PERGAMON

Transport Policy 9 (2002) 117–127

TRANSPORT POLICY

www.elsevier.com/locate/tranpol

An introduction to road vulnerability: what has been done, is done and should be done

Katja Berdica*

Department of Infrastructure, Royal Institute of Technology, Fiskartorpsvagen 15A, 100 44 Stockholm, Sweden

Abstract

Vulnerability in the road transportation system, studied not only from a safety point of view but also as a problem of an insufficient level of service, is proposed as a setting for future transport studies. This relatively new notion is conceptualised by discussing a number of definitions and related concepts, reviewing especially the concept of reliability as a feasible theoretical approach. The paper relates how vulnerability related problems have been addressed so far, current developments and finally what the future should hold in order to provide us with the comprehensive network analysis tool that our complex society calls for. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Vulnerability; Reliability; Road network; Literature review

1. Introduction

To provide all citizens, trade and industry nationwide, with an economically and in the long-term sustainable transport function is more or less the overall objective of most national traffic policies of today—the Swedish one being no exception. In Sweden, six subsidiary goals have been formulated to obtain: (1) an *accessible* transportation system, (2) high *transport quality*, (3) *safety*, (4) a good *environment*, (5) a positive *regional development* and 6) *equality* (Prop 2001/02:20). The focus of this paper is on the road transportation system, for which the fulfilment of all of these points requires not only that physical links *exist* between different destinations, but also that these links are open to traffic. However, extreme weather conditions or other natural phenomena can make large parts of the road network impassable. More rare events like collapsing bridges, as well as every day incidents such as traffic accidents or closures due to maintenance activities, can reduce the capacity of the traffic system. The resulting congestion effects and delays cause serious losses in terms of travel time as well as other costs e.g. in case of disrupted ‘just in time’ deliveries. In order to implement suitable policies to attain various goals there is a need for an overall characterisation of road networks, to gain insight into their propensity to ‘malfunction’, the extent of the resulting consequences, the scope for mitigation measures, etc. So far, this is felt to be missing in transport studies of today. The purpose of this paper is therefore to conceptualise and put into context a relatively new notion, proposed to constitute a framework for such analyses: *vulnerability* in the road transportation system.

The starting point is the fact that road networks from time to time display a reduced level of service due to various reasons. A number of definitions and related concepts are used to fence in the general idea, going on to suggest a basis from which vulnerability issues can be addressed, as well as pointing out the possibilities and problems involved in practical implementation. The final discussion argues for a more systematic consideration of these issues throughout the road infrastructure planning process, all the way from investment planning, through operations management, to optimisation of maintenance activities.

2. How to define vulnerability

2.1. Serviceability versus accessibility

Vulnerability of the road transportation system is here regarded not from a safety point of view, but rather as a problem of reduced accessibility that occurs because of various reasons. Also, emphasis is put on the function of the system rather than the physical network itself, although some of the reasons for discontinuities in the road network are indeed caused by physical failures. Accessibility is however a term of widely varying meaning depending on the context in which it is used. Generally, accessibility is understood as ‘ease of reaching’ and concerns the

* Tel.: +46-8-790-94-16; fax: +46-8-790-67-61.
E-mail address: ktj@infra.kth.se (K. Berdica).

0967-070X/02/$ - see front matter
PII: S0967-070X(02)00011-2

opportunity provided by the transport system for people to take part in a particular activity/set of activities from a given location (Jones, 1981). In this sense accessibility can be increased either by increasing the number of routes to a certain service or by increasing the amount of services reachable on a certain route or within a certain time budget. Accessibility is, according to Jones, a function of mobility, meaning the ability of an individual to move about (i.e. a *potential* to move, *not* the completed movement that Ross (2000) uses in his arguments for the relationship being one of reciprocity) by means of private and/or public transport. Mobility in its turn consists of: (1) the performance/effectiveness of the transport system in connecting spatially separated locations, and (2) individual characteristics influencing the extent to which people are able to make use of the transport system.

This implies that even though accessibility depends on the degree to which the transportation system is functioning, it is actually approaching the issue from the *demand* side. This is of course important when dealing with the quantification of consequences, i.e. evaluating the total delays caused by various events. At this stage it is however more interesting to regard the road network from the *supply* side, meaning the actual existence of a functioning route from one location to another. Therefore, it is in this case better to describe the performance of the road transportation system in terms of serviceability. Definition: The *serviceability* of a link/route/road network describes the possibility to use that link/route/road network during a given time period. Another approach can be to talk about quality in transport, which is a multi-dimensional concept with many different attributes, all of which exist in the context of (quote): " the basic ability of a system to deliver you from where you are, to where you want to be, at the time you want to travel, at a cost…that makes the journey worthwhile" (Goodwin, 1992, p. 661). This is in fact, another more detailed way of describing the previously defined term serviceability.

2.2. Incidents

The events of interest in the case of vulnerability in the road transportation system are the ones causing disturbances in traffic. These events can be of a more or less sudden and/or unpredictable nature, ranging from extremely adverse weather, through physical failures and traffic accidents, to planned road works, as well as intentional harm in the shape of conflicts in labour relations or terrorist actions—all of which have bearing on the performance of the road network. The frequency, predictability, geographical extent, etc. vary greatly both among event categories and among events within categories. What these events have in common though, is that they could have a negative influence on the serviceability of the road network, either on their own or by starting a chain of events resulting in a disturbance. They are therefore referred to as incidents. Definition: An *incident* is an event, which directly or indirectly can result in considerable reductions or interruptions in the *serviceability* of a link/route/road network.

According to this, an incident is hence an event causing reduced capacity and/or increased demand. This is a somewhat more specific usage of the term, since in most dictionaries 'incident' is simply a synonym for 'event'. To distinguish between the two, it is suitable to note that studying vulnerability in the road transportation system is not primarily a question of network optimisation regarding the function during ordinary traffic flow variations or minor changes in weather. Therefore the specification 'considerable reductions or interruptions' is used in the definition earlier. The actual quantification of this level is an important task for future studies and is elaborated later.

