Vulnerability and resilience of transport systems – A discussion of recent research

Lars-Göran Mattsson*, Erik Jenelius

Department of Transport Science, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

ABSTRACT

The transport system is critical to the welfare of modern societies. This article provides an overview of recent research on vulnerability and resilience of transport systems. Definitions of vulnerability and resilience are formulated and discussed together with related concepts. In the increasing and extensive literature of transport vulnerability studies, two distinct traditions are identified. One tradition with roots in graph theory studies the vulnerability of transport networks based on their topological properties. The other tradition also represents the demand and supply side of the transport systems to allow for a more complete assessment of the consequences of disruptions or disasters for the users and society. The merits and drawbacks of the approaches are discussed. The concept of resilience offers a broader socio-technical perspective on the transport system’s capacity to maintain or quickly recover its function after a disruption or a disaster. The transport resilience literature is less abundant, especially concerning the post-disaster phases of response and recovery. The research on transport system vulnerability and resilience is now a mature field with a developed methodology and a large amount of research findings with large potential practical usefulness. The authors argue that more cross-disciplinary collaborations between authorities, operators and researchers would be desirable to transform this knowledge into practical strategies to strengthen the resilience of the transport system.

1. Introduction

Our societies are highly dependent on a number of critical infrastructure systems, including electric power, transport, water supply and sewage handling, information and communication, and banking systems. These systems have gradually become increasingly complex and interdependent. For instance, most of them require electric power, access to computer networks and road connectivity. If the supply of any of these services stops or is drastically reduced, the dependent systems will fail or function at a low level of performance. As one example, in the event of a general power outage, the mobile telephone network will also stop functioning after only a few hours when back-up batteries are emptied. To minimise the costs, infrastructure systems are often designed to work close to their capacity with small margins of reserve capacity and little redundancy. This renders them sensitive to various incidents, technical failures, disruptions, extreme weather, natural disasters, antagonistic actions and other threats. This, in addition to the interdependencies between and within the systems, could lead to serious consequences for society, should a critical component or sub-system fail or break down.

* Corresponding author.

E-mail addresses: lars-goran.mattsson@abe.kth.se (L.-G. Mattsson), erik.jenelius@abe.kth.se (E. Jenelius).

http://dx.doi.org/10.1016/j.tra.2015.06.002
0965-8564/© 2015 Elsevier Ltd. All rights reserved.
The focus of this paper is on one specific critical infrastructure system – the transport system. The road system, in particular, is fundamental to the functioning of society in developed as well as developing countries. Society is not only dependent on the road system for people’s daily mobility and for goods transport, but it also serves as a life-line system for rescuing people and economic values and for repairing and restoring other infrastructure systems when they are disrupted.

The importance of a robust and reliable transport system from an economic and welfare perspective has led to considerable research in order to understand the mechanisms and interrelationships that create its vulnerability, to find ways to make it more robust and resilient, and to mitigate consequences of disturbances and disruptions. The research efforts have accelerated in recent years, the scope has widened and much new knowledge has been added since Berdica (2002) reviewed the road transport vulnerability literature a decade ago. In this article we discuss what has been achieved since then with an emphasis on the most recent contributions, also taking into account other transport systems than roads. The number of studies is too large and dispersed to make a complete review possible. Our aim is instead to select a number of interesting studies and critically describe their methodological approaches, developed tools, research findings and conclusions. By describing the studies in some detail, our hope is that the article will be useful for researchers and practitioners when judging the opportunities and limitations of present research for practical vulnerability and resilience studies.

The rest of the article is organised as follows. The concept of transport vulnerability is discussed in Section 2 and its relationship with resilience in Section 3. Section 4 provides a literature overview of recent transport vulnerability studies with a focus on contributions to the methodology. The overview starts with topological studies of transport networks based on graph theoretical concepts followed by transport system studies, which add behavioural aspects through travel demand and supply models, and ending with a discussion of their respective merits and drawbacks. In Section 5 we discuss the need for viewing vulnerability analysis as part of a broader resilience perspective and add a discussion of some studies on how to improve the response and recovery phases after a disaster. Section 6 concludes the article.

2. Transport system vulnerability

There is no commonly accepted definition of transport system vulnerability. The definition suggested by Berdica (2002, p. 119) is, however, often cited and representative of part of the literature (her emphases): “Vulnerability in the road transportation system is a susceptibility to incidents that can result in considerable reductions in road network serviceability.” This definition is equally valid for other modes of transport. It emphasises that there is an initiating disruptive event, that the fundamental purpose of the transport system is hurt (its ability to provide transport services to the users), and that the adverse consequences are significant. A somewhat shorter formulation with essentially the same meaning is: “Transport system vulnerability is … society’s risk of transport system disruptions and degradations” (Jenelius and Mattsson, 2015, p. 137). Risk is here perceived in accordance with Kaplan and Garrick (1981, p. 409), who suggest that a risk analysis should answer the questions: “What can happen? How likely is that? What are the consequences?” For every conceivable risk scenario this can be formalised as a “triplet”: a scenario description, the probability and the consequences (measure of damage) of that scenario, respectively. Risk is then conceptualised as the set of all possible such triplets. This is a fundamentally richer description of risk than the not uncommon operationalisation of risk as expected consequence: the product of probability and consequence.

The risk concept may be illustrated in the form of a risk curve (see Fig. 1). The scenarios on the horizontal axis are sorted according to increasing severity of the consequence x. The vertical axis indicates the cumulative probability of scenarios with consequences greater than or equal to x during some period of time. There is some controversy about the meaning of probability. Does it represent the (objective) numerical value to which the relative frequency of a specific scenario will tend in repeated experiments or does it represent somebody’s (subjective) degree of (un)certainty about the relative frequency of that scenario? Following Kaplan and Garrick (1981) the risk curve could be consistent with both interpretations by viewing it as the expected frequency.

Fig. 1 also illustrates the distinction between (un)reliability and vulnerability. Although we think that it is meaningful and useful to make this distinction, it should be remembered that it is not possible to draw a precise boundary between these concepts. Vulnerability, as we understand it, is about events that are infrequent and have considerable adverse consequences. It is thus related to the lower right section of the risk curve.

Reliability is often used in risk analysis in a well-specified meaning as “the probability of a device performing its purpose adequately for the period of time intended under the operating conditions encountered” (e.g., in Billington and Allan, 1992). In the transport literature, reliability is used more generally to describe the stability, certainty and predictability of travel conditions. Taylor (2013) and Rasouli and Timmermanns (2014) provide thorough reviews of recent research on travel time (un)reliability and how uncertainty affects travel behaviour. The focus is on the daily variability in travel times and how the traveller by having general (historical) or up-dated (on the route) information on travel time variation can minimise the disutility related to this unreliability and uncertainty. This means that transport unreliability in this sense is related to the upper left section of the risk curve in Fig. 1.

---

1. For a recent critical review on supply chain risk, see Heckmann et al. (2015).
2. See Cats and Jenelius (2014) for an example of mitigation effects of information in public transport.
There is another fundamental difference between the two sections of the risk curve. It is relatively easy to collect data on travel time variation from day to day, or hour to hour, under different weather conditions and other circumstances. This means that a traveller can accumulate experience of these variations and reasonably well build up a picture of the travel time distribution under various conditions (in line with the objective interpretation of probability). Transport authorities can also do so, for example by utilising data from opportunistic sensors like taxi GPS systems that are becoming increasingly available (Jenelius and Koutsopoulos, 2013). The upper left section of the curve can therefore be estimated relatively precisely.

The situation is different for the lower right section. At the far right we have fundamentally unexpected disastrous events – “Black Swans” in the words of Taleb (2010). Since they are politically, publicly and commercially unacceptable, the evolutionary development of society and technology tends to prevent the same kind of event with disastrous consequences from being repeated. In addition, they are often the result of some totally unexpected mechanism and there are therefore virtually no empirical data about their frequencies or consequences. Therefore, it is not possible to base this section of the risk curve on statistical estimates. Rather, it has to rely on our subjective degree of uncertainty about such rare disastrous events (in line with the subjective interpretation of probability).

2.1. Causes of disruptions

Different kinds of disruptions and threats to critical infrastructure, including the transport system, may require different tools of analysis and courses of action for anticipation, prevention, mitigation and restoration.\footnote{See Janić (2015) for an informative account of recent severe disruptive events affecting air transport.} Useful distinctions can be made between internal and external causes of disruption and between accidental events and intentional interferences (see Table 1). This distinction is related to the separation between safety and security in transport.

