Performance Evaluation of Train Moving-Block Control

Giovanni Neglia, Sara Alouf, Abdulhalim Dandoush, Sebastien Simoens,
Pierre Dersin, Alina Tuholukova, Jérôme Billion, Pascal Derouet

Performance Evaluation of Train Moving-Block Control*

Giovanni Neglia[1†], Sara Alouf[1], Abdulhalim Dandoush[2], Sebastien Simoens[3], Pierre Dersin[3], Alina Tuholukova[1], Jérôme Billion[3], Pascal Derouet[3]

[1] Université Côte d'Azur, Inria, France
name.surname@inria.fr
[2] ESME Sudria, France
dandoush@esme.fr
[3] Alstom Transport, France
name.surname@transport.alstom.com

Abstract

In moving block systems for railway transportation a central controller periodically communicates to the train how far it can safely advance. On-board automatic protection mechanisms stop the train if no message is received during a given time window.

In this paper we consider as reference a typical implementation of moving-block control for metro and quantify the rate of spurious Emergency Brakes (EBs), i.e. of train stops due to communication losses and not to an actual risk of collision. Such unexpected EBs can happen at any point on the track and are a major service disturbance.

Our general formula for the EB rate requires a probabilistic characterization of losses and delays. Calculations are surprisingly simple in the case of homogeneous and independent packet losses. Our approach is computationally efficient even when emergency brakes are very rare (as they should be) and can no longer be estimated via discrete-event simulations.

Keywords: Emergency brakes · Communication Based Train Control (CBTC) · European Train Control System (ETCS)

1 Introduction

In order to avoid collisions between consecutive trains traveling on the same track, the track is traditionally divided in fixed sections—called blocks—and only one train at a time is allowed to be in a given block.

The increasing demand for efficient mass transit transport requires to utilize railway infrastructure more efficiently. The improvements of train-sidetrack wireless communications, on board processing and actuators have made possible the introduction in the last 15 years of moving block systems, where blocks are dynamically calculated. Figure 1 schematically illustrates the two different approaches. The moving-block control can reduce the headway taking into account the actual distance between the trains as well as their speeds. It is being deployed as Communication-Based Train Control (CBTC) for urban mass transit system and is under consideration for next generation of European Train Control System (ETCS). This is referred as ETCS level 3 and is currently under standardization.

Moving-block systems require a continuous information exchange (detailed in Sec. 2) between an on board local controller, called the Carborne Controller (CC) and an external ground controller, called the Zone Controller (ZC) because it monitors all the trains in a given zone. Safety-critical messages are exchanged using standard or proprietary radio technologies. If no message is received during a given interval then the

*This is an author version of the 16-page paper that has appeared in the Proceedings of QEST 2016, Quebec City, QC, Canada, August 23-25, 2016.

†Corresponding author.

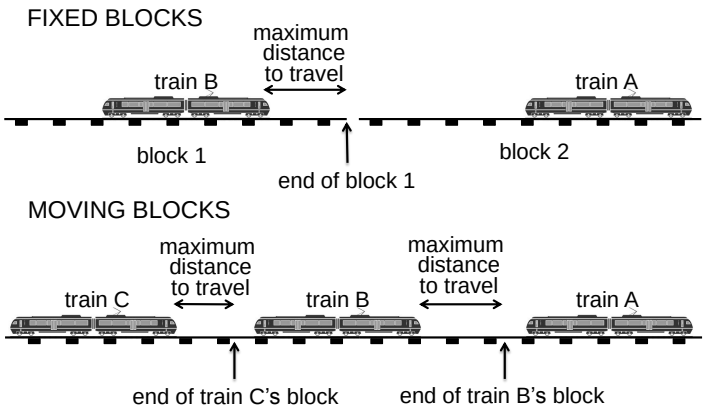

Figure 1: Fixed-block and moving-block operation.

CC will no longer have valid guarantees that train movement is still safe and will trigger an Emergency Brake (EB). It is clearly desirable to limit the frequency of spurious emergency brakes, i.e. emergency brakes that are simply due to losses on the wireless channel and not to a potential collision risk. Indeed spurious emergency brakes can be themselves a cause of danger, with trains potentially blocked in tunnels, risks of passengers disembarking on the tracks, etc. Moreover, a spurious EB can generate legitimate EBs on the following trains on the track, causing in this way major service disturbance. For this reason, the so-called performance based contracts can bind rail transport companies to specify the maximum number of spurious emergency brakes over a given period of time.

In spite of their criticality, the estimation of the rate of spurious EBs is mostly based on historical operational data. This approach strongly limits the possibility to evaluate ahead of time the performance when significant changes are deployed and in particular when new lines based on new technologies are built. It is often required to experimentally adapt different system parameters (e.g. transmission power levels, timer values, ...) after the deployment of the line, and sometimes even to deploy additional trackside equipment (e.g. radio transmitters). These difficulties are often considered one of the reasons for the delay in the standardization of ETCS level 3. For example [8] shows that the official quality of service specifications for the different subcomponents of the ETCS level 3 system can lead to a ridiculously high rate of spurious EBs (one every 30 minutes).

A model-based analysis can then play a fundamental role for a preliminary evaluation of the real performance of moving block control. Some work has been done in this direction following [8], and then considering its abstraction from ETCS level 3 specifications mostly using Stochastic Petri Nets (SPNs) [9, 5, 1, 3, 2]. In particular the approach proposed in [8] to numerically solve the SPN works only under the so-called enabling restriction, i.e. only one transition can be generally distributed and all the others should be exponential random variables. In the more realistic cases, the authors rely then on Monte Carlo simulations of the SPN. The naive simulation approach presented in [8] cannot manage to quantify EB rate smaller than 2 EBs per hour. Importance splitting techniques used in [9] allow to estimate much smaller rates (about 10^{-10} per hour). It is not clear if the computational cost of this numerical approach is insensitive to the packet loss probability p. References [5] and [7] show how UML descriptions can be used to describe the moving block control in ETCS level 3 and can be automatically translated to MoDeST formal language (a process algebra-based formalism) and to SPNs, but they do not solve the problem of quantitative evaluation of such rates when losses are rare. In the very recent paper [2] Carnevali et al. use the tool ORIS to solve numerically the SPN proposed in [8, 9], without the need to rely on Monte Carlo simulations. The tool indeed overcomes the limit of the enabling restriction thanks to recent advancements based on the method of stochastic state classes [6]. Moreover, it allows for a transient analysis of the system. As a case study, the authors consider a toy-example similar to that in [8] leading to very high EB rates. From a preliminary analysis using their tool, it is not clear if more realistic scenarios can be solved in a reasonable amount of time.

