Transportation Research Part C 82 (2017) 314–336

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Performability evaluation of the ERTMS/ETCS – Level 3

Marco Biagi [a], Laura Carnevali [a,*], Marco Paolieri [b], Enrico Vicario [a]

[a] Department of Information Engineering, University of Florence, via di Santa Marta 3, Florence, Italy
[b] Department of Computer Science, University of Southern California, 941 Bloom Walk, Los Angeles, CA 90089, USA

ARTICLE INFO

Article history:
Received 10 March 2017
Received in revised form 25 June 2017
Accepted 5 July 2017
Available online 18 July 2017

Keywords:
European Train Control System
Moving-block signalling
Performability evaluation
Emergency brake
Non-Markovian stochastic Petri nets
Stochastic state classes

ABSTRACT

Level 3 of the ERTMS/ETCS improves the capacity of railways by replacing fixed-block signalling, which prevents a train to enter a block occupied by another train, with moving block signalling, which allows a train to proceed as long as it receives radio messages ensuring that the track ahead is clear of other trains. If messages are lost, a train must stop for safety reasons within a given deadline, even though the track ahead is clear, making the availability of the communication link crucial for successful operation.

We combine analytic evaluation of failures due to burst noise and connection losses with numerical solution of a non-Markovian model representing also failures due to handovers between radio stations. In so doing, we show that handovers experienced by a pair of chasing trains periodically affect the availability of the radio link, making behavior of the overall communication system recurrent over the hyper-period of periodic message releases and periodic arrivals at cell borders. As a notable aspect, non-Markovian transient analysis within two hyper-periods is sufficient to derive an upper bound on the first-passage time distribution to an emergency brake, permitting to achieve a trade-off between railway throughput and stop probability. A sensitivity analysis is performed with respect to train speed and headway distance, permitting to gain insight into the consequences of system-level design choices.

1. Introduction

The European Rail Traffic Management System (ERTMS) (UIC, 1999a,b) aims at improving performance, reliability, safety, cross-border interoperability, and maintenance costs of European railway networks. Developed by the European Union and several industrial manufacturers, it has become a worldwide success, with several installations and investment programs outside Europe (e.g., China, India, Australia).

The ERTMS relies on the European Train Control System (ETCS), an automatic train protection system which continuously supervises the train to ensure that the safe speed and distance are not exceeded. A *track-side* Radio Block Centre (RBC) periodically receives for each train a Position Report (PR) and the results of the integrity check. In turn, the *on-board* European Vital Computer (EVC) of each train receives from the RBC a Movement Authority (MA) specifying the maximum allowed advancement, the maximum allowed speed depending on the track morphology, and possible temporary speed restrictions. This supports safe control of the distance between a *foregoing* train and the *chasing* train that follows: whenever the chasing train reaches the allowed advancement position defined by the most recent MA, an emergency brake is activated by its EVC and the train is safely stopped.

* Corresponding author at: via di Santa Marta 3, 50139 Firenze, Italy.
E-mail addresses: marco.biagi@unifi.it (M. Biagi), laura.carnevali@unifi.it (L. Carnevali), paolieri@usc.edu (M. Paolieri), enrico.vicario@unifi.it (E. Vicario).

http://dx.doi.org/10.1016/j.trc.2017.07.002
0968-090X/

The ERTMS/ETCS is specified at three levels of operation. Level 1 (L1) uses track-side equipments to detect free railway blocks and perform train integrity checks, computes the MA by summing track circuits that are neither occupied nor in out-of-service or exclusion conditions (*fixed-block signalling*), and transmits the MA via transponders (balises) placed at a regular distance (*non-continuous track-to-train* communication). Level 2 (L2) relies on the Global System for Mobile Communications – Railway (GSM-R) (*continuous bidirectional* communication), reducing maintenance costs (balises are used only to transmit train position and line information) while making the availability of the radio channel critical for correct operation. Level 3 (L3) avoids train detection systems by considering the train itself as a moving-block, computing the MA from the minimum safe rear-end of the foregoing train (*moving-block signalling*), further improving line capacity and maintenance costs. While the ERTMS/ETCS-L2 is implemented by most railway lines, the ERTMS/ETCS-L3 is the most promising level in terms of capacity gains and track-side installation savings.

The ERTMS/ETCS has been widely addressed in the literature (Ghazel, 2014), with specific focus on L2 (Flammini et al., 2014, 2006; Abbaneo et al., 2006; Qiu et al., 2014; Hassami and Foord, 2001) and L3 (Zimmermann and Hommel, 2003, 2005; Carnevali et al., 2015; Hermanns et al., 2005; Babczyński and Magott, 2014; Esposito et al., 2003; Neglia et al., 2016). In Ghazel (2014), a subset of the System Requirements Specification of the ERTMS/ETCS is translated into a conditional transition system, supporting automated verification of safety, interoperability, and liveness properties. In Flammini et al. (2014, 2006), sensitivity analysis of a multi-formalism model of the ERTMS/ETCS-L2 evaluates the influence of design parameters on the probability of system-level failure modes, reducing the cost of components while fulfilling RAM (Reliability, Availability, Maintainability) requirements. Static analysis of the software of the trackside subsystem of the ERTMS/ETCS-L2 is performed in Abbaneo et al. (2006), supporting correctness verification, traceability, and refactoring. In Qiu et al. (2014), the ERTMS/ETCS-L2 railway signalling system is modeled as a system of systems using statecharts, and its dependability parameters are evaluated by considering human factors, network failures, and imprecise failure and repair rates. Risk forecasting for the on-board and trackside subsystems of the ERTMS/ETCS-L2 is performed in Hassami and Foord (2001) by leveraging a cause-consequence model and a systematic analysis framework.

In Zimmermann and Hommel (2003, 2005), probabilistic parameters derived from the ERTMS/ETCS-L3 specification are cast in a hierarchical approach supported by the TimeNET Tool (Zimmermann et al., 2006), modeling failures of the radio communication due to burst noise, connection losses, and handovers between radio stations. Though showing that handovers have a major impact on communication failures, modeling is restricted and inter-times between crossing of subsequent cell borders are represented as exponential variables, so as to avoid multiple concurrent generally-distributed timers (enabling restriction) (Ciardo et al., 1994; Choi et al., 1994). Analysis results are integrated within a model of location and MA exchange, using rare event simulation to derive the steady-state probability of an emergency brake, proving that trains cannot follow each other in braking distance as envisaged by the specification.

The approach of Carnevali et al. (2015) develops the results of Zimmermann and Hommel (2003, 2005) by presenting a communication model with multiple concurrent non-exponential timers, computing an upper bound on the first-passage probability that a train is stopped due to a communication failure. Reliability analysis of the ERTMS/ETCS-L3 is also addressed in Hermanns et al. (2005), leveraging discrete-event simulation supported by the Möbius Tool (Courtney et al., 2009). In Babczyński and Magott (2014), probabilistic parameters of Zimmermann and Hommel (2003, 2005) are used to develop a performance statechart model of the ERTMS/ETCS-L3 communication and operation, evaluating the probability of an emergency stop as a function of the distance between chasing trains. A formal methodology based on UML StateCharts is applied in Esposito et al. (2003) to verify the wireless communication protocol against CENELEC requirements, specifically EN 50128 (CENELEC, 2011) and EN 50159 (CENELEC, 2010). In Neglia et al. (2016), an actual implementation of a moving-block system for metro is considered, modeling the periodic communication between an on-board controller and an external ground controller. As typical in the literature on the ERTMS/ETCS-L3, except for (Carnevali et al., 2015), the approach focuses on the steady-state rate of train stops due to communication failures, which is evaluated through a closed-form expression under the simplifying assumption of independent packet losses.

In this paper, we characterize the first-passage time distribution to an emergency brake due to communication failures in the ERTMS/ETCS-L3. To this end, we exploit non-Markovian modeling to account for the regular quasi-periodic sequencing of cell borders, showing that handovers experienced by a pair of chasing trains are dependent events that periodically affect the communication availability (Section 3). Solution is achieved by combining analytic evaluation of failures due to burst noise and connection losses with transient analysis of an approximate model that accounts also for failures due to handovers (Section 4), leveraging Stochastic Time Petri Nets (STPNs) (Vicario et al., 2009) for model specification and the method of stochastic state classes for its analysis (Horváth et al., 2012). Given that behavior is recurrent over the hyper-period of the processes of message generation and cell border arrival, stochastic analysis within two hyper-periods is sufficient to derive an upper bound on the first-passage time distribution to an emergency brake over a time interval of arbitrary duration. A sensitivity analysis is performed with respect to different system parameters, supporting selection of a convenient headway in a trade-off between stop probability and line capacity/throughput (Section 5). Finally, we draw conclusions and discuss next steps (Section 6).

To make the paper self-contained, we recall STPNs and stochastic state classes (Section 2), and, for readability, we report proofs in the Appendix (Section A).

2. Preliminaries

2.1. Stochastic Time Petri Nets

An STPN is a tuple $\langle P, T, A^-, A^+, m_0, F, W, E, U\rangle$ where: P is the set of places; T is the set of transitions; $A^- \subseteq P \times T$ and $A^+ \subseteq T \times P$ are the sets of precondition and postcondition arcs, respectively; $m_0 : P \to \mathbb{N}$ is the initial marking; $F : T \to [0,1]^{[EFT_t, LFT_t]}$ associates each transition t with a Cumulative Distribution Function (CDF) $F(t) : [EFT_t, LFT_t] \to [0,1]$, where $EFT_t \in \mathbb{Q}_{\geqslant 0}$ and $LFT_t \in \mathbb{Q}_{\geqslant 0} \cup \{\infty\}$ are the *earliest* and *latest firing time*, respectively; $W : T \to \mathbb{R}_{>0}$ associates each transition with a weight; E and U associate each transition t with an enabling function $E(t) : \mathbb{N}^P \to \{true, false\}$ and an update function $U(t) : \mathbb{N}^P \to \mathbb{N}^P$, which associate each marking with a boolean value and a new marking, respectively.

A place p is an *input* or an *output* place for a transition t if $\langle p, t\rangle \in A^-$ or $\langle t, p\rangle \in A^+$, respectively. A transition t is *immediate* (IMM) if $EFT_t = LFT_t = 0$ and *timed* otherwise; a timed transition t is *exponential* (EXP) if $F_t(x) = 1 - e^{-\lambda x}$ over $[0, \infty]$ with $\lambda \in \mathbb{R}_{>0}$, and *general* (GEN) otherwise; a general transition t is *deterministic* (DET) if $EFT_t = LFT_t > 0$ and *distributed* otherwise; for each distributed transition t, we assume that F_t is the integral function of a Probability Density Function (PDF) f_t, i.e., $F_t(x) = \int_0^x f_t(y)dy$. IMM, EXP, GEN, and DET transitions are represented by thin black, thick white, thick black, and thick gray bars, respectively; enabling functions, update functions, and weights are annotated next to transitions as "? *expression*", "`place` = *expression*", and "*weight* = *value*", respectively.

The state of an STPN is a pair $\langle m, \tau\rangle$, where m is a marking and $\tau : T \to \mathbb{R}_{\geqslant 0}$ associates each transition with a time-to-fire. A transition is *enabled* by a marking if each of its input places contains at least one token and its enabling function evaluates to true; an enabled transition t is *firable* in a state if its time-to-fire is equal to zero. The next transition t to fire in a state $s = \langle m, \tau\rangle$ is selected among the set of firable transitions $T_{f,s}$ with probability $W(t)/\sum_{t_i \in T_{f,s}} W(t_i)$. When t fires, s is replaced with $s' = \langle m', \tau'\rangle$, where:

- m' is derived from m by: removing a token from each input place of t, which yields an intermediate marking m_{tmp}; adding a token to each output place of t, which yields a second intermediate marking m'_{tmp}; and, applying the update function $U(t)$ to m'_{tmp};
- τ' is derived from τ by: (i) reducing the time-to-fire of each *persistent* transition (i.e., enabled by m, m_{tmp} and m') by the time elapsed in s; (ii) sampling the time-to-fire of each *newly-enabled* transition t_n (i.e., enabled by m' but not by m_{tmp}) according to F_{t_n}; and (iii) removing the time-to-fire of each *disabled* transition (i.e., enabled by m but not by m').

In general, the marking process underlying an STPN is a Generalized Semi-Markov Process (GSMP) (Haas, 2006) where all times-to-fire run with the same speed and token moves are deterministic.

An STPN is a fully stochastic model: given an initial distribution for the vector of times-to-fire τ, the STPN identifies a unique probability space $\langle \Omega, \mathcal{F}, \mathbb{P}\rangle$, where the set of outcomes Ω includes all and only the feasible timed firing sequences of the model. By disregarding the stochastic parameters F and W, the STPN $\langle P, T, A^-, A^+, m_0, F, W, E, U\rangle$ identifies an underlying non deterministic model $\langle P, T, A^-, A^+, m_0, E, U\rangle$, which comprises an instance of the Time Petri Net (TPN) formalism (Lime and Roux, 2003; Berthomieu and Diaz, 1991) and whose set of feasible timed firing sequences coincides with Ω.

2.2. Analysis based on the method stochastic state classes

The method of stochastic state classes (Vicario et al., 2009; Horváth et al., 2012) permits the analysis of STPNs with multiple concurrent GEN transitions. Given a sequence of firings, a *stochastic state class* encodes the marking and the joint PDF of the times-to-fire of the enabled transitions and the absolute elapsed time τ_{age}. Starting from an initial stochastic state class, the *transient tree* of stochastic state classes that can be reached within a time $t_{\max}$ is enumerated, enabling derivation of continuous-time transient probabilities of markings (*forward transient analysis*), i.e., $p_m(t) := P\{M(t) = m\}\ \forall\, 0 \leqslant t \leqslant t_{\max}$, $\forall\, m \in \mathcal{M}$, where $\{M(t), \quad t \geq 0\}$ is the *marking process* describing the marking $M(t)$ of an STPN for each time $t \geqslant 0$ and $\mathcal{M}$ is the set of reachable markings.

If the STPN always reaches within a bounded number of firings a *regeneration*, i.e., a state satisfying the Markov condition, its marking process is a Markov Regenerative Process (MRP) (Choi et al., 1994), and its analysis can be performed enumerating stochastic state classes between any two regenerations. This results in a set of trees that permit to compute a local and a global kernel characterizing the MRP behavior, enabling evaluation of transient marking probabilities through the numerical solution of Markov renewal equations (*regenerative transient analysis*). Trees also permit to compute conditional probabilities of the Discrete Time Markov Chain (DTMC) embedded at regenerations and the expected time spent in any marking after each occurrence of any regeneration (Martina et al., 2016), supporting derivation of steady-state marking probabilities according to the Markov renewal theory (*regenerative steady-state analysis*).

While stochastic state classes support quantitative evaluation of an STPN model, the set Ω of behaviors of the STPN can be identified with simpler and more consolidated means through non-deterministic analysis of the underlying TPN model. In this case, the state space is covered through the method of *state classes* (Berthomieu and Diaz, 1991; Vicario, 2001; Dill, 1989; Lime and Roux, 2003), each made of a marking and a joint support for τ_{age} and the times-to-fire of the enabled tran-

sitions. In this approach, enumeration of state classes starting from an initial marking provides a representation for the continuous set of executions of an STPN, enabling verification of qualitative properties of the model, e.g., guarantee, with certainty, that a marking cannot be reached within a given time bound (*non-deterministic transient analysis*).

The ORIS Tool (ORIS Tool; Carnevali et al., 2011) efficiently implements the method of stochastic state classes, including regenerative transient, regenerative steady-state, and non-deterministic analyses.