Internal threats may originate from mistakes and accidents caused by staff or users, technical failures, components that break down, faulty constructions, overload, etc. They could also be intentional, such as labour market conflicts. Since society is so dependent on transport services, air, rail and public transport systems are all common targets in such conflicts.

External threats may be related to natural phenomena including various degrees of adverse weather and natural disasters: heavy rains, snowfalls, thunderstorms, hurricanes, tornadoes, floods, wildfires, landslides, tsunamis, volcanic eruptions, earthquakes, etc. One specific long-term threat in-between internal and external is global warming, which is partially a consequence of human activities in the transport sector (Koetse and Rietveld, 2009; National Research Council, 2008).

External threats also include antagonistic actions ranging from pranks, sabotage, terrorist actions to acts of war. It is a sad fact that air, train, underground and bus systems have all been targets of several terrorist actions. By their very nature, antagonistic attacks are genuinely difficult to predict both with respect to frequency and location as compared to natural threats and technical and human failures for which it may be possible to collect statistics that are useful for prediction and prevention. Game theory approaches may be a possibility to analyse the vulnerability of transport systems in this respect, however (Bell, 2000; Bell et al., 2008; Chen et al., 2009).

2.2. Impacts of disruptions

The consequences of transport network disruptions can be multifaceted. Accidents, infrastructure collapses and terrorist attacks may lead to injuries and fatalities directly or indirectly. Many everyday disruptions have less severe impacts: a road link may be blocked, some trains have to stop, or some air flights are cancelled for a certain period. Such events will increase the travel time for passengers and delivery times for goods or lead to cancelled trips. This will lead to social or direct economic costs. The costs for making the transport system operational again and for repairing or rebuilding the infrastructure may also be significant.

The transport system is a system of subsystems, and sometimes one subsystem can cover for another and reduce vulnerability to some extent. When the ash eruptions of Eyjafjallajökull in Iceland in April 2010 stopped the air traffic in...
Northern Europe for more than a week, the rail and road transport systems provided alternative options for many travellers. Consequences of disruptions may differ between cities and rural areas. In cities, disruptions affecting one mode of transport may lead to congestion or overcrowding in other modes but the travellers may still be able to reach their destinations. During a public transport strike, for example, empty seats in private cars or biking and walking can be realistic options for some public transport commuters. In rural areas, disruptions in the road system may have more drastic consequences, since there may not exist any realistic alternatives if a critical road is closed.

When capacity utilisation is increased, redundancy is often reduced, and the system becomes more vulnerable (Goldberg, 1975). This is typically the case in many metropolitan areas. Watling and Balijepalli (2012) argue that it is relevant to consider demand growth when assessing road network vulnerability and they have developed a method for separating out the effect of demand growth on the mean, variance and skewness of travel times to identify the most vulnerable links of a network.

It may be noted that an action that reduces vulnerability in one respect may sometimes increase it in another. This can be the case when a new bridge widens the labour and housing markets in a region. The vehicle and train bridge across the Öresund Strait between Malmö in Sweden and Copenhagen in Denmark that was opened in the year 2000 established a faster, more convenient and much more reliable connection than the previous ferry lines between the two countries. In response to the improved transport services, some relocation occurred and the interaction between the countries increased. The number of commuters increased from around 3000 in the year before the bridge was opened to around 16,000 in 2014 (Öresundsinstitutet, 2015). The fact that so many people are now dependent on a bridge that functions reliably has introduced a new kind of vulnerability, induced by the investment that originally reduced the vulnerability.

### 3. A resilient transport system

Resilience is a concept closely related to vulnerability that has also attracted increasing attention in the transport literature. The dictionary definition of resilient refers to the property of a material of “springing back into shape, position, etc. after being stretched, bent or compressed” (Collins Concise Dictionary, 1978). Reggiani (2013) investigates the methodological relationship between network resilience and transport security and stresses the connection between the topological structure of a network and its resilience/vulnerability, an issue we will return to in the next section. She reviews a number of interpretations of resilience and suggestions to measure and enhance it. Holling (1973) introduced resilience in theoretical ecology to allow for a distinction between stability and resilience in ecological systems. Whereas stability is about a system’s ability to return to equilibrium after a perturbation and the speed at which it does so, resilience reflects the system’s capability to persist when exposed to changes or shocks. How large a shock can a system be exposed to without moving to a different state of behaviour? A resilient system may fluctuate between states as long as it is able to absorb shocks, in the sense that its qualitative behaviour does not change. In a resilient ecological system, no species will be driven to extinction by external shocks.

The concept of ecological resilience is similar to static resilience, which Rose (2007) defines as a system’s capability to maintain its function. Wang et al. (2014a, p. 3) suggest a similar definition for supply chain network systems while adding “a capability to predict and assess the damage of possible disruptions”. In contrast, dynamic resilience refers to the rapidity with which a system returns to a state of normal function after a severe perturbation. This is similar to Pimm’s (1984) concept of resilience, which refers to how fast the system returns towards equilibrium after a shock. Reggiani et al. (2015) further deepen the analysis of how resilience and vulnerability can be framed, interpreted and measured, and their relationship with connectivity. These concepts fit nicely into Fig. 2 borrowed from Mc丹iels et al. (2008), which illustrates the effects of decision-making on infrastructure resilience. Ecological or static resilience would then reflect whether the system returns to essentially the same state of function after some external perturbation or shock, while engineering or dynamic resilience refers to the rapidity of this recovery to the full level of function.

We have added conditional vulnerability to the figure. Conditional vulnerability can be seen as the aggregate consequences of a disruption scenario represented by the area between the dotted line in the figure corresponding to the full system function and the relevant curve representing the reduced level of function. The latter curves depend typically on actions of ex ante mitigation or ex post adaptation.

Wang (2015), also drawing on Holling (1973), argues that the concept of resilience should be even more comprehensive to include recovery from disasters, reliability of day-to-day fluctuations in demand and capacity as well as sustainability of the transport system with respect to long-term changes such as climate change.

In safety research, resilience engineering has been launched as a new way of thinking about safety, focussing on proactive processes rather than reactive defences with the goal of increasing the number of things that go right rather than decreasing...
the number that go wrong (Woods and Hollnagel, 2006). Resilience can then be defined as “the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions” (Hollnagel, 2011, p. xxxvi). Compared to vulnerability analysis, resilience engineering offers a much broader socio-technical framework to cope with infrastructure threats and disruptions including preparedness, response, recovery and adaptation (Worton, 2012). Hollnagel (2011, p. xxxvii) phrases this in terms of four cornerstones of resilience: knowing what to do, what to look for, what to expect, and what has happened (see Fig. 3). Vulnerability analysis has essentially dealt with the ability to know what to expect, which is an important prerequisite for adequate proactive actions. The framework of cornerstones indicates the role of vulnerability studies in contributing to the overall goal of strengthening the resilience of the transport system.

4. Vulnerability analysis – knowing what to expect

As indicated, the literature on transport system vulnerability has grown rapidly in the last decades. There are two distinct traditions with limited interaction. The first one could be characterised as topological vulnerability analysis of transport networks. A real transport network is then represented in the form of an abstract network (graph), i.e., an ordered pair $G = (V, L)$ comprising a set $V$ of nodes (or vertices) and a set $L$ of links (or edges). The network could be undirected (no order is assumed between the nodes connected by a link) or directed (there is a start and end node of each link) and unweighted (all links have the same “length”) or weighted (the links may have different “lengths”). The nodes and links in the abstract network may have different counterparts in the real network depending on the application in mind.

The second tradition, which could be called system-based vulnerability analysis of transport networks, represents much more of the structure of the real transport system in the demand and supply models that are applied in the analysis. The transport network is still modelled as an abstract network (graph). Nodes and links then typically correspond to physical intersections and links in the real network. The network is usually weighted with link weights corresponding to actual lengths, travel times, costs or a combination of these in the form of generalised costs. In addition, the interaction between demand and supply is simulated by means of more or less sophisticated comprehensive transport system models (Cascetta, 2009). Travel demand is often modelled in terms of trip generation and attraction, destination choice, and possibly mode choice in which case sub-networks for the different available modes are represented. Route choice may be modelled as network equilibrium taking the interaction with congestion and delays into account. For rail and public transport this may also involve modelling the actual operation of the services in relation to a timetable.

4.1. Topological vulnerability studies of transport networks

The study of topological properties of networks, including transport networks, is a growing research field. Surprisingly little of this research has been published in transport journals, however.

