Assessing rail transport network performance and reliability

R. Raicu & M. A. P. Taylor
Transport Systems Centre, University of South Australia, Australia

Abstract

Network performance in general - and in the computing and telecommunications areas in particular - is a topic of major interest. Much of this research is not in the public domain, but this is beginning to change as international research teams become aware of common problems and issues. In any event, rail systems have certain special characteristics which limit the transferability of general network methods to rail problems or which require additional work to develop relevant interpretations and implementations for use in railway systems analysis. The paper presents results of a research project aimed at developing decision support tools for assessing rail network performance and reliability.
Keywords: rail transport network, network reliability and vulnerability, congestion, delay.

1 Introduction

When analysing network performance an integrated approach considering network infrastructure, operation and transport demand is essential.

The capacity of the transportation network can be evaluated by various measures such as the travel times and the extent of congestion. Network flows are influenced by abnormal events that affect network characteristics and capacity, like disasters, accidents, construction or repair. Ideally, networks should be designed so as to cope with normal fluctuations by offering alternative paths, but planning for abnormal events is much more difficult.

In systems engineering, reliability may be defined as the degree of stability of the quality of service which a system normally offers. In the face of increasing user demands for high levels of service, system reliability is becoming

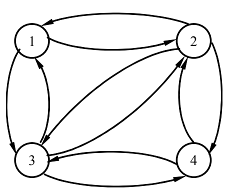
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/CR060081

increasingly important in the planning, construction and operation of transportation networks.

2 Model rail systems as formal networks

The rail transport network is a complex system made of numerous elements with functions determined in the transport process. The description of the rail transport network cannot be limited just to its elements; similar to other complex systems, its structure is very important. The concepts of graph theory are used to formalise the structure of the rail transport network.

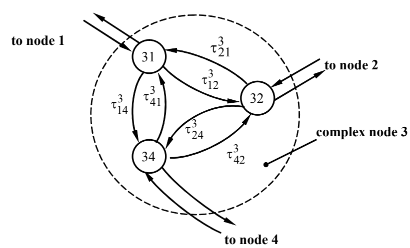
The transport infrastructure network can be presented as a graph (K, D), with K nodes and D arcs. The nodes of the network are just those points from which diverge or to which converge at least three arcs, and in which it is possible to transfer from one arc to any of the other two. Both arcs and nodes are characterised by certain attributes which can only be defined in correlation with the characteristics of the transport means they were designed for. Each arc (paths of the network) can be made of one or more homogenous sections characterised by length, and maximum, respectively minimum speed the transport means can operate.



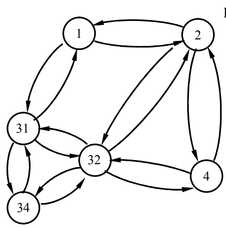
$K = \{1,2,3,4\}$

$D = \{(1,2),(2,1),(1,3),(3,1),(2,3),(3,2),(2,4),(4,2),(3,4),(4,3)\}$

a). Initial graph (node 3 is a complex node)



b). Complex node 3 transformed in simple nodes and corresponding links (arcs)



$K = \{1,2,4,31,32,34\}$

$D = \{(1,2),(2,1),(1,31),(31,1),(2,32),(32,2),(2,4),(4,2),(4,32),$
$(32,4),(31,32),(32,31),(31,34),(34,31),(32,34),(34,32)\}$

c). Modified graph

Figure 1: Obtaining the modified graph of the network.

The transit times cannot be neglected. These nodes, with their significant transit time are considered as being complex nodes. Figure 1 presents an example.

3 Particularities of rail transport networks

3.1 Basic principles

On the relatively simple graph of the network the set of paths (itineraries) which link any two nodes could be indicated. In describing the network the choice of nodes and arcs is essential. Different problems require different degrees of detail regarding the technical state of the network. Some examples of problems to address are: planning economic transport links, freight transport; planning links between urban concentrations, passengers transport; distribute traffic between transport nodes; determine traffic intensity on the sections of the network; planning investment works to increase transport capacity; evaluating resources use to accomplish the forecasted transport task; planning the technical and economical operating indicators of the transport mode and transport system as a whole.

3.2 Structure and contents of the information on the technical state of the rail network

The way the information is organised influences its efficient use for solving concrete problems. In selecting the organisation of information one has to bear in mind that certain parts of the information have to be easily accessible and partial modifications should not require the radical restructuring of the whole system. Thus a hierarchical structure of the information is preferred. For the rail network a three level hierarchical structure is recommended (see Figure 2).