2.3. Risk is probabilities and consequences

Risk is generally associated with something that entails negative consequences for life, health and/or the environment. The definitions of the term vary, but mostly involve a combination of two parts: (1) the probability for an event of negative impact to occur, and (2) the extent of the resulting consequences once this event has taken place. The evaluation of different influencing factors and the consequent rating of risk depend on the point of view taken—the decision maker, the affected user or other interested parties see things differently, i.e. they have different perceptions of risk. Subjects using a definition of risk in terms of consequences tend to give higher ratings than those that focus on probabilities, and the same tendency can be seen when comparing the perspectives 'risks to others/people in general/society' and 'personal risk' (Drottz-Sjöberg, 1991). There is also the aspect of time and space, in which risks 'here-today' tend to be perceived as greater than risks 'there-yesterday/tomorrow' (Westin and Jansson, 1994). Hence, measuring risk involves considerable uncertainty. Even if both probabilities and consequences are theoretically measurable in scientific terms, they in practice involve assessments impaired with various degrees of subjectivity that makes it more difficult to obtain an objective composite of the two.

The earlier mentioned incident categories, causing the road transportation system to malfunction, cover a wide range of combinations where probabilities and consequences are concerned: from minor accidents that happen every now and then, to highly improbable failures (e.g. of a bridge) resulting in serious injuries. This stresses the importance of a risk definition that takes both these aspects into account. Definition: *Risk* is a composite of (1) the probability for an *incident* to occur and (2) the resulting consequences, should the incident occur. One way to explain this is to construct a risk matrix, the principle of which is exemplified in Fig. 1. Of our earlier mentioned examples, the 'everyday minor accident' would be placed in the lower right-hand corner, while a bridge failure would fall into the

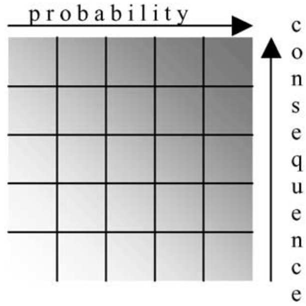

Fig. 1. Risk matrix: combinations of probability and consequence.

region on the top left-hand side. It is also worth noting that it is the diagonal from top left to bottom right that is the front line of interest; events of low probability/low consequence are not a problem, and incidents of high probability/high consequence tend to be taken care of very quickly—either by developing technologies or, as for the Zeppelin air-ships, simply abandoning an idea completely.

By treating risk in this manner, the expected consequences (in a probabilistic sense) could be one way to operationalise risk. However, the combination of probabilities and consequences makes the issue rather complicated and requires considerable effort. An important pilot study in this respect is found in 'Risk Assessment Methods in Road Network Evaluation' (Dalziell and Nicholson, 2001). Here probabilistic risk assessment is used to investigate the potential for, and resulting consequences of, road closures due to adverse weather, seismic activity, and traffic accidents in New Zealand.

2.4. Vulnerability

As is often the case when discussing terminology, there are many alternative ways of understanding the term vulnerability. In one paper the following definition is used (quote): "vulnerability is a susceptibility for rare, though big, risks, while the victims can hardly change the course of events and contribute little or nothing to recovery." (Laurentius, 1994, p. 278). From the road network point of view, this is a rather narrow definition, concentrating on events of catastrophic nature. A better approach is to start by stating that, a vulnerable system is a system susceptible to extreme strains. In a government bill (prop 1996/97:11), handling extreme strains on society is described as the ability to deal with situations characterised by: a deviation from the situation which is considered as 'normal'; sudden occurrence, more or less unpredicted and without warning; basic values being threatened; demand of quick decisions and coordinated/concentrated actions by several authorities. Here the low level of probability and high level of consequence, implied by 'rare, though big, risks' in the previous citation, is omitted. There is, however, an improper emphasis upon the unexpected, considering that in the case of road transports planned actions such as road works can cause serious

Fig. 2. Vulnerability in the road transportation system: wheel of concepts.

disturbances in traffic (Berdica, 1998). Therefore, in the specific case of the road network, a definition of vulnerability that is of a more general nature is suggested here. Definition: *Vulnerability* in the road transportation system is a susceptibility to *incidents* that can result in considerable reductions in road network *serviceability*. These incidents may then be more or less predictable, caused voluntarily or involuntarily, by man or nature. A similar definition, in principle, is found in Nicholson and Du (1994).

By addressing these terms in that order one can see a system of back tracking the issue at hand through a series of sequential definitions, illustrated by the Wheel of Concepts in Fig. 2. Reducing vulnerability can hence be regarded as reducing the risks involved in various incidents, which can be brought about in two ways: (1) by a *fail-safe* approach, meaning that the probability for e.g. a bridge to fail is reduced, or (2) by adopting a *safe-fail* perspective, implying a reduction of the resulting consequences should the failure occur (Holling, 1981). In this example the latter could be achieved by an emergency ferry service. Consequence minimisation is in fact a very important aspect of vulnerability studies, since it is not always suitable to talk about the probabilities of certain incidents, e.g. terrorist actions. In those cases it is necessary to investigate expected effects and possible remedial actions under the assumption that the incident has already taken place.

2.5. 'Neighbouring' terms

There are several other terms that can be regarded as related to the general subject of vulnerability in technical systems that may or may not be relevant in the specific case of vulnerability in the road transportation system. The most common of these are discussed shortly here, mainly in relation to the earlier stated definitions. However, *reliability*—the term that probably first comes to mind—is found to be of special importance and is therefore elaborated later.

In dictionaries, *robustness* is often exemplified in the area of computer technology, i.e. it is the ability of a computer system to cope with errors during execution. In other words, it is the ability of a system to withstand strain. It is easy to draw a parallel and talk about robustness in the road transportation system, as opposed to vulnerability that is defined as a susceptibility to disturbing incidents. Hence the terms can be used as oppositely interchangeable in the present context, i.e. a less vulnerable system can be regarded as more robust and vice versa.

Resilience is often an issue in ecology and expresses the capability of an ecosystem to 'return to normal' after having been disturbed. It is hence a question of stability, involving the factors: (1) maximum disturbance from which the system can recover and (2) speed of recovery (Goldberg, 1975). It is clearly possible to transfer this concept to the road transportation system, with, e.g. 'time to restored serviceability' being a factor of interest. Resilience could also be described as the capability of reaching a new state of equilibrium. However, many of the incidents causing reductions in serviceability have a relatively short duration, never reaching a new equilibrium. This so called transient state still remains to be studied/modelled more thoroughly and recent research shows that microsimulation can be a feasible tool for such analyses (Berdica et al., 2001).

