A NEW MEASURE OF TRAVEL TIME RELIABILITY FOR IN-VEHICLE NAVIGATION SYSTEMS

Ioannis Kaparias
Centre for Transport Studies
Imperial College London
Skempton Building
South Kensington Campus
London SW7 2BU, UK
Tel: +44 20 75945952
Fax: +44 20 75946102
Email: ik00@imperial.ac.uk

Michael G.H. Bell
Centre for Transport Studies
Imperial College London
Skempton Building
South Kensington Campus
London SW7 2BU, UK
Tel: +44 20 75946091
Fax: +44 20 75946102
Email: mghbell@imperial.ac.uk

Heidrun Belzner
BMW Group
Traffic Technology
80788 Munich, Germany
Tel: +49 89 38254267
Fax: +49 89 38230581
Email: Heidrun.Belzner@bmw.de

This paper was submitted to the Journal of Intelligent Transport Systems.

8 August 2008

A NEW MEASURE OF TRAVEL TIME RELIABILITY FOR IN-VEHICLE NAVIGATION SYSTEMS

Ioannis Kaparias[1], Michael G.H. Bell[1], Heidrun Belzner[2]

1 Centre for Transport Studies, Imperial College London, UK
2 Traffic Technology, BMW Group, Munich, Germany

ABSTRACT

This paper introduces a new measure of travel time reliability, for implementation in the dynamic routing algorithm of an intelligent car navigation system. The measure is based on the log-normal distribution of travel time on a link and consists of two indices corresponding to the extreme values of the distribution, such that they reflect the shortest and longest travel times that may be experienced on the link. Through a series of mathematical manipulations, the indices are expressed in terms of the characteristic values of the speed distribution on the link. An expression relating the indices of a route and the indices of the individual links forming it is derived. The accuracy of the measure is then assessed through a field experiment and the results are presented.

Keywords: Transportation, intelligent transportation systems, in-vehicle navigation systems, network reliability.

1 INTRODUCTION

The field of in-vehicle navigation is currently quickly evolving, as navigation systems are, together with the traditional car radio, the most important source of information in the vehicle. Car navigation systems were initially treated as luxury goods and were only offered as premium accessories on high-priced cars. In recent years, however, the market has seen the introduction of portable navigation devices besides the embedded ones, which are available at affordable prices. Forecasts predict that in the next few years, a large number of cars will be equipped with a navigation device, which indicates that car navigation systems are rapidly becoming everyday consumer goods.

The main objective of the routing function of a car navigation system is to compute the fastest route from the current position of the vehicle, established through satellite positioning, to the destination, input by the driver. Apart from the simple routing function, the so-called dynamic routing function, in which current and predicted traffic conditions are taken into account, is gaining importance and is gradually being incorporated in newer navigation systems.

Dynamic routing can be provided by two system architectures, namely the autonomous (or decentralised) and the supported (or centralised) system architectures. In the former, a route based on default estimates of link travel times, complemented in Europe by traffic information about current incidents broadcast by the Traffic Message Channel (TMC), is suggested to the driver, while the entire computation takes place in the vehicle unit itself. The latter on the other hand involves a two-way data exchange between a subscribing guided vehicle and a Traffic Information Centre (TIC), increasing the range and accuracy of information available to subscribers; some of the computation takes place at the TIC and subroutes are transmitted to the vehicle, for example at decision points.

In recent research work (Chen and others, 2005a; 2005b; 2006; Kaparias and others, 2007a; 2007b; Kaparias, 2008; Kaparias and others, 2008), a new dynamic routing algorithm was developed for the autonomous system architecture, aiming to take into account travel time reliability and use it as a second optimisation criterion for finding a set of feasible routes to be suggested to the driver. Reliability was defined as the probability of not encountering congestion along a link or a route and was expressed as a [0,1] continuum. Nonetheless, this definition is not suitable when it comes to implementing it in car navigation systems for two main reasons: firstly, it is not understandable by the driver and it cannot be translated into an expected delay value, which is what the driver is ultimately interested in, and secondly, when it comes to computing the reliability of a route, the measure is dependent on the number of links forming the route, as it expresses the probability of not encountering congestion on any link along the route, which means that the more links there are on the route the lower its reliability.

The aim of this work is to define and implement a better reliability measure to be used in the new dynamic routing algorithm. The distribution of travel times is taken into account and a measure reflecting the maximum possible delay is formulated, based on the descriptive statistics of the distribution. Furthermore, as the distribution of travel times is usually not available, the new measure is re-formulated and expressed in terms of speed. An expression for calculating the reliability of a path based on the reliability of the individual links is also derived. Finally, the accuracy of the new measure is tested through a field experiment, in which the travel times estimated using the measure are compared with actual travel times.

This paper is structured as follows: Section 2 reviews previous work on travel time variability and reliability, their importance to travellers and existing methods of measuring them. The new reliability measure of links is introduced and defined in Section 3, while the mathematical transformations carried out to express it in terms of speed are presented in Section 4. In Section 5 the reliability definition is extended from links to routes and an expression relating the two is derived. Finally, the field experiment carried out to demonstrate the accuracy of the measure is set out in Section 6 and its results are presented, while Section 7 concludes the paper.

2 LITERATURE REVIEW

In this section existing literature on travel time reliability is reviewed. The importance of travel time uncertainty is examined first, followed by a review of studies looking at the distribution of travel times. Then, a discussion of measures of reliability adopted by different studies in the past is given.

2.1 The importance of travel time uncertainty and reliability

The importance of travel time uncertainty has been the objective of many research studies in the past and has therefore been extensively analysed from the traveller's perspective. Many studies have concluded that although travel time is an important factor affecting the traveller's route choice behaviour, travel time variability can be even more important. Travellers are interested in how long it will take them to reach their destination, but are even more concerned with the reliability of their prediction of total travel time. A wrong travel time prediction results in either an early arrival at the destination or in a delay. None of these situations is appreciated by the traveller, with delays usually having more severe consequences for him/her (e.g. late arrival at the workplace) and therefore not being tolerated. Hence, much research has focused on developing methods for modelling travel time reliability.

