Review of Definitions of Travel Time Reliability

Lily Elefteriadou and Xiao Cui

Draft – May 20th, 2005

In recent years, there has been an abundance of studies on travel time reliability. This document summarizes several definitions of reliability, presents some preliminary data analysis findings related to travel time, discusses the advantages and disadvantages of various definitions, and presents conclusions and recommendations.

Reliability Definitions

The concept of reliability is relatively new in the engineering disciplines, but it is becoming increasingly important as part of the engineering design process, the establishment of preventive maintenance programs, and others. In various areas of engineering and manufacturing, there is one definition of reliability generally accepted. For example, Ebeling (1997) defined reliability as "the probability that a component or system will perform a required function for a given period of time when used under stated operating conditions. It is the probability of a non-failure over time." Ebeling further states that the definition must be made specific by providing an unambiguous and observable description of a failure, including the unit of time over which failure will be evaluated. This definition of reliability is widely accepted in engineering. In the area of transportation, on the other hand, there are several different definitions of reliability which have been developed.

In the 1998 California Transportation Plan (Booz-Allen & Hamilton, Inc.,1998; also discussed in NCHRP 311, 2001; Estimation of Reliability, 2000; Reliability as a Measure of Transportation System Performance, 2000; Reliability Measures for Highway Systems and Segments, 2002), reliability is defined as the level of variability between the expected travel time (based on scheduled or average travel time) and the actual travel time experienced. In that definition, the expected travel time is based on scheduled or average travel time, while the actual travel time incorporates the effects of non-recurrent

congestion. The expected travel time is well defined, and is represented by the mean travel time during the period of interest. Likewise, the actual travel time is well defined. The level of variability however is not well defined, i.e., it is not clear what variability level is considered reliable (failure is not defined). In addition, for facilities that are congested for a large portion of the time, the expected travel time would be high. In those cases the difference between the two values (expected –actual travel time) may be small, labeling the facility as "reliable", when it is consistently congested.

In Measures of Effectiveness for Major Investment Studies (Turner et al, 1996; also reviewed in Estimation of Reliability, 2000; Reliability as a measure of Transportation System Performance, 2000; Reliability Measures for Highway Systems and Segments, 2002), trip time reliability is defined as the range of travel times experienced during a large number of daily trips. This definition provides the range of travel times and, similarly to the previous definition, does not define when "failure" has occurred. In addition, it is not related to congestion, or the percent of time the facility operates as expected.

Shaw (2000) suggested the following definition of reliability: "Reliability is generally defined as the operational consistency of a facility over an extended period of time. Reliability has historically been associated with the performance of mechanical equipment or devices. In this context, reliability is defined as the probability of a device performing its purpose adequately for the period of time intended under the stated operating conditions". The definition is similar to the definitions frequently used in reliability engineering.

Florida DOT (2000) developed and documented the Florida Reliability Method. They defined reliability on a highway segment as the percent of travel that takes no longer than the expected travel time plus a certain acceptable additional time. They define three major components of reliability: travel time, expected travel time, and acceptable additional time.

- Travel time - the time it takes a typical commuter to move from the beginning to the end of a corridor.
- Expected travel time- the median travel time across the corridor during the time period being analyzed.
- Acceptable additional time- the amount of additional time, beyond the expected travel time, that a commuter would find acceptable during a commute.

Mathematically, the acceptable travel time can be estimated as follows:

Acceptable TT = $\bar{x} + \Delta$

$\bar{x}$: The median travel time

Δ: Acceptable additional time, expressed as a percentage of median travel time

The percent of reliable travel time is calculated as the percent of travel on a corridor that takes no longer than this acceptable travel time. This definition defines failure clearly and quantitatively, however it relies on the median travel time, which may change over time as a function of demand. Thus this definition does not allow the tracking of reliability over time for a given facility.

TTI (2000) defined reliability and variability separately in their report. Reliability is commonly used in reference to the level of consistency in transportation service; variability is the amount of inconsistency on operating conditions. To quantify the reliability and variability, they defined two measures. A measure of reliability they recommended is the Buffer Time, which is the amount of extra time that must be allowed for the traveler to achieve their destination in a high percentage of the trips. A measure of variability is the average travel time plus one or two standard deviations. Lomax et al. (2004) defined the reliability Buffer Time Index as follows:

Buffer Index (BI) =

[95^{th} percentile confidence travel rate-average travel rate] /[average travel rate] x 100%

Similarly to the Florida Reliability Method which is based on the median travel time, this definition is based on the average travel rate. For similar reasons, then this definition also does not allow the tracking of reliability over time for a given facility.