In electronics/mechanics there is often *redundancy* in the system, meaning a duplication of components so that operations can continue even though part of the equipment fails. The parallel to the road transportation system is obvious: the existence of numerous optional routes/means of transport between origins and destinations can result in less serious consequences in case of a disturbance in some part of the system. This is an important factor in the evaluation of vulnerability in a road transportation system. If traffic is routinely divided between possible routes in everyday use, the redundancy is *active*. A ferry service that is activated only in case of a bridge failure (i.e. a 'reserve') is an example of *passive* redundancy. It is however necessary to remember that the redundancy of the system is dependent upon the nature of the 'threat'. In the electronics case, putting double cables in the same canal is not enough if the canal itself is destroyed e.g. by a bomb assault. In the road transportation system, a serious snowstorm may well disable all alternative routes in a larger area.

3. How to approach vulnerability

3.1. Introducing reliability

So far, there has not been much research on road transportation vulnerability as an explicit subject, although the need for more comprehensive vulnerability analyses obviously grows concurrently with the increasing complexity of modern infrastructures. One fundamental reason could be a want of a well-tried theoretical basis to build on. Let us therefore at this point turn to the concept of *reliability*, in an attempt to close in on the topic of vulnerability. To start with, consider the following, widely accepted definition: 'Reliability is the probability of a device performing its purpose adequately for the period of time intended under the operating conditions encountered.' (Billington and Allan, 1992, p. 6; Wakabayashi and Iida, 1992, p. 30). This leads to the conclusion that reliability studies are generally concerned with probabilities only, which in fact often is the case. D'Este and Taylor, (2001) choose to regard vulnerability as completely separated from reliability due to the mainly probabilistic base of the latter and (quote): "the concept of vulnerability is related to the consequences of link failure, irrespective of the probability of failure". If so, vulnerability would be just a measure of consequences instead of the comprehensive framework for analysis that will be argued here, and in which the composite risk concept is a central idea. However, the formulation of reliability earlier does not necessarily exclude consequences: 'performing its purpose adequately' and 'operating conditions encountered' can well pertain to this aspect. The consequences can hence be included in terms of a *level of performance*. The performance of engineering systems can be expressed by a number of various indices, e.g. expected number of failures within a certain time, average time between failures, expected loss of revenue and/or output, etc. These are more and more commonly referred to simply as 'reliability indices' even though they are not associated with probabilities only, as in the classic case of reliability studies (Billington and Allan, 1992). Reliability theory can well be used for the evaluation of these indices, provided that levels for the unacceptable/inadequate are set first. By choosing this wider interpretation, reliability can be regarded as a complement of vulnerability, expressed in the following. Definition: *vulnerability* in the road transportation system is *reliability*, meaning adequate *serviceability* under the operating conditions encountered during a given time period.

The conclusion is that vulnerability in transportation systems has not yet been considered directly but well-established reliability theory can constitute a tool for addressing such issues, provided that consequences as well as probabilities are taken into consideration. In transportation systems, reliability describes the possibility of successfully travelling from one place to another. So far in the literature, it has been considered in three main aspects: (1) reliability of connectivity (also referred to as *terminal reliability*), meaning the probability of at all reaching a chosen destination, (2) *reliability of travel time*, meaning the probability of reaching a chosen destination within a given time, and (3) *capacity reliability*, meaning the probability of the network being able to 'swallow' a certain amount of traffic. The main difference from traditional systems reliability analysis is the need to identify an acceptable level of service, exchanging the dichotomous states 'functioning–not functioning' with 'serviceability acceptable–not acceptable'. Furthermore, changes in road network reliability can be observed as a result of either fluctuation in traffic *flow* ('normal' conditions;

often referred to as *recurrent* congestion), or fluctuation in *capacity* ('abnormal' conditions; often referred to as *non-recurrent* congestion).

In the following, what has been done in terms of research conducted so far is reviewed, giving account of essential lines of thought pertaining to these matters (see also Berdica, 2000; Schmöcker and Nicholson, 2001).

3.2. Network reliability in general

Wakabayashi and Iida (1992) and Bell and Iida (1997) focus on terminal reliability. The stochastic variable x_i represents the state of, e.g. congestion on link i, with the value of 1 if link i functions and 0 otherwise. The system reliability value is found through the structure function $\varphi(\mathbf{x})$, where $\mathbf{x}$ is the system state vector containing the link state variables. Accordingly, the structure function has the value 1 if the system functions and 0 otherwise. For systems that have—or easily can be broken down into—a series and/or parallel configuration, the structure function is simply $\varphi(\mathbf{x}) = \prod_i x_i$ for a series system and $\varphi(\mathbf{x}) = 1 - \prod_i (1 - x_i)$ for a parallel system, or a combination of these. For more complicated systems the structure function is derived by different methods, of which 'path-and-cut' is the most practical. The system reliability is calculated directly from either: (1) the minimal routes, which is the minimum number of successive links needed to connect a pair of nodes, regarded as series systems in parallel—as long as any one of the routes functions, the system functions; or (2) the minimal cut sets, which is the minimum number of links needed to disconnect a pair of nodes, regarded as parallel systems in series—if any one of the cut sets fails, the system fails.

Link reliability r_i is the expected value of the (assumed) random binary variable x_i and system reliability R is the expected value of the structure function. However, the evaluation of R involves complicated calculations since links can appear in more than one set. Furthermore, large or very complicated networks result in a great amount of calculation. Therefore various heuristic methods for obtaining estimated values of reliability, for instance upper and lower bounds (some using partial route/cut sets), are most interesting. Some other 'problem areas' are identified, e.g. that connectivity reliability studies overlook mutual relationships or dependencies among links, as well as the fact that connectivity reliability does not take flow on links into account.

3.3. Reliability and fluctuation of traffic flow

The main issue in Asakura and Kashiwadani (1991) is a modified traffic assignment model for simulating day-to-day fluctuations of traffic flow during 'ordinary' traffic conditions, that is when capacity is not reduced by road works, accidents, natural disaster, etc. The basic assumption is that travel demand in origin–destination (OD) relations fluctuates stochastically from day-to-day around its (observed/estimated) mean value and that this random variation can be described by two mutually independent, normally distributed random variables. One represents OD-specific variation (e.g. special offer at a certain mall) while the other represents variation that affects the network as a whole (e.g. adverse weather). A correlation parameter of fluctuation (λ) decides how much of each type is included in the total variation. It is easy to show that the correlation coefficient (ρ) for two different OD-demands (on the same day) is determined by this correlation parameter, with the special cases of (1) demand fluctuation being the same for every OD-pair when ($\lambda = 0 \Rightarrow \rho = 1$); and (2) completely independent fluctuating OD-demand when ($\lambda = 1 \Rightarrow \rho = 0$).

