Springer Series in Reliability Engineering

Anatoly Lisnianski
Ilia Frenkel
Alex Karagrigoriou *Editors*

Recent Advances in Multi-state Systems Reliability

Theory and Applications

Springer

Springer Series in Reliability Engineering

Series editor

Hoang Pham, Piscataway, USA

More information about this series at http://www.springer.com/series/6917

Anatoly Lisnianski · Ilia Frenkel
Alex Karagrigoriou
Editors

Recent Advances in Multi-state Systems Reliability

Theory and Applications

Springer

Editors
Anatoly Lisnianski
Planning, Development and Technology Division, The System Reliability Department
Israel Electric Corporation Ltd.
Haifa
Israel

Alex Karagrigoriou
Department of Mathematics
University of the Aegean
Karlovasi, Samos
Greece

Ilia Frenkel
Center for Reliability and Risk Management, Industrial Engineering and Management Department
SCE—Shamoon College of Engineering
Beer Sheva
Israel

ISSN 1614-7839 ISSN 2196-999X (electronic)
Springer Series in Reliability Engineering
ISBN 978-3-319-63422-7 ISBN 978-3-319-63423-4 (eBook)
DOI 10.1007/978-3-319-63423-4

Library of Congress Control Number: 2017946661

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Reliability theory is a multidisciplinary science aiming to develop complex technical and informational systems that are resistant to failures. Traditional reliability theory considered only two possible states for a system and its components—perfect functioning and complete failure. In real world, many components and systems can perform their tasks with various distinctive levels of efficiency usually referred to as performance rates. Such systems are called multi-state systems (MSS). In MSS reliability analysis, the great number of system states that need to be evaluated makes it difficult and often impossible to use traditional binary reliability techniques.

Since the mid-1970s numerous research studies have been published that focus on MSS reliability. Additional experience has also been gathered from industrial settings. Thus, recently MSS reliability has emerged as one of the main fields not only for scientists and researchers but also for engineers and industrial managers.

This book covers the recent developments in multi-state system reliability. It presents new theoretical issues that were not previously presented in the literature, as well as the solutions of important practical problems and case studies illustrating the application methodology.

The book is a collective work by a number of leading scientists, analysts, mathematicians, and engineers who have been working on the front end of reliability science and engineering. All chapters in the book are written by leading researchers and practitioners in their respective fields of expertise and present a plethora of innovative methods, approaches and solutions not covered before in the literature.

Despite the large number of contributing authors, this manuscript presents a continuous story of the modern multi-state system reliability theory and its applications in engineering. An important peculiarity of the book is the presentation some real-world problem solutions.

This book has been divided into two logically contiguous parts.

Part I Modern Mathematical Methods for Multi-state System Reliability Analysis

In this part new theoretical methods are presented that were not published till now.

Chapter "Reliability of a Network with Heterogeneous Components" investigates reliability of network-type systems under the assumption that the network has K types of of i.i.d. components. The method suggested for this purpose is an extension of the D-spectra method to K dimensions and it is based on Monte Carlo simulation.

Chapter "Reliability Analysis of Complex Multi-state System with Common Cause Failure Based on DS Evidence Theory and Bayesian Network" introduces a reliability analysis method for complex MSS with epistemic uncertainty based on Bayesian network and evidence theory.

Chapter "A D-MMAP to Model a Complex Multi-state System with Loss of Units" presents a multi-state model for complex system subject to various types of failures, to which preventive maintenance is applied. The model suggested is using a Markovian Arrival Process with Marked arrivals.

Chapter "Modeling and Inference for Multi-state Systems" focuses on multi-state systems modeled by means of a special type of semi-Markov processes. A special parametrization is proposed for the parameters describing the system, taking into account various types of dependencies of the parameters on the states of the system.

Chapter "Optimizing Availability and Performance of a Two-Unit Redundant Multi-State Deteriorating System" considers a redundant multi-state deteriorating system under maintenance. The system's evolution in time is described through a semi-Markov process and its availability, the expected downtime and the expected cost due to maintenance and unavailability are computed under all possible scenarios.

Chapter "Phase-Type Models and Their Extension to Competing Risks" presents an extension of the phase-type methodology for modeling of lifetime distributions to include the case of competing risks. This is done by considering finite state Markov chains in continuous time with more than one absorbing state, letting each absorbing state correspond to a particular risk.

Chapter "A Study on Repairable Series Systems with Markov Repairable Units" investigates repairable multi-state series systems by using matrix method which has been widely used in aggregated stochastic processes especially in Ion channel modeling and aggregated repairable systems. The formulas for reliability, instantaneous and interval availabilities are given in matrix form for four kinds of repairable series systems.

Chapter "Dynamic Performance of Series Parallel Multi-state Systems with Standby Subsystems or Repairable Binary Elements" presents a method for evaluating performance of multi-state systems with a general series parallel structure. The system components can be either repairable binary elements with given time-to-failure and repair time distributions, or 1-out-of-N warm standby configurations of heterogeneous binary elements characterized by different performances and time-to-failure distributions. Iterative algorithms are presented for determining performance stochastic processes of individual components.

Chapter "Optimal Imperfect Maintenance in a Multi-state System" considers a model with both imperfect preventive and imperfect corrective maintenance actions. A sequential failure limits preventive maintenance policy with infinite planning horizon for both maintenance actions is used to formulate a cost optimization problem.

Chapter "Reliability Evaluation of Non-repairable Multi-state Systems Considering Survival-death Markov Processes" proposes two models of modified "death" Markov processes considering components start-up failures. They are referred to as "survival-death" Markov processes and they differ in that the first model considers only a completely successful or failed start-up, whereas the second model considers also partially successful start-up. L_Z-transform technique is used for evaluating dynamic reliability of non-repairable MSS with start-up failures.

Chapter "Reliability Assessment of Systems with Dependent Degradation Processes Based on Piecewise-Deterministic Markov Process" presents a reliability assessment framework for multi-component systems whose degradation processes are modeled by multi-state and physics-based models. The piecewise-deterministic Markov process modeling approach is employed to treat dependencies between the degradation processes within one component or/and among components.

Chapter "Trade-Off Between Redundancy, Protections, and Imperfect False Targets in Defending Parallel Systems" considers systems that may be destroyed by unintentional impacts or intentional attacks. It studies trade-off between building redundant genuine elements, protections and deploying imperfect false elements in the defense of a capacitated parallel system.

Chapter "Optimal Testing Resources Allocation for Improving Reliability Assessment of Non-repairable Multi-state Systems" studies the testing resources allocation problem for MSSs in order to optimally distribute the limited reliability testing resources to improve the accuracy of reliability estimation/prediction.

Chapter "Topological Analysis of Multi-state Systems Based on Direct Partial Logic Derivatives" deals with the evaluation of influence of the system components on system operation by using Direct Partial Logic Derivatives (DPLDs). It develops a new method for DPLDs computation for multi-state systems that can be decomposed into disjoint modules.

Part II Applications and Case Studies

In this part several solutions are presented for real-world problems from different industrial areas.

Chapters "Short-Term Reliability Analysis of Power Plants with Several Combined Cycle Units" and "Reliability Analysis of a Modified IEEE 6BUS

RBTS Multi-state System" are devoted to power system analysis by using MSS models.

Chapter "Short-term Reliability Analysis of Power Plants with Several Combined Cycle Units" presents an application of *Lz*-transform method to a short-term reliability analysis of power plants consisting of several combined cycle generating units. Each generating unit is presented by 8-state Markov model. Such reliability indices as loss of load probability, availability, expected energy not supplied to consumers, etc. are calculated for the entire power plant.

Chapter "Reliability Analysis of a Modified IEEE 6BUS RBTS Multi-state System" is devoted to performance evaluation of a modified IEEE 6BUS RBTS. It was shown that the recommended modification contributes to the improvement of the system performance by increasing the reliability of operation within the required limits. The modification of the system aims to the reduction of failure rates maintaining the power output specifications.

Chapter "Lz-Transform Approach for Fault Tolerance Assessment of Various Traction Drives Topologies of Hybrid-Electric Helicopter" is devoted to the application of *Lz*-transform approach to fault tolerance assessment of different traction drives topologies of hybrid-electric helicopters.

Chapter "Patient Diagnostic State Evolution During Hospitalization: Developing a Model for Measuring Clinical Diagnostic Dynamics" considers a medical application of MSS reliability theory. It presents a model for measuring clinical diagnostic dynamics during patient hospitalization.

Finally, Chapter "Automated Development of the Markovian Chains to Assess the Availability and Performance of Multi-state Multiprocessor System" presents a method for automated development of the Markov chain for availability and performance assessment of multi-state multiprocessor systems.

We wish to thank all the authors for their insights and excellent contributions to this book. We would like to acknowledge the assistance of all involved in the reviewing process of the book, without whose support this could not have been successfully completed. We wish to thank all who participated in the reviewing process: Prof. Yi Ding, College of Electrical Engineering, Zhejiang University, China; Prof. Alex Karagrigoriou, Department of Mathematics, University of the Aegean, Samos, Greece; Prof. Ilya B. Gertsbakh, Department of Mathematics, Ben-Gurion University, Israel; Prof. Lirong Cui, School of Mathematics and Statistics, Beijing Institute of Technology, Beijing, China; Prof. Dmitry Efrosinin, Johannes Kepler University, Linz, Austria; Prof. Xujie Jia, School of Science, Minzu University of China, Beijing, China; Prof. Waltraud Kahle, Otto-von-Guericke University Madeburg, Germany; Prof. Liudong Xing, University of Massachusetts, Dartmouth, USA; Prof. Bo Henry Lindqvist, Norwegian University of Science and Technology, Trondheim, Norway; Prof. Svetlana Yanushkevich, Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, Canada; Prof. Kjell Hausken, Faculty of Social Sciences, University of Stavanger, Norway; Prof. Yu Liu, School of Mechatronics Engineering, University of Electronic Science and Technology, Chengdu, China; Prof. Gilberto Francisco Martha de Souza, Escuela

Politecnica, Universidad de Sao Paolo, Brazil; Prof. Dmitrii Silerstov, Department of Mathematics, Stockholm University, Sweden; Prof. Maxim Finkelstein, University of the Free State, Bloemfontein, Republic of South Africa; Prof. Suprasad Amari, Indian Institute of Technology, India; Ms.Yahli Chetrit, Industrial Engineering and Management Department, SCE-Shamoon College of Engineering, Israel.

We would like to express our sincere appreciation to Prof. Ilya Gertsbakh from Ben-Gurion University, Israel, for his great impact on book preparation.

It was indeed our pleasure working with the Springer Editorial Team.

Haifa, Israel — Anatoly Lisnianski
Beer Sheva, Israel — Ilia Frenkel
Karlovasi, Greece — Alex Karagrigoriou
May 2017

Contents

Contributors

Vlad Stefan Barbu Laboratoire de Mathématiques Raphaël Salem, Université de Rouen, Saint-Étienne-du-Rouvray, France

Emil Bashkansky Industrial Engineering and Management Department, ORT Braude College of Engineering, Karmiel, Israel

Igor Bolvashenkov Institute of Energy Conversion Technology, Technical University of Munich, Munich, Germany

Lirong Cui School of Management and Economics, Beijing Institute of Technology, Beijing, People's Republic of China

Yi Ding School of Electrical Engineering, Zhejiang University, Hangzhou, China

Stephanie Dietrich Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Ilia Frenkel Center for Reliability and Risk Management, Sami Shamoon College of Engineering, Beer Sheva, Israel

Ilya B. Gertsbakh Department of Mathematics, Ben-Gurion University, Beer-Sheva, Israel

Chuanxin Guo School of Electrical Engineering, Zhejiang University, Hangzhou, China

Hanoch Ben Haim The Israel Electric Corporation, Haifa, Israel

Hans-Georg Herzog Institute of Energy Conversion Technology, Technical University of Munich, Munich, Germany

Hong-Zhong Huang Center for System Reliability and Safety, University of Electronic Science and Technology of China, Chengdu, P. R. China

Tao Jiang School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, China

Waltraud Kahle Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Alex Karagrigoriou Department of Mathematics, University of the Aegean, Samos, Greece

Vyacheslav Kharchenko National Aerospace University “KhAI”, Kharkiv, Ukraine

Lev Khvatskin Center for Reliability and Risk Management, Sami Shamoon College of Engineering, Beer Sheva, Israel

Susanne Hodneland Kjølen Visma Software Labs, Oslo, Norway

Vasilis P. Koutras Department of Financial and Management Engineering, University of the Aegean, Chios, Greece

Miroslav Kvassay Department of Informatics, University of Zilina, Zilina, Slovakia

David Laredo The Israel Electric Corporation, Haifa, Israel

Gregory Levitin The Israel Electric Corporation, Haifa, Israel

Yan-Feng Li Center for System Reliability and Safety, University of Electronic Science and Technology of China, Chengdu, P. R. China

Yan-Fu Li Department of Industrial Engineering, Tsinghua University, Beijing, China

Peng Lin Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China

Yan-Hui Lin School of Reliability and Systems Engineering, Beihang University, Beijing, China

Bo Henry Lindqvist Norwegian University of Science and Technology, Trondheim, Norway

Anatoly Lisnianski The Israel Electric Corporation, Haifa, Israel

Yu Liu School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, China

Sonia Malefaki Department of Mechanical Engineering & Aeronautics, University of Patras, Rio, Patras, Greece

Thomas Markopoulos School of Science and Technology, Hellenic Open University, Patra, Greece

Yariv N. Marmor Industrial Engineering and Management Department, ORT Braude College of Engineering, Karmiel, Israel

Jinhua Mi Center for System Reliability and Safety, University of Electronic Science and Technology of China, Chengdu, P. R. China

Oleksandr Mulyak National University Lviv Polytechnic, Lviv, Ukraine

Rui Peng Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing, China

Weiwen Peng Center for System Reliability and Safety, University of Electronic Science and Technology of China, Chengdu, P. R. China

Agapios N. Platis School of Science and Technology, Hellenic Open University, Patra, Greece; Department of Financial and Management Engineering, University of the Aegean, Chios, Greece

Juan Eloy Ruiz-Castro Faculty of Science, Department of Statistics and Operational Research and IEMath-GR, University of Granada, Granada, Spain

Yoseph Shpungin Software Engineering Department, Sami Shamoon College of Engineering, Beer-Sheva, Israel

Yuanzhang Sun School of Electrical Engineering, Wuhan University, Wuhan, China

Radislav Vaisman School of Mathematics and Physics, The University of Queensland, Brisbane, Australia

Bogdan Volochiy National University Lviv Polytechnic, Lviv, Ukraine

Hui Xiao Department of Management Science, School of Statistics, Southwestern University of Finance and Economics, Chengdu, China

Liudong Xing University of Massachusetts Dartmouth, Dartmouth, MA, USA

He Yi School of Mathematics & Statistics, Beijing Institute of Technology, Beijing, People's Republic of China

Yan Yuan School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China

Elena Zaitseva Department of Informatics, University of Zilina, Zilina, Slovakia

Enrico Zio Chair on Systems Science and the Energetic Challenge, Fondation EDF, at Centrale Supelec, Paris, France; Politecnico di Milano, Milan, Italy

Part I
Modern Mathematical Methods for Multi-state System Reliability Analysis

Reliability of a Network with Heterogeneous Components

Ilya B. Gertsbakh, Yoseph Shpungin and Radislav Vaisman

Abstract We investigate reliability of network-type systems under the assumption that the network has $K > 1$ types of i.i.d. components. Our method is an extension the D-spectra method to K dimensions. It is based on Monte Carlo simulation for estimating the number of system failure sets having k_i components of i-th type, $i = 1, 2, \ldots, K$. We demonstrate our approach on a Barabasi-Albert network with 68 edges and 34 nodes and terminal connectivity as an operational criterion, for $K = 2$ types of nodes or edges as the components subject to failure.

Keywords Network terminal reliability · Several types of components · Two-dimensional spectrum · Monte Carlo simulation · Two-dimensional quantile

1 Introduction

Networks play a major role as critical infrastructures underpinning our societies and economies. Very often networks function in the presence of various disruptions from hacker attacks, natural disasters like earthquakes and natural degradations, as well as unforseen military and terrorist strikes [2, 5, 8, 11, 15, 17]. All these circumstances create growing interest to the problems of network robustness, reliability, and pre-disaster management [2, 4, 5, 15]. Reliability and resilience of network-type structures attracted major attention in the framework of general network theory, see e.g.

I.B. Gertsbakh (✉)
Department of Mathematics, Ben-Gurion University, Beer-Sheva, Israel
e-mail: elyager@bezeqint.net

Y. Shpungin
Software Engineering Department, Sami Shamoon College of Engineering, Beer-Sheva, Israel
e-mail: yosefs@sce.ac.il

R. Vaisman
School of Mathematics and Physics, The University of Queensland, Brisbane, Australia
e-mail: r.vaisman@uq.edu.au

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*, Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_1

[2, 5, 9, 13, 17]. Typically, the basic model of a network functioning in the presence of random "attacks" on its nodes or edges assumed random structure of the network itself, like Poisson or Barabasi-Albert [1, 3, 9] and focused on random and independent removal of nodes and edges. The components subject to failure were assumed to be identical and independent and the network failure criterion was network disintegration or disappearance of the so-called giant component [2, 9, 11]. The research in this direction was successfully advanced in [1] by using the results of percolation theory which provided the threshold value of network components to be removed to cause network failure. The limitation of this approach, however, is that it is not applicable to some other network failure criteria, like loss of terminal connectivity, decrease of the largest network component below some critical size (for finite networks), and network disintegration into critical number of isolated clusters [5, 6].

A very promising direction in the reliability study of network-type structures is the use of so-called signatures, first suggested by Samaniego [12–14]. The essential feature of this approach is that it is based on system structural invariant which depends only on system structure function and does not depend of probabilistic properties (like lifetime distribution) of system components. Despite its elegance and universality with respect to system failure criteria, it has been efficiently applied only to systems consisting of one type i.i.d. or exchangeable components.

The main purpose of the present work is to extend the signature (or so-called D-spectra) approach [5, 6] to network systems consisting of several groups of i.i.d. components. As a principal example to illustrate our approach and its abilities we consider a transportation (or supply) network of realistic size (34 nodes, 68 edges), having as the operational criterion the terminal connectivity. We consider the case when the nodes or the edges are subject to failure. In both cases, the components subjected to failure consist of two different groups of i.i.d. components.

The exposition in the paper is the following. Our approach is an extension of the D-spectra methodology to the case of heterogeneous network. Therefore, we start with a short overview of the D-spectra approach to the systems consisting of one-type components. In this case, the D-spectrum or signature allows to count the number $C(k)$ of failure sets having k failed components. With the knowledge of $C(k)$, system *DOWN* probability can be expressed automatically. Since the case of $K > 2$ groups of independent components is a rather straightforward generalization of the case of $K = 2$ groups of components, we devote the main part of Sect. 3 to the description of our approach to the $K = 2$ case.

When the system has two types of components, the key to the reliability analysis is estimation of the number $C(k, r)$ of so-called (k, r)-failure sets which have k and r failed components of the first and the second type, respectively. $C(k, r)$ are system structural invariants. Similar to the one-dimensional case, the estimation of $C(k, r)$ is made via the so-called *two-dimensional spectrum* which estimates the frequencies of the (k, r)-failure sets in a sample of simulated random permutations. We present an efficient Monte Carlo algorithm for estimating the two-dimensional spectrum.

In Sect. 3, we demonstrate how our approach works for a realistic example of a transportation/supply network with 34 nodes and 68 edges. The network was designed by using Barabasi-Albert preferential attraction method [1]. We consider the case of edge failures and two versions of node failures. We demonstrate how relocation of so-called strong nodes can change network reliability.

Section 4 is devoted to the analysis of the network failure state under a random attack on network nodes by a two-type shocks process. Our analysis allows to define two-dimensional quantile area for the random location of the "hitting point" of network failure. Finally, in the last Sect. 5 we present the formulas generalizing our approach for $K > 2$ types of components and some concluding remarks.

2 The Principal Model: Two Types of Components

2.1 Network Description

Our basic model is a network $N = (V, E, T)$ where V is a set of vertices (nodes), $|V| = n + k$, E is a set of edges (links), $|E| = m$, and T is a set of special nodes called *terminals*, $|T| = k,\ T \subset V$. Components subject to failures are either the links or the nonterminal nodes. Edge failure means that this edge is erased, nonterminal node failure means that all edges incident to this node are erased. In this paper we consider only one form of network *DOWN* state-so-called loss of terminal connectivity which means the network is *DOWN* if not all its terminal nodes are mutually connected.

In this section we consider the case when components subject to failure (nodes or edges) consist of two independent groups of i.i.d. components having lifetime CDF $H_1(t)$ and $H_2(t)$. So, if the edges fail, m_i edges have lifetime CDF $H_i(t)$, $i = 1, 2, \ldots$ and $m_1 + m_2 = m$, and nodes remain absolutely reliable. If the nodes fail, then n_i nodes have i.i.d. lifetimes $H_i(t)$, $i = 1, 2, \ldots$ and $n_1 + n_2 = n$, and edges remain absolutely reliable.

To simplify the exposition, we consider in detail the case of two groups of components in the network. Extension to $K > 2$ groups is straightforward and is left for Sect. 5.

2.2 One Type of Components

Since the case of two-type of component network is almost a straightforward generalization of our method of dealing with the standard one-type case, we remind shortly the basic definitions and principal steps for the "standard" situation where all components have i.i.d. lifetimes with CDF $H(t)$.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be the network component state vector. $x_i = 1/0$ if the i-th component is *up*/*down* respectively.

Network state is determined via a binary function $\varphi(\mathbf{x})$ which is 1 or 0 if the network is *UP* or *DOWN*, respectively. If $\varphi(\mathbf{x}^*) = 0$, $\mathbf{x}^*$ is called a failure vector. If we ignore the order of *up /down* components in this vector, then $\mathbf{x}^*$ determines a *failure set*, i.e. a set of j *down* components and $n - j$ *up* components. For simplicity, we call $\mathbf{x}^*$ *failure set*.

Now define D-spectrum or signature for our network. Let us consider a random permutation of component numbers

$$\pi = (i_1, i_2, \dots, i_n).$$

Suppose that all components are *up* and, moving from left to right, we turn them *down*. The network state is controlled on each step of this destruction process.

Definition 1.1 The ordinal number in the permutation π of the component whose turning down causes network state change from *UP* to *DOWN* is called the *anchor* of this permutation.

Assume that the permutations π are taken randomly and independently from the set of all $n!$ permutations. Then the anchor becomes a discrete random variable with support $\{1, 2, \dots, n\}$.

Definition 1.2 The distribution $\mathbf{f} = (f_1, f_2, \dots, f_n)$ of the anchor is called *D-spectrum* or *signature*, (where "D" stands for *destruction* process of anchor discovery).

Remark 1.1 Historically, the signature was first introduced by Samaniego [11] in a form equivalent to Definition 1.2. Independently, it was described 6 years later in [3] under the term *Internal Distribution*. The authors of [4–6] used the term D-spectra.

Definition 1.3 Denote by Y the discrete random variable with density $\mathbf{f}$. Its *cumulative* distribution function

$$F_0(k) = \sum_{i=1}^{k} f_i$$

is called *cumulative D-spectrum* or cumulative signature.

For networks having more than $n = 7$–8 components the calculation of D-spectra is made by means of an efficient Monte Carlo algorithm, see for example [5, 6]. This algorithm generates a sample of M permutations and estimates the frequency $\widehat{f(k)}$ of anchor appearance on the k-th position.

Denote by $C(k),\ k = 1, \dots, n$, the number of failure sets which have k components *down* and $(n - k)$ remaining components *up*. $C(k)$ is a *combinatorial invariant* of the system. Knowing $C(k)$ and the *up/down* probabilities p and $q = 1 - p$ of network components, we are able to compute system *DOWN* probability as

$$P(DOWN) = \sum_{k=1}^{n} C(k) q^k p^{(n-k)}. \tag{1}$$

Let $H(t)$ be CDF of component lifetime τ: $P(\tau \leq t) = H(t)$. Denote by p the probability that the component is *up* at time t_0. Then

$$p = 1 - H(t_0), q = H(t_0).$$

Therefore, (1) gives the probability that the network is *DOWN* at time t_0. Thus, the probability that system lifetime τ_{sys} does not exceed t_0 is

$$P(\tau_{sys} \leq t_0) = \sum_{k=1}^{n} C(k)[H(t_0)]^k[1 - H(t_0)]^{(n-k)}. \quad (2)$$

The crucial fact in obtaining Eqs. (1) or (2) is the following formula connecting $C(k)$ and $F_0(k)$:

$$C(k) = F_0(k)\frac{n!}{k!(n-k)!}.$$

It can be proved analytically using the formulas of order statistics for random variables with CDF $H(t)$, see [4], or using combinatorial arguments, see for example [5].

2.3 Two Types of Components

Now we turn to the network which has components of two types, namely there are n_1 components of type 1 and n_2 components of type 2. For sake of brevity, we call them x-type and y-type components, respectively, $n_1 + n_2 = n$. These x and y-type components have i.i.d. lifetimes, with CDFs $H_1(t)$ and $H_2(t)$, respectively.

The key to the principal formula (1) is the knowledge of $C(k)$, the number of failure sets with k components *down*. Now, when we have two types of components, we need to know the values of $C(k, r)$, the numbers of failure sets which have k *down* components of x-type and r *down* components of y-type, (the remaining $(n_1 - k)$ and $(n_2 - r)$ components are *up*). Then, the *DOWN* probability for network with two types of components equals

$$P^{\star}(DOWN) = \sum_{0 \leq k \leq n_1} \sum_{0 \leq r \leq n_2} C(k, r) q_1^k p_1^{(n_1-k)} q_2^r p_2^{(n_2-r)},$$

where q_1, q_2 and $p_1 = 1 - q_1, p_2 = 1 - q_2$ are *down* and *up* probabilities for x-type and y-type components, respectively.

Similar to the one-type component systems, $C(k, r)$ are *invariants* depending on system structure function and not depending on component lifetime distributions [5, 14].

2.4 Counting (k, r)-failure Vectors

In case of two types of components, we have to modify the notation for system state vector $x^\star$. Now it will be an ordered sequence of n_1 pairs (x_i, I) for components of x-type and n_2 pairs (y_j, I) for y-type components, where $x_1, \dots, x_{n_1}$ are the names (numbers) of x-components and $y_1, \dots, y_{n_2}$ are the names of y-components. Indicator I will be 1 or 0, if the corresponding component is *up* or *down*, respectively.

Example 1.1 Consider the network shown on Fig. 1. Components subject to failure are the edges, nodes are reliable. The network fails if there no connection between terminals S and T. The network has $n_1 = 3$ components of x-type $x_1 = (b, c)$, $x_2 = (b, T)$, $x_3 = (c, T)$, and two components of y-type—$y_1 = (S, a), y_2 = (a, b)$. Consider, for example, a vector $\mathbf{x}^* = (x_1, 1), (x_2, 0), (x_3, 1), (y_1, 0), (y_2, 0)$. Obviously $\phi(\mathbf{x}^*) = 0$. This failure vector contains one x-edge down and two y-edges down.

Sequential destruction of a random permutation. Consider a random permutation $\pi^\star$ of n_1 x-type pairs mixed randomly with n_2 pairs of y-type. Set $I = 1$ in all pairs, i.e. initially set all components in *up*. Start turning *down* component after component by moving along the permutation from left to right. Check system state on each step and locate the first component (the anchor) when the system goes *DOWN*. Let the first observed failure set has (u, v) components of type 1 and type 2, respectively. Continue turning *down* sequentially all remaining $(n_1 + n_2) - (u + v)$ components in the permutation. Note that on each step appears a new failure set.

Definition 1.4 Random permutation is called of (u, v)-*anchor type* if its anchor produces failure set of type (u, v).

Definition 1.5 Random permutation is called a (k, r)-*generator* if among the failure sets revealed during the destruction process *after* the anchor has been revealed, there is a (k, r)-failure set.

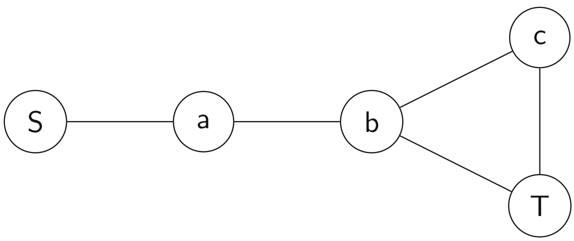


Fig. 1 Network with 5 components. It is *UP* if there is an $S - T$ connection. Edges $(S, a), (a, b)$ are of y-type, the remaining edges are of x-type

Example 1.1 **continued**. Suppose we have the following random permutation before the destruction process starts: $\pi^\star = [(x_1, 1), (x_3, 1)(y_1, 1), (x_2, 1), (y_2, 1)]$. Below are 5 stages of the sequential destruction:

$$
\begin{aligned}
\mathbf{1}: &\ [(x_1, 0), (x_3, 1), (y_1, 1), (x_2, 1), (y_2, 1)].\\
\mathbf{2}: &\ [(x_1, 0), (x_3, 0), (y_1, 1), (x_2, 1), (y_2, 1)].\\
\mathbf{3}: &\ [(x_1, 0), (x_3, 0), (\mathbf{y}_1, 0), (x_2, 1), (y_2, 1)].\\
\mathbf{4}: &\ [(x_1, 0), (x_3, 0), (y_1, 0), (x_2, 0), (y_2, 1)].\\
\mathbf{5}: &\ [(x_1, 0), (x_3, 0), (y_1, 0), (x_2, 0), (y_2, 0)].
\end{aligned}
$$

The anchor is observed on the third step and therefore $\pi^\star$ is of (2,1) anchor-type. Analysing steps 4 and 5, it is seen that $\pi^\star$ is also a (3,1) and (3,2) generator.

Definition 1.6 Denote by $F(k, r)$ the probability that a random permutation is of (k, r) anchor-type **or** is a (k, r)-type generator. Obviously,

$$F(k, r) = \frac{N(k, r)}{(n_1 + n_2)!}, \tag{3}$$

where $N(k, r)$ is the number of permutations which are of (k, r)-anchor type *or* (k, r)-generators. We call the matrix $\|F(k, r)\|_{(n_1+1)\times(n_2+1)}$ the *two-dimensional or 2D-spectrum.*

Definition 1.7 Let $g(k, r)$ be the probability that a random permutation is of (k, r)-anchor type. Obviously,

$$g(k, r) = \frac{A(k, r)}{(n_1 + n_2)!},$$

where $A(k, r)$ is the number of permutations which are of (k, r)-anchor type.

Example 1.1 **continued** Let us determine $N(2, 1)$. All permutations of three x-es and two y-s of type $(x_i, x_j, y_l, x_s, y_z)$ with one y_l on third position and two x-es among the first three positions, produce failure sets of type (2,1). By permuting the first three elements and the remaining two elements, and also by replacing y_1 by y_2 among first three elements, we will have 24 permutations for a fixed pair of x_i, x_j. Since we can choose this pair in three ways, there is a total of $N(2, 1) = 72$ permutations. Among them, there are 8 anchor-type (2,1)-permutations. These permutations must have y_j on the third position, and two x-es on the first two positions, like $\pi = (x_1, x_3, y_1, x_2, y_2)$. There are two ways to exchange the positions of x_1 and x_3, two ways to exchange y_1 by y_2 on the third position, and two ways to exchange components on the fourth and fifth positions. Therefore, for our network, $F(2, 1) = 72/5! = 0.6$ and $a(2, 1) = 8/120 = 0.0666$.

In Table 1 we present the $\|F(k, r)\|$ and $\|g(k, r)\|$ matrices for system shown on Fig. 1.

Table 1 $\|F(k,r)\|$ and $\|g(k,r)\|$ matrices

r	$k=0$	$k=1$	$k=2$	$k=3$	$r=0$	$k=0$	$k=1$	$k=2$	$k=3$
0	0.0	0.0	0.2	0.1	0	0.0	0.0	0.2	0.0333
1	0.4	0.6	0.6	0.4	1	0.4	0.3	0.0666	0.0
2	0.1	0.3	0.6	1.0	2	0	0	0	0

2.5 Counting the Number $C(k,r)$ of (k,r)-failure Sets

Here the main role is played by the following Theorem.

Theorem 1.1

$$C(k,r) = F(k,r)\frac{(n_1+n_2)!}{(k+r)!(n_1+n_2-k-r)!}. \tag{4}$$

Proof From the description of the sequential destruction of random permutation, follows that a (k,r)-failure set is a "compact" block of $(k+r)$ components located at the first $(k+r)$ positions of the permutation (the anchor-type or generated failure set). It is also obvious that one permutation can produce not more than a single (k,r) failure set. Permutations between the members of one such set produce $(k+r)!$ copies of it, and each copy is a failure set. In addition, there are $(n_1+n_2-k-r)!$ permutations of the remaining components. Therefore $N(k,r)$ permutations produce

$$\frac{N(k,r)}{(n+k)!(n_1+n_2-k-r)!}$$

original (k,r) failure sets. Remembering (3), we arrive at the desired formula (4). □

The following Corollary establishes the connection between the cumulative one-dimensional D-spectrum $F_0(k)$ (see Definition 1.3) and the 2D-spectrum.

Corollary 1.1

$$F_0(w) = \sum_{k=0}^{\min(n_1,w)} F(k,w-k).$$

Proof Suppose that we declare n_2 components of y-type to be identical to the components of x-type. Then each (k,r)-failure set becomes a $(k+r)$-failure set in the system having n_1+n_2 identical components. Therefore,

$$\sum_{k=0}^{\min(n_1,w)} C(k,w-k) = C(w),$$

or

$$\sum_{k=0}^{w} F(k, w-k)\frac{(n_1+n_2)!}{w!(n_1+n_2-w)!} = F_0(w)\frac{(n_1+n_2)!}{w!(n_1+n_2-w)!},$$

which proves the Corollary. □

Example 1.1 **continued**. Let us verify $C(2,1)$. By (4), $C(2,1) = 0.6 \cdot 5!/(3!2!) = 6$. Indeed, there are 6 failure sets having two x-type and one y-type component: $[x_1, x_2, y_1,], [x_1, x_3, y_1], [x_2, x_3, y_1], [x_1, x_2, y_2,], [x_1, x_3, y_2], [x_2, x_3, y_2]$.

2.6 Simulation Algorithm for Estimating $F(k, r)$

Algorithm 1 2D-Spectra

Input: n_1 and n_2—the number of x-type and y-type components, respectively. N-number of replications.

Output: $\hat{G}$ and $\hat{F}$ -the estimators of $\|g(k,r)\|$ and $\|F(k,r)\|$, respectively.

1: Set $t = 1$ and let $M_1[i,j]$ and $M_2[i,j]$ be two matrices with $n_1 + 1$ rows and $n_2 + 1$ columns. Put all elements of these matrices to be zero.
2: Generate $\prod_t = (\prod_1^{(t)}, \dots, \prod_{n_1+n_2}^{(t)})$ - a random component permutation.
3: Find the anchor J_t of $\prod_t$.
4: Set K_t and R_t be the number of x-type and y-type components in the first J_t elements of $\prod_t$. Set $M_1[i = K_t + 1, j = R_t + 1] = M_1[i = K_t + 1, j = R_t + 1] + 1$ and $M_2[i = K_t + 1, j = R_t + 1] = M_2[i = K_t + 1, j = R_t + 1] + 1$.
5: Set: $T_1 = K_t$ and $T_2 = R_t$.
6: **for** $i = J_t + 1$ `to` $n_1 + n_2$ **do**
7:     **if** $\prod_{i+1}^{(t)}$ is x-type component **then** set $T_1 := T_1 + 1$,
8:     **else** $T_2 := T_2 + 1$.
9:     **end if**
10:     $M_2[T_1 + 1, T_2 + 1] = M_2[T_1 + 1, T_2 + 1] + 1$.
11: **end for**
12: If $t < N$ set $t = t + 1$ and go to **Step 2**.
13: **return**: $\hat{G} = \|M_1\|/N$, $\hat{F} = \|M_2\|/N$.

Exact calculation of $F(k, r)$, like it was done in Example 1.1, becomes impractical already for n exceeding 6–8. We suggest using a Monte Carlo simulation algorithm for estimation of the $F(k, r)$ probabilities. This algorithm is based on simulating a relatively large (say 1,000,000) random permutations and extracting from them information about the number of failure sets. The algorithm below allows rather efficient and accurate estimation for networks with 50–70 components. Note that each random permutation of size n which has a (k, r)-anchor, produces also $n - (k + r)$ generated failure sets.

This algorithm has been applied to a network with 34 nodes and 68 edges, see Sect. 3. Quite accurate estimates of the G and F matrices were obtained by using $N = 10^6$ replications. The CPU time did not exceed 16 s.

3 Reliability of a Transportation Network

3.1 Description of the Network. Reliable Nodes, Unreliable Edges

The network is shown on Fig. 2. This is a hypothetical geographically oriented road network. It is designed as Barabasi-Albert system [1] with 34 nodes and 68 edges. Centrally located node 31 represents the capital city.

Important strategic objects (e.g. hospitals, supply centers, etc.) are located in terminal nodes 2, 5, 9, 33, 34. Thirteen edges are more reliable roads.

$$(14, 33), (31, 33), (33, 23), (22, 23), (5, 22), (5, 20), (29, 34),$$
$$(34, 14), (15, 14), (34, 31), (5, 31), (20, 31), (20, 29).$$

They form a ring around the capital and also contain several radial roads. These edges in our notation are the "strong" x-type edges. The remaining $68 - 13 = 55$ edges are the y-type edges. We remind that network failure means the loss of terminal connectivity: the network is *DOWN* if at least one of the terminals gets separated from other terminals. Edges can fail as a result of an enemy "attack", natural disaster or heavy road accidents, see [5, 7, 10, 15].

Table 2 presents $P(DOWN)$ calculated by (3) and Algorithm "2D-spectra" for $F(k, r)$ estimation, on the basis of generating $N = 10^6$ random permutations. The results were checked by crude Monte Carlo simulation, based also on 10^6 replications, see P_{cmc}. As it is seen from the table, the relative error is quite small which means that the estimation by our algorithm is very accurate. We see from the table that in order to provide $P(DOWN) \leq 0.05$ it is necessary to have $p_1 \geq 0.7$ for type y and about 0.8–0.9 for strong edges. Very interesting is the fact that increasing strong edge reliability from 0.9 to 0.99 has relatively little effect on $P(DOWN)$.

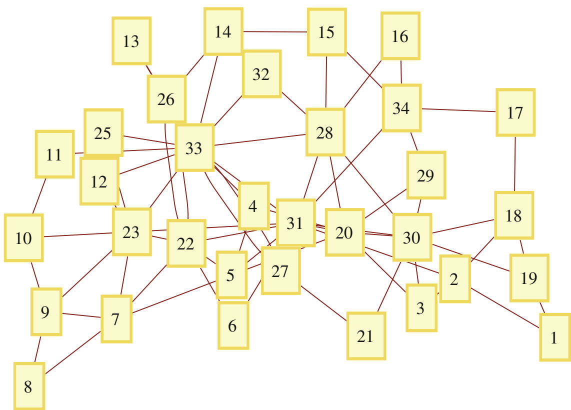


Fig. 2 Transport network with 34 nodes and 68 edges

Table 2 $P(DOWN)$ for edge failure: estimated and simulated values, $N = 1{,}000{,}000$

p_2	p_1	$P(DOWN)$	P_{cmc}	Rel.err. %
0.5	0.6	0.39536	0.39526	0.10
0.5	0.7	0.34776	0.34749	0.13
0.5	0.8	0.32061	0.31854	0.14
0.5	0.9	0.30799	0.30835	0.14
0.5	0.99	0.30615	0.31641	0.15
0.6	0.7	0.16621	0.16551	0.20
0.6	0.8	0.14808	0.14592	0.25
0.6	0.9	0.13862	0.13725	0.25
0.6	0.99	0.13754	0.13415	0.25
0.7	0.8	0.04932	0.04906	0.45
0.7	0.9	0.04551	0.04450	0.45
0.7	0.99	0.04442	0.04400	0.46
0.8	0.9	0.00877	0.00862	1.10
0.8	0.99	0.00844	0.00834	1.10
0.9	0.99	0.00053	0.00049	0.40

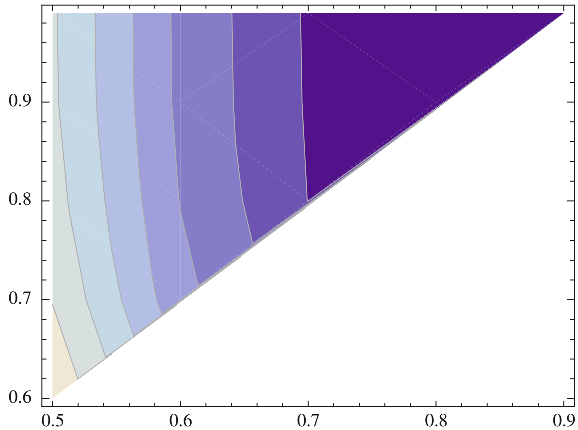


Fig. 3 Contour plot for data of Table 2 (edge failures)

Interesting information is provided by the contour plot on Fig. 3. Area with $P(DOWN) < 0.05$ is shown by deep blue color. The adjacent blue area corresponds to *DOWN* probabilities in the interval [0.05–0.1].

We also investigated the situation with edges are deteriorating in time. It as assumed that strong edge reliability $p_2(t)$ depends on time as $p_1(t) = e^{-t}$, and the remaining edges have $p_2(t) = e^{-2t}$. The numerical results are presented in Table 3.

Table 3 $P(DOWN)$ as a function of time (edge failures)

t	$p_2 = e^{-2t}$	$p_1 = e^{-t}$	$P(DOWN)$
0.1	0.819	0.905	0.0058
0.2	0.801	0.895	0.0086
0.3	0.779	0.882	0.0132
0.4	0.751	0.867	0.0214
0.5	0.716	0.846	0.0371
0.6	0.670	0.819	0.0692
0.7	0.606	0.779	0.1410
0.8	0.513	0.717	0.3142
0.9	0.368	0.607	0.7010
1.0	0.135	0.368	0.9990

3.2 Unreliable Nodes

We also have studied the network reliability when the nodes are subject to failure. Six nodes 20, 22, 23, 28, 30 and 31 are declared to be the x-type1.4 presents the results of the numerical investigation of network reliability. Again it is seen that our algorithm provides quite accurate results with a small relative error. Figure 4 (right) shows the area of parameters (p_1, p_2) where the $DOWN$ probability is smaller than 0.05 (shown by deep blue).