In NCHRP report 398 (1997), reliability is defined as the impact of non-recurrent congestion on the transportation system. In NCHRP report 399 (1998), travel time reliability is defined as a measure of the variability of travel time; it is stated that reliability could be presented as the standard deviation of travel time.

In AASHTO's freight report (2002), reliability is defined as the percent of on-time performance for a given time schedule. This definition is provided for freight transportation. For example, within 1.0 hours of schedule, reliability is defined as being 99 percent on time. This definition is more consistent with the one generally accepted in engineering.

TranSystems (2003) explored some of the definitions that are recommended for travel time reliability. A common definition for reliability was recommended based on the probability of travel times meeting users' expectations. The report stated that different definitions of reliability depend on different viewpoint of users. In mathematical terms, reliability is the probability that a product or service performs adequately over the interval [0, t]. They stated that in other reports and articles, reliability is often defined as a probability or a percentage of acceptable performance, while in transportation it is typically reported as on-time performance.

NCHRP 8-36 (2004) reviewed the definition of travel time reliability in the F-SHRP program. In the F-SHRP Reliability Research Program, it is indicated that travel time reliability can be defined in terms of how travel times vary over time (e.g., hour-to-hour, day-to-day). This concept of variability can be extended to any other travel time-based metrics such as average speeds and delay. In that study the terms travel time variability and reliability are used interchangeably. Freeway travel time reliability metrics mentioned are the Buffer Time Index, 95^{th} percentile of travel times, coefficient of variation of travel times, percent "on-time performance", and travel time window. In Task 39 of NCHRP 8-36, the report concluded that there is no explicit definition of travel time reliability, but that travel time reliability has been used similarly to travel time variability.

The definitions provided above all refer to highway transportation. Other modes however have also defined and used reliability for their purposes. For example, in rail transportation, travel time reliability is defined as the percentage of on-time performance (AASHTO's freight report, 2002). No specific number is given to define "on-time" performance. Also, the US DOT (May, 2005) evaluates the reliability of travel for airlines using the percentage of on-time performance. A flight is considered to be "on-time" if it is not delayed more than 15 minutes beyond the scheduled arrival (see http://www.transtats.bts.gov/HomeDrillChart.asp). Reliability for port operations is similarly defined (TranSystems, 2005).

Preliminary Data Analysis of Travel Times

In an NSF-sponsored project currently underway (Elefteriadou et al, 2005) researchers have been conducting preliminary analyses for travel time data in a freeway corridor in Toronto, Canada. Figure 1 presents a sketch of the facility studied, along with the locations of the loop detectors which provided flow and speed data. The facility is a six lane highway, with three lanes per direction. Figure 2 presents the plot of travel time vs. flow data for Link 2 of the study site (merge area). Note that the travel time is estimated based on the speed at the loop detector, and was not obtained directly for the link.