The results of the simulation model are used to define and estimate two reliability measures. The first is a connection measure (i.e. connectivity), i.e. the probability of travel in an OD-relation without encountering congestion beyond a certain level. If this level of congestion (which can be measured in terms of traffic volume, travel time, generalised travel cost or the like) is exceeded, the link in question is regarded as disconnected. Link congestion fluctuates in proportion to link traffic volume, which has been determined by the simulation model. If the different routes can be regarded as parallel, the probability of an OD-pair being connected is calculated in accordance with the theory presented earlier (Wakabayashi and Iida, 1992; Bell and Iida, 1997). By defining connectivity in terms of *congestion level*, this connection measure describes the probability of finding at least one route without unacceptable congestion between i and j. The second is a time reliability measure, as either (a) the probability of travel time in an OD-relation not exceeding the prescribed travel time, or (b) the travel time bound in an OD-relation for which the probability of not being exceeded is at a prescribed level. Because of the assumed stochastic variation in traffic volume, travel time between OD-pairs fluctuates randomly. This randomness is assumed to follow a normal distribution for which the parameters can be calibrated using the results of the simulations.

It is interesting to note that Asakura (1996) makes a connection between the two reliability measures discussed earlier. In the case of deterioration in a road network with variable flows, travel time reliability is expressed as a function of the ratios between travel times at reduced capacity and those during 'normal' conditions. When the level of degradation is so great that the destination in question cannot be reached, the ratio move towards infinity. This extreme case of travel time reliability is then consistent with the reliability of connectivity concept.

3.4. Improvement of reliability under traffic management

Wakabayashi and Iida (1993) take more or less the same approach as Asakura and Kashiwadani (1991). They propose new indicators of road network performance level, instead of the conventional quantitative (e.g. total length of road per unit area) or static (e.g. average travel time) indicators of road network quality: (1) terminal reliability is the probability that two given nodes are connected with a certain service level for a given time period, and (2) travel time reliability is the probability that travel time between two given nodes will not exceed a given travel time, alternatively treated as the maximum travel time to arrive at a destination with given probability. Traffic variation is assumed to be the main factor influencing both of the above. The difference, however, is the 'abnormal conditions' approach, described in the statement that (quote): "A highly reliable road network provides sure and unfluctuating traffic service by offering drivers alternative routes even when some part of the network is unavailable due to traffic accidents, maintenance or natural disaster." (Wakabayashi and Iida, 1993, p. 29).

Terminal reliability is addressed at link level, and link reliability of connectivity is the probability that demand flow does not exceed reference capacity on a certain link for given time periods. The variation in link flow originates from the variation in OD-flow, following the two assumptions that these are normally distributed and that OD-portions using a certain link are constant. This is explained in a functional model, expressing the coefficients of variation. Link reliability of travel time can then be converted from that of connectivity by derivation of the probability density function for demand traffic volume and using for instance the BPR function as link performance function. Estimations can then be made in two ways, by assuming either independent or correlated link flow variations. By considering traffic flows explicitly, the method proposed can be used to assess the effects of a change in traffic flow on reliability.

3.5. Degradable transportation systems

Nicholson and Du introduce the term 'degradable transportation system' (DTS), since (quote): "transportation systems are subject to degradation as a result of a wide variety of events (e.g. earthquakes, floods, traffic accidents, adverse weather, industrial action, and inadequate maintenance)" (Nicholson and Du, 1997, p. 209). Existing network reliability models are lacking on a number of accounts when considering transportation networks. Most importantly they assume fixed traffic demand, statistically independent component states, and they do not allow for components to have degradation levels between full and zero capacity.

The proposed integrated equilibrium model (Nicholson and Du, 1997), on the other hand, allows for elastic demand, varying levels of degradation, and interdependent component states. Also, it focuses on long duration capacity variations, since there is more scope for traffic to move towards a new equilibrium situation. The network configuration chosen for the DTS is multi-modal, as different transport modes are not necessarily affected equally by an incident. Each link pertains to a single mode, hence there is no modelling of interactions between different types of vehicles. Moreover, this means that route choice also implies mode choice. Both links and nodes are system components of the DTS, although all component degradation is considered to occur within links.

The relationship between a given component state (i.e. link capacity) vector and its corresponding system state vector is defined using a combined model which solves the problems of traffic generation, distribution, modal split and assignment simultaneously within a framework based on the assumption of demand-supply equilibrium. It is assumed that demand in each OD-pair can be formulated explicitly as a function of the generalised travel cost. Various forms of demand function, including the logit, power, exponential, and elastic exponential functions, satisfy these assumptions. The supply function is a multi-variable function and is therefore represented implicitly through the following. Route flows are determined from OD-flows by assuming that individuals choose routes in order to minimise their generalised costs under the usual Wardrop condition on user equilibrium. The travel time on a link is a function of the vehicle flow and the component state, which in turn influences the generalised travel cost on that link. System surplus is then chosen as the performance measure to assess the socio-economic impacts of system degradation. The authors refer to previous work that has shown that the equilibrium OD-flow and link flow vectors, and system surplus, are unique—although the equilibrium route flow vector generally is not (which is also the case in the standard traffic assignment problem). This uniqueness is important when analysing how the system state and thereby its performance is affected by changes in component state, i.e. component degradation. The sensitivity and reliability analysis based on the integrated equilibrium model (Du and Nicholson, 1997)—and hence the model itself—are for a steady-state condition. A future task is to include differences in degradation duration (which as well as degradation occurrence can be regarded as a stochastic variable), since the socio-economic impact of component degradation clearly grows with increasing time for repair/replacement.

3.6. Capacity related reliability

Allowing for varying demand and fluctuating capacity, initiated with the concept of Degradable Transportation Systems, has since been taken up and developed further. Recently the measure of capacity reliability, meaning the probability that a certain traffic demand can be accommodated at an acceptable level of service, has been introduced. It starts with the concept of network reserve capacity, defined as (quote) "the largest multiplier applied to an existing OD-demand matrix that can be allocated to a transportation network without violating the link capacities or exceeding a prespecified volume to capacity ratio (level of service)" (Chen et al., 1999, p. 186). Roadway capacity is modelled as a random variable with defined distributions.