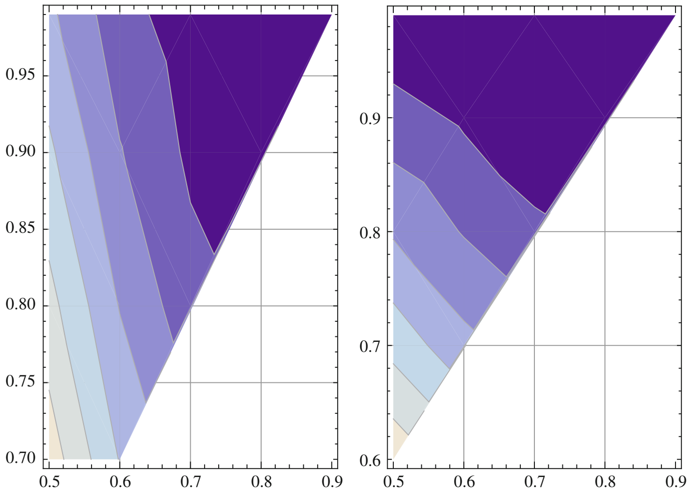


Fig. 4 Contour plots for data of Table 4. Nodes relocated (*left*), original (*right*)

Table 4 $P(DOWN)$ for node failure: estimated and simulated values, N = 1,000,000 runs

p_2	p_1	$P(DOWN)$	P_{cmc}	Rel. err. %	$P(DOWN)^{\star}$
0.5	0.6	0.33666	0.33633	0.14	–
0.5	0.7	0.23340	0.23360	0.18	–
0.5	0.8	0.14375	0.14359	0.24	0.26697
0.5	0.9	0.06719	0.06660	0.30	0.20931
0.5	0.99	0.00647	0.00634	0.12	0.16008
0.6	0.7	0.16476	0.16541	0.22	0.19729
0.6	0.8	0.09654	0.09634	0.30	0.14738
0.6	0.9	0.04240	0.04259	0.47	0.10304
0.6	0.99	0.00394	0.00393	1.60	0.06956
0.7	0.8	0.05727	0.05686	0.40	0.06905
0.7	0.9	0.02374	0.02373	0.60	0.04076
0.7	0.99	0.00220	0.00211	2.12	0.02234
0.8	0.9	0.01030	0.01046	0.95	0.01214
0.8	0.99	0.00083	0.00089	3.40	0.00435
0.9	0.99	0.00024	0.00021	6.40	0.00029

In order to see how influential is the location of the strong nodes, we relocated these nodes to periphery. Now nodes $14, 15, 18, 11, 10, 7$ are declared to be strong nodes of x-type. As it could be expected, the network with relocated strong nodes is less reliable, as it is seen from last column $P(DOWN)^{\star}$ of Table 4, and the contour surface plot on Fig. 4, on the left. Deep blue area shows low $P(DOWN)$ values, and is, therefore, the area of high reliability. It is considerably larger for the original location of the strong nodes (the plot on the right).

4 $||g(k,r)||$ Matrix and "Shock Process" Trajectories

Suppose that the network is subject to a two-dimensional "shock process" which is a random sequence of type "x"-shocks which hit randomly the strong components (strong nodes or strong edges), permuted randomly with type "y"-shocks which hit the weak components.

This process stops when the network fails. As it follows from the definition of the permutation destruction process, the networks fails at the "stopping point" determined by the permutation anchor. The distribution of the location $(x = V, y = U)$ of the "stopping point" is shown on Fig. 5 by means of surface contour plot. In this example the "shocks" kill strong and weak nodes.

The exact probabilistic meaning of this plot is the following. The elements of matrix $G = ||g(k,r)||$ present the conditional probabilities that the shock process stops at coordinate $(V = k, U = r)$, given that network is $DOWN$:

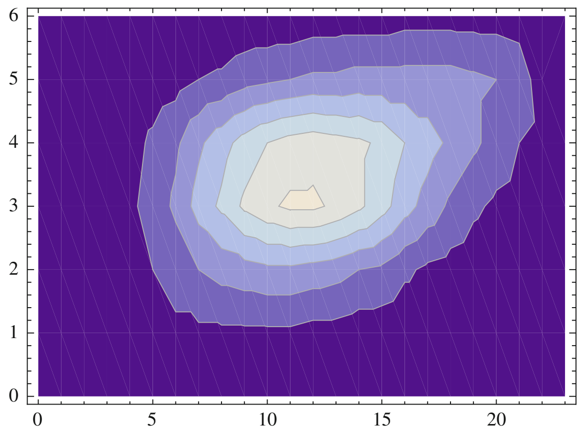


Fig. 5 Contour plot for $\|g(k, r)\|$ matrix

$$g(k, r) = P((V = k, U = r)|DOWN),$$

where V, U are the numbers of strong and weak destroyed components at the stopping point, respectively.

Let us examine the plot on Fig. 5. The horizontal axis is for weak nodes, vertical axis—for strong. By deep blue is shown the area where the trajectory does not stop. Here the trajectory does not stop at all. The adjacent area (light blue) shows points having stopping probability between 0.005 and 0.01. Next area closer to the center shows the points having probabilities between 0.01 and 0.015, and so on. So,the point $g(V = 2, U = 7)$ lies in the probability interval $[0.010, 0.015]$.

The $\|g(k, r)\|$ matrix is a valuable structural characteristic of the network. Let us demonstrate its use by investigating so-called "quantile areas".

Contrary to the definition of a quantile for one-dimensional case, for more dimensions there are many ways to determine the area which has probabilistic mass q, see e.g. [16]. Let us consider here the triangular areas of type $U + V \leq D$. Omitting the routine calculations, we present the following results for $D = 4, 5, 6, 7, 8, 9$:

$$P(U + V \leq 4) = 0.0068, P(U + V \leq 5) = 0.015, P(U + V \leq 6) = 0.030,$$

$$P(U + V \leq 7) = 0.052, P(U + V \leq 8) = 0.084, P(U + V \leq 9) = 0.128.$$

So, for example, the network fails with probability 0.128 if the total number of failed nodes is not more than 9.

Comparing the size of equal quantile areas may serve as an instrument to compare the reliability of alternative structures. For example, structure A is more reliable than structure B if the two dimensional 0.1-quantile area $D_A(q = 0.1)$ for A is larger than the similar area $D_B(q = 0.1)$ for structure B.

5 More Than Two Types of Components—Concluding Remarks

Suppose that the network has $K > 2$ different groups of i.i.d. components. Then the expression for network *DOWN* probability will be a natural extension of (3) to more variables. Denote by n_i the number of i-th type components, $n_1 + n_2 + \cdots + n_K = n$, and let $C(x_1, \ldots, x_K)$ be the number of failure sets having x_i components of i-th type *down*, $i = 1, .., K$. Then

$$P(DOWN) = \sum_{0 \le x_i \le n_i, i=1,\ldots,K} C(x_1, x_2, \ldots, x_K) \prod_{i=1}^{K} q_i^{x_i} \prod_{i=1}^{K} p_i^{n_i - x_i}.$$

The main problem remains estimation of $C(x_1, \ldots, x_K)$, the numbers of failure sets. This can be done in the framework of the above described Algorithm, with obvious modifications. Now the random permutation will have K types of symbols for denoting components of K groups, and now the failure sets of anchor-type and of generated type will have x_i components of i-th type, $i = 1, \ldots, K$ For $K = 3$, for example, the F-matrix will become a three-dimensional cubic matrix.

There are several important issues left outside the scope of the present paper. Let us mention on the first place the investigation of component importance, see e.g. [6]. Similar to the networks with one type of components, for several types of components, importance issues are the key to optimal network design and to the "nomination" of the components to be the "strong" ones.

Very interesting would be also to compare several competing network structures by analyzing their q-quantile "areas", as it was briefly discussed in Sect. 4. We leave these issues for the future research.

Acknowledgements The work of Radislav Vaisman was supported by the Australian Research Council Centre of Excellence for Mathematical & Statistical Frontiers, under grant number CE140100049.

References

1. Barabasi A, Albert R (1999) Emergence of scaling in random networks. Science 286:509–551
2. Daqing L, Qiong Z, Zio E, Havlin S, Kang R (2015) Network reliability analysis based on percolation theory. Eng Reliab Syst Saf 142:1–15
3. ElperinT Gertsbakh I, Lomonosov M (1991) Estimation of network reliability using graph evolution models. IEEE Trans Reliab 40:572–581
4. Gertsbakh I, Shpungin Y (2010) Models of network reliability: analysis, combinatorics and Monte Carlo. CRC Press, Boca Raton
5. Gertsbakh I, Shpungin Y (2011) Network reliability and resilience. Springer briefs in electrical and computer engineering. Springer, Berlin

6. Gertsbakh I, Shpungin Y (2012) Using D-spectra in Monte Carlo estimation of system reliability and component importance. In: FrenkelI LA (ed) Recent advances in system reliability. Springer, London, pp 49–61
7. Hudges J, Healy K (2014) Measuring the resilience of transport infrastructure. New Zealand Transport Agency, Report No. 546
8. Nagurney A, Qlang Q (2007) Robustness of transportation networks subject to degradable links. Lett J Explor Front Phys
9. Newman M (2010) Networks. Oxford University Press
10. Peeta S, Salman S, Gunnec D, Kannan V (2010) Predisaster investment decisions for strengthening a highway network. Comput Oper Res 37:1708–1719
11. Schneider C, Yazdani N, Araujo N, Havlin S, Herrmann H (2013) Toward designing robust coupled networks. Scientific reports
12. Samaniego F (1985) On closure of the IFR under formation of coherent systems. IEEE Trans Reliab 34:69–72
13. Samaniego F (2007) System signatures and their applications in engineering reliability. Springer
14. Samaniego F, Navarro J, Balakrishnan N (2011) Signature-based representation of the reliability of systems with heterogeneous components. J Appl Probab 48:856–867
15. Sanchez-Silva M, Daniels M, Lieras G, Patino D (2005) A transport network reliability model for the efficient assignment of resources. Transp Res Part B: Methodol 39(1):47
16. Sering R (2005) Quantile functions for multivariate approaches and applications. Stat Neerl 56(part B):214–232
17. Sterbenz J, Hutchinson D, Çetinkayaa E, Jabbara A, Rohrera J, Schöllerc M, Smith P (2010) Resilience and survivability in communication networks: strategies, principles and survey of discipline. Comput Netw 54(8):1245–1265

Reliability Analysis of Complex Multi-state System with Common Cause Failure Based on DS Evidence Theory and Bayesian Network

Jinhua Mi, Yan-Feng Li, Weiwen Peng and Hong-Zhong Huang

Abstract With the increasing complexity and larger size of modern advanced engineering systems, the traditional reliability theory cannot characterize and quantify the complex characteristics of complex systems, such as multi-state properties, epistemic uncertainties, common cause failures (CCFs), etc. This chapter focuses on the reliability analysis of complex multi-state system (MSS) with epistemic uncertainty and CCFs. Based on the Bayesian network (BN) method for reliability analysis of MSS, the DS evidence theory is used to express the epistemic uncertainty in system through the state space reconstruction of MSS. An uncertain state, which used to express the epistemic uncertainty is introduced in the new state space. The integration of evidence theory with BN is achieved by updating the conditional probability tables. When the multiple CCF groups (CCFGs) are considered in complex redundant systems, a modified factor parametric model is introduced to model the CCF in systems. An evidence theory based BN method is proposed for the reliability analysis and evaluation of complex MSSs in this chapter. The reliability analysis of servo feeding control system for CNC heavy-duty horizontal lathes (HDHLs) by this proposed method has shown that the presented method has high computational efficiency and strong practical value.

Keywords Complex multi-state system · Multi-state bayesian network · DS evidence theory · Multiple common cause failure groups

J. Mi · Y.-F. Li · W. Peng · H.-Z. Huang (✉)
Center for System Reliability and Safety, University of Electronic Science and Technology of China, Chengdu, P. R. China
e-mail: hzhuang@uestc.edu.cn

J. Mi
e-mail: jinhuami@126.com

Y.-F. Li
e-mail: yanfengli@uestc.edu.cn

W. Peng
e-mail: wwpeng@uestc.edu.cn

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*, Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_2

1 Introduction

The multi-state system (MSS) was firstly proposed by Barlow and Wu, it has been proved that lots of industrial systems are typical MSS, such as electrical power system, pipe transmission system, production and manufacturing system, aerospace system, etc. [1, 5, 9]. Those systems can define the multi-state characteristics of components accurately by analyzing the system failure process and the effect of the change of component performance to the system performance and reliability. Four types of methods can be used for reliability analysis of MSS, including multi-state fault tree method [24], Markov process method [6, 7], Monte-Carlo simulation (MCS) method [14, 31] and universal generation function (UGF) method [4]. The MSS plays a critical role in the reliability analysis and assessment of complex systems and also has extensive application foreground.

The uncertainty caused by lack of data and scarcity of information is one of the most important issues in MSS reliability analysis. When the system state performances and state probabilities cannot be exactly defined and obtained, sometimes the bounds of system states and state probabilities cannot be exactly defined and obtained, so the probability-based methods are no longer applicable for this kind of system. In this situation, the bounds of system states and state probabilities can be expressed by some other data forms, such as linguistic variables. Then some non-probabilistic methods are developed, such as Dempster-Shafer evidence theory (DSET) [27], fuzzy theory [13], probability-box [10, 26], interval theory [17], possibility theory [8], Bayesian method [22, 23], etc. The DS evidence theory has a flexible axiomatic system to describe uncertainty, and also has an independent frame to process uncertainty in system [18, 25]. It has been widely used for uncertainty modeling, quantification, reasoning and management in engineering [3, 29, 30].

There are many researches on Bayesian network (BN) based on evidence theory. Simon et al. [20, 21] analyzed reliability of complex system with epistemic uncertainty by using BN, where evidence theory is used to quantify system uncertainty. Then the evidential networks also have been used for the reliability and performance evaluation of system with imprecise knowledge [19]. Zhao et al. [28] studied the influence of incomplete original parameters and subjective parameters on the reliability of distribution system by using BN and evidence theory. Sallak et al. [16] has developed the combination method of BN and evidence theory for reliability analysis of multi-state system (MSS). It has shown that evidence theory can handle the imprecise information in system, and it can get more useful information than interval analysis method.

This chapter introduces a multi-state BN method for reliability analysis of complex system with CCFGs based on evidence theory. The remainder of this chapter is organized as follows: Firstly, the node definition and BN reasoning of multi-state BN under evidence theory are introduced in Sect. 2. Then, in Sect. 3, when the multiple CCF groups (CCFGs) are considered in complex redundant system, a modified β factor parametric model is introduced to model the CCF in system. This comprehensive method is used to analyze the reliability of an example system and a feeding control system of CNC heavy-duty horizontal lathe (HDHL) in Sect. 4. Finally, some conclusions are presented in Sect. 5.

2 Multi-state Bayesian Network Under Evidence Theory

2.1 The Node Definition of Multi-state Bayesian Network Under Evidence Theory

For a sample BN with three nodes which is shown in Fig. 1, assume that the nodes x_1 and x_2 are three state nodes, the state space is $\Lambda = \{0, 1, 2\}$. Let $x_i = 0, 1, 2$ represent the reliable state, partial failure state and complete failure state of the corresponding component. When the epistemic uncertainty exists in system, an added state $x_i = [0, 1, 2]$ is defined to represent the uncertain state of node x_i. Then the frame of discernment $D = \{0, 1, 2, [0, 1, 2]\}$ is defined under evidence theory, and the basic probability assignment (BPA) is $m\colon 2^D \to [0, 1]$. Its power set can be expressed as

$$2^D = \{m(x = \varnothing) = 0; m(x = 0); m(x = 1); m(x = 2); m(x = [0, 1, 2])\}. \quad (1)$$

For an event $A\colon \{x = 0\}$ on the frame of discernment D, and $B \subseteq A$, the belief function of event A is

$$Bel(A) = \sum_{B \subseteq A} m(A) = m(x = 0). \quad (2)$$

The (2) represents the belief degree of event $A\colon x = 0$. It's the lower bound of belief interval when the probability of uncertain information is not counting in the BPA [21]. Based on the definition of plausibility function, the plausibility function of event A can be gotten by

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B) = m\{x = 0\} + m\{x = [0, 1, 2]\}. \quad (3)$$

Then the interval probability of event A can be calculated by (2) and (3), and it can be expressed as $[P](A) = [Bel(A), Pl(A)]$. Similarly, the interval probabilities of other nonempty events under the frame of discernment D can be computed.

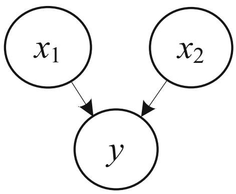


Fig. 1 A sample multi-state BN

When the corresponding components of nodes x_1 and x_2 are parallel or series, the conditional probability table (CPT) of node y under evidence theory can be derived. Then the belief reliability of node y: $Bel(y = 0)$ can be computed by BN reasoning and

$$Bel(y=0) = \sum_{x_1, x_2} Bel(y=0|x_1, x_2) Bel(x_1) Bel(x_2). \tag{4}$$

The plausibility reliability $Pl(y = 0)$ is

$$Pl(y=0) = \sum_{x_1, x_2} Pl(y=0|x_1, x_2) Pl(x_1) Pl(x_2). \tag{5}$$

The practical reliability of node y: $P(y = 0)$ will belongs to interval $[Bel(y = 0), Pl(y = 0)]$.

2.2 The Multi-state Bayesian Network Reasoning Under Evidence Theory

For a multi-state BN with n root nodes, which can be denoted as $x_1, x_2, \ldots, x_n$. Assume that the state number of node x_i and leaf node y are l_i and l_y. The relation between the state probability of leaf node y and root nodes can be expressed as

$$P\left(y = y^j | x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right) = \frac{P\left(y = y^j, x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right)}{P\left(x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right)}, \tag{6}$$

where $1 \le j \le l_y, 1 \le i \le n$ and $1 \le k_i \le l_i$. Suppose that the interval probability of node x_i at state k_i is

$$[P]\left(x_i = x_i^{k_i}\right) = \left[Bel\left(x_i^{k_i}\right), Pl\left(x_i^{k_i}\right)\right], \tag{7}$$

where $Bel(x_i^{k_i})$ and $Pl(x_i^{k_i})$ can be calculated by the CPTs of nodes x_i.

The conditional probability of node y of BN under evidence theory is

$$[P]\left(y = y^j | x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right) = \left[Bel\left(y^j\right), Pl\left(y^j\right)\right]. \tag{8}$$

The mid-value of conditional probability is chosen as the static conditional probability of this node [21], that is

$$P\left(y = y^j | x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right) = \frac{[Bel(y^j) + Pl(y^j)]}{2}. \tag{9}$$

Similarly, the static conditional probability of root node x_i on state $x_i^{k_i}$ is

$$P\left(x_i = x_i^{k_i}\right) = \frac{\left[Bel\left(x_i^{k_i}\right) + Pl\left(x_i^{k_i}\right)\right]}{2}. \quad (10)$$

The probability of node y on j-th state can be gotten by

$$P\left(y = y^j\right) = P\left(y = y^j \middle| x_1 = x_1^{k_1}, \ldots, x_n = x_n^{k_n}\right) P\left(x_1 = x_1^{k_1}\right) \cdots P\left(x_n = x_n^{k_n}\right). \quad (11)$$

For a BN with n root nodes $x_i (i = 1, 2, \ldots, n)$, m non-leaf nodes $y_j (j = 1, 2, \ldots, m)$ and leaf node T. Based on the former reasoning method, the probability of leaf node $T = T_v$ can be expressed as

$$[P](T = T_v) = [Bel(T = T_v), Pl(T = T_v)], \quad (12)$$

where the lower bound $Bel(T = T_v)$ is the belief probability and can be calculated by

$$\begin{aligned} Bel(T = T_v) &= \sum_{x_1, \ldots, x_n, y_1, \ldots, y_m} Bel(x_1, \ldots, x_n, y_1, \ldots, y_m, T = T_v) \\ &= \sum_{\pi(T)} Bel(T = T_v | \pi(T)) \prod_{j=1}^{m} \sum_{\pi(y_1)} Bel\left(y_j \middle| \pi\left(y_j\right)\right) \prod_{i=1}^{n} Bel\left(x_i^{k_i}\right) \\ &= \sum_{\pi(T)} Bel(T = T_v | \pi(T)) \sum_{\pi(y_1)} Bel(y_1 | \pi(y_1)) \times \cdots \\ &\quad \times \sum_{\pi(y_m)} Bel(y_m | \pi(y_m)) \times \cdots \times Bel(x_1 = x_{1,k_1}) \times \cdots \times Bel(x_n = x_{n,k_n}). \end{aligned} \quad (13)$$

The upper bound $Pl(T = T_v)$ is the plausibility probability and can be computed by

$$\begin{aligned} Pl(T = T_v) &= \sum_{x_1, \ldots, x_n, y_1, \ldots, y_m} Pl(x_1, \ldots, x_n, y_1, \ldots, y_m, T = T_v) \\ &= \sum_{\pi(T)} Pl(T = T_v | \pi(T)) \prod_{j=1}^{m} \sum_{\pi(y_1)} Pl\left(y_j \middle| \pi\left(y_j\right)\right) \prod_{i=1}^{n} Pl\left(x_i^{k_i}\right) \\ &= \sum_{\pi(T)} Pl(T = T_v | \pi(T)) \sum_{\pi(y_1)} Pl(y_1 | \pi(y_1)) \times \cdots \\ &\quad \times \sum_{\pi(y_m)} Pl(y_m | \pi(y_m)) \times \cdots \times Pl(x_1 = x_{1,k_1}) \times \cdots \times Pl(x_n = x_{n,k_n}). \end{aligned} \quad (14)$$

The probability of leaf node can be obtained by the former forward reasoning of BN, and the posterior probability of root nodes can be gotten by backward reasoning. When $T = T_v$, the posterior probability of root node $x_i = x_{i,k_i}$ can be computed by

$$[P](x_i = x_{i,k_i} | T = T_v) = \begin{bmatrix} \min(Bel(x_i = x_{i,k_i} | T = T_v), Pl(x_i = x_{i,k_i} | T = T_v)), \\ \max(Bel(x_i = x_{i,k_i} | T = T_v), Pl(x_i = x_{i,k_i} | T = T_v)) \end{bmatrix}, \tag{15}$$

where

$$Bel(x_i = x_{i,k_i} | T = T_v) = \frac{Bel(x_i = x_{i,k_i}, T = T_v)}{Bel(T = T_v)}, \tag{16}$$

$$Pl(x_i = x_{i,k_i} | T = T_v) = \frac{Pl(x_i = x_{i,k_i}, T = T_v)}{Pl(T = T_v)}, \tag{17}$$

where $Bel(x_i = x_{i,k_i}, T = T_v)$ and $Pl(x_i = x_{i,k_i}, T = T_v)$ are the belief and plausibility joint probability of root nodes and leaf node. The root nodes and leaf node of BN reflect the fault causes and fault state properties of system. Therefore, the system state probability can be computed by forward reasoning of BN, which can also realize a quantitative description of system fault states. The backward reasoning of BN can get the posterior probability of fault causes based on system failure state, and also can implement the system failure prediction and judgment, which has certain guiding significance for the reliability improvement of system.

3 Reliability Modeling of System with Multiple CCFGs

3.1 A Modified β Factor Model for CCFGs

Considering the dependent failure caused by interior component physical interactions and human interactions in system, the beta factor parametric model has been widely used for such cases [15]. Assume that P_t is the total failure probability of a component; it can be expanded into an independent contribution P_{ind} and a dependent contribution P_{ccf}, which are functions of time t respectively. When the component is assumed to follow the exponential distribution, λ_t, λ_{ind} and λ_{ccf} are the failure rates of entire system, independent part, and the dependent part respectively. Then the parameter β can be defined as the fraction of the total failure probability attributable to dependent failures [11, 15], and it can be mathematically described as

$$\begin{aligned} \beta = \frac{P_{ccf}}{P_t} = \frac{P_{ccf}}{P_{ind} + P_{ccf}} &= \frac{(1 - \exp(-\lambda_{ccf} \cdot t))}{(1 - \exp(-\lambda_t \cdot t))} \\ &= \frac{(1 - \exp(-\lambda_{ccf} \cdot t))}{(1 - \exp(-\lambda_{ind} \cdot t)) + (1 - \exp(-\lambda_{ccf} \cdot t))}. \end{aligned} \tag{18}$$

The value of β-factor can be obtained by the direct use of field data and experts' experience [11, 12, 15].

In order to present how the beta factor model works, a simple deduction is performed for single component within the FTA model. For a parallel system with two identical components A_1 and A_2, and $P(A_1) = P(A_2) = P_A$, the failure probability of system P_{sys} can be computed as:

$$P_{sys} = P(A_1)P(A_2) = P_A^2 \tag{19}$$

For the basic component A, as shown in Fig. 2, the failure probability of A can be divided into two proportions: independent part and CCF part, and it can be expressed as

$$P_A = P_{A\_ind} + P_{A\_ccf} \tag{20}$$

Adding the CCF part, the failure probability is

$$P_{A\_ccf} = \beta P_A \tag{21}$$

By using the former explicit modeling method, the failure probabilities of component A_1 and A_2 are both divided into independent part and CCF part. Then based on the standard β-factor model and (18), the probability of CCF part can be obtained and $P_{A_1\_ccf} = P_{A_2\_ccf} = \beta P_A$. The two components parallel system also can be further expressed as Fig. 3a. The system failure event Sys can be simplified by using Boolean algebra operation rules and expressed as

$$\begin{aligned} Sys &= A_1A_2 = (A_1\_ind + A_1\_ccf)(A_2\_ind + A_2\_ccf) \\ &= (A_1\_ind + A\_ccf)(A_2\_ind + A\_ccf) \\ &= A_1\_ind \cdot A_2\_ind + A\_ccf \cdot (A\_ccf + A_1\_ind + A_2\_ind) \\ &= A_1\_ind \cdot A_2\_ind + \underbrace{A\_ccf \cdot A\_ccf}_{A\_ccf} + \underbrace{A\_ccf \cdot A_1\_ind}_{0} + \underbrace{A\_ccf \cdot A_2\_ind}_{0} \\ &= A_1\_ind \cdot A_2\_ind + A\_ccf \end{aligned} \tag{22}$$

Finally, the system with consideration of CCF can be simplified and shown as Fig. 3b.

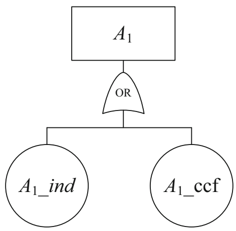
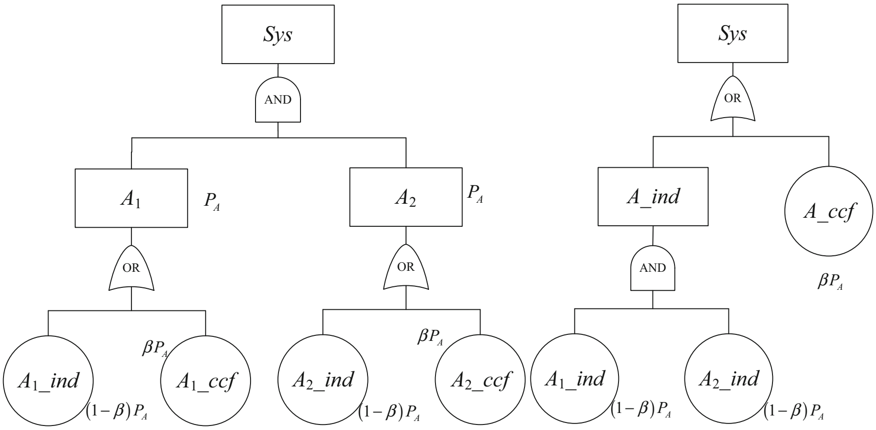


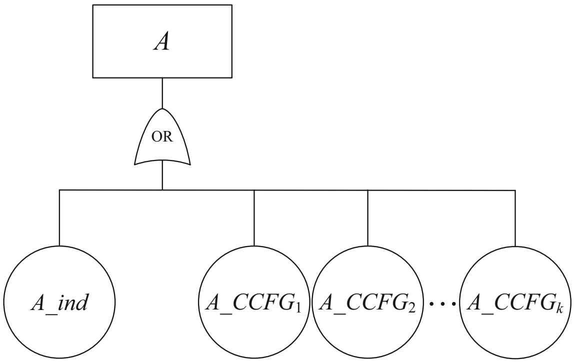
Fig. 2 FTA explicit modeling method for component with CCF



(a)FT of two components parallel system; (b) Simplify FT of system

Fig. 3 FT modeling and simplification of two components parallel system with common cause events

Fig. 4 Explicit modeling of multiple CCFGs within FT



Then the failure probability of system can be obtained by (22) and

$$\begin{aligned} P(Sys) &= P(A_1\_ind \cdot A_2\_ind + A\_ccf) \\ &= P(A_1\_ind) \cdot P(A_2\_ind) + P(A\_ccf) \\ &= (1-\beta)P(A) \cdot (1-\beta)P(A) + \beta P(A) \end{aligned} \tag{23}$$

When a single component fails simultaneously within multiple CCFGs [2, 12], a modified beta factor parametric model is used to express the coupling mechanism. The explicit modeling of component A with multiple CCFGs is shown in Fig. 4.

The failure probability of component A is then given as

$$
\begin{aligned}
P(A) &= P(A\_ind) + P(A\_ccf) \\
&= P(A\_ind) + P(A\_CCFG_1 \cup \cdots \cup A\_CCFG_k) \\
&= P(A\_ind) + P(A\_CCFG_1) + \cdots P(A\_CCFG_k)
\end{aligned} \tag{24}
$$

In this way, the failure probability of component A is divided into CCF parts and independent part as follow

$$
\begin{aligned}
P_{A\_ccf} &= P_{A\_CCFG_1} + P_{A\_CCFG_2} + \cdots + P_{A\_CCFG_k} \\
&= \beta_1 P_A + \beta_2 P_A + \cdots + \beta_k P_A = P_A \sum_{i=1}^{k} \beta_i
\end{aligned} \tag{25}
$$

$$
P_{A\_ind} = (1 - \sum_{i=1}^{k} \beta_i) P_A \tag{26}
$$

3.2 Model Limitation and Solution

Because the beta factors are obtained by expert judgments, there exists the limitation of this modified beta factor parametric model for $(\beta_1 + \beta_2 + \cdots + \beta_k) > 1$. In this case, the failure probability of CCF part is bigger than the probability of total components. To cope with this limitation in this model, a proportional reduction factor (PRF) method [2, 12] is applied in this chapter. The PRF factor is defined as

$$
\mathrm{PRF} = \frac{1}{\sum_{j=1}^{k} \beta_j} \tag{27}
$$

Then a set of new reduced beta factor are generated as

$$
\boldsymbol{\beta} = \left[\beta_1^{'}, \beta_2^{'}, \ldots, \beta_k^{'}\right] = \mathrm{PRF}[\beta_1, \beta_2, \ldots, \beta_k]. \tag{28}
$$

In this way, the failure probability of CCF parts and independent part are rewritten as

$$
P_{A\_ccf} = P_A \sum_{i=1}^{k} \beta_k^{'} \tag{29}
$$

$$
P_{A\_ind} = (1 - \sum_{i=1}^{k} \beta_k^{'}) P_A = 0 \tag{30}
$$

The essence of the PRF method is an equilibrium process of the accumulated common cause parts on β factor, which has weaken the contradiction of the common cause part beyond the total failure probability to some extent. And the PRF method is just one of the methods which can be used to solve this kind of logical contradiction; any other method capable of dealing with such contradictions may be applicable.

3.3 The Bayesian Network Node with CCFGs

When CCFs is considered in system reliability modeling, the failure of system can be divided into independent part and CCF part. The independent part means the fail of system caused by a single cause, and the CCF part represents the simultaneous failure of multiple components which caused by a common coupling mechanism, then those components constitute a CCFG. Component A exists in multiple CCFGs, $(CCFG_1, CCFG_2, \ldots, CCFG_k)$. By using the fault tree explicit modeling of multiple CCFGs in Sect. 3.2, the fault tree can be translated into BN, as shown in Fig. 5.

When the independent failure probability of node A is $P(A_{ind})$, the corresponding β factors of common cause nodes are $\beta_1, \beta_2, \ldots, \beta_k$. When $(\beta_1 + \beta_2 + \cdots + \beta_k) < 1$, the failure probability of node A can be calculated by (23)–(27) and

$$\begin{aligned} P'(A) &= P(A_{ind}) + P\left(A_{ccf}\right) \\ &= P(A_{ind}) + \frac{\sum_{i=1}^{k} \beta_i}{1 - \sum_{i=1}^{k} \beta_i} P(A_{ind}) = \frac{1}{1 - \sum_{i=1}^{k} \beta_i} P(A_{ind}) \end{aligned} \tag{31}$$

When the sum of β factors are larger than 1, that is $(\beta_1 + \beta_2 + \cdots + \beta_k) > 1$, by using the PRF method in Sect. 3.2, the failure probability of node A can be computed by (27)–(30) and

$$P'(A) = P'_{A\_ccf} + P'_{A\_ind} = P_A \sum_{i=1}^{k} \beta'_i + 0 = P_A \cdot PRF \cdot \sum_{i=1}^{k} \beta_i \tag{32}$$

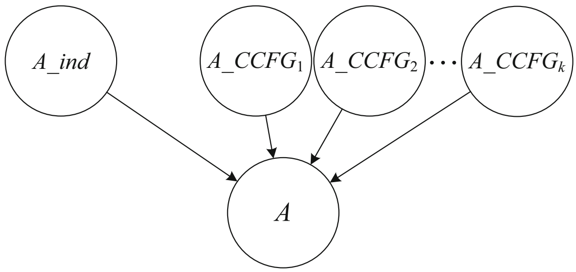


Fig. 5 BN node with CCFGs

4 Reliability Analysis of Feeding Control System for CNC HDHLs with Multiple CCFGs

4.1 Fault Tree Modeling of Feeding Control System

The DL series horizontal lathes are computer numerical control (CNC) types and have the following work axes: X axis of tool head lateral movement, Z axis of tool head longitudinal movement, U_1 axis of left gang tool movement and U_2 axis of right gang tool movement. The functional block diagram of the electrical control and drive system for such DL series horizontal lathes is shown in Fig. 6. The feeding control system include 3 subsystems: X, Z, as well as U_1 and U_2 axes feeding control systems. A signal generated by 611D-type servo driven module (*Mo*) is transmitted through electric wire (*Ew*) to control the motor (*Mt*) in X axis feeding control system. There exists a speed feedback device (*Sf*). The grating scales (*Gr*) feedback the straightness of X axis to *Mo* to adjust the feed speed and direction. The electrical control of Z, U_1 and U_2 axes is almost the same as that of X axis, excepting the difference introduced in Sect. 1. Although U_1 and U_2 axes share a 611D-type servo driven module, they have different current relays (*Re*).

Based on the function analysis and failure mechanism analysis of feeding control system, the "functional failure of feeding control system" has been chosen as the top event in FTA, and the fault tree of feeding control system is built and shown in Fig. 7.

The meanings of the notations in Fig. 7 are as following: T denotes the functional failure of feeding control system; XF, ZF, U_1F and U_2F are the functional failures of X, Z, U_1 and U_2 axes feeding control systems. The basic components of each axes feeding control system include *Gr*, *Sf*, *Ew*, *Mo*, *Mt* and *Re*. Therefore, in the fault tree model, the failure events of basic components are noted by two parts: the code of axes and the code of each component. For example, X^{Ew} represents the *Ew* failure of X axis feeding control system, and the other notations follow the similar interpretations.

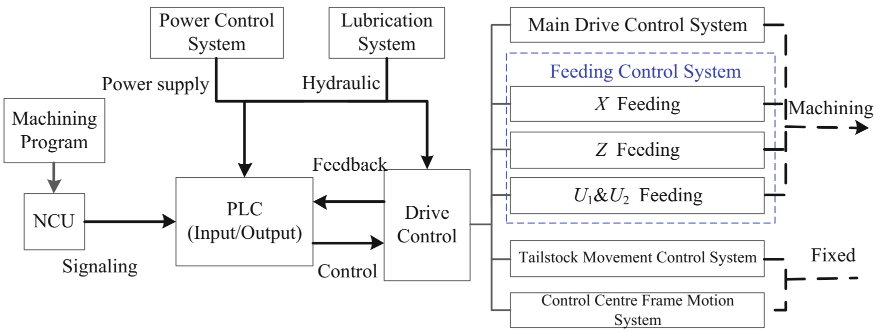


Fig. 6 Functional block diagram of electrical control and drive system for the DL series CNC HDHL

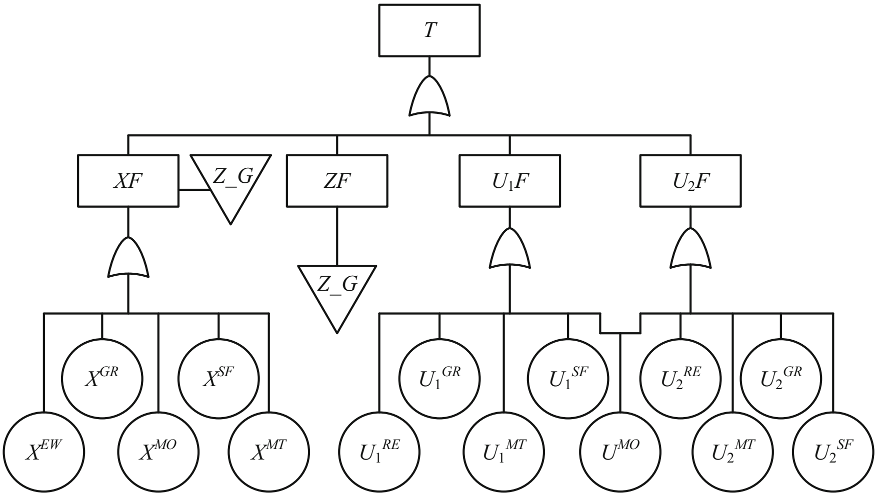


Fig. 7 Fault tree model of the feeding control system

4.2 The BN Modeling and CCFGs Fusion

From Fig. 7, the different subsystems of feeding control system have several identical or similar components, an external shock or interior component physical interactions may cause the failure of those components simultaneously. So when considering the CCF caused by human interactions, system function correlation and environment, the following common cause events or CCFGs will exist in system.

(1) $C^{MO} = \{X^{MO}, Z^{MO}, U^{MO}\}$, which means the motors of different subsystems fail at the same time by one influence factor. Based on expert experience, the common cause factor $\beta^{MO} = 0.1$.
(2) $C^{GR} = \{X^{GR}, Z^{GR}, U_1^{GR}, U_2^{GR}\}, C^{SF} = \{X^{SF}, Z^{SF}, U_1^{SF}, U_2^{SF}\}$ and $\beta^{GR} = 0.2$, $\beta^{SF} = 0.15$.
(3) $C^{EW} = \{X^{EW}, Z^{EW}\}, C^{RE} = \{X^{RE}, Z^{RE}\}$ and $\beta^{EW} = \beta^{RE} = 0.15$.
(4) When X^{MT} exists in multiple CCFGs, and expressed as $CCFG_1^{MT} = \{X^{MT}, Z^{MT}\}, \{X^{MT}, U_1^{MT}\}, \{X^{MT}, U_2^{MT}\}, \{Z^{MT}, U_1^{MT}\}, \{Z^{MT}, U_2^{MT}\}, \{U_1^{MT}, U_2^{MT}\}$; $CCFG_2^{MT} = \{X^{MT}, Z^{MT}, U_1^{MT}\}, \{X^{MT}, Z^{MT}, U_2^{MT}\}, \{Z^{MT}, U_1^{MT}, U_2^{MT}\}$ and $CCFG_3^{MT} = \{X^{MT}, Z^{MT}, U_1^{MT}, U_2^{MT}\}$. The corresponding common cause factors of two components, three components and four components failure simultaneously are $\beta_1^{MT} = 0.25, \beta_2^{MT} = 0.2$ and $\beta_3^{MT} = 0.15$.

The failure rates and failure probabilities of system components at $t = 3000$ h are listed in Table 1. Based on the transformation method of fault tree to BN and the modified β factor model, the fault tree of feeding control system can be transformed to BN and decomposed by explicit modeling method. When CCFs are considered,

Table 1 The failure rates and failure probabilities of components

Code	Failure rate λ (10^{-6}/h)	Failure probability (t = 3000 h)	Code	Failure rate λ (10^{-6}/h)	Failure probability (t = 3000 h)
MO	0.2	0.0006	*MT*	7	0.0208
EW	0.6	0.0018	*SF*	0.5	0.0015
GR	2	0.0060	*RE*	2	0.0060

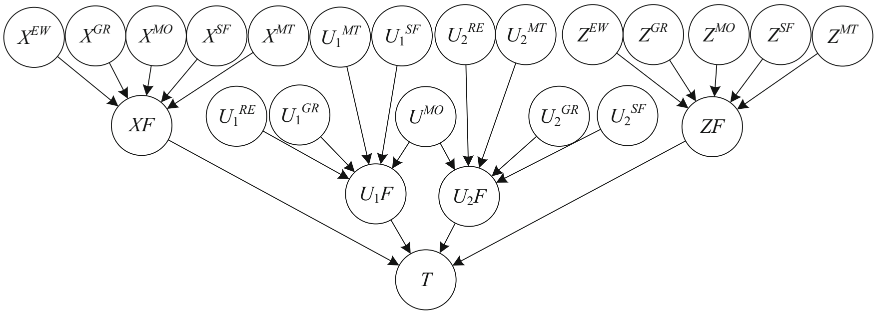


Fig. 8 The system BN with consideration of CCF

the root nodes of BN can be decomposed into independent parts and common cause parts. Then the system BN with consideration of CCFGs can be gotten as Fig. 8.

The BN of Fig. 10 is the same as system BN structure without considering CCF, the difference is the redefinition of the probabilities of root nodes, then CCF of each components can be taken into consideration. The failure probabilities of components in Table 1 are independent probabilities, then the root nodes' actual failure probabilities can be updated by modified β factor model.

For component A which is not included in multiple CCFGs, the updated failure probabilities of this kind of basic components can be calculated by (31) and $P'(EW) = 0.0021, P'(RE) = 0.0071,\ P'(GR) = 0.0075, P'(SF) = 0.0018,\ P'(MO) = 0.0007$. For component MT which is included in multiple CCFGs, the failure probability of MT can be computed by (31) since it does not meet the limitation that the sum of β factors of different CCFGs is larger than 1, then

$$\begin{aligned} P'(MT) &= P(MT_{ind}) + P\left(MT_{ccf}\right) \\ &= P(MT_{ind}) + P\left(\left(CCFG_1^{MT}\right)\right) + P\left(\left(CCFG_2^{MT}\right)\right) + P\left(\left(CCFG_3^{MT}\right)\right) \\ &= P(MT_{ind}) + \left(\frac{\sum_{i=1}^{3}\beta_i^{MT}}{1-\sum_{i=1}^{3}\beta_i^{MT}}\right)P(MT_{ind}) = \frac{1}{1-\sum_{i=1}^{3}\beta_i^{MT}}P(MT_{ind}) \end{aligned} \tag{33}$$

Because $(\beta_1^{MT} + \beta_2^{MT} + \beta_3^{MT}) < 1$, the logical contradiction of modified β factor model is inexistence here. So, the failure probability of this kind of components with consideration of CCFs can be calculated directly, and $P'(MT) = 0.0520$.

