Hierarchical Models of Markov Chains: Optimizations with Limited Pessimism

Martin Kohlík, Hana Kubátová
Department of Digital Design
Faculty of Information Technology
Czech Technical University in Prague
Thákurova 9, 160 00 Prague 6, Czech Republic
{martin.kohlik, hana.kubatova}@fit.cvut.cz

***Abstract*—Dependability models allow calculating the rate of events leading to a hazard state – a situation, where safety of the modeled dependable system (e.g. railway station signalling and interlocking equipment, automotive systems, etc.) is violated, thus the system may cause material loss, serious injuries or casualties. A hierarchical dependability model based on multiple Markov chains allows expressing multiple redundancies made at multiple levels of a system consisting of multiple cooperating blocks. The hazard rates of the blocks are calculated independently and, when combined, they are used to calculate the hazard rate of the whole system. The independent calculations are significantly faster than the calculation of a single model composed of all models of the blocks. The paper shows a method of reducing Markov chains and using them to create hierarchical dependability models and its extensions allowing more accurate results to be achieved. An example study is used to demonstrate the improvements obtained by the extensions when compared to the original method.**

***Index Terms*—Fault tolerant systems, Hierarchical systems, Reliability, Reliability engineering.**

I. INTRODUCTION

Dependability of a system is the ability to avoid *service failures* (situations where the behavior of the system deviates from the correct behavior) that are more frequent and more severe than acceptable. Dependability is an integrating concept that includes Safety, Availability, Reliability, Integrity, and Maintainability [1].

An event causing violation of a system safety will be called a *hazard event*. The rate of hazard events is called *hazard rate*.

Dependability models are models designed to calculate the hazard rate of a system. Models of complex systems consisting of cooperating dependable blocks may be created as coarse-grained and fine-grained. Coarse-grained models are small and simple models allowing exact calculations of hazard rate in a short time. On the other hand, coarse-grained models are inaccurate and do not reflect the internal structure of the system. Fine-grained models are accurate, but they can be too large, and thus the hazard rate calculation is time-consuming. They reflect the internal structure, but they grow rapidly in size when the complexity of a system (e.g. the number of the dependable blocks) increases.

Inexact models may be used to speed up the calculations. Accuracy is not crucial in such cases when we prove that the inexact result is pessimistic. In other words, we must prove that the real system will be safer than the system modeled by the inexact model(s).

This paper presents extensions of the Markov chains-based dependability models reduction method presented in [2]. The reduction allows inexact hierarchical models to be built using multiple linked models reflecting the structure of a system. The extensions presented in this paper allow more accurate results to be achieved for the intended application.

Multi-level hierarchy may be used to describe each level of redundancy independently. The proposed hierarchical models allow to calculate the hazard rate and determine, whether the hazard event can be tolerated/omitted safely.

The reduction allows a trade-off between accuracy and reduction time. The hazard rate, which is calculated by the hierarchical model, is higher than the hazard rate calculated without hierarchy, but the CPU-time spent on its calculation is greatly reduced (ca. 1000 times w.r.t the exact model). The extensions allow more accurate results when specific conditions of the systems are met (e.g. the maximal allowed operational time).

The proposed extensions are demonstrated on a case study system containing multiple (17) identical dependable blocks configured as N-modular redundant system (NMR) in this paper. The hierarchical models use two linked models (a top NMR model and a model of the internal redundancy of the block) containing up to 16 states in total, instead of up to tens of thousands states which are necessarily used for the exact model.

The paper is organized as follows: Section II introduces basic reliability definitions. The reduction procedure is described in Section III and applied on the case study system in Section IV. The results are shown in Section V and Section VI concludes the paper.

II. THEORETICAL BACKGROUND

Reliability function $R(t)$ is a probability that the system will perform its intended operation under specified design limits from time 0 until time t at least [1], [3]. The failure distribution function $F(t)$ is a complementary function to the reliability function, $F(t) = 1-R(t)$.

Manuscript received April XX, 20XX; accepted April XX, 20XX.
This research was has been partially supported by the project SGS13/017/OHK3/1T/18.