The issue of interest is then to estimate the probability of network reserve capacity being greater than, or equal to, required demand for different levels of degradation. However, a common multiplier is used for all OD-pairs to increase travel demand, which is a restriction of this approach. It is also necessary to consider link dependencies when generating capacity degradation for each link.

Chen et al. (1999) state that travel time reliability is obtained as a 'side product' when solving the traffic equilibrium problem to obtain the maximum network reserve capacity. In later work they investigate how the estimation of capacity reliability and travel time reliability is affected by the degree of stochasticity in the route choice model (Chen et al., 2000). Yang et al. (2000) then proposed the combination of travel time and capacity reliability as a comprehensive performance measure of a road network, instead of modelling them separately as has been the case so far. They show that the two reliabilities are closely linked by two reference parameters, namely a travel time threshold and a traffic demand threshold. For a certain demand, the only way to obtain higher capacity reliability is through a higher travel time threshold—i.e. reducing the level of service—while for a given travel time threshold, travel demand needs to be lowered in order to increase capacity reliability.

4. How to study vulnerability

4.1. What has been done

Vulnerability problems have so far mostly been addressed in terms of isolated in-depth-studies of effects of either individual emergency situations or of separate vulnerability related issues. Disturbances in city traffic have often been in focus—their occurrence, type, and consequences in terms of delay. Studies may vary from pure field studies (Kronborg, 1993; Svensson, 1994) to historical data analyses (Golob et al., 1987; Giuliano, 1989). An example of a theoretical extension is found in Ardekani et al. (1997), in which a PC-based tool for identifying diversion routes (based on changes in traffic volume) around roadway disruptions is calibrated against actual observations in the field. Another frequent type of vulnerability related study is 'after-the-emergency' analyses (Savas, 1973; Wolffram, 1981; Romero and Adams, 1995; Swedish National Road Administration, 1999), often going on to more organisational aspects, issues of information and co-ordination of involved parties, etc. Answering these questions is a good beginning, but the piecing together into an overall consideration of vulnerability in the road transportation system is missing.

One factor contributing to this tendency for a piecemeal approach could simply be the difficulty of defining system type—an important issue in starting any systems analysis. Using the terminology in Billington and Allan (1992), *mission oriented* systems accept failure of some components as long as the system as a whole continues to function for the duration of the mission. Reliability can be said to measure to which extent this system is *staying in the operating state*. In systems of the *continuously operating* type, failure is acceptable provided that it is not too frequent or lasts too long. In this case *availability* is a better measure, describing the possibility of *finding the system operating*. However, when working with the road transportation system determining system type is not a straightforward task—it varies with the level of focus. At the level of individual travel demand, the system is mission oriented in that some links/routes may fail but alternative links/routes ensure the successful completion of a trip e.g. to work. At the aggregate level, on the other hand, the system is continuously operating. That is, interruptions in some parts of the network are accepted temporarily (e.g. during a heavy snow storm) if an 'emergency network' provides adequate serviceability at the given time. The road transportation system can hence be described as a continuously operating system consisting of a number of mission oriented subsystems, and functionality is in a way expected at both levels simultaneously.

This brings up the question of how to determine whether the system is operating or not—a question that can have a range of answers, partly because of the duality discussed earlier but also due to geographical setting, objectives of the study in question and/or parties involved, etc. Table 1 presents a selection of performance measures for multimodal transportation systems (Pratt and Lomax, 1996) often used for describing base conditions, identifying problems, aiding in the selection between alternative improvements, evaluating the outcome of various implementations, system monitoring, etc. Travel time is identified as a common thread, in that it exists not only as a direct measure but also as an element of other indicators. In the Highway Capacity Manual, speed is chosen as a simple indicator for the level of service. Speed is directly influenced by traffic flow in that low volumes permit high, steady speed while high volumes cause low, varying speed. In other words, traffic volume should be kept at a lower level than the capacity of the street in order to retain high transport quality. The difference between the two (preferably expressed in terms of actual traffic volume as a proportion of theoretical capacity) is a measure of quality denoted by Goodwin (1992) as the 'quality margin'. In theory, the larger the quality margin the less sensitive the system is to disturbing incidents (although a large quality margin seems to be a rather unrealistic way of reducing system vulnerability since in practice the traffic system is often operated close to capacity). These indicators can measure transport system performance, but vulnerability analysis demands that a level for the acceptable/unacceptable is set first. Also, these measures are static mean values but used to describe a highly dynamic process—traffic. Thus, there is clearly a dimension missing when it comes to discussing what adequate serviceability is and if it is maintained.

Table 1
A selection of performance measures for transportation systems (Source: Pratt and Lomax (1996); author's adaptation)

Measure	Calculated by	Comment
Travel time, tt		att = actual travel time, dtt = desired travel time
Travel rate, tr	tt/segment length = $(\text{speed}^{-1}) - 1$	atr = actual travel rate, dtr = desired travel rate Easier to use than speed in statistical analyses
Delay rate = dr	(atr − dtr = att − dtt)/length	Illustrates the intensity of the congestion problem to travellers
Total delay	dr × people vol. × length = (att − dtt) × people vol.	Illustrates impact of improvements on transportation systems
Relative delay rate	dr/dtr = atr/dtr − 1	Dimensionless measure for comparing systems operation to a *standard*
Delay ratio	dr/atr = 1 − dtr/atr	Dimensionless measure to illustrate the magnitude of the mobility problem in relation to *actual conditions*
Speed of person movement, spm	Passenger vol. × average travel speed	measure of travel efficiency
Corridor mobility index	spm/standard value	Standard value = e.g. one freeway lane operating at capacity with a typical urban vehicle occupancy rate
Accessibility	Average tt to objectives *or* percentage of objectives reachable within a specified time	Specially for joint performance of transportation and land-use
Congested travel	∑(congested segment length × people vol.)	