Considering a connected network $G$, the distance $d_{ij}$ between any pair of nodes $i$ and $j$ ($i \neq j$) could be defined as the shortest distance among all possible routes between the nodes (the number of links forming the route in an unweighted network or the sum of the link distances in a weighted network). A natural efficiency (or rather cost) indicator for the network would then be the average distance across all node pairs, i.e.,

![Fig. 2. Effects of decision-making on resilience (conditional vulnerability and dotted line added by the authors). Source: McDaniels et al. (2008, p. 312)](image-url)
be an abstract network with the same nodes as \( G \) to be zero (as well as of is the subgraph of \( G \) number of links between these neighbours. The average clustering coefficient is then the average of the node clustering coefficient \( C(\mathbf{G}) = \frac{1}{N(N-1)} \sum_{i,j \in G} d_{ij} \)

where \( N \) is the number of nodes in the network. The lower this value is, the faster one can get from one random node to another. If the network is not connected, which may happen in a connected network after a disruption, the measure becomes infinite. However, it could be meaningful to compare two different non-connected networks. \textit{Latora and Marchiori (2001)} introduce an efficiency indicator defined as the average across all node pairs of the reciprocals of the node pair distances, i.e.,

\[
E(\mathbf{G}) = \frac{1}{N(N-1)} \sum_{i,j \in G} \frac{1}{d_{ij}}
\]

This indicator is well-defined also for non-connected networks by defining \( 1/d_{ij} \) to be zero \( (d_{ij} = \infty) \) if \( i \) and \( j \) are not connected.

Let \( G_{id} \) be an abstract network with the same nodes as \( G \) but with a direct link between every node pair with a distance equal to the (Euclidean) distance between the nodes. \textit{Latora and Marchiori (2001)} then define (in this case using Euclidean distance for all links) the \textit{global efficiency} of \( G \) as

\[
E_{\text{glob}}(\mathbf{G}) = E(\mathbf{G})/E(\mathbf{G}_{id})
\]

and the \textit{local efficiency} as

\[
E_{\text{loc}}(\mathbf{G}) = (1/N) \sum_{i \in G} E_{\text{glob}}(\mathbf{G}_i),
\]

where \( G_i \) is the subgraph of \( G \) consisting of all neighbouring nodes of \( i \) and the links between them. Obviously, both indicators are between zero and one. The global efficiency indicates how direct the connections are between all node pairs by comparing the (Euclidean) distance along the shortest route in the network with that of a direct link,\(^4\) while the local efficiency indicates how direct the connections are between the neighbours of a node on average. A “small-world” network (see \textit{Watts and Strogatz, 1998}) is characterised by high values of \( E_{\text{glob}} \) as well as of \( E_{\text{loc}} \). The relationship of this indicator of efficiency to other measures of centrality and to the notion of space syntax (see, e.g., \textit{Hillier, 1996}) is discussed at length in \textit{Porta et al. (2006)}.  

4.1.1 Road networks

Representing intersections as nodes and street segments as links, an urban street network will typically not exhibit small-world properties. However, if it is represented by a suitably defined dual network, this may be the case. \textit{Jiang and Claramunt (2004)} define such a dual network by identifying named streets in a real network with nodes and intersections with links of unit length, respectively. The distance between any two streets in the dual network, \( d_{ij} \), is then the smallest number of intersections along any route connecting them. When applied to three cities of varying size and character, each dual network had short average distance \( C(\mathbf{G}) \) and high average clustering coefficient,\(^5\) which is what characterises small-world networks (\textit{Watts and Strogatz, 1998}).

\textit{Demšar et al. (2008)} study the urban street network of the Helsinki Metropolitan Area in Finland. This network consists of almost 70,000 road links and is analysed as an undirected and unweighted network. The authors argue that cut links\(^6\) or links

\[ c_i = \begin{cases} 1 & \text{if } k_i > 2 \\ 0 & \text{otherwise} \end{cases} \]

across all nodes in the network.

\(^4\) For an alternative interpretation, assume that for all node pairs \( i \) and \( j \) (\( i \neq j \)) the distances \( D_{ij} \) in \( \mathbf{G} \), and the distances \( D_{ij}^D \) in \( \mathbf{G}_{id} \), are all independent exponentially distributed random variables with mean values \( d_i \) and \( d_j \), respectively. Then it is easily shown that \( E_{\text{glob}}(\mathbf{G}) = E(\min_i D_{ij}^D)/E(\min_i D_{ij}), \) where \( E[.] \) denotes expected value.

\(^5\) The average clustering coefficient of a network, similarly to local efficiency, captures to what extent nodes form small, tightly connected groups. Let \( k_i \) be the degree of node \( i \), i.e., the number of links directly connected to the node or, equivalently, the number of immediate neighbours of the node, and let \( l_i \) be the number of links between these neighbours. The average clustering coefficient is then the average of the node clustering coefficient

\[ c_i = \begin{cases} 1 & \text{if } k_i > 2 \\ 0 & \text{otherwise} \end{cases} \]

\(^6\) A cut link in a connected network is a link that if removed will subdivide the network into two non-connected sub-networks.
with a high value of betweenness are the critical elements in the network that should be identified in a vulnerability study. Betweenness is defined as the fraction of all shortest routes between every node pair that passes through the link. It indicates how large a share of the shortest routes will be affected if the link is closed, which gives a rough indication of the consequences of the closure. It should be noted, however, that the same value of betweenness in the abstract network may represent very different consequences in reality in terms of how many travellers that are affected and how large the total increase in travel time will be, depending on the demand and the availability and travel times of alternative routes in the real network.

In a recent study by Duan and Lu (2014), a number of graph topological indicators are calculated for six real city road networks in Asia, Europe and North America. The real road networks are represented as abstract networks at three levels of granularity: segment, stroke and community level. Attacks are simulated by successively removing nodes from the abstract networks randomly or according to their degree or node betweenness. Attacks based on the betweenness strategy tend to be much more harmful. At the segment level the relative size $S$ of the largest remaining connected component of the network is very robust against attacks based on random and degree strategies, attacks based on the betweenness strategy tend to be much more harmful. The level of granularity appears to be important for the robustness of the networks. At the stroke level, the networks are vulnerable to both degree- and betweenness-based attacks. Interestingly, the different cities exhibit quite similar robustness patterns.

4.1.2. Public transport networks

Vulnerability studies of public transport networks have until recently almost exclusively been of a network topological nature. One reason may be that the alternative system-based studies require sophisticated modelling tools that are seldom available. The specific network representation of a public transport system in a topological study varies with what is to be illustrated.

Latora and Marchiori (2001) calculate the previously defined indicators of global and local efficiency for some neural networks, the World Wide Web and the Internet, but also for the Boston subway network. Both efficiency measures are very high for the neural networks and relatively high for the communication networks. For the Boston subway network (with stations as nodes, connections between them as links and with distance equal to Euclidean distance) the global efficiency is very high (0.63) whereas the local efficiency is quite low (0.03). The subway system thus offers travel distances between stations that are only 37% less efficient than an ideal system with direct tunnels between every pair of stations. The fault tolerance, on the other hand, is low: If a station is closed, the remaining connections between its neighbours will be far from direct straight connections as indicated by the low local efficiency. As a way of studying the attack vulnerability of networks, Latora and Marchiori (2005) use the efficiency indicator $E(G)$ to assess how important a single component of a network is, i.e., how much efficiency is reduced should the component be removed. Using travel time instead of physical distance, they find that removing the most important link from the Boston subway network would reduce efficiency by 27.5%.

Recall that the degree of a node in an undirected network is the number of links that connects it directly to other nodes, i.e., its number of immediate neighbours (see footnote 4). If the degree distribution follows a power law, at least asymptotically, i.e., the fraction $p(k)$ of nodes having degree $k$ decreases as $k^{-\lambda}$ for large values of $k$, the network is said to be scale-free. Typically $\lambda$ is between 2 and 3. Scale-free networks are considered resistant to random failures but vulnerable to successful attacks on high-degree nodes. In a study of the public transport system of Beijing, Wu et al. (2004) use an abstract network representation where the origins and destinations of public transport lines are nodes and the lines between them are links. The authors find that this network is scale-free with $\lambda$ equal to 2.24 and that it hence should be robust against random failures, if only its high-degree hub nodes are well protected.

