

# A Discrete Time Markov Chain approach to global risk analysis in railway transportation

P. Cesario<sup>1</sup>, N. Sacco<sup>2,3</sup> & M. Sciutto<sup>1</sup>

<sup>1</sup>*SiConsulting, Via Fieschi 25, 16121 Genoa, Italy*

<sup>2</sup>*NITEL, Via Isonzo 32, 00198, Rome, Italy*

<sup>3</sup>*DIMSET - University of Genoa, Via Montallegro 1, 16145, Genoa, Italy*

## Abstract

Since the earliest days of railways, as traffic was increasing, there has been the need for developing safety devices and procedures with the aim of avoiding or, at least, minimising the likelihood of circulation hazards and their consequences. In this framework, since railways are large and distributed transportation systems, characterised by a high and heterogeneous technological level and by great infrastructure complexity, the problem of assessing the risk is very difficult to cope with. Then, in this paper, the main characteristics of the SIGLO project, defining an analysis process assess risk is described.

*Keywords: railway safety, risk analysis, Discrete Time Markov Chain.*

## 1 Introduction

Since the earliest days of railways, as traffic was increasing, there has been the need for developing safety devices and procedures with the aim of avoiding or, at least, minimising the likelihood of circulation hazards and their consequences. In this framework, in the last years the European Union (EU) has begun to address the railways safety problem by promoting quantitative methodologies for safety level evaluation, and by introducing ad-hoc norms (see, for instance [1] for a detailed description of the point of view of the EU on the safety problem). Then, while in the past the scope of reducing consequences of accidents has been achieved by designing more robust rolling stocks, the challenge of minimising the accident frequency has been mainly faced by continuously introducing new technological devices and “behavioural norms” which ensure an high safety standard. In this



framework, since railways are large and distributed transportation systems, characterised by a high and heterogeneous technological level and by great infrastructure complexity, the problem of assessing the risk is very difficult to cope with. In effect, not only the failures of technological devices, but also human decisions and the relevant errors have to be taken into account.

In this paper, to cope with this problem, the main characteristics of the SIGLO (“Sicurezza GLObale”, meaning “Global Safety” in Italian) project are described. In such a project, realised in collaboration between the Italian railway network holder RFI (“Rete Ferroviaria Italiana”, meaning “Italian Railway Network”) and SiConsulting, the proposal of an analytic methodology to represent, by means of Behavioural Flowcharts (BF), the operations involved in train movement and management, and to assess the relevant hazards by means of Discrete Time Markov Chains (DTMCs). In doing so, each state of the DTMC has been associated with a set of state variables which allows to take into account the characteristics of the trains, such as their speed, the characteristics of infrastructures and of train traffic, that is, pointing out the technological level of the railway system and the trains density. In addition, the interaction with the surrounding environment, such as the influence of fog, rain, trees and bushes on signal visibility and so on, has been modelled. As regards the “causes” of the transition between the states of the DTMC, they have been represented by means of a probabilistic modeling approach which takes into account the probability of human errors.

This paper is organised as follows: after a brief introduction of a brief reference taxonomy in Section 2, in Section 3.1 the main characteristics of SIGLO are described. Then, in Section 3.2, the BF formalism is discussed. Therefore, after having recalled the main properties of Discrete Time Markov Chains in Section 5.1, in Section 5.2 the state-transition representation of a behavioural flowchart is analysed. Finally, Section 6 is devoted to some summarising conclusions.

## 2 Reference taxonomy

In order to clarify the proposal of this paper, it is worth recalling the definitions, hereafter considered:

**Signal Passed At Danger (SPAD)** is the term indicating a train passing a red signal or, more generally, a signal which does not give way to the train;

**Dangers (D)** represent attributes of processes which may potentially cause harm. In railway systems dangers mainly consist of SPAD, that is signal passed when on red;

**Hazards (H)** represent dangers that have a nature and magnitude that is directed to objects, persons, assets, etc;

**Risk (R)** is intended as the measure, under uncertainties, for the severity of a hazard. The risk is a function  $\phi$  (to be determined) of undesired events or accidents frequency  $f$ , and their relevant consequences  $C$ , i.e.  $R = \phi(f, C)$ .