Our approach differs from the related literature in three main aspects. First, rather than moving from

the current proposals for ETCS level 3, we consider as reference an actual implementation of the moving-block system for metro by Alstom, one of the world largest company in the domain of rail transport and signaling. Looking at an actual implementation has led us to identify the importance of the time-slotted operation of the two controllers (the CC and the ZC). Indeed, the most important delay component in the messages' exchange between the CC and the ZC is due to the waiting time for the next clock tick at which the controller can process the message. This waiting time can be equal to hundreds of milliseconds versus the tens of milliseconds due to network delays. This aspect was ignored in the previous literature and we show that has to be addressed to correctly evaluate the system performance. In particular, a consequence of the time-slotted operation is that the EB rate exhibits non-trivial discontinuity as the timer value changes. A second (methodological) difference in comparison to the direction of [8] and follow-ups is that we push as further as possible the probabilistic analysis to derive closed-formula expressions. We derive a general formula for the rate of spurious EBs under general loss and delay processes, and a simple formula for the case of independent and homogeneous packet losses. The analysis allows to better understand the role of the different system parameters. On the contrary, the existing literature only relies on simulations or (in the case of [2]) on the numerical solution of a SPN. In both cases the dependence on the system parameters is hidden. Finally, from the algorithmic point of view, it is not clear if the numerical approaches proposed until now can be practically used to estimate EB rates as low as in this paper. Our guess is that this is probably not the case but, perhaps, for [9] and [2].Indeed our approach does not need to simulate rare sequences of packet losses and is then practically implementable.

The paper is organized as follows. In Sec. 2 we describe our assumptions about the train scenario and the details of the moving-block control including typical values for system parameters. Then in Sec. 3 we describe our general approach to study the system, we show that a worst case analysis is of limited utility (Sec. 3.1.2) and then move to derive a general formula for the EB rate (Sec. 3.1.3) that requires to characterize system delays (Sec. 3.2) and losses. The case of independent and homogeneous packet losses is considered in Sec. 3.3. Some numerical experiments are in Sec. 4. Section 5 concludes the paper and discusses how to extend our approach to more general loss scenarios. The most frequently used acronyms are listed in Table 1. Due to space constraints some of the results are in the companion technical report [4].

Table 1: List of Acronyms

CBTC	Communication Based Train Control
CC	Carborne Controller
DCS	Data Communication Sub-System
EB	Emergency Brake
EOA	End-Of-Authority
ETCS	European Train Control System
LOC	Location report
TM	validity duration Timer of a LOC
ZC	Zone Controller

2 Scenario

Here we describe the specific railway scenario we consider. In our description we will refer to transmission technologies and parameters typical of a urban rail network (and then of a CBTC system), but our following analysis does not depend on these specific implementation details. What is instead required is that the random variables (r.v.s) defined below (train speed, distances between access points, etc.) have bounded support and are lower bounded by a positive constant. For a given r.v. α, we denote by $\alpha_{\min} > 0$ its lower bound and by $\alpha_{\max} < \infty$ its upper bound.[1]

We consider a train moving on an infinitely long track. The train has two WiFi On Board Modems (OBMs) with directional antennas: one is located at the front of the train, the other at the back. We refer

[1] Throughout the paper Greek letters always denote random variables, while capital letters usually denote system parameters.

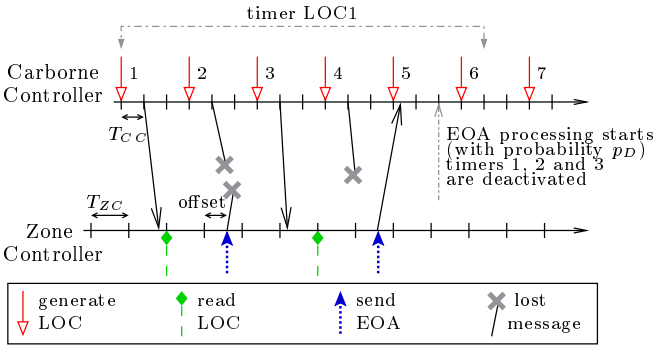

Figure 2: Illustration of LOC-EOA exchanges.

to them respectively as the blue and the red OBMs. Along the track there are pairs of closely-located WiFi Access Points (APs), using the same channel. The pair is called a Trackside Radio Equipment (TRE). Each AP in a TRE is devoted to communicate with one of the two OBMs and is connected to an independent wired network through which the Zone Controller (ZC) can be reached. We also label the APs, the wireless channels and the wired networks blue or red as the corresponding OBM. Hence communications between the train and the ZC are possible through separate paths, each with a single wireless link.

2.1 Train Moving-Block Control

In this section we describe the detailed operation of a moving block system considering as reference the specific CBTC implementation by Alstom.[2]

Figure 2 shows a messages exchange between the on board controller (the CC) and the ground controller (the ZC). Observe that both the controllers operate in discrete time on the basis of clock periods of hundreds of milliseconds. This is due to the fact that they are actually e-out-of-f voting systems where different processors perform in parallel the same calculations and a time-slotted operation simplifies the synchronism of the processors. The clock periods at the ZC and at the CC (respectively T_{ZC} and T_{CC}) are in general different because the subsystems are provided by different vendors and also because they have different computational loads during one period.