Many empirical studies have identified the importance of travel time reliability. In a study by Jackson and Jucker (1981) it was found that there is a trade-off between travel time and reliability for travellers. In another study by Abdel-Aty and others (1995) it was demonstrated that a traveller's route choice is influenced by reliability and also by traffic information, as this is a way to reduce travel time uncertainty. A later study by Lam and Small (2001) concluded that reliability is an important factor affecting route choice and distinguished between genders, showing that women tend to be more risk-averse than men and hence value reliability more.

Acknowledging the importance of reliability, many studies have attempted to incorporate it into a model, such as the work of Noland and Small (1995), who identified that many travellers adopt a safety margin during their morning commute, so as to take uncertainty into account as much as possible, and applied penalties for early and late arrival. In another study by Bates and others (2002), travel time uncertainty was formulated as an additional "schedule disutility" in the traveller's journey time valuation function, proportional to the standard deviation of the travel time distribution, while a study by Liu and others (2004), formulated travellers' route choice as a mixed logit model, containing coefficients representing individual travellers' preferences towards travel time, reliability and cost.

More studies, aiming to model travel time reliability exist in the literature; comprehensive reviews of this topic have been carried out by Bates and others (2001) and Noland and Polak (2002).

2.2 The distribution of travel times

The distribution of travel times at the same demand level has been extensively investigated in the last decades, as it is a very important topic when it comes to modelling travel time variability and reliability. A large number of research studies are available in the literature, attempting to fit the distribution of travel times to one of the continuous probability distributions.

Wardrop (1952) first identified that travel times follow a skewed distribution. This was subsequently verified by Herman and Lam (1974), who carried out an empirical study on travel time variability. The study concluded that travel times indeed follow a skewed distribution and that only the 60% lower values fit a normal distribution well, thus contradicting the assumption made at the time, that travel times were normally distributed. In order to fit the actual distribution of travel times, the log-normal or gamma distribution was suggested.

In later studies, the above finding was further confirmed. Namely, in a small empirical study by Polus (1979), it was concluded that travel times best fit a gamma distribution. A log-normal or gamma distribution was suggested by Dandy and McBean (1984), while studies by Mogridge and Fry (1984), Montgomery and May (1987), Rakha and others (2006) and Chen and others (2007) also derived a log-normal fit.

Despite more recent studies assuming a normal distribution for travel times for reasons of simplicity (Lomax and others, 2003), the skewness of the distribution remains an important characteristic that needs to be considered when analysing travel time variability and reliability. Therefore, it is assumed in this study that travel times follow a log-normal distribution; this assumption is used in Section 3, where the new measure of reliability is defined.

2.3 Existing measures of reliability

A considerable amount of research has focussed on defining adequate measures for quantifying travel time reliability. Most of them use various characteristics of the travel time distribution, such that two types of measures can be identified: measures indicating the probability that a certain link is unusable and measures attempting to quantify the amount of congestion that may be encountered on a link.

The most widely used reliability measure so far belongs to the first category and is the one defined by Bell and Iida (1997), which is expressed as the probability of a link to be uncongested, based on the assumption that the condition of traffic flow on a link is binary, i.e. congested or uncongested. The range of values of this measure is a [0,1] continuum. While this measure is suitable for quantifying the reliability of a network, it has the disadvantage that it does not give any indication on the amount of congestion that may be encountered. Other measures of the first category, such as the ones by Al-Deek and Emam (2006), Chen and others (2007) and Eleftheriadou and Cui (2007), are limited in the same way.

Thus, a number of studies have attempted to quantify the reliability of a link by using the travel time distribution directly, deriving measures of the second category. The first measure was adopted in a study by Polus (1979), where reliability was defined as the inverse of the standard deviation of the link's travel time distribution. The main disadvantage of this measure, however, is the fact that it is not dimensionless. This is also the case of measures developed in further studies, such as Dandy and McBean (1984), who used the 95th percentile travel time, and Lam and Small (2001), who quantified reliability as the difference between the 90th percentile and the median of the travel time distribution.

Some studies developed reliability measures, not only considering the width of the travel time distribution, but also its skewness. For instance, van Lint and van Zuylen (2005) proposed two reliability metrics, based on the 10th, the 50th and the 90th percentiles of the travel time distribution. However, the most important contribution to defining reliability measures has been made by Lomax and others (2001; 2003), who presented a series of measures of reliability and categorised them in three groups, as statistical range measures, buffer time measures and tardy trip indicators. Statistical range measures were defined as presenting information in a relatively "unprocessed format", meaning that they are mainly based on concepts only familiar to statisticians and generally not understandable to ordinary travellers, while buffer time measures were defined as being intended to relate well to the way travellers make decisions, and are therefore understandable by them. They indicate how much extra time a traveller should allow for his/her journey to account for uncertainty in the travel conditions. Finally, tardy trip indicators were defined as representing the unreliability impacts in terms of the amount of late trips. Following the categorisation of the various reliability measures, three measures were recommended as the most appropriate to use, one from each category. These were, the percent variation, the buffer time index and the misery index, defined as:

$$Percent\ Variation = \frac{Standard\ Deviation}{Average\ Travel\ Time} \times 100$$

$$Buffer\ Time\ Index = \frac{95^{\%}\ Travel\ Rate - Average\ Travel\ Rate}{Average\ Travel\ Rate} \times 100$$

$$Misery\ Index = \frac{Av.\ Travel\ Rate\ for\ Worst\ 20\%\ Trips - Av.\ Travel\ Rate}{Av.\ Travel\ Rate}$$

where the term "travel rate" represents travel time per distance unit.

While the advantage of these measures is that they quantify the amount congestion to be encountered such that delay values can be obtained, their drawback is that the range of their values is not a [0,1] continuum, thus requiring major modifications of the new dynamic routing algorithm, for which a reliability measure is needed (Chen and others, 2005a; 2005b; 2006; Kaparias and others, 2007a; 2007b; Kaparias, 2008; Kaparias and others, 2008).

2.4 Summary

In summary, the findings and conclusions of the literature review are:

- Travel time variability is very important to travellers; many travellers are risk-averse and are therefore prepared to choose a longer route if it is more reliable, i.e. if they can be certain that they will arrive on time at their destination. Thus, incorporating this in in-vehicle navigation will be a very useful feature that will advance the current status of the car navigation systems technology.

