Reliability and Safety Models of Transportation Systems - a Literature Review

Franciszek J. Restel
Wroclaw University of Technology, Wroclaw, Poland

Abstract: Transportation systems form the backbone of economy and play an important role in society. Because of the far-reaching effects of disruptions on these systems (social, economic, national defense), they are classified as critical infrastructure systems. Reliability researches on various elements of transportation systems were carried out since the mid-twentieth century. The focus was on vehicles and their components. Infrastructure is an important component of system, in addition to vehicles. The highest level of complexity is characterized by the railway infrastructure. It is natural, that this led to a number of models describing selected issues. In recent years, much attention has been paid to critical infrastructure systems. There have been numerous proposals for the use of graph models in the analysis of resilience and vulnerability of transportation systems. There are many groups of models describing reliability. Some models contain reliability factors only fragmentary. This paper presents experience with reliability models, were tested in research work on the railway transportation system. The inference is not limited to railway transportation system, but generally relates to land transportation systems. The review includes also own models that are dedicated to describing reliability of fixed-track systems, in which processes are determined by a timetable.

Keywords: Railway, Transportation Systems, Reliability, Safety, Models.

1. INTRODUCTION

There is a large ambiguity in nomenclature in the field of the rail transportation system as far as the connecting areas of such terms as reliability, transportation systems and railway engineering are concerned. The most common examples of inconsistencies can be found in the studies of transportation systems that discuss elements of reliability. This problem has been revealed after a preliminary literature review.

Public transport companies show a common tendency to narrow the facets of reliability down to proper execution of transportation tasks - i.e. punctuality [58]. Studies of this aspect are limited to modelling of arrival and departure times [4,58]. Furthermore, punctuality is also the sole aspect mentioned in papers presenting models used for selection of transportation modes in public transport services [5]. The term “travel time reliability” is introduced, denoting punctuality of inter-connected services at transfer nodes. Travel time reliability is, first and foremost, mentioned in analyses that refer to passengers [40]. In [44] travel time reliability is defined as probability of travelling a certain route in a period of time that is shorter than or equal to the assumed travel time.

Most often, however, reliability is understood as punctuality, i.e. an attribute referring to process execution in accordance with the established schedules. Chen [12] enumerates the most commonly appearing terms in the context of reliability in the railway system, by listing the following definitions:

- Reliability - probability of performing a required function by an object under given conditions and in a specified period of time,
- Railway transportation services reliability - participation of trains actually starting at a given time (including reserve trains) in the total number of all scheduled trains in a theoretical timetable. This attribute does not take into account the size of (original and secondary) disruptions,
- Punctuality - participation of timely train arrivals at every station in all arrivals scheduled in a timetable.

Punctuality is considered one of the most important measures of reliability of transportation system processes [28,47,49,67]. In practice, punctuality is not measured for every train arrival and departure to and from every single station, but it is only verified at the destination station of a given run [49].

Vromans [66], however, gives the following measures, as the most popular aspects in determining reliability of a rail transportation system:
- punctuality,
- punctuality of inter-connected services - concerns delays resulting from the loss of inter-connected train service,
- number of cancelled trains,
- average train delay,
- average passenger delay,

Factors that influence punctuality of trains in the rail transportation system were identified in [49,58]:
- number of passengers,
- degree of vehicle occupation in a train,
- use of traffic capacity,
- cancelled services,
- temporary speed limits,
- technical support of the infrastructure,
- train traffic organization.

The authors in [4] take into account specific ranges of delay in punctuality analyses. They quote two minutes as the border value. However, when presenting the results obtained from the study of an actual system, the border value was established as 5 minutes. Some scientific studies also propose the border value as 2.5-minute delays [48]. However, such values refer to rail transportation systems in highly developed societies (e.g. Norway). Such rail transportation systems have become more efficient as a result of significant society development. Along with the development of societies, new types of constraints are put on anthropotechnical systems. One such example are Scandinavian railway services with high punctuality standards and their adverse events research focusing, in fact, on the cases of suicides [53,54]. In [65] Vansteenwegen defines punctuality in the context of a train’s arrival at the destination station. At the same time it is stated that delays up to 5 minutes are a good result of system operation.