The most important CBTC messages are location reports (LOC) and end-of-authority ones (EOA). A LOC is a message periodically transmitted from the on board CC through the Data Communication Sub-System (DCS) to the ground ZC. The message is actually sent twice through the blue and the red networks. The first LOC arriving at the ZC is processed. Each LOC is acknowledged by an EOA message in the reverse direction (again sent through the two networks). The EOA communicates to the CC how far the train can advance. The LOC has a validity duration TM and a timer with such duration is activated at the generation of the LOC. An EOA is said to be valid if the timer of the corresponding LOC has not expired yet. The CC-ZC-CC exchange works as follows.

1. A LOC is generated at the CC every T_{LOC}, multiple of the CC clock period T_{CC}.
2. The LOC (say LOC k) is ready to be emitted and passed to the DCS after a processing delay equal to T_{CC}.
3. The delivery delay introduced by the DCS is a random variable χ_1 with support in $[T_{DCS,\min}, T_{DCS,\max}]$.
4. At the ZC the LOC is available for computing at the next tick of the clock.
5. The computing time at the ZC required to process the LOCs from all the trains in the zone and generate the corresponding EOAs is T_{ZC}.

[2] The parameters' values have been slightly changed and some specific implementation details are hidden to protect Alstom industrial know-how.

6. The EOA k is emitted within the next cycle of the ZC at an offset O depending on the train.
7. The EOA is delivered to the CC after a random delay χ_2, distributed as χ_1, but independent from it.
8. At the CC the EOA gets in a processing queue, at the next tick of the CC clock the most recent EOA present in the queue is processed unless there are higher priority tasks arrived during the same CC clock period (which happens with probability p_D). In any case an EOA processing is not delayed more than an additional CC period.
9. The EOA k is actually processed only if it remains valid until the end of the current CC clock. Once processing starts, all the pending timers for older LOCs (i.e. LOC h for $h \leq k$) are deactivated.
10. If the timer of a LOC is not deactivated before its expiration, the EB procedure is triggered.

In what follows we refer to the k-th LOC and its corresponding EOA as the k-th LOC-EOA exchange, but note that any later EOA can disactivate the timer of the k-th LOC. We say that a LOC-EOA exchange is lost if either the LOC or the EOA does not arrive to destination.

Table 2: Notation and typical values for the variables. In the paper some of the variables appear with subscripts. A subscript b (r) denotes that the variable refers to the blue (red) OBM or network. A subscript L (E) denotes that it refers to a LOC (an EOA).

Symbol	Quantity	Value
T_{ZC}	ZC clock period	378 ms
T_{CC}	CC clock period	225 ms
T_{LOC}	LOC generation period	$3T_{CC}$
TM	validity duration of a LOC	5.5 s
T_{DCS}	transmission delay	$[10, 50]$ ms
τ	positive random component of T_{DCS}	$[0, 40]$ ms
ϕ	positive random component of T_{DCS} for first message to arrive	$[0, 40]$ ms
O	EOA transmission offset	$[0, T_{ZC}]$
ω_{CC}	number of CC ticks an EOA waits until CC processes it	$\{0, 1\}$
p_D	probability that ω_{CC} is 1	0.01
ω_{ZC}	time interval between LOC arrival at ZC and next ZC tick	
σ	time interval between earliest arrival time of a LOC at ZC and next ZC tick	
q_{EB}	emergency brake probability	
r_{EB}	emergency brake rate	
p	packet loss	
$\tilde{p}$	probability to lose a LOC-EOA exchange	
T_k	arrival time of k-th EOA	
γ_k	tick at which k-th EOA is processed	
$\mathcal{D}_k$	event that k-th EOA is late to deactivate the timer of LOC 1	
$\mathcal{T}_k$	event that k-th LOC experiences a timeout	
$\mathcal{L}_k$	event of k-th LOC-EOA exchange loss	

3 Analysis

In this paper we consider that the system is described by a stationary stochastic process and calculate the steady-state rate at which emergency brakes occur (as common to all the related literature but [2]). In particular we consider that the train is moving according to some stationary mobility model and the algorithm described above is running all the time, even after the occurrence of an emergency brake. Ignoring the train stopping time after an EB is a reasonable approximation because we are estimating rare events.

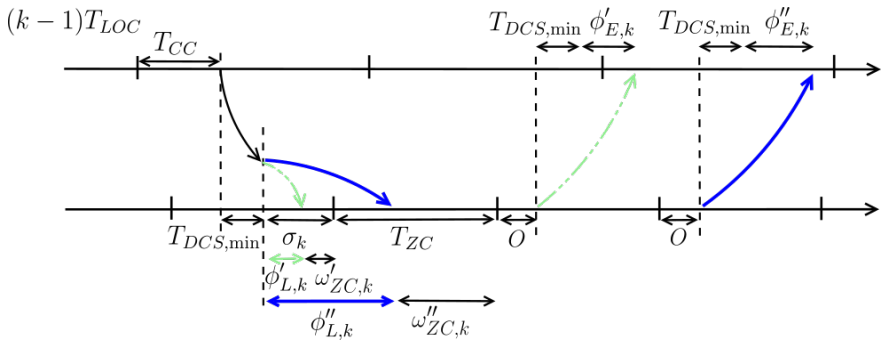

Figure 3: Different delay components of the k-th LOC-EOA exchange for two different values of the LOC transmission delay $\phi'_{L,k}$ and $\phi''_{L,k}$.