- Travel times follow a right-skewed distribution, which is very close to the log-normal or the gamma distribution. It can be therefore safely assumed, that travel times are log-normally distributed.

- A number of reliability measures have been adopted by several studies; however, they all have some important drawbacks making them unsuitable for an in-vehicle navigation algorithm. A new measure therefore needs to be developed.

3 QUANTIFYING LINK RELIABILITY

Having reviewed previous work on measures of travel time reliability, this section presents a new measure, aimed at car navigation systems. This new measure should however, meet a number of requirements, ensuring its accuracy and simplicity. These are presented next, followed by the description of the measure.

3.1 Requirements of a new reliability measure

The first requirement that the new reliability measure needs to meet is the comprehensibility by the driver. Many of the measures that have been used in the literature so far have had the major drawback, that it was not possible to communicate to the driver the reliability of a route, because it was in a non-understandable form. Taking into account the fact that the driver is often interested in how late or how early he/she will arrive at the destination, i.e. the amount of congestion that will be encountered rather than the probability of encountering congestion, the new measure should be designed to be easily convertible to a travel time.

Furthermore, a dimensionless quantity should be chosen to ensure that the new reliability measure is independent of the units used. This would enable the comparison between links and routes, whose characteristics are given in different units. In addition, the new reliability measure should be independent of the length of the link, so that the division of a link into two without any physical changes leaves the reliability measure unaffected. The measure should also be expressed as a [0,1] continuum, such that it can be employed in the new dynamic routing algorithm without any major modifications of the latter (Chen and others, 2005a; 2005b; 2006; Kaparias and others, 2007a; 2007b; Kaparias, 2008; Kaparias and others, 2008). When it comes to calculating the reliability of a route consisting of a series of links, it should be ensured that the measure used takes into account the statistical dependence between them, as it is very likely that congestion on one link will result in more links becoming congested too.

The latter point is covered in Section 5, where a method for computing the reliability of a route is given. The next paragraphs describe the new measure and the way of computing the reliability of links.

3.2 Definition of the reliability indices

As travel times on a link have been assumed to be log-normally distributed, the distribution of the travel time $t(l)$ of a link l is $t(l) \sim Log\text{-}N(\mu(l),[\sigma(l)]^2)$, where $\mu(l)$ and $\sigma(l)$ are the mean and standard deviation of the natural logarithm of travel time respectively. They are connected to the mean $\bar{t}\,(l)$ and variance $var[t(l)]$ of the travel time distribution by the following expressions:

$$\mu(l) = ln(\bar{t}\,(l)) - \frac{1}{2}\,ln\left(1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2}\right)$$

$$\sigma(l) = \sqrt{ln\left(1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2}\right)}$$

Defining the dimensionless travel time variation logarithm as

$$T_{log}(l) = ln\left(1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2}\right) \tag{1}$$

the expressions become:

$$\mu(l) = ln(\bar{t}\,(l)) - \frac{1}{2}\,T_{log}(l) \tag{2}$$

$$\sigma(l) = \sqrt{T_{log}(l)} \tag{3}$$

The standard deviation of the distribution indicates the spread of the travel time values around the mean. Hence, large $\sigma(l)$ values mean that the spread of the logarithms of the travel times around the mean of the logarithms of the travel times $\mu(l)$ is great, and so is the spread of the actual travel times around the mean travel time $\bar{t}\,(l)$, resulting in a greater uncertainty regarding the travel time that is to be experienced on link l. Because travel times are log-normally distributed, a confidence interval of the travel time to be experienced on the link is given by $\{\mu_{geo}(l) / [\sigma_{geo}(l)]^{z\alpha/2}, \mu_{geo}(l)\cdot[\sigma_{geo}(l)]^{z\alpha/2}\}$, where $\mu_{geo}(l) = e^{\mu(l)}$ and $\sigma_{geo}(l) = e^{\sigma(l)}$ are the geometric mean and standard deviation of the travel time distribution respectively, and $z_{\alpha/2}$ is the standard normal distribution tail probability for a confidence coefficient α, corresponding to a confidence level of 1-α. More specifically, for confidence coefficients α = 0.1, α = 0.05 and α = 0.001, corresponding to confidence levels of 90%, 95% and 99% respectively, the tail probabilities are $z_{0.05} = 1.65$, $z_{0.025} = 1.96$ and $z_{0.0005} = 2.58$.

Thus, based on these definitions, the confidence interval of the travel time experienced on link l can be expressed as $\{e^{\mu(l)\ -\ z\alpha/2\ \sigma(l)}, e^{\mu(l)\ +\ z\alpha/2\ \sigma(l)}\} = \{t(l)_{\alpha/2}, t(l)_{1-\alpha/2}\}$. This is an asymmetrical interval around $\bar{t}\,(l)$ giving an indication of what the maximum possible travel time (upped bound) and the minimum possible travel time (lower bound) on the link can be. However, for a link to be characterised as "reliable" or "unreliable", the scale of the maximum and minimum possible travel times with respect to the mean travel time needs to be known, as the same maximum and minimum travel time values can have different effects on different mean values. For example, if the link has a mean travel time of 5 minutes, a maximum time of 35 minutes is very large; however, if the link's mean travel time is 30 minutes, the same maximum additional travel time of 35 minutes is fairly small. Additionally, a dimensionless length-neutral measure of the reliability of the links is required, so as to be able to compare links with each other.

Thus, two reliability indices, namely the lateness index and the earliness index, are defined. The lateness reliability index, defined as

$$r_L(l) = \frac{\bar{t}\,(l)}{t(l)_{1-\alpha/2}} = \frac{\bar{t}\,(l)}{exp[\mu\,(l) + z_{\alpha/2}\,\sigma\,(l)]} = exp[ln(\bar{t}\,(l)) - \mu(l) - z_{\alpha/2}\cdot\sigma(l)] \tag{4}$$

is proposed as a measure of the reliability of links regarding their lateness, i.e. how much later than the expected arrival time the actual arrival may occur (latest possible arrival time). On the other hand, the earliness reliability index, defined as

$$r_E(l) = \frac{t(l)_{\alpha/2}}{\bar{t}\,(l)} = \frac{exp[\mu\,(l) - z_{\alpha/2}\,\sigma\,(l)]}{\bar{t}\,(l)} = exp[-\,ln(\bar{t}\,(l)) + \mu(l) - z_{\alpha/2}\cdot\sigma(l)] \tag{5}$$

is proposed as a measure of the reliability of links regarding their earliness, i.e. how much earlier than the expected arrival time the actual arrival may occur (earliest possible arrival time).