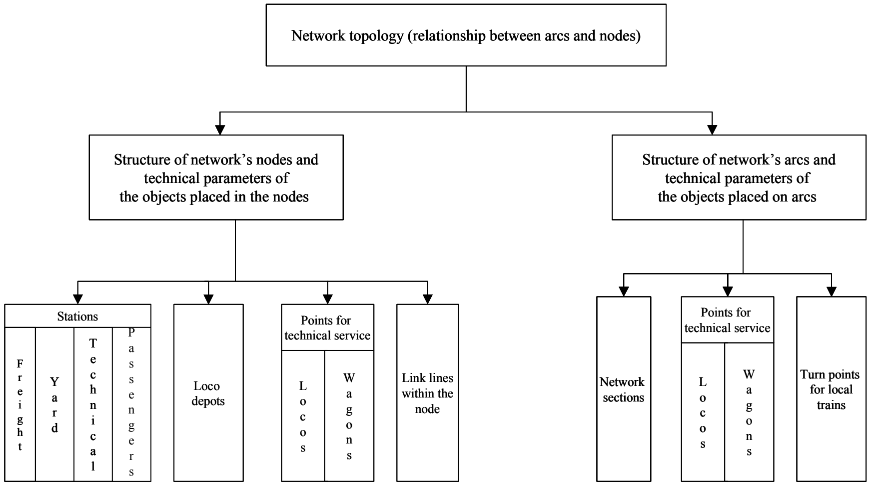


Figure 2: Structure of the information on rail network state.

The first level (superior) of information characterises the whole rail network and contains the list of arcs and nodes and the table of correspondence between them.

The second level of the hierarchic structure comprises information about the managed objects and the interactions between them. It refers to the objects in the nodes. Depending on the type of problem to solve different level of detail on the transportation process is necessary. The nodes of the network generally have a complex structure. One can identify in their structure two types of elements: objects and links between objects. In various transport models just a part of these objects is involved. For example, information about the location of the freight stations is necessary for the detailed evaluation of the material and financial expenses of the transportation operations. Technical stations and yards are included in the models of train formation on the network. The location of passenger stations and technical passenger stations is necessary to describe the passenger trains operation. Information on depots, locos and driver rosters are necessary to evaluate locos and drivers activity.

The third level of the hierarchic structure comprises information on the constructive parameters of the objects. The structure of the parameters allows the most synthetic characterisation of the objects: transport capacity and the cost of technological operations. To solve some planning problems, other characteristics might be necessary as well. Thus, in time the structure of the parameters changes. This means that the form of presentation of the information at the third level has to allow for changes.

4 Reliability of the infrastructure networks

As part of the transport infrastructure, the transport network, similar to other technical infrastructures of the society, can be formally represented by a nonplanar graph with nodes and arcs.

The calculation of reliability between two nodes of the network (connectivity property) is very complex for large networks due to the multiple links (Dupuy [1]). Various mathematical and heuristic models have been developed, and also computer simulation (Bell and Iida [2]). The commonality of all these methods is the fact that in order to establish the probability of having a link between two nodes a structure function is used

$$\Phi(\mathrm{X}) = \begin{cases} 1, & \text{if nodes are linked} \\ 0, & \text{if not} \end{cases}$$

where $\mathrm{X} = (x_1, x_2, \ldots, x_k)$ is a vector with its components representing the operational capacity of different arcs (considering the nodes are ideally operating or that their failures are transferred to the incidental arcs) which represent the link between the points considered, respectively

$$x_i = \begin{cases} 1, \text{if the arc is operational,} \\ 0, \text{if the arc is not operational.} \end{cases}$$

Bell and Iida [2] demonstrate that any network can be formally represented with the aid of paths and minimum sections. By path we understand the succession of arcs which ensures the link between the initial and final point, and by section the combination of arcs which when removed from the network causes the interruption of the link.

The minimum path is that path in which the failure of any of the arcs leads to the interruption of the link.

Each minimum path, A_j can be related to the function $\alpha_j(\mathrm{X}) = \bigcap\limits_{i \in A_j} x_i$, which takes the value 1 if all the arcs of the minimum path are operating normally, in other words if for all $i \in A_j$ the condition $x_i = 1$ is simultaneously achieved.