3. Quantitative modeling of communication failures

In the communication between an RBC and a pair of trains, a PR sent by the foregoing train to the RBC and the corresponding MA sent by the RBC to the chasing train comprise an *end-to-end message*. Transient failures of the GSM-R may cause the loss of end-to-end messages, impairing the *up-link* transmission of PRs or the *down-link* transmission of MAs, due to three causes identified in the ERTMS/ETCS specification:

- *burst noise* causing temporary unavailability of the communication link;
- *connection losses* due to transient failures of devices;
- *handovers* between the communication areas of neighboring Base Transceiver Stations (BTSs), occurring when a train reaches the border of a GSM-R cell.

When multiple consecutive end-to-end messages are lost, the MA of the chasing train is not updated and an emergency brake may occur even when the trains are at safe distance. The maximum number M of consecutive losses that can be tolerated before an emergency brake depends on the headway distance s_h between trains, on their nominal speed v, and on the signalling period T_{msg} in the generation of PRs:

$$s_h \geqslant s_b + (M+2)\, T_{msg}\, v \tag{1}$$

where s_b is the minimum braking distance, and $(M+2)\, T_{msg}\, v$ is the distance covered by the chasing train since the *timestamp* of the PR of the last successful end-to-end transmission, i.e., the time when the foregoing train started the integrity check reported in that PR. For any given values of T_{msg} and v, a greater s_h reduces capacity but makes signalling immune to a larger number M of consecutive losses.

In this section, the three causes of failure are characterized through separated STPN models, and their impact on the communication availability is finally composed with an STPN model of generation and transmission of PRs and MAs. Stochastic parameters are derived following the methodology of Zimmermann and Hommel (2003, 2005), with some changes reflecting amendments introduced in the evolution of the ERTMS/ETCS specification (ERA, 2016, 2014; AA.VV., 2005). While models will be instantiated on selected significant values, parameters can be varied so as to cover most of the specification, under two restrictions that become crucial for the feasibility of the applied solution technique. The first restriction guarantees that, when the foregoing train generates a PR, the last MA sent by the RBC has already been delivered to the chasing train or lost.

Hypothesis 1. The delay τ_{tx} from the dispatch of a PR by the foregoing train until the receipt of the corresponding MA by the chasing train is not greater than the PR generation period T_{msg}.

Given that $\tau_{tx} \leqslant 5.56$ s and $T_{msg} \in [5, 12]$ s (Zimmermann and Hommel, 2003, 2005), the restriction rules out the cases with $T_{msg} \in [5, 5.56]$ s, but still covers most of the specification range with $T_{msg} \in [5.56, 12]$ s. With typical speed $v = 300$ km h^{-1} and braking distance $s_b = 3$ km (Zimmermann and Hommel, 2003, 2005), and with $T_{msg} = 6$ s equal to the minimum integer satisfying Hypothesis 1, the number of tolerated losses M turns out to be 2 if $s_h \in [5, 5.5)$ km, 3 if $s_h \in [5.5, 6)$ km, and 4 if $s_h \in [6, 6.5)$ km. Note that, according to Hypothesis 1, the delay from the timestamp of a PR until the receipt of the corresponding MA, usually termed *packet age* (Zimmermann and Hommel, 2003, 2005), is at most $2\, T_{msg}$.

The second restriction guarantees that the variability in the size of GSM-R cells is not such that one of the two trains can encounter two handovers while the other one remains within the same cell.

Hypothesis 2. The time T_{BTS} needed to cover the minimum distance s_{BTS} between two consecutive BTSs is not lower than the time T_h needed to travel the headway distance s_h, plus a time T_{rec} needed to establish a connection with the next BTS after a cell handover, plus a random time jitter τ_j needed to cover a variable distance s_j accounting for the possibly not regular displacement of BTSs, i.e., $T_h + \tau_j + T_{rec} \leqslant T_{BTS}$.

The latter restriction is satisfied in most of the range of variability of parameters in the actual practice. Based on Zimmermann and Hommel (2003, 2005) and ERA (2014), we consider $T_{rec} = 0.5$ s and $s_{BTS} = 7$ km. With typical speed $v = 300$ km h^{-1} and headway $s_h \in [5, 6]$ km, we obtain $T_{BTS} = s_{BTS}/v = 84$ s and $T_h = s_h/v \in [60, 72]$ s. To satisfy Hypothesis 2, we consider jitter $\tau_j \in [0, 10]$ s, which corresponds to $s_j = v\tau_j \in [0, 0.83]$ km, representing variations in BTS displacement in the order of 10% of s_{BTS}. For a comparative study with respect to nominal parameter values, in the experiments we consider values of v within the range $[200, 300]$ km h^{-1}, without violating Hypothesis 2 while still allowing a margin $s_j = \in [0, 0.83]$ km.

3.1. Model of burst noise

As in Zimmermann and Hommel (2003, 2005), GSM-R failures due to burst noise are represented using the Gilbert-Elliot model, which alternates availability and burst periods of exponential duration. The model is represented by the STPN in Fig. 1, where the EXP transitions `startBurst` and `endBurst` represent the arrival and the termination of a burst, respectively. In AA.VV. (2005), the arrival time of bursts is prescribed to be (i) larger than 20 s in 95% of the cases and (ii) larger than 7 s in 99% of the cases, while the duration of bursts is prescribed to be (iii) lower than 0.8 s in 95% of the cases and (iv) lower than 1 s in 99% of the cases. According to this, rate $\lambda_{\text{burst}} = 0.002565\ \text{s}^{-1}$ is safely defined so as to fit requirement (i), yielding an EXP random variable τ_{burst} *stochastically smaller* than the EXP random variable τ'_{burst} with rate $\lambda'_{\text{burst}} = 0.001436\ \text{s}^{-1}$ that would fit requirement (ii), i.e., $P\{\tau_{\text{burst}} \leqslant x\} \geqslant P\{\tau'_{\text{burst}} \leqslant x\}\ \forall x \in [0, \infty)$. In a similar manner, $\mu_{\text{burst}} = 3.7446\ \text{s}^{-1}$ is safely selected so as to fit requirement (iii), yielding an EXP random variable v_{burst} *stochastically larger* than the EXP random variable v'_{burst} with rate $\mu'_{\text{burst}} = 4.6051\ \text{s}^{-1}$ that would fit requirement (iv), i.e., $P\{v_{\text{burst}} \leqslant x\} \leqslant P\{v'_{\text{burst}} \leqslant x\}\ \forall x \in [0, \infty)$.

While other distributions could be applied to fit the specification, for instance continuous phase-type distributions (Horvath et al., 2000; Longo and Scarpa, 2009), in this specific model EXP distributions achieve a good trade-off between effectiveness and simplicity of approximation.

3.2. Model of connection losses

A connection loss takes some time to be detected and re-established, and it must be canceled and re-tried if it is not successful within a predefined timeout. Following Zimmermann and Hommel (2003, 2005), this behavior is represented by the STPN in Fig. 2, where: connection losses are assumed to occur with exponential inter-arrival times, represented by the EXP transition `loss` with rate $\lambda_{\text{loss}} = 2.77 \times 10^{-8}\ \text{s}^{-1}$, fitting the requirement that the probability of a connection loss per hour be not larger than 10^{-4}; and, the detection delay is represented the DET transition `indicate`, fitting the worst case detection delay $T_{\text{indicate}} = 1$ s.

According to AA.VV. (2005): the timeout is modeled by the DET transition `fail`, fitting the worst case timeout $T_{\text{fail}} = 10$ s prescribed by the specification; and, the result in each trial is represented by the IMM transitions `estP` and `failP`, with weights set equal to 0.9999 and 0.0001, respectively, so as to fit the worst case success probability 0.9999 prescribed by the specification.

Finally, the reconnection time in case of success shall be (i) less than 5 s in 95% of the cases and (ii) less than 7.5 s in 99% of the cases (AA.VV., 2005). These prescriptions are safely modeled by the EXP transition `connect` with rate $\lambda_{\text{connect}} = 0.5991\ \text{s}^{-1}$ fitting requirement (i), yielding an EXP random variable stochastically larger than the one with rate $\lambda'_{\text{connect}} = 0.6140\ \text{s}^{-1}$ that would fit requirement (ii).

3.3. Model of cell handovers

BTSs are installed at regular distance typically every $s_{\text{BTS}} = 7$ km (ERA, 2014), in normal operating conditions trains travel at approximately constant speed, and railway lines are nearly straight. According to this, disconnections due to cell handovers occur in a quasi periodic manner, at different times for the foregoing and the chasing train. To capture the quasi periodic structure of this behavior, the arrival of handovers at the foregoing train is represented as a periodic process with jitter, with an initial phase offset, while the arrival at the chasing train is modeled so as to occur after a delay corresponding to the time spent to cover the headway. Specifically, the period $T_{\text{BTS}} = 84$ s is defined so as to fit the time spent to traverse a cell

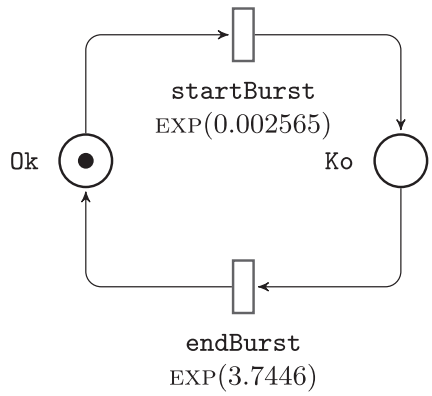


Fig. 1. STPN model of communication failures due to burst noise (time expressed in s).

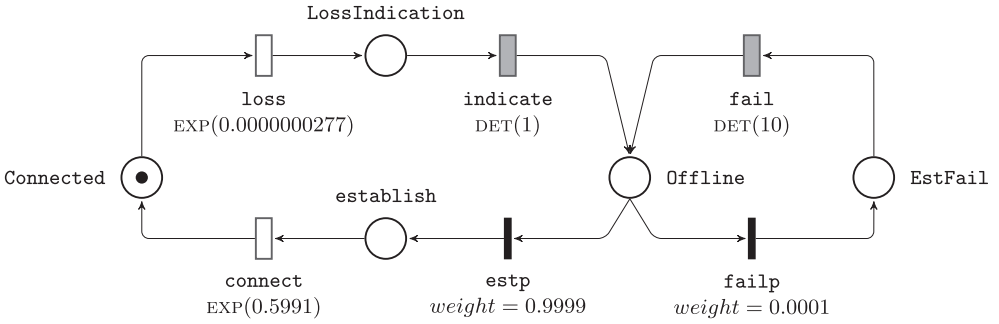


Fig. 2. STPN model of communication failures due to connection losses (time expressed in s).

with size $s_{\text{BTS}} = 7$ km at speed $v = 300$ km h^{-1}; the jitter adds a uniform delay after the start of each period so as to reproduce the variability in the distance between consecutive BTSs; the initial offset is a deterministic value; and, though being a random variable in practice, also the headway delay is modeled as a deterministic value, so that it can be varied (with the initial offset and the nominal speed) in a parametric evaluation, providing better insight on the sensitivity of the ERTMS/ETCS-L3 performance.

The model is represented by the STPN in Fig. 3. A token in places $\texttt{Handover}_{\text{ft}}$ and $\texttt{Handover}_{\text{ct}}$ means that the foregoing train and the chasing train are disconnected for a handover, respectively; the handover of the chasing train occurs with a DET delay after the handover of the foregoing train, represented by transition `headway` whose duration T_{h} corresponds to the time spent to cover the headway s_{h}; for each train, reconnection occurs after a DET delay represented by transitions $\texttt{reconnect}_{\text{ft}}$ and $\texttt{reconnect}_{\text{ct}}$, respectively, fitting the maximum allowed reconnection time of 0.3 s (Zimmermann and Hommel, 2003, 2005). The firing of transition $\texttt{cellBorder}_{\text{ft}}$ or $\texttt{offset}_{\text{ft}}$ brings a token in place `Jitter`, representing the arrival of the foregoing train at the nominal point of a cell border; the actual arrival occurs after a random jitter, modeled by the uniformly distributed transition `jitter`. The DET transition $\texttt{offset}_{\text{ft}}$ fires only once and accounts for the first arrival at a nominal cell border after a delay O_{ft}; whereas, transition $\texttt{cellBorder}_{\text{ft}}$ has no input places and thus fires periodically every 84 s starting from the first arrival, which empties place $\texttt{Offset}_{\text{ft}}$ and thus sets true the enabling function $?\texttt{Offset}_{\text{ft=0}}$.

3.4. Composed model

The STPN of Fig. 4 represents the periodic generation of PRs by the EVC of the foregoing train, the subsequent transmission to the RBC, and then the transmission of the MA from the RBC to the EVC of the chasing train, following the communication model of Zimmermann and Hommel (2003, 2005): the DET transition `genMsg` with value T_{msg} models the periodic execution of integrity checks by the foregoing train and the consequent PR transmission to the RBC; the DET transition `rbc` accounts for the maximum time $T_{\text{RBC}} = 0.5$ s that PR processing at the RBC is supposed to take; and, the GEN transitions `transmitUp` and `transmitDown` represent the time τ_{link} spent to send PRs from the foregoing train to the RBC and MAs from the RBC to the chasing train, respectively, and they are associated with a piecewise uniform distribution over the intervals $[0.55, 0.65)$ s, $[0.65, 1.35)$ s, and $[1.35, 2.55)$ s, with probability 0.95, 0.04, and 0.01, respectively, derived by over-approximating the actual transmission time of a PR/MA with 0.15 s and by fitting the specification that the transmission delay of a PR/MA is between 0.4 s and 0.5 s on average, less than 0.5 s in 95% of the cases, less than 1.2 s in 99% of the cases, and less than 2.4 s in 99.99% of the cases (Zimmermann and Hommel, 2003, 2005). Specifically:

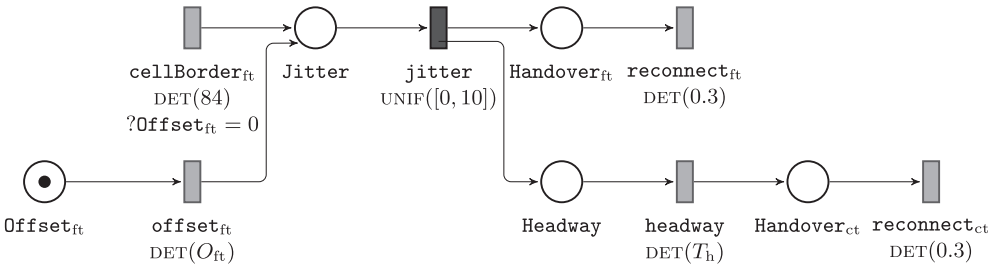


Fig. 3. STPN Model of communication failures due to cell handovers for train speed $v = 300$ km h^{-1} (time expressed in s). Transition `headway` takes a deterministic value $T_{\text{h}} \in [54, 72]$ s and transition $\texttt{offset}_{\text{ft}}$ takes a deterministic value $O_{\text{ft}} \in [0, 84 - T_{\text{h}}]$ s.

