17**(4), 583–599.

Fu, L. (2001), 'An adaptive routing algorithm for in-vehicle route guidance systems with real-time information', *Transportation Research Part B* **35**(8), 749–765.

Fu, L. & Rilett, L. R. (1998), 'Expected shortest paths in dynamic and stochastic traffic networks', *Transportation Research Part B* **32**(7), 499–516.

Gao, S. & Chabini, I. (2006), 'Optimal routing policy problems in stochastic time-dependent networks', *Transportation Research Part B* **40**(2), 93–122.

Hadar, J. & Russell, W. (1969), 'Rules for ordering uncertain prospects', *American Economic Review* **59**, 25–34.

Hall, P. R. (1983), 'Travel outcome and performance: The effect of uncertainty on accessibility', *Transportation Research Part B* **17**(4), 275–290.

Hall, R. W. (1986), 'The fastest path through a network with random time-dependent travel time', *Transportation Science* **20**(3), 182–188.

Kaparias, I., Bell, M., Chen, Y. & Bogenberger, K. (2007), 'Icnavs: a tool for reliable dynamic route guidance', *IET Intellegent Transportation Systems* **1**(4), 225–253.

Lo, H. K., Luo, X. W. & Siu, B. W. Y. (2006), 'Degradable transport network: travel time budget of travelers with heterogeneous risk aversion', *Transportation Research Part B* **40**(9), 792–806.

Loui, R. P. (1983), 'Optimal paths in graphs with stochastic or multidimensional weights', *Communications of the ACM* **26**(9), 670–676.

Miller-Hooks, E. (1997), Optimal Routing in Time-Varying, Stochastic Networks: Algorithms and Implementations, PhD thesis, Department of Civil Engineering, University of Texas at Austin.

Miller-Hooks, E. D. (2001), 'Adaptive least-expected time paths in stochastic, time-varying transportation and data networks', *Networks* **37**(1), 35–52.

Miller-Hooks, E. D. & Mahmassani, H. S. (2000), 'Least expected time paths in stochastic, time-varying transportation networks', *Transportation Science* **34**(2), 198–215.

Miller-Hooks, E. D. & Mahmassani, H. S. (2003), 'Path comparisons for a priori and time-adaptive decisions in stochastic, time-varying networks', *European Journal of Operational Research* **146**(2), 67–82.

Mirchandani, P. B. (1976), 'Shortest distance and reliability of probabilistic networks', *Computers and Operations Research* **3**(4), 347–355.

Mirchandani, P. & Soroush, H. (1987), 'Generalized traffic equilibrium with probabilistic travel times and perceptions', *Transportation Science* **21**(3), 133–152.

Murthy, I. & Sarkar, S. (1996), 'A relaxation-based pruning technique for a class of stochastic shortest path problems', *Transportation Science* **30**(3), 220–236.

Murthy, I. & Sarkar, S. (1998), 'Stochastic shortest path problems with piecewise linear concave linear functions', *Management Science* **44**(11), 125–136.

Nie, Y. (2006), *A Programmer's Manual for Toolkit of Network Modeling (TNM)*, University of California, Davis.

Nie, Y. & Wu, X. (2009*a*), 'Reliable a priori shortest path problem with limited spatial and temporal dependencies', *In Proceedings of the 18th International Symposium on Transportation and Traffic Theory* pp. 169–196.

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

Nie, Y., Wu, X., Dillenburg, J. F. & Nelson, P. C. (2010), 'Reliable route guidance: A case study from chicago', *Transportation Research Part A* **Under revision**.

Polus, A. (1979), 'A study of travel time and reliability on arterial routes', *Transportation* **8**(2), 141–151.

Schrank, D. & Lomax, T. (2011), 2010 urban mobility study, Technical report, Texas Transportation Institute, Texas A&M University, College Station, TX.

Sen, S., Pillai, R., Joshi, S. & Rathi, A. (2001), 'A mean-variance model for route guidance in advanced traveler information systems', *Transportation Science* **35**(1), 37–49.

Shao, H. W. L., Meng, Q. & Tam, M. L. (2006), 'Demand-driven traffic assignment problem based on travel time reliability', *Transportation Research Record* **1985**, 220–230.

Sivakumar, R. & Batta, R. (1994), 'The variance-constrained shortest path problem', *Transportation Science* **28**(4), 309–316.

Uchida, T. & Iida, Y. (1993), Risk assignment: a new traffic assignment model considering risk of travel time variation, *in* C. Daganzo, ed., 'Proceedings of the 12th International Symposium on Transportation and Traffic Theory', Elsevier, Amsterdam, pp. 89–105.

von Neumann, J. & Morgenstern, O. (1967), *Theory of Games and Economic Behavior*, Wiley, New York.

Waller, S. T. & Ziliaskopoulos, A. K. (2002), 'On the online shortest path problem with limited arc cost dependencies', *Networks* **40**(4), 216–227.

Wu, X. & Nie, Y. (2009), 'Implementation issues in approximate algorithms for reliable a priori shortest path problem', *Journal of the Transportation Research Board* **2091**, 51–60.

Wu, X. & Nie, Y. (2011*a*), 'Application of discrete fourier transform in finding reliable shortest paths', *Journal of Transportation Research Board* **In press**.

Wu, X. & Nie, Y. (2011*b*), 'Modeling heterogeneous risk-taking behavior in route choice: A stochastic dominance approach', *Transportation Research Part A* **In press**.

Yin, Y., Lam, W. H. K. & Ieda, H. (2004), 'New technology and the modeling of risk-taking behavior in congested road networks', *Transportation Research Part C* **12**(3-4), 171–192.