4.2. What is done

Vulnerability, exposure, and criticality in various infrastructures are issues that have been more explicitly looked into in recent years—unfortunately so, one might add, when considering the underlying reasons for this increased awareness. The United States Commission on Critical Infrastructure Protection (PDD-63, 1998) and the National Contingency Planning Group in Canada (NCPG, 2000) are only two among several examples. After an initial tendency to focus on mainly computers, telecommunications, and power supply (Haesken et al. 1997; ÖCB, 2000) the area of transport has received increasing attention. Concurrently, the importance of reliability studies—previously mostly associated with operations and management of more technological engineering systems—has also been more explicitly recognised in the transport sector during the 90s, although transport reliability as such has been of interest for a long time (Garrison, 1960). Today, network reliability in transport modelling is an important and growing field of research (Lam, 1999).

Earlier transport reliability studies introduced travel time reliability, terminal (connectivity) reliability and capacity reliability as main performance measures for road networks. Later developments include *travel demand satisfaction reliability* (Lam and Zhang, 2000), described as the probability that the network can satisfy a 'latent travel demand' caused by so called 'latent demand events' such as peak-period traffic or sporting events. It is stated that network reliability can be wrongly estimated if variations due to such temporary additional demands are not taken into account. Another contribution is *encountered reliability* (Bell and Schmöcker, 2001), expressed as the probability of not encountering any degraded links on the least expected cost path between origin and destination. Since users will probably try to avoid using degraded links, the level of information and the users' reaction to this information are important considerations. Apparently, the more 'risk averse' users are, the closer encountered reliability is to terminal reliability—given that link degradation probabilities are known to, and followed by, the users. Another way to model this is the game theory approach proposed by Bell (2000), in which the user seeks a minimum expected trip cost path while an 'evil entity' tries to maximise the expected trip cost by degrading a critical link in the network. The Nash-equilibrium point (at which neither user nor assailant can 'improve' their situation) provides worst-case link failure probabilities that can then be used to calculate other reliability measures.

Fluctuation of flow and fluctuation of capacity are generally identified as important factors when it comes to maintaining an acceptable level of performance, since reduced supply (capacity) on one link often results in increased demand (flow) on another. The interaction between demand and supply depends very much on individual road users' perceptions, e.g. of congestion levels and expected travel times, and also what aspects of travel time they regard as important. In order to correctly assess road network performance and decide what should be regarded as adequate levels of service—as well as to estimate the effects of possible mitigation measures—increased knowledge of these behavioural aspects of transport modelling is needed. Noland et al. (1998) present a simulation model that determines congestion effects of travel time uncertainties resulting from capacity reducing incidents such as traffic accidents, etc. They found that the extra travel costs resulting from increased travel times were in fact not as great as those connected to an increasing probability of arriving at the 'wrong' time, so called scheduling costs. Noland (1997) suggests that instead of increasing capacity (an often very capital-intensive solution) it may be more effective to reduce expected costs of travel by reducing travel time uncertainty e.g. by providing information on travel time variance and changing congestion conditions.

How users perceive and respond to travel time variability (generally in terms of departure time choice) has been the subject for both theoretical model development and empirical analysis in the past several years, a comprehensive overview of which was published recently (Noland and Polak, 2002). Other research deals more directly with how drivers react to information of varying detail and accuracy, given to them at various stages of their journey—Mahmassani (1990), Mahmassani and Jayakrishnan (1991) and Mahmassani and Chen (1993) are only some examples. This is all the more important considering recent developments in information technology and that advanced traveller's information systems (ATIS) are often exemplified as possible alleviators of congestion problems. While spatial and environmental constraints often make it impossible to solve capacity shortages by increasing capacity, AITS could allow us to use existing capacity more efficiently. Also, real-time information systems should enable quicker and more effective rerouting of traffic e.g. in case of blockages. However, it is important to recognise and further look into the fact that information will not automatically lead to improved traffic conditions (Mahmassani and Jayakrishnan, 1991) and if all drivers are given and also act on this information, conditions could actually worsen (Mahmassani and Chen, 1993). Another important issue in the development of behavioural models is the lack of actual observations of behaviour and corresponding measurements of traffic conditions. One solution that has been tried is conducting interactive laboratory-like experiments (Mahmassani, 1990).

4.3. What should be done

It is plain that road network reliability is on the rise as an important issue in society today. Three main areas of research are identified in Iida (1999): the development of a modelling framework capable of reliability investigations, the establishment of a traffic management system capable of providing high road performance, and new procedures for evaluation and optimisation of road network planning, construction and management which take reliability issues into account. Now, this seems to be comprehensive enough, so why not just leave the notion of vulnerability and concentrate on reliability instead?

It is clear that road vulnerability encompasses all of the above, but we also feel that vulnerability in road transport concerns more than this. Bearing the previous definitions and discussions in mind, road vulnerability analysis regards the road network as a whole and involves identifying a spectrum of incidents, collecting data on probabilities and consequences to estimate risks, performing various studies and experiments to set values for desirable/acceptable serviceability, as well as investigating and assessing the effects of possible mitigation measures and improvement strategies. We admit that the reliability concept could probably be stretched and modified to include these steps as well but it is always hazardous to fill old words with new meanings. Vulnerability as such is of course not completely new to our vocabulary either, but at least of such novelty in the area of transport that we feel comfortable with suggesting vulnerability analysis of road networks as an all-embracing framework using the different aspects of transport reliability studies as its foremost tools.

Vulnerability analysis is not introduced as a downright measure but more as a way of characterising a road network, basically as a means for being 'wise after the event in advance'. As such it should be integrated in all the different stages of the infrastructure process, adding a valuable dimension in investment planning, optimisation of road maintenance, and daily operations management. A new road may be built in a different corridor when taking overall vulnerability aspects into account; including risks and consequences of different incidents in a network could indicate a different order of maintenance prioritisation; vulnerability analysis would aid in recommending optional routes in case of road works or real time incident management, as well as for constructing contingency plans or for other emergency related issues. It can also be of help in assessing possible outcomes and overall effects of various strategic policies to avoid so-called problem migration (Hass-Klau et al., 1998). A main point is to identify possible vulnerability related problems at the right (which often means the earliest possible) level, since proactive measures are often preferable to reactive ones from an economic point of view.

Many question marks still need to be straightened out before reaching the structure for implementation of this comprehensive network analysis method. One main issue is that common transport modelling tools are often based on equilibrium assumptions on a macroscopic level. This is not suitable for detailed modelling of traffic under extraordinary conditions such as road closures, and continued development of tools e.g. microsimulation with these issues in mind is important (Goodwin, 1999; Berdica et al., 2001; Nicholson et al., 2001). Data on actual redistribution of traffic during disturbances is needed for the calibration and validation of these models, but such information is scarce today. An increased knowledge of how users—both private and business—react to and evaluate unreliability is of utmost importance, both for developing theoretical models and for incorporating these impacts into cost-benefit analyses (Small et al., 1995; Koorey and Mitchell, 2000).