Using equations (2) and (3), both reliability indices can be expressed in terms of the mean and variance of the travel time distributions. Thus

$$r_L(l) = exp[ln(\bar{t}\,(l)) - \mu(l) - z_{\alpha/2}\cdot\sigma(l)] \Leftrightarrow$$
$$r_L(l) = exp[ln(\bar{t}\,(l)) - \big(ln(\bar{t}\,(l)) - \tfrac{1}{2}\cdot T_{log}(l)\big) - z_{\alpha/2}\cdot\sqrt{T_{log}(l)}] \Leftrightarrow$$
$$r_L(l) = exp[\tfrac{1}{2}\cdot T_{log}(l) - z_{\alpha/2}\cdot\sqrt{T_{log}(l)}] \tag{6}$$

and

$$r_E(l) = exp[-\, ln(\bar{t}\,(l)) + \mu(l) - z_{\alpha/2}\cdot\, \sigma(l)] \Leftrightarrow$$
$$r_E(l) = exp[-\, ln(\bar{t}\,(l)) + \left(ln(\bar{t}\,(l)) - ½ \cdot T_{log}(l)\right) - z_{\alpha/2}\cdot\sqrt{T_{log}(l)}] \Leftrightarrow$$
$$r_E(l) = exp[-\, ½\cdot T_{log}(l) - z_{\alpha/2}\cdot\sqrt{T_{log}(l)}] \tag{7}$$

$r_L(l)$ takes values ranging from 0 to 1, as the minimum value that the denominator in equation (4) may take is $\bar{t}\,(l)$. Lower values indicate low lateness reliability and high values show that the link is reliable. In the extreme cases, when $r_L(l)$ is close to 0, the maximum possible travel time is much longer than the mean travel time and hence the link is extremely unreliable in terms of lateness; on the other hand, when $r_L(l)$ is close to 1, the maximum travel time is approximately equal to the mean travel time, which means that almost no deviation from the mean travel time is to be expected, making the link very reliable in terms of lateness.

Similarly, the value of $r_E(l)$ also ranges between 0 and 1, as $\bar{t}\,(l)$ is the maximum value the nominator in equation (5) can take. Small values indicate that the minimum possible travel time is much smaller than the mean, implying hence a low reliability in terms of earliness, as a much earlier arrival than the one predicted is possible. Conversely, values close to 1 indicate that there is small variation of travel times and hence the link is reliable in terms of earliness.

Of course, the accuracy and reliability of both reliability indices strongly depends on the confidence level chosen. If a very high confidence level is chosen, the estimates are very reliable and one can be certain that the travel time will lie between the calculated maximum and minimum travel times; however, the confidence interval will be large and, in order to account for even the extreme values, the indices will not be representative of the non-extreme travel time values. On the other hand, if a lower confidence level is chosen, the indices will reflect the non-extreme travel times very well; however, most of the abnormally occurring long and short travel times will not be considered.

Since neither early nor late arrival at the destination are tolerated, both indices have to be taken into account when providing route guidance; however, due to the asymmetry of the interval, one of the two travel time bounds, maximum and minimum, is more critical, because it deviates more from the mean and implies therefore a greater level of uncertainty and hence, a lower reliability index. As the distribution in question is log-normal, and hence right-skewed, the maximum possible deviation from the mean is the upper bound of the interval, i.e. $t(l)_{1-\alpha/2}$, and the critical index is thus $r_L(l)$.

The tolerance level of early and late arrivals, though, is most likely going to be different for the two cases and, depending on their socioeconomic situation, drivers will be less tolerant to the one than to the other. This needs to be reflected in the implementation of the reliability indices in the routing function of a navigation system by applying appropriate weights to the indices, such that early or late arrivals are avoided more. This, however, extends beyond the scope of this work and therefore it is assumed in this paper that early and late arrivals are equally not tolerated.

It should be noted here, that travel time, and hence the mean and variance of its distribution, are not constant but time-dependent, which means that they vary between different times of the day, days of the week and times of the year, according to the varying demand levels on the road network. A procedure along the lines the so-called 'Flow Speed Model' method, introduced by Sung and others (2000) and modified by Kaparias and others (2007a) for deriving time-dependent travel times, has been developed for the calculation of time-dependent earliness and lateness indices (Kaparias, 2008). Nevertheless, a detailed description is beyond the scope of this paper and the reader is referred to the appropriate references for further information.

4 CALCULATING LINK RELIABILITY FROM LINK SPEED

Although in a theoretical framework one usually works with travel time distributions, this data is rarely available in this form and is usually expressed as speed measurements. Although the conversion of a speed measurement to a travel time measurement is simple, given that the link's length is constant, carrying out this procedure can be very inefficient, as a large number of measurements will need to be converted. Besides that, it is possible that the actual speed measurements will not be given, and only descriptive speed statistics will be provided. Thus, working with the speed distribution and attempting to relate the reliability of speed with the reliability of travel time, the relationship of the travel time and speed variances needs to be investigated.