4.3 Reliability Analysis of Feeding Control System by Using DSET Based BN

As the main power take-off components of horizontal lathe, the work state of motors will affect the processing efficiency directly. Therefore, in this chapter, there exists an intermediate state between the perfect work state and failure state of the motors of DL series horizontal lathes, called derating working state. So the state space of motors can be expressed as $\{0, 1, 2\}$, where, 0 is the perfect working state, 1 is the derating working state and 2 represents failure state. The other components of system are all considered as two-state component. Due to the complexity of system structure and the coupling relation between components, only a little reliability data are available, an uncertain state [0 ,1, 2] is induced to the state space to represent the uncertainty of system. Assume that the life of all components obey exponential distribution, the basic components state probabilities of feeding control system can be obtained by literature research and experts experience and listed in Table 2.

By using the BN node definition and probability reasoning method introduced in Sects. 2.1 and 2.2, the conditional probability table (CPT) of non-leaf nodes of BN in Fig. 8 can be gotten. Table 3 is the CPT of non-leaf nodes XF, ZF, U_1F and U_2F. Then the system BN model under evidence theory can be shown as Fig. 9, and the CPT of leaf node T is shown in Table 4. By using the multi-state BN reasoning method under Evidence theory in Sect. 2, the belief probabilities and plausibility probabilities of non-leaf nodes XF, ZF, U_1F and U_2F can be obtained and listed in Table 5.

The belief and plausibility probabilities of leaf node T can be calculated by (13) and (14), and

Table 2 The state probabilities of components at $t = 3000$ h with CCF

Component	State			
	0	1	2	[0,1,2]
MO	0.9993	–	0.0007	–
EW	0.9979	–	0.0021	–
GR	0.9925	–	0.0075	–
MT	0.9304	0.0089	0.0520	0.0087
SF	0.9982	–	0.0018	–
RE	0.9929	–	0.0071	–

Table 3 The CPT of non-leaf nodes under evidence theory

X^{EW} Z^{EW} U^{RE}	X^{GR} Z^{GR} U^{GR}	X^{MO} Z^{MO} U^{MO}	X^{SF} Z^{SF} U^{SF}	X^{MT} Z^{MT} U^{MT}	(OR) XF, ZF, U_1F, U_2F					
					Bel			*Pl*		
					0	1	2	0	1	2
0	0	0	0	0	1	0	0	1	0	0
0	0	0	0	1	0	1	0	0	1	0
0	0	0	0	2	0	0	1	0	0	1
0	0	0	0	[0,1,2]	0	0	0	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2	2	2	2	2	0	0	1	0	0	1
2	2	2	2	[0,1,2]	0	0	1	0	0	1

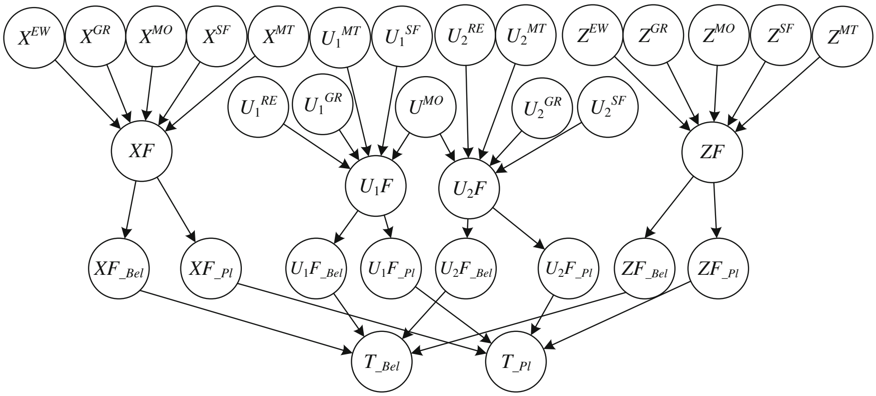


Fig. 9 System BN model under evidence theory

$$
\begin{aligned}
Bel(T=0) = P(T\_Bel=0) &= \sum_{XF,\,ZF,\,U_1F,\,U_2F} Bel(XF, ZF, U_1F, U_2F, T=0) \\
&= \sum_{XF,\,ZF,\,U_1F,\,U_2F} Bel(T=0|XF, ZF, U_1F, U_2F) \prod_{i=1}^{n} Bel\left(x_i^{k_i}\right) \\
&= \sum_{XF,\,ZF,\,U_1F,\,U_2F} Bel(T=0|XF, ZF, U_1F, U_2F) Bel(XF) \\
&\quad \cdot Bel(ZF) Bel(U_1F) Bel(U_2F)
\end{aligned}
\tag{34}
$$

Table 4 The CPT of leaf node T under evidence theory

XF	ZF	U_1F	U_2F	(OR) T					
				$T_{\_Bel}$			$T_{\_Pl}$		
				0	1	2	0	1	2
0	0	0	0	1	0	0	1	0	0
0	0	0	1	0	1	0	0	1	0
0	0	0	2	0	0	1	0	0	1
0	0	1	0	0	1	0	0	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2	2	2	1	0	0	1	0	0	1
2	2	2	2	0	0	1	0	0	1

Table 5 The state belief and plausibility probabilities of non-leaf nodes of BN

Node	State					
	Bel			Pl		
	0	1	2	0	1	2
XF	0.919180	0.008793	0.063432	0.927775	0.017388	0.072027
ZF	0.919180	0.008793	0.063432	0.927775	0.017388	0.072027
U_1F	0.914575	0.008749	0.068125	0.923127	0.017301	0.076677
U_2F	0.914575	0.008749	0.068125	0.923127	0.017301	0.076677

$$\begin{aligned} Pl(T=0) = P(T\_Pl=0) &= \sum_{XF, ZF, U_1F, U_2F} Pl(XF, ZF, U_1F, U_2F, T=0) \\ &= \sum_{XF, ZF, U_1F, U_2F} Pl(T=0|XF, ZF, U_1F, U_2F) \prod_{i=1}^{n} Pl\left(x_i^{k_i}\right) \\ &= \sum_{XF, ZF, U_1F, U_2F} Pl(T=0|XF, ZF, U_1F, U_2F) \\ &\quad \cdot Pl(XF)Pl(ZF)Pl(U_1F)Pl(U_2F) \end{aligned} \tag{35}$$

Then the state belief probabilities and plausibility probabilities of leaf node T under epistemic uncertainty can be calculated, and the results of system state probabilities when considering the influence of CCFs and without CCFs are listed in Table 6. In order to illustrate the influence of epistemic uncertainty to system, the uncertain state of component MT is classified as perfect work state 0. Then the state probabilities of system at t = 3000 h are calculated and listed in Table 6.

Based on the previous assumption, the lifetime of components obey exponential distribution, and the derating work state is regarded as perfect working state. From the belief and plausibility probability of feeding control system at state 2 in Table 6, it has shown that the failure probability interval and failure rate interval of system at t = 3000 h is [0.232005, 0.280306] and [8.7991×10^{-5}, 1.0964×10^{-4}]/h respectively when consider the influence of epistemic uncertainty and CCFGs. When the

Table 6 The state probabilities of leaf node *T*

Leaf node *T*		Considering CCFGs		
State		0	1	2
Epistemic uncertainty	Belief prob.	0.706706	0.027431	0.232005
	Plausibility Prob.	0.733514	0.056553	0.280306
Ignore uncertainty	State prob.	0.733514	0.028204	0.238282
Leaf node *T*		Without considering CCFGs		
Epistemic uncertainty	Belief prob.	0.808964	0.030368	0.119782
	Plausibility prob.	0.838640	0.062523	0.161703
Ignore uncertainty	State prob.	0.838640	0.031195	0.122977

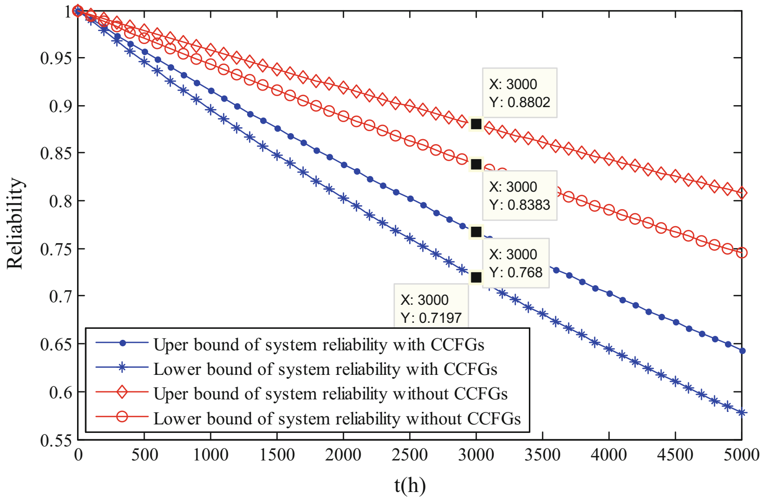


Fig. 10 The contrast curves of the influence of epistemic uncertainty and CCFGs to system reliability

CCFGs are ignored, the system failure probability interval will be [0.119782, 0.161703], and failure rate interval is $[4.2529 \times 10^{-5}, 5.8794 \times 10^{-5}]$/h. The contrast curves of system reliability with consideration of CCF are also obtained and shown in Fig. 10. From Fig. 10 we know that when the influence of uncertainty is ignored, the failure probability and failure rate of system are 0.238282 and 9.072629×10^{-5}/h. And when the CCF and uncertainty are both ignored, the corresponding failure probability and failure rate of feeding control system are 0.122977 and 4.374068×10^{-5}/h. Finally, the contrast curves of system reliability with epistemic uncertainty are shown in Fig. 11.

Based on the system function analysis and failure mechanism analysis, this section built an fault tree model of the feeding control system of a DL series horizontal lathe. The evidence theory is introduced to quantify the epistemic

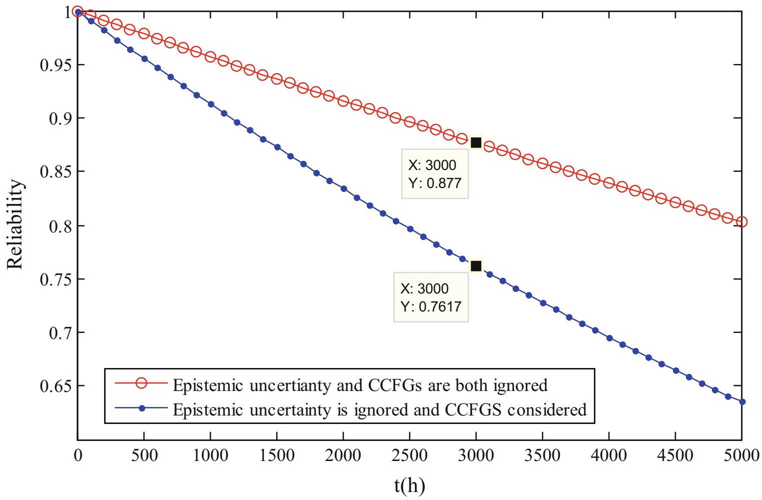


Fig. 11 The contrast curves of the influence of CCFGs to system reliability without considering epistemic uncertainty

uncertainty caused by lack of data and information in this system, and BN model is combined to realize the system reliability indexes calculation. A modified β factor model is used to model the CCFGs existed in system. From Table 6 and Fig. 10, when the influence of epistemic uncertainty to system is considered, system reliability interval at t = 3000 h will be [0.808964, 0.838640] without consider CCFs, and when the influence of CCFs is also considered, the reliability interval will be [0.706706, 0.733514]. This shows that CCFs has evident effect on system reliability. The system state probabilities in Table 6 when the epistemic uncertainty is ignored are between the corresponding belief probabilities and plausibility probabilities, which verify the accuracy of results. This chapter provides an effective method for reliability analysis of complex system under epistemic uncertainty and CCFGs.

5 Conclusions

This chapter introduces a reliability analysis method for complex MSS with epistemic uncertainty based on BN and evidence theory. The epistemic uncertainty of system is quantified through adding an uncertain state of root nodes in multi-state BN, and then the state space is constructed. The belief function and plausibility function are defined under evidence theory. Based on the BN forward reasoning, the system reliability and failure probability can be computed. The case study has confirmed the feasibility of this comprehensive method, and realized a quantitative analysis of system failure state. The backward reasoning can get the posterior probability of failure causes based on the system failure state, and provide guidance for prediction of system failure types.

CCF is an important failure mode in complex systems, so the reliability analysis of MSS with consideration of both epistemic uncertainty and CCF are also studied in this chapter. When CCFGs exist in system, a modified β factor model is introduced and integrated with evidence theory based BN, and realize the state expression and probability reasoning for complex system with epistemic uncertainty and CCFGs. The reliability analysis of the feeding control system of DL series HDHLs by this method has shown that, the proposed comprehensive method has high computing efficiency and strong practical value.

Acknowledgements This research was partially supported by the National Science and Technology Major Project of China under the contract number 2013ZX04013-011, and the Open Project of Traction Power State Key Laboratory of Southwest Jiaotong University under the contract number TPL 1410.

References

1. Gu YK, Li J (2012) Multi-state system reliability: a new and systematic review. Proc Eng 29:531–536
2. Kančev D, Čepin M (2012) A new method for explicit modelling of single failure event within different common cause failure groups. Reliab Eng Syst Saf 103:84–93
3. Kohlas J, Monney PA (2013) A mathematical theory of hints: an approach to the Dempster-Shafer theory of evidence. Springer Science & Business Media
4. Levitin G (2005) The universal generating function in reliability analysis and optimization. Springer, Berlin
5. Li YF, Zio E (2012) A multi-state model for the reliability assessment of a distributed generation system via universal generating function. Reliab Eng Syst Saf 106:28–36
6. Lisnianski A, Elmakias D, Laredo D, Ben Haim H (2012) A multi-state Markov model for a short-term reliability analysis of a power generating unit. Reliab Eng Syst Saf 98(1):1–6
7. Liu YW, Kapur KC (2006) Reliability measures for dynamic multi-state nonrepairable systems and their applications to system performance evaluation. IIE Trans 38(6):511–520
8. Lorini E, Prade H (2012) Strong possibility and weak necessity as a basis for a logic of desires. In: Working chapters of the ECAI workshop on weighted logics for artificial intelligence, Montpellier, France, pp 99–103
9. Massim Y, Zeblah A, Benguediab M, Ghouraf A, Meziane R (2006) Reliability evaluation of electrical power systems including multi-state considerations. Electr Eng 88(2):109–116
10. Mehl CH (2013) P-Boxes for cost uncertainty analysis. Mech Syst Signal Process 37(1–2): 253–263
11. Mi J, Li YF, Huang HZ, Liu Y, Zhang X (2013) Reliability analysis of multi-state systems with common cause failure based on Bayesian networks. Eksploatacja i Niezawodnosc—Maint Reliab 15(2):169–175
12. Mi J, Li YF, Peng W, Yang Y, Huang HZ (2016) Fault tree analysis of feeding control system for computer numerical control heavy-duty horizontal lathes with multiple common cause failure groups. J Shanghai Jiaotong Univ (Science) 21(4):504–508
13. Mula J, Poler R, Garcia-Sabater JP (2007) Material requirement planning with fuzzy constraints and fuzzy coefficients. Fuzzy Set Syst 158(7):783–793

14. Ramirez-Marquez JE, Coit DV (2005) Composite importance measures for multi-state systems with multi-state components. IEEE Trans Reliab 54(3):517–529
15. Rausand M (2011) Common-Cause Failures. Risk assessment. Wiley, Hoboken, NJ, pp 469–495
16. Sallak M, Schön W, Aguirre F (2013) Reliability assessment for multi-state systems under uncertainties based on the Dempster-Shafer theory. IIE Trans 45(9):995–1007
17. Sankararaman S, Mahadevan S (2011) Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. Reliab Eng Syst Saf 96(7):814–824
18. Shah H, Hosder S, Winter T (2015) Quantification of margins and mixed uncertainties using evidence theory and stochastic expansions. Reliab Eng Syst Saf 138:59–72
19. Simon C, Weber P (2009) Evidential networks for reliability analysis and performance evaluation of systems with imprecise knowledge. IEEE Trans Reliab 58(1):69–87
20. Simon C, Weber P, Levrat E (2007) Bayesian networks and evidence theory to model complex systems reliability. J Comput 2(1):33–43
21. Simon C, Weber P, Evsukoff A (2008) Bayesian networks inference algorithm to implement Dempster Shafer theory in reliability analysis. Reliab Eng Syst Saf 93(7):950–963
22. Soundappan P, Nikolaidis E, Haftka RT, Grandhi R, Canfield R (2004) Comparison of evidence theory and Bayesian theory for uncertainty modeling. Reliab Eng Syst Saf 85(1): 295–311
23. Troffaes MCM, Walter G, Kelly D (2014) A robust Bayesian approach to modeling epistemic uncertainty in common-cause failure models. Reliab Eng Syst Saf 125:13–21
24. Xue J (1985) On multistate system analysis. IEEE Trans Reliab 34(4):329–337
25. Yang JP, Huang HZ, Liu Y, Li YF (2015) Quantification classification algorithm of multiple sources of evidence. Int J Inf Tech Decis 14(5):1017–1034
26. Yang X, Liu Y, Zhang Y, Yue Z (2015) Hybrid reliability analysis with both random and probability-box variables. Acta Mech 226(5):1341–1357
27. Zhang Z, Jiang C, Wang GG, Han X (2015) First and second order approximate reliability analysis methods using evidence theory. Reliab Eng Syst Saf 137:40–49
28. Zhao S, Wang H, Cheng D (2010) Power distribution system reliability evaluation by DS evidence inference and Bayesian network method. In: IEEE 11th international conference on probabilistic methods applied to power systems pp 654–658
29. Zhou J, Liu L, Guo J, Sun L (2013) Multisensory data fusion for water quality evaluation using Dempster-Shafer evidence theory. Int J Distrib Sens 1–6
30. Zhou Q, Zhou H, Zhou Q, Yang F, Luo L, Li T (2015) Structural damage detection based on posteriori probability support vector machine and Dempster-Shafer evidence theory. Appl Soft Comput 36:368–374
31. Zio E, Podofillini L, Levitin G (2004) Estimation of the importance measures of multi-state elements by Monte Carlo simulation. Reliab Eng Syst Saf 86(3):191–204

A D-MMAP to Model a Complex Multi-state System with Loss of Units

Juan Eloy Ruiz-Castro

Abstract A complex multi-state system subject to different types of failures and preventive maintenance, with loss of units, is modelled by considering a discrete marked Markovian arrival process. The system is composed of K units, one online and the rest in cold standby. The online unit is submitted to different types of failures and when a non-repairable failure occurs the corresponding unit is removed. Several internal degradation states are considered which are observed when a random inspection occurs. This unit is subject to internal repairable failure, external shocks and preventive maintenance. If one internal repairable failure occurs, the unit goes to the repair facility for corrective repair, if a major degradation level is observed by inspection, the unit goes to preventive maintenance and when one external shock happens, this one may produce an aggravation of the internal degradation level, cumulative external damage or external extreme failure (non-repairable failure). Preventive maintenance and corrective repair times follow different distributions. The system is modelled in transient regime and relevant performance measures are obtained. All results are expressed in algorithmic and computational form and they have been implemented computationally with MATLAB and R. A numerical example shows the versatility of the model.

Keywords Reliability • Complex multi-state systems • Phase type distribution • Marked Markovian arrival process (MMAP)

1 Introduction

The failure of a reliability system may provoke severe economic and human damage. Accordingly, system redundancy and preventive maintenance are implemented in order to enhance reliability and availability. In engineering terms,

J.E. Ruiz-Castro (✉)
Faculty of Science, Department of Statistics and Operational Research
and IEMath-GR, University of Granada, Campus Fuentenueva s/n., Granada, Spain
e-mail: jeloy@ugr.es

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_3

redundancy in a system means duplicating its critical components or functions in order to increase reliability. Redundant systems arouse considerable interest in reliability literature. In this respect, Huang and Xu [6] presented a general closed-form expression for the lifetime reliability of load-sharing k-out-of-n:G hybrid redundant systems in which components are initially configured as active units. Recently, a warm standby system with multiple types of failures was developed by Ruiz-Castro [18].

Preventive maintenance refers to regular, routine maintenance to help keep equipment up and running, thus avoiding the financial and time costs involved in corrective maintenance. Designing a maintenance system and its corresponding logistic support is a very complex process in which the aim is to find the optimal level of maintenance, effective relations among different maintenance procedures and appropriate implementation. The question of determining an effective preventive maintenance policy for a specific reliability system is one that has aroused much research interest. In this respect, Nakagawa [11] conducted a detailed study of maintenance policies for reliability systems, emphasising mathematical formulation and optimisation techniques. New approaches to the reliability modeling of systems with preventive maintenance subject to shared loads are given by Liu et al. [9].

In recent years, the classical binary-state reliability theory (perfect functionality and complete failure) has been greatly extended by multi-state systems. A multi-state system is one containing components with different performance levels and several failure modes, each with different effects on system performance. Lisnianki el al. [8] performed a comprehensive analysis of multi-state systems, and Eryilmaz [4] has studied the concepts of mean residual life and mean past lifetime for multi-state systems. Measures for single-unit multi-state systems and multi-state k-out-of-n: G systems under the assumption that the degradation in systems and components follows an acyclic Markov process which has a discrete state space have also been evaluated. Mi et al. [10] extended the universal generating function based on belief function theory to conduct a reliability analysis of multi-state systems with epistemic uncertainty.

Multi-state complex systems can be analysed in different ways in reliability modelling. Typical failure time distributions such as Weibull and Gamma are commonly found in reliability literature, but in order to facilitate the analytical approach, phase-type (PH) distributions play an important role. This class of distribution was introduced by Neuts [12, 14] and has been applied in fields such as reliability and queuing theory. Markov and semi-Markov processes have also been used to study the behaviour over time of reliability systems with preventive maintenance [3, 17, 20]. Ruiz-Castro [19] considered Markov counting and reward processes in a computational form to analyse the performance and profitability of a complex system with and without preventive maintenance.

Many reliability systems have inputs to the system over time, such as a repairable failure, a non-repairable failure, preventive maintenance or an external shock. When a multi-state system is considered, the number of events over time can be modelled through a Markovian arrival process (MAP) which is a generalisation of the Poisson process. The set of Markovian arrival processes is versatile.

Any stochastic counting process can be closely approximated by a sequence of MAPs. The MAPs were introduced by Neuts [13]. A well-structured introduction of MAPs, comparing them with the Poisson process, is given in [5]. BMAPs and MMAPs are extensions of MAPs when arrivals occur in batch and for marked arrivals, respectively. MAPs have been widely used in several fields such as reliability or queuing theory [1, 2, 7, 15, 18].

Reliability systems are usually studied in the continuous case; nevertheless, not all systems can be continuously monitored, and they must be observed at certain epochs, due to the internal structure of the system, the need for periodic inspections, etc. A system can also be subject to periodic inspections, with the state of the system being identified at discrete times. Reliability systems that evolve in discrete time have been modelled by considering a Markovian structure. Thus, Ruiz-Castro [16] modelled a complex reliability system using a MAP in discrete time taking into account several types of failure and preventive maintenance to determine whether preventive maintenance was economic and profitable from a performance standpoint.

The aim of this paper is to model a complex system, subject to several types of failure that evolves in discrete time through a Marked Markovian Arrival Process. The system may undergo repairable and/or non-repairable failures as a consequence of internal wear or external shock. In reliability modelling, it is usually considered that when a unit undergoes a non-repairable failure, it will be replaced by a new one. In this paper, we assume a redundant system with a general number of units, in which a unit is removed without replacement when a non-repairable failure occurs. Preventive maintenance is introduced, in conjunction with random inspection. If major internal or external damage is observed on inspection, the unit is sent to the repair facility for preventive maintenance. The system is modelled in an algorithmic and computational form and it has been implemented computationally with MATLAB.

This paper is organised as follows. The system and its modelling are described in Sect. 2. The MMAP that governs the system is given in Sect. 3. The Sect. 4 is focused on expressing the transient distribution in an algorithmic form. Measures are given in Sect. 5 and a numerical example shows the versatility of the model in Sect. 6. Finally, conclusions are given in Sect. 7.

2 The System and the Model

We assume a cold standby system, composed of a general number of units, K. At any time, the system is operational if at least one unit is operational, and the online unit is subject to different events. Several internal operational stages are considered with respect to possible internal operational failure. These stages are partitioned into two well-differentiated groups: minor and major, corresponding to a low and high risk of failure, respectively. The failure may be repairable or non-repairable. There is one repair facility, staffed by one repairperson. The online unit is exposed to external shocks. If a shock occurs, the online unit may undergo one of three consequences: internal stage aggravation, cumulative external damage or extreme external failure. The cumulative external damage increases each time an external shock happens, and when it reaches a given threshold, the unit fails. In the same way as in the internal case, the cumulative external damage is partitioned into minor and major external stages. If the online unit fails as a consequence of an external shock, a non-repairable failure is always produced. After a non-repairable failure, the unit is removed and the system continues working with one less unit. If only one unit remains and a non-repairable failure occurs, then the system is reinitialised completely. To avoid major damage, preventive maintenance is introduced, in conjunction with random inspection, of which periodic inspection is a particular case. During inspection, the internal and the cumulative external stages are observed. If a major failure is observed, the unit is sent to the repair facility for preventive maintenance. One of three types of preventive maintenance will be performed, according to whether only internal major damage, only external cumulative major damage or both are observed. The system is based on the following assumptions.

Assumption 1 The internal operational time of the online unit is *PH*-distributed with representation $(\boldsymbol{\alpha}, \mathbf{T})$. The number of operational states is equal to n, and these are partitioned in minors (the first n_1 states) and majors states (states $n_1 + 1, \ldots, n$).

Assumption 2 When one internal failure occurs, it can be repairable or non-repairable. The probability of undergoing a repairable or non-repairable failure from a transient state is given by the column vectors $\mathbf{T}_r^0$ y $\mathbf{T}_{nr}^0$, respectively.

Assumption 3 Events that produce failures of the online unit due to external shocks occur according to a phase type renewal process. If the online place is busy, the unit undergoes the effect of this shock. The time between two consecutive events is *PH* distributed with representation $(\boldsymbol{\gamma}, \mathbf{L})$. The order of the matrix L is equal to t.

Assumption 4 An external shock may produce one of three different effects on the online unit: external cumulative damage, aggravation of internal degradation or extreme failure (non-repairable failure).

Assumption 5 External damage can pass through an indeterminate number of external degradation states. The number of external degradation states is equal to d, and these are partitioned in minors (the first d_1 states) and major states (states $d_1 + 1, \ldots, d$). If the external degradation states is i, then the external shock changes to state j with probability d_{ij}. These probabilities are contained in the matrix $\mathbf{D}$. A cumulative external damage threshold is reached from the external damage states after an external shock through the probability column vector $\mathbf{D}^0$. If it occurs then the unit undergoes a non-repairable failure. Initially, previously to an external shock, the unit is in external degradation state 1 (no damage due to external shock). The initial distribution for external damage when one unit occupies the online place initially is $\boldsymbol{\omega} = (1, 0, \ldots, 0)_{1\mathrm{x}d}$.

Assumption 6 An external shock may directly produce extreme failure (non-repairable failure). It occurs with a probability equal to ω^0.

Assumption 7 One external shock can produce modification of the internal degradation state. If the internal degradation state is i, then the external shock changes this one to state j with probability w_{ij}. These probabilities are including in the matrix $\mathbf{W}$. An internal repairable failure can occur due to this fact from any performance state with a probability column vector $\mathbf{W}^0$.

Assumption 8 The corrective repair time when the online unit fails is *PH* distributed with representation ($\boldsymbol{\beta}^{\mathbf{1}}$, $\mathbf{S_1}$). The order of this matrix is equal to z_1.

Assumption 9 When the online unit undergoes a non-repairable failure then it is removed.

Assumption 10 While the online place is busy by a unit, random inspections can occur. The time between two consecutive inspections is *PH* distributed with representation ($\boldsymbol{\eta}$, $\mathbf{M}$). The order of the matrix $\mathbf{M}$ is equal to ε.

Assumption 11 If inspection observes only major internal failure then the preventive maintenance time is *PH* distributed with representation ($\boldsymbol{\beta}^{\mathbf{2}}$, $\mathbf{S_2}$). The order of this matrix is equal to z_2.

Assumption 12 If inspection observes only major external cumulative failure then the preventive maintenance time is *PH* distributed with representation ($\boldsymbol{\beta}^{\mathbf{3}}$, $\mathbf{S_3}$). The order of this matrix is equal to z_3.

Assumption 13 If inspection observes both major internal and external failure then the preventive maintenance time is *PH* distributed with representation ($\boldsymbol{\beta}^{\mathbf{4}}$, $\mathbf{S_4}$). The order of this matrix is equal to z_2.

Assumption 14 When the system is composed of only one unit and this unit undergoes a non-repairable failure, then the system is replaced by a new, identical one with K units.

2.1 The State Space

The system described above is governed by a vector Markov process. The state space E is composed of the macro-states $E = \left\{E^K, E^{K-1}, \ldots, E^1\right\}$, where E^k contains the phases when there are k units in the system for $k = 1, \ldots, K$. At the same time, these macro-states are composed by new macro-states according to the number of units in the repair facility. Then, $E^k = \left\{E^{k,0}, E^{k,1}, \ldots, E^{k,k-1}, E^{k,k}\right\}$ where $E^{k,r}$ contains the phases when the system has k operational units and r of them are in the repair facility for $r = 0, \ldots, k$. It is important to consider the order of the units in the repair facility, according to arrival, because the type of repair needed (corrective or preventive) determines the repair time. Then, the macro-state $E^{k,r}$, for $r = 1, \ldots, k$, is composed of the following macro-states $E^{k,r} = \left\{ E^{k}_{i_1, i_2, \ldots, i_r};\ i_l = 1, 2, 3, 4,\ l = 1, \ldots, r \right\}$. The macro-state $E^{k}_{i_1, i_2, \ldots, i_r}$ contains the phases when there are k operational units, of which r are in the repair facility and the order of repair of these units are given by $i_1, \ldots, i_r$, where 1 indicates that the unit underwent a repairable failure, 2 inspection has observed only major internal failure, 3 inspection has observed only major external cumulative failure and 4 when inspection has observed major internal and external cumulative failure.

Finally, the phases of these latest macro-states are given by

$$E^{k,0} = \{(i, j, u, m); 1 \le i \le n, 1 \le j \le t, 1 \le u \le d, 1 \le m \le \varepsilon\}, \text{ for } k = 1, \ldots, K,$$

$$E^{k}_{i_1, i_2, \ldots, i_r} = \{(i, j, u, m, a); 1 \le i \le n, 1 \le j \le t, 1 \le u \le d, 1 \le m \le \varepsilon, 1 \le a \le z_{i_1}\},$$
for $k = 1, \ldots, K$ and $r = 1, \ldots, k - 1$,

$$E^{k}_{i_1, i_2, \ldots, i_k} = \{(j, a); 1 \le j \le t, 1 \le a \le z_{i_1}\}, \text{ for } k = 1, \ldots, K,$$

where i denotes the phase of the internal operational time, j the phase of the shock external time, u the cumulative external damage, m the phase of the inspection time and a the phase of the repair time. This state space has been built in this way to minimize the computational cost of the modeling process.

2.2 Analyzing Events

When a complex system is submitted to several types of events is important to analyze the behavior of these ones to avoid or delaying economical or catastrophic failures. The online unit is subject to several types of events, which cause failures. The system is modeled in a well structured form to analyze these events on the online unit.

Six different events over the online unit are considered; a repairable internal failure of the online unit (A), inspection with response major revision for only internal damage ($B1$), major revision for only cumulative external damage ($B2$), major revision for both internal damage and cumulative external damage ($B3$) and non-repairable failure (C). Finally, C is also taken into account when it causes a renewal of the system.

The transition probabilities associated to these events are modeled in a well structured form. Previous to show them, some auxiliary matrices are introduced.

The matrices $\mathbf{U}_l$ and $\mathbf{V}_l$, for $l = 1, 2$, are square matrices of order n and d respectively, whose element (s, t) is given by,

$$\mathbf{U}_1(s,t)=\begin{cases}1; & 1\le s=t\le n_1\\ 0; & \text{otherwise}\end{cases},\quad \mathbf{U}_2(s,t)=\begin{cases}1; & s=t>n_1\\ 0; & \text{otherwise}\end{cases},$$
$$\mathbf{V}_1(s,t)=\begin{cases}1; & 1\le s=t\le d_1\\ 0; & \text{otherwise}\end{cases},\quad \mathbf{V}_2(s,t)=\begin{cases}1; & s=t>d_1\\ 0; & \text{otherwise.}\end{cases}$$

These matrices are applied when a minor of major damage is observed by inspection ($l=1$ and $l=2$ respectively), for internal damage (matrix **U**) or cumulative external damage (matrix **V**).

Throughout the paper, given a matrix **A** we denote $\mathbf{A}^0$ to the matrix $\mathbf{A}^0=\mathbf{e}-\mathbf{A}\mathbf{e}$, where **e** is a column vector of ones with appropriate order. The vector $\mathbf{e}_a$ denotes a vector of ones with order a. A matrix of zeros with appropriate order is denoted by **0** and the function $I_{\{\}}$ is the indicatory function.

Next, the behavior of the online unit is described.

Internal Repairable failure of the online unit (A)

The online unit may undergo a repairable failure from any operational internal state due to wear or external shock. In the first case, it occurs because the repairable internal failure is produced ($\mathbf{T}_r^0$). In the second case, if an external shock occurs, but no external failure is produced ($\mathbf{D}\mathbf{e}\omega(1-\omega^0)$), one internal failure may happen because this shock modifies the internal behavior ($\mathbf{T}\mathbf{W}^0$). In any case, the inspection ($\mathbf{e}_\varepsilon\eta$) and the external damage are reinitialized for the new unit. The unit is replaced by one standby unit with initial distribution $\boldsymbol{\alpha}$.

The matrix that governs this transition is given by

$$\mathbf{H}_1=\left[\mathbf{T}_r^0\alpha\otimes\mathbf{L}\otimes\mathbf{e}\omega+\left(\mathbf{T}_r^0\alpha+\mathbf{T}\mathbf{W}^0\alpha\right)\otimes\mathbf{L}^0\boldsymbol{\gamma}\otimes\mathbf{D}\mathbf{e}\omega\left(1-\omega^0\right)\right]\otimes\mathbf{e}_\varepsilon\eta.$$

If the online unit that fails is the last operational one and no unit is available from the repair facility, then no units will occupy the online place at the next time point. Then,

$$\mathbf{H}_1'=\left[\mathbf{T}_r^0\otimes\mathbf{L}\otimes\mathbf{e}_d+\left(\mathbf{T}_r^0+\mathbf{T}\mathbf{W}^0\right)\otimes\mathbf{L}^0\boldsymbol{\gamma}\otimes\mathbf{D}\mathbf{e}\left(1-\omega^0\right)\right]\otimes\mathbf{e}_\varepsilon.$$

Inspection observes only major internal damage of the online unit ($B1$)

While the online unit is working, an inspection may take place ($\mathbf{M}^0\eta$). The inspection will only reveal major internal damage, and may take place either following ($\mathbf{U}_2\mathbf{T}\mathbf{W}\mathbf{e}\alpha\otimes\mathbf{L}^0\boldsymbol{\gamma}\otimes\mathbf{V}_1\mathbf{D}\mathbf{e}\omega(1-\omega^0)$), or in the absence ($\mathbf{U}_2\left(\mathbf{e}-\mathbf{T}^0\right)\alpha\otimes\mathbf{L}\otimes\mathbf{V}_1\mathbf{e}\omega$) of, a previous external shock. The matrix that governs it is given by

$$\mathbf{H}_2^1=\left[\mathbf{U}_2\left(\mathbf{e}-\mathbf{T}^0\right)\alpha\otimes\mathbf{L}\otimes\mathbf{V}_1\mathbf{e}\omega+\mathbf{U}_2\mathbf{T}\mathbf{W}\mathbf{e}\alpha\otimes\mathbf{L}^0\boldsymbol{\gamma}\otimes\mathbf{V}_1\mathbf{D}\mathbf{e}\omega\left(1-\omega^0\right)\right]\otimes\mathbf{M}^0\eta.$$

Preventive maintenance is performed after major inspection only if there are at least two operational units in the system.

Inspection observes only major cumulative external damage of the online unit (*B*2)

This situation is similar to the previous one but in this case the cumulative external damage observed by inspection is major. In this case,

$$\mathbf{H}_2^2 = \left[\mathbf{U}_1\left(\mathbf{e}-\mathbf{T}^0\right)\alpha \otimes \mathbf{L} \otimes \mathbf{V}_2\mathbf{e}\omega + \mathbf{U}_1\mathbf{TWe}\alpha \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{V}_2\mathbf{De}\omega\left(1-\omega^0\right)\right] \otimes \mathbf{M}^0\eta$$

Inspection observes major internal and external damage of the online unit (*B*3)

In this case, inspection observes major internal damage and major cumulative damage due to external shocks. The matrix that governs this event is given by

$$\mathbf{H}_2^3 = \left[\mathbf{U}_2\left(\mathbf{e}-\mathbf{T}^0\right)\alpha \otimes \mathbf{L} \otimes \mathbf{V}_2\mathbf{e}\omega + \mathbf{U}_2\mathbf{TWe}\alpha \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{V}_2\mathbf{De}\omega\left(1-\omega^0\right)\right] \otimes \mathbf{M}^0\eta.$$

Non-repairable failure (*C*)

While the online unit is working, an internal non-repairable failure may occur, due to wear from any operational state ($\mathbf{T}_{nr}^0$) or as a consequence of an external shock. This situation is considered when an external shock causes an extreme failure (ω^0) or when the cumulative external threshold is reached ($\mathbf{D}^0$). In any case, the operational time of the online unit, the cumulative external damage and the inspection time are all reinitialised ($\alpha, \omega, \boldsymbol{\eta}$). The matrix when there is at least one operational unit to occupy the online place is

$$\mathbf{H}_3 = \left[\mathbf{T}_{nr}^0\alpha \otimes \left[\mathbf{L} \otimes \mathbf{e}\omega + \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{De}\omega\left(1-\omega^0\right)\right] + \mathbf{e}\alpha \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \left(\mathbf{e}\omega\omega^0 + \mathbf{D}^0\omega\left(1-\omega^0\right)\right)\right] \otimes \mathbf{e}\eta.$$

If there are no operational units to occupy the online place then

$$\mathbf{H}_3' = \left[\mathbf{T}_{nr}^0 \otimes \left[\mathbf{L} \otimes \mathbf{e} + \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{De}\left(1-\omega^0\right)\right] + \mathbf{e} \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \left(\mathbf{e}\omega^0 + \mathbf{D}^0\left(1-\omega^0\right)\right)\right] \otimes \mathbf{e}.$$

No events (*O*)

Finally, there are several situations which describe transition without any of the above events taking place. These transitions occur when no failures are produced or when an inspection only reveals minor damage. In both cases, there may be an external shock without a failure being producing. The matrix is

$$\begin{aligned}\mathbf{H}_0 = &\left[\mathbf{T} \otimes \mathbf{L} \otimes \mathbf{I} + \mathbf{TW} \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{D}\left(1-\omega^0\right)\right] \otimes \mathbf{M} \\ &+ \left[\mathbf{U}_1\mathbf{T} \otimes \mathbf{L} \otimes \mathbf{V}_1 + \mathbf{U}_1\mathbf{TW} \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{V}_1\mathbf{D}\left(1-\omega^0\right)\right] \otimes \mathbf{M}^0\eta.\end{aligned}$$

To optimise the reliability of the system, preventive maintenance is only carried out when there are operational units to occupy the online place. Then, if there is no failure and an inspection takes place, revealing major damage but when there is no operational unit in standby, the unit continues working, and no event is recorded. The matrix that governs this transition is

$$\mathbf{H}_2^{'} = [\mathbf{U}_1\mathbf{T} \otimes \mathbf{L} \otimes \mathbf{V}_2\mathbf{I} + \mathbf{U}_2\mathbf{T} \otimes \mathbf{L} \otimes \mathbf{I} + \mathbf{U}_1\mathbf{TW} \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{V}_2\mathbf{D}(1-\omega^0) + \mathbf{U}_2\mathbf{TW} \otimes \mathbf{L}^0\boldsymbol{\gamma} \otimes \mathbf{D}(1-\omega^0)] \otimes \mathbf{M}^0\eta.$$

3 The Markovian Arrival Process with Marked Arrivals

The system described in Sect. 2 is governed by the Markovian Arrival Process with marked arrivals (MMAP) with representation $(\mathbf{D}^O, \mathbf{D}^A, \mathbf{D}^{B_1}, \mathbf{D}^{B_2}, \mathbf{D}^{B_3}, \mathbf{D}^C, \mathbf{D}^{FC})$, where $\mathbf{D}^Y$ denotes the matrix associated to the event *Y*, being *FC* the non-repairable failure of the online unit event by producing a replacement of the system.

The matrices $\mathbf{D}^Y$ are composed of matrix blocks according to the macro-states E^k, for $k \leq K$. Thus,

$$\mathbf{D}^Y = \begin{pmatrix} \mathbf{R}_p^{K,Y} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_p^{K-1,Y} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}_p^{2,Y} & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{R}_p^{1,Y} \end{pmatrix}, \quad \text{for } Y = O, A, B_1, B_2, B_3,$$

$$\mathbf{D}^Y = \begin{pmatrix} \mathbf{0} & \mathbf{R}_s^{K,Y} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_s^{K-1,Y} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & & \ddots & \mathbf{R}_s^{2,Y} \\ \mathbf{0} & \mathbf{0} & \cdots & & \mathbf{0} \end{pmatrix},$$

$$\text{for } Y = C \text{ and } \mathbf{D}^Y = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & & \ddots & \mathbf{0} \\ \mathbf{R}_s^{1,Y} & \mathbf{0} & \cdots & & \mathbf{0} \end{pmatrix}$$

for $Y = FC$, where $\mathbf{R}_p^{k,Y}$ is the matrix that contains the transition probabilities between the macro-states $E^k \rightarrow E^k$ (*k* units in the system and a non-repairable failure is not produced) and the event *Y* occurs. On the other hand, if a non-repairable failure occurs, then the transition probabilities between the macro-states $E^k \rightarrow E^{k-1}$ is given by the matrix blocks $\mathbf{R}_s^{k,Y}$ when the event *Y* is produced.

Next, the matrix blocks for repairable failure are shown in detail. The rest are given in the Appendix.

Matrix $\mathbf{R}_p^{k,A}$

The matrix block $\mathbf{R}_p^{k,A}$ contains the transition probabilities when the system is composed of k units, when a non-repairable failure does not occur and when a repairable failure occurs. This matrix is composed of matrix blocks according to some of the transitions between the macro-states $E^{k,r} \rightarrow E^{k,r}$ and $E^{k,\ r} \rightarrow E^{k,r+1}$. The elements of the matrix $\mathbf{R}_p^{k,A}$ for k = 1, …, K are given by

$$\mathbf{R}_p^{k,A} = \begin{pmatrix} \mathbf{0} & \mathbf{D}_{01}^{k,A} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{11}^{k,A} & \mathbf{D}_{12}^{k,A} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{22}^{k,A} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{D}_{k-2,k-2}^{k,A} & \mathbf{D}_{k-2,k-1}^{k,A} & \mathbf{0} \\ \mathbf{0} & \cdots & & & \mathbf{D}_{k-1,k-1}^{k,A} & \mathbf{D}_{k-1,k}^{k,A} \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

The matrices $\mathbf{D}_{l,h}^{k,A}$ are again formed by matrix blocks by considering the transitions between the macro-states $E_{i_1,\ldots,i_l}^{k} \rightarrow E_{j_1,\ldots,j_h}^{k}$. Finally, the matrices that contains the phases between these macro-states are denoted by $\mathbf{D}_{l,h}^{k,A}(j_1, \ldots, j_h; i_1, \ldots i_l)$. If l or h is equal to zero then only one queue appears. Only transitions different to zero are shown.