Different aspects of these issues are already being addressed in different research projects and will undoubtedly continue to be of interest in the near future. Still, there is always a risk of running into a catch 22 situation: if no one expresses a demand for this kind of analysis tools, further research and development in this area is eventually pointless—and if a sound theoretical and methodological base is not established, no one will find any use for this kind of analysis tools. We suggest that vulnerability in the road transportation system should be brought out and recognised as a crucial part in the infrastructure sector, and that this over-all notion be the meeting point for all the different strands of transport reliability research and other related issues. A first step has already been taken by the Swedish government in their statement that reliable transports are an important aspect in the national policy goal of providing good quality transports (Prop 2001/02:20). Similar ideas are detectable in the British Government's transport strategy (Goodwin, 1999).

5. Conclusion

Road vulnerability as such has not been in focus for long, despite the fundamental importance of our road networks—not only as a direct means of 'normal' transport but also in assisting/maintaining/repairing other infrastructures and for evacuation in case of disaster. The road transportation system is hence one of our most important lifelines. Its wide spread and lack of supervision makes it difficult to manage, but at the same time lends it a certain robustness. Vulnerability analysis of road networks is regarded as a framework using the different aspects of transport reliability studies as its foremost tools. Road vulnerability analysis could thereby be regarded as a hub for the whole battery of transport studies needed to gain the insights necessary to describe how well our transport systems work in different respects, what steps to take and what policies to implement in order to reach desired goals. This paper gives the notion of vulnerability a more coherent structure, developing it not so much as a quantitative measure in itself, but rather as a way of thinking. We believe that adopting this perspective all through the planning, construction and management stages of the road infrastructure process will contribute to the public economy and welfare in society as a whole. The achievement of this, however, calls for cooperation on all levels, between all parties involved. This paper is an attempt to form a starting point that introduces concepts and views necessary to facilitate such a dialogue.

Acknowledgements

We would like to thank the Swedish Agency for Innovation Systems (former Swedish Transport and Communications Research Board) and the Swedish Agency for Civil Emergency Planning for their financial support.

References

Ardekani, S.A., Kazmi, A.M., Ahmadi, M.S., 1997. A PC-based tool for roadway incident management. Proceedings of the Conference on Traffic Congestion and Traffic Safety in the 21st Century, Chicago, June 8–11, 1997, pp. 361–367.

Asakura, Y., 1996. Reliability measures of an origin and destination pair in a deteriorated road network with variable flows. Proceedings of the fourth Meeting of the EURO Working Group in Transportation.

Asakura, Y., Kashiwadani, M., 1991. Road network reliability caused by daily fluctuation of traffic flow. Proceedings of the 19th PTRC Summer Annual Meeting (Seminar G), Brighton, pp. 73–84.

Bell, M.G.H., 2000. A game theory approach to measuring the performance reliability of transportation networks. Transportation Research—B 34 (6), 533–545.

Bell, M.G.H., Iida, Y., 1997. Transportation Network Analysis. Wiley, Chichester, West Sussex.

Bell, M.G.H., Schmöcker, J.D., 2001. Estimating the terminal reliability of degradable transport networks, TRISTAN (IV), Azores.

Berdica, K., 1998. Vulnerability in the Road Transportation System—Overview and Availability of Historical Data (in Swedish). TRITA-IP AR 98-70, Department of Infrastructure and Planning, Royal Institute of Technology, Stockholm.

Berdica, K., 2000. Analysing Vulnerability in the Road Transportation System—Putting Theory into Practice Using Sweden as an Example. TRITA-IP FR 00-76, Department of Infrastructure and Planning, Royal Institute of Technology, Stockholm.

Berdica, K., Andjic, Z., Nicholson, A.J., 2001. Simulation Road Traffic Interruptions—Does It Matter What Model We Use? Proceedings of First International Symposium on Transport Network Reliability, Kyoto University, July/August 2001.

Billington, R., Allan, R.N., 1992. Reliability Evaluation of Engineering Systems, Concepts and Techniques. Plenum Press, New York.

Chen, A., Yang, H., Lo, H.K., Tang, W.H., 1999. Capacity related reliability for transportation networks. Journal of Advanced Transportation 33 (2), 183–200.

Chen, A., Tatineni, M., Lee, D.-H., Yang, H., 2000. The effect of the route choice models on estimating network capacity reliability. Presented at Transportation Research Board 79th Annual Meeting, Washington, DC, January 2000.

Dalziell, E.P., Nicholson, A.J., 2001. Risk and impact of natural hazards on a road network. ASCE Journal of Transportation Engineering 127 (2), 159–166.

D'Este, G.M., Taylor, M.A.P., 2001. Network vulnerability: an issue for regional, national and international strategic transport networks. Proceedings of first International Symposium on Transport Network Reliability, Kyoto University, July/August 2001.

Drottz-Sjöberg, B.-M., 1991. Perception of risk, studies of risk and hazards 1/1991, Centre for Risk Research, Stockholm.

Du, Z.P., Nicholson, A.J., 1997. Degradable transportation systems: sensitivity and reliability analysis. Transportation Research—B 31 (3), 225–237.

Garrison, W.L., 1960. Connectivity of the interstate highway system. Papers and Proceedings of the Regional Science Association, vol. 6, pp. 121—137.

Giuliano, G., 1989. Incident characteristics, frequency, and duration on a high volume urban freeway. Transportation Research—A 23A (5), 387–396.

Goldberg, M.A., 1975. On the inefficiency of being efficient. Environment and Planning A 7, 921–939.

Golob, T.F., Recker, W.W., Leonard, J.D., 1987. An analysis of the severity and incident duration of truck-involved freeway accidents. Accident Analysis and Prevention 19 (5), 375–395.

Goodwin, P.B., 1992. A quality margin in transport. Traffic Engineering and Control 33 (12), 661–665.

Goodwin, P.B., 1999. Transformation of transport policy in Great Britain. Transportation Research Part A 33 (7-8), 655–669.

Haesken, O.M., Olsen, T.G., Fridheim, T., 1997. Protection of the Society (POS)—Final Report (in Norwegian). FFI/Rapport-97/01459, Forsvarets forskningsinstitutt, Kjeller, Norway.