We denote by $\mathcal{L}_k$ the event that the exchange k is lost, $\mathcal{T}_k$ the event that the k-th LOC experiences a timeout and $\bar{A}$ the complement of set A. The k-th LOC experiences a timeout if the k-th exchange is lost and the later EOAs do not arrive or arrive too late, then $\mathcal{T}_k \subset \mathcal{L}_k$[3]. We observe that a sequence of consecutive timeouts generates a single EB and then a timeout for a given LOC, say it LOC 1, is counted as an EB only if the previous LOC 0 does not experience a timeout. The probability q_{EB} that a random LOC experiences an emergency break is then $q_{EB} = \Pr(\bar{\mathcal{T}}_0 \cap \mathcal{T}_1)$ that does not depend on the specific pair of LOCs considered because the process is stationary. Moreover, under the condition that LOC 1 experiences a timeout, LOC 0 experiences a timeout if and only if the corresponding exchange is lost, because later EOAs are not able to block the timer of LOC 1 and a fortiori the timer of LOC 0. Then $\bar{\mathcal{T}}_0 \cap \mathcal{T}_1 = \bar{\mathcal{L}}_0 \cap \mathcal{T}_1$ and the rate of emergency brakes is

$$r_{EB} = \frac{q_{EB}}{T_{LOC}} = \frac{\Pr(\bar{\mathcal{L}}_0 \cap \mathcal{T}_1)}{T_{LOC}}. \tag{1}$$

3.1 EB Probability

In this section we first derive some simple bounds for q_{EB}. The bounds will reveal to be too loose to be practically used, but they are nevertheless useful for the subsequent analysis. We conclude the section with a general formula for the EB rate, whose terms will be calculated in the following sections. We report numerical values corresponding to the typical scenario presented in Sec. 2.

3.1.1 Minimum and maximum LOC-EOA round trip times.

We calculate the minimum and the maximum time between the generation of a LOC and the instant T when the corresponding EOA is available for computation at the CC. Consider a LOC generated at time 0. Its EOA arrives at the CC at time (see also Fig. 3):

$$T = T_{\min} + \phi_L + \phi_E + \omega_{ZC} + O, \tag{2}$$

where $T_{min} = T_{CC} + 2T_{DCS,\min} + T_{ZC} = 623$ ms, ω_{ZC} is the time interval between the arrival of the LOC at the ZC and the next ZC tick and ϕ_L and ϕ_E are the random components of the transmission delays respectively for the first LOC and the first EOA to arrive at destination.

The earliest arrival time $T_{\min} + O$ occurs when the LOC and the EOA experience the minimum travel times on the DCS (i.e. $\phi_L = \phi_E = 0$) and the LOC is available for computing at the ZC immediately before a ZC tick (i.e. $\omega_{ZC} = 0$).

The latest arrival time $T_{\max} + O$ occurs when the LOC and the EOA experience the maximum travel time on the DCS (i.e. $\phi_L = \phi_E = T_{DCS,\max} - T_{DCS,\min}$) and the LOC is available for computing at the ZC immediately after a ZC tick. In this case the LOC will wait an additional T_{ZC} before being processed (i.e. $\omega_{ZC} = T_{ZC}$). Hence $T_{\max} = T_{CC} + T_{DCS,\max} + T_{ZC} + T_{ZC} + T_{DCS,\max} = 1081$ms.

[3] In this paper $A \subset B$ denotes that A is a subset of B, not necessarily proper.

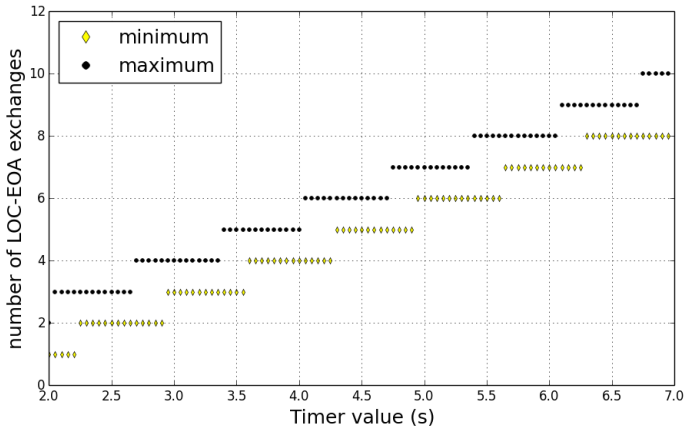

Figure 4: Minimum and maximum number of LOC-EOA exchanges for $O = 50$ ms, calculated through Eqs. (4) and (3).

3.1.2 Number of potential LOC-EOA exchanges before a TimeOut.

Even if a LOC or an EOA is lost, the EOAs corresponding to following LOCs could still deactivate its timer and then the emergency brake would be prevented. In this section we calculate how many LOC-EOA exchanges can happen between the generation of a LOC and the expiration of the corresponding timer, i.e. how many other EOAs can have a chance to block the timer.

Let us consider that the first LOC is generated at time $t = 0$, then its timer would expire at time $t = TM$. The maximum number $n_{\max}$ of LOC-EOA exchanges can be calculated considering that i) the last potentially useful EOA arrives in the shortest time possible and ii) it is immediately processed by the following CC tick, which is the last one before the timer expires.

The last potential useful EOA arrives at $(n_{\max} - 1)T_{LOC} + T_{\min} + O$ and it can then be processed at $T_{CC} \lceil ((n_{\max} - 1)T_{LOC} + T_{\min} + O) / T_{CC} \rceil$. The CC tick just before the timer expires occurs at time $T_{CC} \lfloor TM/T_{CC} \rfloor$, We determine $n_{\max}$ by imposing that $\left\lceil \frac{(n_{\max}-1)T_{LOC}+T_{\min}+O}{T_{CC}} \right\rceil = \left\lfloor \frac{TM}{T_{CC}} \right\rfloor$,[4] and we can manipulate this equality as in [4], to obtain:

$$n_{\max} = 1 + \left\lfloor \frac{TM - \left\lceil \frac{T_{\min}+O}{T_{CC}} \right\rceil T_{CC}}{T_{LOC}} \right\rfloor . \tag{3}$$

Similarly the minimum number $n_{\min}$ of LOC-EOA exchanges can be calculated considering that i) the last potentially useful EOA arrives in the longest time possible and ii) it is processed 2 CC ticks later in correspondence of the last tick before the timer expires. Then we determine $n_{\min}$ by imposing that $\left\lceil \frac{(n_{\min}-1)T_{LOC}+T_{\max}+O}{T_{CC}} \right\rceil = \left\lfloor \frac{TM}{T_{CC}} \right\rfloor - 1$, and proceeding as above we obtain:

$$n_{\min} = 1 + \left\lfloor \frac{TM - \left(\left\lceil \frac{T_{\max}+O}{T_{CC}} \right\rceil + 1 \right) T_{CC}}{T_{LOC}} \right\rfloor . \tag{4}$$

The difference between $n_{\max}$ and $n_{\min}$ depends on the timer TM and also on the offset. For the typical values in Table 2 they differ by at most 2 exchanges, i.e. $n_{\max} \leq n_{\min} + 2$. Figure 4 shows $n_{\min}$ and $n_{\max}$ for different values of the timer TM and an offset $O = 50$ ms. It also shows that the difference of two exchanges is achieved for some values of TM.