The minimum section is that section for which the rehabilitation of the operational capacity of only one arc leads to the rehabilitation of the operational capacity of the whole network. Each minimum section B_k can be related to the logical function $\beta_k(\mathrm{X}) = \bigcup\limits_{i \in B_k} x_i$, which takes the 0 value if all the elements belonging to the section fail, and takes the 1 value if at least one element is operating normally.

In general, the network can have several paths and minimum sections. Any network can be represented as a parallel connection of minimum paths or a serial connection of minimum sections. It could be that in a network one and the same element appears in two or more paths or minimum sections. In other words, some minimum paths or/and sections are dependent. This means that network reliability cannot be calculated based on a mathematical model built on the independency of events hypothesis.

It is however possible to evaluate the maximum and minimum limits of the probability of operation without rejection for any size network, with

$$1 - \prod_{j=1}^{Q} \left(1 - \prod_{i \in A_j} P_i \right) \geq P[\Phi(\mathrm{X}) = 1] \geq \prod_{k=1}^{S} \left[1 - \prod_{i \in B_{k\,j}} (1 - P_i) \right] \tag{1}$$

where P_i is the probability of operation without rejection of element i (e.g., in the transport network case, the number of daily hours without congestion divided by the daily hours the arc of the network is open to the traffic),

Q – number of minimum paths,

S – number of minimum sections.

It can be demonstrated Raicu et al. [3] that in order to simplify the calculations with expression (1) is possible to eliminate the minimum paths and sections with a great number of elements, without substantially altering the precision of the result.

The methods used to evaluate network reliability (irrespective of the nature of the flows) are different from the methods used to calculate the reliability of products. Transport networks, as oppose to telecommunication networks for example have few essential particularities. Some of these particularities are presented hereafter.

5 Particular aspects of transport network reliability

5.1 Reliability in relation to connectivity

The particularity consists in that the effects of congestion on the basis on which the good operation probabilities of the arcs are estimated are also sensed in the nodes of the network affecting the qualitative indicators of transport (time, security, regularity).

5.2 Reliability in relation to capacity

It is defined as being the probability of a transport network to ensure the successful transfer (according to the estimated indicators) of the assigned traffic flows, even when one or several arcs of the network have lost their functionality. It assumes that the reserve of capacity for each arc remains greater than the temporal non-uniformity of the traffic flow or, in a different formulation, or in fact in a different hypothesis – the loss of capacity due to the partial functional degradation of the network cannot be greater than the reserve of capacity.

This characterisation of reliability is necessary to evaluate the transport network. With respect to connectivity the network can ensure the link for any relation between nodes, but this does not mean that the integral transfer of flow is possible in all relations or in a particular one, if one or several arcs have lost their functionality.

A network with good reliability from the point of view of connectivity may not be reliable from the point of view of capacity. In order to characterise the network the concept of vulnerability Taylor and D'Este [4], Berdica [5] or that of operational reliability Raicu [6] has to be employed.

5.3 Reliability in relation to transit time

Coming back to the problem presented before and assuming those reserves of capacity would ensure the operational reliability of the network, it is obvious that the alternative routes suppose greater transit times than the initial ones, maybe even unacceptable for the users.

Therefore, when estimating transit time reliability admissible limits have to be prescribed.

Examining the predicted route as a succession of homogenous sections for which the available statistics provide data about the daily succession of traffic regimes is possible to build the transition diagram of the Markov process associated to the state of the homogenous sections.

In the simplest manner the states of the route encountered by a vehicle when moving in a certain relation can be binary quantified – congested C, respectively not congested, L.

If the probabilities of going from one state to the other are estimated the signal graph can be built (Figure 3) and can evaluate the probability of the route to be in the non congested state after a number of periods, if at the reference moment it was in the same state.

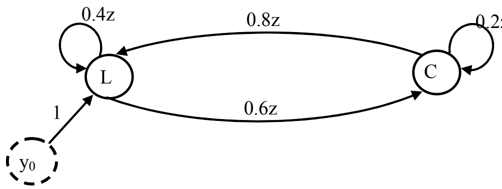


Figure 3: Signal graph associated to a route (the transforms Z of the passing probabilities are written on the arcs).

To solve the problem a dummy node y_0 has been introduced linked to L (“not congested” state) through a dummy arc of transmittance 1.

