Epistemic parametric uncertainties in availability assessment of a Railway Signalling System using Monte Carlo simulation

S. Qiu, M. Sallak & W. Schön

Computer Engineering Department, Compiegne University of Technology, Heudiasyc Laboratory, UMR 7253, CNRS, Research Center of Royallieu, France

Z. Cherfi-Boulanger

Department of Mechanics, Compiegne University of Technology, Roberval Laboratory, UMR 7337, CNRS, Research Center of Royallieu, France

ABSTRACT: In this paper, firstly, we propose a modeling of a Railway Signalling System in Statechart and the system’s dynamic behavior is analyzed in this model. This system is the ERTMS/ETCS (European Rail Traffic Management System/European Train Control System) Level 2 whose performance is evaluated in terms of the availability. Secondly, the epistemic uncertainties (imprecision) are introduced into the transition rates of the system and handled by a methodology based on two-phase Monte Carlo simulation. In the two-phase Monte Carlo simulation, epistemic variables are sampled in the outer loop and the model of the system is executed in the inner loop. The originality of this work lies in modeling the dynamic behavior of the ERTMS/ETCS Level 2 and proposing a methodology based on two-phase Monte Carlo simulation to evaluate the availability of the system considering epistemic parametric uncertainties.

1 INTRODUCTION

Train is widely used as a form of public transport in the world. In railway systems design, safety is the most important factor that designers have to consider. In this paper, a railway signalling system, European Rail Traffic Management System (ERTMS), is studied. This railway signalling system is modeled and evaluated by one of its safety parameters: availability. Availability is usually used to evaluate a repairable system and is a function of reliability and maintainability. Reliability is a function of failure rate and maintainability is a function of repair rate. In our knowledge, there isn’t much work based on the ERTMS railway platforms. In Hermanns et al. (2005), StoCharts has been applied to model the European Train Control System (ETCS) which is a part of the ERTMS and evaluate the dependability of the train radio system. StoCharts is a predictable QoS (Quality of Service)-oriented extension of UML Statechart diagrams and it supports stochastic process modeling. StoCharts is useful in QoS modeling but lacks tool support, so it’s translated into MODEST which is a formal language to describe stochastic timed systems. Flammini et al. (2006) have modeled the failures of the ERTMS by Fault Trees and Bayesian Networks. They have instantiated models with realistic parameters and evaluated the system unavailability. In Vernez and Vuille (2009), an approach based on the functional Failure Mode, Effects and Criticality Analysis (FMECA) is used to optimize the dependability of the railway signalling system. This basic model FMECA considers the operational procedures, alarm systems, environmental and human factors, as well as operation in degraded mode and is implemented on a commercial software tool. This approach assesses the global risk level and availability level of the complex railway signalling system and identifies its vulnerabilities. In Lalouette et al. (2010), a new approach which uses Coloured Petri Nets to evaluate the dependability is applied to study the European railway signalling system superposed of the French system. This approach evaluates the dependability of a range of hazards that may be encountered during the operational life cycle of a system instead of arbitrarily chosen mission profiles. So far, there isn’t a complete model which describes the entire architecture of the ERTMS and at the same time analyzes its dynamic behavior. We are interested in the composition of the whole system, the communication among its components as well as its dynamic behavior, so we model this railway signalling system in UML Statechart. UML Statechart is a finite-state machine in Computer
Science and it provides researchers a dynamic view of system models.

None of the above researches has considered uncertainties in their models. Nilsen and Aven (2003) think the conception model uncertainty is commonly related to deviations between the real world and its representation in models. For these deviations, there may be two sources: the limitation of modeler’s knowledge and the deliberate simplification introduced by the modeler. Indeed, during the last years, the reliability and risk assessments community has recognized that there are different sources/types of uncertainties that play important roles in reliability and risk evaluation (see Winkler (1996) and Aven and Nøkland (2010)). In Aven (2011), uncertainties are usually divided into two types: aleatory uncertainty which is represented by the probability models and frequentist probabilities, and epistemic uncertainty which expresses the lack of knowledge about the true values of the frequentist probabilities and parameters of probability models. The distinction is important because epistemic uncertainties can be reduced by acquiring knowledge on the studied system, whereas aleatory uncertainties cannot. Furthermore, some works have proven that uncertainties in reliability and risk assessments are mainly epistemic (see Drouin et al. (2009)). Keep in mind that there are other points of view as to how to distinguish sources/types of uncertainty (see Dubois (2010) and Blockley (2012)). In this paper, only epistemic uncertainties are studied.