2. RELIABILITY IN RAIL TRANSPORTATION SYSTEM

From the very beginning of reliability studies in rail transportation system studies mainly focused on vehicles. Apart from testing mathematical models, operational data analyses were conducted concerning damage. The research was narrowed down to a statistical analysis and conclusions drawn therefrom. Rail vehicles consist of systems, assemblies, subassemblies, etc. which may be subject to failure resulting in the entire vehicle becoming non-operational. Thus when analyzing their reliability decomposition is performed in order to design the models more precisely [59]. Operational studies into vehicle failure rates are used, among others, to determine optimal inspection and repair intervals.

A crucial aspect in the railway engineering practice is determination of the state of infrastructure and relating safety improving actions to it. [3]. From such a perspective introduction of train speed limit is considered an alternative to technical service. Both the choice of when to introduce a speed limit and performance of technical service entail costs. A crucial group of studies in this aspect involve the so-
called Life Cycle Costs analyses [41]. In [23,29] decision models for determining the right time for technical service of infrastructure at minimum total costs have been designed.

The above-mentioned models take into account the costs of a planned technical service, speed limitation costs, however, they do not include any aspects of reliability. The costs related to adverse events are not considered. Another group of tasks performed in rail transportation system includes a dispatcher’s actions once disruptions occur. One of the possible methods is to regulate the speed by running trains in order to minimize unplanned stopovers (particularly perceived by passengers) and reduce energy consumption [16]. In this approach no structural process changes are proposed, but merely changes in process parameters. In fact, these aspects are related to re-organization of traffic after occurrence of disruptions whose aim is to minimize further propagation of disruptions [10,62,63].

A more detailed insight is provided by an analysis of consequences of original and the related secondary damage (excluding the traffic impact) [60]. Chen in [12] presented train services reliability and punctuality models. The train service reliability model is composed of three sections:

- reliability of technical facilities constituting the system (independence of events occurring in subsystems was assumed),
- interactions between the transport modes subsystem and other subsystems,
- intensity of restoring the traffic after interruption caused by an adverse event.

The authors [33] proposed a disrupted train traffic management support model. The model is based on the costs of rail traffic reorganization or cancellation of trains in the context of railway employees (engine drivers and train traffic service staff).

The probability of delay suppression is directly related to the so-called “resistant timetables” [36] and resilience [22], i.e. an ability of the system to regain functionality after an event. Vromans [66] narrows the term of rail transportation system reliability down to reliability of transport services. Time disruptions constitute the only aspect that is taken into account. The author identifies potential causes of disruptions, however, he assumes that the time disruptions have only one source. The paper focuses on max-plus algebra used for assessment of timetables (cf. [27,28]) in the context of time reserve in a timetable. A support model for decision-making during traffic control in the case of disruptions was presented in [1]. In fact, a crucial impact of infrastructure on propagation of disruptions was noticed. Decision variables include the moments when a k-event begins and finishes.

The previous groups show a tendency to narrow the subject down to one specific aspect. The next group of aspects has a completely different approach in comparison to the before-mentioned ones. They include research in which the impact of catastrophic events on system operation is analyzed. Railway services are in this respect considered Critical Infrastructure System (CIS), whereas conducted analyses focus on serious events with dire consequences. A CIS description contains graph models in which the most basic ones are modelling simply the relations between junctions. The more advanced models include traffic capacity of the edges, travel time, traffic control at junctions, as well as the mode of power supply [18,19]. In this aspect the term - reliability of railway network infrastructure traffic capacity - is introduced [72].