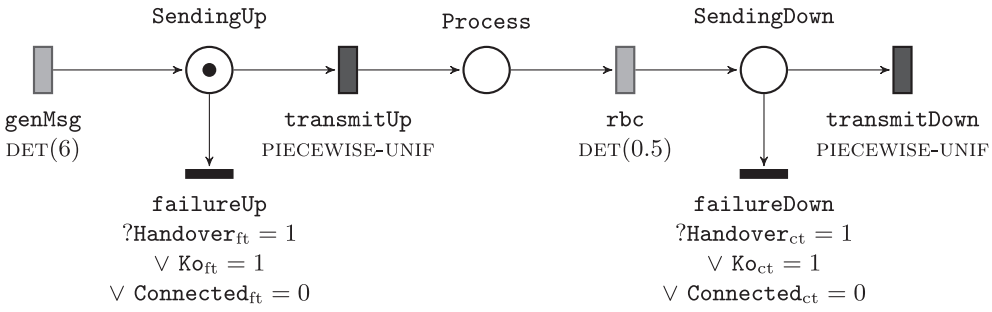


Fig. 4. STPN model of end-to-end message processing and transmission (time expressed in s). Transitions `transmitUp` and `trasmitDown` have a piecewise uniform PDF defined in Eq. (2). The enabling functions of transitions `failureUp` and `failureDown` compose the STPN with the models of disconnections due to burst noise (Fig. 1), connection losses (Fig. 2), and handovers (Fig. 3), for the up-link and down-link communication (with places and transitions subscripted as "`ft`" and "`ct`", respectively).

$$f_{\texttt{transmitUp}}(x) = f_{\texttt{transmitDown}}(x) = \begin{cases} 9.5 & \text{if } 0.55 \leqslant x < 0.65 \\ 0.057143 & \text{if } 0.65 \leqslant x < 1.35 \\ 0.008333 & \text{if } 1.35 \leqslant x < 2.55 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The IMM transitions `failureUp` and `failureDown` account for losses occurring in the down-link and up-link connection, respectively. The enabling functions on the IMM transitions, "$?\texttt{Handover}_{\texttt{ft}} = 1 \vee \texttt{Ko}_{\texttt{ft}} = 1 \vee \texttt{Connected}_{\texttt{ft}} = 0$" for the uplink and "$?\texttt{Handover}_{\texttt{ct}} = 1 \vee \texttt{Ko}_{\texttt{ct}} = 1 \vee \texttt{Connected}_{\texttt{ct}} = 0$" for the downlink, compose the STPN with the models of Figs. 1–3 characterizing the times of unavailability of the connection due to burst noise, connection losses, and handover disconnections, respectively. In so doing, the IMM transition `failureUp` models the loss of a PR due to a failure in the communication from the foregoing train to the RBC, caused by the crossing of a cell border by the foregoing train (represented by a token in place $\texttt{Handover}_{\texttt{ft}}$) or due to a burst of noise in the up-link communication (represented by a token in place $\texttt{Ko}_{\texttt{ft}}$) or due to a connection loss in the up-link communication (represented by the absence of a token in place $\texttt{Connected}_{\texttt{ft}}$). In a similar manner, the IMM transition `failureDown` models the loss of a MA due to a failure in the communication from the RBC to the chasing train.

4. Quantitative evaluation of communication failures

We characterize the first passage time distribution $p(M,t)$ of the event E_M that M consecutive end-to-end messages are lost, due to a failure in the up-link transmission of a PR from the foregoing train to the RBC or in the down-link transmission of an MA from the RBC to the chasing train:

$$p(M,t) := P\{E_M \text{ occurs within } [0,t)\} \tag{3}$$

To simplify the derivation, we evaluate the probability $\phi(M,k)$ that E_M occurs within the time interval $[0, kT_{\text{msg}})$, i.e., within the time instant that precedes the dispatch of the PR of the $(k+1)$-th end-to-end message by the foregoing train:

$$\phi(M,k) := P\{E_M \text{ occurs within } [0,kT_{\text{msg}})\} = p(M,kT_{\text{msg}}) \tag{4}$$

Values of $\phi(M,k)$ provide a bounded approximation of $p(M,t)$ for any $t \in \mathbb{R}_{\geqslant 0}$, given that:

$$\phi\left(M, \left\lfloor \frac{t}{T_{\text{msg}}} \right\rfloor\right) \leqslant p(M,t) \leqslant \phi\left(M, \left\lceil \frac{t}{T_{\text{msg}}} \right\rceil\right) \tag{5}$$

Note that the laxity of the bound of Eq. (5) is practically negligible due to the fact that, by design, the expected time of E_M is significantly larger than T_{msg}.

In principle, $\phi(M,k)$ might be evaluated by directly solving the flat model that integrates the STPNs of Figs. 1–4. However, stochastic transient analysis implemented in the ORIS Tool (ORIS Tool), either forward or regenerative, would not be practically feasible due to the complexity of behaviors resulting from multiple concurrent GEN timers with overlapping activity intervals. This complexity is also evident from the results of non-deterministic analysis of the underlying TPN, which was stopped after the enumeration of more than 300,000 state classes, including at least two enabled GEN transitions in each class and allowing an unbounded number of firings without encountering a regeneration in each feasible behavior.

Moreover, approaches based on the approximation of GEN transitions, either completely exploiting phase-type distributions (Horvath et al., 2000; Longo and Scarpa, 2009), or allowing at most one expolynomially distributed timer in each marking (Lindemann and Thümmler, 1999), appear not well suited due to the many concurrent GEN transitions with firmly bounded support. Though avoiding model approximation, stochastic simulation would suffer the different order of

magnitude of durations, the presence of rare events, and the need to perform the evaluation for a time span covering the transmission of M consecutive end-to-end messages. This complexity was already addressed in Zimmermann and Hommel (2003, 2005) for a model with a simpler structure, requiring rare-event simulation to derive measures of interest.

Given that burst noise, connection losses, and cell handovers are due to independent physical phenomena, their impact on the GSM-R communication can be separately assessed. To this end, communication unavailability due to each failure type is evaluated either analytically or exploiting numerical solution of the corresponding STPN model by the method of stochastic state classes implemented in the ORIS Tool; note that the above mentioned approaches (Horvath et al., 2000; Longo and Scarpa, 2009; Lindemann and Thümmler, 1999) could be applied as well to solve the individual STPN model of each failure type, though approximation of GEN transitions would require to achieve a trade-off between result accuracy and computational complexity. Evaluation will show that: multiple losses due to burst noise can be regarded as independent events (Section 4.1); the effect of connection losses on the communication availability is negligible with respect to the impact of burst noise (Section 4.2); and, for values of M with practical relevance, no more than 2 out of M consecutive end-to-end messages can be lost due to handovers, so that at least $M-2$ out of M consecutive failures are due to burst noise or connection losses (Section 4.3). Combination of all these results enables a compositional strategy for feasible and accurate evaluation of the first passage time distribution to an emergency brake due to any cause of communication failure (Section 4.4).

4.1. Burst noise

Losses of end-to-end messages due to burst noise can be accurately approximated as independent events:

Lemma 4.1. *For any set of m not necessarily consecutive end-to-end messages, the probability that burst noise causes the loss of all the m end-to-end messages is:*

$$p_{\text{burst}}(m) = P_{\text{burst}}^{m} \tag{6}$$

where P_{burst} is the probability that burst noise causes the loss of a single end-to-end message.

In turn, P_{burst} is derived from the probability $P_{\text{link,burst}}$ that burst noise impairs the up-link or the down-link transmission:

Lemma 4.2.

$$P_{\text{burst}} = 2P_{\text{link,burst}} - P_{\text{link,burst}}^{2} \tag{7}$$

where $P_{\text{link,burst}}$ is the probability that burst noise impairs either the up-link or the down-link transmission, and it can be upper-bounded as:

$$P_{\text{link,burst}} \leqslant \widetilde{P}_{\text{link,burst}} = \frac{\lambda_{\text{burst}}}{\lambda_{\text{burst}}+\mu_{\text{burst}}} + \frac{\mu_{\text{burst}}}{\lambda_{\text{burst}}+\mu_{\text{burst}}}\left(1 - e^{-\lambda_{\text{burst}}\max\{\tau_{\text{link}}\}}\right) \tag{8}$$

where τ_{link} is the random variable representing the duration of message transmission up-link or down-link, distributed according to the PDF defined in Eq. (2), so that $\max\{\tau_{\text{link}}\} = 2.55$ s is thus the maximum duration of the up-link or down-link transmission.

Finally, according to Lemmas 4.1 and 4.2, $p_{\text{burst}}(m)$ can be upper-bounded as:

$$p_{\text{burst}}(m) \leqslant \tilde{p}_{\text{burst}}(m) = \left(2\widetilde{P}_{\text{link,burst}} - \widetilde{P}_{\text{link,burst}}^{2}\right)^{m} \tag{9}$$

To provide an intuition about the order of magnitude, by substituting values of parameters, we obtain $\widetilde{P}_{\text{burst,link}} = 0.00719$, which yields $P_{\text{burst}} = p_{\text{burst}}(1) \leqslant 0.01432$, $p_{\text{burst}}(2) \leqslant 0.205 \cdot 10^{-2}$, $p_{\text{burst}}(3) \leqslant 0.294 \cdot 10^{-5}$, and $p_{\text{burst}}(4) \leqslant 0.421 \cdot 10^{-7}$.

4.2. Connection losses

As opposed to burst noise, connection losses have a settling time that is not negligible with respect to the time that may elapse between sequential PR transmissions, and losses in consecutive end-to-end message are thus positively correlated:

Lemma 4.3. *The probability $p_{\text{loss}}(m)$ that connection losses impair the transmission of m not necessarily consecutive end-to-end messages is upper-bounded by the probability of m consecutive losses, and can thus be estimated as:*

$$p_{\text{loss}}(m) \leqslant P_{\text{loss}} \cdot \sum_{q \in Q} U_q \cdot u_{\text{loss},q}((m-2)T_{\text{msg}} + \delta) \tag{10}$$

where: P_{loss} is the probability that connection losses impair the transmission of a single end-to-end message; $Q = \{lossIndication, establish, estFail\}$ is the set of the considered initial markings; U_q is the steady state probability of marking q; $u_{\text{loss},q}(t)$ is the transient unavailability of the radio channel due to connection losses conditioned to the hypothesis that the initial marking of the model of Fig. 2 is q; and, δ is the minimum time between two consecutive transmissions from the RBC.

As in the case of burst noise, P_{loss} is derived from the probability $P_{\text{link,loss}}$ that connection losses impair the up-link or the down-link transmission, which are independent events referring to failures of different devices on board of the foregoing and the chasing train:

Lemma 4.4. *The probability that connection losses impair the transmission of an end-to-end message is:*

$$P_{\text{loss}} = 2P_{\text{link,loss}} - P_{\text{link,loss}}^2 \tag{11}$$

where $P_{\text{link,loss}}$ is the probability that connection losses impair either the up-link or the down-link transmission, and it can be upper-bounded as:

$$P_{\text{link,loss}} \leqslant \widetilde{P}_{\text{link,loss}} = U_{\text{loss}} + (1 - U_{\text{loss}})\left(1 - e^{-\lambda_{\text{loss}} \max\{\tau_{\text{link}}\}}\right) \tag{12}$$

where U_{loss} is the steady-state value of communication unavailability due to connection losses, and τ_{link} is the random variable representing the duration of message transmission up-link or down-link, distributed according to the PDF defined in Eq. (2).

Finally, according to Lemmas 4.3 and 4.4, $p_{\text{loss}}(m)$ can be upper-bounded:

$$p_{\text{loss}}(m) \leqslant \tilde{p}_{\text{loss}}(m) = (2\widetilde{P}_{\text{link,loss}} - \widetilde{P}_{\text{link,loss}}^2) \cdot \sum_{\mathrm{q} \in \mathrm{Q}} U_{\mathrm{q}} \cdot u_{\text{loss,q}}((m-2)T_{\text{msg}} + \delta) \tag{13}$$

Also in this case, to provide an intuition about the order of magnitude, by substituting parameter values, we obtain $\widetilde{P}_{\text{loss,link}} = 0.145 \cdot 10^{-6}$, which yields $P_{\text{loss}} = p_{\text{loss}}(1) \leqslant 0.289 \cdot 10^{-6}$, $p_{\text{loss}}(2) \leqslant 0.344 \cdot 10^{-7}$, $p_{\text{loss}}(3) \leqslant 0.106 \cdot 10^{-8}$, and $p_{\text{loss}}(4) \leqslant 0.291 \cdot 10^{-10}$. Note that, for values of interest of M comprised between 2 and 4, the probability that m out of M consecutive messages are not delivered to the chasing train due to connection losses is by far lower than the probability that they are lost due to burst noise.

4.3. Cell handovers

Cell handovers are the most probable events causing the loss of an end-to-end message, as already pointed out in Zimmermann and Hommel (2003, 2005). However, non-deterministic transient analysis of the composition of models in Figs. 3 and 4 proves, with certainty, that no more than 2 out of M consecutive end-to-end messages can be lost due to handovers for values of M with practical relevance.

The composed model is shown in Fig. 5, where the IMM transitions `failureUp` and `failureDown` of Fig. 4 are replaced with the IMM transitions `handoverUp` and `handoverDown`, respectively, associated with enabling functions "$?\texttt{Handover}_{\text{ft}} = 1$" and "$?\texttt{Handover}_{\text{ct}} = 1$", respectively, so as to remove references to the models of burst noise and connection losses, which are analyzed separately. By Hypotheses 1 and 2, the model behavior is periodic over the *hyper-period HP*, defined as the least common multiple of the period T_{msg} of message generation and the period T_{BTS} of cell border arrival. Although the model behavior repeats itself on a period of duration *HP*, the analysis of the model state space in the interval

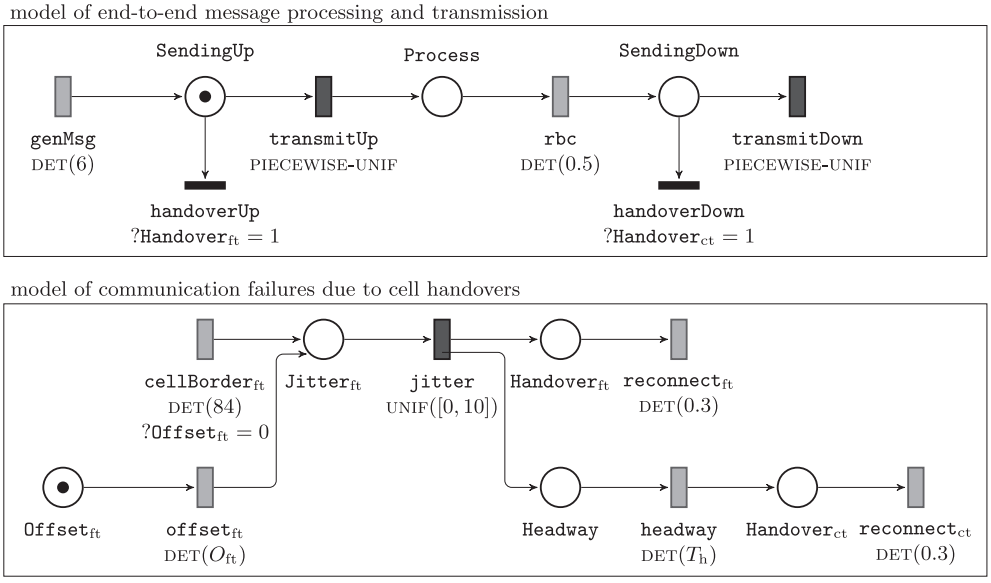


Fig. 5. STPN model accounting for end-to-end message processing and transmission, and for communication failures due to cell handovers (time expressed in s).

$[0, HP]$ would not permit to observe *border effects* over the start of a hyper-period, i.e., losses in a sequence of M consecutive end-to-end messages spanning over two consecutive hyper-periods. Therefore, non-deterministic transient analysis is performed over the interval $[0, 2HP]$ to prove that no more than 2 out of M consecutive end-to-end messages can be lost due to handovers for $M \leqslant HP/T_{\text{msg}}$.

Lemma 4.5. *If $M \leqslant HP/T_{\text{msg}}$, cell handovers affect at most 2 out of M consecutive end-to-end messages.*

Given that $HP/T_{\text{msg}} = 14$, the hypothesis $M \leqslant HP/T_{\text{msg}}$ of Lemma 4.5 is largely verified for values of M with practical relevance, which in fact range between 2 and 4.

4.4. Combined effect of the different causes of communication failures

We characterize the first-passage time distribution of an emergency brake triggered by the loss of M consecutive end-to-end messages. To this end, we develop the derivation through the following steps:

- We derive an upper bound $\tilde{\phi}(M,k)$ on $\phi(M,k)$ through numerical solution of a *model of message processing, transmission, and loss*, accounting for burst noise, connection losses, and cell handovers.
- We derive an upper-bound $\tilde{\Phi}$ on the probability Φ that M consecutive losses occur *within a hyper-period*.
- We extend the evaluation *beyond a hyper-period*, deriving an upper-bound $\tilde{\beta}(M,H)$ on the first-passage probability that M consecutive end-to-end messages are lost within H hyper-periods.