### 3 The SIGLO methodology

In this section, the main characteristics of SIGLO will be described. More particularly, the attention will be focused on the project specifications, on the behavioural flowcharts modelling formalism, and on the risk computing procedure.

#### 3.1 The project specifications

As introduced above, the main aim of SIGLO is to provide a quantitative methodology for computing the safety risk related to train circulation taking into account both the train and infrastructure safety devices, human actions, and possible human errors.

In detail, the purpose of such a project is to provide a tool for:

1. Improving the accuracy in economic evaluation of timeslots in the timetable of different lines;
2. Computing the risk of railway circulation, and thus evaluating the safety level of railway lines due to their technological setup;
3. Demonstrating the adequacy of the implementation of different technologies, depending on the different circulation conditions and on the train frequency;
4. Demonstrating the continuous improvement of railway safety.

In order to do so, a functional description is used for describing all the possible operations that a railway system is able to do. Then, from this point of view, a railway system may be defined at a general level as follows.

**Definition 1.** A railway system is a transportation system which realises the *getting on/off of passengers*, the *load/unload of freights* and the relevant *transportation* by means of trains travelling on railways from an origin towards a destination in an a-priori determined time interval, without “harming” them.

Therefore, in such a framework, any railway operation/function is characterised by a sequence of:

**actions**, that is the behaviours followed by the human actors, or even by a technological devices fulfilling a-priori defined specific rules;

**events**, that is all the occurrences which do not depend directly on the actors' actions, by might be a consequence of them.

Then, any function is modelled with the aim of pointing out the actor performing actions or events, their sequence, their consequences, and the involved tools or facilities. Moreover, the functional representation is able to show both the events/actions trajectory or the state sequences. Finally it is worth saying that the considered functions have to be able to put in evidence the:

**nominal behaviours**, that is all the normal a-priori identified behaviours representing railway “physiological” operations;

**degraded behaviours**, that is all the “pathological” behaviours which are, anyway, predictable and correctly managed by actors;



**faulty behaviours**, that is all the “pathological” behaviours which may be unpredictable or, even when predictable, not correctly managed by operators.

Finally, the proposed representation methodology has allow hierarchical representation of the different functions, that is to be able to represent the railway operations at any of the following levels of details:

**Function level:** This is the most general level, consisting of the function list;

**System Level:** At this level, the model of the different operations is derived with a top-down process, that is regardless to the real implementation of railway safety technological devices and norms;

**Subsystem Level:** At this level, the model of the different operations is derived with a bottom-up process, that is starting from the real implementation of railway safety technological devices and norms;

**Application Level:** This level is the application of the subsystem models to a particular case of analysis, consisting in the set of functions needful to perform a specific “mission”, such as for instance the transportation of group of people  $X$  from the origin  $A$  to the destination  $B$ .

### 3.2 Behavioural flowcharts

The above consideration on what a behavioural model should be able to manage, leads to introduce flowcharts as the simplest representation of railway processes. In fact, such a graphical and modular formalism, is sufficiently general to be able to take into account all the elements and characteristics above described. Moreover, any flowchart process is built by means of blocks coming from a limited set, without depending on the considered representation level. In facts, it is possible to model any kind of railway process involving different actors, only by means of the blocks representing a generic process (graphically depicted as a squared box) and a decision (graphically depicted as a diamond box), respectively.

## 4 Flow-chart modelling a train approaching a signal

With the aim of clarifying the above introduced modelling approach, in this section the flowchart representing the locomotive driver actions and behaviour is presented. In order to do so, consider the simple example of a train approaching a red signal, which has not been passed by the train. The relevant signal layout, as it normally is in Italian Railways, is depicted in Fig. 1.

In such a representation it is worth noting the red signal and the yellow “warning” signal announcing the red one. Then, in this case, due to the norms presently in force in Italian Railways, the locomotive drivers, once having seen the yellow warning signal must start slowing down the train speed, in order to stop it before the red signal. Note that, due to the low braking power of trains, when a train approaches a red signal at a normal circulation speed  $V_{\text{normal}}$ , the locomotive drivers are normally not able to stop the train before the signal itself.