[4] This assumes $n_{max} > 1$. The first EOA needs to be valid until the end of the CC clock during which it is processed and then its processing time should start the latest at the tick number $\left\lfloor \frac{TM - T_{CC}}{T_{CC}} \right\rfloor$.

The two values $n_{\min}$ and $n_{\max}$ allow us to provide respectively upper and lower bounds for the EB probability and then for the EB rate, but these bounds can be too loose for practical uses. We are going to show it in the simple case when packet losses on the two wireless blue and red channels are independent Bernoulli random variables with parameter p. In this case a LOC or an EOA message is received with probability $1-p^2$ and the probability $\tilde{p}$ to lose a LOC-EOA exchange is then $\tilde{p} = 1-(1-p^2)^2$. An emergency brake requires that the exchange 0 is not lost. Moreover the EB will necessarily occur if the $n_{\max}$ following LOC-EOA exchanges are lost (even if the $(n_{\max}+1)$-th EOA arrives, it will be after the timer expiration) and cannot occur unless $n_{\min}$ exchanges are lost (the first $n_{\min}$ EOA cannot arrive late even in the worst case). It follows that

$$(1-\tilde{p})\tilde{p}^{n_{\max}} \leq q_{EB} \leq (1-\tilde{p})\tilde{p}^{n_{\min}}. \tag{5}$$

With the values in Table 2 the upper bound can be up to $\tilde{p}^{-2}$ times larger than the lower bound. A typical value for the packet loss probability is $p = 5\%$, and then $\tilde{p} \approx 0.5\%$ and the ratio of the two bounds is almost 4×10^4. In this case, as we are going to show later, the upper bound can be too pessimistic and practically of no utility to set the parameter TM. For this reason a more refined analysis is required.

3.1.3 Exact Formula

LOC 1 is generated at time $t = 0$ and then the k-th LOC is generated at $(k-1)T_{LOC}$. The k-th EOA is the EOA corresponding to the k-th LOC. The timer of LOC 1 would expire at time $t = TM$. Remember that $\mathcal{L}_k$ denotes the event that the k-th LOC-EOA exchange is lost. Let $\mathcal{D}_k$ denote the event that the k-th EOA arrives too late to deactivate the timer of LOC 1. The two events are disjoint, i.e. $\mathcal{L}_k \cap \mathcal{D}_k = \emptyset$. LOC 1 experiences a timeout if and only if all the following exchanges are lost or their EOAs arrive too late, i.e.

$$\mathcal{T}_1 = \bigcap_{k=1}^{\infty} (\mathcal{L}_k \cup \mathcal{D}_k) = \bigcap_{k=1}^{n_{\max}} (\mathcal{L}_k \cup \mathcal{D}_k), \tag{6}$$

where the last equality follows from the fact that only the first $n_{\max}$ exchanges have a possibility to stop the timer ($\Pr(\mathcal{L}_k \cup \mathcal{D}_k) = 1$ for $k > n_{\max}$).

Due to timing constraints EOAS cannot arrive out of order. A consequence is that if the k-th EOA arrives too late to deactivate the timer of LOC 1, no later EOA will be able to deactivate it. In particular later EOAs will be lost or will arrive too late, i.e. $\mathcal{D}_k \subset \mathcal{D}_{k'} \cup \mathcal{L}_{k'}$ for all $k' \geq k$. This simple relation allows us to conclude [4] that for any m

$$\bigcap_{k=1}^{m} (\mathcal{L}_k \cup \mathcal{D}_k) = \bigcup_{k=1}^{m} \left(\mathcal{D}_k \cap \left(\bigcap_{h=1}^{k-1} \mathcal{L}_h \right) \right) \cup \left(\bigcap_{h=1}^{m} \mathcal{L}_h \right). \tag{7}$$

We can now move to calculate q_{EB}. From Eqs. (6) and (7), it follows that

$$\begin{aligned} q_{EB} &= \Pr\left(\bar{\mathcal{L}}_0 \cap \mathcal{T}_1\right) = \Pr\left(\bar{\mathcal{L}}_0 \cap \bigcap_{k=1}^{n_{\max}} (\mathcal{L}_k \cup \mathcal{D}_k)\right) \\ &= \Pr\left(\bar{\mathcal{L}}_0 \cap \left(\bigcup_{k=1}^{n_{\max}} \left(\mathcal{D}_k \cap \left(\bigcap_{h=1}^{k-1} \mathcal{L}_h \right) \right) \cup \left(\bigcap_{h=1}^{n_{\max}} \mathcal{L}_h \right) \right) \right). \end{aligned} \tag{8}$$

This expression can be simplified observing that the first $n_{\min} - 1$ EOAs cannot arrive late ($\Pr(\mathcal{D}_k) = 0$ for $k \leq n_{\min}$)

$$q_{EB} = \Pr\left(\bar{\mathcal{L}}_0 \cap \left(\bigcup_{k=n_{\min}+1}^{n_{\max}} \left(\mathcal{D}_k \cap \left(\bigcap_{h=1}^{k-1} \mathcal{L}_h \right) \right) \cup \left(\bigcap_{h=1}^{n_{\max}} \mathcal{L}_h \right) \right) \right) \tag{9}$$