Hence, the transition matrix when the system is composed of k units, all of them in operational state, and one repairable failure occurs is given by

$$\mathbf{D}_{01}^{k,A}(1) = \left(\mathbf{H}_1 I_{\{k>1\}} + \mathbf{H}_1' I_{\{k=1\}}\right) \otimes \boldsymbol{\beta}^1,$$

where the transition probability for the online unit depends on the number of units in the system and after failure the unit begins its repairing with initial distribution $\boldsymbol{\beta}^1$.

If there are k units in the system of which l are in the repair facility, one internal repairable failure occurs, one repair does not happen, then the transition probability matrix is

$$\mathbf{D}_{l,l+1}^{k,A}(i_1, \ldots, i_l, 1; i_1, \ldots, i_l) = \left(\mathbf{H}_1 I_{\{l<k-1\}} + \mathbf{H}_1' I_{\{l=k-1\}}\right) \otimes \mathbf{S}_{i_1},$$

for $l = 1, \ldots, k-1; i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$. The unit that is being repaired keeps on repairing governed by the matrix $\mathbf{S}_{i_1}$.

If the system is composed of k units, with l of them in the repair facility, at the next time will have also l units in the repair facility if a repair occurs (a repairable failure happens).

Then,

$$\mathbf{D}_{l,l}^{k,A}(i_2, \ldots, i_l, 1; i_1, \ldots, i_l) = \mathbf{H}_1 \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2},$$

for $l = 1, \ldots, k-1; i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$. The unit that is being repaired finishes its repairing ($\mathbf{S}_{i_1}^0$) and the next unit in queuing enters to be repaired with the corresponding initial distribution ($\boldsymbol{\beta}^{i_2}$).

4 The Transient Distribution

The transition probabilities have been worked out in a computational and algorithmic form by considering matrix blocks. From the MMAP defined in the previous section, the transition probability matrix of the Markov process that governs the system is given by

$$\mathbf{P} = \mathbf{D}^O + \mathbf{D}^A + \mathbf{D}^{B1} + \mathbf{D}^{B2} + \mathbf{D}^{B3} + \mathbf{D}^C + \mathbf{D}^{FC}.$$

This transition probability matrix can be expressed as follows,

$$\mathbf{P} = \begin{pmatrix} \mathbf{R}_p^K & \mathbf{R}_s^K & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_p^{K-1} & \mathbf{R}_s^{K-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}_p^2 & \mathbf{R}_s^2 \\ \mathbf{R}_s^1 & \cdots & \cdots & \mathbf{0} & \mathbf{R}_p^1 \end{pmatrix},$$

where $\mathbf{R}_p^k = \sum_Y \mathbf{R}_p^{k,Y}$ and $\mathbf{R}_s^k = \sum_Y \mathbf{R}_s^{k,Y}$. Thus, $\mathbf{R}_p^k$ is the matrix that contains the transition probabilities between the macro-states $E^k \to E^k$ and $\mathbf{R}_s^k$ the transition probabilities between the macro-states $E^k \to E^{k-1}$. The matrix $\mathbf{R}_s^1$ contains the transition probabilities between the macro-states $E^1 \to E^K$. It occurs when there is only one unit in the system and a non-repairable failure occurs. The system is reinitialized by a new one with K units.

If the matrix $\mathbf{P}_{ij}^{(n)}$ denotes the matrix block with the transition probabilities between the macro-states $E^i \to E^j$ after n steps, then this one can be worked out in a recursive way through the matrix blocks defined above as

$$\mathbf{P}_{ij}^{(1)} = I_{\{i=j\}} \mathbf{R}_p^i + I_{\{j=i-1\}} \mathbf{R}_s^i + I_{\{i=1,j=K\}} \mathbf{R}_s^1,$$

$$\mathbf{P}_{ij}^{(n)} = \mathbf{P}_{ij}^{(n-1)} \mathbf{R}_p^j + \mathbf{P}_{i,j+1}^{(n-1)} \mathbf{R}_s^{j+1} + I_{\{j=K\}} \mathbf{P}_{i1}^{(n-1)} \mathbf{R}_s^1 + I_{\{j=1\}} \mathbf{P}_{i1}^{(n-1)} \mathbf{R}_p^1 \quad ; \quad n \geq 2.$$

Given the initial distribut5ion ϕ, partitioned according to the macro-states E^k such as $\phi = \{\phi^K, \phi^{K-1}, \ldots, \phi^1\}$, the transient distribution can be got as

$P\{$the system occupies the macro-state E^k at time $n\} = p^n_{E^k} = \sum_{i=1}^{K} \phi_i \mathbf{P}^{(n)}_{ik}$.

If initially the system is composed of K units then $p^n_{E^k} = \phi_K \mathbf{P}^{(n)}_{K,k}$, and $P\{$the system occupies the macro-state $E^{k,r}$ at time $n\} = p^n_{E^{k,r}}$ will be calculated by considering the corresponding part of the probability vector $p^n_{E^k}$.

5 Measures

Several measures associated to this system have been worked out.

5.1 Availability

The availability is the probability that at time ν the system will be operational. It is given by

$$A(\nu) = 1 - \sum_{k=1}^{K} p^{\nu}_{E^{k,k}} \cdot \mathbf{e}_{4^{k-2}t(z_1+z_2+z_3+z_4)}.$$

5.2 Reliability

Two different reliabilities are defined in this section: the probability of the system being reinitialised for the first time after time ν, and the probability of the system being non-operational the first time after time ν. Each probability is the reliability function of a phase-type distributions with representations $(\phi, \mathbf{P}_1)$ and $(\phi_*, \mathbf{P}_*)$ respectively, being $\mathbf{P}_1 = \mathbf{D}^O + \mathbf{D}^A + \mathbf{D}^{B1} + \mathbf{D}^{B2} + \mathbf{D}^{B3} + \mathbf{D}^C$ and $\mathbf{P}_* = \mathbf{D}^O_* + \mathbf{D}^A_* + \mathbf{D}^{B1}_* + \mathbf{D}^{B2}_* + \mathbf{D}^{B3}_* + \mathbf{D}^C_* + \mathbf{D}^{FC}_*$ where * indicates that the matrices contained in those are restricted to the macro-states $E^{k,r}$, for $k = 1, \ldots, K$ and $r = 0, \ldots, k-1$.

5.3 Mean Sojourn Times

The phase of the system changes while it is working. It is interesting to obtain information about the mean time elapsed in certain situations, up to a certain time.

Mean operational time up to time ν

The mean time that the system is operational up to time ν is given by

$$O^{\nu} = 1 + \nu - \sum_{k=1}^{K} \sum_{m=0}^{\nu} p^{m}_{E^{k,k}} \mathbf{e} \quad , \text{ for } k = 1, \ldots, K \text{ and } r = 0, 1, \ldots, k.$$

Mean working time up to a certain time ν

The mean time that the repairperson is occupied, according to the type of work being done, up to a certain time, is calculated. The repairperson may be working on four different types of repairs, corrective, internal preventive maintenance, external preventive maintenance and both internal and external preventive maintenance.

Mean working time on corrective repair up to a certain time ν

Up to a certain time ν, the repairperson has been working on corrective repair a mean time equal to

$$\begin{aligned}\psi^{\nu}_{corr} = & \sum_{m=0}^{\nu} p^{m}_{E^{1,1}} \mathbf{e}_{1:t \cdot z_1} + \sum_{m=0}^{\nu} \sum_{k=2}^{K} \sum_{r=1}^{k-1} p^{m}_{E^{k,r}} \mathbf{e}_{1:4^{r-1} n \cdot t \cdot d \cdot \varepsilon \cdot z_1} \\ & + \sum_{m=0}^{\nu} \sum_{k=2}^{K-1} p^{m}_{E^{k,k}} \mathbf{e}_{1:4^{k-1} t \cdot z_1} + \sum_{m=0}^{\nu} p^{m}_{E^{K,K}} \mathbf{e}_{1:4^{K-2} t \cdot z_1}.\end{aligned}$$

Mean working time on just internal preventive maintenance up to a certain time ν

$$\begin{aligned}\psi^{\nu}_{prev1} = & \sum_{m=0}^{\nu} p^{m}_{E^{1,1}} \mathbf{e}_{t \cdot z_1 + 1 : t \cdot (z_1 + z_2)} + \sum_{m=0}^{\nu} \sum_{k=2}^{K} \sum_{r=1}^{k-1} p^{m}_{E^{k,r}} \mathbf{e}_{4^{r-1} n \cdot t \cdot d \cdot \varepsilon \cdot z_1 + 1 : 4^{r-1} n \cdot t \cdot d \cdot \varepsilon \cdot (z_1 + z_2)} \\ & + \sum_{m=0}^{\nu} \sum_{k=2}^{K-1} p^{m}_{E^{k,k}} \mathbf{e}_{4^{k-1} t \cdot z_1 + 1 : 4^{k-1} t \cdot (z_1 + z_2)} + \sum_{m=0}^{\nu} p^{m}_{E^{K,K}} \mathbf{e}_{4^{K-2} t \cdot z_1 + 1 : 4^{K-2} t \cdot (z_1 + z_2)}.\end{aligned}$$

Mean working time on just external preventive maintenance up to a certain time ν

$$\begin{aligned}\psi^{\nu}_{prev2} = & \sum_{m=0}^{\nu} p^{m}_{E^{1,1}} \mathbf{e}_{t \cdot (z_1 + z_2) + 1 : t \cdot (z_1 + z_2 + z_3)} + \sum_{m=0}^{\nu} \sum_{k=2}^{K} \sum_{r=1}^{k-1} p^{m}_{E^{k,r}} \mathbf{e}_{4^{r-1} n \cdot t \cdot d \cdot \varepsilon \cdot (z_1 + z_2) + 1 : 4^{r-1} n \cdot t \cdot d \cdot \varepsilon \cdot (z_1 + z_2 + z_3)} \\ & + \sum_{m=0}^{\nu} \sum_{k=2}^{K-1} p^{m}_{E^{k,k}} \mathbf{e}_{4^{k-1} t \cdot (z_1 + z_2) + 1 : 4^{k-1} t \cdot (z_1 + z_2 + z_3)} + \sum_{m=0}^{\nu} p^{m}_{E^{K,K}} \mathbf{e}_{4^{K-2} t \cdot (z_1 + z_2) + 1 : 4^{K-2} t \cdot (z_1 + z_2 + z_3)}.\end{aligned}$$

Mean working time on internal-external preventive maintenance up to a certain time ν

$$\begin{aligned}\psi^{\nu}_{prev3} = & \sum_{m=0}^{\nu} p^{m}_{E^{1,1}} \mathbf{e}_{t\cdot(z_1+z_2+z_3)+1:t\cdot(z_1+z_2+z_3+z_4)} \\ & + \sum_{m=0}^{\nu}\sum_{k=2}^{K}\sum_{r=1}^{k-1} p^{m}_{E^{k,r}} \mathbf{e}_{4^{r-1}n\cdot t\cdot d\cdot\varepsilon\cdot(z_1+z_2+z_3)+1:4^{r-1}n\cdot t\cdot d\cdot\varepsilon\cdot(z_1+z_2+z_3+z_4)} \\ & + \sum_{m=0}^{\nu}\sum_{k=2}^{K-1} p^{m}_{E^{k,k}} \mathbf{e}_{4^{k-1}t\cdot(z_1+z_2+z_3)+1:4^{k-1}t\cdot(z_1+z_2+z_3+z_4)} \\ & + \sum_{m=0}^{\nu} p^{m}_{E^{K,K}} \mathbf{e}_{4^{K-2}t\cdot(z_1+z_2+z_3)+1:4^{K-2}t\cdot(z_1+z_2+z_3+z_4)}.\end{aligned}$$

5.4 Mean Number of Events

The system has been modeled by considering a MMAP to express the results in a well structured form and for calculating the mean number of events, described in Sect. 3, up to a certain time. This mean number of events is given by

$$\Gamma^{\nu}_{Y} = \phi \sum_{m=1}^{\nu} \mathbf{P}^{(m-1)}\mathbf{D}^{Y}\mathbf{e}, \quad Y = A, B1, B2, B3, C, FC.$$

6 A Numerical Example

This section highlights the value of preventive maintenance by comparing similar systems with and without this maintenance. We assume a system composed of three units, one online and two in cold standby. The phase type distributions embedded in the system and the mean times are given in Tables 1 and 2.

When an internal failure occurs, it may be repairable or non-repairable. The probability of either case occurring, from a transient state, is given by the column vectors $\mathbf{T}^0_r = (0,0,0.045)'$ and $\mathbf{T}^0_{nr} = (0,0,0.005)'$. We assume that one external shock can produce a fatal extreme non-repairable failure with probability 0.3 and the matrix that governs the transitions between cumulative external damage is

Table 1 Internal failure, external shock and inspection phase type distributions

Internal failure time	External shock time	Inspection time
$\boldsymbol{\alpha} = (1,0,0)$ $\mathbf{T} = \begin{pmatrix} 0.99 & 0.01 & 0 \\ 0.002 & 0.99 & 0.008 \\ 0 & 0 & 0.95 \end{pmatrix}$ Mean time: 270	$\boldsymbol{\gamma} = (1,0)$ $\mathbf{L} = \begin{pmatrix} 0.98 & 0.002 \\ 0.98 & 0.005 \end{pmatrix}$ Mean time: 55.5741	$\boldsymbol{\eta} = (1,0)$ $\mathbf{M} = \begin{pmatrix} 0.7 & 0.1 \\ 0.6 & 0.2 \end{pmatrix}$ Mean time: 5

Table 2 Corrective and preventive maintenance phase type distributions

Corrective repair time	Only internal preventive maintenance time	Only external preventive maintenance time	Both internal and external preventive maintenance time
$\boldsymbol{\beta}^1 = (1, 0)$ $\mathbf{S}_1 = \begin{pmatrix} 0.9 & 0.05 \\ 0.3 & 0.65 \end{pmatrix}$ Mean time: 20	$\boldsymbol{\beta}^2 = (1, 0)$ $\mathbf{S}_2 = \begin{pmatrix} 0.05 & 0.005 \\ 0.05 & 0.005 \end{pmatrix}$ Mean time: 1.0582	$\boldsymbol{\beta}^3 = (1, 0)$ $\mathbf{S}_3 = \begin{pmatrix} 0.005 & 0.005 \\ 0.005 & 0.005 \end{pmatrix}$ Mean time: 1.0101	$\boldsymbol{\beta}^4 = (1, 0)$ $\mathbf{S}_4 = \begin{pmatrix} 0.04 & 0.02 \\ 0.01 & 0.05 \end{pmatrix}$ Mean time: 1.0638

Table 3 Mean operational time and mean working times for both models (without preventive maintenance in parenthesis)

Time (ν)	O^ν	ψ^ν_{corr}	ψ^ν_{prev1}	ψ^ν_{prev2}	ψ^ν_{prev3}
50	50.9984 (50.9977)	0.1441 (0.3734)	0.0544	0.0888	0.0040
100	100.9630 (100.9467)	0.6228 (2.1177)	0.1555	0.2593	0.0150
200	200.5377 (200.3538)	2.0386 (7.1788)	0.3518	0.5953	0.0395
500	496.6584 (495.2728)	8.3530 (22.7648)	0.8403	1.4347	0.1011
1000	988.7855 (985.6309)	20.0059 (48.5891)	1.6094	2.7540	0.1968
∞	0.9843 (0.9808)	0.0233 (0.0517)	0.0015	0.0026	0.0002

$$\mathbf{D} = \begin{pmatrix} 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.66 \end{pmatrix}.$$

When an external shock is produced, the cumulative threshold damage is reached from any transient state according to the column vector $\mathbf{D}^0 = (0, 0, 0, 0.34)'$. In this case a non-repairable failure occurs.

Some measures, described in Sect. 5, have been applied. The mean operational time up to a certain time and the mean time that the repairperson is working in corrective and preventive maintenance up to a certain time for both models are given in Table 3. The stationary measures are given in the last row of this table. Thus, the proportional operational time is 0.9843 for the case with preventive maintenance and 0.9808 otherwise.

Table 4 Mean number of events up to a certain time (without preventive maintenance in parenthesis)

Time (ν)	Γ_A^ν	Γ_{B1}^ν	Γ_{B2}^ν	Γ_{B3}^ν	Γ_C^ν	Γ_{FC}^ν
50	0.0137 (0.0388)	0.0515	0.0880	0.0038	0.2708 (0.2791)	0.0026 (0.0026)
100	0.0421 (0.1486)	0.1471	0.2569	0.0141	0.5324 (0.5741)	0.0183 (0.0194)
200	0.1178 (0.4092)	0.3325	0.5895	0.0372	1.0058 (1.1298)	0.1051 (0.1194)
500	0.4397 (1.1877)	0.7942	1.4205	0.0950	2.2098 (2.4879)	0.6222 (0.7292)
1000	1.0224 (2.4789)	1.5210	2.7265	0.1850	4.1358 (4.6109)	1.5893 (1.8230)
∞	0.00113 (0.00256)	0.0015	0.0026	0.00018	0.00384 (0.00439)	0.00188 (0.00219)

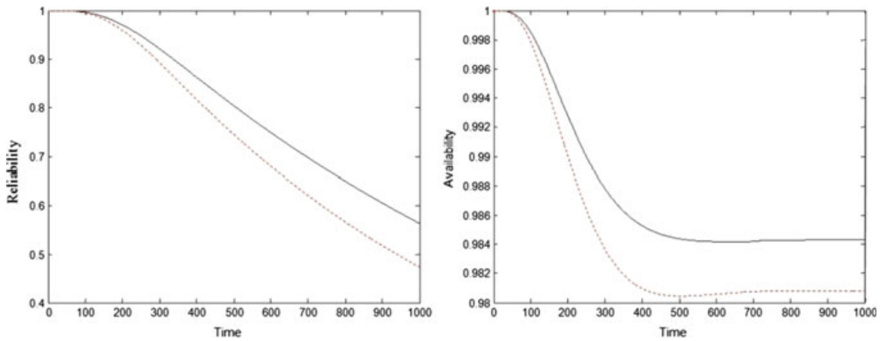


Fig. 1 Reliability and availability of the systems with (*continuous line*) and without (*dashed line*) preventive maintenance

The mean number of events has also been worked out from Sect. 5.4. The results are given in Table 4. The last row shows the stationary measures, proportion of occurrence of each event per unit of time.

The availability and reliability defined in Sect. 5.1 and Sect. 5.2, respectively, have been plotted in Fig. 1. The mean time up to first time that all units are in the repair facility is equal to 1597.8010 and 1275.1894 for the cases with and without preventive maintenance respectively.

7 Conclusions

This paper presents a complex multi-state model subject to various types of failure, to which preventive maintenance is applied, modelled using a Markovian Arrival Process with Marked arrivals in an algorithmic and computational form. This model is developed for any number of units and it is subject to possible internal failure and/or external shocks which may provoke non-repairable failure. When the latter occurs, the unit is not replaced by a new one, but is removed and the system continues working with one unit less. Random inspections are conducted and preventive maintenance is performed if major internal or external damage is observed.

The MMAP enables us to express this modelling and its associated measures in a well structured form. Furthermore, this method enables us to determine the transient distribution, the number of different events and the mean number of events. A numerical example, comparing similar systems with and without preventive maintenance, illustrates the versatility of the model.

Acknowledgements This chapter is partially supported by the Junta de Andalucía, Spain, under the grant FQM-307 and by the Ministerio de Economía y Competitividad, España, under Grant MTM2013-47929-P and by the European Regional Development Fund (ERDF).

Appendix

The MMAP that governs the system has been introduced in Sect. 3. Some matrix blocks have been described in that section, the rest are given in this Appendix.

Matrix $\mathbf{R}_p^{k,O}$

The elements of the matrix $\mathbf{R}_p^{k,O}$ for $k = 1, \ldots, K$ are given by

$$\mathbf{R}_p^{k,O} = \begin{pmatrix} \mathbf{D}_{00}^{k,O} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_{10}^{k,O} & \mathbf{D}_{11}^{k,O} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{21}^{k,O} & \mathbf{D}_{22}^{k,O} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{D}_{k-1,k-2}^{k,O} & \mathbf{D}_{k-1,k-1}^{k,O} & \mathbf{0} \\ \mathbf{0} & \cdots & & & \mathbf{D}_{k,k-1}^{k,O} & \mathbf{D}_{kk}^{k,O} \end{pmatrix},$$

where

$$\mathbf{D}_{00}^{k,O} = \mathbf{H}_0 + \mathbf{H}_2^{'} I_{\{k=1\}},$$

$\mathbf{D}_{l,l}^{k,O}(i_1, \ldots, i_l; i_1, \ldots, i_l) = \left(\mathbf{H}_0 + \mathbf{H}_2' I_{\{l=k-1\}}\right) \otimes \mathbf{S}_{i_1}, \quad$ for $\quad l = 1, \ldots, k-1;$ $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l,$

$$\mathbf{D}_{k,k}^{k,O}(i_1, \ldots, i_k, ; i_1, \ldots, i_k) = \left(\mathbf{L} + \mathbf{L}^0 \boldsymbol{\gamma}\right) \otimes \mathbf{S}_{i_1},$$

for $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, k; k \neq K, \ i_k = 1, k = K,$

$$\mathbf{D}_{10}^{k,O}(i_1) = \mathbf{H}_0 \otimes \mathbf{S}_{i_1}^0 \quad ; \quad i_1 = 1, 2, 3, 4\,; k \geq 2,$$

$$\mathbf{D}_{l,l-1}^{k,O}(i_2, \ldots, i_l; i_1, \ldots, i_l) = \mathbf{H}_0 \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2} \quad ; \quad l = 2, \ldots, k-1; \ i_s = 1, 2, 3, 4 \text{ for } s = 1, \ldots, l,$$

$$\mathbf{D}_{k,k-1}^{k,O}(i_2, \ldots, i_k; i_1, \ldots, i_k) = \boldsymbol{\alpha} \otimes \left(\mathbf{L} + \mathbf{L}^0 \boldsymbol{\gamma}\right) \otimes \boldsymbol{\omega} \otimes \boldsymbol{\eta} \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2},$$

for $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, k; k \neq K, \ i_k = 1, k = K,$

$$\mathbf{D}_{10}^{1,O}(1) = \boldsymbol{\alpha} \otimes \left(\mathbf{L} + \mathbf{L}^0 \boldsymbol{\gamma}\right) \otimes \boldsymbol{\omega} \otimes \boldsymbol{\eta} \otimes \mathbf{S}_1^0.$$

Matrix $\mathbf{R}_p^{k,B_i}$

The elements of the matrix $\mathbf{R}_p^{k,B_i}$ for $i = 1, 2, 3$ and for $k = 2, \ldots, K$ are given by

$$\mathbf{R}_p^{k,B_i} = \begin{pmatrix} \mathbf{0} & \mathbf{D}_{01}^{k,B_i} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{11}^{k,B_i} & \mathbf{D}_{12}^{k,B_i} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{22}^{k,B_i} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{D}_{k-2,k-2}^{k,B_i} & \mathbf{D}_{k-2,k-1}^{k,B_i} & \mathbf{0} \\ \mathbf{0} & \cdots & & & \mathbf{D}_{k-1,k-1}^{k,B_i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{R}_p^{1,B_i} = \mathbf{0}$$

$$\mathbf{D}_{01}^{k,B_i}(2) = \mathbf{H}_2^i \otimes \boldsymbol{\beta}^2,$$

$$\mathbf{D}_{l,l+1}^{k,B_i}(i_1, \ldots, i_l, 2; i_1, \ldots, i_l) = \mathbf{H}_2^i \otimes \mathbf{S}_{i_1} \quad ; \quad l = 1, \ldots, k-2\,; i_s = 1, 2, 3, 4 \text{ for } s = 1, \ldots, l,$$

$$\mathbf{D}_{l,l}^{k,B_i}(i_2, \ldots, i_l, 2; i_1, \ldots, i_l) = \mathbf{H}_2^i \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2} \quad ; \quad l = 1, \ldots, k-1\,; i_s = 1, 2, 3, 4 \text{ for } s = 1, \ldots, l.$$

Matrix $\mathbf{R}_p^{k,C}$

The elements of the matrix $\mathbf{R}_s^{k,C}$ for $k = 2, \ldots, K$ are given by

$$\mathbf{R}_s^{k,C} = \begin{pmatrix} \mathbf{D}_{00}^{k,C} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_{10}^{k,C} & \mathbf{D}_{11}^{k,C} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{21}^{k,C} & \mathbf{D}_{22}^{k,C} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{D}_{k-1,k-2}^{k,C} & \mathbf{D}_{k-1,k-1}^{k,C} & \mathbf{0} \\ \mathbf{0} & \cdots & & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$\mathbf{D}_{00}^{k,C} = \mathbf{H}_3 I_{\{k>1\}}$,

$$\mathbf{D}_{l,l}^{k,C}(i_1, \ldots, i_l; i_1, \ldots, i_l) = \left(I_{\{k \neq K-1\}}\mathbf{H}_3 + I_{\{k=K-1\}}\mathbf{H}_3^{'}\right) \otimes \mathbf{S}_{i_1} \quad ; \quad \text{for } l = 1, \ldots, k-1,$$

and $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$,

$$\mathbf{D}_{l,l-1}^{k,C}(i_2, \ldots, i_l; i_1, \ldots, i_l) = \mathbf{H}_3 \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2} \quad ; \quad l = 2, \ldots, k-1;\ i_s = 1, 2, 3, 4 \text{ for } s = 1, \ldots, l,$$

$$\mathbf{D}_{10}^{k,C}(i_1) = \mathbf{H}_3 \otimes \mathbf{S}_{i_1}^0 \quad ; \quad i_s = 1, 2, 3, 4 \text{ for } s = 1, \ldots, l.$$

Matrix $\mathbf{R}_s^{k,C}$

The elements of the matrix $\mathbf{R}_s^{k,C}$ for $k = 2, \ldots, K$ are given by

$$\mathbf{R}_s^{k,C} = \begin{pmatrix} \mathbf{D}_{00}^{k,C} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_{10}^{k,C} & \mathbf{D}_{11}^{k,C} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{21}^{k,C} & \mathbf{D}_{22}^{k,C} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{D}_{k-1,k-2}^{k,C} & \mathbf{D}_{k-1,k-1}^{k,C} & \mathbf{0} \\ \mathbf{0} & \cdots & & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where

$$\mathbf{D}_{00}^{k,C} = \mathbf{H}_3 I_{\{k>1\}};$$

$\mathbf{D}_{l,l}^{k,C}(i_1, \ldots, i_l; i_1, \ldots, i_l) = \left(I_{\{k \neq K-1\}}\mathbf{H}_3 + I_{\{k=K-1\}}\mathbf{H}_3^{'}\right) \otimes \mathbf{S}_{i_1}$,
for $l = 1, \ldots, k-1; i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$,
$\mathbf{D}_{l,l-1}^{k,C}(i_2, \ldots, i_l; i_1, \ldots, i_l) = \mathbf{H}_3 \otimes \mathbf{S}_{i_1}^0 \boldsymbol{\beta}^{i_2}$,
for $l = 2, \ldots, k-1$; $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$,
$\mathbf{D}_{10}^{k,C}(i_1) = \mathbf{H}_3 \otimes \mathbf{S}_{i_1}^0$, for $i_s = 1, 2, 3, 4$ for $s = 1, \ldots, l$.
Matrix $\mathbf{R}_p^{1,FC}$

The elements of the matrix $\mathbf{R}_s^{1,FC}$ are given by

$$\mathbf{R}_s^{1,FC} = \begin{pmatrix} \mathbf{D}_{00}^{1,C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \text{where } \mathbf{D}_{00}^{1,C} = \mathbf{H}_3$$

References

1. Breuer L (2002) An EM algorithm for batch Markovian arrival processes and its comparison to a simpler estimation procedure. Ann Oper Res 112:123–138
2. Chakravarthy SR (2001) The batch Markovian arrival process: a review and future work. In: Krishnamoorthy A, Raju N, Ramaswami V (eds) Advances in probability and stochastic processes. Notable Publications, pp 21–49
3. Dui H, Si S, Zuo M, Sun S (2015) Semi-Markov process-based integrated importance measure for multi-state systems. IEEE Trans Reliab 64(2):754–765
4. Eryilmaz S (2010) Mean residual and mean past lifetime of multi-state systems with identical components. IEEE Trans Reliab 59(4):644–649
5. He QM (2014) Fundamental of matrix analytic methods. Springer, New York
6. Huang L, Xu Q (2010) Lifetime reliability for load-sharing redundant systems with arbitrary failure distributions. IEEE Trans Reliab 59(2):319–330
7. Latouche G, Remiche MA, Taylor P (2003) Transient Markovian arrival processes. Ann Appl Probab 13:628–640
8. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
9. Liu B, Xie M, Kuo W (2015) Reliability modeling and preventive maintenance of load-sharing systems with degrading components. IIE Trans 48(8):699–709
10. Mi J, Li YF, Liu Y, Yang Y, Huang H (2015) Belief universal generating function analysis of multi-state systems under epistemic uncertainty and common cause failures. IEEE Trans Reliab 64(4):1300–1309
11. Nakagawa T (2005) Maintenance theory of reliability. Springer, London
12. Neuts M (1975) Probability distributions of phase type. Liber Amicorum Prof. Emeritus H. Florin. Department of Mathematics, University of Louvain, Belgium, pp 183–206
13. Neuts M (1979) A versatile Markovian point process. J Appl Probab 16:764–779
14. Neuts M (1981) Matrix geometric solutions in stochastic models. An algorithmic approach. John Hopkins, University Press, Baltimore
15. Ramírez-Cobo P, Lillo R, Wiper M (2010) Nonidentifiability of the two-state Markovian arrival process. J Appl Probab 47:630–649
16. Ruiz-Castro JE (2015) A multi-state cold standby system modelled through a marked Markovian arrival process. In: Karagrigoriou A, Oliveira T, Skiadas CH (eds) Statistical, stochastic and data analysis methods and applications. ISAST: International Society for the Advancement of Science and Technology, pp 3–18
17. Ruiz-Castro JE (2015) A preventive maintenance policy for a standby system subject to internal failures and external shocks with loss of units. Int J Syst Sci 46(9):1600–1613
18. Ruiz-Castro JE (2016) Complex multi-state systems modelled through marked Markovian arrival processes. Eur J Oper Res 252:852–865
19. Ruiz-Castro JE (2016) Markov counting and reward processes for analyzing the performance of a complex system subject to random inspections. Reliab Eng Syst Saf 145:155–168
20. Sheu S, Zhang Z (2013) An optimal age replacement policy for multi-state systems. IEEE Trans Reliab 62(3):722–735

Modeling and Inference for Multi-state Systems

Vlad Stefan Barbu and Alex Karagrigoriou

Abstract In this work we are focused on multi-state systems modeled by means of a special type of semi-Markov processes. The sojourn times are seen to be independent not necessarily identically distributed random variables and assumed to belong to a general class of distributions closed under extrema that includes, in addition to some discrete distributions, several typical reliability distributions like the exponential, Weibull, and Pareto. A special parametrization is proposed for the parameters describing the system, taking thus into account various types of dependencies of the parameters on the the states of the system. We obtain maximum likelihood estimators of the parameters and plug-in type estimators are furnished for the basic quantities describing the semi-Markov system under study.

Keywords Multi-state system · Reliability theory · Survival analysis · Semi-Markov processes · Parameter estimation · Time-varying model · Scale parameter

1 Introduction

Technical and technological systems assuming multiple possible states are known as multi-state systems (MSS). Any system that is allowed to assume a finite number of performance rates can be modeled by means of a multi-state system.

Such modeling approaches, which are more realistic and provide more accurate representations of engineering systems, are much more complex and present major difficulties in system definition and performance evaluation. MSS reliability has received a substantial amount of attention in the past four decades with basic concepts being introduced in the 70s by [1–4]. Extensions and generalizations can

V.S. Barbu
Laboratoire de Mathématiques Raphaël Salem, Université de Rouen,
Avenue de l'Université, BP.12, UMR 6085, F76801 Saint-Étienne-du-Rouvray, France
e-mail: barbu@univ-rouen.fr

A. Karagrigoriou (✉)
Department of Mathematics, University of the Aegean, Samos, Greece
e-mail: alex.karagrigoriou@aegean.gr

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_4

be found in [5–7]. Essential achievements that were attained up to the mid 1980s are reflected in [8, 9], where one can find the state of the art in the field of MSS reliability. For theoretical advances and significant applications in MSS reliability theory in recent years, the reader is referred to [10–12]. For general references on continuous-time semi Markov systems and associated reliability topics, one can see [13–17].

Consider a process defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with state space $E = \{1, 2, \ldots, N\}$. For example, state "N" is associated with nominal performance of the system and state "1" is associated with total failure. Markov processes represent typical tools for modelling such a system. In this work we focus on multi state systems that we model by means of semi-Markov processes, which generalize typical Markov jump processes by allowing general distributions for sojourn times [13]. For this reason, the semi-Markov processes are more adapted for reliability studies (and for applications in general).

The sojourn times in a given state are assumed to belong to a general class of distributions, cf. Relation (2). The interest of this distribution class is twofold. First, it is worth noticing that the class is closed under extrema (cf. [18]) and secondly it unifies under a single umbrella, not only some discrete distributions but also and more importantly, several typical reliability distributions like the exponential, Weibull, Rayleigh and Pareto distributions.

In this chapter we consider a special case of the semi-Markov system introduced in [19]. In that article, the system under study depended on some parameters a_{ij}, with i, j belonging to the state space E. In the present work, the dependence of the parameters a_{ij} on the states i and j is made explicit through a function $g(i, j)$. We also provide several examples of such a function g that could be of interest in different modeling situations, according to the application under study. It is important to indicate that time- and state-varying parameters become quite popular since more and more systems are subject to dynamic changes. Note that the problem of time- and state-varying parameters has received in recent years, increased attention (see e.g. [20–23]) because of an ever-growing body of evidence that typical assumptions of stable parameters often appear invalid.

The chapter is organized as follows. Some preliminaries regarding semi-Markov processes are presented in Sect. 2; here we also describe the special case of the semi-Markov system introduced in [19], developed in a multi-state system framework. In Sect. 3 we describe the class of distributions considered in this work and we propose a special parametrization of these distributions. In Sect. 4 the likelihood function and the associated maximum likelihood estimators of the parameters of interest are provided.

2 A Special Case of Semi-Markov Multi-state Systems

As previously mentioned, we assume that the random system has finite state space $E = \{1, \ldots, N\}$, $N < \infty$ and its time evolution is governed by a stochastic process $Z = (Z_t)_{t \in \mathbb{R}_+}$. Let us denote by $S = (S_n)_{n \in \mathbb{N}}$ the successive time points when state changes in $(Z_t)_{t \in \mathbb{R}_+}$ occur and by $J = (J_n)_{n \in \mathbb{N}}$ the successive visited states at these time points. Set also $X = (X_n)_{n \in \mathbb{N}}$ for the successive sojourn times in the visited states. Thus, $X_n = S_n - S_{n-1}$, $n \in \mathbb{N}^*$, and, by convention, we set $X_0 = S_0 = 0$.

Let us first recall the definition of a Markov renewal and semi-Markov process (cf. [13]). If $(J, S) = (J_n, S_n)_{n \in \mathbb{N}}$ satisfies the relation

$$\begin{aligned} &\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \le t | J_0, \ldots, J_n; S_1, \ldots, S_n) \\ &= \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \le t | J_n), j \in E, t \in \mathbb{R}_+, \end{aligned}$$

then

- (J, S) is called a *Markov renewal process* (MRP);
- $Z = (Z_t)_{t \in \mathbb{R}_+}$ is called a *semi-Markov process* (SMP) associated to (J, S), where

$$Z_t := J_{N(t)} \quad \Leftrightarrow \quad J_n = Z_{S_n},$$

  with

$$N(t) := \max\{n \in \mathbb{N} \mid S_n \le t\}, \ t \in \mathbb{R}_+, \tag{1}$$

  the counting process of the number of jumps in the time interval $(0, t]$. Thus, Z_t gives the state of the system at time t.

If $(J_n, S_n)_{n \in \mathbb{N}}$ is a MRP, it can be immediately checked that $(J_n)_{n \in \mathbb{N}}$ is a Markov chain, called the *embedded Markov chain*.

All along this work we assume that the SMP (or equivalently, the MRP) is regular, irreducible and positive-recurrent (see, e.g., [13, 24, 25] for more details on these notions).

A SM model is characterized by its *initial distribution* $\alpha = (\alpha_1, \ldots, \alpha_N)$

$$\alpha_j := \mathbb{P}(J_0 = j), \ j \in E,$$

and by the *semi-Markov kernel*

$$Q_{ij}(t) := \mathbb{P}(J_n = j, X_n \le t | J_{n-1} = i).$$

Let us also introduce the *transition probabilities* of the embedded Markov chain $(J_n)_{n \in \mathbb{N}}$,

$$p_{ij} := \mathbb{P}(J_n = j | J_{n-1} = i) = \lim_{t \to \infty} Q_{ij}(t),$$

and the *conditional sojourn time distribution functions*

$$\begin{aligned} W_{ij}(t) &:= \mathbb{P}(S_n - S_{n-1} \le t | J_{n-1} = i, J_n = j) \\ &= \mathbb{P}(X_n \le t | J_{n-1} = i, J_n = j). \end{aligned}$$

Observe that

$$Q_{ij}(t) = p_{ij} W_{ij}(t).$$

In the sequel, we will consider a special case of semi-Markov system introduced in [19]. As it will be seen in the next section, we will consider here a particular parametrization of this system.

Let us assume that we have at our disposal a collection of positive random variables T_{ij}, that can be seen as potential times spent in state i before moving (directly) to state j. We denote by $F_{ij}(t;\theta_{ij})$ its cumulative distribution function (cdf), where θ_{ij} is the m-dimensional parameter involved in the underlying distribution. We assume that the distribution of T_{ij} is absolutely continuous with respect to the Lebesgue measure; an associated density is denoted by $f_{ij}(t;\theta_{ij})$.

The dynamic of the system is as follows: the next state to be visited after state i is the one for which T_{il} is the minimum, $l \in E$. This is the way the next state to be visited, say j, is "chosen", namely $j = argmin_{l\in E}(T_{il})$. Thus, for our semi-Markov system, the semi-Markov kernel becomes

$$\begin{aligned}
Q_{ij}(t) &= \mathbb{P}(\min_l T_{il} \le t \ \text{ \& the min occurs for } j | J_{n-1} = i) \\
&= \mathbb{P}(\min_l T_{il} \le t, T_{ij} \le T_{il}, \forall l | J_{n-1} = i) \\
&= \mathbb{P}(\min_l T_{il} \le t | J_{n-1} = i, J_n = j) \times \mathbb{P}(T_{ij} \le T_{il}, \forall l | J_{n-1} = i) \\
&= p_{ij} W_i(t),
\end{aligned}$$

where

$$p_{ij} = \mathbb{P}(J_n = j | J_{n-1} = i) = \mathbb{P}(T_{ij} \le T_{il}, \forall l | J_{n-1} = i)$$

and

$$\begin{aligned}
W_{ij}(t) &= \mathbb{P}(S_n - S_{n-1} \le t | J_{n-1} = i, J_n = j) \\
&= \mathbb{P}(\min_l T_{il} \le t | J_{n-1} = i, J_n = j) \\
&= \mathbb{P}(\min_l T_{il} \le t | J_{n-1} = i) =: W_i(t), \ \text{ independent of } j,
\end{aligned}$$

which represents the cdf of the sojourn time in state i (unconditional to the next state to be visited). Note that

$$\sum_j Q_{ij}(t) = W_i(t).$$

Let us assume that $W_i(t)$ is absolutely continuous w.r.t. the Lebesgue measure and has a density denoted by $f_i(t)$.

As we will be dealing in the sequel with parametric inference, whenever a quantity of interest will depend on a parameter $\theta \in \Theta \subset \mathbb{R}^m$, we may set this parameter as an argument. For instance, if $Q_{ij}(t)$ depends on some parameter θ, we could denote it by $Q_{ij}(t;\theta)$.

Our intention is to provide estimators of p_{ij}, $W_i(t)$, and $Q_{ij}(t)$ under a general class of distributions, with a specific parametrization. This class of distributions and the corresponding parametrization are presented and discussed in the next section.

3 Parametric Specification of the System

The type of distributions considered for the random variables T_{ij} are first presented in this section. Then, a specific parametrization is considered. More specifically, we consider the case where the distributions $F_{ij}(\cdot;\theta_{ij})$, $i,j = 1,\dots,N$, are of the same functional form but with different parameters, i.e., we are focusing on independent but not necessarily identically distributed (inid) random variables. Nonidentically but independently distributed random variables are usually not easy to deal with. But, when these belong to families of random variables closed under maxima or minima then elegant expressions of various statistical characteristics such as order statistics are possible. A member of such a class of distribution functions with parameter a is assumed to verify the following distributional form

$$F(x;a) = 1 - (1 - F(x;1))^a \,. \tag{2}$$

Let us assume that $F(x;a)$ is absolutely continuous w.r.t. the Lebesgue measure and let us denote its density by $f(x;a)$, namely

$$f(x;a) = a\,(1 - F(x;1))^{a-1} f(x;1). \tag{3}$$

The following result states that the minimum order statistic from an *inid* random sample from the above class has a distribution belonging to the same class.

Lemma 1 (cf. [18]) *Let $X_1,\dots,X_N$ be inid random variables such that $X_i \sim F(x;a_i)$ which belongs to class (2). Then the distribution function $F^{(1)}$ of the minimum order statistic $X_{(1)}$ belongs also to (2).*

It is worth noticing that examples of distribution that belong to class (2) are the geometric distribution, the Pareto distribution, the Weibull distribution and its special cases like the exponential, the Rayleigh and the Erlang truncated exponential.