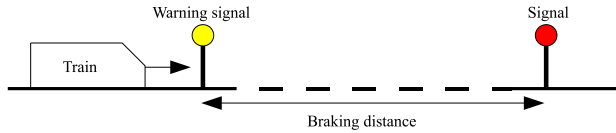


Figure 1: Signal state in the considered example.

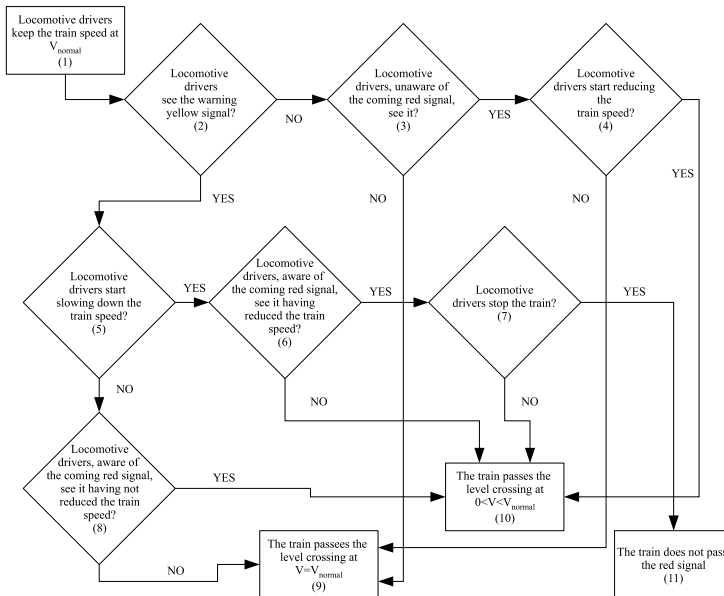


Figure 2: Flowchart representing a train approaching a red signal and the relevant warning yellow signal.

Then, the relevant flowchart representing the possible actions of locomotive drivers is depicted in Fig. 2.

In this model the final conditions consist of the safe behaviour or the unsafe Signal Passed At Danger (SPAD) condition. More precisely, the final conditions are given by the block:

- 9, which represents the worse SPAD occurrence, that is a SPAD characterised by a signal passed at the train speed typical of the normal “free” circulation  $V_{normal}$ ;
- 10, representing a SPAD, although the train passes the signal at the speed less the one characterising the normal circulation  $V_{normal}$ ;
- 11, which represents the safe condition of red signal not passed.

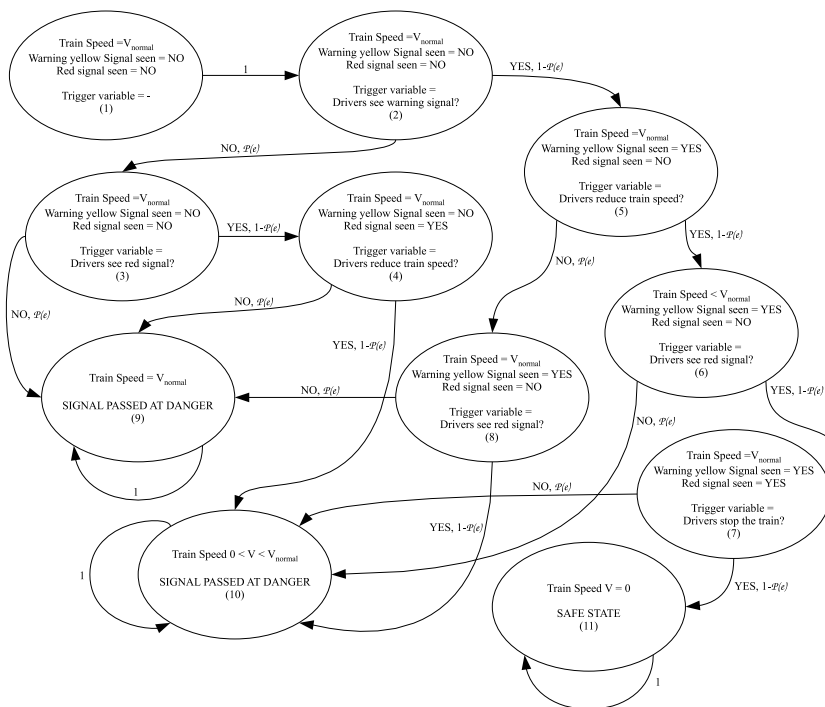


Figure 3: State-transition model of the behavioural flowchart of Fig. 2.