Angeloudis and Fisk (2006) study the degree distribution of twenty subway networks selected among the largest in the world with states as nodes and the connections between stations as links. The authors simulate attacks on the stations in each of these subway networks, which turn out to be very robust not only to random attacks but also to attacks on high degree stations. A coordinated attack on several stations is required to achieve serious fragmentation of the subway network. Han and Liu (2009) analyse ten Chinese subway systems represented as unweighted undirected networks, again with stations as nodes and tracks between them as links. They conclude that with this topology the networks are not scale-free and are not of the small-world type. They study the vulnerability of the networks by simulating the consequences, measured as the relative size $S$ of the largest remaining connected component of the networks, when exposed to random error or attack scenarios on their nodes (cf., Duan and Lu, 2014). Several attack strategies are tested and the most effective one is to remove nodes according to their betweenness recalculated after each removal. Whereas $S$ is as low as 0.1 when 10% of the nodes are removed according this attack strategy, $S$ is around 0.7 when the same fraction of nodes is removed randomly. This illustrates the robust-yet-fragile property of subway systems.

Regarding the representation of a public transport system as a network, there are at least four useful ways, as has been clarified by von Ferber et al. (2009). The most obvious one is L-space, where each station is a node and there is a link between any two nodes if they correspond to two consecutive stations of at least one public transport line. In B-space all stations and all lines are represented as nodes, and each line node is linked to all station nodes that it services. There are no direct links between nodes of the same kind. In P-space the nodes are the stations and there is a link between any two stations if they are

---

7 Node betweenness is the fraction of all shortest routes between every pair of nodes that include the node.

8 $S$ is defined as the number of nodes in the largest component divided by the number of nodes in the original network.
by at least one common line. Finally, in C-space the nodes are the lines with a link between any two of them if they have a station in common. The average length of the shortest paths, the cost indicator $C(G)$, is in L-space the average minimum number of stops on a line between any two stations; in P-space it is the average minimum number of transfers needed to travel between any two stations; and in C-space it is the average minimum number of transfers needed to pass between any two lines. The degree has interesting interpretations in the different spaces as well. The degree of a station in L-space is the number of stations that can be reached by one stop; in B-space the degree of a station node is the number of lines servicing it while the degree of a line node is the number of stations on the line; in P-space the degree of a station is the number of stations that can be reached without changing line; and in C-space the degree of a line is the number of other lines to which one can transfer directly.

In a companion paper (Berche et al., 2009), the vulnerability of fourteen large public transport networks to random disruptions and intentional attack scenarios on nodes is studied. Both the efficiency indicator $E(G)$ and the relative size of the largest remaining connected component $S$ are tested as performance indicators when nodes are removed. In the random scenario, nodes are removed randomly; in the attack scenarios, nodes are removed in the order of their importance to the performance of the network. Several importance indicators are evaluated including node degree and centrality indicators such as betweenness. These indicators are either calculated only once for the initial network or recalculated after each node removal. Comparing, for example, the public transport network of Paris as represented in L-space, a removal of 10–12% of the nodes according to their degree or betweenness recalculated after each removal is required to reduce the S indicator by half. For a random attack, on the other hand, it is necessary to remove 37% of the nodes.

In Berche et al. (2012), removal of links is also studied. The degree of a link is then defined as the sum of the degree of its two end nodes minus 2. Moreover, a robustness measure is introduced as the normalised area $A$ under the curve $S(c)$ when the fraction $c$ of removed nodes or links goes from 0 to 1, i.e., $A = 100 \frac{1}{c} S(c) dc$. A general observation is that a link-based random or attack scenario has less severe consequences than its equivalent node scenario. This should not be surprising since the removal of a node also closes all links connected to it, while the opposite is not true. Comparing a selection of thirteen public transport networks, for which nodes or links are removed according to their degree value recalculated after each removal, they find that for node removal the $A$ measure varies between 3 and 11, whereas for link removal it varies between 9 and 47. Accepting $A$ as a robustness measure, a more surprising finding is that removing links in the order of their initial level of degree or betweenness may be a more effective attack strategy than removing links in the order of the successively recalculated levels. von Ferber et al. (2012) apply the same methodology to the public transport networks of London and Paris. They find that the Paris network is significantly more robust than the London network in almost all respects.

Derrible and Kennedy (2010) propose a robustness indicator related to the number of cycles in a network and apply it to thirty-three railway systems in the world. According to this criterion, the subways of Tokyo and Seoul are particularly robust but also those of Madrid, Paris, Osaka, London and Moscow.

### 4.2. System-based vulnerability studies of transport networks

In the decade following the review by Berdica (2002), a substantial number of system-based vulnerability studies of transport networks have been reported. This research has usually been published in transport journals. Murray et al. (2008) and Wang et al. (2014a,b) provide thorough methodological overviews. Heckmann et al. (2015) complement these overviews by reviewing supply chain vulnerability studies and Khademi et al. (2015) by reviewing vulnerability studies in the perspective of their usefulness for disaster risk reduction. We will restrict our review to a selection of recent research of particular interest from a methodological or application point of view; some classical examples will also be discussed briefly.

#### 4.2.1. Road networks

Pioneering contributions, although not expressed as vulnerability studies, occurred in the early operations research tradition. They include Garrison (1960), who used graph-theoretical concepts to analyse the structure of the U.S. Interstate Highway system, and Wollmer (1964), Ratliff et al. (1975) and Ball et al. (1989), who developed algorithms to find the most “vital” link(s) or node(s) in a network, i.e., the link(s) or node(s) that when removed would increase the shortest path or reduce the maximum flow between a given node pair of origin and destination the most. These ideas of finding hot spots in an infrastructure system are equally relevant for today’s vulnerability studies.
Another research tradition has emanated from transport policy and planning studies and from the development of modelling tools to support such studies. The transport system is not a system that operates under stable demand and supply conditions although the first generation of travel demand models often assumed so. Recent research on transport system reliability has focused on the performance of the system under demand or capacity fluctuations (Taylor, 2013). A number of reliability measures have been developed. Terminal or connectivity reliability is the probability that there is still a connection between a pair of nodes when one or more links are removed (e.g., Bell and Iida, 1997; Wakabayashi and Iida, 1992); travel time reliability is the probability that a trip can be completed within a specified time interval (e.g., Clark and Watling, 2005; Yang et al., 2000); and capacity reliability is the probability that a network can accommodate a specified level of travel demand (e.g., Chen et al., 2002; Yang et al., 2000). Connors and Watling (2014) introduce the related concept of demand vulnerability to capture the impact of unknown future changes in demand on level of service and how it could be analysed at a comprehensive level by network aggregation.

One of the first published transport studies that explicitly mentions vulnerability is Nicholson and Du (1997). The authors argue for the use of an integrated traffic equilibrium model to identify the socio-economic impacts of degradation of transport system components. At the time, several natural disasters and terrorist attacks indicated that the transport systems were vulnerable and research on infrastructure protection was stimulated. It was recognised that additional approaches to those applied in the field of transport reliability were needed to adequately analyse and assess such scenarios with severe consequences (Berdica, 2002; D’Este and Taylor, 2003; Nicholson, 2003). Since then a large amount of theoretical and applied system-based transport vulnerability research has emerged.

Recent research has progressed in different directions. Some studies have developed frameworks or metrics for evaluating transport network vulnerability. Murray-Tuite and Mahmassani (2004) propose a vulnerability index that accounts for traffic flow, link capacities, travel times and the availability of alternative routes. Jenelius et al. (2006) suggest the use of increase in generalised travel time/cost (which can be seen as a demand-weighted generalisation of the cost indicator $C(G)$) together with unsatisfied demand as measures of the reduced level of service following the disruption of a link. The authors use this basic idea to define measures of link importance and regional exposure, and apply them to the road network of northern Sweden. Scott et al. (2006) independently use the same idea when they propose a Network Robustness Index to identify important links in a highway network. Their index for a link is defined as the increase in user equilibrium travel time that is incurred when the link is closed. In Sullivan et al. (2010) this index is modified to allow the level of link disruption to be lower than 100%. The authors also suggest an aggregation of individual link values to a network robustness index that allows for comparisons between different transport networks. A similar importance measure, although based on the reciprocals of the travel costs, is defined by and applied in the evaluation of network robustness in Nagurney and Qiang (2007, 2012). This is the natural demand-weighted generalisation of the topological efficiency indicator $E(G)$.