Namely, for link l of length $\lambda(l)$, normally distributed speed $\upsilon(l)$ and log-normally distributed travel time $t(l)$, it is $t(l) = \lambda(l) / \upsilon(l)$; when it comes to the mean travel time, $\bar{t}\,(l) = \lambda(l) / \bar{\upsilon}_s(l)$, where $\bar{\upsilon}_s(l)$ represents space-mean speed, which is defined as the harmonic mean of the speed distribution. The travel time variation logarithm is thus affected, such that

$$T_{log}(l) = ln\left(1 + \frac{var\left(\frac{\lambda(l)}{\upsilon(l)}\right)}{\left(\frac{\lambda(l)}{\bar{\upsilon}_s(l)}\right)^2}\right) = ln\left(1 + \frac{[\lambda(l)]^2 \cdot var\left(\frac{1}{\upsilon(l)}\right)}{[\lambda(l)]^2 \cdot \left(\frac{1}{\bar{\upsilon}_s(l)}\right)^2}\right) \Leftrightarrow$$

$$T_{log}(l) = ln\left(1 + [\bar{\upsilon}_s(l)]^2 \cdot var\left(\frac{1}{\upsilon(l)}\right)\right) \tag{8}$$

In order to evaluate $T_{log}(l)$, the variance of the distribution of the inverse of speed needs to be known. Performing a first degree Taylor series expansion of the $1/\upsilon(l)$ term around $E[\upsilon(l)]$ (the expected value of the speed distribution, which is the time-mean speed $\bar{\upsilon}(l)$), the following expression is obtained:

$$\frac{1}{\upsilon(l)} \cong \frac{1}{E[\upsilon(l)]} - (\upsilon(l) - E[\upsilon(l)]) \cdot \frac{1}{(E[\upsilon(l)])^2} \tag{9}$$

Taking the expected values of both sides of the expression:

$$E\left(\frac{1}{\upsilon(l)}\right) \cong \frac{1}{E[\upsilon(l)]} - E[(\upsilon(l) - E[\upsilon(l)])] \cdot \frac{1}{(E[\upsilon(l)])^2} = \frac{1}{E[\upsilon(l)]} \tag{10}$$

Subtracting equation (9) from (10):

$$\frac{1}{\upsilon(l)} - \frac{1}{E[\upsilon(l)]} \cong -(\upsilon(l) - E[\upsilon(l)]) \cdot \frac{1}{(E[\upsilon(l)])^2} \Leftrightarrow$$
$$\left(\frac{1}{\upsilon(l)} - \frac{1}{E[\upsilon(l)]}\right)^2 \cong (\upsilon(l) - E[\upsilon(l)])^2 \cdot \frac{1}{(E[\upsilon(l)])^4} \Leftrightarrow$$
$$E\left(\left(\frac{1}{\upsilon(l)} - \frac{1}{E[\upsilon(l)]}\right)^2\right) \cong E((\upsilon(l) - E[\upsilon(l)])^2) \cdot \frac{1}{(E[\upsilon(l)])^4} \Leftrightarrow$$

$$var\left(\frac{1}{\upsilon(l)}\right) \cong \frac{var[\upsilon(l)]}{[\bar{\upsilon}(l)]^4} \tag{11}$$

Hence, substituting the variance term in (8) with the result derived in (11):

$$T_{log}(l) \cong ln\left(1 + [\bar{\upsilon}_s(l)]^2 \cdot \frac{var[\upsilon(l)]}{[\bar{\upsilon}(l)]^4}\right) = ln\left(1 + \frac{[\omega(l)]^2 \cdot var[\upsilon(l)]}{[\bar{\upsilon}(l)]^2}\right) \tag{12}$$

where $\omega(l) = \bar{v}_s(l)/\bar{v}(l)$ is the ratio of the space mean speed over the time mean speed. Since the time mean speed is always greater than or equal to the space mean speed, $\omega(l)$ only takes values in the range 0 to 1. Using the finding of Rakha and Wang (2005), resulting from Wardrop's (1952) formula relating time-mean speed and space-mean speed under the assumption of homogeneous traffic, i.e.

$$\bar{\upsilon}_s(l) \cong \bar{\upsilon}(l) - \frac{var[\upsilon(l)]}{\bar{\upsilon}(l)}$$

equation (12) can be reformulated as

$$T_{log}(l) \cong ln\left(1 + \frac{var[\upsilon(l)]}{[\bar{\upsilon}(l)]^2} \cdot \left(1 - \frac{var[\upsilon(l)]}{[\bar{\upsilon}(l)]^2}\right)^2\right) \tag{13}$$

which can be useful when the actual speed measurements are not given and only the time-mean speed and variance of the distribution are provided.

Expression (12) or (13) can then be substituted into equations (6) and (7), so as to calculate the earliness and lateness indices in terms of the speed distribution. Using equation (11), it can also be found that:

$$var[t(l)] = var\left(\frac{\lambda(l)}{\upsilon(l)}\right) = [\lambda(l)]^2 \cdot var\left(\frac{1}{\upsilon(l)}\right) \cong [\lambda(l)]^2 \cdot \frac{var[\upsilon(l)]}{[\bar{\upsilon}(l)]^4} \tag{14}$$

5 CALCULATING ROUTE RELIABILITY

So far, the definition of the reliability indices of individual links is fairly simple; however, the problem becomes slightly more complicated when it comes to computing the reliability indices of a route consisting of a series of links. Namely, the lateness and earliness indices of route p, consisting of n links are:

$$r_L(p) = \frac{\bar{t}\,(p)}{t(p)_{1-\alpha/2}} = exp[½\cdot T_{log}(p) - z_{\alpha/2}\cdot\sqrt{T_{log}(p)}] \tag{15}$$

and

$$r_E(p) = \frac{t(p)_{\alpha/2}}{\bar{t}\,(p)} = exp[-½\cdot T_{log}(p) - z_{\alpha/2}\cdot\sqrt{T_{log}(p)}] \tag{16}$$

respectively, where the path travel time variation logarithm is

$$T_{log}(p) = ln\left(1 + \frac{var[t(p)]}{[\bar{t}\,(p)]^2}\right).$$

While the mean travel time of the path can be easily calculated, simply by adding up the mean travel times of the elements forming the path, such that

$$\bar{t}\,(p) = \sum_{i=1}^{n} \bar{t}\,(l_i) \tag{17}$$

the issue that arises is the calculation of the total travel time variance. As empirically demonstrated by Rakha and others (2006), the most accurate method for computing the travel time variance of a route is to compute the expected coefficient of variation as the conditional expectation over all realisations of the various links that make up the route and thus assume that the route's coefficient of variation is the mean coefficient of variation over all links, such that

$$var[t(p)] = \frac{[\bar{t}\,(p)]^2}{n^2}\cdot\left(\sum_{i=1}^{n}\frac{\sqrt{var[t(l_i)]}}{\bar{t}\,(l_i)}\right)^2 \tag{18}$$