Once built the flowchart representation, the probability to reach the above mentioned states may be computed by building the relevant state-transition diagram, as described in the next section.

#### 4.1 State-transition representation of behavioural flowcharts

The above described flowcharts are not suitable for analytically computing the probability of reaching an unsafe condition. To cope with this problem, a state-transition diagram biunivocally corresponding to a behavioural flowchart. In effect, although representing a discrete system by means of a such formalism is less immediate with respect to the above mentioned sections, it allows one to easily compute the probability of reaching some particular states. For a more detailed description of the above mentioned formalism, the reader may refer to [2].

Then, consider the state-transition model of Fig. 3 representing the behavioural flowchart of Fig. 2. In such a representation each state, depicted as an ellipse, correspond to a block of the behavioural flowchart. Then, as regards the transitions exiting from any state, they represent the action or error leading towards another

state. In addition, each transition is associated with the probability of the relevant action or error.

## 5 The hazard probability assessment

Once described a railway function as a state-diagram, the probability to reach a particular state may be obtained by noting that it fulfils the Markov conditions. Such a consideration, in fact, allows one to apply all the Markov Chains' analytical background and properties to compute the above mentioned probabilities.

In order to do so, in the next sections the basic definitions of such a formalism are recalled and the probability to reach states 9, 10, and 11 of the flowchart of Fig. 2 is then computed.

### 5.1 Basic on Discrete Time Markov Chain

In this section, Discrete Time Markov Chains are briefly recalled. DTMCs are a graphical-analytical formalism for representing Discrete Event Systems (DES), that is a class of systems characterised by a discrete state space and a set of admissible events which drive the DES dynamics (see for instance [3] for details).

For what concerns the main definitions and properties of DTMCs, consider the following definition.

**Definition 2.** Consider a DES  $S$  and its discrete state space  $X$ ,  $\dim(X) = m$ . Then, let  $\xi_n = x_n$ ,  $x_n \in X$  be the state of  $S$  after  $n$  events, that is after  $n$  state changes. Then,  $S$  is a DTMC if and only if the probability of being in a generic  $\{\xi_{n+1} = x_{n+1}\}$  after  $(n + 1)$  events fulfils the Markov property

$$\begin{aligned} \mathcal{P}\{\xi_{n+1} = x_{n+1} | \xi_i = x_i | \xi_{n-1} = x_{n-1} | \dots | \xi_0 = x_0\} \\ = \mathcal{P}\{\xi_{n+1} = x_{n+1} | \xi_n = x_n\} \end{aligned} \quad i = 0, 1, 2, \dots \quad (1)$$

indicating the probability that the state  $\xi_{n+1}$  assumes a certain value  $x_{n+1}$  after  $(n + 1)$  events depends only on the probability that  $S$  is in the state  $\xi_n = x_n$  after  $n$ , and not on the whole state trajectory  $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$ .  $\square$

The second term of Eq. (1) is said *one-step transition probability* and may indicated by

$$p_{i,j} = \mathcal{P}\{\xi_{n+1} = x_i | \xi_n = x_j\}, \quad \forall x_i, x_j \in X.$$

Therefore, indicating with  $\pi_i(n) = \mathcal{P}\{\xi_n = x_i\}$  the probability that  $S$  is in the state  $\xi_{n+1} = x_i$  after  $(n + 1)$  events, and keeping in mind the total probability equation, it is possible to state the relation

$$\begin{aligned} \pi_i(n + 1) &= \sum_{j | x_j \in X} \mathcal{P}\{\xi_{n+1} = x_i | \xi_n = x_j\} \mathcal{P}\{\xi_n = x_j\} \\ &= \sum_{j | x_j \in X} p_{i,j} \pi_j(n). \end{aligned} \quad (2)$$