Equation (9) can be read as follows: a timeout occurs if there is a sequence of $n_{\min}$, $n_{\min}+1$ up to $\ldots n_{\max}-1$ exchanges lost and the following EOA arrives late or if all the $n_{\max}$ exchanges are lost. These events are

disjoint, because $\mathcal{D}_k \cap \mathcal{L}_k = \emptyset$, and then we can conclude:

$$q_{EB} = \sum_{k=n_{\min}+1}^{n_{\max}} \Pr\left(\mathcal{D}_k \cap \left(\bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{k-1} \mathcal{L}_h\right)\right) + \Pr\left(\bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{n_{\max}} \mathcal{L}_h\right) \tag{10}$$

$$= \sum_{k=n_{\min}+1}^{n_{\max}} \Pr\left(\mathcal{D}_k \,\middle|\, \bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{k-1} \mathcal{L}_h \cap \bar{\mathcal{L}}_k\right) \Pr\left(\bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{k-1} \mathcal{L}_h \cap \bar{\mathcal{L}}_k\right) + \Pr\left(\bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{n_{\max}} \mathcal{L}_h\right). \tag{11}$$

The last equality holds because $\mathcal{D}_k = \mathcal{D}_k \cap \bar{\mathcal{L}}_k$. The reason why we introduce the additional set $\bar{\mathcal{L}}_k$ will be clear in the following sections, where we will move to characterize delays and losses in order to compute the terms appearing in Eq. (11). We denote this sequence of loss events as $S_{\mathcal{L},k} \triangleq \bar{\mathcal{L}}_0 \cap \bigcap_{h=1}^{k-1} \mathcal{L}_h \cap \bar{\mathcal{L}}_k$.

As observed, for the typical values in Table 2 it is $n_{\max} \leq n_{\min} + 2$ and then there are at most 3 terms in Eq. (11).

3.2 Delay

In this section we characterize the event $\mathcal{D}_k$. In particular, we are interested to evaluate the probabilities $\Pr(\mathcal{D}_k \mid S_{\mathcal{L},k})$ appearing in Eq. (11). To this purpose we will study in detail the different components that determine if the k-th EOA arrives before or after the expiration of the timer of the first LOC.

Again, assume that LOC 1 is generated at time 0. If the k-th exchange LOC-EOA is not lost, then the arrival time of the k-th EOA is

$$T_k = T_{\min,k} + \phi_{L,k} + \phi_{E,k} + \omega_{ZC,k} \tag{12}$$

where $T_{\min,k} = T_{CC} + 2T_{DCS,\min} + T_{ZC} + (k-1)T_{LOC} + O$ and the random variables $\omega_{ZC,k}$, $\phi_{L,k}$, $\phi_{E,k}$ represent the same quantities as those in Eq. (2), but are referred to the k-th exchange rather than to the first one. The EOA is processed at the tick

$$\gamma_k \triangleq \left\lceil \frac{T_k}{T_{CC}} \right\rceil + \omega_{CC,k}, \tag{13}$$

where $\omega_{CC,k}$ represents the processing delay at the CC expressed in number of ticks. According to the description in Sec. 2.1 $\omega_{CC,k}$ can assume value 0, if the EOA is going to be processed at the first CC tick after T_k, or value 1, if it is going to be processed at the following tick. We are going to characterize the Bernoulli r.v. $\omega_{CC,k}$ soon, for the moment we observe that the EOA arrives too late if $\gamma_k > \frac{TM}{T_{CC}}$ i.e. the EOA starts being processed after the expiration of the timeout. Then, the event $\mathcal{D}_k$ can be expressed as $\mathcal{D}_k = \bar{\mathcal{L}}_k \cap \left\{\gamma_k > \frac{TM}{T_{CC}}\right\}$, and

$$\Pr\left(\mathcal{D}_k \,\middle|\, S_{\mathcal{L},k}\right) = \Pr\left(\gamma_k > \frac{TM}{T_{CC}} \,\middle|\, S_{\mathcal{L},k}\right), \tag{14}$$

because $\bar{\mathcal{L}}_k \subset S_{\mathcal{L},k}$. In order to calculate this probability we now move to consider each source of randomness in γ_k.

3.2.1 Processing delay at the CC.

Observe that $\omega_{CC,k}$ is independent of the arrival time of the k-th EOA T_k, as well as on arrival of any other EOA. In fact the queuing delay for the k-th EOA depends only on higher-priority traffic and not on the previous EOAs (that may or not being present in the processing queue), because only the most recent EOA is processed. It follows that $\omega_{CC,k}$ is independent of the event $\bigcap_{h=1}^{k-1} \mathcal{L}_h$ and its conditional distribution is equal to the a priori distribution provided in Sec. 2.1, i.e. $\omega_{CC,k}$ in Eq. (14) is a Bernoulli random variable with parameter p_D. While $\omega_{CC,k}$ as introduced is defined only when the k-th exchange is not lost, we can define it for any k as an independent Bernoulli random variable with parameter p_D. It can then be interpreted as the processing delay experienced by an hypothetical EOA arriving at a given time. The distribution of $\omega_{CC,k}$ does not depend on k and is independent of $S_{\mathcal{L},k}$.

3.2.2 Processing delay at the ZC.

Going back to Eq. (12), the random variable $\omega_{ZC,k}$ is dependent on the relative position of the ticks of the two clocks but also on the value of $\phi_{L,k}$. In fact the later the LOC arrives at the ZC (the larger $\phi_{L,k}$) the less the LOC has to wait until the next ZC tick (the smaller $\omega_{ZC,k}$), unless the LOC arrives so late that it misses the first available ZC tick and needs to wait for the next one. While we cannot get rid completely of this dependence, it is simpler to reverse it. With reference to Fig. 3, we express T_k with this equivalent expression:

$$T_k = T_{\min,k} + \sigma_k + \mathbb{1}_{\phi_{L,k}>\sigma_k} T_{ZC} + \phi_{E,k} \tag{15}$$

where σ_k denotes the time interval between the earliest possible instant at which the k-th LOC could be received at the ZC and the next ZC tick and $\mathbb{1}_{\phi_{L,k}>\sigma_k}$ is a Bernoulli random variable indicating if the random component of the communication delay will cause the LOC to miss this ZC tick and then to wait for the following one. It can be easily verified that σ_k depends on the specific LOC we are considering because the two clock periods are different. Then coherently with the idea that, in order to evaluate q_{EB}, the first LOC is chosen at random, σ_k is a random variable. Observe that the variable σ_k is independent of the loss processes and in particular of $S_{\mathcal{L},k}$. Moreover, it is independent of communication delays (i.e. of the variables $\phi_{L,k}$, $\phi_{E,k}$) and of processing delay at the ZC (i.e. of $\omega_{CC,k}$). Our next task is to determine σ_k's distribution.