The reason for the method being accurate is the fact that it seems to empirically compensate for the losses in accuracy caused by the existence of covariances between successive links on a path, resulting from phenomena such as "blocking back". Using equation (18), the route travel time variation logarithm becomes

$$T_{log}(p) = ln\left(1 + \frac{\frac{[\bar{t}\,(p)]^2}{n^2}\cdot\left(\sum_{i=1}^{n}\frac{\sqrt{var[t(l_i)]}}{\bar{t}\,(l_i)}\right)^2}{[\bar{t}\,(p)]^2}\right) \Leftrightarrow$$

$$T_{log}(p) = ln\left(1 + \left(\frac{1}{n}\cdot\sum_{i=1}^{n}\frac{\sqrt{var[t(l_i)]}}{\bar{t}\,(l_i)}\right)^2\right) \Leftrightarrow$$

$$T_{log}(p) = ln\left(1 + \left(\frac{1}{n}\cdot\sum_{i=1}^{n}\sqrt{\frac{var[t(l_i)]}{[\bar{t}\,(l_i)]^2}}\right)^2\right) \tag{19}$$

Reformulating equation (1) for the link travel time variation logarithm:

$$T_{log}(l) = ln\left(1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2}\right) \Leftrightarrow$$

$$exp[T_{log}(l)] = 1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2} \Leftrightarrow$$
$$exp[T_{log}(l)] - 1 = \frac{var[t(l)]}{[\bar{t}\,(l)]^2}.$$

Hence, equation (19) becomes:

$$T_{log}(p) = ln\left(1 + \left(\frac{1}{n} \cdot \sum_{i=1}^{n} \sqrt{exp[T_{log}(l_i)] - 1}\right)^2\right) \tag{20}$$

Substituting expression (20) into equations (15) and (16) enables the calculation of the earliness and lateness indices for route p.

A point that should be mentioned here is the fact that time-dependence is considered in equation (20), by calculating $T_{log}(l_i)$ for each link l_i of route p using the mean and variance values of the travel time at the time of arrival at l_i rather than at the start of the trip. Time-dependent link mean travel time and variance values are calculated, as was mentioned in Section 3, using the method by Kaparias (2008) and are then input into equation (20).

6 EXPERIMENTAL RESULTS

In order to assess the accuracy of the reliability measure developed here, a field experiment is carried out. This involves finding a route for each one of a number of different origin-destination pairs in a specified test network for given departure times and estimating the expected time of arrival (ETA), the earliness and lateness indices and hence the so-called reliable time of arrival (RTA), which represents a time window, during which the arrival time is expected to lie. A vehicle is then driven along each route and the actual arrival time is recorded and compared to the RTA.

The test network is the London Congestion Charging zone, which is approximately 7km long and 5km wide and covers most of the Central London area. Link speed data is provided through floating vehicles in the form of individual measurements for the main arteries of the network over a three-month period in 2006 and is then aggregated to derive link speed distributions. For the links for which data is not available, values are simulated based on the available values, following the procedure used in the study by Kaparias and others (2008). Weekdays are aggregated into 15-minute intervals and the mean and standard deviation is calculated for each interval. Using the formulae of Section 4 earliness and lateness values are computed for each link, using a confidence level of 90%.

Using ARIAdNE, successor of ICNavS (Kaparias and others, 2007b), a purpose-developed software tool for implementing the new dynamic routing algorithm (Chen and others, 2005a; 2005b; 2006; Kaparias and others, 2007a; 2007b; Kaparias, 2008), the fastest route is generated for each of 72 selected origin-destination pairs and their ETA and RTA are computed. For comparison purposes, the same origin-destination pair is input into a conventional navigation system and its ETA is recorded. Thus, the following times are recorded: departure time (DT), conventional ETA (CETA), ARIAdNE ETA (AETA), ARIAdNE earliest RTA (AERTA) and ARIAdNE latest RTA (ALRTA). A vehicle is then driven along each route and the actual time of arrival (ATA) is compared with the expected ones from ARIAdNE and from the conventional system.

From the arrival times, the following travel times are deduced: the conventional expected travel time (CETT = CETA – DT), the ARIAdNE expected travel time (AETT = AETA – DT), the ARIAdNE earliest and latest reliable travel times (AERTT = AERTA – DT and ALRTT = ALRTA – DT) and the actual travel time (ATT = ATA – DT). In order to make the derived travel times over all measurements comparable with each other, the fractions of those to the ATT are calculated. For each of the resulting ratios a distribution is obtained from the total of the runs of the experiment; by ranking the distribution and by dividing the rank of each value by the total number of runs, cumulative probabilities are obtained, which result in a cumulative distribution. Then, plotting the cumulative distributions of all four travel times in a single graph, a visual comparison between them becomes possible.

The complete table of measured arrival times and the resulting table of travel times for each individual run are not included in the paper due to their size. For more detailed descriptions of the experimental conduct procedure and analysis method, and a presentation of the complete table of results, the reader is referred to the study by Kaparias (2008). Here, a table displaying the mean, standard deviation and extreme values of the CETT/ATT, the AERTT/ATT, the AETT/ATT and the ALRTT/ATT distributions is shown (Table 1), as well as the cumulative distribution plots of those obtained by ranking them and dividing by the total number of observations (Figure 1).

[Table 1 goes here]

[Figure 1 goes here]

From the distributions it can be seen that the mean AETT/ATT value is approximately 0.95, which shows that, on average, the AETT is only 5% shorter than the ATT. As expected, the AERTT always underestimates the ATT by being on average approximately 47% shorter than it (average AERTT/ATT value of 0.53); however, in unexpectedly short trips without any unpredictable delays resulting in early arrivals, the ATT becomes close to the AERTT, with a maximum observed AERTT/ATT value of approximately 0.82 (i.e. the AERTT is 18% shorter than the ATT) in the experiment. On the other hand, as expected, the ALRTT always overestimates the ATT by being on average approximately 53% longer. However, in trips where a combination of unpredictable delays arise, the ALRTT nears the ATT; in the experiment, the lowest ALRTT/ATT value is exactly 1, which shows that in one observation, the ATT is equal to the ALRTT.