4.4.1. A model of message processing, transmission, and loss

The first-passage probability $\phi(M,k)$ that event E_M (i.e., the event that M consecutive end-to-end messages are lost) occurs within the time interval $[0, kT_{\text{msg}})$ can be equivalently defined as:

$$\phi(M,k) := P\{\exists h \in [1,k] \text{ such that } E_M \text{ occurs within } \left[(h-1)T_{\text{msg}}, hT_{\text{msg}}\right)\} \tag{14}$$

By the law of total probability:

$$\phi(M,k) = \sum_{h=M}^{k} P\{E_{M,h} \,|\, \neg E_{M,i} \forall i \in [1,h)\} \tag{15}$$

where $E_{M,h}$ is the event that E_M occurs within the time interval $[(h-1)T_{\text{msg}}, hT_{\text{msg}})$ (i.e., the event that the h-th end-to-end message is the M-th end-to-end message of a sequence of M consecutively lost end-to-end messages), and $\neg E_{M,h}$ is the event that $E_{M,h}$ does not occur. According to this, $E_{M,h} \,|\, \neg E_{M,i} \forall i \in [1,h)$ is the event that the loss of the h-th end-to-end message yields the first occurrence of event E_M. This condition occurs if and only if no sequence of M consecutive end-to-end messages was lost within $[0, (h-M)T_{\text{msg}})$ and all the $M-1$ end-to-end messages sent within the time interval $[(h-M)T_{\text{msg}}, (h-1)T_{\text{msg}})$ were lost:

Lemma 4.6. *The loss of the h-th end-to-end message causes the first occurrence of E_M if and only if: (i) $\neg E_{M,i} \forall i \in [1, h-1]$ and (ii) all the $M-1$ messages sent within $[(h-M)T_{\text{msg}}, (h-1)T_{\text{msg}})$ were lost.*

Moreover, the successful delivery of the $(h-M)$-th end-to-end message is a necessary condition for the loss of the h-th end-to-end message to yield the first occurrence of event E_M:

Lemma 4.7. *The loss of the h-th end-to-end message can cause the first occurrence of E_M only if the $(h-M)$-th end-to-end message was correctly delivered.*

By Lemmas 4.6 and 4.7 and the results of Sections 4.1,4.2,4.3, an upper bound on $\phi(M,k)$ can be derived by: (i) considering that event $E_{M,h}$ is yielded if cell handovers cause $m \leqslant 2$ losses among the M end-to-end messages sent within $[(h-M)T_{\text{msg}}, hT_{\text{msg}})$, while burst noise or connection losses impair the remaining $M-m$ end-to-end messages and (ii) neglecting, for burst noise and connection losses, the conditioning that no sequence of M consecutive end-to-end messages is lost within $[0, (h-1)T_{\text{msg}})$:

Lemma 4.8. *$\phi(M,k)$ is upper-bounded by the following expression $\forall k > 0$:*

$$\phi(M,k) \leqslant \sum_{h=M}^{k} \sum_{m=0}^{2} p_{\text{handover}}(M,h,m) \cdot \sum_{n=0}^{M-m} p_{\text{burst}}(n) \cdot p_{\text{loss}}(M-m-n) \tag{16}$$

where $p_{\text{handover}}(M,h,m) = P\{H^m_{M,h} \,|\, Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1, h-M]\}$; $H^m_{M,h}$ is the event that m end-to-end messages are lost due to cell handovers within $[(h-M)T_{\text{msg}}, hT_{\text{msg}})$; Z_{h-M} is the event that the $(h-M)$-th end-to-end message is delivered; and, $p_{\text{handover}}(0) = 1$, $p_{\text{burst}}(0) = 1$, and $p_{\text{loss}}(0) = 1$.

By Lemma 4.8, an upper-bound $\tilde{\phi}(M,k)$ on $\phi(M,k)$ can be conveniently evaluated through numerical solution of a model that extends the STPN of Fig. 5 with M places acting as *watch variables* that keep memory of which among the last M end-to-end messages have been lost due to cell handovers. Without loss of generality, Fig. 6 shows the model for $M = 4$:

- Place $\mathtt{W}_0$ models the watch variable associated with the last end-to-end message: it is empty if the message has been lost, and it contains a token otherwise (i.e., if neither the PR nor the MA of that end-to-end message have been lost, though they may have not been delivered yet).
  Place $\mathtt{W}_m$, with $m \in [1, M-1]$, represents the watch variable associated with the $(m+1)$-th to last end-to-end message: it is empty if the message has been lost, and it contains a token otherwise (i.e., if both the PR and the MA of that end-to-end message have been correctly delivered).
- At each PR generation, the update function of `genMsg` (i.e., "$\mathtt{W}_{M-1} = \mathtt{W}_{M-2}, \ldots, \mathtt{W}_1 = \mathtt{W}_0, \mathtt{W}_0 = 1$") assigns the marking of places $\mathtt{W}_0, \ldots, \mathtt{W}_{M-2}$ to places $\mathtt{W}_1, \ldots, \mathtt{W}_{M-1}$, respectively, and assigns a token to place $\mathtt{W}_0$, loosing memory of whether the $(M+1)$-th to last end-to-end message has been delivered or lost, and starting to keep memory of the last end-to-end message. Whenever a handover impairs the up-link or the down-link transmission, the update function of `handoverUp` and `handoverDown` (i.e., "$\mathtt{W}_0 = 0$") flushes the watch variable place referring to the last end-to-end message.
  Hence, in any marking $n \in \mathcal{M}$ where $\mathcal{M}$ is the set of reachable markings of the model, the number of end-to-end messages lost due to handovers among the last M end-to-end messages is:

$$m = M - \sum_{j=0}^{M-1} n(\mathtt{W}_j) \tag{17}$$

- When an end-to-end message is delivered or lost due to a handover, a token arrives in place `Check` (through the post-condition arc from `transmitDown` to `Check` or through the update function "$\mathtt{Check} = 1$" of `handoverUp` and `handoverDown`), enabling the IMM transitions `failure` and `noFailure`. These transitions have weight f and $1-f$, where f is the probability that $M-m$ out of the last M end-to-end messages are lost due to burst noise or connection losses, not conditioned on the assumption that no sequence of M consecutive losses occurred before:

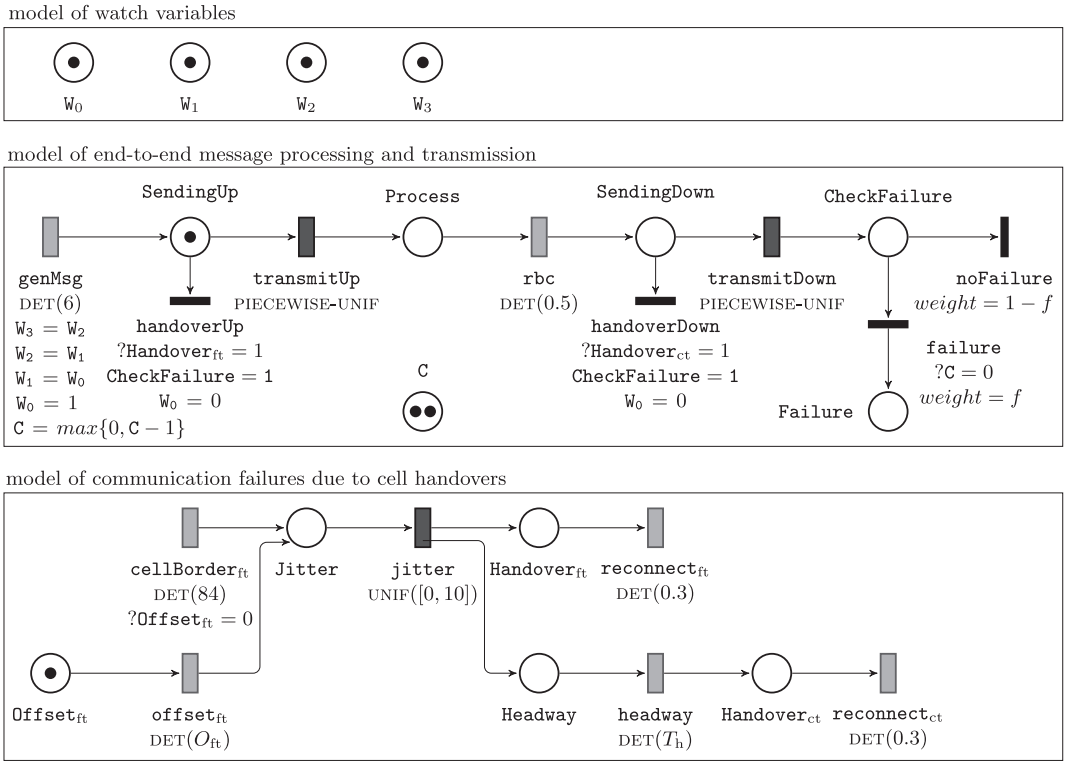


Fig. 6. STPN model of end-to-end message processing, transmission, and loss, including $M = 4$ places $\mathtt{W}_0, \ldots, \mathtt{W}_3$ keeping memory of the number m of end-to-end messages that were lost due to handovers among the last M messages (see Eq. (17)), and with two IMM transitions `failure` and `noFailure` with weight f and $(1-f)$ (see Eq. (18)), respectively, over-approximating the probability that the remaining $M-m$ messages were lost due to burst noise or connection losses (time expressed in s).

$$f = \sum_{n=0}^{M-m} \tilde{p}_{\text{burst}}(n) \cdot \tilde{p}_{\text{loss}}(M-m-n) \tag{18}$$

Note that transition `failure` cannot fire before the PR of the M-th end-to-end messages has been generated: in fact, `failure` is enabled only if place C is empty; place C contains $M-2=2$ tokens in the initial marking; and, transition `genMsg` reduces by one the number of tokens in C at each firing, until the place is empty (through the update function $\mathtt{C} = max\{0, \mathtt{C}-1\}$).

By Hypotheses 1 and 2, the model reaches a regeneration within a bounded number of firings at least at the end of each hyper-period. According to this, the model can be analyzed through regenerative transient analysis with absorbing condition $\mathtt{Failure}=1$, reducing the evaluation of the probability measure of paths that encounter M consecutive losses *within* $[0, kT_{\text{msg}})$ to the evaluation of the probability measure of any marking such that place `Failure` contains a token *at* time $(kT_{\text{msg}})^{-}$:

$$\tilde{\phi}(M,k) = \sum_{i \in \mathcal{M}^a} p_i((kT_{\text{msg}})^{-}) \tag{19}$$

where $\mathcal{M}^a$ is the set of reachable markings of the model that satisfy the absorbing condition $\mathtt{Failure}=1$, and $(kT_{\text{msg}})^{-}$ is the time instant that precedes the dispatch of the PR of the $(k+1)$-th end-to-end message.

4.4.2. Evaluation within a hyper-period

We define Φ as the probability that M consecutive end-to-end messages are lost within a hyper-period of duration HP, normalized with respect to the probability of behaviors within a hyper-period. To take into account border effects over the start of a hyper-period, Φ can be conveniently expressed as a function of probabilities $\phi(M,k)$ computed on the second hyper-period $[HP, 2HP]$:

$$\Phi := \frac{\phi\left(M, \frac{2HP}{T_{\text{msg}}}\right) - \phi\left(M, \frac{HP}{T_{\text{msg}}}\right)}{1 - \phi\left(M, \frac{HP}{T_{\text{msg}}}\right)} \tag{20}$$

Given that transient probabilities of markings of the model shown in Fig. 6 within the time interval $[HP, 2HP]$ are not conditioned by the initial values of watch variables (Lemma 4.9 and Corollary 4.1), an upper bound on Φ can be derived from probabilities $\tilde{\phi}(M,k)$ for $k \in [HP/T_{\text{msg}}, 2HP/T_{\text{msg}}]$, which can be computed through regenerative transient analysis of the model up to time $2HP$ (Lemma 4.10).

Specifically, at multiples of T_{msg}, the marking of places $\mathtt{W}_{M-2}, \ldots, \mathtt{W}_0$ is assigned to places $\mathtt{W}_{M-1}, \ldots, \mathtt{W}_1$, respectively, and a token is assigned to place $\mathtt{W}_0$, loosing memory of the previous marking of place $\mathtt{W}_{M-1}$. In so doing, for any time $t_0 \geqslant 0$ and $k \in \mathbb{N}$ s.t. $k < M$, memory of the marking of places $\mathtt{W}_{M-k}, \ldots, \mathtt{W}_{M-1}$ at time t_0 is lost at time $\lceil t_0/T_{\text{msg}} \rceil T_{\text{msg}} + kT_{\text{msg}}$:

Lemma 4.9. *$\forall t_0 \geqslant 0, \forall k \in \mathbb{N}$ s.t. $k < M, \forall m \in \mathcal{M}_{P \setminus \Omega_{k+1}^{M-1}}, \forall w \in \mathcal{M}_{\Omega_{M-k-1}^{M-1}}$:*

$$P\left\{M_{P \setminus \Omega_{k+1}^{M-1}}(t) = m \,|\, M_{\Omega_{M-k-1}^{M-1}}(t_0) = \omega\right\} = P\left\{M_{P \setminus \Omega_{k+1}^{M-1}}(t) = m\right\} \tag{21}$$

where $t \geqslant (\lfloor t_0/T_{\text{msg}} \rfloor + 1)\, T_{\text{msg}} + kT_{\text{msg}}$; $\Omega_i^{M-1} = \{W_j\}_{j=i}^{M-1}$ if $i \in \mathbb{N}$ s.t. $i \leqslant M$; $\Omega_M^{M-1} = \varnothing$; P is the set of places of the model shown in Fig. 6; and, $\mathcal{M}_Q$ is the set of reachable markings for places in $Q \subseteq P$.

As a consequence of Lemma 4.9, transient probabilities of markings of the model shown in Fig. 6 at any time $t \geqslant HP$ are not conditioned by the initial values of watch variables:

Corollary 4.1. *$\forall t \geqslant HP, \forall m \in \mathcal{M}, \forall w \in \mathcal{M}_{\Omega_0^{M-1}}$:*

$$P\{M_P(t) = m \,|\, M_{\Omega_0^{M-1}}(0) = \omega\} = P\{M_P(t) = m\} \tag{22}$$

As a consequence of Corollary 4.1, values of $\tilde{\phi}(M,k)$ for any $k \geqslant HP/T_{\text{msg}}$ are not conditioned by the initial values of watch variables, and can be used to derive an upper bound on Φ:

Lemma 4.10. *An upper-bound $\tilde{\Phi} \leqslant \Phi$ can be computed as:*

$$\tilde{\Phi} = \frac{\tilde{\phi}\left(M, \frac{2HP}{T_{\text{msg}}}\right) - \tilde{\phi}\left(M, \frac{HP}{T_{\text{msg}}}\right)}{1 - \tilde{\phi}\left(M, \frac{HP}{T_{\text{msg}}}\right)} \tag{23}$$

Furthermore, for each scenario $f = \langle \mathcal{H}_m, \mathcal{B}_n, \mathcal{L}_{M-m-n} \rangle$ where a sequence of M consecutive losses is caused by m failures due to cell handovers (with $m \leqslant M$), n failures due to burst noise (with $n \leqslant M-m$), and $M-m-n$ failures due to connection losses, the contribution of f to $\tilde{\Phi}$ can be easily evaluated as:

$$\tilde{\Phi}_f = \frac{\psi(m) \cdot p_{\text{burst}}(n) \cdot p_{\text{loss}}(M-m-n)}{\sum_{n=0}^{M-m} p_{\text{burst}}(n) \cdot p_{\text{loss}}(M-m-n)} \tag{24}$$

where $\psi(m)$ is the probability that m messages are lost due to handovers within a hyper-period of duration HP, normalized with respect to the probability of behaviors spanning over a hyper-period:

$$\psi(m) = \frac{\sum_{i \in \mathcal{M}_m^a} p_i\left(\left(\frac{2HP}{T_{\text{msg}}}\right)^{-}\right) - \sum_{i \in \mathcal{M}_m^a} p_i\left(\left(\frac{HP}{T_{\text{msg}}}\right)^{-}\right)}{1 - \tilde{\phi}\left(M, \frac{HP}{T_{\text{msg}}}\right)} \tag{25}$$

where in turn, $\mathcal{M}_m^a$ is the set of reachable markings of the model shown in Fig. 6 that satisfy the absorbing condition $\texttt{Failure} = 1 \wedge \sum_{i=0}^{M-1} \mathbb{W}_i = M - m$.