3. FUNCTIONAL RELIABILITY MODELS

In [71] Zamojski presents a functional reliability model of a discrete transportation system. A discrete transportation system is further defined as a relation in the Cartesian product of transportation task sets (theoretical and actually performed), means of transport, infrastructure, a dispatcher and time. The described time is the so-called network chronicle [71]. A resource set called functional system configuration is assigned for the purpose of input task execution. An input task is defined as a priori by a dispatcher based on available resources. An output task is the actual execution of a transportation task. Reliability characteristics are attributed solely to vehicles.
In [20] a functional reliability model of track availability control system was presented. The model is based on the following three layers:

- a traffic process along the track system,
- an operational process of train traffic control devices,
- a decision-making process in management of train traffic control devices.

Layer 1 concerns arrangement of tracks and vehicle flow with possible disruptions. The second layer (operational process layer) ensures ordering of routes into admissible and inadmissible at a given moment. It defines whether two train courses can be run at the same time without a risk of collision. The third layer is based on a decision model generating a series of decisions that control the processes based on available information from the two remaining layers. Three types of adverse events related to availability-identifying devices have been distinguished:

- “catastrophic” (accidental) damage causing irreversible changes in the attributes of the system, whose removal is possible only by replacing technical elements,
- disappearing damage, resulting from a temporary overrun of work parameters can be “cancelled” by an employee,
- wear-out damage, whose removal is possible only through replacement of technical elements.

In [51] a railway network simulation model was presented. Two railway lines with a common middle section were divided into equal sections. Every section can assigned one of the six states of degradation, where 0 is the incapacity state. The intensity values describing transitions between the states are constant. For the purposes of the model, traffic volume and time spent by a train for travelling each section were defined. The degradation model has been completed by possible line speed values (nominal speed 100 km/h, reduced speed - 80 or 60 km/h) by obtaining an 18-state model allowing for determination of a daily delay on a rail network.

4. MODELS BASED ON STATE-TRANSITION GRAPHS

One of the first reliability appraisals of rail vehicle elements were based on Markov processes, a reliability structure analysis and a fault tree. In the recent years Markov processes have been also applied in modelling of degradation and critical damage of the track structure [17]. The authors in [11] discuss an analysis of reliability and safety of rail traffic control device components in the aspect of safe failures. The paper focuses on an analysis of a junction reliability and safety analysis involving special microcomputers used for calculations, called transputers. To this end, a 9-state Markov model was developed. The model does not present the state of safety faults. The states with unrevealed errors or errors leading to similar output signals (potentially erroneous) form security threats.

In [70] a functional and reliability model of a discrete transportation system was presented. The model, similarly to the previous functional and reliability models narrows down in the aspect of reliability to a vehicle subsystem, since Markov processes have been used in it solely to describe reliability of vehicles. A series of assumptions regulating the functional part has been assumed in the model. The reliability part has been described by the Markov model in which states are defined by the number of damaged vehicles. Transportation system has been described by a matrix differential equation for n-recipient:

\[
\frac{d}{dt} p^{(n)}(t) = p^{(n)}(t)q^{(n)}
\]

where:

- \( p^{(n)}(t) \) - state probability vector,
- \( q^{(n)} \) - transition intensity matrix.

The authors [14] presented a Markov model of tram reliability systems. Six possible reliability states have been assumed for vehicles:

- \( TTB \) – availability state,
- \( ÑTP \) – partial availability due to damage to one out of two driving modules,
- $TT\bar{B}$ – vehicle failure due to damage to the braking module,
- $T\bar{T}B$ – vehicle failure due to damage to two drive modules,
- $T\bar{T}B$ – vehicle failure state due to damage to one drive module and a breaking module,
- $TT\bar{B}$ – vehicle failure state due to non-serviceability of two drive modules and a braking module.

The authors emphasized that the time required for a tram repair has a non-exponential distribution and assumed Erlang-3 distribution. Due to the connections with exponential distribution it is possible to represent Erlang-3 distribution in Markov model. One repair state with Erlang-3 distribution is replaced by three states described by the same exponential distributions. Due to the characteristics of the actual system it was assumed that at a given time, there can be maximum three fully or partially damaged vehicles. It was shown that for the analyzed tram transportation system the use of exponential distributions for a description of repair times shows higher availability of the system than in reality. The use of approximation of Erlang-3 distribution allowed for a more detailed modelling of the system, more convergent with actual results.