The above results suggest that ARIAdNE's output, employing the reliability measure introduced in this paper, is fairly accurate, as not only the AETT is a good estimate to the ATT, but also the AERTT and the ALRTT are accurate lower and upper bound estimates to it. It is important to note that the bounds are not exceeded in any run of the experiment. Another feature to note is the fact that the curve of the AETT/ATT cumulative distribution is S-shaped, indicating a high concentration of values around the mean rather than around the extremes (also indicated by the low standard deviation value). More specifically, the curve shows that around 75% of the values are equal to or lower than the ATT, while around 70% are +/-20% of it. As opposed to these observations the results obtained by the conventional system do not seem to exhibit the same level of accuracy. Namely, the mean CETT/ATT value is approximately 0.49, while the maximum and minimum values are 0.28 and 0.79. This implies that the CETT constantly underestimates the ATT, on average being 51% shorter than it. It is interesting to note that in none of the runs of the experiment did the CETT estimate the ATT correctly, with the closest observed value being around 79% of the ATT. The CETT/ATT distribution is thus significantly biased.

It could be argued, that the bias of the CETT/ATT distribution is a result of the conventional system using less accurate data than ARIAdNE. While ARIAdNE makes use of floating vehicle data records, reflecting actual measured speeds, the conventional system uses default estimates of travel times based on the speed limits provided in the map database. An approximate method to eliminate the bias would be to multiply the distribution by a specific factor, so as to shift the cumulative distribution curve to the right and bring it around the ATT line. Thus, by multiplying the CETT/ATT distribution by 2, the unbiased CETT/ATT (UCETT/ATT) distribution is obtained and is added to the plot (Figure 2).

[Figure 2 goes here]

With the bias eliminated, the UCETT/ATT distribution can be directly compared to the AETT/ATT distribution. The most important observation that can be made by examining the graph is that the AETT/ATT curve is S-shaped, as mentioned above, while the UCETT/ATT curve is more straight-shaped, with the exception of the upper extreme. This means that while the AETT/ATT distribution has a high concentration of values around the mean, this is not the case for the UCETT/ATT distribution, which again implies that the estimates calculated by ARIAdNE are more accurate than the ones calculated by the conventional system, suggesting that the reliability measure defined here is accurate.

7 CONCLUSIONS

In this paper, a new reliability measure, consisting of two reliability indices based on the log-normal distribution, was defined, meeting a number of imposed requirements, such as comprehensibility by the traveller and “dimensionlessness”. It was then expressed in terms of the distribution of speeds, such that it could be possible to evaluate it when the given data consists of speed measurements rather than travel time measurements. A method for calculating the reliability of a route using the reliabilities of the links forming the route was also derived. With the help of a field experiment, the measure was tested against real traffic conditions and the results suggested that it reflects the delays and early arrivals that may be encountered fairly accurately.

As reliability was found to be an important factor affecting a driver’s route choice, its incorporation into a car navigation system will offer significant advantages, not only to the individual user of the system, but also implicitly to all users of the road network. The proposed reliability indices will enable navigation systems to provide more accurate information about arrival times to the drivers, as well as to identify locations with potential adverse traffic conditions and avoid them. Besides the direct benefit of travel time savings that the new measure will offer to the users of navigation systems, its use will also reduce the total travel time lost in the entire road network, thus resulting in a decrease in fuel consumption, vehicle wear and tear and CO_2 emissions.

Future research will concentrate on further testing of the new measure in routing applications for car navigation systems. Together with ongoing research on routing algorithms, a prototype reliable vehicle navigation system is to be developed and tested further. More field trials are scheduled to be carried out, so that the advantages of the new system can be demonstrated and validated.

ACKNOWLEDGEMENTS

The authors would like to thank BMW AG for sponsoring this work, ITIS Holdings UK for supplying traffic data and PTV AG for supplying the test network for the field experiment. Also, special thanks are given to K. Zavitsas for his help during the conduct of the experiment.

REFERENCES

Adbel-Aty, A., Kitamura, R. and Jovanis, P. P. (1995). Investigating the effect of travel time variability on route choice using repeated measurement stated preference data, *Transportation Research Record*, **1493**, 39-45.

Al-Deek, H. and Emam, E. B. (2006). New methodology for estimating reliability in transportation networks with degraded link capacities, *Journal of Intelligent Transportation Systems*, **10**, 117-129.

Bates, J., Black, I., Fearon, J., Gilliam, C. and Porter, S. (2002). Supply models for use in modelling the variability of journey times on the highway network, AET - Proceedings of the European Transport Conference 2002, Cambridge, UK.

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

Bell, M. G. H. and Iida, Y. (1997). Network reliability, in: Transportation network analysis Chichester Wiley and Sons, 179-192.

Chen, K., Yu, L., Guo, J. and Wen, H. (2007). Characteristics analysis of road network reliability in Beijing based on the data logs from taxis, TRB 2007 Annual Meeting CD-ROM, Washington DC, USA.

Chen, Y., Bell, M. G. H. and Bogenberger, K. (2005a). Reliable pre-trip multi-path planning and dynamic adaptation for a centralized road navigation system, ITSC 2005 - 8th International IEEE Conference on Intelligent Transportation Systems, Vienna, Austria.

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

Chen, Y., Kaparias, I., Bell, M. G. H. and Bogenberger, K. (2005b). Reliable autonomous route guidance by a constrained A* search considering intersection delays, The Reliability of Traveling and the Robustness of Transport Systems, Delft, The Netherlands: TRAIL Research School, 69-92.

Dandy, G. C. and McBean, E. A. (1984). Variability of individual travel time components, *ASCE Journal of Transportation Engineering*, **110**, 340-357.

Eleftheriadou, L. and Cui, X. (2007). A framework for defining and estimating travel time reliability, TRB 2007 Annual Meeting CD-ROM, Washington DC, USA.

Herman, R. and Lam, T. (1974). Trip time characteristics of journeys to and from work, Proceedings 6th International Symposium on Transportation and Traffic Theory, Sydney, New South Wales, Australia: 57-85.

Jackson, W. B. and Jucker, J. V. (1981). An empirical study of travel time variability and travel choice behaviour, *Transportation Science*, **16**, 460-475.

Kaparias, I. (2008). Reliable dynamic in-vehicle navigation, *PhD Thesis*, Imperial College London.

Kaparias, I., Bell, M. G. H. and Belzner, H. (2008). Field testing of a reliable in-vehicle navigation algorithm, 10th International Conference on Applications of Advanced Technologies in Transportation, Athens, Greece.