Finally, if  $\pi(n)$  (resp.,  $\pi(n+1)$ ) is the vector gathering the probabilities  $\pi_i(n)$  (resp.,  $\pi_i(n+1)$ ),  $\forall i | x_i \in X$ , it is possible to rewrite Eq. (2) as the vector

$$\pi(n+1) = \mathbf{P}\pi(n). \quad (3)$$

being  $\mathbf{P}$  the *transition matrix*

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & \dots & p_{1,m} \\ \vdots & \ddots & \vdots \\ p_{m,1} & \dots & p_{m,m} \end{bmatrix}. \quad (4)$$

Finally, by iteration the Eq. (6) it is possible to state the dynamic equation

$$\pi(n+1) = \mathbf{P}^n \pi(0), \quad (5)$$

where  $\pi(0)$  is the initial probability vector. Such a relation is particularly useful for computing the probability of the final states, that is

$$\pi(\infty) = \lim_{n \rightarrow \infty} \mathbf{P}^n \pi(0), \quad (6)$$

## 5.2 DTMC model of behavioural flowchart

In this section, it is shown how to use DTMC for computing the hazard probability introduced in Sec. 3.2. Then, in order to compute probability to reach an unsafe state, the behavioural flowcharts have to be represented by a state-transition diagram. In doing so, any element of the flowchart is transposed into a state, whereas the arrows of the flowchart are transposed into transitions. As said, such a diagram is, in effect, a DTMC fulfilling the property of Eq. (1), so that the above mentioned probability can be computed by means of Eq. (3).

In such a DTMC, the probability error  $\mathcal{P}(e)$  represent the probability that both the locomotive drivers make an error at a time, that is

$$\mathcal{P}(e) = \mathcal{P}_1(e) \cdot \mathcal{P}_2(e) = 10^{-3} \cdot 10^{-3} = 10^{-6}, \quad (7)$$

being  $\mathcal{P}_i(e) = 10^{-3}$ ,  $i = 1, 2$ , the probability that a single locomotive driver makes an error.

Therefore, the relevant transition matrix is

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathcal{P}(e) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{P}(\bar{e}) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathcal{P}(\bar{e}) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{P}(\bar{e}) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathcal{P}(\bar{e}) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{P}(e) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{P}(e) & \mathcal{P}(e) & 0 & 0 & 0 & \mathcal{P}(e) & 1 & 0 & 0 \\ 0 & 0 & 0 & \mathcal{P}(\bar{e}) & 0 & \mathcal{P}(e) & \mathcal{P}(e) & \mathcal{P}(\bar{e}) & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{P}(\bar{e}) & 0 & 0 & 0 & 1 \end{bmatrix},$$

where  $\mathcal{P}(\bar{e}) = 1 - \mathcal{P}(e)$  is the probability that drivers make the right choice.





By means of Eq. (3) it is easy to compute the probabilities

$$\begin{aligned}\pi_9(\infty) &= 3 \cdot 10^{-12} \\ \pi_{10}(\infty) &= 3.9999991 \cdot 10^{-6} \\ \pi_{11}(\infty) &= 9.99996000006 \cdot 10^{-1}.\end{aligned}$$

indicating that the hazard conditions characterised probability, whereas the probability to reach the safe state is

$$\begin{aligned}\mathcal{P}(SPAD) &= \pi_9(\infty) + \pi_{10}(\infty) \simeq 4 \cdot 10^{-6} \\ \mathcal{P}(safe) &= \pi_{11}(\infty) = 9.99996000006 \cdot 10^{-1}.\end{aligned}$$

## 6 Conclusions

In this paper, the main characteristics and the aim of the SIGLO project have been described. As said, such a project has developed a graphical/analytical tool, based in flowchart and DTMC representation, for computing the hazard probabilities. In effects, once computed these probabilities, SIGLO also provides the relevant risk by taking into account, for instance, the frequency of trains on lines and other relevant variables. Anyway, such aspects, which are based on “classical” event tree analysis, have not been described here.

## References

- [1] Høj, N.P. & Kröger, W., Risk analyses of transportation on road and railway from a european perspective. *Safety Science, Pergamon*, **40(20)**, pp. 337–357, 2002.
- [2] Cassandras, C.G. & Lafortune, S., *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, 1999.
- [3] Meyn, S.P. & Tweedie, R.L., *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993.