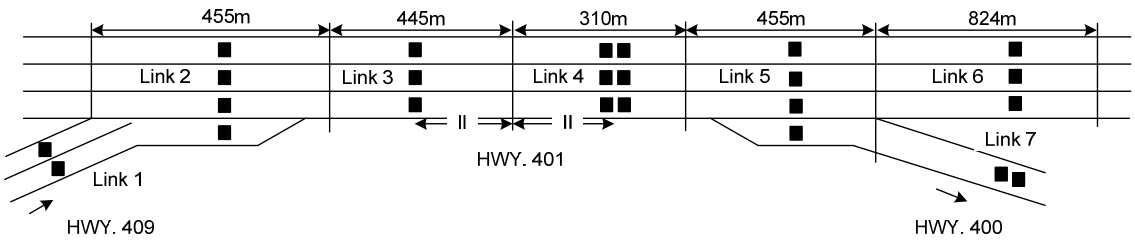

Figure 1 – Sketch of Study Area

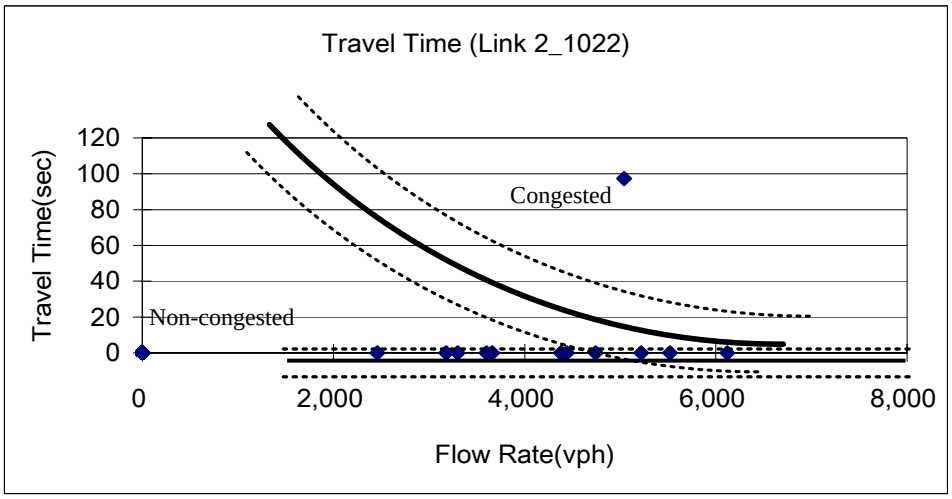

Figure 2 – Travel Times as a Function of Flow

Figure 3 presents the two travel time distributions (congested and non-congested travel times) for link 2, and for flow rates ranging from 4000 – 4500 vph, while Figure 4 presents the same distributions for flow rates ranging from 4500-5000 vph.

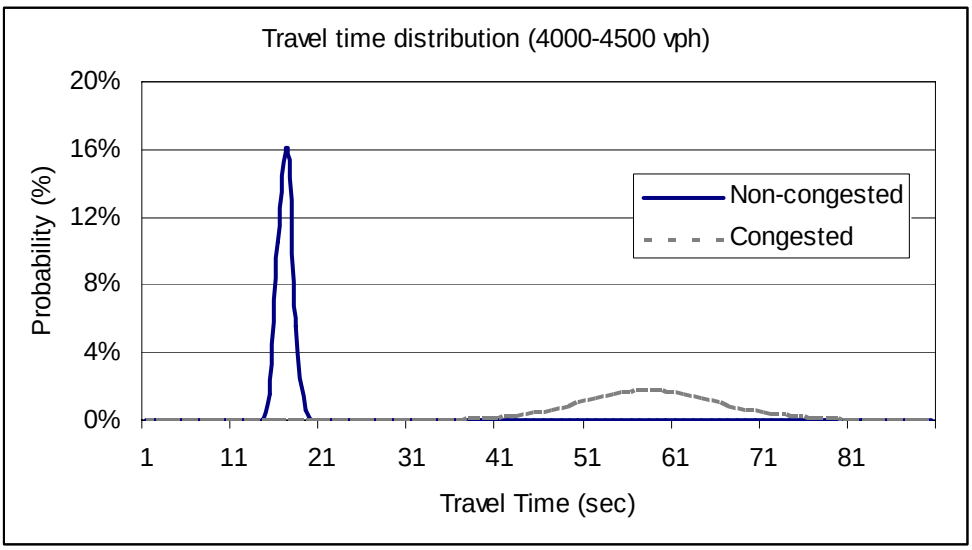

Figure 3 – Travel Time Distributions for 4000-4500 vph

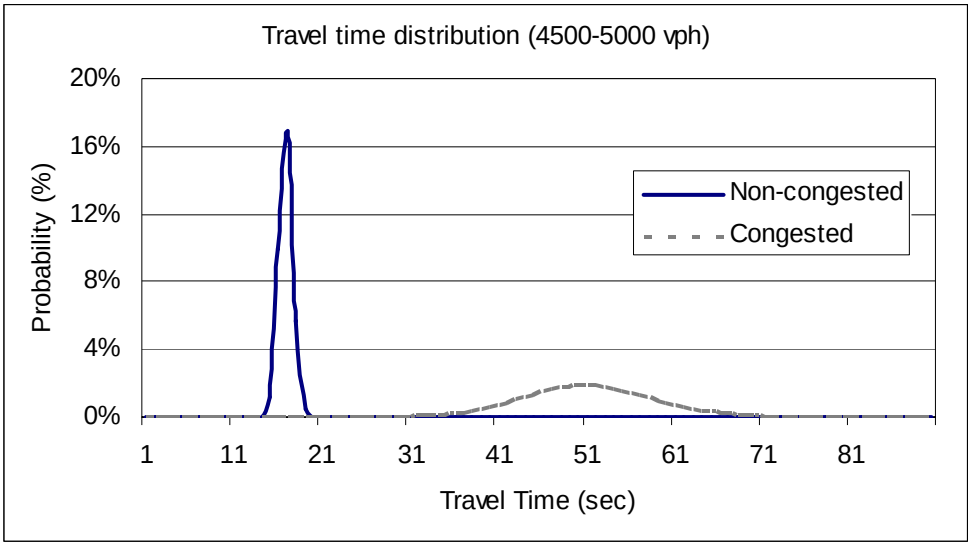

Figure 4 – Travel Time Distributions for 4500-5000 vph

As shown, the variability for non-congested travel times is much smaller than that for congested travel times. It should be noted that the flow rates represent discharge flows, and not demands. It is possible that travel time could be determined based on demand flow rates, but it is very difficult to obtain those in the field from the subject section.