In [9] the Markov model was used to analyze the impact of various instances of faults on the availability of the railway transportation system. A four-state model in which the following states were specified has been proposed:
- (1) - availability state,
- (2) - safe failures occurred in the system which do not hold train traffic.
- (3) – hazardous errors occurred in the system - safety threat state with running train traffic,
- (4) – train traffic stoppage in the system in order to remove the threat.

Introduction of state (2) was caused by continuation of traffic in the actual system after occurrence of safe damages. Observations of the railway transportation system show that assessment whether damage is safe or whether it depends on the traffic supervisor. Thus introduction of state (3) is justified. The authors of the model did not predict the state of safety loss, by introducing interchangeable a threat removal state (4).

A frequent subject of tests is an element of railway transportation system. The example can be analyses of technical facilities belonging to the infrastructure. In [32] a 40-state model of reliability and safety of the railway crossing security equipment. Apart from intensity of damage and repairs, the model also includes detection of failures which was the reverse of the time between inspections. The first model was simplified to 21 states.

In [47] Markov processes were used to analyze reliability and safety of the combined transportation system (road-rail). In [46] a three-phase Markov model of intermodal transportation system was presented. The phase structures are identical. Introduction of phases results from the change of transportation mode in the case of intermodal transportation. The three phases represent road, rail and road transport. The author used the model for determination of availability function for intermodal transportation depending on the time proportion of individual phases. By analogy, Markov processes were applied for a multimodal transportation system [34]. The authors of the paper introduce the term of functional reliability defined as probability of supplying the right amount of load in the time not exceeding the exponential time.

Other areas in which Markov processes are used include:
- a train aggregation analysis in a selected subsystem [37],
- a reliability analysis of rail vehicles with their operational systems.

Markov chains have found application e.g. in:
- Bayesian analysis of collective bus transport routing [35],
- modelling of toll collection systems [68],
- technical service scheduling with the use of fault trees [45],
- in modelling of an aging system, including correction services [38].
5. SAFETY IN TRANSPORTATION SYSTEMS

Safety is related to maintenance of the system state that prevents occurrence of adverse events, such as [43]:
- death,
- body injuries,
- tangible property loss,
- natural environment loss.

In [61] a method of barrier identification based on the fault tree was presented. The method is based on the so-called Swiss cheese model, in which the holes must overlap so that an arrow can go through them (for safety failure to take place). A risk situation, which is a peak event is modelled by a classic fault tree with AND and OR logic gates. Once the tree is drawn up the first logic gates are searched for starting from the peak event. The event above a given OR gate is directly related to one barrier and used for determination of a barrier.

In the case of railway transportation system an event tree and a fault tree can be used in adverse event risk analyses. In [2] such an analyses was expanded by addition of risk influencing factors. The problem was shown on an example of a single-track line, for which a peak event was a collision of two trains coming from two different directions. Barriers which aimed at preventing occurrence of peak events were catalogued and then a tree of events leading to the barrier faults were drawn up. Operational risk influencing factors were attributed to the basic events.

The studies [6,7,56] explore security engineering in rail transportation system design. Articles discuss the issue of ERTMS (European Rail Traffic Management System) implementation, which in their structure also contain a unified European communications standard GSM-Rail. The problem of security at ERTMS implementation is all the more crucial if we take into account lack of experience in operation of such a system in conditions corresponding to the implementation (for Dutch railways in 2003). The basis for discussions [6] is risk identification for the implemented system. The [7] publication summarizes literature which introduces risk analysis components for a newly designed railway system (ERTMS). For security appraisal of the ETCS system, a slightly simplified ERTMS variant, Functional Hazard Assessment method was proposed in [55].