A technical advantage of the efficiency index applied by Nagurney and Qiang (2007, 2012) as compared to the travel cost indicator applied by Jenelius et al. (2006) is that the former is also well-defined for non-connected networks. On the other hand, the interpretation is not as clear and intuitive as for the cost indicator. In practice, a network will only be non-connected for a certain time after the disruption of a cut link, and the increase in travel cost will in fact be finite. The consequences will depend on the duration of the closure, which is thus an additional important parameter of a vulnerability scenario. To incorporate this idea, Jenelius (2009, 2010) suggests that the travel time increase following the closure of a link should, for those who have no alternative routes, be calculated as the time they have to wait until the link opens again, and for those who have alternative routes as the minimum of that time and the additional travel time along the shortest alternative route. It should be noted that the longer the closure duration, the more important are cut links relative to other links (Jenelius and Mattsson, 2015). This is particularly relevant when comparing the disruption performance of road networks in sparsely versus densely populated areas.

Rupi et al. (2014) suggest an alternative way of generalising the concept of link importance that also serves the purpose of associating a cost to cut links. They first suggest that the importance of a link should have two components: the average daily traffic on the link if there are no disruptions (local importance) and the increase in total travel cost in the study area if the link is closed (global importance). In the case of a cut link, the global importance is modified by adding a cost proportional to the unsatisfied demand. Finally, both components are scaled so they vary between 0 and 1 for the studied set of links and weighed together. It can be noted that local importance is a kind of demand-weighted betweenness measure. The authors then demonstrate the usefulness of this generalised importance measure by applying it to a road network with a large share of cut links, namely that of the mountainous province of Bolzano in northern Italy.

Taylor et al. (2006) propose the use of various measures of diminished accessibility to evaluate the impacts of link failures and analyse the Australian road network at different levels of spatial resolution. Sohn (2006) also applies measures of accessibility reduction in an assessment of critical links in Maryland in the U.S. under flood damage. Chen et al. (2007) propose the use of a combined travel demand model incorporating trip generation, destination, mode, and route choices to assess the long-term equilibrium effects of the closure of one or more links. The consequences are calculated as the decrease of a utility-based accessibility measure. A similar accessibility-based measure is proposed by Taylor (2008). As a variant of an accessibility index, Taylor and Susilawati (2012) employ a remoteness index that is used to identify the most critical road links in a rural area in south east Australia, i.e., the links that if closed would increase the index the most. Snelder et al. (2012) develop a framework for robustness analysis that synthesises and extends many of the ideas that hitherto have been applied in vulnerability studies. In contrast to previous studies, Balijepalli and Oppong (2014) focus on road networks in urban areas. They compare the performance of several vulnerability indices usually applied to sparse regional networks.
and suggest a new index more relevant to urban networks by considering the importance of the road links for the urban network and illustrate its computation in the event of a flooding in the city of York in England.

Nyberg and Johansson (2013) combine spatial population data with data about the location of high forest stands near roads in a case study of road vulnerability to storm-felled trees in southern Sweden. This enables the authors to calculate the exposure of, for example, elderly in rural areas to reduced accessibility to certain critical services. Another approach to determining road link criticality for reaching (emergency) service facilities is suggested by Novak and Sullivan (2014). The criticality is then based on an index called critical closeness accessibility, which is related to the closeness concept in graph theory.

Some studies relate vulnerability indices to indicators based on transport network and/or travel pattern characteristics. Jenelius (2009) finds in a case study of the Swedish road network that regional importance, i.e., the expected increase in total network travel time should a random link in the region be closed, can largely be explained by network structure and average traffic load in the region. Regional user exposure, i.e., the expected average increase in travel time per trip starting in the region should a random link in the whole network be closed, can largely be explained by the network structure and the average user travel time in the region. El-Rashidy and Grant-Muller (2014) suggest the use of linear regression to establish a relationship between link importance, i.e., the increase in total network travel time following a link closure, and a number of attributes associated with that link. To illustrate the applicability of their approach, the authors apply it to a simplified road network of Delft in the Netherlands. With four attributes dependent on link capacity, flow, length, free flow speed, and congestion density, and by modifying link importance to account for the consequences of cut links, they achieve an $R^2$-value of 0.91, which is similar to the level of explained variation in Jenelius (2009). Knoop et al. (2012) study the correlation between a similar set of link-based indicators of link importance and the true delays of link closures as calculated by a dynamic traffic assignment model for the networks of Delft and Rotterdam in the Netherlands. In slight contrast to El-Rashidy and Grant-Muller (2014), the authors conclude that even a combined indicator based on linear regression cannot predict the importance of a link satisfactorily.

Another direction of research deals with modelling and computational aspects of vulnerability analysis. He and Liu (2012) model the day-to-day development of traffic flows after an unexpected network disruption based on past experience and beliefs about the future conditions. Having access to behavioural field data from the 2007 collapse of the I-35W Mississippi River Bridge in Minneapolis, Minnesota, they are able to calibrate their model to represent the traffic flow dynamics after a real-world network disruption. In a study of the Swiss road network, Erath et al. (2009) handle the computational challenges associated with full-scan analyses of congested networks by restricting the calculations to a sub-network surrounding each disrupted link. A similar approach is used by Chen et al. (2012) to identify important links in the Hong Kong road network under stochastic demand with a reliability-based user equilibrium model. Luathep et al. (2011) study the possibility of using sensitivity analysis as a way of reducing the computational requirements of a full-scan vulnerability analysis of single-link disruptions. Recently, Bíl and Vodák (2015) have reported some progress on finding critical combinations of links by means of a simulated annealing algorithm.

Still another direction of research develops mathematical modelling and optimisation techniques to identify worst-case scenarios, or best responses to such scenarios. Matisziw and Murray (2009) use an integer programming formulation to identify the most severe disruptions of a given number of links in the truck transport network of Ohio, USA. Bell et al. (2008) integrate a macroscopic traffic assignment model in a game-theoretic framework to determine routing strategies for the shipment of VIPs in London under the risk of antagonistic attacks. Lou and Zhang (2011) identify network design strategies for optimising the resilience of transport networks under random and targeted attacks. Yates and Sanjeevi (2013) develop a variant of the shortest path network interdiction problem to analyse attacks on critical infrastructure, using the road network to reach targets. The problem is represented as a two-player game, in which the attacker seeks the path of maximum non-detection between any entry and target node, while the defender allocates a limited number of sensors that reduce the non-detection probabilities to minimise the value of the maximum non-detection path. The model is applied to a subset of the California highway network.

Further, many papers put most emphasis on the vulnerability evaluation itself. Notably, several authors have assessed the economic impacts of earthquakes disrupting the road network using integrated transport network and multiregional trade models (Cho et al., 2001; Ham et al., 2005; Kim et al., 2002; Tatano and Tsuchiya, 2008). Bono and Gutiérrez (2011) use network theory and GIS-based analysis to evaluate the impacts of disrupted roads on accessibility following the earthquake that hit Haiti in 2010. Suárez et al. (2005) study the impacts of flooding and climate change on the urban transport system of the Boston Metro Area, using a four-step (i.e., trip generation, destination, mode and route choice) transport modelling system. A similar approach is used by Berdica and Mattsson (2007) to evaluate the societal impacts of bridge closures in Stockholm, Sweden. Dalziell and Nicholson (2001) identify a range of plausible disruption hazards for the Central North Island road network of New Zealand. Specific scenarios are then generated through Monte Carlo simulation, and the impacts including both restoration work and the economic losses of the travellers are evaluated. Following this assessment, a cost-benefit analysis is performed to identify the most efficient options for mitigating the identified vulnerabilities.

Many disruptions in transport networks caused by natural events such as snowstorms, floods and earthquakes will typically affect a wider geographical area. The approaches in the studies considered so far, however, are only applicable to disruptions affecting a single or a pre-defined number of links in a network. In contrast, Jenelius and Mattsson (2012) present a grid-based approach to analyse area-covering disruptions. More specifically, the authors represent an area-covering disruption with a square cell, within which all links are assumed to be closed during a certain period, and they calculate the
resulting increase in total travel time. By moving the cell systematically across the study region, they can identify the most critical areas. The impacts are largely determined by the amount of traffic affected by the closure of the links in a cell, which in turn is mainly determined by the population density. This is in contrast to single-link disruptions, where the redundancy of the surrounding network is also of importance (see also Jenelius and Mattsson, 2015). Mitsakis et al. (2014) also consider events with area-wide extension. They suggest an integrating framework for determining optimal transport network adaptation strategies to handle extreme weather and climate change impacts. Another aspect of interest is how people change their travel and location behaviour in response to extreme weather events. Lu et al. (2014a, 2014b) report travel and residence and job location change choice behaviour under flooding and extreme weather based on questionnaire data collected in Bangladesh. Road disruption, isolation by water, and flood frequency were found to affect behaviour significantly.