4.4.3. Evaluation beyond a hyper-period

Due to the periodic behavior of the model of Fig. 6, an upper-bound $\tilde{\beta}(M,H)$ on the probability that M consecutive end-to-end messages are lost within H hyper-periods can be derived as:

$$\tilde{\beta}(M,H) = \begin{cases} 0 & \text{if } H = 0 \\ \tilde{\beta}(M,H-1) + (1 - \tilde{\beta}(M,H-1)) \cdot \tilde{\Phi} & \text{if } H > 0 \end{cases} \tag{26}$$

Specifically, $\tilde{\beta}(M,H)$ is derived as the probability that M consecutive end-to-end messages are lost within $H-1$ hyper-periods (i.e., $\tilde{\beta}(M,H-1)$) plus the probability that M consecutive end-to-end messages are not lost within $H-1$ hyper-periods and are then lost during the H-th hyper-period (i.e., $(1 - \tilde{\beta}(M,H-1)) \cdot \tilde{\Phi}$). Solving the recurrence of Eq. (26) yields a geometric distribution with parameter $\tilde{\Phi}$:

$$\tilde{\beta}(M,H) = 1 - (1 - \tilde{\Phi})^H \tag{27}$$

where $\tilde{\Phi}$ is derived through regenerative transient analysis of the model Fig. 6 within the interval $[0, 2HP]$, as illustrated in Section 4.4.2. In so doing, Eq. (27) yields an upper-bound on the first-passage probability of M consecutive message losses over a time interval of arbitrary duration, which can be computed without the need to extend regenerative transient analysis of the model Fig. 6 beyond the time limit $2HP$.

Moreover, the result of Eq. (27) also permits to evaluate an upper-bound $\tilde{p}_{\text{stop}}$ on the long-run probability p_{stop} that a train is stopped due to an emergency brake:

$$p_{\text{stop}} \leqslant \tilde{p}_{\text{stop}} = \frac{\tau_{\text{recovery}}}{\tau_{\text{recovery}} + \tau_{\text{brake}}} \tag{28}$$

where $\tau_{\text{brake}} = HP/\tilde{\Phi}$ is the average time to the next emergency brake, computed as the mean value of the geometric distribution of Eq. (27), and τ_{recovery} is the average time needed to restart the train after an emergency brake, which is considered equal to 15 min in Zimmermann and Hommel (2003, 2005).

5. Experimental evaluation

A sensitivity analysis is performed with respect to: the train speed v; the headway delay T_{h} between trains; and, the offset O_{ft} of the foregoing train at the first cell border. Note that v and T_{h} comprise dependent parameters, possibly under the control of railway operators; however, for each value of v, there is a narrow range of reasonable values for T_{h}. According to this, to provide better insight on the system behavior, in the sensitivity analysis we take into account also the number M of consecutive communication failures triggering an emergency brake, regarding T_{h} as the headway delay that tolerates a given number of consecutive losses.

In the following, we present results of the evaluation during a hyper-period (Section 5.1) and during a time interval of arbitrary duration (Section 5.2), assuming train speed $v = 300$ km h^{-1} in both cases; then, we present results obtained for different values of train speed (Section 5.3) To this end, the approach combines analytic evaluation with numerical solution of STPN models presented in Section 3. Specifically:

- we derive $\tilde{\beta}(M,H)$ and $\tilde{p}_{\text{stop}}$ (i.e., upper-bound on the first-passage probability that M consecutive losses occur within H hyper-periods and on the long-run probability that a train is stopped due to an emergency brake, respectively) from $\tilde{\Phi}$ (i.e., upper bound on the probability that M consecutive end-to-end messages are lost within a hyper-period) according to Eqs. (27) and (28), respectively;
- in turn, $\tilde{\Phi}$ is computed according to Eq. (23) using $\tilde{\phi}(M, HP/T_{\text{msg}})$ and $\tilde{\phi}(M, 2HP/T_{\text{msg}})$ (i.e., upper bound on the probability that M consecutive losses occur within $[0, HPT_{\text{msg}})$ and within $[0, 2HPT_{\text{msg}})$, respectively), which are derived according to Eq. (19) from transient marking probabilities of the STPN of Fig. 6, computed by regenerative transient analysis of the model through the ORIS Tool;

- moreover, the weights of the IMM transitions `failure` and `noFailure` in the STPN of Fig. 6 are computed according to Eq. (18) from $\tilde{p}_{\text{burst}}(m)$ (i.e, upper bound on the probability that m not necessarily consecutive end-to-end messages are impaired by burst noise) and $\tilde{p}_{\text{loss}}(m)$ (i.e, upper bound on the probability that m not necessarily consecutive end-to-end messages are impaired by connection losses);
- finally, $\tilde{p}_{\text{burst}}(m)$ is derived according to Eq. (9) from stochastic parameters of the STPN of Fig. 1, while $\tilde{p}_{\text{loss}}(m)$ is derived according to Eq. (13) from steady-state and transient marking probabilities of the STPN of Fig. 2, which are computed by performing regenerative steady-state analysis and regenerative transient analysis through the ORIS Tool, respectively.

Experiments were performed on a single core of a 2.67 GHz Intel Xeon E5640 with 32 GB RAM.

5.1. Evaluation of communication failures during a hyper-period with train speed $v = 300$ km h^{-1}

Regenerative transient analysis of the model of message processing, transmission, and loss shown in Fig. 6 is performed using parameters introduced in Section 3.3, which are also recalled here for convenience: message generation period $T_{\text{msg}} = 6$ s; RBC processing time $T_{\text{RBC}} = 0.5$ s; time $T_{\text{BTS}} = 84$ s spent to cover the minimum distance between neighboring BTSs; time $T_{\text{rec}} = 0.3$ s needed to establish a connection with the next BTS after a handover; random transmission time τ_{link}, distributed over $[0.55.2.55]$ s according to the PDF of Eq. (2); and, random jitter τ_{j} uniformly distributed over $[0, 10]$ s, accounting for the possibly not regular displacement of BTSs. In addition, the analysis is performed for different values of:

- the number M of consecutive losses causing an emergency brake, typically comprised between 2 and 4;
- the time T_{h} needed to cover the headway distance s_{h}, which takes values according to Eq. (1);
- the offset O_{ft} representing the minimum time after which the foregoing train crosses the first nominal cell border, which takes values within the interval $[0, T_{\text{msg}})$ so as to account for any possible synchronization delay between the period T_{msg} of message generation and the period T_{BTS} of cell border arrival.

Specifically, for each value of O_{ft} within the set $\{0, 1, 2, 3, 4, 5\}$ s, the analysis is repeated for $M = 2$ and $T_{\text{h}} \in \{60, 62, 64\}$ s (corresponding to $s_{\text{h}} \in \{5, 5.17, 5.33\}$ km, respectively), $M = 3$ and $T_{\text{h}} \in \{66, 68, 70\}$ s (corresponding to $s_{\text{h}} \in \{5.5, 5.67, 5.83\}$ km, respectively), $M = 4$ and $T_{\text{h}} = 72$ s (corresponding to $s_{\text{h}} = 6$ km). Note that values of T_{h} larger than 72 s are not considered not to violate Hypothesis 2.

Table 1 shows the upper-bound $\tilde{\Phi}$ on the probability that M consecutive losses occur *within a hyper-period*, computed through Eq. (23). Results show that, regardless of the value of O_{ft}, $\tilde{\Phi}$ decreases by one order of magnitude as M increases by 1, which, according to Eq. (5), corresponds to an increase of 6 s of T_{h} and to an increase of 0.5 km of s_{h}; specifically, $\tilde{\Phi}$ is in the order of 10^{-1}, 10^{-2}, and 10^{-3} for M equal to 2, 3, and 4, respectively. Moreover, for assigned values of M and T_{h}, controlling O_{ft} would permit to reduce $\tilde{\Phi}$, though it would remain in the order of 10^{-1}, 10^{-2}, and 10^{-3} for M equal to 2, 3, and 4, respectively; in particular, a maximum reduction of nearly 10^{-2}, 10^{-3}, and 10^{-4} could be achieved for M equal to 2, 3, and 4, respectively.

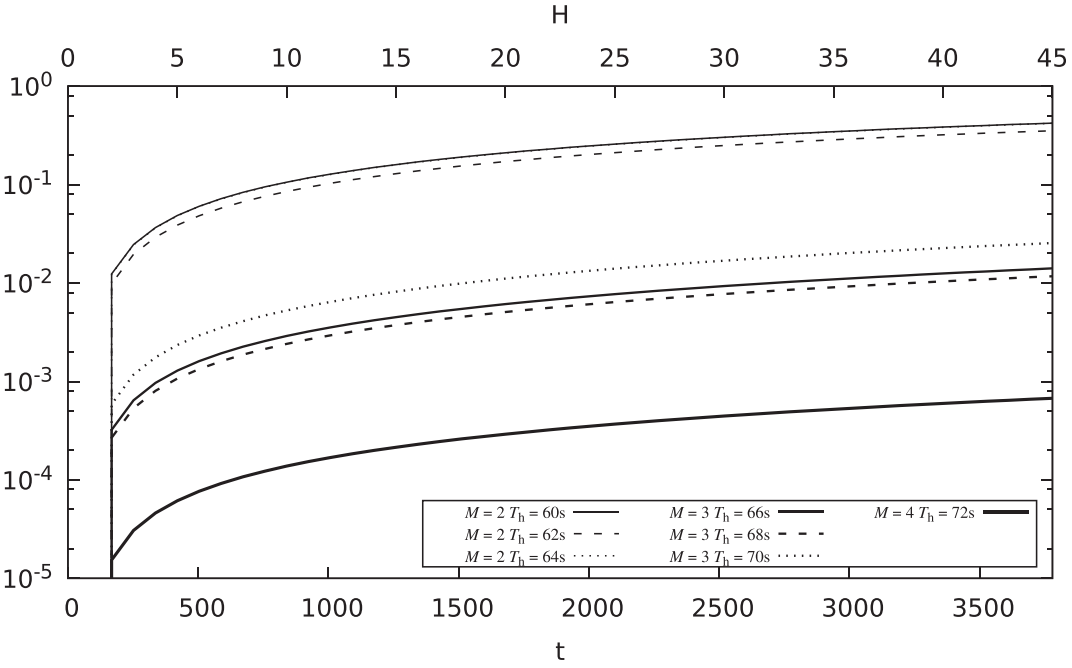
Furthermore, for any combination of parameters M, T_{h}, and O_{ft}, the evaluation of the weight of each failure scenario on $\tilde{\Phi}$ (according to Eq. (24)) proves that cell handovers have a major impact on the communication unavailability with respect to burst noise and connection losses, and, secondarily, that connection losses have a negligible impact with respect to burst noise. For instance, for the case with $M = 4$, $T_{\text{h}} = 72$ s, and $O_{\text{ft}} = 0$ s, the scenarios $\langle \mathcal{H}_2, \mathcal{B}_2, \mathcal{L}_0 \rangle$, $\langle \mathcal{H}_1, \mathcal{B}_3, \mathcal{L}_0 \rangle$, and $\langle \mathcal{H}_0, \mathcal{B}_4, \mathcal{L}_0 \rangle$ (i.e., 2 failures due to cell handovers and 2 failures due to burst noise, 1 failure due to cell handovers and 3 failures due to burst noise, and 4 failures due to burst noise, respectively), account for nearly 72.519%, 23.982%, and 3.478% of the cases with 4 consecutively lost end-to-end messages, respectively. The remaining 0.021% of the cases is ascribed to scenarios including at least 1 failure due to a connection loss.

Table 1
Evaluation within a hyper-period for the case with $v = 300$ km h^{-1}: the upper-bound $\tilde{\Phi}$ on the probability that M consecutive end-to-end messages are lost in a hyper-period of duration $HP = 84$ s, computed for different values of $M, T_{\text{h}}, O_{\text{ft}}$.

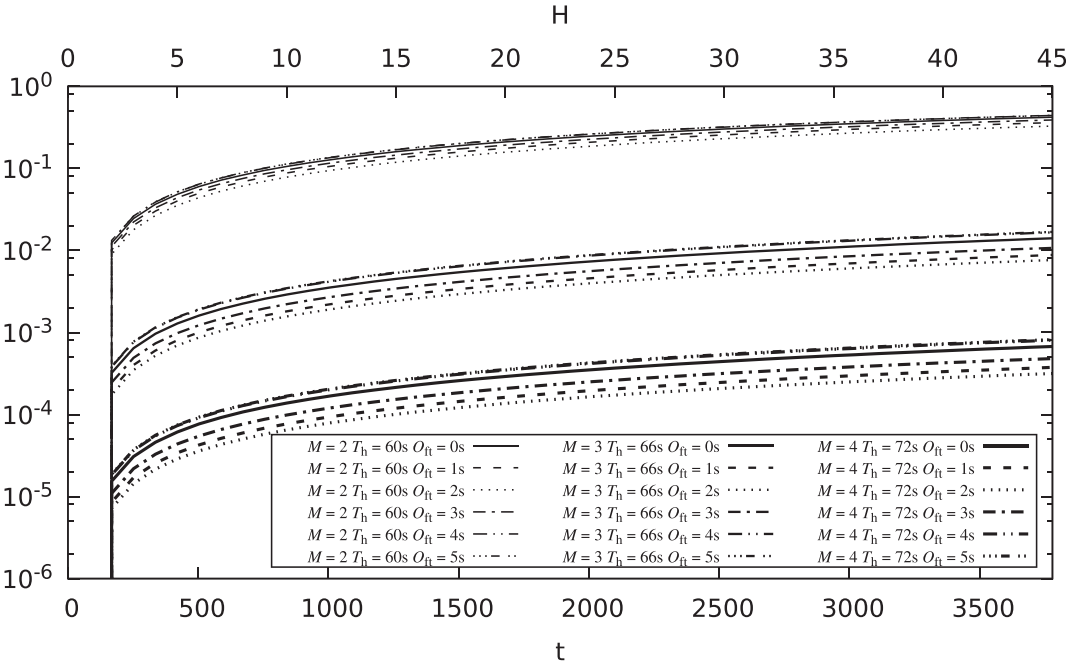
O_{ft}	$M = 2$ $T_{\text{h}} = 60$ s $s_{\text{h}} = 5$ km	$M = 2$ $T_{\text{h}} = 62$ s $s_{\text{h}} = 5.17$ km	$M = 2$ $T_{\text{h}} = 64$ s $s_{\text{h}} = 5.33$ km	$M = 3$ $T_{\text{h}} = 66$ s $s_{\text{h}} = 5.5$ km	$M = 3$ $T_{\text{h}} = 68$ s $s_{\text{h}} = 5.67$ km	$M = 3$ $T_{\text{h}} = 70$ s $s_{\text{h}} = 5.83$ km	$M = 4$ $T_{\text{h}} = 72$ s $s_{\text{h}} = 6$ km
0 s	0.123E−1	0.983E−2	0.123E−1	0.322E−3	0.268E−3	0.586E−3	0.153E−4
1 s	0.102E−1	0.864E−2	0.107E−1	0.201E−3	0.193E−3	0.335E−3	0.853E−5
2 s	0.899E−2	0.114E−1	0.115E−1	0.174E−3	0.343E−3	0.349E−3	0.719E−5
3 s	0.112E−1	0.132E−1	0.127E−1	0.246E−3	0.385E−3	0.374E−3	0.109E−4
4 s	0.132E−1	0.132E−1	0.107E−1	0.381E−3	0.387E−3	0.332E−3	0.183E−4
5 s	0.132E−1	0.127E−1	0.112E−1	0.386E−3	0.375E−3	0.416E−3	0.186E−4

5.2. Evaluation of communication failures beyond a hyper-period with train speed $v = 300$ km h^{-1}

Fig. 7a plots the bound $\tilde{\beta}(M,H)$ on the first-passage probability that M consecutive end-to-end messages have been lost within H hyper-periods for $O_{\text{ft}} = 0$ s and for increasing values of M and T_{h}; conversely, Fig. 7b shows $\tilde{\beta}(M,H)$ for increasing values of O_{ft} with $M = 2$ and $T_{\text{h}} = 60$ s, with $M = 3$ and $T_{\text{h}} = 66$ s, and with $M = 4$ and $T_{\text{h}} = 72$ s; in both cases, we consider a



(a) $\tilde{\beta}(M,H)$ computed for $O_{\text{ft}} = 0$ s.