In our research, systems are modeled in UML Statechart which consists of states and transitions. In Statechart models, two kinds of epistemic uncertainties may exist: epistemic state uncertainty which exists in the states and epistemic parametric uncertainty which exists in the transition rates. State uncertainty means that there is imprecision about the states of systems. In other words, sometimes we can’t confirm the states of systems. This is caused by the lack of information about components of systems and represents the ignorance about the states of systems. In this paper, epistemic state uncertainty isn’t treated, and the focus is the epistemic parametric uncertainty. Parametric uncertainty means that there exists imprecision in some parameters, because these parameters’ values are fixed but their accurate values are unknown. The values of parameters of models usually come from the statistics, from systems which have the similar functionality or from the experts’ opinions. The values from these sources aren’t accurate. This makes parametric uncertainty analysis necessary in modeling systems. Many researchers have studied the parametric uncertainties while modeling systems. Liang et al. (2001) propose three ways to model the parametric uncertainty in the model: reliability bounds, confidence intervals and probability distributions. Then they derive the confidence interval of system reliability from the confidence intervals of parameters by the second-order approximation and the normal approximation. The proposed analytic method is validated by the Monte Carlo simulation method. Sen et al. (2006) study the problem of model checking Interval-valued Discrete-time Markov Chains (IDTMC) for which the accurate transition probabilities are unknown. Two interpretations for the uncertainty in the transition probabilities are considered: Uncertain Markov Chains whose transition probabilities lie within the interval range given in the IDTMC and Interval Markov Decision Process in which the uncertainty is considered as being resolved through non-determinism. When the initial and transition probabilities of a finite Markov chain aren’t well known, the finite Markov chain can be considered as a basic uncertainty model. The imprecise Markov chains and their limit behavior are studied in Cooman et al. (2009).

In this paper, a methodology based on two-phase Monte Carlo simulation is proposed to deal with epistemic parametric uncertainties which are characterized by probability distributions. The epistemic variables are sampled in the outer loop and the model of the system is executed in the inner loop.

This paper is organized as follows: Section 2 presents the UML Statechart and the methodology based on two-phase Monte Carlo simulation. Section 3 applies this methodology on the railway signalling system ERTMS/ETCS Level 2 considering epistemic parametric uncertainties. Section 4 contains conclusion and prospect of the future work.

2 EPISTEMIC PARAMETRIC UNCERTAINTY IN AVAILABILITY STUDIES

2.1 Statechart

UML (Unified Modeling Language) is a graphical modeling language in the field of object-oriented software engineering. It’s used in modeling and development of software systems. UML 2.2 has 14 kinds of diagrams and they are divided into two categories: structure diagrams and behavior diagrams. UML state machine is one type of behavior diagrams. It uses the states and state transitions to describe the behavior of systems. It specifies the sequences of states that systems go through because of the occurrences of events, and their corresponding actions.

In fact, as a graphical tool, the Statechart has been widely used in researches. Some research-
ers use Statechart to model the complex systems. A computer controlled Railway Interlocking system can specify precisely the rules which ensure the safety of routes for trains. As a formal method, Statechart is used to give precise specifications of the Railway Interlocking system in Banci et al. (2004). To develop tools and techniques which can check automatically the conformance of the railway equipment to the operation requirements, Statechart is used to model the ERTMS/ETCS specifications in Herranz et al. (2011). Statechart is also used in the reliability studies. Safety criteria analysis is an important step in system design. Two canonical intermediate representations of Statechart specification which are suitable for safety criteria analysis are presented in Pap et al. (2005).