Ho et al. (2013) suggest a continuum modelling approach as an alternative way to find the most critical areas of degradation in a study region. The vulnerability analysis is formulated as a bi-level problem, where a continuum traffic equilibrium model is used to determine travel costs and traffic flows at the lower level given the location of network degradations from the upper level. At the upper level, the most vulnerable locations for degradation are found to be those that reduce an accessibility index most, based on traffic information from the lower level. This bi-level problem is solved by iterating between the two levels. A city’s land use pattern affects the use of its road network. Zhao et al. (2014) develop a bid-rent approach to find efficient land-use adaptation strategies to mitigate the vulnerability in the long run.

To improve liveability, many cities consider reducing the access of cars to the road network in its central areas. Ortigosa and Menendez (2014) study how different road removal strategies affect traffic performance in a city assuming a stylised grid network and uniform travel demand between node pairs. Links are removed according to three different strategies under the restriction that all nodes should remain connected: links are randomly removed, or peripheral links or central links are most likely to be removed. As expected, total travel time and distance travelled increase and average speed decreases at an increasing rate as more links are removed. The traffic performance deteriorates fastest with the central removal strategy and slowest with the peripheral strategy. For the central strategy, the total travel time is almost doubled when 15% of the links are removed.

4.2.2. Rail and public transport networks

Rail and public transport networks are generally more sensitive to disruptions than road networks. Subway and rail networks are usually rather sparse and there are often limited possibilities to redirect trains if some link is disrupted. Moreover, if a train breaks down, this will knock on delays on subsequent trains. It is therefore surprising that relatively little research has been devoted to demand-based rail and public transport vulnerability analysis. Some recent exceptions are considered in the following.

Rodríguez-Núñez and García-Palomares (2014) are interested in characterising the importance of the links in a subway system. They assume that link travel times and an OD travel demand matrix are known and that travellers choose the fastest route in the network to reach their destinations. The closure of a link can have two distinctly different outcomes: (1) the network is separated into two non-connected components, or (2) some travellers have to make a detour to reach their destinations. Following Jenelius et al. (2006), they define the importance of a link in case (1) as the unsatisfied demand, i.e., the number of trips that cannot be carried out and in case (2) as the increase in average travel time assuming that affected travellers make the fastest possible detour. An application to the Madrid subway system shows that the closure of the most important link that separates the network into two components leads to an unsatisfied demand for a normal weekday of 116,000 trips (4.7% of total demand). Closing the most important non-cut link increases total travel time by 7.2%. Also closing the second-most important non-cut link increases total travel time by 13.5%, and when the five most important non-cut links are closed, the accumulated increase is 43.6%. De-Los-Santos et al. (2012) perform a similar network scan evaluation while also considering the case of a replacement service for the closed link.

Li and Kim (2014, p. 8) introduce the robustness concept survivability “as the ability of a network to maintain its topological and functional state when a certain level of disruptions on stations occurs simultaneously”. They operationalise survivability as a combination of system connectivity loss and passenger flow loss when one or more hub stations are assumed to be disrupted. In an application to the Beijing subway, they conclude that its survivability after the disruption of one hub is quite good. If two or more hubs were disrupted, the severity of the consequences would be quite dependent on the precise choice of disrupted hubs. This indicates that the Beijing subway may be quite vulnerable to informed attacks.

Cats and Jenelius (2014) introduce a dynamic, stochastic and multimodal notion of public transport network vulnerability, accounting for interactions between supply and demand and the accumulated effect of disruption on system performance. Candidate critical links are then identified by extending the measures of betweenness centrality and link importance to a dynamic-stochastic setting from the perspectives of both operators and passengers. The criticality of a link is evaluated as the reduction in welfare (considering travel time, number of transfers, etc.) due to a capacity reduction of the link. Their findings suggest that the dynamic betweenness centrality measure is a better indicator for disruption impact on passenger welfare than the static betweenness centrality, which is based solely on network topology. The authors also study the mitigating impact of real-time information provision, and find that it may have a significant positive influence, although counter-examples also exist due to cascading effects. Cats and Jenelius (2015) study the possibility of reducing vulnerability by increasing the capacity on lines that can serve as alternatives when critical links are disrupted, and propose a methodology for identifying the lines where capacity increases are the most effective.
Chandra and Quadrifoglio (2013) develop a methodology for identifying critical links for demand responsive feeder services in a stylised grid street network. Assuming uniform demand between the node pairs in the network, they derive an approximate analytical expression for the increase in total demand-weighted travel distance if one or more links in the street network are removed. Not surprisingly, removal of central links leads to the largest increase in demand-weighted distance. The conclusions are similar to the study by Ortigosa and Menendez (2014) for road traffic.

Railway systems are particularly vulnerable to natural hazards and other threats because of their lack of excess capacity, limited possibilities to reroute trains and the presence of single line tracks. Hong et al. (2015) propose a methodology to study the vulnerability of the Chinese railway system under floods and the effectiveness of alternative mitigation strategies. This study is particularly interesting because it estimates disruption probabilities. The authors generate flood event scenarios based on flood statistics for the past 30 years. Based on these data, they estimate link disruption probabilities. These are then used in a Monte Carlo simulation study to calculate the average number of interrupted trains by province and time period. Finally, the effectiveness of alternative strategies in selecting links for maintenance to reduce their disruption probabilities is evaluated. Selecting links according to the product of original link disruption probability and the number of trains using the link (a kind of betweenness measure) turned out to be the most effective mitigation strategy.

This study is particularly interesting because it estimates disruption probabilities. The authors generate flood event scenarios based on flood statistics for the past 30 years. Based on these data, they estimate link disruption probabilities. These are then used in a Monte Carlo simulation study to calculate the average number of interrupted trains by province and time period. Finally, the effectiveness of alternative strategies in selecting links for maintenance to reduce their disruption probabilities is evaluated. Selecting links according to the product of original link disruption probability and the number of trains using the link (a kind of betweenness measure) turned out to be the most effective mitigation strategy.

Peterson and Church (2008) develop a framework for modelling rail freight transport vulnerability in much the same way as Jenelius et al. (2006) and Scott et al. (2006) for road networks. The authors apply a simple routing model (a slight generalisation of the classical shortest-path algorithm) to find the best way of routing a shipment in the original network and rerouting it after a certain link has been disrupted. Using this as the basis for the estimation of the detour costs of a disrupted link, they analyse the impacts for rail freight to and from the State of Washington in the U.S. due to the disruption of an important bridge in the State of Idaho. A similar approach is employed by Du et al. (2014) in a full-scan analysis of the multiple logistics transport network of the island of Hokkaido, Japan.

Gedik et al. (2014) extend the tools for vulnerability analysis of freight trains under disruptions. They formulate an integer maximisation problem where an antagonist seeks to maximise the sum of train transport and delay costs by destroying a limited number of nodes in the network given that an operator reroutes the trains on the remaining network to minimise the increase in these costs. This is a bi-level mixed integer problem and the authors suggest an approach whereby they relax the formulation to make it computationally feasible. They apply their model to a real coal transport network in the U.S. One conclusion is that increasing the number of disruptions escalates the delay cost dramatically, whereas the transport cost remains more or less the same.

4.2.3. Waterway and air transport networks

The redundancy of inland waterway networks is typically very low and they are therefore particularly vulnerable to disruptive events. Baroud et al. (2014) propose two stochastic importance measures to quantify the initial impact of a disruption in an inland waterway network and its ability to recover functionality after the disruption. The authors show how the measures can be used to prioritise resilience improvement activities in the network. They apply their methodology to a case study of the Mississippi River Navigation System in the U.S. In another U.S. case study, DiPietro et al. (2014) perform an in-depth analysis of the consequences of potential failure of heavily used locks and dams in the Monongahela River waterway system in Pennsylvania. They focus on the impacts on shipments of coal from mines to power plants of a potential but not unlikely failure of an outdated critical dam, which could close the navigation on the river for a year. They carefully assess potential mode shift responses, market-wide responses and financial impacts of this infrastructure disruption. The paper presents a valuable contribution to the methodology of assessing the wider societal impacts of a vulnerability scenario.