A crucial aspect in the assessment is identification of Safety Integrity Levels (SIL). In the draft of the European standard [24] a simplified SIL table was used for railway traffic control, communication, data processing equipment and electronic systems significantly influencing the safety of rail transportation. In [8] representing railway transportation risk levels in the form of a table was suggested (Table 1). In this way unacceptable risk, acceptable risk areas and a border area were obtained. Combinations of frequency and consequences in borderline risk areas were then used for drawing up a risk table based on the SIL table. The proposed table presents frequency of event occurrence and consequences divided into A to E. Group A represents events which can be classified as fail-safe. The remaining groups have been divided in terms of energy accumulated during the event:
- B - concerns consequences of events during shunting,
- C - concerns events at low linear speed values,
- D - events at medium linear speed values,
- E - concerns consequences of events at high speed values.

<table>
<thead>
<tr>
<th>Table 1: Permissible Risk Levels Table. Prepared on the Basis of [8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure rate, per hour of use</td>
</tr>
<tr>
<td>--------------------------------</td>
</tr>
<tr>
<td>&gt; 10^{-5}</td>
</tr>
<tr>
<td>10^{-5}</td>
</tr>
<tr>
<td>3 \cdot 10^{-7}</td>
</tr>
<tr>
<td>10^{-8}</td>
</tr>
<tr>
<td>10^{-9}</td>
</tr>
</tbody>
</table>

Probabilistic Safety Assessment and Management PSAM 12, June 2014, Honolulu, Hawaii
In [15] risk assessment of hazardous load transportation was conducted by using events/vehicle-kilometers as a measure of event occurrence intensity. In [39] a model used in a risk analysis of hazardous material rail transportation was presented. The risk is determined as the product of intensity of derailment of carriages used for transport of hazardous materials, operational work related to transportation of hazardous materials, conditional probability of hazardous material release after derailment and consequences of hazardous material release from the carriage.

6. HUMAN FACTORS

The literary sources point out that the human factor dominates during the occurrence of hazards [30,64]. In the case of standard rail traffic control devices a situation may occur in which the security system will have to be circumvented to enable further operation of train traffic. Emergency traffic operation is conditioned by improper system operation which constitutes a possibility for unreliability of security [21]. In [26] a model for human error occurrence probability during operation of rail traffic security system was presented:

\[ HEP = HEP_{gen} \cdot [(EPC - 1) \cdot EFF + 1] \]

where:
- \( HEP \) - human error probability,
- \( HEP_{gen} \) - intensity of human error probability in atypical situations,
- \( EPC \) - coefficient describing whether work conditions can contribute to a human error,
- \( EFF \) - coefficient describing efficiency of error-conducive conditions in error occurrence.

Increased behavioral and cognitive load has an impact on traffic safety (more than 90% of accidents in the rail transportation system occur after taking over the responsibility by a human [56], whereas disasters in transport occur in around 80% of cases due to a human error [31]. A human factor exists in the entire system not only in direct operation of trains by railway traffic control stations [69].

A weighty problem in the use of the rail transportation system is the SPAD (Signal Passed At Danger) phenomenon. The most common cause of this type of events is the machine operator’s error. The [25] source says that in 2011 more than 45% of accidents were caused by a train passing a stop signal in a dangerous way. SPAD events were examined qualitatively in [50] by using the so-called SHEL model. SHEL is an acronym which stands for Software-Hardware-Environment-Liveware. The software is understood as a collection of principles, guidelines of operation and other practices which define cooperation of the system elements. The hardware represents all technical facilities belonging to the system. The environment represents social, political and economic influence on the system from the outside. The liveware represents a collection of people in the system and their influence on the executed processes.

Due to the crucial impact of those events on occurrence of safety failure they are examined in detail by using various methods (e.g. Bayesian networks [42]).

In the context of a human factor a science dealing with safety culture in anthropotechnical systems should be noted. Models of the problems were synthetically introduced in [13].