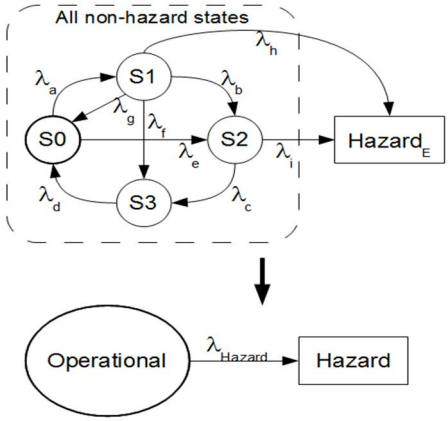

Fig. 1. Illustrative example of dependability model reduction

A failure rate is defined as the probability that a system will fail at a time t, assuming that it has survived until that time. Hazard rate (λ) is defined as a constant failure rate [3].

The presented hierarchical models used to calculate the hazard rate of the system are based on non-renewable Markov chains [4]. Markov chains (MCs) are able to model systems whose events are defined by continuous intensity rates. All events in the system modeled by MC must satisfy the Markovian property – the future of the system is based on its present state only.

A non-renewable MC contains hazard and non-hazard states. There are paths from each non-hazard state leading to a hazard state and there are no paths leading from a hazard state to a non-hazard state.

III. Dependability Models Reduction

The hazard rate cannot be calculated directly from a general MC due to non-constant failure rate in most cases. One way to achieve a constant failure rate is to create a simple MC containing two states and one hazard rate only.

The approximation of a general MC by a simple one (*reduction*) is made by merging all non-hazard states of a general MC into a single state that is called *Operational* in this paper (see Fig. 1). The merge is feasible, because there is no need to distinguish among the non-hazard states in the hazard rate calculation. The reduced model contains a new hazard rate λ_{Hazard} – the hazard rate substituting all hazard rates in the exact model.

The hazard rate of the reduced model is calculated as a pessimistic value meeting the condition called the **main requirement** in this paper. The main requirement is met when (1) is valid.

$$\forall t: F_R(t) \geq F_E(t) \tag{1}$$

$F_E(t)$ is the failure distribution function of the non-reduced (exact) model and $F_R(t)$ is the failure distribution function of the reduced model equal to the probability of the hazard state.

The reduction always leads to the same reduced model, thus the function $F_R(t)$ is calculated in (2).

$$F_R(t) = p_{Hazard} = 1 - e^{-\lambda_{Hazard} * t} \tag{2}$$

The drawback of the reduction is the loss of accuracy, because $F_E(t)$ can have any shape in general and $F_R(t)$ has always an exponential shape, thus they are not equal.

The main requirement is met only for a given time limit value t_{limit}. This way leads to more accurate solutions, but it can be used only when it is guaranteed that the modeled system will be replaced/repaired before t_{limit} is reached. If the t_{limit} is exceeded, the calculated hazard rate cannot be used.

The details about reduction algorithm have been presented in [2].

This paper is focused on the reduction using three different ways to set the limit time value t_{limit}:

1. Time-limited reduction – uses t_{limit} itself
2. Probability-limited reduction – uses a *limit* probability (see (3)).

$$F_E(t_{limit}) = F_R(t_{limit}) = limit \tag{3}$$

3. Hazard rate-limited reduction – uses a λ_{Limit} hazard rate (see (4)).

$$F_E(t_{limit}) = 1 - e^{-\lambda_{Limit} * t_{limit}} [= F_R(t_{limit})] \tag{4}$$

The main issue of using the reduction to calculate a hazard rate of a multi-level hierarchical model is using of $F_E(t)$ during the reduction. The calculation of $F_E(t)$ is time-consuming, thus it is necessary to avoid it to keep the main advantage of the reduction – the speedup of the calculation.

The calculation of $F_E(t)$ of the whole system can be avoided by using the reduction with the same t_{limit} during the reduction of all levels of the model. If the same t_{limit} is used, the $F_R(t)$ of the top-level model will meet the main requirement exactly up to t_{limit}, thus the most accurate valid result is produced.

The reduction of multi-level hierarchical model using limit probability or limit hazard rate requires balancing of limit values for individual levels of the model. The two-level hierarchical model reduction is accomplished by algorithm shown in Fig. 2. The multi-level hierarchy model will use the same principle. The iterative algorithm is used to find t_{limit} optimal for their limit value in both cases. The algorithms are similar, except for details enclosed in ** (specific for the probability limit) and // (specific for the hazard rate limit).

The end of the main loop is determined by *minStep* parameter. This parameter is used to balance the accuracy and the duration of the reduction.