Advantages and Disadvantages of Existing Travel Time Reliability Definitions

In summary, the definitions of travel time reliability found in the literature fall in two categories:

a) Definitions, based on the concept of reliability used in manufacturing and other engineering disciplines: Those define reliability as the probability of a non-failure over time. Those definitions require that the failure be clearly defined in quantitative terms. For example, one can specify that failure is defined as the condition of operating speed falling below a certain boundary value, i.e., 40 mph, or 10 mph below the speed limit, etc.
b) Definitions based on the concept of variability of travel time. Those definitions focus on the perspective of the traveler. They define reliability as the "unpredictability" of travel times, and they are intended to be used by travelers

budgeting time for their trip. Those definitions use some measure of central tendency (mean, median, etc.) and a measure of dispersion (standard deviation, acceptable additional time, etc.) to assist travelers in the time budgeting for their trip.

The first type of definition is more appropriate for tracking the performance of a facility over time, and is closely related to congestion. Because it uses a fixed boundary value to define failure, an agency can use this definition to evaluate trends in the performance of a facility or a network. Those types of definitions can be used to estimate the frequency of the presence of congestion within a year, and changes, as well as the rate of changes, over time. Travelers, however would not find such a measure of particular relevance when trying to estimate the required travel time between an origin and a destination.
The second type of definition would be more useful to travelers (or freight company operators), as it can provide an indication of the expected travel time, and the variability of that travel time. Graphs such as those shown in Figures 3 and 4 could be developed, and then interpreted to be presented to the general public, so as to provide travelers with an indication of the expected variability on a given link, or route. This type of definition, however, would not be appropriate for use by an agency to monitor their performance, because these definitions are based on the mean (or mode) of travel time, both of which may change over time. These definitions are not based on a constant value to be used as a benchmark by the operating agency.

In summary, it appears that two different performance measures related to reliability could be developed: a) A performance measure which would be appropriate for use by an agency to monitor the performance of various facilities, and b) a performance measure appropriate to be ultimately provided to travelers for estimating travel times between a given origin and a given destination.

REFERENCES

AASHTO, “Transportation Invest in America, the Bottom Line”, American Association of State Highway and Transportation Officials, 2000

Ebeling, C.E., “Introduction to Reliability and Maintainability Engineering”, McGraw-Hill Companies Inc., 1997

Elefteriadou, L., G. Natarajan, P. Johnson, J. Yeon, “Travel Time Route Reliability” On-going research, NSF, 2005

Florida DOT, “The Florida Reliability Method in Florida’s Mobility Performance Measures Program”, Florida Department of Transportation, 2000

Jackson, D.D., “Reliability as a Measure of Transportation System Performance”, Texas A&M University, 2000

Lomax, T., Turner, T., and Margiotta, R., “Monitoring Urban Roadways in 2002: Using Archived Operations Data for Reliability and Mobility Measurement”, Texas Transportation Institute, College Station, 2004

Margiotta, R. “Congestion and Reliability, Agency and User Perspectives”, presented as the 84th TRB Annual Meeting, Washington, D.C., 2005

NCHRP 8-36 Interim Report, “Guide to Effective Freeway Performance Measurement”, Cambridge Systematics, Inc., 2004

Lomax, T. et al, “NCHRP Report 398, Quantifying Congestion: Final Report”, Transportation Research Board, National Research Council, Washington, D.C., 1997

NCHRP Report 399, “Multimodal Corridor and Capacity Analysis Manual”, Transportation Research Board, National Research Council, Washington, D.C., 1998

Reynolds, Smith & Hills, Inc., “Estimation of Reliability, Task 1- Analyze and Compare Reliability Estimation Techniques”, Reynolds, Smith & Hills, Inc, 2000

Shaw, T., “Draft Estimation of Reliability for Florida’s Mobility Performance Measures Program”, Florida Department of Transportation Statistics Office, 2000

Shaw, T., “Reliability Measures for Highway Systems and Segments”, TranSystems Corporation, 2002

TranSystems, “Executive Summary, Full-Scale EMT Demonstration”, TranSystems Corporation, 2005

TTI, “Possible Reliability Measures”, Texas Transportation Institute, 2000

US DOT, “Air Travel Customer Report, May 2005”, US DOT, Aviation Customer Protection Division, May 2005