Mason’s rule is used to evaluate the transmittance. According to this rule, the transmittance between two nodes i and j of the graph is:

$$T_{ij} = \sum \Pi_k \frac{\Delta_k}{D}$$

where Π_k is the transmittance of the direct route k between i and j,

$$D = 1 - \sum_i L_i + \sum_{i \neq j} L_i \times L_j - \sum_{i \neq j \neq r} L_i \times L_j \times L_r$$

where L_i, L_j, L_r are the transmittances of the loops of the graph, if $L_i \times L_j \times L_r \times ...$ are achieved only for the non-adjacent loops (those that do not have a common node),

Δ_k has a similar expression as D, except it refers to the loops left after the nodes of the direct route k of Π_k transmittance have been eliminated.

For the graph In Figure 3, D=1, $\Pi_{y_0-L} = 1$ with $\Delta_1 = 1 - 0.2 \cdot z$. Because $D = 1 - \left(0.4z + 0.2z + 0.48z^2\right) + 0.08z^2$, results

$$T_{y_0-L} = \frac{1-0.2z}{1-0.6z-0.4z^2} = 1 + 0.4z + 0.64z^2 + 0.544z^3 + 0.6824z^4 + 0.62704z^5 + 0.649184z^6$$
$$+0.6403264z^7 + 0.64386944z^8 + ...$$

The coefficients of z^k represent the probability that after a number of k periods the system (analysed route) is in “not congested” state, L. It can be noted that the value of the coefficients of z^k after a number of periods start to stabilise which confirms the fact that state L is a recurring event with a finite average recurrence time.

5.4 Reliability – propriety of the networks

The networks of the technical infrastructures of the society, in general, and the transport networks in particular are topologically and structurally characterised through properties like connexity, connectivity, homogeneity, isotropy, nodality, etc. These properties, as oppose to others like length, density, etc outline the links between the networks and the territorial system they serve [1, 3].

Using a simulation model (written in Visual FoxPro) Raicu et al. [3], the reliability matrix in relation to connectivity is obtained, even for large networks in relatively short computational time.

6 Analysis of rail transport network performance

The specific aim of this research is to develop a suite of decision support tools for assessing network performance and reliability. The tools necessary to analyse and assess existing rail networks, and to plan and evaluate proposed infrastructure improvements and access provision using congestion pricing models will be developed. The outcome will be a set of high-level tools for strategic planning of the long-term development of rail networks.

A particular emphasis is the study of congestion and methods to identify the extent and spread of congestion in a network.

6.1 Congestion/delay in the rail network

The relationship between delay and congestion has been explored. We consider that congestion exists if the passage of one train causes delays to other trains in the system. The level of congestion would then depend on the extent of the delays. Some congestion may be inevitable as a consequence – it is inefficient to design a railway system in which each train has its own, uniquely assigned track. It will be determined if the level of congestion is acceptable or not.

The first observation that should be made is that congestion generates delays, but not all the delays are a consequence of the congestion. Initial delays may arise through some exogenous factor or event, such as a mechanical breakdown. The consequence of these delays may then be the occurrence of congestion.

6.2 Congestion maps

The main sources of train delays are functions of: timetabling; infrastructure or train performance; poor scheduling.

The University of South Australia has developed train scheduling software, internally known as TPOD (Train Plans On Demand) that is able to produce train plans quickly and effectively (see Figure 4). TPOD takes a set of train requests on a particular rail network and minimises the aggregate delays whilst moving the trains from their origins to their destinations Eitzen [7]. By using TPOD to generate the train plans, the poor scheduling factor can be alleviated. If we assume timetabling is largely market driven and therefore largely fixed then delays are an indicator of infrastructure bottlenecks. Delays can be costed (e.g. on an annual basis) yielding a dollar value for capital expenditure justification. Then, the infrastructure can be upgraded to reduce bottlenecks and retest via TPOD to quantify the impact of new infrastructure and confirm cost/benefit (see Figure 5).

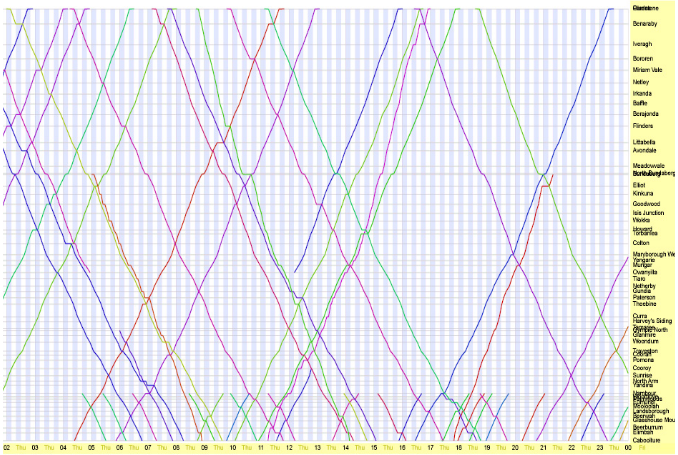


Figure 4: Example of TPOD model output.