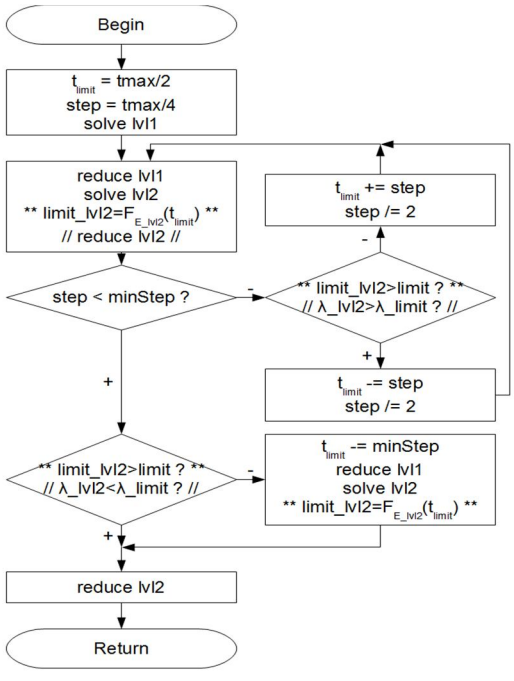

Fig. 2. The Reduction flowchart

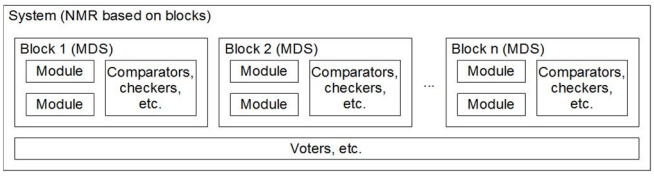

Fig. 3. Block diagram of a case study system

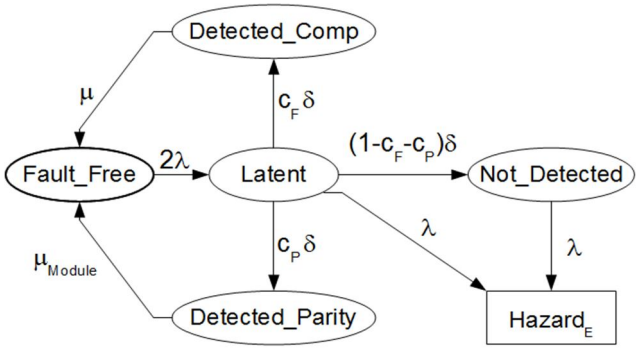

Fig. 4. Dependability model of the Modified duplex system block used to calculate the exact model failure distribution function.

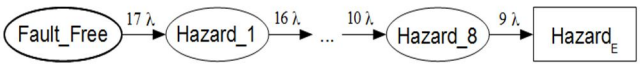

Fig. 5. Dependability model of N-modular redundant system used to calculate the exact model failure distribution function.

IV. CASE STUDY SYSTEM

A case study system uses dependable blocks connected in N-modular Redundant (NMR) system. The system is reduced using a two-level reduction.

The system is based on blocks using Modified duplex system redundancy [5]. Each dependable block contains two independent copies of functional modules, thus the safety of the blocks using these redundancies cannot be violated by a single fault. The structure of the system is shown in Fig. 3.

A. Modified Duplex System

Modified Duplex System (MDS) is based on two independent modules with parity checkers attached [5]. The parity checkers are able to detect some faults. The remaining faults are detected by comparators attached to the outputs of both modules.

The MDS is designed to utilize the reconfiguration ability of a field-programmable gate array (FPGA). FPGA is an integrated circuit designed to be configured by a customer or a designer after manufacturing. A part of the FPGA affected by a fault can be repaired by reconfiguration in tenths or hundreds of milliseconds.

The dependability model of MDS used in this paper is constructed using the following assumptions:

– Two faults will never occur at the same time.

– When a fault occurs in one module, the parity checker attached to this module may detect the fault. The parity checker needs not cover all possible faults. If the fault is detected by the parity checker, the affected module is repaired. If the fault is not detected by the parity checker, it may be detected by comparators. Both modules have to be repaired in such case, because the faulty module cannot be identified.

– If another fault occurs before the repair is completed, the safety of the block can be violated. This double-fault situation is considered as a hazard state.