Figure 1 is an illustration of a statechart. The system has three states. First of all, it enters into State1. State1 has entry, do and exit actions. When trigger1 is triggered and at the same time the guard is satisfied, the system passes from State1 to State2. This transition includes also an effect which is an action. The effect is executed when the transition fires. State2 has two regions and these two regions are executed in parallel. In other words, the State2.1 and State2.2 are both in active when the system enters into State2. When trigger2 is triggered, the system enters into State3. When trigger3 is triggered, the system returns to State1.

We’d like to model the dynamic behavior of systems. The dynamic behavior of systems is best described by finite-state machines. The advantages of Statechart lie in the facts that it is considered as the finite-state machine and it introduces new concepts such as the hierarchy of states and orthogonal regions that avoid the problem of combinatorial explosion which is encountered in finite-state automata. It also extends the actions that depend on systems’ states, and entry, do, exit activities. UML Statechart remains the benefits of traditional finite-state machines and enriches them. That’s why UML Statechart is chosen as our modeling language.

2.2 Methodology for epistemic parametric uncertainty analysis

This section proposes the methodology based on two-phase Monte Carlo simulation for epistemic parametric uncertainty analysis and applies it on a binary component considering epistemic uncertainty in the failure rate.

A binary component can be in either of two states (working or failed) at any given time. Given that the failure rate and the repair rate are constant, the component availability can be derived by the probabilistic approach. Figure 2(a) is the Markov chain of the binary component. “0” denotes the working state and “1” denotes the failed state. $\lambda$ is the failure rate and $\mu$ is the repair rate. Figure 2(b) is the corresponding statechart of this binary component in Stateflow. Stateflow is a toolbox of Matlab. It provides a design environment to model systems via Statechart.

The epistemic uncertainty is introduced into the $\lambda$ of the binary component. It means $\lambda$ is fixed but its accurate value is unknown. In our numerical example, $\lambda$ belongs to the interval $[\lambda_0, \lambda] = [0.015 \text{ h}^{-1}, 0.045 \text{ h}^{-1}]$ and $\mu = 0.07 \text{ h}^{-1}$. The epistemic uncertainty in $\lambda$ is characterized by a probability distribution.

In the two-phase nested Monte Carlo simulation shown in Figure 3, values of $\lambda$ are randomly selected from the probability distribution of the epistemic uncertainty in the outer loop. In the inner loop, for each selected $\lambda$, the model is executed many times and an average availability is calculated. Two kinds of probability distributions of the epistemic uncertainty are considered: uniform

![Figure 1. Illustration of a statechart.](image1)

![Figure 2. Models of a binary component.](image2)

![Figure 3. Two-phase nested Monte Carlo simulation.](image3)
distribution and normal distribution. For the normal distribution, two kinds of confidence intervals are discussed separately: 95% confidence interval and 99% confidence interval.

For the uniform distribution, the imprecision of $\lambda$ is uniformly distributed on $[0.015 \, h^{-1}, 0.045 \, h^{-1}]$.

For the normal distribution, the imprecision of $\lambda$ is normally distributed: $\lambda \sim N(\text{mean}, \sigma^2)$. If its 95% confidence interval is $[\text{mean} - 1.96 \times \sigma, \text{mean} + 1.96 \times \sigma] = [0.015 \, h^{-1}, 0.045 \, h^{-1}]$, we have mean $= 0.03 \, h^{-1}$ and the standard deviation $\sigma = 0.0077 \, h^{-1}$.

For the normal distribution, if $\lambda$’s 99% confidence interval is $[\text{mean} - 2.58 \times \sigma, \text{mean} + 2.58 \times \sigma] = [0.015 \, h^{-1}, 0.045 \, h^{-1}]$, we have mean $= 0.03 \, h^{-1}$ and the standard deviation $\sigma = 0.0058 \, h^{-1}$.