(b) $\tilde{\beta}(M,H)$ computed for $M = 2$ and $T_{\text{h}} = 60$ s, $M = 3$ and $T_{\text{h}} = 66$ s, $M = 4$ and $T_{\text{h}} = 72$ s.

Fig. 7. Evaluation over multiple hyper-periods for the case with $v = 300$ km h^{-1}: the upper-bound $\tilde{\beta}(M,H)$ on the probability that M consecutive end-to-end messages are lost within H hyper-periods of duration $HP = 84$ s (time t expressed in s).

time limit equal to $H = 45$ hyper-periods, which corresponds to a time interval of nearly 1 h, given that the hyper-period duration is $HP = 84$ s and the train speed is $v = 300$ km h^{-1}.

As shown in Fig. 7a, given a value of O_{ft}, $\tilde{\beta}(M,H)$ significantly decreases by one order of magnitude as T_h increases by 6 s, which corresponds to an increase of M by 1. Specifically, at $H = 45$, $\tilde{\beta}(M,H)$ is in the order of $4.3 \cdot 10^{-1}$, $1.5 \cdot 10^{-2}$, and $6.9 \cdot 10^{-4}$ for $T_{\text{h}} = 60$ s ($M = 2$), $T_{\text{h}} = 66$ s ($M = 3$), and $T_{\text{h}} = 72$ s ($M = 4$), respectively; hence, a reduction of $\tilde{\beta}(M,H)$ by one order of magnitude is expected when T_{h} is further increased with a step of 6 s beyond the limit of 72 s. According to this, results of Fig. 7a could be used to select the delay T_{h} between a pair of chasing trains so as to achieve a trade-off between the utilization of the railway line and the probability that a train is stopped within 1 h; in turn, the delay T_{h} corresponds to a headway distance $s_{\text{h}} = T_{\text{h}}\, v$, which determines the maximum number M of consecutive tolerated losses according to Eq. (1).

Plots in Fig. 7b show that, for assigned values of T_{h} and M, O_{ft} substantially impacts on $\tilde{\beta}(M,H)$. Specifically, significant variations of $\tilde{\beta}(M,H)$ can be observed for $T_{\text{h}} = 66$ s and $T_{\text{h}} = 72$ s, though values remain within the same order of magnitude as O_{ft} varies. In so doing, if the synchronization delay between the integrity check on board a train and the arrivals at cell borders could be controlled, results of Fig. 7b could permit to select the value of O_{ft} that minimizes the probability that a train is stopped within 1 h.

Overall, results of Fig. 7 point out that the headway distance should be not lower than 6 km not to impair the railway throughput with too frequent emergency stops, which corresponds to a limit of $M = 4$ tolerated losses in consecutive end-to-end messages transmitted across the radio channel.

Similar considerations are made for the upper-bound $\tilde{p}_{\text{stop}}$ on the long-run probability that a train is stopped, which decreases by one order of magnitude from 0.11685 to 0.00344 up to 0.00016 as T_{h} increases from 60 s to 66 s up to 72 s, respectively (considering $O_{\text{ft}} = 0$ s in all cases), and exhibits variations in the same order of magnitude as O_{ft} varies while M and T_{h} remain fixed. Overall, $\tilde{p}_{\text{stop}}$ turns out to be nearly two orders of magnitude lower than the steady-state stopping probability computed in Zimmermann and Hommel (2003) for packet age equal to $2T_{\text{msg}} = 12$ s, anyway confirming that the headway distance must be significantly larger than the nominal braking distance of 3 km envisaged by the ERTMS/ETCS specification.

5.3. Evaluation of communication failures for different values of train speed

We consider three additional values of train speed in the sensitivity analysis, selected so as to satisfy Hypothesis 2 with a non-negligible jitter τ_{j}, and to guarantee that message releases and arrivals at cell borders have harmonic periods. Specifically, we consider 280 km h^{-1}, 233.33 km h^{-1}, and 200 km h^{-1}, which yield a period T_{BTS} of arrivals at cell borders equal to 90 s, 108 s, and 126 s, respectively. The braking distance s_{b} decreases with v, and it can be estimated from the case with $v = 300$ km h^{-1}, by considering a train length of 4.1 km and a position error of 0.02 km as in Zimmermann and Hommel (2003, 2005). Then, a lower bound on the headway distance s_{h} is obtained from v, s_{b}, and M according to Eq. (1), while an upper bound is derived so as to satisfy Hypothesis 2 while allowing a jitter $s_{\text{j}} \in [0, 0.83]$ km on the distance between consecutive BTSs. Finally, the headway delay and the time jitter are easily derived as $T_{\text{h}} = s_{\text{h}}/v$ and $\tau_{\text{j}} = s_{\text{j}}/v$, respectively. For instance, for the case with $v = 280$ km h^{-1}, we consider $s_{\text{h}} \in \{4.565, 4.744, 4.9\}$ km for $M = 2$ (corresponding to $T_{\text{h}} \in \{59, 61, 63\}$ s, respectively), $s_{\text{h}} \in \{5.031, 5.211, 5.367\}$ km for $M = 3$ (corresponding to $T_{\text{h}} \in \{65, 67, 69\}$ s, respectively), and $s_{\text{h}} = 5.497$ km for $M = 4$ (corresponding to $T_{\text{h}} = 71$ s).

Values of $\tilde{\Phi}$, $\tilde{\beta}(M,H)$, and $\tilde{p}_{\text{stop}}$ computed for the same values of M and O_{ft} but for different values of v turn out to be in the same order of magnitude, due to the fact that a reduction in speed corresponds to a reduction in the headway distance. Conversely, the impact of cell handovers on the communication availability decreases with the train speed, due to the fact that cell borders are encountered less often. Specifically, for the case with $v = 280$ km h^{-1}, $M = 4$, $T_{\text{h}} = 71$ s, and $O_{\text{ft}} = 0$ s, the scenarios $\langle \mathcal{H}_2, \mathcal{B}_2, \mathcal{L}_0 \rangle$, $\langle \mathcal{H}_1, \mathcal{B}_3, \mathcal{L}_0 \rangle$, and $\langle \mathcal{H}_0, \mathcal{B}_4, \mathcal{L}_0 \rangle$ (i.e., 2 failures due to cell handovers and 2 failures due to burst noise, 1 failure due to cell handovers and 3 failures due to burst noise, and 4 failures due to burst noise, respectively), account for nearly 51.064%, 42.123%, and 6.772% of the scenarios with $M = 4$ consecutively lost end-to-end messages, respectively (the remaining 0.041% of the cases is due to scenarios with at least a connection loss). When v decreases, it is never the case that 2 out of M consecutive end-to-end messages are lost due to handovers, so that $\langle \mathcal{H}_1, \mathcal{B}_3, \mathcal{L}_0 \rangle$ and $\langle \mathcal{H}_0, \mathcal{B}_4, \mathcal{L}_0 \rangle$ account for 86.847% and 13.089% of the cases if $v = 233.33$ km h^{-1}, respectively, and for 83.168% and 16.766% of the cases if $v = 200$ km h^{-1}, respectively.

Fig. 8 plots the upper-bound $\tilde{\beta}(M, 45)$ on the probability that M consecutive end-to-end messages are lost within 45 hyper-periods as a function of the headway distance s_{h}. As the speed decreases, the same bound on the stop probability can be guaranteed by a shorter headway distance; conversely, given a headway distance, the stop probability significantly decreases with the speed, at least by one order of magnitude. All curves alternate discontinuities and segments that are close to plateaus: the former account for an increase in the headway distance that yields an increase by 1 in the number M of consecutive tolerated losses, whereas the latter correspond to intervals of the headway distance where M remains constant. The curves for $v = 200$ km h^{-1} and $v = 233.33$ km h^{-1} exhibit two longer plateaus in the intervals $[3.5, 4.3]$ km and $[4.3, 5]$ km, respectively, so that the stop probability does not significantly decrease as M is increased by 1: in both cases, the effect

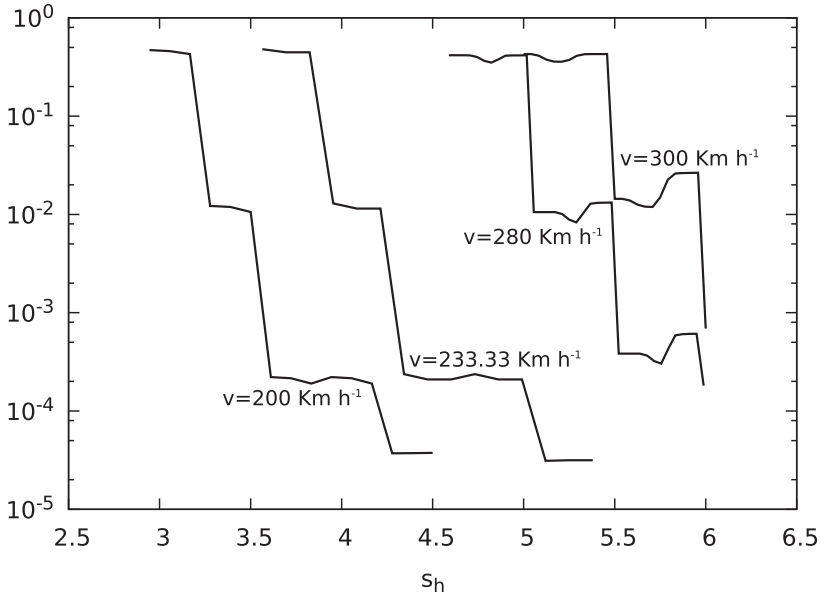


Fig. 8. The upper-bound $\bar{\beta}(M, 45)$ on the probability that M consecutive end-to-end messages are lost within 45 hyper-periods as a function of the headway distance s_h (expressed in km), computed for different values of train speed v, assuming $O_{ft} = 0$ s.

of an increase of M is balanced by the fact that scenarios where 2 out of M end-to-end messages are lost due to cell handovers become possible, which occurs for $s_h \geqslant 4$ km and $s_h \geqslant 4.7$ km in the cases with $v = 200$ km h^{-1} and $v = 233.33$ km h^{-1}, respectively.

Note that plateaus are not flat but rather exhibit significant variations of the stop probability, which increase in width with the train speed, though remaining in the same order of magnitude. This effect is due to the synchronization of the two trains in the crossing of cell borders: in fact, for a given train speed, variations in the headway distance affect the probability that both trains cross a cell border within a time interval of duration $M \cdot T_{msg}$. Overall, these results support the definition of high-level strategies aimed at controlling the headway distance so as to minimize the impact of handovers according to the displacement of BTSs, supporting the achievement of a trade-off between the line capacity/throughput and the stop probability.

6. Conclusions

The ERTMS/ETCS-L3 has not been deployed yet, and it may take years before having it fully operational. Therefore, modeling and analysis efforts are currently oriented to assess the specification requirements and to evaluate both operational and degraded scenarios, with the aim of providing insight on the range of possible design choices and supporting the development of an early deployment demonstrator. To this end, we evaluate emergency stops in the ERTMS/ETCS-L3 due to communication failures, using stochastic parameters derived in Zimmermann and Hommel (2003, 2005), with some changes reflecting amendments introduced in the evolution of the ERTMS/ETCS specification. The approach combines analytic evaluation of failures due to burst noise and connection losses with numerical solution of a non-Markovian model accounting also for failures due to cell handovers. The model represents periodic arrivals at cell borders by a pair of chasing trains and periodic message releases, leveraging regenerative transient analysis within two hyper-periods to derive the first-passage time distribution to an emergency stop over a time interval of arbitrary duration. This result also permits to compute the long-run probability that a train is stopped.

Experimental results prove that cell handovers mainly affect the availability of the GSM-R, though their impact decreases with the train speed, given that cell borders are crossed less often; conversely, the effects of connection losses on the GSM-R communication are negligible, also with respect to those of burst noise. The obtained results also suggest that the headway distance must be significantly larger than the braking distance to effectively limit the expected number of spurious emergency stops. More specifically, for any headway distance, a lower stop probability can be achieved by decreasing the train speed, supporting the definition of high-level policies for optimization of the line capacity/throughput. On the other hand, for each train speed, a convenient headway distance can be selected so as to minimize the impact of cell handovers, supporting the development of further strategies for the network control.

Based on the insight gained from the model construction, further investigation could be focused on the evaluation of a scenario where the RBC periodically transmits an MA to the chasing train using the most recent PR received from the foregoing train (rather than acting as a transponder that sends an MA to the chasing train in reaction to a PR received from the

foregoing train), which would reasonably reduce the stop probability. Moreover, the proposed solution could be easily adapted to any system that features a radio-based moving-block technology, notably metro-railways, where a high frequency of trains has to be guaranteed. The approach also provides the basis to extend the evaluation to the case of multiple trains running on the same railway line, where a spurious emergency brake causes a cascade of non-spurious stops. More generally, a challenge would be the application of the approach to vehicle platooning, where automated control of speed and distance aims at improving highway throughput while avoiding traffic congestion.

Acknowledgements

The authors would like to thank Francesco Flammini for fruitful discussions on the development of the ERTMS/ETCS standard.

Appendix A. Proofs

A.1. Burst noise

Lemma 4.1. *For any set of m not necessarily consecutive end-to-end messages, the probability that burst noise causes the loss of all the m end-to-end messages is:*

$$p_{\text{burst}}(m) = P_{\text{burst}}^{m} \tag{6}$$

where P_{burst} is the probability that burst noise causes the loss of a single end-to-end message.

Proof. Transient unavailability of the radio channel due to burst noise can be easily derived from marking probabilities of the model of Fig. 1 through forward transient analysis, i.e., $u_{\text{burst},\mathtt{q}}(t) = \sum_{n \in \mathcal{M}_{\text{burst}}} n(\mathtt{ko}) p_n(t)$, where $\mathtt{q} \in \{\mathtt{ok}, \mathtt{ko}\}$ is the initial marking and $\mathcal{M}_{\text{burst}}$ is the set of reachable markings. Both for $\mathtt{q} = \mathtt{ok}$ and $\mathtt{q} = \mathtt{ko}$, after a time equal to the shortest PR generation period T_{msg} (i.e., 6 s), $u_{\text{burst},\mathtt{q}}(t)$ is in the range $\pm 0.00003\%$ of the steady-state value. Moreover, burst noise affects the up-link and the down-link transmission in an independent manner, due to the headway distance between trains and the independence of transmission devices. Based on these results, subsequent not necessarily consecutive losses of end-to-end messages can be accurately approximated in product form as independent events. □

Lemma 4.2.

$$P_{\text{burst}} = 2P_{\text{link,burst}} - P_{\text{link,burst}}^{2} \tag{7}$$

where $P_{\text{link,burst}}$ is the probability that burst noise impairs either the up-link or the down-link transmission, and it can be upper-bounded as:

$$P_{\text{link,burst}} \leqslant \widetilde{P}_{\text{link,burst}} = \frac{\lambda_{\text{burst}}}{\lambda_{\text{burst}} + \mu_{\text{burst}}} + \frac{\mu_{\text{burst}}}{\lambda_{\text{burst}} + \mu_{\text{burst}}} \left(1 - e^{-\lambda_{\text{burst}} \max\{\tau_{\text{link}}\}}\right) \tag{8}$$

where τ_{link} is the random variable representing the duration of message transmission up-link or down-link, distributed according to the PDF defined in Eq. (2), so that $\max\{\tau_{\text{link}}\} = 2.55$ s is thus the maximum duration of the up-link or down-link transmission.