The model shown in Fig. 4 is used to calculate the exact model failure distribution function $F_E(t)$ of the MDS block.

The model, its states and rates, is described in [2].

B. N-modular Redundancy

N-modular Redundancy (NMR) is based on N identical blocks and voter. This voter is able to compare all outputs of the blocks and use majority voting to produce a single output. If less than half of the blocks fails, the voter is able to produce correct output. If more than half of the blocks fails, the voter will produce incorrect output – this situation is considered as a hazard state. The erroneous blocks cannot be identified, thus there is no repair/recovery system.

The system containing 17 blocks is used in this paper. The model shown in Fig. 5 is used to calculate the exact model failure distribution function.

The details about reduction of both MDS block and NMR system are shown in [2].

V. RESULTS

We have presented results of the reduction (including the speedup of the reduction when compared to the exact solution and more NMR systems containing from 3 to 17 blocks) of the case study system in [2]. Additionally to [2], here we present the results of the time-limited and hazard rate-limited reductions in this paper. The probability-limited reduction presented in [2] used $t_{limit} = 10^{12}$ for low-level (MDS) model, thus the results have been more pessimistic than results presented in this paper.

The duration of the exact solution of the case study system is ca. 750 s [2], all results presented in this paper are acquired in ca. 1 s.

TABLE I. TIME-LIMITED REDUCTION.

Limit time [10^3 h]	Hazard rate [h^{-1}]
100	855.0×10^{-12}
150	57.34×10^{-9}
200	514.3×10^{-9}
250	1.530×10^{-6}
300	3.155×10^{-6}
350	5.127×10^{-6}
400	7.146×10^{-6}
450	8.990×10^{-6}
500	10.53×10^{-6}
550	12.49×10^{-6}
$\rightarrow \infty$ [1)]	23.85×10^{-6}

[1)] The reduction with limit 10^{12} is used.

Table I contains results of time-limited reduction. The first column contains the *limit time* – the maximal allowed operational time of the system. If this time is exceeded, the hazard rate (in the second column) will not be pessimistic when compared to the exact solution.

The plot in Fig. 6 shows failure distribution functions when time limit $t_{limit} = 200 \times 10^3$ h. The horizontal axis of the plot represents the time of operation measured in hours, the vertical axis represents the failure distribution function. The thick dashed line represents the exact model failure distribution function, the gray line represents the reduced model failure distribution function calculated using the full reduction and the black line represents the reduced model

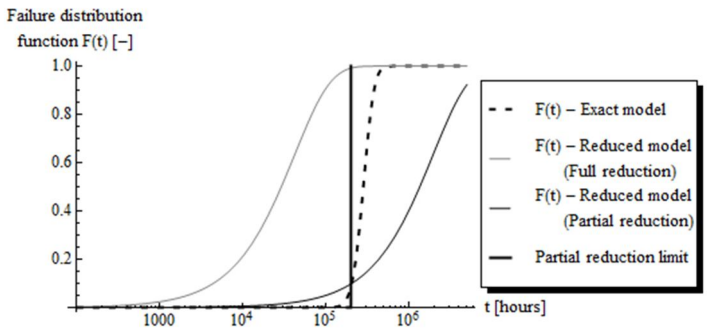

Fig. 6. Comparison of exact, full reduction, and time-limited reduction failure distribution functions

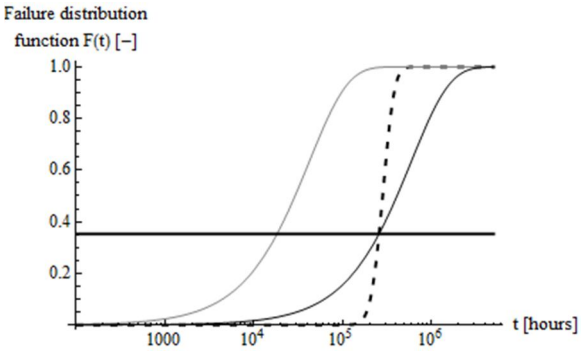

Fig. 7. Comparison of exact, full reduction, and probability-limited reduction failure distribution functions

failure distribution function calculated using the reduction. The vertical line represents the time limit t_{limit}.

TABLE II. PROBABILITY-LIMITED REDUCTION.