Kaparias, I., Bell, M. G. H., Bogenberger, K. and Chen, Y. (2007a). An approach to time-dependence and reliability in dynamic route guidance, *Transportation Research Record*, **2039**, 32-41.

Kaparias, I., Bell, M. G. H., Chen, Y. and Bogenberger, K. (2007b). ICNavS: A tool of reliable dynamic route guidance, *IET Intelligent Transport Systems*, **1**, 225-233.

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

Liu, H. X., Recker, W. and Chen, A. (2004). Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data, *Transportation Research A*, **38**, 435-453.

Lomax, T., Schrank, D., Turner, S. and Margiotta, R. (2003). Selecting travel reliability measures, *Texas Transportation Institute*, 474360-1.

Lomax, T., Turner, S. and Margiotta, R. (2001). Monitoring urban roadways in 2000: Using archived operations data for reliability and mobility measurement, *Federal Highway Administration - Operations Core Business Unit*, FHWA-OP-02-029.

Mogridge, M. and Fry, S. (1984). The variability of car journeys on a particular route in central London, *Traffic Engineering and Control*, **25**, 510-511.

Montgomery, F. O. and May, A. D. (1987). Factors affecting travel times on urban radial routes, *Traffic Engineering and Control*, **28**, 452-458.

Noland, R. B. and Polak, J. W. (2002). Travel time variability: a review of theoretical and empirical issues, *Transport Reviews*, **22**, 39-54.

Noland, R. B. and Small, K. A. (1995). Travel-time uncertainty, departure time choice, and the cost of morning commutes., *Transportation Research Record*, **1493**, 150-158.

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

Rakha, H., El-Shawarby, I., Arafeh, M. and Dion, F. (2006). Estimating path travel-time reliability, ITSC 2006 - 9th International IEEE Conference on Intelligent Transportation Systems, Toronto, Canada: 236-241.

Rakha, H. and Wang, Z. (2005). Estimating traffic stream space mean speed and reliability from dual- and single-loop detectors, *Transportation Research Record*, **1925**, 38-47.

Sung, K., Bell, M. G. H., Seong, M. and Park, S. (2000). Shortest paths in a network with time-dependent flow speeds, *European Journal of Operational Research*, **121**, 32-39.

van Lint, J. W. C. and van Zuylen, H. J. (2005). Monitoring and predicting freeway travel time reliability - Using width and skew of day-to-day travel time distribution, *Transportation Research Record*, **1917**, 54-62.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers*, **1**, 325-378.

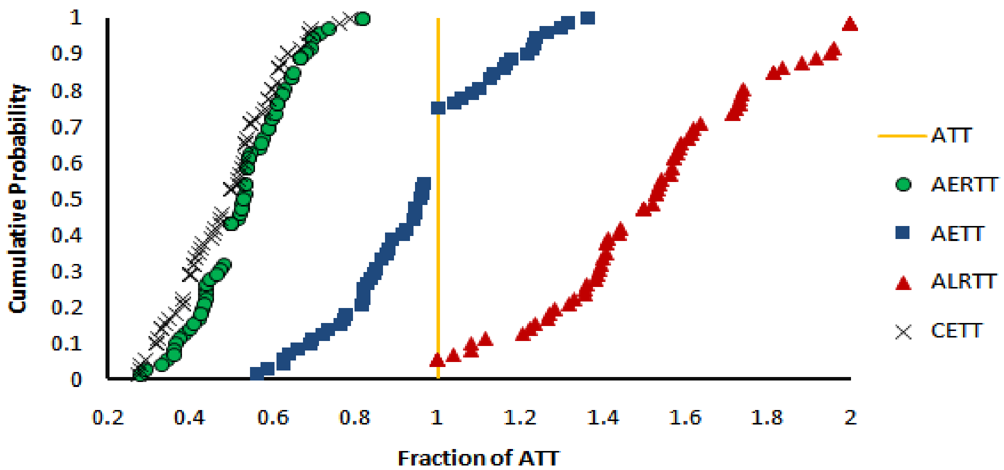

Figure 1: Cumulative distributions

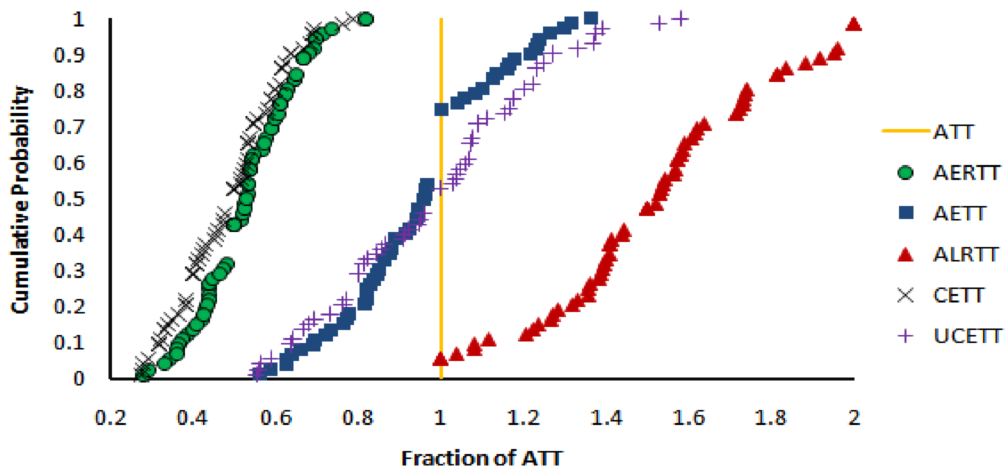

Figure 2: Cumulative distributions, including UCETT/ATT

Table 1: Mean, standard deviation and extreme values of travel time ratio distributions

	Conv.	**ARIAdNE**		
	CETT/ATT	**AERTT/ATT**	**AETT/ATT**	**ALRTT/ATT**
Mean	0.4922	0.5325	0.9492	1.5254
Std. Dev.	0.1228	0.1164	0.1815	0.2764
Minimum	0.2778	0.2813	0.5625	1.0000
Maximum	0.7895	0.8182	1.3636	2.0455