Janić (2015) presents a methodology for estimating the resilience, fragility and costs of an air transport network exposed to a large-scale disruptive event. The methodology is interesting in general because, with some modifications, it could be applied to other transport networks. Resilience is understood here as “the ability to withstand and stay operational . . . during the impact of a disruptive event” excluding actions in the recovery phase. The author operationalises the resilience of an airport in different ways. The most obvious measure is the realised on-time and, possibly, delayed flights to and from an airport during the disruption as a proportion of the planned flights. As such, this measure is a kind of opposite to a typical vulnerability measure. The fragility of an airport is defined as the reduction in the aggregate resilience across all airports in the network in the event that the considered airport is closed. The applicability of the developed methodology is demonstrated by a study of part of the U.S. air transport network during Hurricane Sandy in October 2012. Interestingly, the methodology could also be applied a priori in a “what-if” scenario approach.

4.3. Discussion of the approaches

We have identified two broad approaches in vulnerability analysis of transport systems that have essentially developed separately from each other. They each have different advantages and disadvantages.

The main advantage of the topological approach is its limited data hungerness. The methodology relies on a mathematically elegant, well elaborated and rigorous theory. At the most basic level, a transport system is represented as an undirected and unweighted network, which can be done in different ways depending on the specific application in mind. Besides the obvious way of representing intersections/stations/airports as nodes and road/track segments/air routes as links, there are other useful representations frequently used for rail and public transport systems for demonstrating small-world behaviour. Vulnerability scenarios are simulated by removing nodes or links randomly or according to some attack strategy...
based on some centrality measure such as their degree or betweenness values calculated initially or after each removal. The performance of the network after the removal of nodes/links is often evaluated as the change in the network efficiency $E(G)$, i.e., in the average of the reciprocals of the shortest distances between all node pairs in the network. An alternative or complementary performance indicator is the relative size $S$ of the largest connected component of the network, or, as an aggregate indicator, the area under this size curve $S(c)$ as a function of the fraction $c$ of removed nodes or links. A general conclusion is that transport networks are usually quite robust with respect to random removal of nodes/links but more vulnerable to attacks especially when nodes are removed based on betweenness recalculated after each removal. The networks are also less vulnerable with respect to link removal when the removal is based on attack strategies.

The straightforwardness of the methodology and the limited need for data make it computationally realistic to study the performance of transport networks under the successive removal of all nodes/links. This systematic removal of all nodes/links is rarely possible in system-based vulnerability studies. It is also realistic to perform the extensive calculations necessary to compare different transport networks with each other as well as with other real or theoretical networks. Zhang et al. (2015) explore this possibility by systematically studying the resilience of seventeen principal network topologies in terms of throughputs, connectivity, and compactness in a number of numerical experiments.

The limitation of the approach is linked to its strength: the simplistic way in which the transport system is represented as an abstract network does not allow for a very realistic description of the behavioural responses to a disruption. One somewhat technical example of this is the use of average distance $C(G)$ and efficiency $E(G)$ to represent the performance of a real physical transport network when this is represented by an abstract network. If a new network is formed by simply joining two sub-networks, average distance will increase and efficiency will decrease as compared to these values for the sub-networks although the performance has not necessarily deteriorated. Calculating, for example, these indicators for the joined road networks of Sweden and Denmark after the opening of the Öresund bridge would indicate a higher average distance and a lower efficiency than calculating the same indicators separately for each national network. The reason is that distances between very remote node pairs with limited interaction will also be included in the calculation for the joined network. It should then be noted that the effect on efficiency will be less pronounced than the effect on average distance. By weighting the node distances or their reciprocals with the relative travel demand as suggested by Jenelius et al. (2006), a size-neutral performance indicator may be defined. Then demand data are also needed in addition to the more easily acquired topological network data.

In reality, the consequences of a vulnerability scenario will depend on the duration of the disruption, how many travellers are affected, to what extent they can reroute their trips, change destinations, modes, or postpone the trips, what information they have about the alternatives to minimise the negative consequences of the disruption, etc. Moreover, with a topological approach it is not possible to capture dynamic effects of the disruption such as increased congestion on alternative routes and the behavioural responses to that. When a cut link in a subway or rail network is disrupted, it is highly unrealistic to assume, as is implicitly done, that each sub-network can continue to function as before. Although topological vulnerability studies can provide important general insights and indicate various structural weaknesses in transport networks, they are in some respects too simplistic to be useful for assessing actual policy actions in a specific transport system.

System-based transport vulnerability studies can be seen as an attempt to overcome some of these limitations of topological studies. This comes at the price of requiring more information such as travel demand and supply data as well as calibrated behavioural models, by which travellers’ responses to disruptions and their further repercussions on other travellers can reasonably well be predicted. This means that the methodology of system-based vulnerability studies will be less uniform than topological studies. What can be studied in a concrete situation will depend on the availability of data and models. A plethora of approaches have therefore been developed that capture various aspects of the vulnerability of transport systems. The computational efforts of a systematic removal of combinations of nodes/links are presently prohibitively large. Performing a full-scan analysis of a real large-scale network is only realistic for single-link/node removal. Even the full-scan analysis of the removal of all combinations of two links/nodes would be computationally infeasible. This makes it difficult to study how the vulnerability of a network develops over time and to compare the vulnerability of a transport network with other kinds of networks. This is easier to do within the topological approach.

Dehghani et al. (2014) represent a rare example of a vulnerability study that compares a topological-based measure with a system-based measure. The authors set up a hypothetical road network, where each link has a given length and can be closed independently of the other links with a certain probability. An OD matrix for the travel demand between the nodes in the network is randomly generated. As topological vulnerability measure they apply the expected (by applying the link closure probabilities) relative decrease in the efficiency indicator $E(G)$ and as system-based indicator the expected relative increase in the demand-weighted modification of the average distance indicator $C(G)$, or, equivalently, in vehicle miles travelled (VMT). They simulate a large number of disruption scenarios with varying average closure probability. The interesting and surprising result is that both indicators, although S-shaped and having very different scales, behave quite similarly as a function of the average closure probability. This indicates that topological vulnerability indicators can at least in certain cases be useful alternatives to the more data-demanding system-based indicators.

Compared to topological vulnerability studies, the richer set-up of system-based studies allows more intuitive impact measures to be defined, capturing a wider spectrum of consequences of disruptive events. These measures may range from the simplest ones of increase in travel time or generalised travel cost, or cancelled travel options, over more general measures of accessibility to comprehensive economic measures of consumer surplus or financial impacts. It is then important to incorporate how travellers value their (increased) travel time in different situations. One aspect that has received
surprisingly little attention in the literature is the connection between urban structure and vulnerability. Some studies have looked at where the most critical links in an urban street network are located (Chandra and Quadrifoglio, 2013; Ortigosa and Menendez, 2014) and possibilities to mitigate vulnerability by land-use strategies (Zhao et al., 2014). We also discussed how a bridge that replaces unreliable ferry lines can increase the economic integration in the region and in the end induce a new kind of vulnerability. Another issue that would deserve more attention in the future is equity aspects of vulnerability. Jenelius (2010) provides an example of a spatial equity study but social and economic equity aspects are equally important to analyse.

Many disruptive events affecting transport systems have consequences over an extended area, affect more than one system or may have cascading effects on other depending systems. Blizzards and heavy storms have a geographical extension and often disrupt more than one mode of transport and possibly other infrastructure systems. There are some studies attempting to analyse geographically extended effects (Jenelius and Mattsson, 2012) and interrelationships within and between systems with applications to track-based systems (Johansson et al., 2011; Zhang et al., 2014) and to road systems (Hémond and Robert, 2010; Hsieh and Feng, 2014). Further methodological development in this respect seems to be an interesting avenue of future research. This is particularly true for the analysis of dependencies and cascading effects in track-based transport systems such as railway and subway systems.

Most of the studies in the literature analyse conditional vulnerability, i.e., what the consequences would be given a certain disruption in the studied transport network. There are few attempts to estimate the probability of disruptions. Alexakis et al. (2014) represent one exception where GIS techniques and expert assessments are combined in a hazard assessment model to predict the probability of landslides for an area in Cyprus. The previously discussed study by Hong et al. (2015) represents another approach that is applied to railways exposed to flooding.

5. Knowing what to do – towards resilience analysis

Returning to resilience and Hollnagel’s (2011) four cornerstones, vulnerability analysis deals primarily with knowing what to expect. Resilience, on the other hand, offers a broad socio-technical perspective that emphasises how the anticipating ability of vulnerability analysis must interact with the monitoring (knowing what to look for), responding (knowing what to do) and learning (knowing what has happened) abilities in order to contribute to a more resilient system. While there is a substantial literature on transport system vulnerability, as our selective review has indicated, the literature on transport system resilience is less extensive. This is evident from Faturechi and Miller-Hooks’ (2014) comprehensive overview of transport infrastructure system performance during disasters (see also Khademi et al., 2015). The authors find a large body of research on assessment of critical components in the transport systems (vulnerability studies) but much less on disaster management. In the latter case, there is a concentration on the pre-disaster phases of mitigation and preparedness rather than on the post-disaster phases of response and recovery (Özdamar et al., 2014, and references therein are examples of studies of the latter kind, however).