7. SPATIAL REPRESENTATION OF THE RAIL TRANSPORTATION SYSTEM STATES

A model of rail transportation system can be presented as a process described by states, which is divided into a specified number of subsets. The most commonly distinguished subset is the availability and failure subset. Defining of states from the rail transportation system perspective is a complex problem therefore they were divided into constituent parts. When analyzing reliability of the system, its operational availability (in execution of transportation tasks) should be the first issue examined. In this aspect the set of states has been divided into subsets:
- availability of the system in execution of transportation tasks,
- partial availability of the system in execution of transportation tasks,
- unserviceability of the system in performance of transport tasks.

Consequences of incorrect operation are introduced in the second dimension:
- none,
- small,
- big.

In the third dimension organizational consequences of unwanted events (delays) are introduced:
- none,
- small,
- big.

The possibility of unwanted event occurrence and spreading of the consequences depends on the intensity of the system operation expressed as operational work [57] or transportation work (transportation load) [52]. As a result a super cube state defining model was developed, for state defining in four dimensions (Figure 1 and Figure 2).

Figure 1: State defining cube – three dimensions

Figure 2: Fourth dimension of the super cube

8. CONCLUSION

Research on the reliability and safety of rail transportation system can be divided into the following groups of models:
- Safety in rail transport:
  - risk models,
  - models describing the human factor,
  - models of the formation of accidents (including the failure of barriers)
- Reliability models based on ST graphs:
  - system models taking into account the dangerous failures,
  - system models taking into account only secure failures,
  - models of selected subsystems,
- Functional-reliability models:
  - system models with breakable infrastructure,
  - system models with breakable fleet,
  - models of breakable subsystem,
- Reliability models of transport processes:
  - models with random streams of applications and operating,
  - models with random unwanted events,
  - models with random travel time,
  - models with unwanted events occurring according to scenario.

Risk models are used to assess the occurrence and consequences of unwanted events. Models used for transport systems take into account the intensity of use of the system (volume of traffic), the probability of system corruption, improper use of the system (for example, against the technical rules of conduct) and human factors. Models describing human factors determine the unreliability of man in different circumstances the systems condition. The improper use of the system (conscious or unconscious errors) is taking into account and the process which will lead to the loss of security is examined. Models of accidents and analysis of the barriers, include physical dependence of trains movement, improper use of the system, human factors and risks. The process which leads to that failure of safety is determined.

Reliability models of rail transport system based on state-transition graphs (ST) with dangerous failures, include the failure of rolling stock or the general unreliability of the system. States of improper use of the system and the unreliability of security are modelled. Another group are functional and reliability models. These are primarily functional models of systems with selected reliability characteristics. Calculations of such models are most often performed using computer simulation. Analyzed system models take into account the failure of the infrastructure or of the rolling stock.

The next group are models of reliability of transport processes. Models of mass service system include random streams of applications and service. The unreliability of the system is contained in the random streams. Models with random driving time also apply to capacity and punctuality. The unreliability of the system is included in the randomness of driving time. In these models, interference in the process of transport are taken into account only in the form of delays.

Models of unwanted events according to the scenario relate primarily to examine the effects depending on the event. The effects are most often reported to the functions of the cost. These models are related to subsystems cycle cost analysis, using time reserves, losses due to interference and reorganize the movement. A common feature is skipping the mishap. In the other hand, the traffic and the intensity of use of the system are taken into account. Most models of critical infrastructures fall into this group.

In conclusion it should be noted that in previous studies of the reliability and safety of rail transport system developed models that consider individual characteristics. Security models do not take into account such features as motor disturbances or physical dependence of trains movement. Reliability models based on graphs ST do not include those features. In addition, limited to include only damages the individual subsystems or the system as a black box. Inventoried models of functional reliability show all functional aspects, taking into account only unreliability of one subsystem or component. Reliability models of the transport process focus on the assessment of the timetable, however, aspects of security and misuse are ignored.

Due to the nature of the rail transport system and the lack of appropriate models, it was developed a model of reliability and safety, which include the following features:

- The intensity of use of the system,
- Dependence on trains movement,
- The failure of the infrastructure and rolling stock,
- Providing disrupted the transport process,
- Improper use of the system,
- The human factor,
- Threats and security failure, process which will lead to loss of security.

References

[24] European Standard: EN50129