Proof. Due to the headway distance between trains and the independence of transmission devices, burst noise affects the up-link and the down-link transmission in an independent manner. Hence, the probability P_{burst} that burst noise impairs the up-link transmission of a PR, or the down-link transmission of an MA, or both, can be derived as $P_{\text{burst}} = P\{\text{burst up-link}\} + P\{\text{burst down-link}\} - P\{\text{burst up-link}\} \cdot P\{\text{burst down-link}\}$. Given that the transmission time is distributed according to the PDF of Eq. (2) both up-link and down-link, $P\{\text{burst up-link}\} = P\{\text{burst down-link}\} = P_{\text{burst,link}}$, which yields $P_{\text{burst}} = 2P_{\text{link,burst}} - P_{\text{link,burst}}^{2}$.

The process of burst noise can be considered in steady-state when it affects transmitted messages, given that it is in an unknown state at any previous time and it is independent of connection losses, cell handovers, and message transmission. According to this, $P_{\text{link,burst}} = U_{\text{burst}} + (1 - U_{\text{burst}}) \int_0^{\tau_{\text{link}}} \lambda_{\text{burst}} e^{-\lambda_{\text{burst}} t} dt$, where U_{burst} is the steady-state value of connection unavailability due to burst noise, so that the first addend is the probability that burst noise is occurring when transmission starts, while the second one accounts for the probability that the channel is available when transmission starts, but burst noise occurs before the transmission is completed. Given that τ_{link} is a random variable with a bounded support, $P_{\text{burst,link}}$ can be safely upper-bounded by replacing τ_{link} with its maximum value $\max\{\tau_{\text{link}}\}$, which yields

$P_{\text{burst,link}} \leqslant U_{\text{burst}} + (1 - U_{\text{burst}}) \int_0^{\max\{\tau_{\text{link}}\}} \lambda_{\text{burst}} e^{-\lambda_{\text{burst}} t} dt$. Therefore, Eq. (8) is obtained by substituting $U_{\text{burst}} = \lambda_{\text{burst}}/(\lambda_{\text{burst}} + \mu_{\text{burst}})$ and solving the definite integral. □

A.2. Connection losses

Lemma 4.3. *The probability $p_{\text{loss}}(m)$ that connection losses impair the transmission of m not necessarily consecutive end-to-end messages is upper-bounded by the probability of m consecutive losses, and can thus be estimated as:*

$$p_{\text{loss}}(m) \leqslant P_{\text{loss}} \cdot \sum_{\mathtt{q} \in Q} U_{\mathtt{q}} \cdot u_{\text{loss},\mathtt{q}}((m-2)T_{\text{msg}} + \delta) \tag{10}$$

where P_{loss} is the probability that connection losses impair the transmission of a single end-to-end message; $Q = \{lossIndication, establish, estFail\}$ is the set of the considered initial markings; U_q is the steady state probability of marking q; $u_{\text{loss},q}(t)$ is the transient unavailability of the radio channel due to connection losses conditioned to the hypothesis that the initial marking of the model of Fig. 2 is q; and, δ is the minimum time between two consecutive transmissions from the RBC.

Proof. Transient unavailability of the radio channel due to connection losses is derived from marking probabilities of the model of Fig. 2 through forward transient analysis, i.e., $u_{\text{loss},\mathtt{q}}(t) = 1 - \sum_{n \in \mathcal{M}_{\text{loss}}} n(\mathtt{connected}) p_n(t)$, where $\mathtt{q}$ is the initial marking and $\mathcal{M}_{\text{loss}}$ is the set of reachable markings. Fig. 9 plots $u_{\text{loss},\mathtt{q}}(t)$, showing that its settling time is significantly larger than the shortest PR generation period T_{msg} (i.e., 6 s) for each possible initial marking $\mathtt{q} \in \{\mathtt{connected}, \mathtt{lossIndication}, \mathtt{establish}, \mathtt{estFail}\}$. According to this, connection losses affecting the transmission of m out of M consecutive messages are not independent events for practically significant values of M.

While the foregoing train generates PRs with period T_{msg}, the starting time of MA transmissions in the down-link from the RBC is subject to the stochastic variability of transmission and processing delays. However, the minimum time between the first and the m-th MA transmission is not lower than $(m-2)T_{\text{msg}} + \delta$, where δ is the minimum time between two consecutive transmissions from the RBC. Since $u_{\text{loss},\mathtt{q}}(t)$ is monotonic decreasing for any initial marking $\mathtt{q} \in \{\mathtt{lossIndication}, \mathtt{establish}, \mathtt{estFail}\}$ that represents a failure state, $p_{\text{loss}}(m)$ is upper-bounded by the probability $r_{\text{loss}}(m)$ of m consecutive losses, with the first $m-1$ transmissions being equidistant with time-step T_{msg} and the last one occurring after δ, i.e., $p_{\text{loss}}(m) \leqslant r_{\text{loss}}(m)$.

In turn, $r_{\text{loss}}(m) = P_{\text{loss}} \cdot u_{\text{loss}}((m-2)T_{\text{msg}} + \delta)$, where the first factor is the probability that an initial connection loss has occurred, while $u_{\text{loss}}(t)$ is the transient unavailability of the radio channel (not conditioned on the initial state), so that the second factor accounts for the probability that the channel is still unavailable when the transmission of the m-th MA is started. The process of connection losses can be considered in steady-state when it affects transmitted messages, given that it is in an unknown state at any previous time and it is independent of burst noise, cell handovers, and message transmission. According to this, $u_{\text{loss}}(t)$ is derived by weighting $u_{\text{loss},\mathtt{q}}(t)$ with the steady-state probability of each marking $\mathtt{q}$ that represents a failure state, i.e., $u_{\text{loss}}(t) = \sum_{q \in Q} U_{\mathtt{q}} \cdot u_{\text{loss},q}(t)$, which finally yields Eq. (10). □

Lemma 4.4. *The probability that connection losses impair the transmission of an end-to-end message is:*

$$P_{\text{loss}} = 2P_{\text{link,loss}} - P_{\text{link,loss}}^2 \tag{11}$$

where $P_{\text{link,loss}}$ is the probability that connection losses impair either the up-link or the down-link transmission, and it can be upper-bounded as:

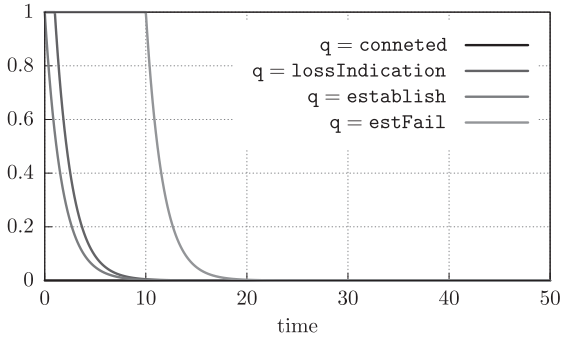


Fig. 9. Transient unavailability $u_{\text{loss},\mathtt{q}}(t)$ of the radio channel due to connection losses, for each possible initial marking $\mathtt{q}$ (time expressed in s).

$$P_{\text{link,loss}} \leqslant \widetilde{P}_{\text{link,loss}} = U_{\text{loss}} + (1 - U_{\text{loss}})\left(1 - e^{-\lambda_{\text{loss}} \max\{\tau_{\text{link}}\}}\right) \tag{12}$$

where U_{loss} is the steady-state value of communication unavailability due to connection losses, and τ_{link} is the random variable representing the duration of message transmission up-link or down-link, distributed according to the PDF defined in Eq. (2).

Proof. Also connection losses affect the up-link and the down-link transmission in an independent manner, due to the headway distance between trains and the independence of transmission devices. Hence, the probability P_{loss} that a connection loss affects the up-link transmission of a PR, or the down-link transmission of an MA, or both, can be computed as $P_{\text{loss}} = P\{\text{loss up-link}\} + P\{\text{lossdown} - \text{link}\} - P\{\text{loss up-link}\} \cdot P\{\text{lossdown} - \text{link}\}$. Given that the up-link and the down-link transmission time are distributed according to the PDF of Eq. (2), $P\{\text{loss up-link}\} = P\{\text{lossdown} - \text{link}\} = P_{\text{loss,link}}$, which yields $P_{\text{loss}} = 2P_{\text{link,loss}} - P^2_{\text{link,loss}}$.

As already pointed out, the process of connection losses can be considered in steady-state when it affects transmitted messages. Therefore, $P_{\text{loss,link}} = U_{\text{loss}} + (1 - U_{\text{loss}}) \int_0^{\tau_{\text{link}}} \lambda_{\text{loss}} e^{-\lambda_{\text{loss}} t} dt$, where U_{loss} is the steady-state value of connection unavailability due to connection losses, so that the first added is the probability that a connection loss is occurring when transmission starts, while the second one accounts for the probability that the channel is available when transmission starts, but a connection loss occurs before the transmission is completed. Given that the model of Fig. 2 reaches a regeneration after each firing, U_{loss} can be derived through regenerative steady-state analysis as the steady-state probability of any marking such that place `connected` is empty. According to the above remarks, $P_{\text{loss,link}}$ is upper-bounded by $U_{\text{loss}} + (1 - U_{\text{loss}}) \int_0^{\max\{\tau_{\text{link}}\}} \lambda_{\text{loss}} e^{-\lambda_{\text{loss}} t} dt$, which yields the inequality of Eq. (12). □

A.3. Cell handovers

Lemma 4.5. *If $M \leqslant HP/T_{\text{msg}}$, cell handovers affect at most 2 out of M consecutive end-to-end messages.*

Proof. Let Γ be the tree of state classes enumerated by nondeterministic transient analysis of the TPN underlying the model of Fig. 5 until time $2HP$. Let Ω be the set of paths in Γ that: (*i*) start from the initial class or from a class where the fired transition is `genMsg`; (*ii*) end with a state class reached through the firing of `genMsg`; and, (*iii*) include $M - 1$ intermediate firings of `genMsg`. If $M \leqslant HP/T_{\text{msg}}$, then any path in Ω does not include more than 2 firings of `failureUp` or `failureDown`, i.e., at any time, at most 2 out of the last M end-to-end messages can be been lost due to cell handovers. □

A.4. Combined effect of the different causes of communication failures

Lemma 4.6. *The loss of the h-th end-to-end message causes the first occurrence of E_M if and only if: (i) $\neg E_{M,i} \forall i \in [1, h-1]$ and (ii) all the $M - 1$ messages sent within $[(h - M)T_{\text{msg}}, (h-1)T_{\text{msg}})$ were lost.*

Proof. (If) Condition (ii) is sufficient to guarantee that event E_M occurs within $[(h-1)T_{\text{msg}}, hT_{\text{msg}})$, and condition (i) guarantees that this is the first one. (Only if) condition (ii) is necessary to guarantee that event E_M occurs within $[(h-1)T_{\text{msg}}, hT_{\text{msg}})$, and condition (i) is trivially necessary for the first occurrence of E_M be at $t \geqslant (h-1)T_{\text{msg}}$. □

Lemma 4.7. *The loss of the h-th end-to-end message can cause the first occurrence of E_M only if the $(h - M)$-th end-to-end message was correctly delivered.*

Proof. If the $(h - M)$-th end-to-end message was lost, either some end-to-end message is delivered within $[(h - M)T_{\text{msg}}, (h-1)T_{\text{msg}})$ or event E_M occurs within $(h-1)T_{\text{msg}}$; in both cases, according to Lemma 4.6, the loss of the h-th end-to-end message cannot cause the first occurrence of E_M. □

Lemma 4.8. *$\phi(M, k)$ is upper-bounded by the following expression $\forall k > 0$:*

$$\phi(M,k) \leqslant \sum_{h=M}^{k} \sum_{m=0}^{2} p_{\text{handover}}(M,h,m) \cdot \sum_{n=0}^{M-m} p_{\text{burst}}(n) \cdot p_{\text{loss}}(M-m-n) \tag{16}$$

where $p_{\text{handover}}(M,h,m) = P\{H^m_{M,h} | Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1, h-M]\}$; $H^m_{M,h}$ *is the event that* m *end-to-end messages are lost due to cell handovers within* $[(h-M)T_{\text{msg}}, hT_{\text{msg}})$; Z_{h-M} *is the event that the* $(h-M)$*-th end-to-end message is delivered; and,* $p_{\text{handover}}(0) = 1, p_{\text{burst}}(0) = 1$, *and* $p_{\text{loss}}(0) = 1$.

Proof. According to the results of Sections 4.1,4.2,4.3, event $E_{M,h}$ occurs if handovers cause $m \leqslant 2$ losses among the M end-to-end messages sent within $[(h-M+1)T_{\text{msg}}, hT_{\text{msg}})$, while burst noise or connection losses impair the remaining $M-m$. Hence, by the law of total probability, Eq. (15) can be rewritten as:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} \sum_{n=0}^{M-m} P\{H^m_{M,h} \wedge B^n_{M,h} \wedge L^{M-m-n}_{M,h} | \neg E_{M,i} \forall i \in [1,h)\} \tag{29}$$

where $B^n_{M,h}$ is the event that n end-to-end messages are lost due to burst noise within the time interval $[(h-M+1)T_{\text{msg}}, hT_{\text{msg}})$ and $L^{M-m-n}_{M,h}$ is the event that $M-m-n$ end-to-end messages are lost due to connection losses within the time interval $[(h-M+1)T_{\text{msg}}, hT_{\text{msg}})$. According to Lemmas 4.6 and 4.7, the set Θ of all behaviors that yield the first occurrence of E_M within $[(h-1)T_{\text{msg}}, hT_{\text{msg}})$ includes all and only the behaviors for which the $(h-M)$-th end-to-end message was delivered and no sequence of M consecutive end-to-end messages was lost within $[0, (h-M)T_{\text{msg}})$:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} \sum_{n=0}^{M-m} P\{H^m_{M,h} \wedge B^n_{M,h} \wedge L^{M-m-n}_{M,h} | Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1,h-M]\} \tag{30}$$

where Z_{h-M} is the event that the $(h-M)$-th end-to-end message was delivered. According to the definition of conditional probability, $\phi(M,k)$ can be expressed as:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} \sum_{n=0}^{M-m} \frac{P\{H^m_{M,h} \wedge B^n_{M,h} \wedge L^{M-m-n}_{M,h} \wedge Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1,h-M]\}}{P\{Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1,h-M]\}} \tag{31}$$

Given that failures due to burst noise, connection losses and cell handovers are independent of each other, and losses due to burst noise are approximated as independent events according to Lemma 4.1, we obtain:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} \sum_{n=0}^{M-m} \frac{P\{H^m_{M,h} \wedge L^{M-m-n}_{M,h} \wedge W_{M,h}\} \cdot P\{B^n_{M,h}\}}{P\{W_{M,h}\}} \tag{32}$$

where $W_{M,h} = Z_{h-M} \wedge \neg E_{M,i} \forall i \in [1,h-M]$. By the chain rule, $\phi(M,k)$ can be expressed as:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} \sum_{n=0}^{M-m} \frac{P\{W_{M,h}\} \cdot P\{H^m_{M,h} | W_{M,h}\} \cdot P\{L^{M-m-n}_{M,h} | W_{M,h} \wedge H^m_{M,h}\} \cdot P\{B^n_{M,h}\}}{P\{W_{M,h}\}} \tag{33}$$