Hass-Klau, C., Godwin, P., Cairns, S., 1998. Better use of road capacity—what happens to the traffic? Public Transportation International 47 (5), 30–33.

Holling, C.S., 1981. Resilience in the Unforgiving Society. Report R-24, Institute of Animal Resource Ecology, UBC, Vancouver.

Iida, Y., 1999. Basic concepts and future directions of road network reliability analysis. Journal of Advanced Transportation 33 (2), 125–134.

Jones, S.R., 1981. Accessibility Measures: A Literature Review. TRRL Report 967, Transport and Road Research Laboratory, Crowthorne, Berkshire.

Koorey, G.F., Mitchell, T.J., 2000. The Effect of Link Reliability on Benefit/Cost Ratio. Report No 159, Transfund New Zealand, Wellington.

Kronborg, P., 1993. Disturbances in Stockholm's traffic, a Field Study of Greater Stockholm During Two Weeks (in Swedish). MR78, Transport Research Institute, Stockholm. pp. 65–76.

Lam, W.H.K., 1999. Special issue: Network reliability and transport modelling. Journal of Advanced Transportation 33 (2) editorial.

Lam, W.H.K., Zhang, N., 2000. A new concept of travel demand satisfaction reliability for assessing road network performance. Presented at the Matsuyama Workshop on Transport Network Analysis, August 2000.

Laurentius, G., 1994. The vulnerability of the city. In Weissglas, G. (Ed), Planning a High Resilience Society, Geographical Reports no. 11, Umeå.

Mahmassani, H.S., 1990. Dynamic models of commuter behaviour: experimental investigation and application to the analysis of planned traffic disruptions. Transportation Research Part A 24A (6), 465–484.

Mahmassani, H.S., Chen, P.S.T., 1993. An investigation of the reliability of real-time information for route choice decisions in a congested traffic system. Transportation 20 (2), 157–178.

Mahmassani, H.S., Jayakrishnan, R., 1991. System performance and user response under real-time information in a congested traffic corridor. Transportation Research Part A 25A (5), 293–307.

NCPG, Canadian infrastructures and their dependencies, National Contingency Planning Group, March 2000.

Nicholson, A.J., Du, Z.P., 1994. Improving network reliability: a framework. Proceedings of 17th Australian Road Research Board Conference 17(6), 1–17.

Nicholson, A.J., Du, Z.P., 1997. Degradable transportation systems: an integrated equilibrium model. Transportation Research—B 31 (3), 209–223.

Nicholson, A., Berdica, K., Andjic, Z., 2001. Comparing Traffic Models: Two Case Studies. Transport Engineering in Australia, vol. 7. Institution of Engineers, Sydney, Australia.

Noland, R.B., 1997. Commuter responses to travel time uncertainty under congested conditions: expected costs and the provision of information. Journal of Urban Economics 41, 377–406.

Noland, R.B., Small, K.A., Koskenoja, P.M., Chu, X., 1998. Simulating travel reliability. Regional Science and Urban Economics 28, 535–564.

Noland, R.B., Polak, J.W., 2002. Travel time variability: a review of theoretical and empirical issues. Transport Reviews 22 (1), 39–54.

ÖCB, 2000. The Infrastructure Assignment—About Vulnerability in the Technical Infrastructure. Swedish Agency for Civil Emergency Planning, Stockholm, 2000.

PDD-63, May 22, 1998. The Clinton administration's policy on critical infrastructure protection. Presidential Decision Directive, 63.

Pratt, R.H., Lomax, T.J., July 1996. Performance measures for multi-modal transportation systems. Transportation Research Record 1518, July 1996, pp. 85—93.

Prop, 1996/97:11. Preparedness Against Extreme Peacetime Strains on Society (in Swedish). Regeringens proposition 1996/97:11, Stockholm, 1996.

Prop, 2001/02:20. Infrastructure for Long-Term sustainable Transport Systems (in Swedish). Regeringens proposition 2001/02:20, Stockholm, 2001.

Romero, P.J., Adams, J.L., 1995. Economic impact of the North Ridge earthquake. The Northridge, California, Earthquake of 17 January 1994, Woods, M.C., Seiple, W.D. (Eds.). California Department of Conservation, Division of Mines and Geology, special publication, 116, 263–271.

Ross, W., 2000. Mobility and accessibility: the yin and yang of planning. World Transport Policy and Practice 6 (2), 13–19.

Savas, E.S., 1973. The political properties of crystalline H_2O: planning for snow emergencies in New York. Management Science 20 (2), 137–145.

Schmöcker, J.D., Nicholson, A.J., 2001. Assessing transport reliability: malevolence and user knowledge, personal communication.

Small, K.A, Noland, R.B., Koskenoja, P., 1995. Socio-Economic Attributes and Impacts of Travel Reliability: A Stated Preference Approach. MOU-117 Final Report, UCB-ITS-PRR-95-36, University of California, Irvine.

Svensson, H.E., 1994. Disturbances in Göteborg's Traffic. A Field Study of Göteborg During One Week (in Swedish). MR87, Transport Research Institute, Stockholm.

Swedish National Road Administration, 1999. Handling of Crises and Great Peacetime Strains (in Swedish). Publication 1999:12, Enheten för statlig väghållning, Borlänge.

Wakabayashi, H., Iida, Y., 1992. Upper and lower bounds of terminal reliability in road networks: an efficient method with Boolean algebra. Journal of Natural Disaster Science 14 (1), 29–44.

Wakabayashi, H., Iida, Y., 1993. Improvement of terminal reliability and travel time reliability under traffic management. Proceedings of Third International Conference on Applications of Advanced Technologies in Transportation Engineering, Seattle, American Society of Civil Engineers, 25–28 August 1993, pp. 211–217.

Westin, K., Jansson, B., 1994. Assessment of risk in society systems. In: Weissglas, G. (Ed), Planning a High Resilience Society, Geographical Reports no. 11, Umeå.

Wolffram, B., 1981W. Tjörn bridge—an accident that made great demands and resulted in fast construction. Väg-och vattenbyggaren 4, 41–44.

Yang, H., Lo, K.K., Tang, W.H., 2000. Travel time versus capacity reliability of a road network. Presented at Transportation Research Board 79th Annual Meeting, Washington, DC, January 2000.