Limit probability [-]	Limit time [10^3 h]	Hazard rate [h^{-1}]
10^{-4}	102	1.113×10^{-9}
10^{-3}	123	8.160×10^{-9}
10^{-2}	155	69.73×10^{-9}
10^{-1}	209	590.2×10^{-9}
0.35	257	1.694×10^{-6}
0.6	302	3.155×10^{-6}
0.95	417	7.505×10^{-6}
0.99	479	9.754×10^{-6}
0.999	562	12.49×10^{-6}
1	$\rightarrow \infty$ [1)]	23.85×10^{-6}

[1)] The reduction limit 10^{12} is used.

Table II contains results of probability-limited reduction. The first column contains *limit probability* – the maximal allowed probability of failure of the system. This probability is reached at operational limit time shown in the second column. The hazard rate of the system, when the limit is applied, is shown in the third column.

The plot in Fig. 7 shows failure distribution functions when probability limit *limit* = 0.35. The axes represent the time of operation measured in hours and the failure distribution function (they are identical to the axes used in the previous plot). The horizontal line represents the probability limit.

Table III contains results of hazard rate-limited reduction. The first column contains *limit hazard rate* – the maximal allowed hazard of the system. The solution is pessimistic up to operational limit time shown in the second column. The hazard rate of the system, when the limit is applied, is shown in the third column.

TABLE III. HAZARD RATE-LIMITED REDUCTION.

Limit hazard rate [h^{-1}]		Limit time [10^3 h]
5×10^{-8}	50×10^{-9}	141
10^{-7}	100×10^{-9}	158
2×10^{-7}	200×10^{-9}	174
5×10^{-7}	500×10^{-9}	200
10^{-6}	1×10^{-6}	229
2×10^{-6}	2×10^{-6}	240
5×10^{-6}	5×10^{-6}	347
10^{-5}	10×10^{-6}	479
2×10^{-5}	20×10^{-6}	891

VI. CONCLUSIONS

The presented reduction of hierarchical dependability models based on Markov chains has been used to calculate the hazard rates of safety-critical systems in this paper. The model uses redundancy – the Modified duplex system method as low-level redundancy and the N-modular redundancy as high-level redundancy.

The results indicate that the reduction can improve the calculated hazard rate of the system significantly when compared to the full reduction.

The reduction decreases the hazard rate of the case study system ca. 40 times (from ca. 23×10^{-6} h^{-1} – one failure per ca. 5 years – to ca. 5×10^{-7} h^{-1} – one failure per ca. 200 years), but the system must be replaced/repaired before 200,000 hours of operation (ca. 22 years). If the replacement/repair before 150,000 hours (ca. 17 years) of operation is guaranteed, the hazard rate can be decreased to ca. 5×10^{-8} h^{-1} (the probability of a system failure, until 150000 hours of operation is reached, is ca. 1%).

The reduction can be bounded by the limit probability, too. This reduction type is suitable if the overall system failure probability cannot exceed a limit (the limit may be specified by some standard etc.)

The last but not least reduction method is specified by a limit hazard rate. Many safety-critical systems (e.g. railway station signalling and interlocking equipment, automotive systems, etc.) have maximal allowed hazard rate. Hazard rate limited reduction can find a maximal theoretical operational time of such systems thus the regular replacements may be performed less often.

REFERENCES

[1] A. Avižienis, J.-C. Laprie, B. Randell, C. Landwehr: "Basic Concepts and Taxonomy of Dependable and Secure Computing", *IEEE Transactions on Dependable and Secure Computing*, Vol. 1, No. 1, January–March 2004.

[2] M. Kohlík, H. Kubátová: "Markov chains hierarchical dependability models: Worst-case computations", in *proceedings of 14th Latin American Test Workshop*. Cordoba, Argentina, 2013. ISBN 978-1-4799-0595-9, p. 6.

[3] P. Hoang: *System Software Reliability, Chapter 2: System Reliability Concepts*, Springer Series in Reliability Engineering, Springer London (2007), pp. 9–75.

[4] A. L. Reibman, M. Veeraraghavan: Reliability modeling: an overview for system designers, *IEEE Transactions on Computer*, vol. 24, no. 4, (1991), pp. 49–57.

[5] P. Kubalík, H. Kubátová: Dependable design technique for system-on-chip, *Journal of Systems Architecture*, Vol. 2008, no. 54, (2008) pp. 452–464. ISSN 1383-7621.