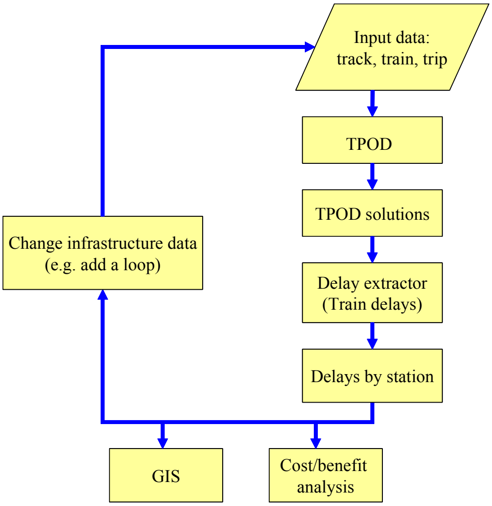


Figure 5: Delay and congestion analysis- system architecture.

A procedure to extract delay information from TPOD (Train Plans on Demand) was designed and programmed. The procedure captures train delays from a TPOD network solution and can display them by station, by time for individual trains or any combination of trains. (Figure 6) The delay data can be exported to a spreadsheet or other planning software for cost/benefit analysis or used to simulate the network working under new conditions for planning purposes and congestion analysis.

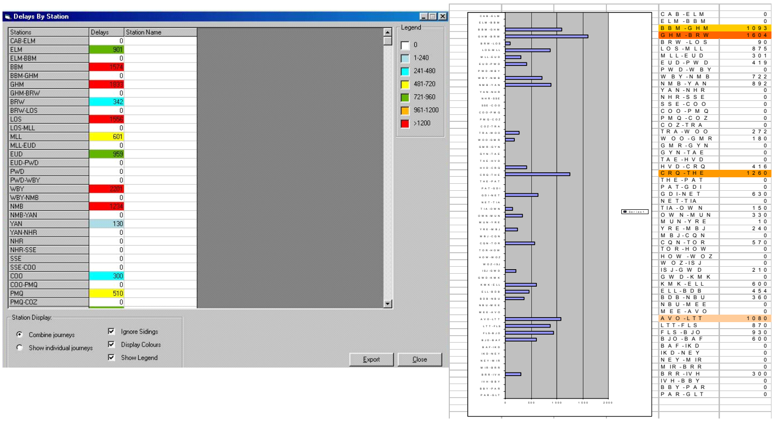


Figure 6: Example of delays by station (output from the delay extractor).

7 Conclusions

This research will lead to the development of expertise in the formulation and application of methods for the analysis of network reliability, identification of vulnerable locations, and assessment of level of performance of intact and degraded networks. This will allow planner and managers to determine the sensitivity of overall network performance to variations in local capacity and local operations, and to identify congestion points and their impact on network performance. These new tools will assist managers and planners to develop insights into network performance under different operating scenarios and objective evaluation of proposed infrastructure development proposals.

References

[1] Dupuy, G., *Systèmes, réseaux et territoires: principe de reseautique territorial.* Paris, Presses de l'ENCP, 1985.

[2] Bell, M.G.H., Iida, Y., *Network Reliability. Transportation Network Analysis.* John Wiley & Sons, pp.179-192, 1997.

[3] Raicu, R., Raicu, S., and Popa, M., The influence of transportation networks reliability on city logistics. In *Proceedings of the IV International Conference on City Logistics*, 12-14 July, Langkawi, pp 117-131, 2005.

[4] Taylor, M.A.P., D'Este, G.M., Critical infrastructure and transport network vulnerability. *The Second International Symposium on Transportation Network Reliability (INSTR)*, Christchurch, New Zealand, August, pp.96-103, 2004.

[5] Berdica, K., An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy*, **9(2)**, pp.117-127, 2002.

[6] Raicu, S., Asupra fiabilitatii sistemului de transport feroviar. *Revista Transporturilor*, no. 5-6, pp.145-151, 1991.

[7] Eitzen, G., TPOD Documentation Draft, University of South Australia, 2004.