In the present problem, the outer loop contains 60 iterations and the inner loop simulation contains 5000 iterations with mission time of 500 hours. When the sample time is 1 hour, the entire simulation takes 12.16 hours on a computer DELL Precision M4600 (Processor: Intel(R) Core(TM) i7-2820QM CPU @ 2.30 GHz; RAM: 8 G; System type: 64-bit operating system). The upper bounds and the lower bounds of the binary component availability of the three different cases are drawn in Figure 4. We find each curve converges to a constant after 50 hours, and then the curve fluctuates around this constant. The thick solid lines show the interval of the component availability when the imprecision of $\lambda$ is uniformly distributed on $[0.015 \, h^{-1}, 0.045 \, h^{-1}]$. The dotted lines show the interval of the component availability when the imprecision of $\lambda$ is normally distributed and its 95% confidence interval is $[0.015 \, h^{-1}, 0.045 \, h^{-1}]$.

Figure 4. Upper bounds and lower bounds of the availability of a binary component considering epistemic parametric uncertainty.

### 3 APPLICATION

In the above section, a methodology based on two-phase Monte Carlo simulation is proposed to evaluate the availability interval of a binary component considering epistemic uncertainty in the failure rate. In this section, this methodology is applied on a railway signalling system: ERTMS/ETCS Level 2.

#### 3.1 ERTMS/ETCS level 2

The ERTMS is a platform supported by Europe to guarantee the interoperability across different countries and manufacturers by creating a single Europe-wide standard for train control and command systems. It is made up of two components: European Train Control System (ETCS), a standard for train control systems, and GSM-R, an international wireless communications standard for railway communication and applications.

The ETCS has three levels. These different levels are distinguished by the different Trackside and Onboard ETCS equipment and the different technologies of information transmission. The ETCS Level 1 superimposes on the existing signalling system. The information transmission from track to train borne system depends totally on the balises which are installed in the track. The driver operates the train according to the lineside signals. In ETCS Level 2, the information transmission is done by the radio. The authority and track description are displayed directly in the cab for the driver, so the lineside signals are no longer needed. The balises are used as positioning beacons to help the train to determine its position via sensors. In ETCS Level 3, the train integrity checking is done by the train itself, so the track circuits are no longer needed. The balises are used to update position information and transmit position and integrity data back to the interlocking via GSM-R.

Because the ERTMS/ETCS Level 3 is currently under development and the ERTMS/ETCS Level 2 is widely implemented in Europe, we take the ERTMS/ETCS Level 2 as our research object. Figure 5(a) describes the architecture of our model. This railway signalling system consists of three parts: Onboard system, Trackside system and GSM-R system. Figure 5(b) shows its hierarchical structure.

The Onboard system is equipped in the train and serves to control train movements. It receives the information comes from the Trackside system to create a “braking curve”. The train should respect this speed profile in order to slow down or brake before stop signals or emergencies. It also receives telegrams from balises and sends Position Reports which contain for example the train
position and operating mode to the Trackside system via GSM-R. In the Onboard system, the following modules are considered: RTM (Radio Transmission Module), BTM (Balise Transmission Module), TIU (Train Interface Unit), DMI (Driver Machine Interface) and EVC (European Vital Computer).

If the driver doesn’t give a correct operation in time, the Onboard system will call automatically the braking procedure and operate the train borne equipment by the TIU interface.

The Trackside system does train routing, collects track circuit occupation status, detects train position and sends the correct speed profiles to trains. The Trackside system contains the IXL (Interlocking) and the separation system. The separation system is made up of the RBCs (Radio Block Centers) and Eurobalises.

GSM (Global System for Mobile Communications) is a standard for mobile communications. GSM-R is an international wireless communications standard for railway communication and applications. The direction of communication is decided by the frequency of GSM-R messages. For the direction “Train to Track”, the frequency of GSM-R messages is between 876 MHz and 880 MHz. For the direction “Track to Train”, the frequency of GSM-R messages is between 921 MHz and 925 MHz.

Figure 6 represents the statechart of the entire ERTMS/ETCS Level 2. Figures 7, 8 and 9 show the statecharts of the three systems which make up the ERTMS/ETCS Level 2. These statecharts describe the communication between Onboard system and Trackside system via GSM-R in the presence of degradations and failures. As shown in the Figure 6, the railway signalling system consists of three systems: Onboard system, Trackside system and GSM-R system which work in parallel.