Let us now assume that the random variables T_{ij} considered in the previous section belong to the class (2), with the corresponding parameters a_{ij}, i.e., the corresponding cumulative distributions $F(t;a_{ij})$ verify (2). Moreover, we assume that we have a parametrization for a_{ij} that makes explicit the dependence on the states i and j. To be more specific, let us assume that a_{ij} has the expression

$$a_{ij} := a_{\infty}\left(1 - e^{g(i,j)/e_1}\right), \tag{4}$$

where a_{∞} and e_1 are real parameters, while $g(i,j)$ is a known function of states i and j, depending on certain parameters. Typical examples of $g(i,j)$ can be obtained by considering

$$g(i,j) := c_1 i^{k_1} j^{l_1} + c_2 i^{k_2} j^{l_2}, \tag{5}$$

where $c_m, k_m, l_m, m = 1, 2$, are real parameters. Examples of such a function g that could be of interest in different modeling situations, according to the application under study, could be:

$$g(i,j) = i + j, \tag{6}$$

$$g(i,j) = c_1 i + c_2 j, \text{with } c_1 + c_2 = 1, \tag{7}$$

$$g(i,j) = \sqrt{ij}, \tag{8}$$

$$g(i,j) = (ij)^c, c \in \mathbb{R}. \tag{9}$$

Remark 1 1. Note that this parametrization is done by analogy with a framework considered in [20], where the times between two successive failures are assumed to be inid random variables distributed according to a cumulative distribution $F(x; a_i)$ belonging to the class (2) with different scale parameters a_i. These parameters are assumed to be time varying; one type of variation along time proposed in that article is of the type $a_i = a_{\infty}\left(1 - e^{-t_i/e_1}\right), i = 1, 2, \ldots$, where $t_1, t_2, \ldots$ are observed successive failure times. Nonetheless, note that, in the present chapter, the variation is on both states i and j, while in [20] the variation is along time.

2. Note that, if we consider a semi-Markov system with only one state ($E = \{1\}$), we are in the framework of [20], where $S_n, n = 1, 2, \ldots$, are the successive failure times of a system, $S_n < S_{n+1}$, and $S_0 := 0$, while $X_n := S_n - S_{n-1}, n = 1, 2, \ldots$, are the times between two successive failures. It is clear that, in this case, there is no state variation anymore and a modeling like the one proposed in [20] would be appropriate.

Under these conditions, the following result concerning the main semi-Markov characteristics can be proved. For notational convenience, we set $F(t) := F(t; 1)$, $f(t) := f(t; 1)$ and $Q_{ij}\left(t; a_{ik}; k = 1, \ldots, N\right) := Q_{ij}(t)$.

Proposition 1 *(cf. [19]) Under the setup of this section, the following results hold:*

$$Q_{ij}(t) = \frac{a_{ij}}{\sum\limits_{k \in E} a_{ik}}\left[1 - (1 - F(t))^{\sum\limits_{k \in E} a_{ik}}\right], \tag{10}$$

$$p_{ij} = \frac{a_{ij}}{\sum\limits_{k \in E} a_{ik}}, \tag{11}$$

$$W_i(t) = 1 - [1 - F(t)]^{\sum_{j=1}^{N} a_{ij}} \tag{12}$$

and

$$f_i(t) = \sum_{j=1}^{N} a_{ij} (1 - F(t))^{\sum_{j=1}^{N} a_{ij}} \frac{f(t)}{1 - F(t)}. \tag{13}$$

4 Maximum Likelihood Estimation

In this section we consider the problem of obtaining the maximum likelihood estimators of the parameters of the system (a_∞ and e_1, or, equivalently, a_{ij}). Then we will get the corresponding plug-in estimators of the main quantities defining the semi-Markov system.

Basically, two important statistical settings could be considered: either we start with one sample path, or with several sample paths. In both cases, it can be assumed that the sample paths are complete or that the sojourn time in the last visited state can be right censored (lost to follow-up, for instance). In the sequel we consider the most general case, that is the one of several sample paths with possible censored last sojourn time. The other cases can be obtained from the one we present, as a particular case; we will also give some details on this point.

Given L sample paths of a semi-Markov process censored at time M, $\left\{ j_0^{(l)}, x_1^{(l)}, j_1^{(l)}, x_2^{(l)}, \dots, j_{N^l(M)}^{(l)}, u_M^{(l)} \right\}$, $l = 1, \dots, L$, then the associated likelihood is

$$\mathscr{L} = \left(\prod_{i \in E} \alpha_i^{N_{i,0}(L)} \right) \left(\prod_{i,j \in E} p_{ij}^{\sum_{l=1}^{L} N_{ij}^{(l)}(M)} \right) \times \\ \times \left(\prod_{l=1}^{L} \prod_{i \in E} \prod_{k=1}^{N_i^{(l)}(M)} f_i(x_i^{(l,k)}) \right) \prod_{i \in E} \prod_{k=1}^{N_{i,M}(L)} \left(1 - W_i(u_i^{(k)}) \right), \tag{14}$$

where we set

- $N_{i,0}^{(L)} := \sum_{l=1}^{L} \mathbb{1}_{\{J_0^{(l)} = i\}}$: the number of sample paths starting in state i;
- $N_i^{(l)}(M)$: the number of visits to state i up to time M of the lth trajectory, $l = 1, \dots, L$;
- $N_i(L, M) := \sum_{l=1}^{L} N_i^{(l)}(M)$: the total number of visits to state i up to time M along the L trajectories;
- $N_{ij}^{(l)}(M)$: the number of transitions from state i to state j up to time M during the lth trajectory, $l = 1, \dots, L$;

- $N_{ij}(L,M) := \sum_{l=1}^{L} N_{ij}^{(l)}(M)$: the total number of transitions from state i to state j up to time M along the L trajectories;
- $x_i^{(l,k)}$: the sojourn time in state i during the kth visit, $k = 1, \ldots, N_i^{(l)}(M)$ of the lth trajectory, $l = 1, \ldots, L$;
- $u_M^{(l)} := M - S_{N^l(M)}$ is the observed censored time of the lth trajectory;
- $N_{i,M}(L) = \sum_{l=1}^{L} \mathbb{1}_{\{J^{(l)}_{N^l(M)}=i\}}$ is the number of visits of state i, as last visited state, over the L trajectories; note that $\sum_{i \in E} N_{i,M}(L) = L$;
- $u_i^{(k)}$ is the observed censored sojourn time in state i during the kth visit, $k = 1, \ldots, N_{i,M}(L)$.

Note that, for $L = 1$, the likelihood given in (14) reduces to the likelihood of 1 trajectory. Note also that, if the censoring time M in a certain trajectory l is a jump time, then for the corresponding observed censored time we have $u_M^{(l)} = 0$. Consequently, the contribution to the likelihood of the associated term will be equal to 1. For this reason, if no censoring is involved, the uncensored likelihood can be obtained as a particular case of (14).

For the class of distributions given in (2), the likelihood takes the form

$$\begin{aligned}\mathscr{L} &= \left(\prod_{i\in E} \alpha_i^{N_{i,0}^{(L)}}\right)\left(\prod_{l=1}^{L}\prod_{i,j\in E} a_{ij}^{N_{ij}^{(l)}(M)}\right)\times \\ &\times \prod_{l,i,k}\left[\left(1-F\left(x_i^{(l,k)}\right)\right)^{\sum_{j\in E} a_{ij}}\left(\frac{f\left(x_i^{(l,k)}\right)}{1-F\left(x_i^{(l,k)}\right)}\right)\right]\times \\ &\times\left(\prod_{i\in E}\prod_{k=1}^{N_{i,M}(L)}\left(1-F\left(u_i^{(k)}\right)\right)^{\sum_{j\in E} a_{ij}}\right), \end{aligned} \tag{15}$$

where a_{ij} has been given in (4). Consequently, the log-likelihood has the expression

$$\begin{aligned}\log(\mathscr{L}) &= \log\left(\prod_{i\in E} \alpha_i^{N_{i,0}^{(L)}}\right) + \sum_{l=1}^{L}\sum_{i,j\in E} N_{ij}^{(l)}(M)\log(a_{ij}) \\ &+ \sum_{l,i,k}\left(\sum_{j\in E} a_{ij}\right)\log\left(1-F\left(x_i^{(l,k)}\right)\right) + \log\prod_{l,i,k}\left(\frac{f\left(x_i^{(l,k)}\right)}{1-F\left(x_i^{(l,k)}\right)}\right) \\ &+ \sum_{i\in E}\sum_{k=1}^{N_{i,M}(L)}\left(\sum_{j\in E} a_{ij}\right)\log\left(1-F\left(u_i^{(k)}\right)\right). \end{aligned} \tag{16}$$

Using $a_{ij} = a_\infty \left(1 - e^{g(i,j)/e_1}\right)$, taking the derivatives of $\log(\mathscr{L})$ with respect to a_∞ and e_1 we obtain the critical equations:

$$\frac{\partial \log \mathscr{L}}{\partial a_\infty} = \sum_{l=1}^{L} \sum_{i,j\in E} N_{ij}^{(l)}(M) \frac{1}{a_\infty} + \sum_{i\in E} \sum_{l=1}^{L} \sum_{k=1}^{N_i^{(l)}(M)} \sum_{j\in E} \left(1 - e^{g(i,j)/e_1}\right) \log\left(1 - F\left(x_i^{(l,k)}\right)\right)$$
$$+ \sum_{i\in E} \sum_{k=1}^{N_{i,M}(L)} \sum_{j\in E} \left(1 - e^{g(i,j)/e_1}\right) \log\left(1 - F\left(u_i^{(k)}\right)\right) = 0, \tag{17}$$

$$\frac{\partial \log \mathscr{L}}{\partial e_1} = \sum_{l=1}^{L} \sum_{i,j\in E} N_{ij}^{(l)}(M) \frac{e^{g(i,j)/e_1}}{1 - e^{g(i,j)/e_1}} \frac{g(i,j)}{e_1^2}$$
$$+ \sum_{l,i,k} \sum_{j\in E} \left(a_\infty e^{g(i,j)/e_1} \frac{g(i,j)}{e_1^2} \right) \log\left(1 - F\left(x_i^{(l,k)}\right)\right)$$
$$+ \sum_{i\in E} \sum_{k=1}^{N_{i,M}(L)} \sum_{j\in E} \left(a_\infty e^{g(i,j)/e_1} \frac{g(i,j)}{e_1^2} \right) \log\left(1 - F\left(u_i^{(k)}\right)\right) = 0. \tag{18}$$

Equation (17) provides an explicit expression of a_∞ in terms of e_1

$$a_\infty = \tag{19}$$
$$-\frac{\sum_{l=1}^{L} \sum_{i,j\in E} N_{ij}^{(l)}(M)}{\sum_{i\in E} \sum_{j\in E} \left(1 - e^{g(i,j)/e_1}\right) \left[\sum_{l=1}^{L} \sum_{k=1}^{N_i^{(l)}(M)} \log\left(1 - F\left(x_i^{(l,k)}\right)\right) + \sum_{k=1}^{N_{i,M}(L)} \log\left(1 - F\left(u_i^{(k)}\right)\right) \right]}.$$

This expression replaced in Eq. (18) provides an equation in e_1 that has to be solved numerically. Thus we obtain the corresponding MLEs $\widehat{a}_\infty(L, M)$ and $\widehat{e}_1(L, M)$ and also the corresponding plug-in estimator of a_{ij},

$$\widehat{a}_{ij}(L, M) = \widehat{a}_\infty(L, M) \left(1 - e^{g(i,j)/\widehat{e}_1(L,M)}\right). \tag{20}$$

Consequently, using Proposition 1, we get the plug-in estimators of the main quantities that define the semi-Markov system, namely p_{ij}, $W_i(t)$ and $Q_{ij}(t)$:

$$\widehat{p}_{ij}(L, M) = \frac{\widehat{a}_{ij}(L, M)}{\sum_{l\in E} \widehat{a}_{il}(L, M)} = \frac{N_{ij}(L, M)}{N_i(L, M)}, \tag{21}$$

$$\widehat{W}_i(t; L, M) = \left[1 - (1 - F(t))^{\sum_{j\in E} \widehat{a}_{ij}(L,M)} \right] \tag{22}$$

and

$$\widehat{Q}_{ij}(t;L,M) = \frac{\widehat{a}_{ij}(L,M)}{\sum\limits_{k\in E}\widehat{a}_{ik}(L,M)}\left[1-(1-F(t))^{\sum\limits_{k\in E}\widehat{a}_{ik}(L,M)}\right]. \tag{23}$$

Note also that, once we have obtained the estimators of the basic quantities associated to a multi-state semi-Markov system, we can immediately obtain estimators of the associated reliability indicators, following the lines presented in [19].

Remark 2 A more general framework may be considered if some or all of the parameters involved in the function $g(\cdot,\cdot)$ are assumed to be unknown. In such a case, the appropriate derivatives of the loglikelihood in (16) with respect to $c_m, k_m, l_m,\ m = 1, 2$, should be considered and the normal equations in addition to (17) and (18) should include the derivatives with respect to extra unknown parameters. In this more general setting, the system of equations has to be solved numerically for the estimators of the parameters to be obtained.

5 Concluding Remarks

In many settings the challenge is to determine if and where the parameters of the underlying model change their value. The rationales for time-varying parameter models may be several. For instance, the *true* coefficients themselves can often be viewed directly as the outcome of a stochastic process. Furthermore, even when the underlying parameters are stable, situations arise in which a time-varying coefficient approach will prove to be effective. More considerations could be provided on this topic. The present chapter deals with the problem in a general setting where a general class of distributions is considered with state-varying parameters. In particular in a multi-state system modeled by means of a special type of semi-Markov process, the parameters involved are assumed to be affected by the present state as well as the state to be visited and the likelihood together with the parameter estimates are provided under various dependency types of the parameters involved on the states of the system.

Acknowledgements The research work of Vlad Stefan Barbu was partially supported by the projects *XTerM—Complex Systems, Territorial Intelligence and Mobility* (2014–2018) and *MOUSTIC—Random Models and Statistical, Informatics and Combinatorial Tools* (2016–2019) within the Large Scale Research Networks from the Region of Normandy, France. The research work of Alex Karagrigoriou was completed as part of the research activities of the *Laboratory of Statistics and Data Analysis* of the Department of Mathematics of the University of the Aegean.

References

1. Barlow R, Wu A (1978) Coherent systems with multi-state components. Math Oper Res 3:275–281
2. El-Neveihi E, Proschan F, Sethuraman J (1978) Multistate coherent systems. J Appl Probab 15:675–688
3. Murchland J (1975) Fundamental concepts and relations for reliability analysis of multistate systems. In: Barlow R, Fussell J, Singpurwalla N (eds) Reliability and fault tree analysis: theoretical and applied aspects of system reliability. SIAM, Philadelphia, pp 581–618
4. Ross S (1979) Multivalued state component systems. Ann Probab 7:379–383
5. Block H, Savits T (1982) A decomposition of multistate monotone system. J Appl Probab 19:391–402
6. Hudson J, Kapur K (1982) Reliability theory for multistate systems with multistate components. Microelectron Reliab 22:1–7
7. Natvig B (1982) Two suggestions of how to define a multistate coherent system. Adv Appl Probab 14:434–455
8. El-Neveihi E, Proschan F (1984) Degradable systems: a survey of multistate system theory. Comm Stat Theory Method 13:405–432
9. Natvig B (1985) Multi-state coherent systems. In: Jonson N, Kotz S (eds) Encyclopedia of statistical sciences, vol 5. Wiley, New York, pp 732–735
10. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
11. Lisnianski A, Levitin G (2003) Multi-state system reliability: assessment. Optimization and applications. World Scientific, Singapore
12. Natvig B (2011) Multistate systems reliability. Theory with applications. Wiley, New York
13. Limnios N, Oprisan G (2001) Semi-Markov processes and reliability. Birkhauser, Boston
14. Limnios N, Ouhbi B (2003) Empirical estimators of reliability and related functions for semi-Markov systems. In: Lindqvist B, Doksum K (eds) Mathematical and statistical methods in reliability 7. World Scientific, Singapore, pp 469–484
15. Limnios N, Ouhbi B (2006) Nonparametric estimation of some important indicators in reliability for semi-Markov processes. Stat Methodol 3:341–350
16. Ouhbi B, Limnios N (1996) Non-parametric estimation for semi-Markov kernels with application to reliability analysis. Appl Stoch Model Data Anal 12:209–220
17. Ouhbi B, Limnios N (1999) Non-parametric estimation for semi-Markov processes based on its hazard rate functions. Stat Infer Stoch Process 2(2):151–173
18. Balasubramanian K, Beg Bapat R (1991) On families of distributions closed under extrema. Sankhya Indian J Stat 53:375–388
19. Barbu V, Karagrigoriou A, Makrides A (2016) Semi-Markov modelling for multi-state systems. Method Comput Appl Probab
20. Barbu V, Karagrigoriou A, Makrides A (2017) Statistical inference for a general class of distributions with time varying parameters, submitted
21. Jiang R (2016) Weibull model with time-varying scale parameter for modeling failure processes of repairable systems, In: Frenkel I, Lisnianski A (eds) Proceedings of the second international symposium on stochastic models in reliability engineering. Life science and operations management (SMRLO16). Beer Sheva, Israel, pp 292–295, IEEE CPS, 15–18 Feb 2016, 978-1-4673-9941-8/16 (2016). doi:10.1109/SMRLO.2016.54
22. Yang T, (2011) Measurement of yield distribution: a time-varying distribution model. Agricultural and Applied Economics Association, Annual Meeting, 24–26 July 2011. Pittsburgh, Pennsylvania
23. Zhu Y, Goodwin B, Ghosh S (2010) Modeling yield risk under technological change: dynamic yield distributions and the U.S. crop insurance program. Dissertation Paper, Department of Agricultural and Resource Economics, North Carolina State University
24. Limnios N, Ouhbi B, Sadek A (2005) Empirical estimator of stationary distribution for semi-Markov processes. Comm Stat Theory Method 34(4):987–995

25. Moore H, Pyke R (1968) Estimation of the transition distribution of a Markov renewal process. Ann Inst Stat Math 20:411–424

Optimizing Availability and Performance of a Two-Unit Redundant Multi-state Deteriorating System

Sonia Malefaki, Vasilis P. Koutras and Agapios N. Platis

Abstract The most of the contemporary large scale technological systems are functioning under multiple stages of degradation, from their perfect state to their total failure. The study of the performance and the availability of multi-stage systems is of great importance since their deterioration and/or failure may lead to important losses. Under a proper inspection and maintenance policy, it is feasible the operation of the system to be improved significantly. Our main goal is to model multi-state systems with redundancy and to identify the optimal maintenance policies. The system is inspected periodically. Depending on the condition of the system, either no action takes place or maintenance is carried out, either minimal or major. The proposed model takes also into account the scenario of imperfect and failed maintenance. The asymptotic behaviour of the system is studied and optimization problems for the asymptotic availability, the downtime cost and the expected cost due to maintenance and unavailability, with respect to inspection intervals, are formulated and solved.

Keywords Multi-state systems · Redundancy · Preventive maintenance · Dependability measures · Operational cost · Optimization

1 Introduction

Over the last years the design of large scale, complex and accurate technological systems is of great importance basically due to the rapid development of technology and the continuously increasing demand on various important industrial applications.

S. Malefaki (✉)
Department of Mechanical Engineering & Aeronautics, University of Patras,
26504 Rio, Patras, Greece
e-mail: smalefaki@upatras.gr

V.P. Koutras · A.N. Platis
Department of Financial and Management Engineering, University of the Aegean,
82100 Chios, Greece
e-mail: v.koutras@fme.aegean.gr

A.N. Platis
e-mail: platis@aegean.gr

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_5

Initially, these systems have been modelled and studied unrealistically as two–state systems (up/functioning state and down/failure state). In fact, several of these systems are functioning under multiple states of degradation, from their perfect state to their total failure [1]. Moreover, it has been observed that the deterioration process of these systems not only depends on the operational time of the system but also on the state of the system [2–4]. As a result, complex technological systems can operate in various degradation levels but with very low efficiency before their total failure. These systems are called multi-state deterioration systems and were introduced by Murchland [5] and they have been studied extensively the last decades (see for example [1, 3, 6–11] and the references therein). Moreover, the study of the performance and the availability of multi-state deterioration systems are of great importance since their deterioration and/or failure may lead to important economical and social losses [12, 13].

In order to improve the operation of a system and increase its availability and reliability, redundancy can be introduced. One of the most commonly used types of redundancy is the standby redundancy. In a standby system, apart from the functioning units, there is a number of standby units as backups, in order to replace any component after its failure. There are three types of standby, i.e. cold, warm and hot standby. In cold standby case, the inactive components cannot fail during the standby period but they need time in order to enter in a functioning mode. On the other hand, when time to response to a failure is of great importance a hot standby strategy should be adopted. In hot standby case, the standby units are ready for use, thus they can enter immediately into a active state. The drawback in this case is that since the standby unit are active before being used, they have an increased failure rate at the time that the system control is switched to them. Finally, the warm standby is an intermediate case, meaning that the inactive units have lower failure rates than the hot standby case and they can enter into functioning state faster than the cold standby [14]. In this study, a two-unit cold standby system is proposed under the assumption that the switching process does not last too long and it is perfect.

For further improvement of the operational time and conditions of a multi-state system, maintenance actions can be adopted [15, 16]. All the necessary actions for keeping and/or restoring a system in an acceptable functioning condition or even extending its lifetime are considered as maintenance of a system. These actions can be divided into corrective and preventive maintenance actions. Corrective maintenance actions take place just after a total failure of the system and restore the system in an operational condition. Contrarily to corrective, preventive maintenance actions take place regularly at predefined time intervals during the operational time of the system in order to restore the system in a better deterioration state or even to its perfectly functioning state [8]. Under normal circumstances, preventive maintenance is more effective than corrective, since its main aim is to keep the system available and avoid undesirable failures that incur cost [17].

There are two main types of preventive maintenance, condition based and time based maintenance [18]. Time-based maintenance is carried out at specific time intervals independently of the system' s state. On the other hand, the condition based maintenance depends on the system's state. The system is inspected regularly and depending on its state, either it is left without maintenance, if it operates in an acceptable level, or minimal/major maintenance takes place restoring the system in a previous deterioration level or in an as good as new state respectively, if the system does not operate in an efficient deterioration level.

Although preventive maintenance has been adopted for improving the performance of a system, it incurs cost. Thus, an appropriate preventive maintenance schedule that manages to reduce the total operational cost and improve the availability of the system should carefully designed and implemented. A lot of research effort has been conducted to this direction (see for example [19–22] and the references therein). A recent review paper on optimal maintenance policies is [23].

In the current work a system with two identical units is considered, in which condition based preventive maintenance is performed. One unit is operational and the other one is in cold standby. Each unit experiences k deterioration levels before its total failure. Among these deterioration levels, it is assumed that in the g initial levels, each unit operates almost perfectly, a degraded state is assumed when each unit operates from $g+1$ to b deterioration level and an extensive degradation state is assumed when each unit operates from $b+1$ to k deteriorated level. The system is inspected in constant time intervals and when the deterioration level of the system is lower than g, no maintenance action takes place. On the other hand, if the detected deterioration level is between g and b, minimal maintenance is performed. Finally, when the deterioration is higher than level b, major maintenance action is triggered. Apart from being perfectly implemented, the minimal and the major maintenance can be imperfect or even failed restoring the system not in the desirable level but in a higher deterioration level or even in a total failure. Our main aim is to determine an inspection policy and thus an optimal maintenance policy that optimize system's dependability and/or performance measures.

The rest of the paper is organized as follows: In Sects. 2 and 3, the proposed model is described in detail. The necessary Semi-Markov theory is presented in Sect. 4. The main dependability and performance measures are presented in Sect. 5 and in Sect. 6 the optimization aspect is considered. In Sect. 7 some numerical results are presented and the paper concludes by providing a short discussion.

2 Description of a Two-Unit System with Maintenance

The proposed system consists of two identical units, one operational and one in cold standby. It is assumed that the system starts to operate in its perfect functioning state, i.e. the functioning and also the standby unit are in their perfect state. In order to improve the operation of the system and to extend its lifetime, condition based maintenance is proposed. Thus, the system is inspected regularly and whenever certain conditions are met, the proper maintenance action is initiated.

2.1 Two-Unit System with Minimal and Major Maintenance

Since the units are assumed to be identical, initially the operating flow of the functioning unit is presented and thereafter the combination of both units is described. The functioning unit starts to operate in its perfect state. Without loss of generality, we assume that each unit, apart form being at a perfect functioning state, can experience two levels of degradation before a total failure occurs. To capture real life conditions, it is assumed that the level of deterioration of each unit increases with respect to its operational time. From each deterioration level the unit may enter an inspection state where the level of deterioration is identified, or it may enter the next deterioration state or even a total failure, mainly due to external unpredicted conditions. As far as the inspection process is concerned, it is assumed that inspection is carried out within a constant time interval of short duration. Depending on the condition of the unit, from the inspection state it may return to its previous deterioration state if it can still operate in an acceptable level, or it may enter a maintenance state. The type of maintenance to be carried out depends on the level of unit deterioration. More specifically, if the deterioration is negligible and the unit can still operate on its perfect functioning state, no maintenance action takes place. In this case after inspection, the unit reenters the aforementioned state. In case that the unit operates in its first deterioration level, after inspection, it enters a minimal maintenance state. Correspondingly, if inspection identifies that the unit operates in its second deterioration level, major maintenance actions are triggered.

Minimal maintenance, when performed, restores the unit to its previous deterioration level. Nevertheless, it is assumed that minimal maintenance can be imperfect as well. In the latter case, maintenance fails to be properly implemented and the unit returns to the same deterioration state (first deterioration state). Besides imperfect minimal maintenance, the case where minimal maintenance is badly performed, mainly due to human factor, is also taken into account. In this case the unit enters a worse deterioration state or even the total failure state. In such cases, the minimal maintenance action is characterized as failed. Note that, as already mentioned, minimal maintenance actions have minor impact on units upgrading but this is also the case for its cost.

On the other hand, major maintenance is an expensive action of great impact on unit's deterioration state. Likewise minimal maintenance, major maintenance can be perfect, imperfect or failed too. When major maintenance is properly implemented, it restores the unit to its initial perfect functioning state. In an imperfect major maintenance action the unit returns back to any better state or to the deterioration state where the inspection was initiated but not to the perfect state. More specifically, in case of an imperfect major maintenance the unit does not manage to return to an as good as new state but enters any state with lower or equal deterioration level with the one where the inspection initiated. Finally, in the case of failed major maintenance, the unit enters a worse deterioration state instead of a better one. In the studied model, this means actually that the unit enters the total failure state. When the operating unit experiences a total failure, a repair process is initiated to restore the unit to its perfect state.

Taking into account the operation flow of each unit, the behavior of the entire two-unit standby system can be described. The fundamental assumption of the model is that the system's control is switched to the standby unit after a total failure of the primary unit or when a maintenance action of the primary unit starts, if and only if the standby unit is at an operational condition. However, it is meaningless to switch system control to the standby unit if this unit is under maintenance or repair. Thus, the switching process takes place only when the secondary unit is at the fully operational state or at a deterioration state. As far as switching is concerned, we assume that it is perfect.

Note that additional assumptions need to be listed. It is assumed that when the maintenance of the standby unit is successfully completed, the unit becomes available and consequently is ready to take system control whenever the primary unit fails or needs to be maintained. If maintenance of the standby unit fails, the unit remains at the non-operational failure state and thus it cannot take system control unless it is repaired. On the other hand, if the maintenance of the standby unit is imperfect and the unit enters either the first or the second deterioration state, it is still in an operational mode and it can take the control of the system if needed. Similarly, when the standby unit is repaired after a failure it becomes available and ready to takeover when the primary unit will enter a non-operational state. Note also, that the inspection states are considered as non-operational states, but due to the fact that the sojourn times in these states are significantly shorter that the rest of the sojourn times, it is assumed that the system control is not switched to the standby unit when the primary unit enters an inspection state.

Each system state is denoted by a pair (j, i), where index j denotes the condition of the first unit while index i denotes the condition of the second one with:

$$i, j \in \{O,\ S,\ D_1^O,\ D_1^S,\ D_2^O,\ D_2^S,\ I_0,\ I_1,\ I_2,\ m,\ M,\ F\}$$

where

- O unit is operational
- S unit is in a standby mode
- D_j^c unit is at its j-th deterioration level and in operational condition c $c = O$ if is operational, $c = S$ if is standby
- I_j inspection state of unit when there is on j deterioration level, $j \in \{0, 1, 2\}$
- m minimal maintenance state of unit
- M major maintenance state of unit
- F unit is in the total failure state

In Fig. 1, the concept of the multi-state two-unit deteriorating system is presented through two parts of a state transition diagram. Nevertheless, in Appendix, Table 3 with all possible states and transitions of the system is given in order to fully describe system's evolution.

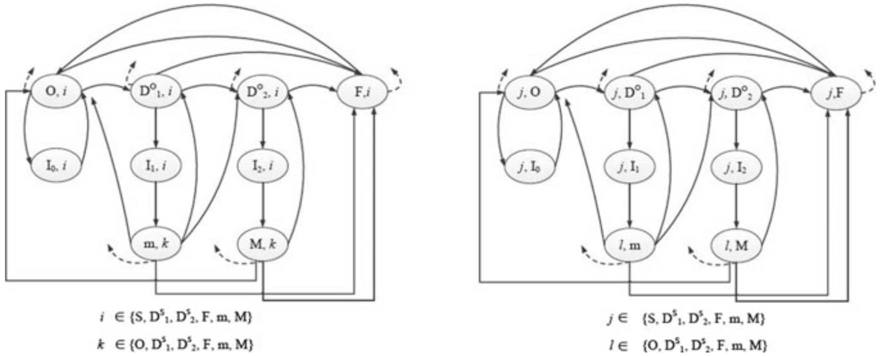


Fig. 1 Parts of the state transition diagram for a two-unit system with minimal and major maintenance actions

As mentioned above, neither the effects nor the cost of performing minimal or major maintenance are the same. Minimal maintenance has less effects and lower cost though major maintenance has more efficient effects but higher cost. Thus, usually in the relevant literature the aim is distinguish a trade-off between performing minimal or major maintenance. Having already presented a model with both actions, it is also interesting to examine the behaviour of the system when only one type of maintenance is performed.

2.2 *Two-Unit System with Minimal or Major Maintenance Only*

Initially, a multi-state two-unit system with only minimal maintenance actions is considered. Units' evolution is exactly the same, as presented in the previous section. The difference of this model consists in the fact that minimal maintenance is performed whenever maintenance is needed, i.e. in both deterioration levels. Perfect, imperfect and failed minimal maintenance actions have exactly the same effects as previously described. Fig. 2 and Table 4 in Appendix, provides a description of the proposed model with solely minimal maintenance actions.

Correspondingly, a multi-state two-unit system with only major maintenance actions is also considered. Once again, the effects of perfect, imperfect and failed major maintenance are exactly the same as previously described. It is worth mentioning though that major maintenance is assumed to be implemented only at the second deterioration level of each unit. Figure 3 and Table 5 in Appendix, provides a description of the proposed model with solely major maintenance actions.

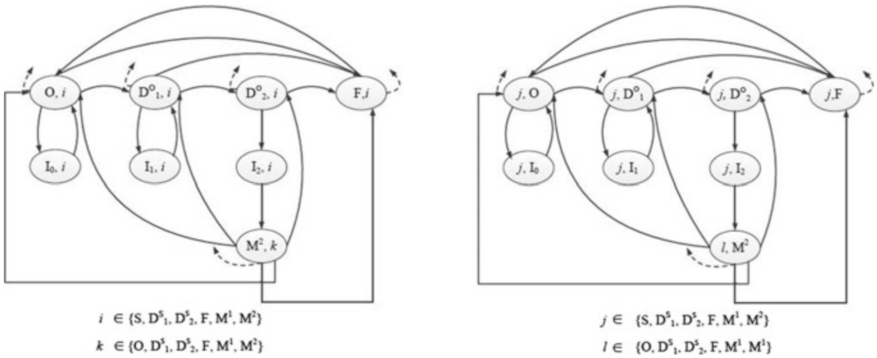


Fig. 2 Parts of the state transition diagram for a two-unit system with only minimal maintenance actions

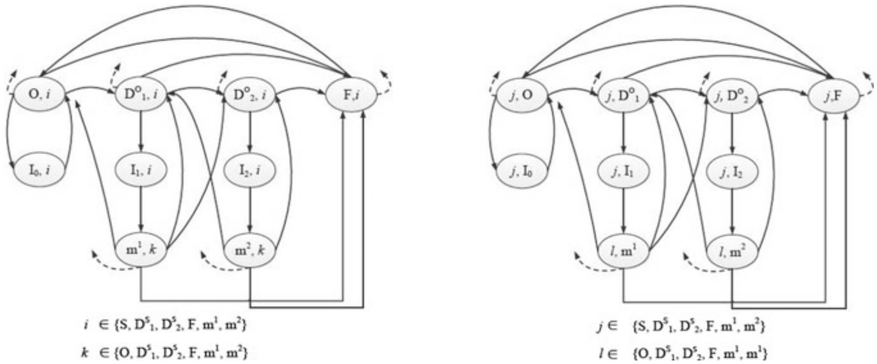


Fig. 3 Parts of the state transition diagram for a two-unit system with only major maintenance actions

3 Model Assumption and Sojourn Time Distributions

Modelling the deterioration, the failure, the maintenance and the repair and/ or replacement process with proper distribution is of great importance and has gained a lot of research effort in the recent literature.

A realistic assumption for the inspection times is to be considered constant and equals to T. Thus, the sojourn time distribution at any deterioration level given that the system enters the inspection state can be modelled by the unit-step function:

$$u(t-T) = \begin{cases} 0 \text{ for } t < T \\ 1 \text{ for } t \geq T. \end{cases}$$

For the response time of the inspection process there are several suggestions in the literature from negligible response time [24], to constant or exponential [8, 10, 25] and Erlang distributed times [26]. In the current work response time is assumed to be exponentially distributed but negligible comparing with the rest of the times. Note that, when the operating unit is under inspection, the cold stand-by unit does not turn into functioning mode. Thus, the inspection states are considered down states for the whole system. The deterioration process is usually assumed exponentially distributed [10, 25, 27, 28]. Moreover, Weibull [18, 29, 30] or hypo-exponential [28, 31] distributions have also been used for modelling the deterioration process. Without loss of generality, in the current work, exponential distributions are assumed for the deterioration process with state-dependent failure rates, i.e. the higher the deterioration level the more probable is for the system to enter into a worse deterioration level. The time to failure of the functioning unit of the system is also exponentially distributed. Usually the corrective maintenance, either minimal or maximal, and the repair times are assumed exponentially distributed [10, 25, 27]. In a less realistic case, the duration of maintenance has been considered negligible (see for example [32]). Moreover, several distributions, as lognormal [33], Erlang [11] and many others, have been used for modelling the maintenance and the repair times. Constant maintenance and repair times have been also used [8, 29]. In the current work exponential distributions have been assumed for the maintenance and the repair times. Finally, the imperfect and the failed, either minimal or major maintenance are assumed exponentially distributed.

Under these assumptions, the system's evolution in time is described by a semi–Markov process $\{Z(t),\ t \geq 0\}$. Thus, the asymptotic behaviour of the system and its performance and dependability analysis will be studied under the semi-Markov theory.

4 Semi-Markov Modelling

In order to define the semi-Markov process which describes the evolution in time of the studied system, initially some other important stochastic processes will be defined.

Roughly speaking, a stochastic process is a probabilistic model for a system that evolves randomly in time. If the system is observed continuously in time and $X(t)$ being its state at time t, then it is described by the continuous time stochastic process $(X(t),\ t \in R_+)$ with state space $(E, \mathcal{E})$. If the state space is countable and moreover the equality

$$P(X(t+s) = j | X(u), u \leq s) = P(X(t+s) = j | X(s)),$$

for all $s, t \in R_+ \times R_+$, and $j \in E$, holds true, $X(t)$ is a Markov process.

Let also $0 = S_0 \leq S_1 \leq \cdots \leq S_n \leq S_{n+1} \leq \ldots$. If the two-dimensional process $(X_n,\ S_n)_{n\in N}$ satisfies the following equality

$$\begin{aligned} P(X_{n+1} = j,\ S_{n+1} - S_n \leq t | X_0, \ldots, X_n, S_1, \ldots, S_n) \\ = P(X_{n+1} = j,\ S_{n+1} - S_n \leq t | X_n), \end{aligned}$$

for all $t \in R_+, j \in E$ and $n \in N$ is called Markov Renewal Process (MRP) with transition kernel

$$Q_{ij}(t) = P(X_{n+1} = j,\ S_{n+1} - S_n \leq t|X_n)$$

(see for example [34, 35]).

The stochastic process $\{Z(t), t \geq 0\}$ which is associated to the MRP $(X_n,\ S_n)_{n\in N}$ through the equality

$$Z(t) = X_n \text{ if } S_n \leq t < S_{n+1} \text{ for } t \geq 0$$

is a semi-Markov process that represents the state of the MRP at an arbitrary time point. Thus, $Z(t)$ gives the system states at time t and S_i's represent the jump times of the semi-Markov process.

The transition probabilities of the Markov chain X are equal to

$$p_{i,j} = Q_{i,j}(\infty) = \lim_{t\to\infty} Q_{i,j}(t),\ \text{for } i,j \in E.$$

Moreover, let $\xi_n = S_n - S_{n-1},\ n \geq 1$, then the semi-Markov kernel can be rewritten as

$$Q_{ij}(t) = P(X_{n+1} = j, \xi_{n+1} \leq t|X_n = i) = p_{ij} \cdot F_{ij}(t),\ i,j \in E$$

where $F_{ij}(t) = P(\xi_{n+1} \leq t|X_n = i, X_{n+1} = j)$ is the conditional distribution function of the sojourn time in state i given that the next visited state is the state j with $j \neq i$.

The non-zero elements of the transition probability matrix of the Markov chain X, the embedded Markov chain of $Z(t)$, is given by the formula

$$p_{ij} = \int_0^\infty \prod_{l\in L_{i-j}} \bar{F}_{i,l}(t) dF_{ij}(t),$$

where L_{i-j} is the set of all the states of the system that state i is connected with except from state j.

The steady state probabilities of the semi-Markov process $Z(t)$ are given by

$$\pi_j = \frac{v_j \cdot m_j}{\sum_{i\in E} v_i \cdot m_i},\ j \in E,$$

where $\mathbf{v} = \{v_i,\ i \in E\}$ the invariant probabilities of the embedded Markov chain $\{X_n\}_{n\in N}$ and

$$m_i = \int_0^\infty \prod_{l\in L_i} \bar{F}_{i,l}(t) dt,$$

the mean sojourn time in state i [36, 37].

Having computed the asymptotic probabilities of the SMP, the asymptotic performance and availability of the studied system can be derived and are also of great importance.

5 Asymptotic Dependability and Performance Measures

For the stochastic system with state space E described by a semi–Markov process $\{Z(t),\ t \geq 0\}$ lets consider a partition $U,\ D$ of E, i.e. $E = U \cup D$ with $U \cap D \neq \emptyset$. The set U contains the up states (functioning states) of the system and the set D contains the down states (not functioning states) of the system.

5.1 Asymptotic Availability

The availability of the system can be defined as the probability of the system to be in a functioning state at time t

$$Av(t) = P(Z(t) \in U).$$

Thus the asymptotic availability is

$$Av = Av(\infty) = \lim_{t \to \infty} Av(t) = \sum_{i \in U} \pi_i.$$

Besides its dependability expressed by the asymptotic availability index, the performance of such a system is also of prior importance, since preventive maintenance although is carried out to avoid a total failure it incurs downtime and cost. To model and measure systems performance, two performability indicators are considered. The first one concerns system downtime, though the second one concerns overall operational cost.

5.2 Expected Downtime Cost

The two-unit multi-state deteriorating system is considered down when none of the units is in an operational state. Each unit is considered as non operational when inspection is performed, maintenance is implemented or finally it is under a repair process. Combining all the possible states of the system where no unit is operating, subset D of system down states is constructed.

In real-life systems the total time that a system is unavailable to operate is of critical importance. Maintenance, repair and other actions schedules are carefully designed in order to minimize the total system downtime. Consequently, usually we are interested in the total time that a system is non operational during a specified time period. The mean value of a random variable expressing the cumulative time spent in system down states in hours per year, and in steady state is considered to express the expected downtime [38, 39]:

In order to express the aforementioned performance measure, let

$$w_1(i) = \begin{cases} 8760 \text{ if } i \in D \\ 0 \quad \text{ if } i \in U \end{cases}$$

be a reward function concerning the hours per year. Let also

$$g_1(Z(t)) = \sum_{i \in E} w_1(i) \cdot I(Z(t) = i).$$

be the downtime reward rate a time t, where $I(Z(t) = i)$ is the indicator function and $Z(t)$ the systems state at time t. The total Expected Downtime Cost (EDC) in hours per year and in the steady-state can be then derived as follows:

$$\begin{aligned} EDC(T) = \lim_{t\to\infty} E(g_1(Z(t))) &= \lim_{t\to\infty} E\left(\sum_{i\in E} w_1(i)\cdot I(Z(t)=i)\right) \\ &= \lim_{t\to\infty} \sum_{i\in E} w(i)\cdot E\left(I(Z(t)=i)\right) \\ &= \lim_{t\to\infty} \sum_{i\in E} w_1(i)\cdot P(Z(t)=i) = 8760\cdot \sum_{i\in D} \pi_i. \end{aligned}$$

where π_i is the steady state probability for the proposed two-unit multi-state deteriorating system of the semi-Markov process $\{Z(t),\ t \geq 0\}$.

5.3 Expected Cost Due to Maintenance and Unavailability

Another measure that is of prior importance when determining or estimating the optimal inspection policy for similar systems, is the expected cost related to the maintenance, repair and inspection and additionally to the unavailability of the system [39]. The meaning of such a measure consists firstly in the idea that any action taken incurs a cost. Thus, inspection, maintenance, either minimal or maintenance and finally repair, are actions that cost. Furthermore, for systems which are designed to operate continuously in time, for example technological systems in industry, the cost of not being available is also of critical importance. Independently of the reason why such a system is unavailable, the fact of unavailability itself incurs a cost.

To model such a performance measure [38], let $w_2(i) = d_i \cdot C_i$ be a reward function where

$$d_i = \begin{cases} 1 \text{ if } i \in D \\ 0 \text{ if } i \in U \end{cases}$$

Moreover let C_i be a cost function including inspection, maintenance, repair and unavailability costs. The cost function can defined as:

$$C_i = \begin{cases} C_m & \text{if in state } i \text{ a unit is under minimal maintenance} \\ C_M & \text{if in state } i \text{ a unit is under major maintenance} \\ C_R & \text{if in state } i \text{ a unit is under repair} \\ C_{ins} & \text{if in state } i \text{ a unit is under inspection} \\ C_D & i \in D \end{cases},$$

where C_m is the minimal maintenance cost, C_M is the major maintenance cost, C_R is the repair cost, C_{ins} is the cost of inspection and C_D is the cost due to system unavailability. Note that, since inspection and maintenance are planned actions, their corresponding costs are considerably lower than the cost of an unplanned failure and the consequent repair. Moreover, due to its nature, minimal maintenance cost is lower than major maintenance cost though the inspection cost is much lower than maintenance and repair costs. Finally, since in system down state, either inspection or maintenance or repair is carried out, the additional unavailability cost C_D that is added actually to the aforementioned action costs should be considered low enough.