Given the value $\sigma_1 = s_1$ for the first LOC, the values of the other r.v.s σ_k for $k > 1$ are uniquely determined, let $\sigma_k = s_k$. Assuming that T_{ZC} and T_{LOC} are commensurable numbers and choosing an opportune unit so that their values can be expressed as integers, in [4] we show that the possible values for s_k are the values s in $[0, T_{ZC})$ for which the following Diophantine equation in m and n admits integer solutions:

$$mT_{ZC} - nT_{LOC} = s - s_1. \tag{16}$$

The study of this equation in [4] leads to the conclusions that s_k assumes all and only the values in the set $S = \{\tilde{s} + iM,\ i = 0, 1, \dots q_{ZC} - 1\}$ where M is the greatest common divisor of T_{ZC} and T_{LOC}, $T_{ZC} = q_{ZC}M$ and $\tilde{s} = s_1 \% M$. For example for the typical values we consider ($T_{ZC} = 378$ ms, $T_{LOC} = 675$ ms) it is $M = 27$, $q_{ZC} = 14$. Moreover, the sequence s_n is periodic with period q_{ZC} and then assumes the q_{ZC} values in S only once during each period. When we consider that the first LOC is a LOC selected at random, we conclude then that the variable σ_k is a uniform random variable over the set $S = \{\tilde{s} + kM, k = 0, 1, \dots q_{ZC} - 1\}$.[5]

3.2.3 Communication delays.

In order to completely characterize the probability in Eq. (14), we need to discuss the two random variables $\phi_{L,k}$ and $\phi_{E,k}$. Remember that $\phi_{L,k}$ is the delay experienced by the "fastest" of the two LOC packets conditional on one of them arriving at the ZC. Let $\tau_{r,L}$ denote the random component of the delay experienced by the k-th LOC packet transmitted on the red network if it is not lost (we omit for simplicity the dependence on k). We can similarly introduce $\tau_{b,L}$, $\tau_{r,E}$ and $\tau_{b,E}$. These delays are independent and identically distributed random variables with Cumulative Distribution Function (CDF) $F_\tau(t)$. In particular, under the typical values in Sec. 2.1 they have support $[0, 40]$ ms.

3.3 Independent losses

As an application of Eq. (11) we consider the case when packet losses are independent and homogeneous and Eq. (11) reduces to an easy-to-calculate exact formula. The independence allows to write:

$$\Pr\left(\mathcal{D}_k \,\middle|\, S_{\mathcal{L},k}\right) = \Pr\left(\mathcal{D}_k \,\middle|\, \bar{\mathcal{L}}_k\right) = \Pr\left(\gamma_k > \frac{TM}{T_{CC}}\right) \triangleq d(k), \tag{17}$$

where γ_k is a function of the independent r.v.s $\omega_{CC,k}$, σ_k (already characterized in the previous section) and $\phi_{L,k}$ and $\phi_{E,k}$, whose CDF $F_\phi(t)$ can be easily derived by conditioning on the number of packets arriving at the ZC/CC:

$$F_\phi(t) = \frac{(1-p)^2}{1-p^2}\left(1 - (1 - F_\tau(t))^2\right) + \frac{2(1-p)p}{1-p^2} F_\tau(t) = \frac{F_\tau(t)\left(2 - F_\tau(t)(1-p)\right)}{1+p}.$$

[5] The analysis can be easily adapted to take into account the effect of clocks' frequency-shift [4].

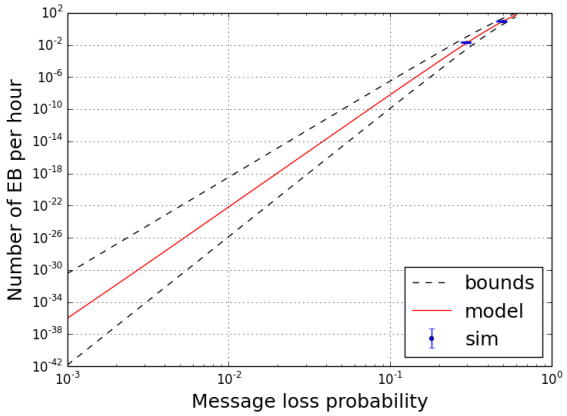

Figure 5: Number of emergency brakes per hour when $TM = 5.5$ s.

Our definition of $d(k)$ stresses that $\Pr(\gamma_k > TM/T_{CC})$ is a function of k, but this happens because of the constant $T_{\min,k}$, while the distributions of the r.v.s $\omega_{ZC,k}$, $\sigma_{CC,k}$, $\phi_{L,k}$ and $\phi_{E,k}$ do not depend on k.

Finally, by developing the terms $\Pr(S_{\mathcal{L},k})$ in Eq. (11), we obtain

$$q_{EB} = \sum_{k=n_{\min}+1}^{n_{\max}} d(k)\tilde{p}^{k-1}(1-\tilde{p})^2 + \tilde{p}^{n_{\max}}(1-\tilde{p}), \tag{18}$$

where $\tilde{p} = 1-(1-p^2)^2$ is the probability that an exchange is lost.

4 Numerical Experiments

In this section we validate Eq. (18) through discrete-event simulations of the system, for which we have developed an ad-hoc Python simulator. The scenario tested by discrete-event simulations matches that described in Sec. 2 and considered in our analysis. For constant system parameters and the support of random variables, we have considered the typical values indicated in Table 2.