Above all, these three systems enter into the state "Waiting". If the variable "Start" is true, all these systems enter into the state “Normal”. In the state “Normal”, Onboard system and Trackside system communicate with each other via GSM-R. At the beginning, Onboard system is in the state “Calculation”, Trackside system is in the state “CollectInfoCalculation” and GSM-R system is in the state “CollectMessage”. When an event SignalFromTrack comes and at the same time the frequency of GSM-R messages is not less than 900 MHz, that’s to say the Trackside system sends information to the Onboard system. At this time, the Onboard system enters into the state “Receive”, the Trackside system enters into the state “Send” and the GSM-R system enters into the state “Track2Train”. When an event EndSendToTrain comes, Onboard system goes back to the state “Calculation”, Trackside system goes back to the state “CollectInfoCalculation” and GSM-R system goes back to the state “CollectMessage”. The operation for the information transmission from Onboard system to Trackside system is similar.

The Onboard system has a degraded state. When an event Operation comes, if the operator is available, the system enters into the state “OperationByOperator”, otherwise the system will enter into the state “OperationByComputer” which is a substate of the state “Degraded_OnBoard”. When the EndOperation comes, the system goes back to the state “Calculation” if the operator is unavailable, otherwise the Onboard system returns to the state “Normal”.

Each system has a state of failure. This state of failure consists of two kinds of failures. The first one is the “ErrorStateOfNet”. A variable “network_failed” is used to show the state of the whole network. It's modeled by the statistics and once it's true, systems will all enter into the state “ErrorState”. This failure will be repaired and
systems will return to the state “**CorrectState**” when an event **RepairNet** arrives. The second one is the “**OrderOfErrorOfNet**”. When the rail traffic controller discovers an abnormality in the network communication, he can give an order of error **ErrorTrain2Track** or **ErrorTrack2Train** immediately to interrupt the network and make all the systems enter into the state “**OrderOfErrorOfNet**”. This failure can be repaired by corresponding repair events like **RepairSend_OB**, **RepairReceive_
In the outer loop of the two-phase nested Monte Carlo simulation, values of transition rates are randomly selected from the probability distributions of the epistemic uncertainties. In the inner loop, for selected transition rates, the model is executed many times and an average availability is calculated. Two kinds of probability distributions of the epistemic uncertainties are considered: uniform distribution and normal distribution. For the normal distribution, two kinds of confidence intervals are discussed separately: 95% confidence interval and 99% confidence interval.

For the uniform distribution, the imprecision of each of these four transition rates is uniformly distributed on \([\text{transition rate}, \text{transition rate}]\).

For the normal distribution, the imprecision of each of these four transition rates is normally distributed: \(\text{transition rate} \sim N(\text{mean}, \sigma^2)\). If their 95% confidence intervals are \([\text{mean} - 1.96 \times \sigma, \text{mean} + 1.96 \times \sigma] = [\text{transition rate}, \text{transition rate}]\), the means and the standard deviations of these four transition rates are calculated as follows:

\[
\begin{align*}
\text{mean}_{\lambda_1} & = 9.3885 \times 10^{-8} \text{ h}^{-1} \\
\sigma_{\lambda_1} & = 2.395 \times 10^{-9} \text{ h}^{-1} \\
\text{mean}_{\mu_1} & = 0.6 \text{ h}^{-1} \\
\sigma_{\mu_1} & = 0.153 \text{ h}^{-1} \\
\text{mean}_{\lambda_2} & = 0.0001 \text{ h}^{-1} \\
\sigma_{\lambda_2} & = 2.551 \times 10^{-5} \text{ h}^{-1} \\
\text{mean}_{\mu_2} & = 0.6 \text{ h}^{-1} \\
\sigma_{\mu_2} & = 0.153 \text{ h}^{-1}
\end{align*}
\]
For the normal distribution, if transition rates’ 99% confidence intervals are \([\text{mean} -2.58\times\sigma, \text{mean} +2.58\times\sigma]=[\text{transition rate}, \text{transition rate}]\), the means and the standard deviations of these four transition rates are calculated as follows