To define the total expected cost due to maintenance and unavailability per unit time, let

$$g_2(Z(t)) = \sum_{i \in E} w_2(i) \cdot I(Z(t) = i).$$

be the cost rate a time t. Then the total Expected Cost due to Maintenance and Unavailability (ECMU) per unit time in the steady-state for the proposed model can be derived as:

$$\begin{aligned} ECMU(T) = \lim_{t \to \infty} E(g_2(Z(t))) &= \lim_{t \to \infty} E\left(\sum_{i \in E} w_2(i) \cdot I(Z(t) = i)\right) \\ &= \lim_{t \to \infty} \sum_{i \in E} w_2(i) \cdot E\left(I(Z(t) = i)\right) \\ &= \lim_{t \to \infty} \sum_{i \in E} w_2(i) \cdot P(Z(t) = i) \\ &= \sum_{i \in D} w_2(i) \cdot \pi_i. \end{aligned}$$

6 Optimization Problems

All of the aforementioned dependability and performance measures need to be optimized with respect to the inspection and hence maintenance policy which is expresses by the inspection interval T. Thus, asymptotic availability, EDC and ECMU can be considered as functions of the inspection interval T. Then, the following optimization problems arise.

6.1 Maximization of the Asymptotic Availability

Initially the maximization of the asymptotic availability is of great importance, thus the following optimization problem should be solved

$$\max_T Av(T) = \max_T \sum_{\pi \in U} \pi_i \qquad (1)$$
$$T_{min} < T < T_{max},$$

where T_{max} is a predefined data-dependent time period which is considered reasonable as inspection interval and T_{min} is a predefined minimum inspection time period since it is assumed that inspection cannot be performed instantaneously when the system enters a deterioration state.

6.2 Minimization of the Overall Cost

Alternatively, the minimization of EDC is also a very important and challenging problem, i.e.

$$\min_T EDC(T) \qquad (2)$$
$$T_{min} < T < T_{max},$$

while for minimizing ECMU, the corresponding optimization problem is:

$$\min_T ECMU(T) \qquad (3)$$
$$T_{min} < T < T_{max},$$

There are cases that we are not interested in optimizing only either the asymptotic availability or the EDC and the ECMU independently, but in optimizing them simultaneously. Thus, the need to use multi-objective optimization methods is arises.

6.3 Multi-objective Optimization Using Weighted Sum Approach

To decide on a policy that can optimize simultaneously the total expected cost due to maintenance and unavailability and system asymptotic availability the following optimization problem is solved:

$$\begin{aligned} &\min_{T} ECMU(T) \\ T_{min} &< T < T_{max}, \\ &Av(T) \geq Av_0 \end{aligned} \tag{4}$$

Note that availability participates into the optimization as a constraint. The designer should decide on in order to guarantee the desired level of system asymptotic availability.

7 Numerical example

In order to illustrate the theoretical results presented in the previous sections, a system with two identical units, one functioning and one in cold standby is considered. Each unit works under three levels. The system is inspected periodically in T time units. Without loss of generality, exponential distributions are assumed for the rest of the transitions, whose failure rates are presented in Table 1.

Initially, we are interested in identifying the effects of inspection policies on the measures of interest. In Fig. 4, the behaviour of the asymptotic availability with respect to the inspection interval T is shown, where minimal and major maintenance actions can be implemented. As it can be observed, the availability increases with the increase of T up to T_0 and from this value onwards it gets quite stable. This is due to the fact that as delayed the system enters inspection states it spends more time to operational states and thus the availability is higher.

This is also the case when only major maintenance is allowed to be performed. However, when only minimal maintenance is performed availability achieves its higher value for a certain T, the optimal one, but thereafter it reduces with the increase of T. This is because in this case, there should be a trade-off between inspecting as rarely the system as possible but on the other hand, since minimal maintenance manages to restore the system only to the previous deterioration state which is more susceptible to deterioration, an often inspection could help to avoid a total failure which incurs a high enough sojourn time in a non-operational state. The aforementioned observation can be seen in Fig. 5. Moreover, by Fig. 5 it is shown that the system with minimal and major maintenance actions reveals a lower maximum availability value than the system with only major or only minimal maintenance. But as T increases, the system with minimal and major maintenance provides higher

Table 1 The failure rates of the conditional sojourn time distributions in time units^{-1}

Transition	Failure rates	
From perfect state to first deterioration level (deterioration)	λ_1	1/120
From first to second deterioration level (deterioration)	λ_2	1/100
From second deterioration level to total failure (deterioration)	λ_3	1/80
Sudden failure	λ_{sf}	0.0001
Minimal maintenance action	λ_m	1.0
Major maintenance action	λ_M	0.2
Imperfect minimal maintenance action in first deterioration level	λ_{mD1}	0.02
Imperfect minimal maintenance action in second deterioration level	λ_{mD2}	0.01
Imperfect major maintenance action in first deterioration level	λ_{MD1}	0.025
Imperfect major maintenance action in second deterioration level	λ_{MD2}	0.015
Failed minimal maintenance action	λ_{Fm}	0.001
Failed major maintenance action	λ_{FM}	0.001
Repair action	λ_R	0.025
Time in inspection state	λ_{ins}	1/10

availability comparing with the system with only minimal but lower availability than the system with only major maintenance. The behavior of availability indexes lies on the fact that firstly the proposed model with only major maintenance actions has less non-operational states comparing with the rest of the models. Additionally, the model with minimal maintenance spends less time at the non-operational maintenance states than the model with minimal and major maintenance, since in the first one there is only minimal maintenance which lasts less than major.

As far as the expected downtime is concerned, its behaviour with respect to the inspection interval for the model with minimal and major maintenance is shown in Fig. 6. The total expected downtime decreases as the inspection interval decreases since the system enters inspection and maintenance states, which are down states, less often.

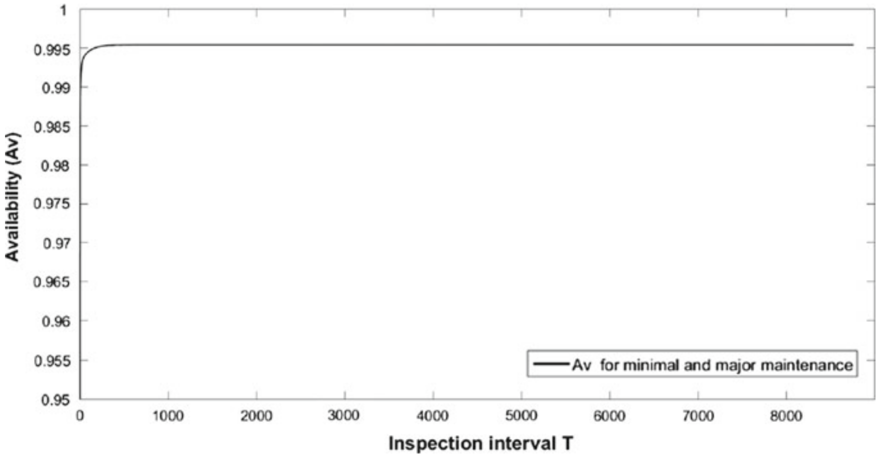


Fig. 4 Asymptotic availability for the two-unit multi-state deteriorating system with minimal and major maintenance actions

Similar is the behaviour of EDC for the models with major maintenance and minimal maintenance only, as it can be observed in Fig. 7. However, note that the model with minimal maintenance only reveals lower downtime than the rest of the models due to the reason already discussed concerning the reduced time spent in minimal maintenance states. Moreover, the model with major maintenance only, reveals less downtime than the model with minimal and major maintenance, since there are less maintenance states and consequently the system will spend less time in maintenance down states.

Additionally, the behaviour of index ECMU with respect to T is presented in Fig. 8 for the model with minimal and major maintenance actions. The increase of T causes an increase of ECMU since delayed inspection is responsible for letting the system without maintenance, a fact that makes the system susceptible to total failures which incur importantly higher cost than maintenance does.

This is also the case for the model with only minimal and only major maintenance, as it can be seen in Fig. 9. It is worth mentioning that the model with only major maintenance actions reveals the lowest cost. This is clearly due to the fact that maintenance is decided to be performed only at the second deterioration levels and thus there fewer states that incur cost. Additionally, it is obtained that when only minimal maintenance is performed, the system incurs higher cost comparing with the case of performing both minimal and major maintenance actions. This can be explained by the fact that minimal maintenance alone, although it incurs less cost, cannot effectively prevent total failures that incur importantly higher cost.

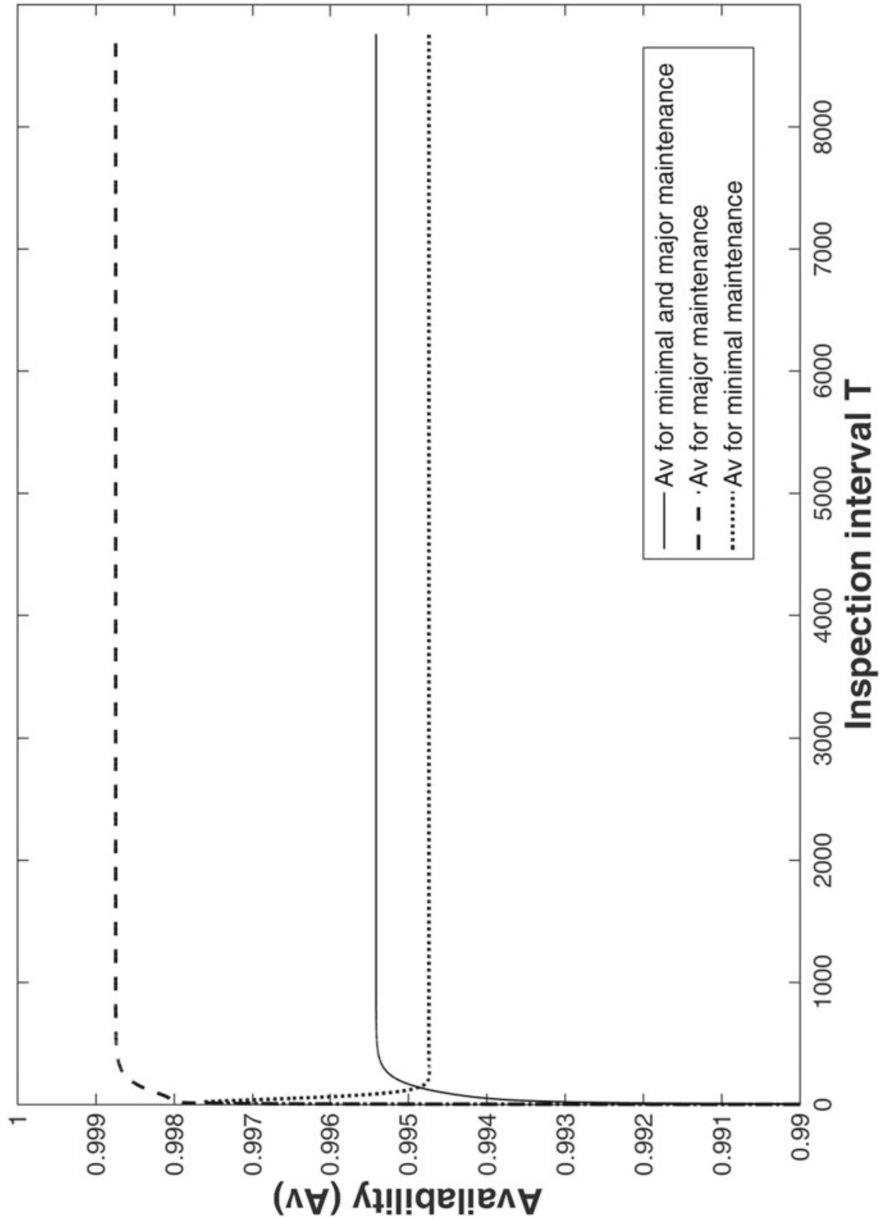


Fig. 5 Asymptotic availability comparison

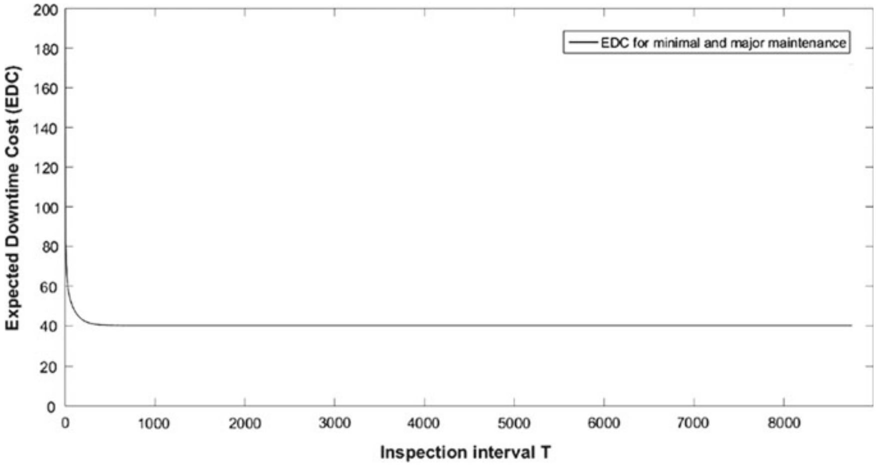


Fig. 6 Expected downtime for the two-unit multi-state deteriorating system with minimal and major maintenance actions

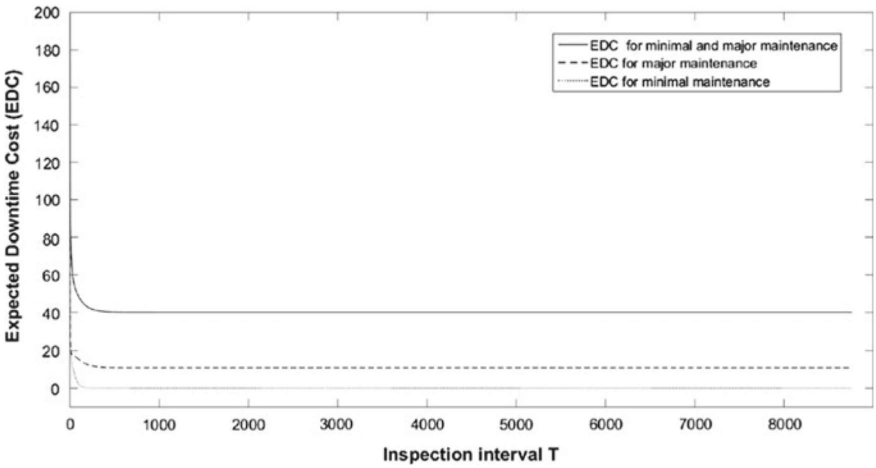


Fig. 7 Expected downtime comparison

Besides the behaviour of the measures of interest with respect to the inspection interval T, it is also interesting to examine the effects of model parameters upon availability, downtime and operational cost. Note that the rate of a sudden failure λ_{sf}, the rates of deterioration λ_1, λ_2, λ_3 and the repair rate λ_R are the parameters which affect the most Av, EDC and ECMU for all the proposed models.

In Fig. 10, the effect of the sudden failure rate λ_{sf} on asymptotic availability is presented. As it can be observed, an increased sudden failure rate causes a decrease of the availability since the system enters more often the total failure state. Correspondingly by Fig. 11 it is obtained that as the sudden failure rate λ_{sf} increases the

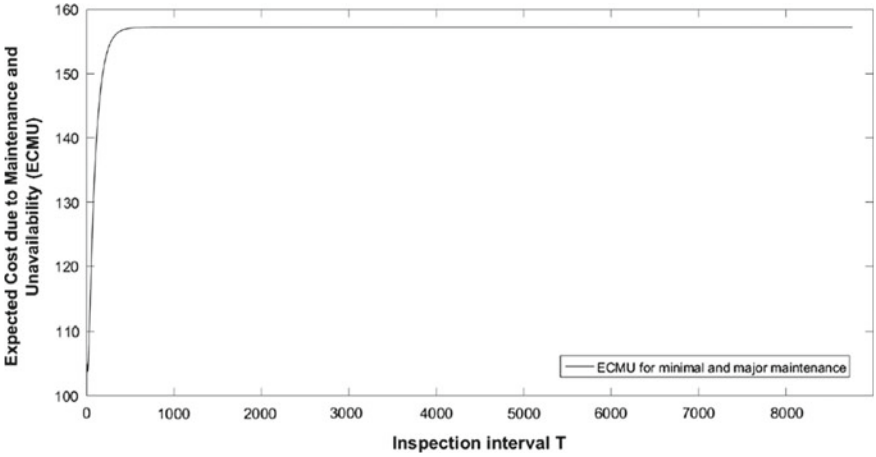


Fig. 8 Expected cost due to maintenance and unavailability for the two-unit multi-state deteriorating system with minimal and major maintenance actions

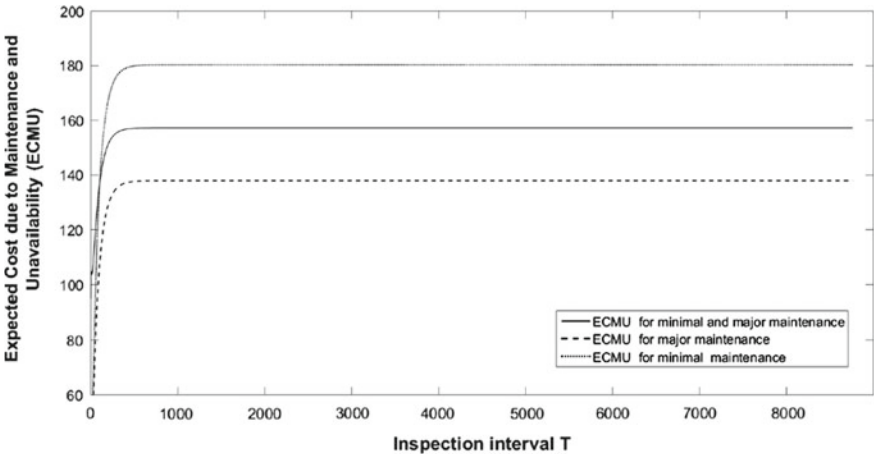


Fig. 9 Expected cost due to maintenance and unavailability comparison

downtime increases as well since the system enters more frequently the total failure state which incurs a higher mean sojourn down time. This is also the case for the expected cost due to maintenance and unavailability which increases, since the total failure state incurs a much higher cost, with the increase of λ_{sf} as it is shown in Fig. 12.

The effects of the deterioration rates λ_1, λ_2 and λ_3 on system's availability are shown in Fig. 13. As it can be observed, the increase of the deterioration rates causes a decrease on availability since the system enters faster and more often the total

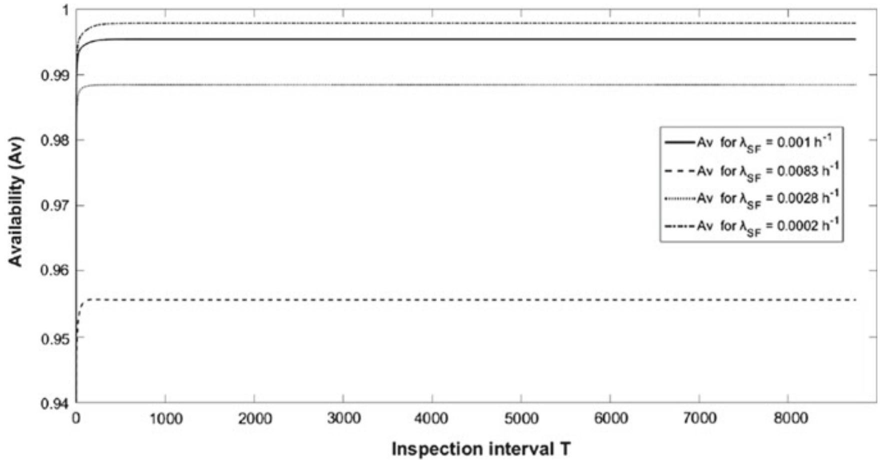


Fig. 10 Effect of sudden failure rate λ_{sf} on asymptotic availability

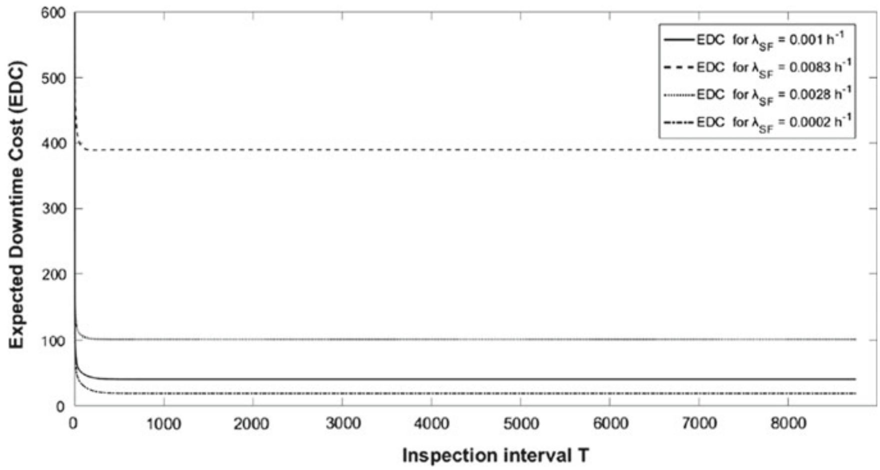


Fig. 11 Effect of sudden failure rate λ_{sf} on expected downtime

failure state. Similarly, as it can be observed in Figs. 14 and 15 respectively, the increase of the deterioration rates results in increasing not only the downtime but the total operational cost as well.

Finally, the effects of repair process duration on availability, expected downtime and total expected cost are presented in Figs. 16, 17 and 18 respectively. As it can be observed, the increase of time to repair results an increased sojourn time in the total failure state which and as consequence the system remains for longer a the

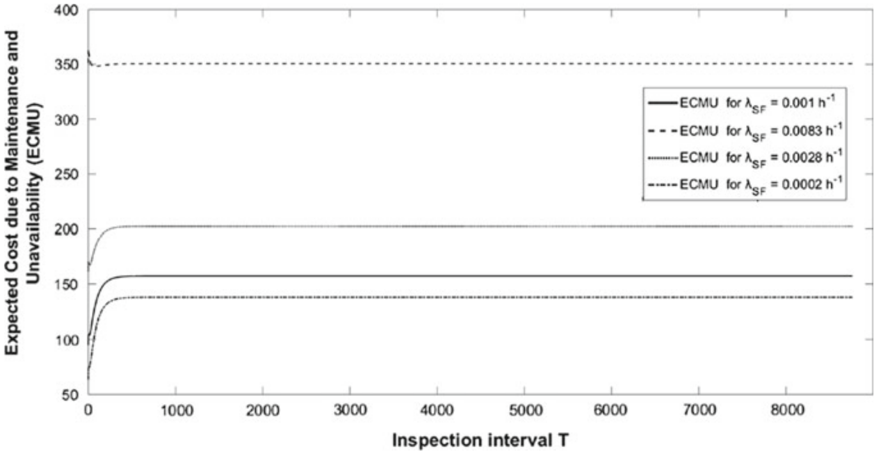


Fig. 12 Effect of sudden failure rate λ_{sf} on expected cost due to maintenance and unavailability

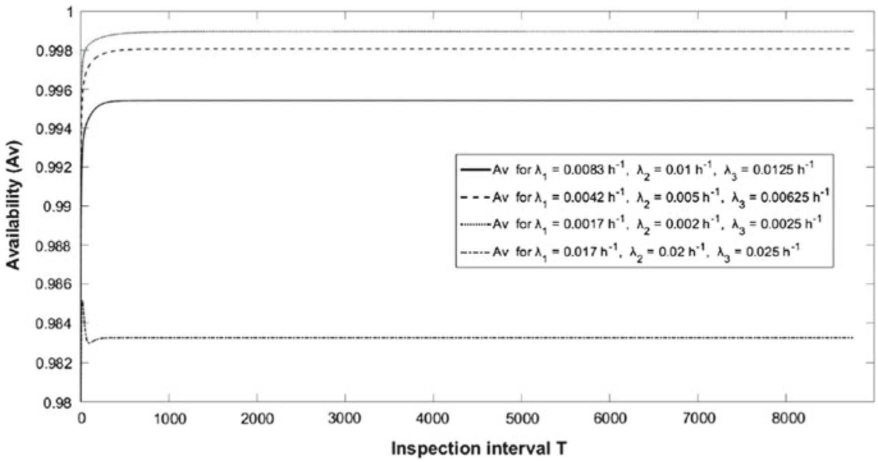


Fig. 13 Effect of the deterioration rates λ_1, λ_2 and λ_3 on asymptotic availability

total failure state resulting thus in an increased availability measure, an increased downtime and cost.

Apart from modelling a two-unit redundant multi-state deteriorating system with maintenance, the endmost aim of this work, as already highlighted, is to provide the system designer with the appropriate theoretical and methodological framework in order to schedule an inspection policy that optimizes the measures of interest, either by introducing and solving separate optimization problems for each measure or by considering a multi-objective optimization problem. Thus, in Table 2, the results of the corresponding optimization problems are provided. Since the data used are not

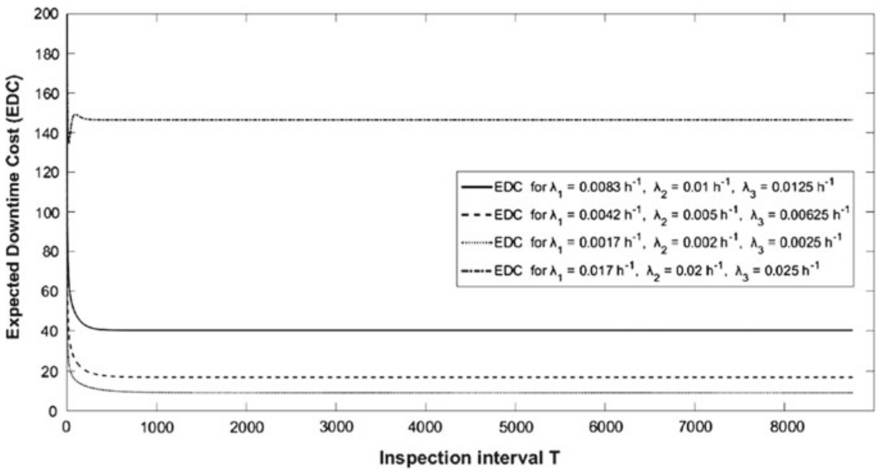


Fig. 14 Effect of the deterioration rates λ_1, λ_2 and λ_3 on expected downtime

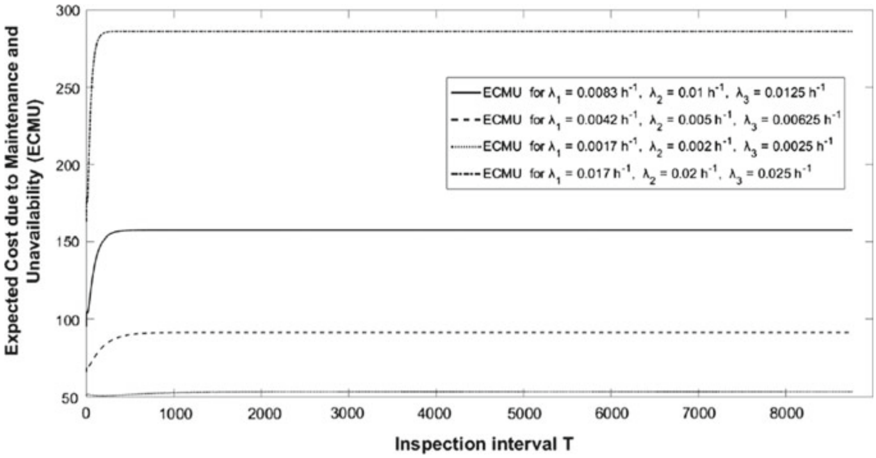


Fig. 15 Effect of the deterioration rates λ_1, λ_2 and λ_3 on expected cost due to maintenance and unavailability

taken form a real life system, the designer can implement the parameters for the system that he intends to improve into the presented analysis and conclude with the optimal inspection policies, like the ones presented in Table 2. In Table 2, the results of optimization problems (1), (2), (3) and (4) are given for all maintenance scenarios (minimal and major, only minimal, only major).

From Table 2 it is obtained that a relatively low inspection interval benefits system availability, especially in the case of performing only minimal or only major maintenance. This is logical since often inspection and thus maintenance prevents a

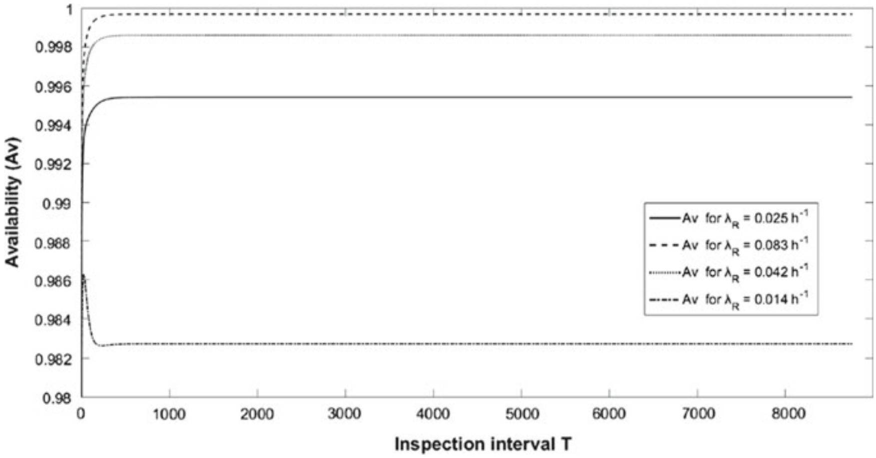


Fig. 16 Effect of the repair rate λ_R on asymptotic availability

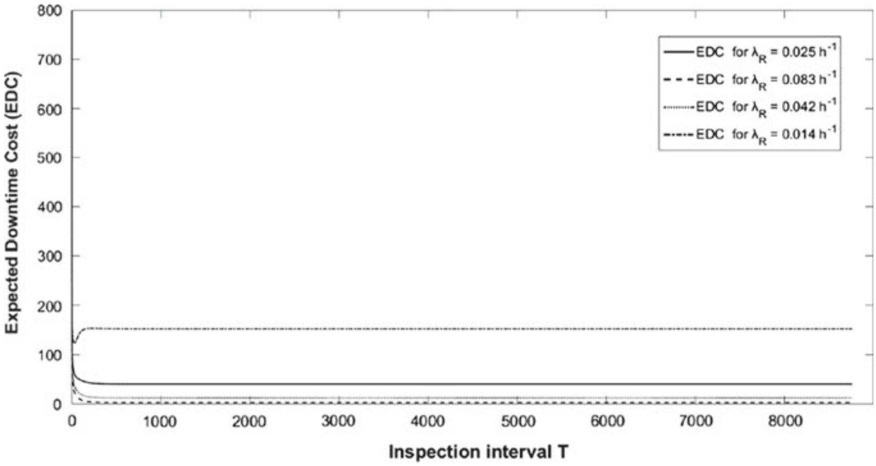


Fig. 17 Effect of the repair rate λ_R on expected downtime

total failure and consequently the system remains available (in an operational state) for longer. On the other hand, the total downtime is optimized for relatively high T, since a high T indicates a delay on entering the maintenance states which incur downtime. The total operational cost though is optimizes when the inspection interval is too short and thus maintenance is implemented more often. This is also obvious

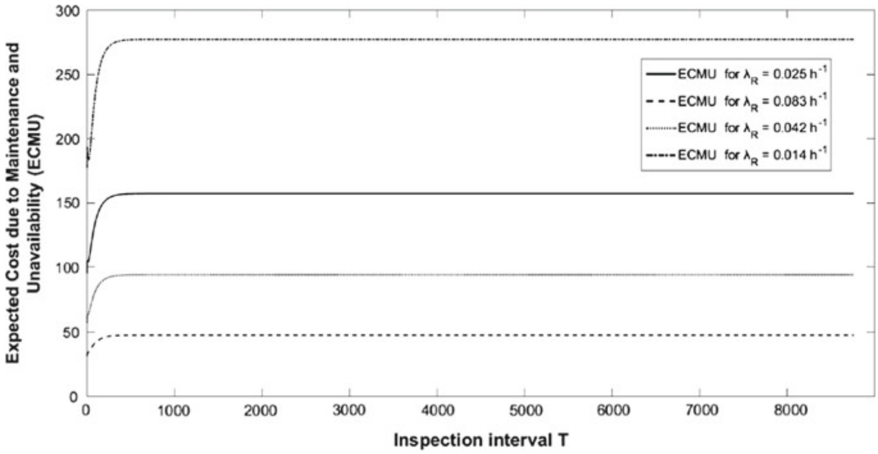


Fig. 18 Effect of the repair rate λ_R on expected cost due to maintenance and unavailability

Table 2 Optimization problems results

Minimal and major maintenance									
Problem (5.1)		Problem (5.2)		Problem (5.3)		Problem (5.4)			
T^*	Av^*	T^*	EDC^*	T^*	$ECMU^*$	Av_0	T^*	$ECMU^*$	Av^*
329.078	0.9953	1127.94	40.18	0.1	94.98	0.9	0.162	95.48	0.9001
						0.99	11.828	103.73	0.9919
						0.995	167.149	148.17	0.995
Minimal maintenance									
18.732	0.9976	886.324	9.65	1.485	40.41	0.9	1.486	40.41	0.9879
						0.99	1.829	40.42	0.99
						0.995	4.145	40.9	0.995
Major maintenance									
33.17	0.9980	1070.25	10.95	0.281	13.45	0.9	0.281	13.45	0.9436
						0.99	1.734	16.52	0.99
						0.995	3.753	21.06	0.95

since frequently performed maintenance prevents total failures which incur a much higher cost. Finally, as far as multi-objective optimization results are concerned, it seems that the optimal T is mostly affected by the operational cost rather than the availability. As it can be obtained, relatively low values of T, which benefit ECMU are given as solution of the multi-objective problems since they can provide availability that satisfies the corresponding constraints.

8 Conclusions and Future Work

In this work a two-unit redundant multi-state deteriorating system under maintenance is modelled and studied. The system consists in two units, one operational and one in a cold standby mode. The deterioration level of the operational unit is identified through an inspection process and when it reaches some predefined levels, the maintenance action to be implemented is decided based on this level. When the operational unit undergoes maintenance or repair, the system control is switched automatically to the standby unit, if the standby unit is operational. Three different scenarios are presented where both minimal and major maintenance can be triggered or solely minimal or solely major maintenance can be performed. The evolution in time of the system is described through a semi-Markov process and its availability, the expected downtime and the expected cost due to maintenance and unavailability are computed under all scenarios.

The aim of this work consists in innovatively present the theoretical and methodological framework of such a redundant multi-state deteriorating system with maintenance. The important issue in such system consists in determining the optimal time to perform any type of maintenance. Thus, optimization problems are formulated and solved with respect to the unit inspection interval, aiming in optimizing systems dependability and performance measures. Apart from modelling such a system, we also take into account the fact that maintenance of any kind can be imperfect of failed as well. To the best of our knowledge, a multi-state deteriorating system with all the aforementioned characteristics is innovatively introduced into the relevant literature. Thus the contribution of this work consists in providing the system designer with all the appropriate modelling and optimization framework in order to schedule maintenance in a manner that can improve not only system availability and downtime but the total operational cost too.

The current work can be extended in the future in many ways. Initially, the assumption of exponentially distributed times can be relaxed and the effect of the chosen distributions on systems availability downtime and cost can be investigated. Additionally, as far as redundancy is concerned, the effects of considering a hot or a warm standby unit can be also investigated. Finally, an important issue which is authors future intentions is to relax the assumption of perfect switching among the two units and examine how an imperfect switch can affect systems dependability and performance.

Appendix

In the next tables the systems states and all the transitions for the proposed two-unit systems are presented analytically. (See Tables 3, 4 and 5).

Table 3 System state and transitions for the two-unit system with minimal and major maintenance actions

No	From state	To state
1	O, S	$(I_0, S), (D^O{}_1, S), (F, O)$
2	$O, D^S{}_1$	$(I_0, D^S{}_1), (D^O{}_1, D^S{}_1), (F, D^O{}_1)$
3	$O, D^S{}_2$	$(I_0, D^S{}_2), (D^O{}_1, D^S{}_2), (F, D^O{}_2)$
4	O, m	$(I_0, m), (D^O{}_1, m), (F, m),$ $(O, S), (O, D^S{}_1), (O, D^S{}_2), (O, F)$
5	O, M	$(I_0, M), (D^O{}_1, M), (F, M),$ $(O, S), (O, D^S{}_1), (O, D^S{}_2), (O, F)$
6	O, F	$(I_0, F), (D^S{}_1, F), (F, F), (O, S)$
7	$D^O{}_1, S$	$(I_1, S), (D^O{}_2, S), (F, O)$
8	$D^O{}_1, D^S{}_1$	$(D^O{}_2, D^S{}_1)\ (I_1, D^S{}_1)\ (F, D^O{}_1)$
9	$D^O{}_1, D^S{}_2$	$(I_1, D^S{}_2), (D^O{}_2, D^S{}_2), (F, D^O{}_2)$
10	$D^O{}_1, m$	$(I_1, m)\ (D^O{}_2, m), (F, m),$ $(D^O{}_1, S), (D^O{}_1, D^S{}_1),), (D^O{}_1, D^S{}_1), (D^O{}_1, F)$
11	$D^O{}_1, M$	$(I_1, m)\ (D^O{}_2, M), (F, M),$ $(D^O{}_1, S), (D^O{}_1, D^S{}_1), (D^O{}_1, D^S{}_1), (D^O{}_1, F)$
12	$D^O{}_1, F$	$(D^O{}_2, F), (I_1, F), (F, F), (D^O{}_1, S)$
13	$D^O{}_2, S$	$(I_2, S), (F, O)$
14	$D^O{}_2, D^S{}_1$	$(I_2, D^S{}_1), (F, D^O{}_1)$
15	$D^O{}_2, D^S{}_2$	$(I_2, D^S{}_2), (F, D^O{}_2)$
16	$D^O{}_2, m$	$(I_2, m), (F, m),$ $(D^O{}_2, S), (D^O{}_2, D^S{}_1), (D^O{}_2, D^S{}_2), (D^O{}_2, F)$
17	$D^O{}_2, M$	$(I_2, M), (F, M),$ $(D^O{}_2, S), (D^O{}_2, D^S{}_1), (D^O{}_2, D^S{}_2), (D^O{}_2, F)$
18	$D^O{}_2, F$	$(I_2, F), (F, F), (D^O{}_2, S)$
19	I_0, S	(O, S)
20	$I_0, D^S{}_1$	$(O, D^S{}_1)$
21	$I_0, D^S{}_2$	$(O, D^S{}_2)$
22	I_0, m	(O, m)
23	I_0, M	(O, M)
24	I_0, F	(O, F)
25	I_1, S	(m, O)
26	$I_1, D^S{}_1$	$(m, D^O{}_1)$
27	$I_1, D^S{}_2$	$(m, D^O{}_2)$
28	I_1, m	(m, m)

(continued)

Table 3 (continued)

No	From state	To state
29	I_1, M	(m, M)
30	I_1, F	(m, F)
31	m, O	(S, O), ($D^S{}_1$, O), ($D^S{}_2$, O), (F, O),
		(m, I_0), (m, $D^O{}_1$), (m, F)
32	m, I_0	(m, O)
33	m, $D^O{}_1$	(S, $D^O{}_1$), ($D^S{}_1$, $D^O{}_1$), ($D^S{}_2$, $D^O{}_1$),
		(F, $D^O{}_1$), (m, I_1), (m, $D^O{}_2$), (m, F)
34	m, I_1	(m, m)
35	m, m	(O, m), ($D^O{}_1$, m), ($D^O{}_2$, m), (F, m),
		(m, O), (m, $D^O{}_1$), (m, $D^O{}_2$), (m, F)
36	m, $D^O{}_2$	(S, $D^O{}_2$), ($D^S{}_1$, $D^O{}_2$), ($D^S{}_2$, $D^O{}_2$),
		(F, $D^O{}_2$), (m, I_2), (m, F)
37	m, I_2	(m, M)
38	m, M	(O, M), ($D^O{}_1$, M), ($D^O{}_2$, M), (F, M),
		(m, O), (m, $D^O{}_1$), (m, $D^O{}_2$), (m, F)
39	m, F	(O, F), ($D^O{}_1$, F), ($D^O{}_2$, F), (F, F), (m, O)
40	M, O	(S, O), ($D^S{}_1$, O), ($D^S{}_2$, O), (F, O),
		(M, $D^O{}_1$), (M, I_1), (M, F)
41	M, I_0	(M, O)
42	M, $D^O{}_1$	(S, $D^O{}_1$), ($D^S{}_1$, $D^O{}_1$), ($D^S{}_2$, $D^O{}_1$), (F, $D^O{}_1$),
		(M, I_1), (M, $D^O{}_2$), (M, F)
43	M, I_1	(M, m)
44	M, $D^O{}_2$	(S, $D^O{}_2$), ($D^S{}_1$, $D^O{}_2$), ($D^S{}_2$, $D^O{}_2$),
		(F, $D^O{}_2$), (M, I_2), (M, F)
45	M, I_2	(M, M)
46	M, m	(O, m), ($D^O{}_1$, m), ($D^O{}_2$, m), (F, m),
		(M, O), (M, $D^O{}_1$), (M, $D^O{}_2$), (M, F)
47	M, M	(O, M), ($D^O{}_1$, M), ($D^O{}_2$, M), (F, M),
		(M, O), (M, $D^O{}_1$), (M, $D^O{}_2$), (M, F)
48	M, F	(O, F), ($D^O{}_1$, F), ($D^O{}_2$, F), (F, F),
		(M, O)
49	I_2, S	(M, O)
50	I_2, $D^S{}_1$	(M, $D^O{}_1$)
51	I_2, $D^S{}_2$	(M, $D^O{}_2$)

(continued)

Table 3 (continued)

No	From state	To state
52	I_2, m	(M, m)
53	I_2, M	(M, M)
54	I_2, F	(M, F)
55	F, O	(S, O),
		(F, $D^O{}_1$), (F, I_0), (F, F)
56	F, $D^O{}_1$	(S, $D^O{}_1$),
		(F, $D^O{}_2$), (F, I_1), (F, F)
57	F, $D^O{}_2$	(S, $D^O{}_2$), (F, F), (F, I_2)
58	F, m	(O, m),
		(F, O), (F, $D^O{}_1$), (F, $D^O{}_2$), (F, F)
59	F, M	(O, M), (F, O), (F, $D^O{}_1$), (F, $D^O{}_2$), (F, F)
60	F, F	(O, F), (F, O)
61	F, I_0	(F, O)
62	F, I_1	(F, m)
63	F, I_2	(F, M)
64	S, O	(S, I_0), (S, $D^O{}_1$), (O, F)
65	S, $D^O{}_1$	(S, $D^O{}_2$), (S, I_1), (O, F)
66	S, $D^O{}_2$	(O, F), (S, I_2)
67	S, I_0	(S, O)
68	S, I_1	(O, m)
69	S, I_2	(O, M)
70	$D^S{}_1$, O	($D^S{}_1$, I_0), ($D^S{}_1$, $D^O{}_1$), ($D^O{}_1$, F)
71	$D^S{}_1$, $D^O{}_1$	($D^S{}_1$, I_1), ($D^S{}_1$, $D^O{}_2$), ($D^O{}_1$, F)
72	$D^S{}_1$, $D^O{}_2$	($D^S{}_1$, I_2), ($D^O{}_1$, F)
73	$D^S{}_1$, I_0	($D^S{}_1$, O)
74	$D^S{}_1$, I_1	($D^O{}_1$, m)
75	$D^S{}_1$, I_2	($D^O{}_1$, M)
76	$D^S{}_2$, O	($D^S{}_2$, I_0), ($D^S{}_2$, $D^O{}_2$),
		($D^O{}_2$, F)
77	$D^S{}_2$, $D^O{}_1$	($D^S{}_2$, I_1), ($D^S{}_2$, $D^O{}_2$),
		($D^O{}_2$, F)
78	$D^S{}_2$, $D^O{}_2$	($D^S{}_2$, I_2), ($D^O{}_2$, F)
79	$D^S{}_2$, I_0	($D^S{}_2$, O)
80	$D^S{}_2$, I_1	($D^O{}_2$, m)
81	$D^S{}_2$, I_2	($D^O{}_2$, M)