Nicholson (2007, p. 350) uses similar categories of options for reducing road network unreliability according to when they occur:

- “reduction (i.e., identifying and reducing risks via infrastructure improvement)
- readiness (i.e., training civil defence staff and installing warning systems)
- response (i.e., reacting to emergency situation in the short term)
- recovery (i.e., minimising the socio-economic impacts in the long term).”

Nicholson (2007) notes that the primary research focus in the past has been on reduction options, which is the aim of vulnerability studies. He currently sees a shift in interest towards readiness, response and recovery options. To enhance resilience in transport systems, he identifies a need for research in areas such as organisational planning for hazard events, decision support tools for prioritising physical response and recovery of infrastructure networks, and suitable legal and contractual frameworks in the event of major disasters. As one example of such research, Rowan et al. (2014) report on efforts to develop indicators for transport managers that will help them assess climate change vulnerabilities in their transport system. This research also indicates the importance of stakeholder engagement and collaboration for incorporating institutional knowledge into the assessment.

Imran et al. (2014) discuss the complex issue of measuring resilience in a case study of Manawatu in New Zealand. They rely on a methodological framework developed by Bruneau et al. (2003) for seismic resilience of communities, in which actions to improve resilience may aim at reducing the probability, consequences, or duration of events that severely hurt the function of the system. The authors suggest a conceptual definition of a measure of seismic resilience that has general applicability to infrastructure systems including transport systems. Let \( Q(t) \) be the quality of an infrastructure system at time \( t \) expressed as percentage of full quality during normal function. The loss of resilience due to a disastrous event at time \( t_0 \) until the system is fully recovered at time \( t_1 \) is then

\[
R = \int_{t_0}^{t_1} [100 - Q(t)] dt.
\]
This measure corresponds to what we termed conditional vulnerability in Fig. 2, which illustrates the difficulty of strictly separating the definitions of different related concepts. Bruneau et al. (2003) suggest that in specific applications a number of resilience measures should be defined in terms of robustness, redundancy, resourcefulness and rapidity as well as technical, organisational, social and economic dimensions (see Reggiani, 2013, for a more extensive discussion of these issues). In their case study, Imran et al. (2014) identify ways of improving robustness (strengthening roads to withstand landslides and floods better), redundancy (increasing the number of alternate routes), resourcefulness (strengthening the ability to identify, prioritise and address problems) and rapidity (increasing the capacity of rapid recovery and restoration after road closures). Cox et al. (2011) extend the concepts of Bruneau et al. (2003) and develop operational metrics for the resilience of passenger transport systems to terrorism. Their measure of direct static economic resilience (DSER) is the relative avoidance of maximum economic loss that a particular disastrous event could bring about. In a case study of the July 2005 London Underground and bus bombings, they find that 77.4% of the total reduction in journeys on attacked modes were offset by increases in alternate modes during the four months following the attacks. This illustrates the ability of the transport system to absorb the shock of a terrorist attack.

There are several studies focusing on optimising the recovery and restoration process after a disaster or severe disruption in a transport system. Kepaptsoglou et al. (2014) apply a bi-level network design model for planning highway operations for the surviving network in the recovery period. In an application to the Peloponnesian region of Greece, they consider lane reversal and shoulder use for managing day-to-day trips after disruptions due to wildfires or an earthquake. Matisziw et al. (2010) study trade-offs between the minimisation of system cost and the maximisation of system flow in restoring a disrupted infrastructure network. Aksu and Özdamar (2014) and Özdamar et al. (2014) also focus on the post-disaster phase and define a Debris Clearance Scheduling Model that schedules the road restoration in a region in order to maximise the total weighted earliness of all cleared paths.

When a disturbance or disruption occurs in the railway system, it is necessary in real time to reschedule the timetable, the rolling stock and the crews. Recovery models and algorithms for real-time disturbance and disruption management have recently been subject to considerable research effort. Cacchiani et al. (2014) provide an extensive review of this research. Developed models have only been applied in experimental situations, and although there are a lot of promising results, the research is far from being applied in practice. Andersson et al. (2013) focus on strategies to decrease sensitivity to disruptions by increasing timetable robustness. They compare various robustness measures and suggest a new one that enables critical points in the timetable to be identified. These are the points in time and space where adding runtime margin could be particularly effective in increasing timetable robustness. Nielsen et al. (2012) present an approach for disruption management of railway rolling stock to provide train dispatchers with decision support tools to mitigate the consequences of disturbances. Chen and Miller-Hooks (2012) define a resilience indicator for intermodal transport networks. They then formulate a mathematical model for optimising recovery activities after a disaster to maximise the resilience of the intermodal network subject to budget constraints.

Khademi et al. (2015) present a recent study that is particularly interesting because of its comprehensive approach to transport network vulnerability/resilience analysis for the case of a catastrophic event. They notice that most transport vulnerability studies deal with the pre-hazard phase and most often only consider a single link failure. In contrast, this study stresses the need to analyse changes in travel demand and behaviour in the response phase of a disaster when emergency trips have to be prioritised and many roads are impassable. Their specific case is a possible earthquake in Tehran in Iran. They propose a comprehensive conceptual model for transport network vulnerability analysis. In performing this analysis, a number of brainstorming sessions involving stakeholders and researchers were held to find out the best way of planning for the post-disaster situation of travelling. To support this planning work, formal network analysis methods were also developed.

Although there is an increasing interest in studying all aspects of transport system resilience including preparedness, response, recovery and adaptation, the amount of published research is still limited compared with that on vulnerability, which is usually more focused on what can be done in the pre-hazard phase. On the other hand, transport vulnerability research has reached a certain level of maturity and sophistication. However, there are only limited signs that vulnerability research has actually been used by planners, practitioners and operators. For the research to be more generally applied in practice, it is in the authors’ opinion necessary to strengthen the cross-disciplinary collaborations with responsible authorities, operators and other stakeholders for mutual learning and transferring of knowledge. The earthquake study by Khademi et al. (2015) and some international studies on climate change impacts on the transport system may be role models in this respect (e.g., National Research Council, 2008; the EU projects WEATHER and MOWE-IT).9 Without such transdisciplinary activities the full potential of the research reviewed in this article may not be achieved.

6. Concluding remarks

Research on vulnerability and resilience of transport systems is now a well-established field, and a large number of articles have been published in the international research literature. At the conceptual level, there is still no consensus on the precise definition of vulnerability and resilience and their relationships to related concepts such as risk, fragility, survivability, reliability and robustness. There are two distinct traditions in vulnerability studies. One has its roots in graph theory and

---

studies the vulnerability of transport networks based on their topological properties. The other tradition attempts to represent the demand and supply side of the transport system and travellers’ responses to disturbances and disruptions by the use of more or less sophisticated models of the transport system. One useful property of the first approach is that it only requires definitional network data and allows for detailed analysis of different attack strategies and also comparisons with other very different kinds of networks. The second approach requires extensive data about demand and supply aspects of the studied transport system and the availability of models for simulating the consequences of disruptive events. The computational efforts of the latter approach are larger; on the other hand, it provides a more complete description of the consequences of different vulnerability scenarios for the travellers and society.

The concept of resilience is intended to capture a system’s capacity to maintain its function after a major disruption or disaster. It may also include the rapidity with which the system returns to a state of normal operation after such an event. Compared to vulnerability, the literature on transport system resilience is more limited, with relatively few studies related to the post-disaster phases of response and recovery. Transport vulnerability and resilience research has accumulated considerable knowledge of large potential practical usefulness. In the authors’ opinion, cross-disciplinary collaboration between researchers, authorities, operators and other stakeholders could be a way of achieving the mutual learning and transfer of information that would enable this knowledge to be transformed into practical strategies to strengthen the resilience of the transport system.

Acknowledgements

The authors are grateful for comments from participants at the NECTAR International Conference, Ponta Delgada, Azores, 2013, and in the Second Workshop of Deans Forum on Resilience Engineering, MINES ParisTech, Paris, 2013. The authors are also grateful to two anonymous reviewers for valuable comments and suggestions. The first author gratefully acknowledges financial support from the Swedish Transport Administration and the Swedish Fortification Agency.

References