\[
\begin{align*}
\text{mean}_{\lambda_1} &= -9.3885 \times 10^6 \text{ h}^{-1} \\
\sigma_{\lambda_1} &= 1.8195 \times 10^6 \text{ h}^{-1} \\
\text{mean}_{\mu_{1}} &= 0.6 \text{ h}^{-1} \\
\sigma_{\mu_{1}} &= 0.11628 \text{ h}^{-1} \\
\text{mean}_{\mu_{2}} &= 0.0001 \text{ h}^{-1} \\
\sigma_{\mu_{2}} &= 1.938 \times 10^{-5} \text{ h}^{-1} \\
\text{mean}_{\mu_{3}} &= 0.6 \text{ h}^{-1} \\
\sigma_{\mu_{3}} &= 0.11628 \text{ h}^{-1}
\end{align*}
\]

In the present problem, the outer loop contains 30 iterations and the inner loop simulation contains 1000 iterations with mission time of 3 years. When the sample time is 1 hour, the entire simulation takes 26.45 hours on a computer DELL Precision M4600 (Processor: Intel(R) Core(TM) i7-2820QM CPU @ 2.30 GHz 2.30 GHz; RAM: 8 G; System type: 64-bit operating system). The upper bounds and the lower bounds of the system availability of the three different cases are drawn in Figure 10. The thick solid lines show the interval of the system availability when the imprecision of these four transition rates is normally distributed and their given 99% confidence intervals are \([\text{transition rate}, \text{transition rate}]\). The dotted lines show the interval of the system availability when the imprecision of these four transition rates is uniformly distributed and their given 99% confidence intervals are \([\text{transition rate}, \text{transition rate}]\). The solid lines show the interval of the system availability when the imprecision of these four transition rates is normally distributed and their given 99% confidence intervals are \([\text{transition rate}, \text{transition rate}]\). Figure 10(a) shows the availability intervals in the short term. We find each curve converges quickly to a constant, and then the curve fluctuates around this constant. Figure 10(b) shows the availability intervals in the long term. Each curve fluctuates around a constant.

![Figure 10. Upper bounds and lower bounds of the availability of ERTMS/ETCS level 2 considering epistemic parametric uncertainties.](image)

3.3 Sensitivity analysis

The aim of sensitivity analysis is to estimate which kind of probability distribution of epistemic uncertainties has the most significant influence on the availability of the whole railway signalling system. This influence is measured by the distance between the minimum of the upper bound and the maximum of the lower bound. For the uniform distribution, the distance is 0.00081. For the normal distribution, if the given intervals are their 95% confidence intervals, the distance is 0.00144. If the given intervals are their 99% confidence intervals, the distance is 0.00080. Obviously, when the interval of the imprecision is given, the normal distribution of 95% confidence interval presents more imprecision than the uniform distribution and the normal distribution of 99% confidence interval. The uniform distribution presents almost the same imprecision with the normal distribution of 99% confidence interval.

4 CONCLUSION

A railway signalling system ERTMS/ETCS Level 2 is modeled in Statechart so that its dynamic behavior is analyzed in this model. When modeling systems in Statechart, two kinds of epistemic uncertainties may exist in the model and only the epistemic parametric uncertainties are introduced into our model. In this background, a methodology based on two-phase Monte Carlo simulation is proposed to evaluate the availability interval of the system considering epistemic parametric uncertainties. We did also a sensitivity analysis to measure which kind of the probability distribution of the uncertainties has the most significant influence on the system availability. In the future, we’d like to model this railway signalling system
in VBS (Valuation-Based Systems) and introduce belief functions theory to analyze epistemic uncertainties.

ACKNOWLEDGEMENT

This work was carried out and funded in the framework of the Labex MS2T. It was supported by the French Government, through the program “Investments for the future” managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02).

REFERENCES