Table 4 System state and transitions for the two-unit system with minimal maintenance actions

No	From state	To state
1	O, S	(I_0, S), $(D^O{}_1, S)$, (F, O)
2	O, $D^S{}_1$	$(I_0, D^S{}_1)$, $(D^O{}_1, D^S{}_1)$, $(F, D^O{}_1)$
3	O, $D^S{}_2$	$(I_0, D^S{}_2)$, $(D^O{}_1, D^S{}_2)$, $(F, D^O{}_2)$
4	O, m^1	(I_0, m^1), $(D^O{}_1, m^1)$, (F, m^1), (O, S), $(O, D^S{}_1)$, $(O, D^S{}_2)$, (O, F)
5	O, m^2	(I_0, m^2), $(D^O{}_1, m^2)$, (F, m^2), $(O, D^S{}_1)$, $(O, D^S{}_2)$, (O, F)
6	O, F	(I_0, F), $(D^O{}_1, F)$, (F, F), (O, S)
7	$D^O{}_1$, S	(I_1, S), $(D^O{}_2, S)$, (F, O)
8	$D^O{}_1$, $D^S{}_1$	$(D^O{}_2, D^S{}_1)$, $(I_1, D^S{}_1)$, $(F, D^O{}_1)$
9	$D^S{}_1$, $D^O{}_1$	$(D^S{}_1, D^O{}_2)$, $(D^S{}_1, I_1)$, $(D^O{}_1, F)$
10	$D^O{}_1$, $D^S{}_2$	$(D^O{}_2, D^S{}_2)$, $(I_1, D^S{}_2)$, $(F, D^O{}_2)$
11	$D^S{}_1$, $D^O{}_2$	$(D^O{}_1, F)$, $(D^S{}_1, I_2)$
12	$D^O{}_1$, m^1	(I_1, m^1), $(D^O{}_2, m^1)$, (F, m^1), $(D^O{}_1, S)$, $(D^O{}_1, D^S{}_1)$, $(D^O{}_1, D^S{}_2)$, $(D^0{}_1, F)$
13	$D^O{}_1$, m^2	(I_1, m^2), (D_2, m^2), (F, m^2), $(D^O{}_1, D^S{}_1)$, $(D^O{}_1, D^S{}_2)$, $(D^O{}_1, F)$
14	$D^O{}_1$, F	$(D^O{}_2, F)$, (I_1, F), (F, F), $(D^O{}_1, S)$
15	$D^O{}_2$, S	(I_2, S), (F, O)
16	$D^O{}_2$, $D^S{}_1$	$(I_2, D^S{}_1)$, $(F, D^O{}_1)$
17	$D^S{}_2$, $D^O{}_1$	$(D^S{}_2, I_1)$, $(D^S{}_2, D^O{}_2)$, $(D^O{}_2, F)$
18	$D^O{}_2$, $D^S{}_2$	$(I_2, D^S{}_2)$, $(F, D^O{}_2)$
19	$D^S{}_2$, $D^O{}_2$	$(D^S{}_2, I_2)$, $(D^O{}_2, F)$
20	$D^O{}_2$, m^1	(F, m^1), (I_2, m^1), $(D^O{}_2, S)$, $(D^O{}_2, D^S{}_1)$, $(D^O{}_2, D^S{}_2)$, $(D^O{}_2, F)$
21	$D^O{}_2$, m^2	(F, m^2), (I_2, m^2), $(D^O{}_2, D^S{}_1)$, $(D^O{}_2, D^S{}_2)$, $(D^O{}_2, F)$
22	$D^O{}_2$, F	(F, F), (I_2, F), $(D^O{}_2, S)$
23	I_0, S	(O, S)
24	I_0, $D^S{}_1$	$(O, D^S{}_1)$
25	I_0, $D^S{}_2$	$(O, D^S{}_2)$
26	I_0, m^1	(O, m^1)

(continued)

Table 4 (continued)

No	From state	To state
27	$(m^1, D^O{}_1)$	$(S, D^O{}_1)$, $(D^S{}_1, D^O{}_1)$, $(D^S{}_2, D^O{}_1)$, $(F, D^O{}_1)$,
		$(m^1, D^O{}_2)$, (m^1, I_1), (m^1, F)
28	I_0, m^2	(O, m^2)
29	I_0, F	(O, F)
30	I_1, S	(m^1, O)
31	I_1, $D^S{}_1$	$(m^1, D^O{}_1)$
32	I_1, $D^S{}_2$	$(m^1, D^O{}_2)$
33	I_1, m^1	(m^1, m^1)
34	I_1, m^2	(m^1, m^2)
35	I_1, F	(m^1, F)
36	I_2, S	(m^2, O)
37	I_2, $D^S{}_1$	$(m^2, D^O{}_1)$
38	I_2, $D^S{}_2$	$(m^2, D^O{}_2)$
39	I_2, m^1	(m^2, m^1)
40	I_2, m^2	(m^2, m^2)
41	I_2, F	(m^2, F)
42	m^1, O	(S, O), $(D^S{}_1, O)$, $(D^S{}_2, O)$, (F, O),
		$(m^1, D^O{}_1)$, (m^2, I_0), (m^1, F)
43	m^2, O	$(D^S{}_1, O)$, $(D^S{}_2, O)$, (F, O),
		$(m^2, D^O{}_1)$, (m^2, I_0), (m^2, F)
44	F, O	(S, O), $(F, D^O{}_1)$, (F, I_0), (F, F)
45	F, $D^O{}_1$	$(S, D^O{}_1)$, $(F, D^O{}_2)$, (F, I_1), (F, F)
46	F, $D^O{}_2$	$(S, D^O{}_2)$, (F, F), (F, I_2)
47	F, m^1	(O, m^1),
		(F, O), $(F, D^O{}_1)$, $(F, D^O{}_2)$, (F, F)
48	F, m^2	(O, m^2),
		$(F, D^O{}_1)$, $(F, D^O{}_2)$, (F, F)
49	F, F	(O, F),
		(F, O)
50	S, O	(S, I_0), $(S, D^O{}_1)$, (O, F)
51	$D^S{}_1$, O	$(D^S{}_1, I_0)$, $(D^S{}_1, D^O{}_1)$, $(D^O{}_1, F)$
52	$D^S{}_2$, O	$(D^S{}_2, I_0)$, $(D^S{}_2, D^O{}_1)$, $(D^O{}_2, F)$
53	(m^1, F)	(O, F), $(D^O{}_1, F)$, $(D^O{}_2, F)$, (F, F),
		(m^1, O)
54	(m^1, I_0)	(m^1, O)
55	S, $D^O{}_1$	$(S, D^O{}_2)$, (S, I_1), (O, F)

(continued)

Table 4 (continued)

No	From state	To state
56	m^2, $D^O{}_1$	($D^S{}_1$, $D^O{}_1$), ($D^S{}_2$, $D^O{}_1$),(F, $D^O{}_1$),
		(m^2, $D^O{}_2$), (m^2, I_1), (m^2, F)
57	S, $D^O{}_2$	(S, I_2), (O, F)
58	m^2, $D^O{}_2$	(m^2, I_2), (m^2, F),
		($D^S{}_1$, $D^O{}_2$), ($D^S{}_2$, $D^O{}_2$), (F, $D^O{}_2$)
59	(m^1, $D^O{}_2$)	(S, $D^O{}_2$), ($D^S{}_1$, $D^O{}_2$), ($D^S{}_2$, $D^O{}_2$), (F, $D^O{}_2$),
		(m^1, I_2), (m^1, F)
60	S, I_0	(S, O)
61	$D^S{}_1$, I_0	($D^S{}_1$, O)
62	$D^S{}_2$, I_0	($D^S{}_2$, O)
63	m^2, I_0	(m^2, O)
64	F, I_0	(F, O)
65	S, I_1	(O, m^1)
66	$D^S{}_1$, I_1	($D^O{}_1$, m^1)
67	$D^S{}_2$, I_1	($D^O{}_2$, m^1)
68	m^2, I_1	(m^2, m^1)
69	F, I_1	(F, m^1)
70	S, I_2	(O, m^2)
71	$D^S{}_1$, I_2	($D^O{}_1$, m^2)
72	$D^S{}_2$, I_2	($D^O{}_2$, m^2)
73	m^2, I_2	(m^2, m^2)
74	F, I_2	(F, m^2)
75	m^2, F	(m^2, O),
		(O, F), ($D^O{}_1$, F), ($D^O{}_2$, F), (F, F)
76	(m^2, m^1)	($D^O{}_1$, m^1), ($D^O{}_2$, m^1), (F, m^1),
		(m^2, O), (m^2, $D^O{}_1$), (m^2, $D^O{}_2$), (m^2, F)
77	(m^2, m^2)	($D^O{}_1$, m^2), ($D^O{}_2$, m^2), (F, m^2),
		(m^2, $D^O{}_1$), (m^2, $D^O{}_2$), (m^2, F)
78	(m^1, m^1)	(O, m^1), ($D^O{}_1$, m^1), ($D^O{}_2$, m^1), (F, m^1),
		(m^1, O), (m^1, $D^O{}_1$), (m^1, $D^O{}_2$), (m^1, F)
79	(m^1, m^2)	(O, m^2), ($D^O{}_1$, m^2), ($D^O{}_2$, m^2), (F, m^1),
		(m^1, $D^O{}_1$), (m^1, $D^O{}_2$), (m^1, F)
80	(m^1, I_1)	(m^1, m^1)
81	(m^1, I_2)	(m^1, m^2)

Table 5 System state and transitions for the two-unit system with major maintenance actions

No	From state	To state
1	O, S	$(I_0, S), (D^O{}_1, S), (F, O)$
2	$O, D^S{}_1$	$(I_0, D^S{}_1), (D^O{}_1, D^S{}_1), (F, D^O{}_1)$
3	$O, D^S{}_2$	$(I_0, D^S{}_2), (D^O{}_1, D^S{}_2), (F, D^O{}_2)$
4	O, M	$(I_0, M), (D^O{}_1, M), (F, M)$
		$(O, S), (O, D^S{}_1), (O, D^S{}_2), (O, F)$
5	O, F	$(I_0, F), (D^O{}_1, F), (F, F), (O, S)$
6	$D^O{}_1, S$	$(I_1, S), (D^O{}_2, S), (F, O)$
7	$D^O{}_1, D^S{}_1$	$(D^O{}_2, D^S{}_1), (I_1, D^S{}_1), (F, D^O{}_1)$
8	$D^S{}_1, D^O{}_1$	$(D^S{}_1, D^O{}_2), (D^S{}_1, I_1), (D^O{}_1, F)$
9	$D^O{}_1, D^S{}_2$	$(D^O{}_2, D^S{}_2), (I_1, D^S{}_2), (F, D^O{}_2)$
10	$D^S{}_1, D^O{}_2$	$(D^O{}_1, F), (D^S{}_1, I_2)$
11	$D^O{}_1, M$	$(I_1, M), (D^O{}_2, M), (F, M),$
		$(D^O{}_1, S), (D^O{}_1, D^S{}_1), (D^O{}_1, D^S{}_2), (D^O{}_1, F)$
12	$D^O{}_1, F$	$(D^O{}_2, F), (I_1, F), (F, F), (D^O{}_1, S)$
13	$D^O{}_2, S$	$(I_2, S), (F, O)$
14	$D^O{}_2, D^S{}_1$	$(I_2, D^S{}_1), (F, D^O{}_1)$
15	$D^S{}_2, D^O{}_1$	$(D^S{}_2, I_1), (D^S{}_2, D^O{}_2), (D^O{}_2, F)$
16	$D^O{}_2, D^S{}_2$	$(I_2, D^S{}_2), (F, D^O{}_2)$
17	$D^S{}_2, D^O{}_2$	$(D^S{}_2, I_2), (D^O{}_2, F)$
18	$D^O{}_2, M$	$(F, M), (I_2, M),$
		$(D^O{}_2, S), (D^O{}_2, D^S{}_1), (D^O{}_2, D^S{}_2), (D^O{}_2, F)$
19	$D^O{}_2, F$	$(F, F), (I_2, F), (D^O{}_2, S)$
20	I_0, S	(O, S)
21	$I_0, D^S{}_1$	$(O, D^S{}_1)$
22	$I_0, D^S{}_2$	$(O, D^S{}_2)$
23	I_0, M	(O, M)
24	I_0, F	(O, F)
25	I_1, S	$(D^O{}_1, S)$
26	$I_1, D^S{}_1$	$(D^O{}_1, D^S{}_1)$
27	$I_1, D^S{}_2$	$(D^O{}_1, D^S{}_2)$
28	I_1, M	$(D^O{}_1, M)$
29	I_1, F	$(D^O{}_1, F)$
30	I_2, S	(M, O)
31	$I_2, D^S{}_1$	$(M, D^O{}_1)$
32	$I_2, D^S{}_2$	$(M, D^O{}_2)$
33	$I_2, D^S{}_2$	$(M, D^O{}_2)$
34	I_2, M	(M, M)
35	I_2, F	(M, F)

(continued)

Table 5 (continued)

No	From state	To state
36	M, O	(S, O), ($D^S{}_1$, O), ($D^S{}_2$, O), (F, O),
		(M, $D^O{}_1$), (M, I_0), (M, F)
37	M, $D^O{}_1$	(S, $D^O{}_1$), ($D^S{}_1$, $D^O{}_1$), ($D^S{}_2$, $D^O{}_1$),
		(F, $D^O{}_1$), (M, $D^O{}_2$), (M, I_1), (M, F)
38	M, $D^O{}_2$	(M, I_2), (M, F),
		(S, $D^O{}_2$), ($D^S{}_1$, $D^O{}_2$), ($D^S{}_2$, $D^O{}_2$), (F, $D^O{}_2$)
39	M, I_0	(M, O)
40	M, I_1	(M, $D^O{}_2$)
41	M, I_2	(M, M)
42	M, F	(O, F), ($D^O{}_1$, F), ($D^O{}_2$, F),
		(F, F), (M, O)
43	F, O	(S, O), (F, $D^O{}_1$), (F, I_0), (F, F)
44	F, $D^O{}_1$	(S, $D^O{}_1$), (F, $D^O{}_2$), (F, I_1), (F, F)
45	F, $D^O{}_2$	(S, $D^O{}_2$), (F, F), (F, I_2)
46	F, M	(O, M),
		(F, O), (F, $D^O{}_1$), (F, $D^O{}_2$), (F, F)
47	F, F	(O, F), (F, O)
48	S, O	(S, $D^O{}_1$), (S, I_0), (O, F)
49	S, $D^O{}_1$	(S, $D^O{}_2$), (S, I_1), (O, F)
50	S, $D^O{}_2$	(S, I_2), (O, F)
51	S, I_1	(S, O)
52	S, I_1	(S, $D^O{}_1$)
53	S, I_2	(O, M)
54	$D^S{}_1$, I_0	($D^S{}_1$, O)
55	F, I_0	(F, O)
56	F, I_1	(F, $D^O{}_1$)
57	F, I_2	(F, M)
58	$D^S{}_1$, I_1	($D^S{}_1$, $D^O{}_1$)
59	$D^S{}_2$, I_1	($D^S{}_2$, $D^O{}_1$)
60	$D^S{}_1$, I_2	($D^O{}_1$, M)
61	$D^S{}_2$, I_2	($D^O{}_2$, M)
62	$D^S{}_1$, O	($D^S{}_1$, I_0), ($D^S{}_1$, $D^O{}_1$), ($D^O{}_1$, F)
63	$D^S{}_2$, O	($D^S{}_2$, I_0), ($D^S{}_2$, $D^O{}_1$), ($D^O{}_2$, F)
64	(M, M)	(O, M), ($D^O{}_1$, M), ($D^O{}_2$, M), (F, M),
		(M, O), (M, $D^O{}_1$), (M, $D^O{}_2$),(M, F)

References

1. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
2. Chawla R, Kumar G (2013) Condition based maintenance modeling for availability analysis of a repairable mechanical system. Int J Innov Eng Technol 2(2):371–379
3. Li W, Pham H (2005) Reliability modeling of multi-state degraded systems with multi-competing failures and random shocks. IEEE Trans Reliab 54(2):297–303
4. Shafiee M, Finkelstein M (2015) An optimal age-based group maintenance policy for multi-unit degrading systems. Reliab Eng Syst Saf 134:230–238
5. Murchland J (1975) Fundamental concepts and relations for reliability analysis of multistate systems. In: Barlow RE, Fussell JB, Singpurwalla N (eds) Reliability and fault tree analysis: theoretical and applied aspects of system reliability. SIAM, Philadelphia, pp 581–618
6. El-Neweihi E, Proschan F (1984) Degradable systems: a survey of multistate system theory. Commun Stat Theory Methods 13:405–432
7. Barlow RE, Wu AS (1978) Coherent systems with multi-state components. Math Operat Res 3:275–281
8. Chen D, Cao Y, Trivedi KS, Hong Y (2003) Preventive maintenance of multi-state system with phase-type failure time distribution and non-zero inspection time. Int J Reliab Qual Saf Eng 10(3):323–344
9. Malefaki S, Koutras VP, Platis AN (2014) Optimizing the availability and the operational cost of a periodically inspected multi-state deteriorating system with condition based maintenance policies. In: 2014 Ninth international conference on availability, reliability and security (ARES), 403–411:8–12
10. Platis AN, Koutras VP, Malefaki S (2014) Achieving high availability levels of a deteriorating system by optimizing condition based maintenance policies. In: Steenbergen SMDJM, van Gelder PHAJM, Vrouwenvelder A (eds) Safety, reliability and risk analysis: beyond the Horizon. Taylor & Francis Group, pp 829–837
11. Ross SM (1979) Multivalued state component systems. Ann Prob 7:379–383
12. Aghezzaf E-H, Khatab A (2016) Phuoc Le Tam: optimizing production and imperfect preventive maintenance plannings integration in failure-prone manufacturing systems. Reliab Eng Syst Saf 145:190–198
13. Blischkem WR, Murthy DNPL (2003) Case studies in reliability and maintenance. Wiley, USA
14. Coita DW (2003) Maximization of system reliability with a choice of redundancy strategies. IIE Trans 35(6):535–543
15. Lisnianski A, Ding Y, Frenkel I, Khvatskin L (2007) Maintenance optimization for multi-state aging systems. Proceedings of 5-th international conference on mathematical methods in reliability, methodology and practice, Glasgow, United Kingdom
16. Natvig B (2011) Probabilistic modeling of monitoring and maintenance. In: Multi-state systems reliability theory with applications. Wiley, UK
17. Amari SV, McLaughlin Lm Pham H (2006) Cost-effective condition-based maintenance using markov decision processes, In: Reliability and maintainability symposium 2006 RAMS '06. Annual 464–469:23–26
18. Lapa CMF, Pereira CMNA, de Barros MP (2006) A model for preventive maintenance planning by genetic algorithms based in cost and reliability. Reliab Eng Syst Saf 91(2):233–240
19. Kapur PK, Kapoor KR (1980) Joint optimum preventive maintenance and repair limit replacement policies. IEEE Trans Reliab R-29(3):279–281
20. Naga Srinivasa Rao P, Achutha Naikan VN (2006) A condition-based preventive maintenance policy for Markov deteriorating systems. Int J Perform Model 2(2):175–189
21. Shey SH, Chang CC, Chen YL, Zhe GZ (2015) Optimal preventive maintenance and repair policies for multi-state systems. Reliab Eng Syst Saf 47:78–87
22. Thein S, Chang YS, Makatsoris C (2012) A study of condition based preventive maintenance model for repairable multi stage deteriorating system. Int J Adv Logist 1(1):83–102

23. Ding SH, Kamaruddin S (2015) Maintenance policy optimization-literature review and directions. Int J Adv Manuf Technol 76(5):1263–1283
24. Wie X, Yiguang H, Trivedi KS (2005) Analysis of a two level software rejuvenation policy. Reliab Eng Syst Saf 87(1):13–22
25. Malefaki S, Koutras VP, Platis AN (2016) Multi-State deteriorating system dependability with maintenance using monte carlo simulation. In: Proceedings of SMRLO'16 2016: second international symposium on stochastic models in reliability engineering, Life science and operations management, February 2016, SCE—Shamoon College of Engineering, Beer Sheva, Israel, pp 61–70, IEEE CPS, 978-1-4673-9941-8/16. doi:10.1109/SMRLO.2016.54
26. Sarkar J, Sarkar S (2010) Availability of a periodically inspected system under perfect repair. J Stat Plan Inference 91(1):77–90
27. Chen D, Trivedi KS (2002) Closed-form analytical results for condition-based maintenance. Reliab Eng Syst Saf 76(1):43–51
28. Sim SH (1988) Optimal preventive maintenance with repair. IEEE Trans Reliab 37(1):92–96
29. Chen D, Trivedi KS (2001) Analysis of periodic preventive maintenance with general system failure distribution. In: Proceedings 2001 pacific rim international symposium on dependable computing, 2001. pp 103–107
30. Cheng C-Y, Chen MC (2003) The periodic preventive maintenance policy for deteriorating systems by using improvement factor model. Int J Appl Sci Eng 114–122
31. Sim SH, Endrenyi J (1993) A failure-repair model with minimal and major maintenance. IEEE Trans Reliab 42(1):134–140
32. Pellegrin C (1992) Choice of a periodic on-condition maintenance policy. Int J Prod Res 30(5):1153–1173
33. Kline MB (1984) Suitability of the lognormal distribution for corrective maintenance repair times. Reliab Eng 9(2):65–80
34. Kulkarni VG (1995) Modeling and analysis of stochastic systems. Chapman & Hall, London
35. Limnios N (1997) Dependability analysis of Semi-Markov systems. Reliab Eng Syst Saf 55(3):203–207
36. Limnios N, Oprian G (2001) Semi-Markov processes and reliability. Birkhuser, Boston
37. Taga Y (1963) On the limiting distributions in Markov renewal processes with finitely many states. Ann Inst Statist Math 15(1):1–10
38. Koutras VP, Malefaki S, Platis AN (2017) Optimization of the dependability and performance measures of a generic model for multi-state deteriorating systems under maintenance. Reliab Eng Syst Saf (accepted)
39. Platis AN, Drosakis EG (2009) Coverage modeling and optimal maintenance frequency of an automated restoration mechanism. IEEE Trans Reliab 58(3):470–475

Phase-Type Models and Their Extension to Competing Risks

Bo Henry Lindqvist and Susanne Hodneland Kjølen

Abstract We present an extension of the phase-type methodology for modeling of lifetime distributions to include the case of competing risks. This is done by considering finite state Markov chains in continuous time with more than one absorbing state, letting each absorbing state correspond to a particular risk. The special structure of Coxian phase-type models is considered in particular. The chapter emphasizes the use of phase-type models in statistical modeling and inference for survival and competing risks data.

Keywords Phase-type distribution · Coxian distribution · Competing risks · Identifiability

1 Introduction

Phase-type distributions represent the time to absorption for a finite state Markov chain in continuous time. The simplest examples are mixtures and convolutions of exponential distributions and phase-type distributions have therefore received much attention in applied probability, in particular in queuing theory. Here they generalize the celebrated Erlang distribution. Nowadays, phase-type distributions are applied in various areas such as reliability analysis and medical statistics.

In its generality, the class of phase-type distributions is both flexible and conceptually simple to work with. Interestingly, the class of phase-type distributions is dense in the sense that any lifetime distribution can be approximated arbitrarily close by a phase-type distribution. For a comprehensive introduction to the topic we refer to Neuts [16], while a shorter and very useful introduction is given by Aalen [1].

B.H. Lindqvist (✉)
Norwegian University of Science and Technology, Alfred Getz Vei 1,
7491 Trondheim, Norway
e-mail: bo.lindqvist@ntnu.no

S.H. Kjølen
Visma Software Labs, Karenlyst Allé 56, 0277 Oslo, Norway
e-mail: susanne.kjolen@gmail.com

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_6

The potential usefulness of phase-type distributions in statistical modeling and inference has more recently been revealed in the literature. In statistical applications there seems to be a particular interest in the use of so-called Coxian phase-type models, first suggested by Cox [8]. These are models for phenomena where the units go through stages (phases) in a specified order, and may transit to the absorbing state (corresponding to the event of interest) at any stage. Coxian phase-type models have recently been successfully applied in health care studies ([10, 11, 15, 18]). The problem of fitting more general phase-type distributions to lifetime data has also been considered in the literature, both in a frequentist setting using the EM-algorithm (Asmussen et al. [3]), and in a Bayesian setting using MCMC (Bladt et al. [5]).

The main purpose of the present chapter is to give the necessary tools and results in order to extend the phase-type methodology to include competing risks. The latter concept has been introduced for cases where one in addition to a lifetime have information about the specific cause of failure or death. The classical examples of competing risks consider individuals subjected to multiple causes of death. A famous example is due to David Bernoulli who around 1760 studied the problem of how to disentangle the risk of dying from smallpox from other causes. In cancer research one may consider both the age at onset of cancer and the cancer type. In reliability engineering, one may observe both the time to breakdown of a mechanical component and the root cause, for example vibration or corrosion. An introduction to the theory can be found in, e.g., Lawless [13, Chap. 9].

The basic ingredient in a competing risks phase-type model is a finite state Markov chain in continuous time with more than one absorbing state, where each absorbing state corresponds to a particular risk. Expressions for cause specific hazard functions, cumulative incidence functions etc. can now be given in terms of the transition matrix of the underlying Markov chain. Special structures like Coxian models may still be studied in the competing risks framework. Statistical inference for competing risks using phase-type models is of particular interest in the chapter. This extends approaches in the literature for ordinary phase-type models, and some basic aspects of this extension will be emphasized by studying simple examples involving Coxian models.

2 Phase-Type Distributions

A phase-type distribution can be described in terms of a Markov process $\{X(t); t \geq 0\}$, say, where the system moves through some or all of K transient states, or phases, before moving to a single absorbing state $K + 1$. The time of absorption, T, is then said to have a phase-type distribution. A simple illustration is given in Fig. 1, where $K = 7$ and state $K + 1 = 8$ is absorbing (state 9 will be considered later).

A Coxian phase-type distribution is obtained when all the transitions from the transient states are either from i to $i + 1$ or to the absorbing state $K + 1$, see Fig. 2. The resulting restriction on the permitted transitions is in fact not as strong as it may

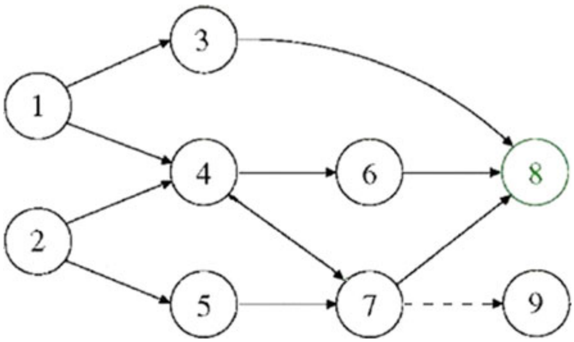


Fig. 1 The state-space and permitted transitions of an absorbing Markov chain, with absorbing state(s) 8 (8 and 9)

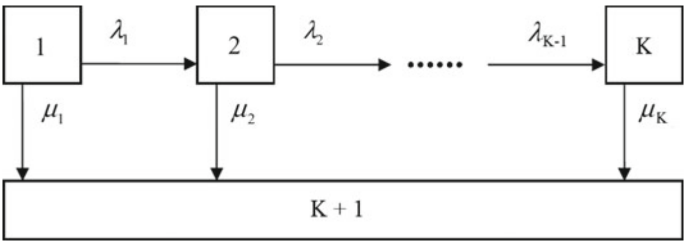


Fig. 2 A coxian phase-type model

look, since any phase-type distribution based on a Markov chain where each move is to a higher numbered state, can be brought on Coxian form (see, e.g., O'Cinneide [17]).

2.1 Model Specification

The infinitesimal transition matrix **A** of the Markov chain producing the phase-type distribution is a $(K+1) \times (K+1)$ matrix given on block form as

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \boldsymbol{\ell} \\ \mathbf{0} & 0 \end{bmatrix}. \quad (1)$$

Here **Q** is the $K \times K$ matrix corresponding to the transitions between the transient states; $\boldsymbol{\ell}$ is the $K \times 1$ vector defining direct transition intensities from the transient states to the absorbing state; while **0** is a $1 \times K$ vector of zeros. Letting $\mathbf{P}(t)$ be the matrix of transition probabilities $P_{ij}(t) = P(X(t) = j|X(0) = i)$ it is well known (e.g., Ross [19, Chap. 5]) that

$$\mathbf{P}(t) = e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \mathbf{A}^i,$$

and it is then straightforward to show that (1) implies

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell} \\ \mathbf{0} & 1 \end{bmatrix}.$$

From this we obtain an expression for the cumulative distribution function of T,

$$F(t) = P(T \leq t) = P(X(t) = K + 1) = \mathbf{p}\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\boldsymbol{\ell}.$$

Here $\mathbf{p}$ is the $1 \times K$-vector with entries $p_i = P(X(0) = i)$ for $i = 1, \ldots, K$, which defines the initial distribution of the Markov chain.

3 Classical Competing Risks

In survival analysis one basically considers the time to failure, T, of a unit. Suppose now that the unit can experience any one of k competing failure causes. Then for each unit one observes both the time to failure, T, and the cause of failure, $C \in \{1, 2, \ldots, k\}$. The pair (T, C) is the observation in the case of competing risks.

In the so called latent failure time approach to competing risks one assumes that the k causes are represented by potential failure times $T_1, \ldots, T_k$, where one only observes the smallest time, $T = \min_j T_j$ and its index $C = \arg\min_j T_j$.

3.1 Distributional Properties of Competing Risks

The joint distribution of the observed pair (T, C) is completely specified by the subdistribution functions and their derivatives, the subdensities,

$$F_j(t) = P(T \leq t, C = j), \quad f_j(t) = F_j'(t).$$

The interpretation of $F_j(t)$ is as the probability of failing from cause j before time t. In biostatistics literature, the $F_j(t)$ are also called cumulative incidence functions.

As an extension of the concept of hazard function of a lifetime distribution, one considers the cause-specific hazard functions,

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P\left(t < T \leq t + \Delta t, C = j | T > t\right)}{\Delta t} = \frac{f_j(t)}{\bar{F}(t)}.$$

The interpretation is that $\lambda_j(t)$ is the failure rate from cause j conditional on survival up to time t.

3.2 The Identifiability Problem of Competing Risks

Consider here the latent failure time approach. The main interest is often in the joint and marginal distributions of the latent failure times $T_1, \ldots, T_k$. The classical problem of competing risks is, however, that the distribution of the observable pair (T, C) in general does not determine the distribution of the latent failure times. In fact, several different joint distributions of $T_1, \ldots, T_k$ will give rise to same distribution of (T, C). This *non-identifiability* property was noted by Cox [9] and formalized by Tsiatis [21]. The main result of Tsiatis is that for a given set of sub-distribution functions $F_j(t)$, there is always a unique (proxy) model with independent T_j yielding these $F_j(t)$.

Biostatisticians have for several decades abandoned the latent failure time approach and claim that statistical conclusions from data only should be based on observable (i.e., identifiable) quantities like the cumulative incidence functions and the cause-specific hazard functions.

4 Phase-Type Models for Competing Risks

As already indicated in Fig. 1, a Markov chain may have more than one absorbing state. In the figure, both states 8 and 9 are absorbing. If we let T be the time of absorption, and C be the identity of the absorbing state, then it is seen that (T, C) is of the form of the observation of a competing risks case.

More generally, consider the general setup of Sect. 2 where the Markov process $\{X(t); t \geq 0\}$ moves among the K transient states before it is absorbed in state $K + 1$. Suppose now instead that there are $m > 1$ absorbing states, named $K + 1, K + 2, \ldots, K + m$, say. Letting T be the time of absorption (in any one of the absorbing states), and letting the cause C represent the state where absorption occurs, by defining $C = K + j$ if $X(T) = K + j$; $j = 1, 2, \ldots, m$, the pair (T, C) can be viewed as an observation from a classical competing risks process with possible causes $K + 1, \ldots, K + m$.

The Coxian phase-type model can now in a straightforward manner be extended to the competing risks case by allowing transitions to any of the m absorbing states $K + 1, \ldots, K + m$ from each of the transient states. The case $m = 2$ is illustrated in Fig. 3.

4.1 Model Specification for Phase-Type Based Competing Risks

By extending the matrix (1) to encompass m absorbing states, we obtain the infinitesimal matrix of the modified Markov process to be the $(K + m) \times (K + m)$ matrix given on block form as

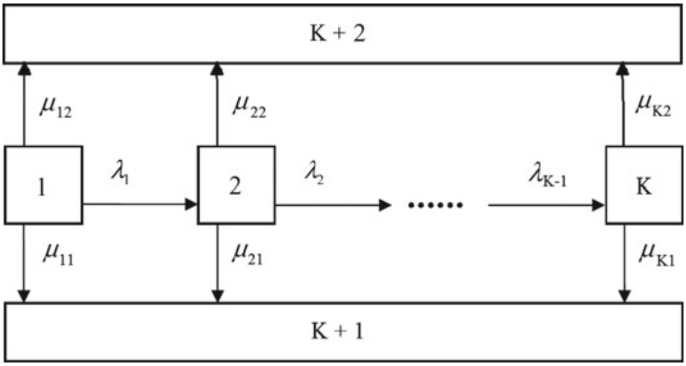


Fig. 3 A coxian phase-type model for $m = 2$ competing risks

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \mathbf{L} \\ \mathbf{0}_1 & \mathbf{0}_2 \end{bmatrix}. \tag{2}$$

As before, $\mathbf{Q}$ is the $K \times K$ matrix corresponding to the transitions between the transient states. The vector $\boldsymbol{\ell}$ is now replaced by the $K \times m$ matrix $\mathbf{L}$ which contains transition intensities from the transient states to the absorbing states. Further, $\mathbf{0}_1$ and $\mathbf{0}_2$ are, respectively, $m \times K$ and $m \times m$ matrices of zeros.

It is rather straightforward to show that (2) implies that the matrix of transition probabilities $P_{ij}(t)$ is given by

$$\mathbf{P}(t) = \begin{bmatrix} e^{\mathbf{Q}t} & \mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\mathbf{L} \\ \mathbf{0}_1 & \mathbf{I} \end{bmatrix}, \tag{3}$$

where $\mathbf{I}$ is the $K \times K$ identity matrix. From (3) we obtain expressions for the subdistribution functions, given by

$$F_j(t) = P(T \le t, C = j) = P(X(t) = j) = \mathbf{p}\mathbf{Q}^{-1}(e^{\mathbf{Q}t} - \mathbf{I})\mathbf{L}\mathbf{v}_j \tag{4}$$

for $j = 1, \ldots, m$. By differentiation we get the subdensities

$$f_j(t) = F_j'(t) = \mathbf{p}e^{\mathbf{Q}t}\mathbf{L}\mathbf{v_j}. \tag{5}$$

In these formulas, $\mathbf{p}$ is the K-vector defining the initial distribution of the Markov chain. It is often natural to assume $p_1 = 1$. Further, $\mathbf{v}_j$ is the m-vector with jth element equal to 1 and the rest equal to 0.

Finally, the cause-specific hazard rate is given by

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(T \le t + \Delta t, C = j | T > t)}{\Delta t} = \frac{F_j'(t)}{P(T > t)} = \frac{\mathbf{p}e^{\mathbf{Q}t}\mathbf{L}\mathbf{v}_j}{\mathbf{p}e^{\mathbf{Q}t}\mathbf{1}_K} \tag{6}$$

(e.g., Braarud [7]). Here $\mathbf{1}_K$ is a K-vector of all 1s.

5 Statistical Inference in Coxian Phase-Type Models

In the present and next section we consider the problem of statistical inference in phase-type models, for data containing survival times only as well as for competing risks data.

Suppose one has a sample of n independent units, where for the ith unit one observes the lifetime T_i and, if applicable, a cause of failure, C_i. The task is to fit phase-type models to the data, where Coxian models will be considered first.

In practice, some of the lifetimes may be censored. Most of the methods to be considered are able to handle censorings, but this problem will not be pursued in the following. This also applies to the inclusion of covariates in the data.

5.1 Coxian Survival Models

Faddy, Graves and Pettitt [10] and McGrory, Pettitt and Faddy [15] considered, respectively, maximum likelihood estimation and Bayesian estimation for Coxian models. In the latter article, the authors used a reversible jump MCMC in their analysis, thus including also K as a parameter in the model. The authors analyze an example dataset comprising lengths of hospital stays of a sample of patients collected from two Australian hospitals to produce a model for a patient's expected length of stay. In particular, posterior distributions for the number of phases and the regression parameters were produced.

In the former article, Faddy et al. [10] considered different variations of Coxian models (Fig. 2), in particular an interesting model assuming $\mu_1 = \mu_2 = \cdots = \mu_{K-3} = 0$ and $\lambda_1 = \lambda_2 = \cdots \lambda_{n-3} = \mu_{n-2} + \lambda_{n-2} = \lambda$. Such a structure corresponds to a gamma-distributed component from the first $n-2$ phases, which makes the model more flexible.

Slud and Suntornchost [20] advocated the use of parametric models based on phase-type distributions with a low number, say 3–8, of parameters. A main conclusion of [20] is that simple phase type models can do almost as well as nonparametric methods, where the latter are commonly the preferred choices in biostatistics and partly in reliability analysis.

Motivated by the above mentioned articles and correspondng conclusions, we shall next show by examples how corresponding statistical analyses can be made with Coxian competing risks models.

5.2 Model 1: Coxian Competing Risks Model with $K = 2$ Transient States and $m = 2$ Absorbing States

This model is illustrated in Fig. 4. The corresponding infinitesimal intensity matrix is

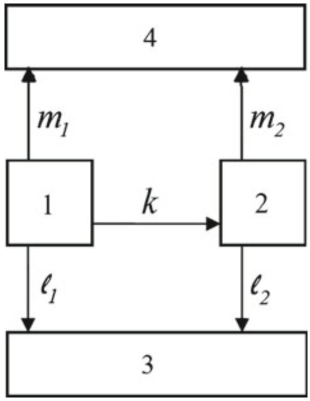


Fig. 4 Model 1

$$\mathbf{A} = \begin{matrix} & 1 & 2 & 3 & 4 \\ 1 & -a & k & l_1 & m_1 \\ 2 & 0 & -b & l_2 & m_2 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \end{matrix},$$

where $a = l_1 + m_1 + k$, $b = l_2 + m_2$. Hence

$$\mathbf{Q} = \begin{bmatrix} -a & k \\ 0 & -b \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} l_1 & m_1 \\ l_2 & m_2 \end{bmatrix}.$$

The subdistribution functions, subdensities and cause-specific hazard functions are found from (4) to (6), and for cause 3 they are, respectively,

$$\begin{aligned} F_3(t) &= \frac{(1 - e^{-at})\, l_1 - \frac{k(e^{-bt} - e^{-at}) l_2}{a-b}}{a} - \frac{k\left(e^{-bt} - 1\right) l_2}{ab} \\ f_3(t) &= \left[l_1 - \frac{k l_2}{a-b} \right] e^{-at} + \frac{k l_2}{a-b} e^{-bt} \\ \lambda_3(t) &= \frac{l_1 (a-b) e^{-at} + k\left(e^{-bt} - e^{-at}\right) l_2}{(a-b)e^{-at} + k\left(e^{-bt} - e^{-at}\right)}. \end{aligned} \tag{7}$$

Note that the formulas are valid only when $a \neq b$. The case $a = b$ follows by taking limits as $b \to a$ (see next subsection). The corresponding formulas for cause 4 are similar, replacing the l_j by m_j.

5.3 Parametric Identifiability of Model 1

Suppose we will use Model 1 in a competing risks case where the pair (T, C) is observed, where $C = 3$ and $C = 4$ represent the absorbing states. In order to estimate the five parameters of the model consistently, we will need the model to be identifiable. This problem will next be considered in detail below, presumably indicating the flavor of the problem also for larger models.

The general identifiability problem of competing risks has been described in Sect. 3.2. The problem is that the underlying probability mechanism is not necessarily identifiable from observations of the pair (T, C). For the present model, the question of identifiability is the following: Does the distribution of the pair (T, C) determine the five parameters of the model, k, l_1, l_2, m_1, m_2?

The functions $f_3(t)$ and $f_4(t)$ are from (7) necessarily given on the form

$$f_3(t) = A_3 e^{-\lambda_1 t} + B_3 e^{-\lambda_2 t}, \;\; f_4(t) = A_4 e^{-\lambda_1 t} + B_4 e^{-\lambda_2 t}.$$

Knowing the distribution of (T, C) means that $\lambda_1, \lambda_2, A_3, B_3, A_4, B_4$ are known to us. We may then without loss of generality assume that $\lambda_1 < \lambda_2$ (the case when they are equal will be treated separately). To identify the parameters of the model, we need to consider two cases, $a < b$ and $a > b$.

Consider first the case $a < b$. By (7) we must have

$$a = \lambda_1, \;\; b = \lambda_2.$$

Further,

$$l_1 - \frac{kl_2}{a-b} = A_3, \;\; \frac{kl_2}{a-b} = B_3, \;\; m_1 - \frac{km_2}{a-b} = A_4, \;\; \frac{km_2}{a-b} = B_4$$

From this it is straightforward to show that the five parameters of the model are uniquely given by

$$\begin{aligned} l_1 &= A_3 + B_3, \qquad l_2 = \frac{(\lambda_1 - \lambda_2)B_3}{\lambda_1 - A_3 - B_3 - A_4 - B_4}, \\ m_1 &= A_4 + B_4, \qquad m_2 = \frac{(\lambda_1 - \lambda_2)B_4}{\lambda_1 - A_3 - B_3 - A_4 - B_4}, \\ k &= \lambda_1 - A_3 - B_3 - A_4 - B_4. \end{aligned}$$

Suppose then that $a > b$. Then obviously $a = \lambda_2$ and $b = \lambda_1$ and it can be shown that the five parameters are uniquely given by

Table 1 Two different sets of parameter values corresponding to the subdensities (8)–(9)

Assumption	k	l_1	m_1	l_2	m_2
$a < b$	1	2	1	3	2
$a > b$	2	2	1	5/2	3/2

$$l_1 = A_3 + B_3, \qquad l_2 = \frac{(\lambda_2 - \lambda_1)A_3}{\lambda_2 - A_3 - B_3 - A_4 - B_4},$$
$$m_1 = A_4 + B_4, \qquad m_2 = \frac{(\lambda_2 - \lambda_1)A_4}{\lambda_2 - A_3 - B_3 - A_4 - B_4},$$
$$k = \lambda_2 - A_3 - B_3 - A_4 - B_4.$$

As an example, suppose we have "observed" that

$$f_3(t) = 5e^{-4t} - 3e^{-5t}, \tag{8}$$
$$f_4(t) = 3e^{-4t} - 2e^{-5t}. \tag{9}$$

By using the above results we conclude that there are exactly two different sets of parameters that give the above functions $f_3(t)$ and $f_4(t)$, see Table 1.