Figure 5 shows the EB rate versus different values of the packet loss probability p for $TM = 5.5$ s. The red solid curve is obtained through Eq. (18). Simulation results obtained by the Python simulator for selected values of p are reported as 95% confidence intervals in blue. About the computational time, Eq. (18) requires a few seconds on a current commodity PC. On the same machine the Python simulator is able to simulate roughly 10^4 hours of train operation in one hour. It follows a rate of the order of 10^{-4} EBs per hour requires roughly 100 hours to be estimated with a precision of 1% through the Python simulator. It is clear that lower EB rates are out of reach for the Python simulator.

Figure 5 also shows the black dashed curves that plot the functions $(1-\tilde{p})\tilde{p}^{n_{\min}}/T_{LOC}$ and $(1-\tilde{p})\tilde{p}^{n_{\max}}/T_{LOC}$ and that correspond to the upper and lower bound in Eq. (5) in presence of independent Bernoulli packet losses with probability p. We observe that the produced bounds are very loose.

As a final application of our methodology, Fig. 6 shows the expected number of emergency brakes per hour for different values of the timer TM, $O = 50$ ms and packet loss probability $p = 0.3$. The theoretical values calculated from Eqs. (18) and (1) (red dots) are compared with the bounds (black dashed lines). The figure shows that the simple upper bound can be orders of magnitude larger than the actual value. We now discuss the discontinuities appearing in the EB rate curve. From Eq. (18) we observe that the EB probability exhibits discontinuities only if $n_{\min}$, $n_{\max}$ or the functions $d(k)$ do. The small gaps of the EB rate correspond indeed to changes in the values $n_{\min}$ or $n_{\max}$ as it is revealed by the corresponding jumps of the bounds. The other gaps correspond to changes of the functions $d(k)$. We remember that $d(k) = \Pr(\gamma_k > TM/T_{CC})$, where γ_k is an integer. Then $d(k)$ does not depend on TM as far as $h \leq TM/T_{CC} < h+1$ for some integer

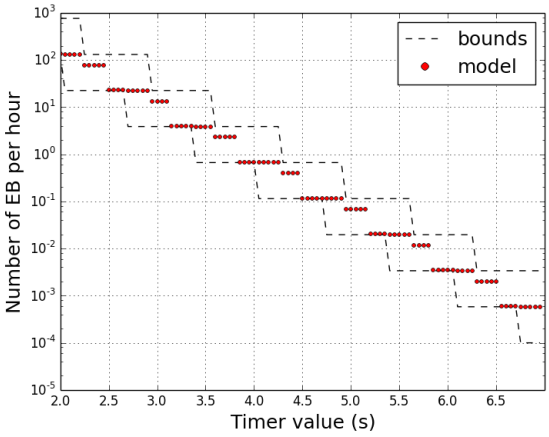

Figure 6: Rate of emergency brakes when $O = 50$ ms and $p = 0.3$.

h. Indeed, it can be checked that the other discontinuities in the curve (when neither $n_{\min}$ nor $n_{\max}$ change) correspond to integer values of TM/T_{CC}. This high sensitivity to the timer value is not only easily revealed by our numerical method, but well explained by our theoretical analysis.

5 Conclusion

In this paper we study the moving block control to quantify the rate of spurious EBs. Differently from existing literature, our starting point is not the current recommendation for the future ETCS level 3, but an actual implementation for metro. Equation (11) characterizes the EB rate in a general stationary setting, but it requires to compute the probability to observe specific patterns of packet losses, that can be a difficult task in general. Nevertheless, in the simple case of independent and homogeneous packet losses, the equation reduces to a simple analytical formula whose computational cost does not depend on the loss probability value. The formula can then be used to quantify extremely rare events (as emergency brakes should be). We are currently working to study more general loss scenarios, where losses are strongly correlated and time-variant. Our current results are in [4] and rely on a Monte Carlo approach to efficiently sample from the stationary distribution of the system.

This work is partially funded by the Inria-Alstom virtual lab.

References

[1] T. Babczyński and J. Magott. Dependability and safety analysis of ETCS communication for ERTMS level 3 using performance statecharts and analytic estimation. In *Proc. of 9th DepCoS-RELCOMEX*, volume 286 of *Advances in Intelligent Systems and Computing*, pages 37–46. 2014.

[2] L. Carnevali, F. Flammini, M. Paolieri, and E. Vicario. Non-markovian performability evaluation of ERTMS/ETCS level 3. In *Computer Performance Engineering (Proc. of EPEW 2015)*, volume 9272 of *Lecture Notes in Computer Science*, pages 47–62. 2015.

[3] F. Flammini, S. Marrone, M. Iacono, N. Mazzocca, and V. Vittorini. A multiformalism modular approach to ERTMS/ETCS failure modeling. *International Journal of Reliability, Quality and Safety Engineering*, 21(1):1450001 (29 pages), 2014.

[4] Giovanni Neglia et al. Performance evaluation of train moving-block control. Research Report RR-8917, Inria, May 2016.

[5] H. Hermanns, D. N. Jansen, and Y. S. Usenko. From StoCharts to MoDeST: A comparative reliability analysis of train radio communications. In *Proc. of WOSP '05*, pages 13–23, 2005.

[6] A. Horváth, M. Paolieri, L. Ridi, and E. Vicario. Transient analysis of non-Markovian models using stochastic state classes. *Performance Evaluation*, 69(7-8):315–335, 2012.

[7] J. Trowitzsch and A. Zimmermann. Using UML state machines and petri nets for the quantitative investigation of ETCS. In *Proc. of Valuetools '06*, 2006.

[8] A. Zimmermann and G. Hommel. A train control system case study in model-based real time system design. In *Proc. of IPDPS '03*, 2003.

[9] A. Zimmermann and G. Hommel. Towards modeling and evaluation of ETCS real-time communication and operation. *Journal of Systems and Software*, 77(1):47–54, 2005.