Given that connection losses are independent of cell handover, it holds that:

$$\phi(M,k) = \sum_{h=M}^{k} \sum_{m=0}^{2} P\{H^m_{M,h} | W_{M,h}\} \cdot \sum_{n=0}^{M-m} P\{B^n_{M,h}\} \cdot P\{L^{M-m-n}_{M,h} | W_{M,h}\} \tag{34}$$

By Lemma 4.1, $P\{B^n_{M,h}\} = p_{\text{burst}}(m)$; moreover, by Lemma 4.3, $P\{L^{M-m-n}_{M,h} | W_{M,h}\} \leqslant p_{\text{loss}}(M-m-n)$. According to this, $\phi(M,k)$ can be finally upper-bounded as:

$$\phi(M,k) \leqslant \sum_{h=M}^{k} \sum_{m=0}^{2} p_{\text{handover}}(M,h,m) \cdot \sum_{n=0}^{M-m} p_{\text{burst}}(n) \cdot p_{\text{loss}}(M-m-n) \qquad \square \tag{35}$$

Lemma 4.9. $\forall t_0 \geqslant 0, \forall k \in \mathbb{N}$ *s.t.* $k < M, \forall m \in \mathcal{M}_{P \setminus \Omega^{M-1}_{k+1}}, \forall w \in \mathcal{M}_{\Omega^{M-1}_{M-k-1}}$:

$$P\{M_{P \setminus \Omega^{M-1}_{k+1}}(t) = m | M_{\Omega^{M-1}_{M-k-1}}(t_0) = \omega\} = P\{M_{P \setminus \Omega^{M-1}_{k+1}}(t) = m\} \tag{21}$$

where $t \geqslant (\lfloor t_0/T_{\text{msg}} \rfloor + 1)T_{\text{msg}} + kT_{\text{msg}}$; $\Omega^{M-1}_i = \{W_j\}^{M-1}_{j=i}$ *if* $i \in \mathbb{N}$ *s.t.* $i \leqslant M$; $\Omega^{M-1}_M = \varnothing$; P *is the set of places of the model shown in Fig. 6; and,* $\mathcal{M}_Q$ *is the set of reachable markings for places in* $Q \subseteq P$.

Proof. On the one hand, the behavior of the model of watch variables does not condition the behavior of the model of message processing and transmission and the model of communication failures due to handovers, as the marking of $\mathbb{W}_0, \ldots, \mathbb{W}_{M-1}$ does not condition the enabling of any transition in the other two models.

On the other hand, the behavior of the model of message processing and transmission and the model of communication failures due to handovers conditions the behavior of the model of watch variables through the update functions of

genMsg, handoverUp, and handoverDown. Specifically, at multiples of T_{msg}, the marking of $\mathbb{W}_{i-1}$ is assigned to $\mathbb{W}_i$ $\forall i = 1, \ldots, M-1$, and a token is assigned to $\mathbb{W}_0$. At the first multiple of T_{msg} following t_0, which is $(\lfloor t_0/T_{msg} \rfloor + 1)T_{msg}$, memory of the marking of $\mathbb{W}_{M-1}$ at time t_0 is lost, and the marking of $\mathbb{W}_0, \ldots, \mathbb{W}_{M-2}$ at time t_0 is encoded in $\mathbb{W}_1, \ldots, \mathbb{W}_{M-1}$, respectively; at time $(\lfloor t_0/T_{msg} \rfloor + 1)T_{msg} + T_{msg}$, memory of the marking of $\mathbb{W}_{M-2}$ at time t_0 is lost, and the marking of $\mathbb{W}_0, \ldots, \mathbb{W}_{M-3}$ at time t_0 is encoded in $\mathbb{W}_2, \ldots, \mathbb{W}_{M-1}$, respectively; at time $(\lfloor t_0/T_{msg} \rfloor + 1)T_{msg} + kT_{msg}$ with $k < M-1$, memory of the marking of $\mathbb{W}_{M-k-1}, \ldots, \mathbb{W}_{M-1}$ at time t_0 has been lost, and the marking of $\mathbb{W}_0, \ldots, \mathbb{W}_{M-k-2}$ at time t_0 is encoded in $\mathbb{W}_{k+1}, \ldots, \mathbb{W}_{M-1}$, respectively; and so on, until the marking of $\mathbb{W}_0$ at time t_0 is lost at time $(\lfloor t_0/T_{msg} \rfloor + 1)T_{msg} + (M-1)T_{msg}$.

According to this, at any time $t \geqslant (\lfloor t_0/T_{msg} \rfloor + 1)T_{msg} + kT_{msg}$, the transient probability of any marking $m \in \mathcal{M}_{P \setminus \Omega_{k+1}^{M-1}}$ is not conditioned by the marking of places $\mathbb{W}_{M-k-1}, \ldots, \mathbb{W}_{M-1}$ at time t_0, i.e., $P\{M_{P \setminus \Omega_{k+1}^{M-1}}(t) = m \,|\, M_{\Omega_{M-k-1}^{M-1}}(t_0) = \omega\} = P\{M_{P \setminus \Omega_{k+1}^{M-1}}(t) = m\}$. □

Corollary 4.1. $\forall t \geqslant HP, \forall m \in \mathcal{M}, \forall w \in \mathcal{M}_{\Omega_0^{M-1}}$:

$$P\{M_P(t) = m \,|\, M_{\Omega_0^{M-1}}(0) = \omega\} = P\{M_P(t) = m\} \tag{22}$$

Proof. By Lemma 4.9, at any time $t \geqslant (M-1)T_{msg}$, transient probabilities of reachable markings are not conditioned by the initial values of watch variables, i.e., $P\{M_P(t) = m \,|\, M_{\Omega_0^{M-1}}(0) = \omega\} = P\{M_P(t) = m\}$ $\forall t \geqslant (M-1)T_{msg}, \forall m \in \mathcal{M}, \forall w \in \mathcal{M}_{\Omega_0^{M-1}}$, which in fact descends from Eq. (21) for $k = M-1$ and $t_0 = 0$ s. In turn, $t \geqslant (M-1)T_{msg}$ is implied by $t \geqslant HP$, given that $t/T_{msg} \geqslant HP/T_{msg} = 84\text{ s}/6\text{ s} = 14$, and values of M with practical relevance range between 2 and 4. Therefore, Eq. (22) is satisfied $\forall t \geqslant HP$. □

Lemma 4.10. *An upper-bound $\tilde{\Phi} \leqslant \Phi$ can be computed as:*

$$\tilde{\Phi} = \frac{\tilde{\phi}\left(M, \frac{2HP}{T_{msg}}\right) - \tilde{\phi}\left(M, \frac{HP}{T_{msg}}\right)}{1 - \tilde{\phi}\left(M, \frac{HP}{T_{msg}}\right)} \tag{23}$$

Proof. By Lemma 4.8, $\phi(M, HP/T_{msg}) \leqslant \tilde{\phi}(M, HP/T_{msg})$; hence, $1 - \phi(M, HP/T_{msg}) \geqslant 1 - \tilde{\phi}(M, HP/T_{msg})$. Moreover, given that the behavior of the model of Fig. 6 is periodic over the hyper-period HP, $\phi(M, 2HP/T_{msg}) - \phi(M, HP/T_{msg}) \leqslant \tilde{\phi}(M, 2HP/T_{msg}) - \tilde{\phi}(M, HP/T_{msg})$. Therefore, $\Phi \leqslant \tilde{\Phi}$. □

References

AA.VV., 2005. GSM-R Interfaces: Class 1 Requirements.

Abbaneo, C., Flammini, F., Lazzaro, A., Marmo, P., Mazzocca, N., Sanseviero, A., 2006. UML based reverse engineering for the verification of railway control logics. In: International Conference on Dependability of Computer Systems. IEEE, pp. 3–10. http://dx.doi.org/10.1109/DEPCOS-RELCOMEX.2006.55.

Babczyński, T., Magott, J., 2014. Dependability and safety analysis of ETCS communication for ERTMS level 3 using performance statecharts and analytic estimation. In: International Conference on Dependability and Complex Systems, pp. 37–46. http://dx.doi.org/10.1007/978-3-319-07013-1_4.

Berthomieu, B., Diaz, M., 1991. Modeling and verification of time dependent systems using time Petri nets. IEEE Trans. Softw. Eng. 17 (3), 259–273. http://dx.doi.org/10.1109/32.75415.

Carnevali, L., Flammini, F., Paolieri, M., Vicario, E., 2015. Non-markovian performability evaluation of ERTMS/ETCS level 3. In: European Workshop on Computer Performance Engineering, pp. 47–62. http://dx.doi.org/10.1007/978-3-319-23267-6_4.

Carnevali, L., Ridi, L., Vicario, E., 2011. A framework for simulation and symbolic state space analysis of non-Markovian models. In: International Conference on Computer Safety, Reliability, and Security, pp. 409–422.

CENELEC, 2010. EN 50159:2010, Railways Applications. Communication, Signalling and Processing Systems. Safety-Related communication in transmission systems.

CENELEC, 2011. EN 50128:2011, Railway Applications. Communication, Signalling and Processing Systems. Software for railway control and protection systems.

Choi, H., Kulkarni, V.G., Trivedi, K.S., 1994. Markov regenerative stochastic Petri nets. Perform. Eval. 20 (1–3), 337–357. http://dx.doi.org/10.1016/0166-5316(94)90021-3.

Ciardo, G., German, R., Lindemann, C., 1994. A characterization of the stochastic process underlying a stochastic Petri net. IEEE Trans. Softw. Eng. 20 (7), 506–515. http://dx.doi.org/10.1109/32.297939.

Courtney, T., Gaonkar, S., Keefe, K., Rozier, E., Sanders, W.H., 2009. Möbius 2.3: an extensible tool for dependability, security, and performance evaluation of large and complex system models. In: IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 353–358. http://dx.doi.org/10.1109/DSN.2009.5270318.

Dill, D., 1989. Timing assumptions and verification of finite-state concurrent systems. In: Proceedings of the Workshop on Computer Aided Verification Methods for Finite State Systems. http://dx.doi.org/10.1007/3-540-52148-8_17.

ERA, 2014. Study on Migration of Railway Radio Communication System from GSM-R to Other Solutions.

ERA, 2016. ERTMS/ETCS - System Requirements Specification.

Esposito, R., Lazzaro, A., Marmo, P., Sanseviero, A., 2003. Formal verification of ERTMS Euroradio safety critical protocol. In: Symposium on Formal Methods for Railway Operation and Control Systems. doi:10.1.1.120.1262.

Flammini, F., Marrone, S., Iacono, M., Mazzocca, N., Vittorini, V., 2014. A multi-formalism modular approach to ERTMS/ETCS failure modeling. Int. J. Reliab., Qual. Safety Eng. 21 (1). http://dx.doi.org/10.1142/S0218539314500016.

Flammini, F., Marrone, S., Mazzocca, N., Vittorini, V., 2006. Modelling structural reliability aspects of ERTMS/ETCS by fault trees and bayesian networks. European Safety & Reliability Conference, vol. 6. http://dx.doi.org/10.1142/S0218539314500016.

Ghazel, M., 2014. Formalizing a subset of ERTMS/ETCS specifications for verification purposes. Transport. Res. Part C: Emerg. Technol. 42, 60–75. http://dx.doi.org/10.1016/j.trc.2014.02.002.
Haas, P.J., 2006. Stochastic Petri Nets: Modelling, Stability, Simulation.
Hassami, A.G., Foord, A.G., 2001. Systems safety-a real example (European rail traffic management system, ERTMS). In: International Conference on Human Interfaces in Control Rooms, pp. 327–334. http://dx.doi.org/10.1049/cp:20010484.
Hermanns, H., Jansen, D.N., Usenko, Y.S., 2005. From StoCharts to MoDeST: a comparative reliability analysis of train radio communications. In: International Workshop on Software and Performance. ACM, pp. 13–23. http://dx.doi.org/10.1145/1071021.107102.
Horváth, A., Paolieri, M., Ridi, L., Vicario, E., 2012. Transient analysis of non-Markovian models using stochastic state classes. Perf Eval 69 (7-8), 315–335.
Horvath, A., Puliafito, A., Scarpa, M., Telek, M., 2000. Analysis and evaluation of non-Markovian stochastic petri nets. In: Proceedings of the International Conference on Computer Performance Evaluation, pp. 171–187. http://dx.doi.org/10.1007/3-540-46429-8_13.
Lime, D., Roux, O.H., 2003. Expressiveness and analysis of scheduling extended time Petri nets. In: 5th IFAC Conference on Fieldbus and Their Applications (FET 2003). Elsevier Science, Portugal.
Lindemann, C., Thümmler, A., 1999. Transient analysis of deterministic and stochastic Petri nets with concurrent deterministic transitions. Perform. Eval. 36-37 (1–4), 35–54. http://dx.doi.org/10.1016/S0166-5316(99)00020-6.
Longo, F., Scarpa, M., 2009. Applying symbolic techniques to the representation of non-markovian models with continuous PH distributions. In: Proceedings of the European Performance Engineering Workshop on Computer Performance Engineering EPEW '09. Springer, Berlin, Heidelberg, pp. 44–58. http://dx.doi.org/10.1007/978-3-642-02924-0_4.
Martina, S., Paolieri, M., Papini, T., Vicario, E., 2016. Performance evaluation of Fischer's protocol through steady-state analysis of Markov regenerative processes. In: Modelling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS). IEEE, pp. 355–360. http://dx.doi.org/10.1109/MASCOTS.2016.7.
Neglia, G., Alouf, S., Dandoush, A., Simoens, S., Dersin, P., Tuholukova, A., Billion, J., Derouet, P., 2016. Performance evaluation of train moving-block control. In: Agha, G., Van Houdt, B. (Eds.), International Conference on Quantitative Evaluation of Systems, pp. 348–363. http://dx.doi.org/10.1007/978-3-319-43425-4_23.
ORIS Tool. Homepage <http://www.oris-tool.org>.
Qiu, S., Sallak, M., Schon, W., Cherfi-Boulanger, Z., 2014. Modeling of ERTMS level 2 as an SoS and evaluation of its dependability parameters using statecharts. IEEE Syst. J. 8 (4), 1169–1181. http://dx.doi.org/10.1109/JSYST.2013.2297751.
UIC, 1999a. ERTMS/ETCS Systems Requirements Specification.
UIC, 1999b. ERTMS/ETCS RAMS System Requirements Specification.
Vicario, E., 2001. Static analysis and dynamic steering of time dependent systems using time Petri nets. IEEE Trans. Softw. Eng. 27 (1), 728–748. http://dx.doi.org/10.1109/32.940727.
Vicario, E., Sassoli, L., Carnevali, L., 2009. Using stochastic state classes in quantitative evaluation of dense-time reactive systems. IEEE Trans. Softw. Eng. 35 (5), 703–719. http://dx.doi.org/10.1007/978-3-642-24270-0_30.
Zimmermann, A., Hommel, G., 2003. A train control system case study in model-based real time system design. In: International Parallel and Distributed Processing Symposium. IEEE, pp. 118–126. http://dx.doi.org/10.1109/IPDPS.2003.1213234.
Zimmermann, A., Hommel, G., 2005. Towards modeling and evaluation of ETCS real-time communication and operation. J. Syst. Soft. 77 (1), 47–54. http://dx.doi.org/10.1016/j.jss.2003.12.039.
Zimmermann, A., 2012. Modeling and evaluation of stochastic Petri nets with TimeNET 4.1. In: Int. ICST Conf. on Performance Evaluation Methodologies and Tools, pp. 54–63.