If $a = b$ in Model 1, then it follows by taking the limit as $b \to a$ in (7) that

$$f_3(t) = l_1 e^{-at} + kl_2 te^{-at}, \quad f_4(t) = m_1 e^{-at} + km_2 te^{-at}.$$

The observed $f_2(t)$ and $f_3(t)$ are hence necessarily of the form

$$f_3(t) = C_3 e^{-\lambda t} + D_3 te^{-\lambda t}, \quad f_4(t) = C_4 e^{-\lambda t} + D_4 te^{-\lambda t}.$$

It follows immediately that $l_1 = C_3$ and $m_1 = C_4$ and from this that $l_2, m_2.k$ are uniquely given as well.

5.4 *Identifiability of Coxian Phase-Type Models*

It follows from the above that the parameters of Model 1 are identifiable in the case when $a = b$, but that, in the case where $a \neq b$, an additional assumption on the relative size of a and b has to be made as part of the prior specification of the model.

In a practical application of Model 1 we might assume that state 2 involves a more severe condition for the unit than state 1. As a result of this, the transition rates to the absorbing states are expected to be higher from state 2 than from state 1. In this case it might be reasonable to assume that $a < b$, in which case we have an identifiable model.

For Model 1, it is seen that the eigenvalues of Q determine the exponents of the exponentials in the expressions for the subdensity functions $f_3(t)$ and $f_4(t)$. More generally, considering the general Coxian competing risks model in Fig. 3, it is seen that the eigenvalues of $\mathbf{Q}$ are

$$\rho_i = \mu_{i1} + \mu_{i2} + \lambda_i$$

for $i = 1, 2, \ldots, K$, where by convention we put $\lambda_K = 0$. It follows that the subdensity function f_{K+1} is of the form

$$f_{K+1}(t) = \sum_{i=1}^{K} A_i e^{-\rho_i t}$$

provided the ρ_i are all different, and similarly for $f_{K+2}(t)$. If, say, ρ_i has multiplicity $r > 1$, then terms of the form $Ct^j e^{-\rho_i t}$ for $j = 1, 2, \ldots, r-1$ are included in the functions. Motivated by the study of Model 1, we may in general have to consider all permutations of the ρ_i in order to check identifiability.

We close the discussion on identifiability by noting that identifiability problems may also occur in ordinary Coxian models with a single absorbing state. To be explicit, consider Fig. 4 and assume that $m_1 = m_2 = 0$. Then $f_3(t)$ in (7) is simply the density function of the time T to absorption in state 3. We will now show by an example that two different parameterizations can give rise to the same density function. Namely, let (l_1, l_2, k) be given by either $(2, 4, 1)$ or $(2, 3, 2)$. In both cases we obtain the density function of T equal to $f(t) = 6e^{-3t} - 4e^{-4t}$. Hence the model is not identifiable from data on lifetimes T. Again, a prior choice of whether $l_1 + k$ is greater or smaller than l_2 has to be made.

5.5 *Case Study: Pneumonia on Admission to Intensive Care Unit ([4])*

Beyersmann, Allignol and Schuhmacher [4] present data for 747 patients at an intensive care unit, where the purpose is to examine the effect of hospital-acquired infections. The data set contains information on pneumonia status on admission, time of intensive care unit stay and 'intensive care unit outcome', either hospital death or alive discharge.

In order to increase flexibility compared to Model 1 we build on the earlier mentioned idea of Faddy et al. [10] by adding two states. The model is presented in Fig. 5, and we shall denote it by Model 2. We may think of the extended model as adding to the waiting time in state 3 a gamma-distributed length of time. Note that the approach of [10] would make the additional assumption that $l_1 + m_1 + k_1 = k_0$. In this case the total waiting time in state 3 is gamma-distributed.

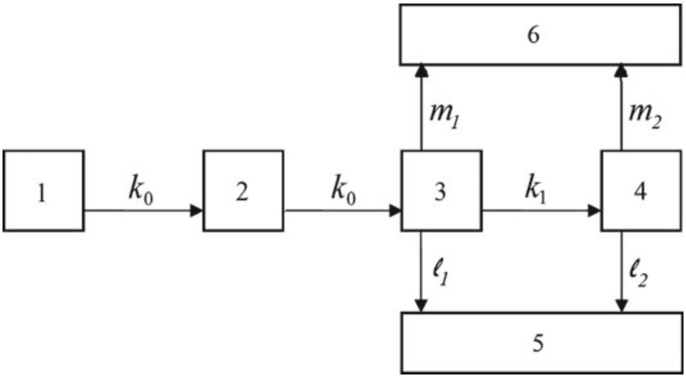


Fig. 5 Model 2

Table 2 Estimated parameters from Model 2 for the pneumonia data of [4]

	k_0	k_1	l_1	l_2	m_1	m_2
Pneumonia	0.079642	0.064075	0.102990	0.015845	0.479752	0.000569
No pneumonia	0.749985	0.072420	0.009248	0.008576	0.155579	0.047151

We made separate analyses for patients with and without pneumonia at admission. The absorbing states 5 and 6 correspond to, respectively, hospital death and discharge from hospital. The estimates of the parameters are given with one line for each analysis in Table 2.

In order to evaluate the results, we present in Fig. 6 plots of the cumulative incidence functions obtained from Model 2, together with nonparametric estimates found by using the Aalen-Johansen estimators (see, e.g., Borgan [6]). The parametric estimates seem to perform very well, a conclusion which confirms the findings and suggestions of [20] as reported earlier.

6 Statistical Inference for General Phase-Type Distributions

Asmussen, Nerman and Olsson [3] presented a general approach to estimation of phase-type distributions from lifetime data. Their idea was to consider the class of phase-type distributions, for a fixed K, as a multi-parameter exponential family. Since one then obviously is in the setting of incomplete observations, they suggested to implement the EM algorithm. Lindqvist [14] gave some details on how to extend the approach of [3] to the competing risks case.

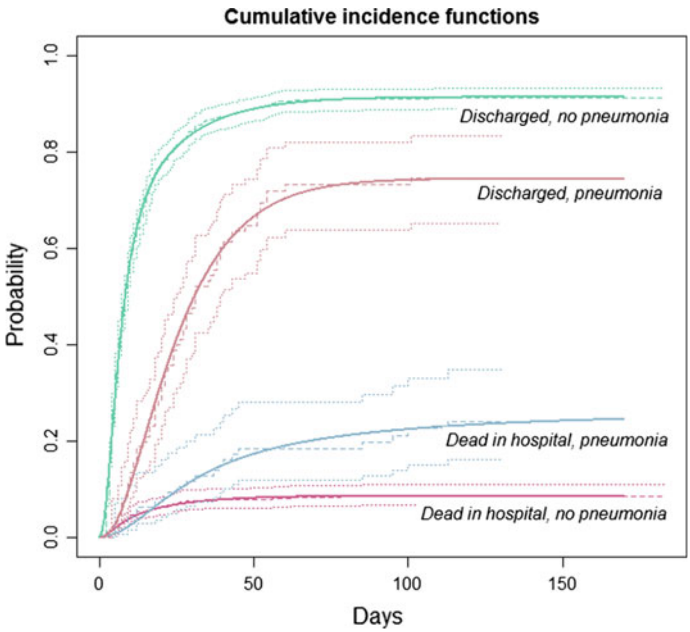


Fig. 6 Estimated cumulative incidence functions for the pneumonia-data of [4]. The smooth estimates are based on model 2 and are compared to nonparametric estimates with corresponding pointwise confidence intervals for each curve

Bladt, Gonzalez and Lauritzen [5] considered Bayesian estimation of phase-type distributions, constructing a Gibbs sampler which draws phase-type parameters from their posterior distribution. They reported as a main advantage of their method, that the uncertainty of estimates of complex functionals of the phase-type distributions could easily be obtained. It is not so clear, on the other hand, how to do this for the EM-algorithm approach. Aslett and Wilson [2] have improved on the method of Bladt et al. [5], and also provide an R-package for practical computation. The approach of Bladt et al. [5] was extended to the competing risks case by Laache [12].

In practice, lifetime data will typically include measured covariates. Most of the methods considered above can be extended to this case, and in some cases this is already a feature of the methods (e.g., [10] and [15]). Lindqvist [14] presented some ideas on a general approach for inclusion of covariates.

References

1. Aalen OO (1995) Phase type distributions in survival analysis. Scandinavian J Statistics 22(4):447–463
2. Aslett, LJM, Wilson SP (2011) Markov chain Monte Carlo for inference on phasetype models. ISI

3. Asmussen S, Nerman O, Olsson M (1996) Fitting phase-type distributions via the EM algorithm. Scandinavian J Statistics 23(4):419–441
4. Beyersmann J, Allignol A, Schumacher M (2011) Competing risks and multistate models with R. Springer Science & Business Media
5. Bladt M, Gonzalez A, Lauritzen SL (2003) The estimation of phase-type related functionals using Markov chain Monte Carlo methods. Scandinavian Actuarial J 2003(4):280–300
6. Borgan O (1998) Three contributions to the encyclopedia of biostatistics: The Nelson-Aalen. Kaplan-Meier and Aalen-Johansen estimators, Encyclopedia of Biostatistics
7. Braarud IH (2012) Phase type and competing risk models in survival analysis. Department of Mathematical Sciences, Norwegian University of Science and Technology, Project work
8. Cox DR (1955) A use of complex probabilities in the theory of stochastic processes. In: Mathematical proceedings of the Cambridge philosophical society vol 51, No 02. Cambridge University Press, pp 313–319
9. Cox DR (1959) The analysis of exponentially distributed life-times with two types of failure. J Royal Stat Soc Series B (Methodological) 21:411–421
10. Faddy MJ, Graves N, Pettitt AN (2009) Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and lognormal distributions. Value Health 12:309–314
11. Gillespie J, McClean S, Scotney B, Garg L, Barton M, Fullerton K (2011) Costing hospital resources for stroke patients using phase-type models. Health Care Manag Sci 14(3):279–291
12. Laache CH (2014) Phase-type inference on competing risks models with covariates, using MCMC methods. Master's thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology
13. Lawless JF (2003) Statistical models and methods for lifetime data. Wiley-Interscience, Hoboken, NJ
14. Lindqvist, B. (2016). Phase-type models for competing risks. In: Frenkel I, Lisnianski A (eds) Proceedings of the second international symposium on stochastic models in reliability engineering, life science and operations management (SMRLO16). Beer Sheva, Israel, pp 37–40, 15–18 February 2016, IEEE CPS, 978-1-4673-9941-8/16. doi:10.1109/SMRLO.2016.17
15. McGrory CA, Pettitt AN, Faddy MJ (2009) A fully Bayesian approach to inference for Coxian phase-type distributions with covariate dependent mean. Comput Stat Data Anal 53(12):4311–4321
16. Neuts MF (1981) Matrix-geometric solutions in stochastic models. Johns Hopkins University Press, Baltimore/London
17. O'Cinneide CA (1989) On non-uniqueness of representations of phase type distributions. Commun Stat—Stoch Models 5:247–259
18. Payne K, Marshall AH, Cairns KJ (2011) Investigating the efficiency of fitting Coxian phase-type distributions to health care data. IMA J Manag Math. doi:10.1093/imaman/dpr008
19. Ross SM (1983) Stochastic processes. Wiley, New York
20. Slud EV, Suntornchost J (2014) Parametric survival densities from phase-type models. Lifetime Data Anal 20(3):459–480
21. Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci 72(1):20–22

A Study on Repairable Series Systems with Markov Repairable Units

He Yi and Lirong Cui

Abstract Consider a repairable series system consisting of n units and each of them follows a Markov process with finite state and continuous time. Under independence assumption among units, the repairable series system has been widely studied by using the Markov process method and *Lz*-transform method. However, both methods have faced the problem of state exploration although some approximation methods have been used. Thus, it is still an interesting and significant problem to be explored. In this chapter, we investigate repairable series systems by using matrix method which has been widely used in aggregated stochastic processes especially in ion channel modeling and aggregated repairable systems. The formulas for reliability, instantaneous and interval availabilities are given in matrix form for four kinds of repairable series systems, general repairable series system, general repairable series system with neglected failures, phased-mission repairable series system and phased-mission repairable series system with neglected failures, respectively. Numerical examples are shown to illustrate the results for the four kinds of systems and present how matrix method is used to solve the problem of state exploration in the *Lz*-transform method. Finally, the conclusions and some future possible applications are given.

Keywords Repairable series system · Markov process · Aggregated stochastic processes · Neglected failures · Phased-mission

---

H. Yi
School of Mathematics & Statistics, Beijing Institute of Technology,
Beijing, People's Republic of China
e-mail: 1215026334@qq.com

L. Cui (✉)
School of Management and Economics, Beijing Institute of Technology,
Beijing, People's Republic of China
e-mail: lirongcui@bit.edu.cn

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_7

1 Introduction

Multi-state systems (MSSs) have been widely studied because components in real word often have more states than up and down. When describing the behaviors of MSSs, traditional Markov method has a problem of state exploration in solving corresponding differential functions and a new method named as universal generating function (UGF) was proposed by Ushakov [1]. UGF method investigated behaviors of the components by Markov method independently and then system behaviors were represented by Ushakov's universal generating operator (UGO). Levitin [2] comprehensively investigated the UGF technique and applied it to many different kinds of binary systems and MSSs.

UGF technique can only be used to study steady-state performance of MSSs theoretically because generating function is defined only for random variables. Then *Lz*-transform was introduced by Lisnianski [3] to analysis dynamic MSS reliability for discrete-state continuous-time Markov process. Frenkel et al. [4] applied *Lz*-transform to a real case study of aging refrigeration system to prove the efficiency of the method. Based on *Lz*-transform, Lisnianski et al. [5] used Ushakov's UGO to reduce computational burden in reliability importance evaluation for components. Based on inverse *Lz*-transform, Lisnianski and Ding [6] developed a special method to calculate reliability function of multi-state power system.

MSSs also take an important role in research of ion channel theory, and an efficient matrix method was first proposed [7] and refined [8] by Colquhoun and Hawkes to describe performance of aggregated Markov repairable system. Ball et al. [9] generalized the matrix method for aggregated semi-Markov repairable system and some new results of ion channel behavior were given. Cui et al. [10] introduced a Markov repairable system with changeable state whose performance depends on the immediately preceding state and several reliability indexes were calculated by matrix method. Hawkes et al. [11] presented a Markov repairable system under alternating regimes and availability of the system and some distributions were given by using matrix method. In this chapter, matrix method was used to describe behaviors of components in MSSs in place of Markov method in *Lz*-transform which simplified *Lz*-transform method by aggregation of states and solved the problem of state exploration to some extent.

In practice, brief events are often too short to be detected because of limitation of devices and they always have ignorable effects on performance of MSSs. For example, performance data of an aircraft engine system are often recorded by several sensors in it periodically, maybe ten times per second. Obviously, failure (working) times less than the period can be recorded at most as an outlier point and they often have ignorable effects on performance of the system in such a short time. This kind of problem which is called time interval omission has been discussed for many kinds of MSSs. Hawkes et al. [12] first derived exact distribution of open times and shut times for single ion channel with assumption of Markov process and time interval omission. Ball [13] further considered number of bursts in single ion channel for Markov and semi-Markov process with time interval omission and properties of two measures were given to describe degree of temporal clustering of bursts. Zheng et al. [14] discussed a single unit Markov repairable system with repair time interval omission and availability of the system was given for both constant and random critical repair time. Bao and Cui [15] studied a series Markov system with failures neglected and delayed, and instantaneous availability and steady availability were presented as reliability measures.

In some cases, systems have different missions and performance during different periods. For example, ion drive system of space shuttles, which is made up of peripheral processor unit (PPU), ion engines and propel valves, is a phased-mission system that have identical system structure but different failure modes and failure rates in different phases. Warships use different systems or devices for different missions, which means the whole warship system is a phased-mission system that have different system structures and different failure modes. Early in 1958, Quandt [16] considered estimation of parameters for a two phased linear regression system, especially the switching point. Later, Goldfeld and Quandt [17] further investigated a Markov model for switching regressions. This kind of phased mission problems have been widely studied [11, 18, 28]. Recently, Shen et al. [19] considered a dynamic system operating under a cycle of k continuous-time Markov processes and availability and distribution of the first uptime were derived. Li et al. [20] discussed non-repairable cyclic phased-mission systems with multiple-failure modes for semi-Markov system and some reliability measures were given.

In the field of reliability, availability is an important measure to describe performance of systems. Since single point availability is not enough for practical requirements, joint availability was first proposed by Baxter [21] to describe availability at multi-point. Interval availability was first proposed by Sericola [22] according to definition of guaranteed availability [23, 24] for computer systems. Csenki investigated joint availability [25] for semi-Markov processes by integral equations and derived closed form of joint interval availability [26] for Markov processes by induction method. Cui et al. [27] further considered mixed multi-point-interval availabilities for Markov repairable system and some properties of them were given.

However, as one of the most popular methods in the research of MSSs, *Lz*-transform method hasn't consider problem of time interval omission and phased-mission, and interval availability which is a very important reliability measure can't be given by *Lz*-transform. In this chapter, four kinds of repairable series systems will be investigated by matrix method and their reliability, instantaneous availability and interval availability will be given, respectively. To illustrate how these reliability measures are calculated, some numerical examples are given in this chapter.

The first system in this chapter is general repairable series system which works if all its units work and fails once one of its units fails. The second system is general repairable series system with neglected failures which works if failures of all its units are not found and fails once a failure of its units is found. The third system is phased-mission repairable series system, which works if the unit in use works and fails once the unit in use fails. The fourth system is phased-mission repairable series system with neglected failures, which works if failures of the unit in use are not found and fails once a failure of the unit in use is found. These systems widely exist in practice since time interval omission and phased-mission are very common problems for repairable series systems.

To illustrate how matrix method solves the problem of state exploration, a magnet resonance inspection (MRI) water cooling system in [5] is considered. As mentioned in [5], the water cooling system is used to remove heat from MRI scanners to guarantee their efficiency. The water cooling system can be described as a repairable series system consisting of four independent units—chillers subsystem, heat exchanger subsystem, pump subsystem and electrical board subsystem, and each of the system follows a Markov process.

With too many states, it's very complex to get reliability of the water cooling system by *Lz*-transform [6]. First, Ushakov's UGO should be used four times to get *Lz*-transforms of the four units. Next, Ushakov's UGO should be used again to get *Lz*-transform of the entire system in which $3\times4\times2\times3=72$ terms should be considered. Then genetic algorithm should be used to generate a matrix with dimension 5×5 in which 20 unknown elements should be given. Finally, the reliability is given by solving a set of five differential functions. However, when calculating reliability of the system by matrix method, we only need to get reliabilities of the units by simple matrix calculation and then reliability of the system is given by their product. By aggregation of states, MSSs are simplified to binary state system that is how matrix method solved the problem of state exploration to some extent.

The rest of this paper is organized as follows. Model assumptions and definitions of the four kinds of repairable series systems are given in Sect. 2. Reliability, instantaneous availability and interval availability are calculated for the four systems in Sects. 3, 4, 5 and 6, respectively. Numerical examples are shown in Sect. 7 to illustrate results of the four systems and present how matrix method is used to solve the problem of state exploration. Finally, some conclusions are presented in Sect. 8.

2 Model Assumptions

Consider a repairable series system consisting of n independent units and each of them follows a Markov process $\{X_i(t), t\geq 0\}$ with continuous time and finite state space. State spaces $S_i(i=1,\ldots,n)$ of the n units can be divided into two subspaces, respectively, that is $S_i=W_i\cup F_i$, $(i=1,\ldots,n)$, W_i for working states and F_i for failure states. Then infinitesimal generator $\mathbf{Q}^{(i)}(i=1,\ldots,n)$ of each unit can be divided into $\begin{pmatrix} \mathbf{Q}^{(i)}_{W_iW_i} & \mathbf{Q}^{(i)}_{W_iF_i} \\ \mathbf{Q}^{(i)}_{F_iW_i} & \mathbf{Q}^{(i)}_{F_iF_i} \end{pmatrix}$, respectively. Assume that all the units are in their best working states at the beginning and $\boldsymbol{\pi}^{(i)}_{W_i}$ is the initial probability vector of unit

i with dimension $1 \times |W_i|$. In this chapter, four kinds of repairable series (general repairable series system, general repairable series system with neglected failures, phased-mission repairable series system and phased-mission repairable series system with neglected failures) are defined, and their reliability, instantaneous availability and interval availability are given, respectively.

Definition 1 The first system—general repairable series system is a repairable system which works if all its units work and fails once one of its units fails.

Remark For example, if the system has three units and unit 1 has two states, $S_1 = \{1, 2\}$, and $W_1 = \{1\}, F_1 = \{2\}$, unit 2 has two states, $S_2 = \{3, 4\}$, and $W_2 = \{3\}, F_2 = \{4\}$, and unit 3 has three states, $S_3 = \{5, 6, 7\}$, and $W_3 = \{5\}, F_3 = \{6, 7\}$, where state 6 and state 7 mean different failure modes or different performance levels lower than need, then possible paths for the first system can be shown in Fig. 1. As shown in Fig. 1, the system is working until unit 2 fails. Before unit 2 is repaired, unit 1 fails, and before unit 1 is repaired, unit 3 fails. Then the system begins to work when unit 3 is repaired because at that time the other two units all work. This system is a general series system whose units follow Markov processes with continuous time and finite state space.

Definition 2 The second system—general repairable series system with neglected failures is a repairable system which works if failures of all its units are not found and fails once a failure of its units is found. In this system, assume that failures of unit i can be found after a critical time τ_i which means failure time being less than τ_i can never be found for unit i and failure time being longer than τ_i can be found after it happens for unit i.

Remark For the same example, paths for the second system are shown in Fig. 2. As shown in Fig. 2, a dotted line at the left endpoint of a double arrow and a solid one

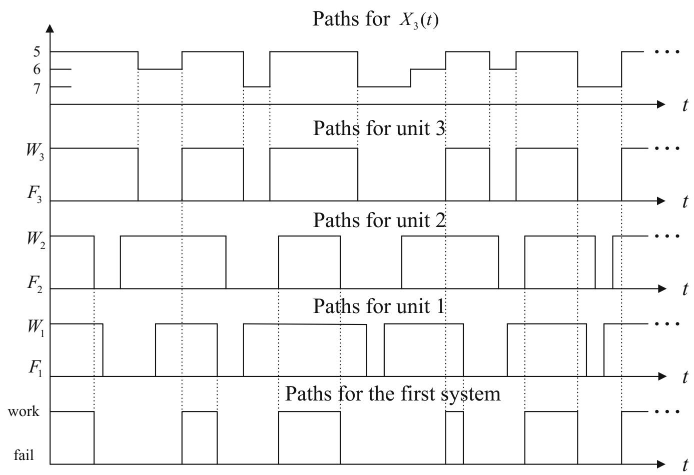


Fig. 1 Paths for the first system

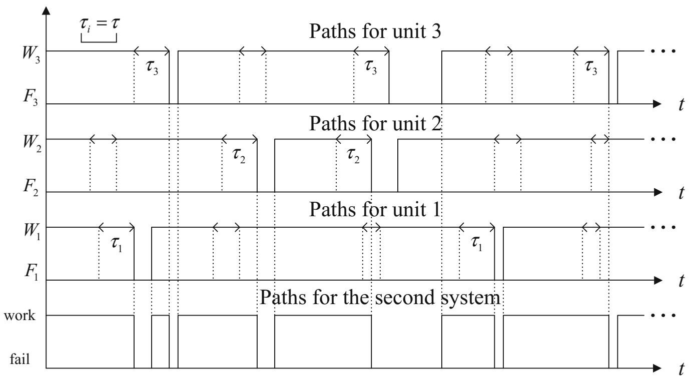


Fig. 2 Paths for the second system

at its right endpoint mean the time a failure happens and the time it is found, respectively, and two dotted lines means the failure is too short to be found. The first failure of the unit 1 is longer than τ_1 , so it's found τ_1 after it happens. The first failure of unit 2 is less than τ_2, so it can never be found. This system is a general series system with time interval omission problem considered for failures and its only difference with the first system is that failures are all delayed which causes some short failures never be found.

Definition 3 The third system—phased-mission repairable series system is a repairable system which works if the unit in use works and fails once the unit in use fails. In this system, assume that unit i is in use during $\left[mT + \sum_{j=1}^{i-1} T_j,\ mT + \sum_{j=1}^{i} T_j\right]$, $m = 0, 1, \ldots, \infty$, where $T = T_1 + T_2 + \cdots + T_n$.

Remark For the same example, paths for the third system are shown in Fig. 3. As shown in Fig. 3, in every period, the system uses unit 1 first for T_1, and then unit 2 for T_2, unit 3 for T_3 and so on. In the first phase, unit 1 is in use and it fails for one time, so the system fails for one time during $[0, T_1]$. This system is a series system with phased-mission problem considered and it uses the units periodically which means in each phase, performance of the system is decided only by performance of the unit in use.

Definition 4 The fourth system—phased-mission repairable series system with neglected failures is a repairable system which works if failures of the unit in use are not found and fails once a failure of the unit in use is found. In this system, assume that unit i is in use during $\left[mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j\right]$, $m = 0, 1, \ldots, \infty$, and failures of unit i can be found after a critical time τ_i which means failure time being less

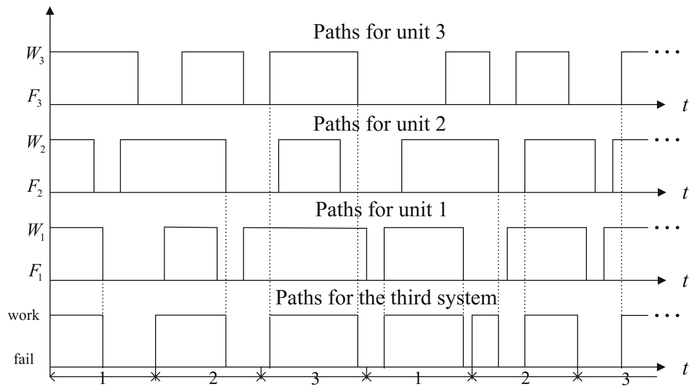


Fig. 3 Paths for the third system

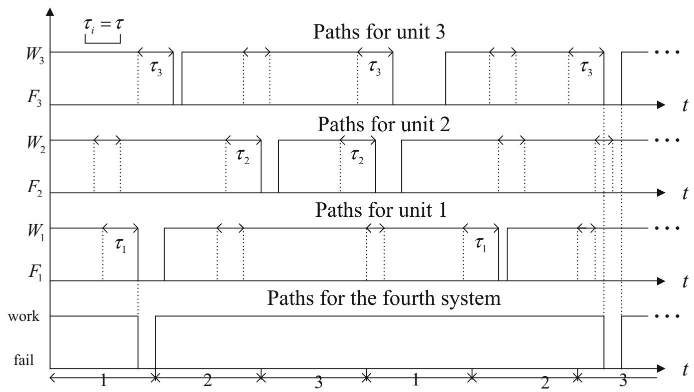


Fig. 4 Paths for the fourth system

than τ_i can never be found for unit i and failure time being longer than τ_i can be found after it happens for unit i.

Remark For the same example, paths for the fourth system are shown in Fig. 4. As shown in Fig. 4, in the first phase, unit 1 is in use and its first failure is found, so the system fails for one time during $[0, T_1]$. In the second phase, unit 2 is in use and its failures are not found, so the system doesn't fail during $[T_1, T_1 + T_2]$. This system is a series system with both phased-mission problem and time interval omission problem considered and its only difference with the third system is that failures are all delayed which causes some short failures never be found.

3 General Repairable Series System

Theorem 1 Reliability of general repairable series system can be given as follows,

$$R(t) = \prod_{i=1}^{n} R_i(t),$$

where $R_i(t) = \boldsymbol{\pi}_{W_i}^{(i)} e^{\mathbf{Q}_{W_iW_i}^{(i)} t} \mathbf{u}_{W_i}^{(i)}$.

Proof Denote reliability of unit i as $R_i(t)$ $(i = 1, 2, \ldots, n)$, respectively. For general repairable series system, it's easy to find $R(t) = \prod_{i=1}^{n} R_i(t)$ under independent assumption. From [8], we have, $R_i(t) = \boldsymbol{\pi}_{W_i}^{(i)} \mathbf{P}_{W_iW_i}^{(i)}(t) \mathbf{u}_{W_i}^{(i)} = \boldsymbol{\pi}_{W_i}^{(i)} e^{\mathbf{Q}_{W_iW_i}^{(i)} t} \mathbf{u}_{W_i}^{(i)}$, where $\boldsymbol{\pi}_{W_i}^{(i)}$ is the initial probability vector of unit i with dimension $1 \times |W_i|$, $\mathbf{P}_{W_iW_i}^{(i)}(t) = e^{\mathbf{Q}_{W_iW_i}^{(i)} t}$ is a matrix with dimension $|W_i| \times |W_i|$ whose element $p_{lk}^{(i)}(t), l, k \in W_i$ is the probability that unit i is working during $[0, t]$ and is in state k at time t condition on initial state l at time 0 and $\mathbf{u}_{W_i}^{(i)}$ is a vector with dimension $|W_i| \times 1$ whose elements are all 1.

□

Theorem 2 Instantaneous availability of general repairable series system can be given as follows,

$$A(t) = \prod_{i=1}^{n} A_i(t),$$

where $A_i^*(s) = \boldsymbol{\pi}_{W_i}^{(i)} [s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)} - \mathbf{Q}_{W_iF_i}^{(i)} (s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)})^{-1} \mathbf{Q}_{F_iW_i}^{(i)}]^{-1} \mathbf{u}_{W_i}^{(i)}$.

Proof Denote instantaneous availability of unit i as $A_i(t)$ $(i = 1, 2, \ldots, n)$, respectively, and then it's easy to find that the instantaneous availability of the system is $A(t) = \prod_{i=1}^{n} A_i(t)$. Denote the conditional probability $P(X_i(t) = k | X_i(0) = l)$ as $T_{lk}^{(i)}(t)$ where $X_i(0) = l$ means unit i is in state l at time 0 and $X_i(t) = k$ means unit i is state k at time t. Then the matrix $\mathbf{T}^{(i)}(t) = (T_{lk}^{(i)}(t))$ with dimension $|S_i| \times |S_i|$ can be divided into $\begin{pmatrix} \mathbf{T}_{W_iW_i}^{(i)}(t) & \mathbf{T}_{W_iF_i}^{(i)}(t) \\ \mathbf{T}_{F_iW_i}^{(i)}(t) & \mathbf{T}_{F_iF_i}^{(i)}(t) \end{pmatrix}$. It is easy to find that

$$\mathbf{T}_{W_iW_i}^{(i)}(t) = e^{\mathbf{Q}_{W_iW_i}^{(i)} t} + \int_0^t \int_0^{t-u} \mathbf{T}_{W_iW_i}^{(i)}(u) \mathbf{Q}_{W_iF_i}^{(i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} s} \mathbf{Q}_{F_iW_i}^{(i)} e^{\mathbf{Q}_{W_iW_i}^{(i)} (t-u-s)} ds du,$$

because there are two cases for the event $X_i(t) = k, k \in W_i$ to happen condition on $X_i(0) = l, l \in W_i$, one case is that the system never leave subspace W_i and the other one is that the system leave subspace W_i for at least one time. Then the instantaneous availability of the system is given as $A_i(t) = \boldsymbol{\pi}_{W_i}^{(i)} \mathbf{T}_{W_iW_i}^{(i)}(t) \mathbf{u}_{W_i}^{(i)}$, where $\boldsymbol{\pi}_{W_i}^{(i)}, \mathbf{u}_{W_i}^{(i)}$ are defined in Theorem 1. By taking Laplace transform, we have,

$$\begin{aligned}
\mathbf{T}_{W_iW_i}^{(i)*}(s) &= \left(s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)}\right)^{-1} + \mathbf{T}_{W_iW_i}^{(i)*}(s)\mathbf{Q}_{W_iF_i}^{(i)}\left(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)}\right)^{-1}\mathbf{Q}_{F_iW_i}^{(i)}\left(s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)}\right)^{-1} \\
&= \left(s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)}\right)^{-1}\left[\mathbf{I} - \mathbf{Q}_{W_iF_i}^{(i)}(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)})^{-1}\mathbf{Q}_{F_iW_i}^{(i)}(s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)})^{-1}\right]^{-1} \\
&= \left[s\mathbf{I} - \mathbf{Q}_{F_iW_i}^{(i)} - \mathbf{Q}_{W_iF_i}^{(i)}(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)})^{-1}\mathbf{Q}_{F_iW_i}^{(i)}\right]^{-1},
\end{aligned}$$

and then $A_i^*(s) = \boldsymbol{\pi}_{W_i}^{(i)}\mathbf{T}_{W_iW_i}^{(i)*}(s)\mathbf{u}_{W_i}^{(i)} = \boldsymbol{\pi}_{W_i}^{(i)}\left[s\mathbf{I} - \mathbf{Q}_{W_iW_i}^{(i)} - \mathbf{Q}_{W_iF_i}^{(i)}(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)})^{-1}\mathbf{Q}_{F_iW_i}^{(i)}\right]^{-1}\mathbf{u}_{W_i}^{(i)}$.

□

Remark Another proof can be given by [8]. From [8], we have $\mathbf{T}^{(i)}(t) = e^{\mathbf{Q}^{(i)}t}$ and its Laplace-transform $\mathbf{T}^{(i)*}(s) = \left(s\mathbf{I} - \mathbf{Q}^{(i)}\right)^{-1}$, and then $\mathbf{T}_{W_iW_i}^{(i)*}(s)$ can be given by inverse of general partitioned matrix.

Theorem 3 Interval availability of general repairable series system can be given as follows,

$$A[a, b] = \prod_{i=1}^{n} A_i[a, b],$$

where $A_i[a, b] = \boldsymbol{\pi}_{W_i}^{(i)}\mathbf{T}_{W_iW_i}^{(i)}(a)e^{\mathbf{Q}_{W_iW_i}^{(i)}(b-a)}\mathbf{u}_{W_i}^{(i)}$.

Proof Denote interval availability of unit i as $A_i[a, b]$ which is the probability that unit i is in working state during [a, b]. Then interval availability of the system $A\,[a, b]$ can be given as $A[a, b] = \prod_{i=1}^{n} A_i[a, b]$, because it is the probability that the system is in working state during [a, b] which is true only when all the units are working during [a, b]. It is easy to find that $A_i[a, b] = \boldsymbol{\pi}_{W_i}^{(i)}\mathbf{T}_{W_iW_i}^{(i)}(a)e^{\mathbf{Q}_{W_iW_i}^{(i)}(b-a)}\mathbf{u}_{W_i}^{(i)}$, where $\boldsymbol{\pi}_{W_i}^{(i)}, e^{\mathbf{Q}_{W_iW_i}^{(i)}(b-a)}, \mathbf{u}_{W_i}^{(i)}$ are defined in Theorem 1 and $\mathbf{T}_{W_iW_i}^{(i)}(a)$ is defined in Theorem 2.

□

4 General Repairable Series System with Neglected Failures

Theorem 4 Reliability of repairable series system with neglected failures can be given as follows,

$$\widetilde{R}(t) = \prod_{i=1}^{n} \widetilde{R}_i(t),$$

where $\widetilde{R}_i^*(s) = \boldsymbol{\pi}_{W_i}^{(i)} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)*}(s) \left\{ \mathbf{u}_{W_i}^{(i)} + \mathbf{Q}_{W_iF_i}^{(i)} \left[\mathbf{I} - e^{-(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)})\tau_i} \right] \left(s\mathbf{I} - \mathbf{Q}_{F_iF_i}^{(i)} \right)^{-1} \mathbf{u}_{F_i}^{(i)} \right\}.$

Proof For repairable series system with neglected failures, denote the probability that failures of unit i are not found during $[0, t]$ as $\widetilde{R}_i(t)$, which means unit i doesn't fail during $[0, t]$ or all its failure are repaired in τ_i. Similar to Theorem 1, we have $\widetilde{R}(t) = \prod_{i=1}^{n} \widetilde{R}_i(t)$. Denote as $\tilde{p}_{lk}(t)$ the conditional probability that failures of unit i are not found during $[0, t]$ and unit i is in state k at time t condition on initial state l at time 0. The matrix $\widetilde{\mathbf{P}}^{(i)}(t) = (\tilde{p}_{lk}^{(i)}(t))$ with dimension $|S_i| \times |S_i|$ can be divided into $\begin{pmatrix} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t) & \widetilde{\mathbf{P}}_{W_iF_i}^{(i)}(t) \\ \widetilde{\mathbf{P}}_{F_iW_i}^{(i)}(t) & \widetilde{\mathbf{P}}_{F_iF_i}^{(i)}(t) \end{pmatrix}$. Then we have, $\widetilde{R}_i(t) = \boldsymbol{\pi}_{W_i}^{(i)} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t) \mathbf{u}_{W_i}^{(i)} + \boldsymbol{\pi}_{W_i}^{(i)} \widetilde{\mathbf{P}}_{W_iF_i}^{(i)}(t) \mathbf{u}_{F_i}^{(i)}$, where $\boldsymbol{\pi}_{W_i}^{(i)}, \mathbf{u}_{W_i}^{(i)}$ are defined in Theorem 1 and $\mathbf{u}_{W_i}^{(i)}$ is a vector with dimension $|F_i| \times 1$ whose elements are all 1 From [8], we have,

$$\widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t) = e^{\mathbf{Q}_{W_iW_i}^{(i)} t} + \int_0^t \int_0^{\min(t-u, \tau_i)} e^{\mathbf{Q}_{W_iW_i}^{(i)} u} \mathbf{Q}_{W_iF_i}^{(i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} s} \mathbf{Q}_{F_iW_i}^{(i)} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t-u-s) \, ds du,$$

$$\widetilde{\mathbf{P}}_{W_iF_i}^{(i)}(t) = \int_0^{\min(t, \tau_i)} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t-s) \mathbf{Q}_{W_iF_i}^{(i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} s} ds, \ \widetilde{\mathbf{P}}_{F_iW_i}^{(i)}(t) = \int_0^{\min(t, \tau_i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} s} \mathbf{Q}_{F_iW_i}^{(i)} \widetilde{\mathbf{P}}_{W_iW_i}^{(i)}(t-s) \, ds,$$

$$\tilde{\mathbf{P}}_{F_iF_i}^{(i)}(t) = e^{\mathbf{Q}_{F_iF_i}^{(i)} t} I_{\{t < \tau_i\}} + \int_0^{\min(t, \tau_i)} \int_0^{\min(t-s, \tau_i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} u} \mathbf{Q}_{F_iW_i}^{(i)} \tilde{\mathbf{P}}_{W_iW_i}^{(i)}(t-u-s) \mathbf{Q}_{W_iF_i}^{(i)} e^{\mathbf{Q}_{F_iF_i}^{(i)} s} duds.$$

By taking Laplace transform, we have,

$$\begin{aligned}
\tilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s) &= (s\mathbf{I}-\mathbf{Q}^{(i)}_{W_iW_i})^{-1} + (s\mathbf{I}-\mathbf{Q}^{(i)}_{W_iW_i})^{-1}\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}] \\
&\quad (s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\tilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s) \\
&= \{\mathbf{I}-(s\mathbf{I}-\mathbf{Q}^{(i)}_{W_iW_i})^{-1}\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\}^{-1}(s\mathbf{I}-\mathbf{Q}^{(i)}_{W_iW_i})^{-1} \\
&= \{s\mathbf{I}-\mathbf{Q}^{(i)}_{W_iW_i}-\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\}^{-1}, \\
\tilde{\mathbf{P}}^{(i)*}_{W_iF_i}(s) &= \tilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}, \\
\tilde{\mathbf{P}}^{(i)*}_{F_iW_i}(s) &= [\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\tilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s), \\
\tilde{\mathbf{P}}^{(i)*}_{F_iF_i}(s) &= [\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1} \\
&\quad \{\mathbf{I}+\mathbf{Q}^{(i)}_{F_iW_i}\tilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\}.
\end{aligned}$$

And then $\widetilde{R}^*_i(s)=\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)\left\{\mathbf{u}_{W_i}+\mathbf{Q}^{(i)}_{W_iF_i}\left[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}\right]\left(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i}\right)^{-1}\mathbf{u}^{(i)}_{F_i}\right\}.$ □

Theorem 5 Instantaneous availability of general repairable series system with neglected failures can be given as follows,

$$\widetilde{A}(t)=\prod_{i=1}^{n}\widetilde{A}_i(t),$$

where $\widetilde{A}^*_i(s)=\boldsymbol{\pi}^{(i)}_{W_i}\tilde{\mathbf{T}}^{(i)*}_{W_iW_i}(s)\left\{\mathbf{u}^{(i)}_{W_i}+\mathbf{Q}^{(i)}_{W_iF_i}\left[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}\right]\left(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i}\right)^{-1}\mathbf{u}^{(i)}_{F_i}\right\}.$

Proof For repairable series system with neglected failures, denote the probability that failures of unit i are not found at time t as $\widetilde{A}_i(t)$, which means unit i is in working state at t or its last failure before time t happens after $t-\tau_i$. Similar to Theorem 2, we have $\widetilde{A}(t)=\prod_{i=1}^{n}\widetilde{A}_i(t)$. Denote as $\widetilde{T}^{(i)}_{lk}(t)$ the conditional probability that failures of unit i are not found at time t and unit i is in state k at time t condition on initial state l at time 0. The matrix $\widetilde{\mathbf{T}}^{(i)}(t)=(\widetilde{T}^{(i)}_{lk}(t))$ with dimension $|S_i|\times|S_i|$ can be divided into $\begin{pmatrix}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(t) & \widetilde{\mathbf{T}}^{(i)}_{W_iF_i}(t)\\ \widetilde{\mathbf{T}}^{(i)}_{F_iW_i}(t) & \widetilde{\mathbf{T}}^{(i)}_{F_iF_i}(t)\end{pmatrix}$. Then we have, $\widetilde{A}_i(t)=\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(t)\mathbf{u}^{(i)}_{W_i}+\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iF_i}(t)\mathbf{u}^{(i)}_{F_i}$, where $\boldsymbol{\pi}^{(i)}_{W_i}$, $\mathbf{u}^{(i)}_{W_i}$ are defined in Theorem 1 and $\mathbf{u}^{(i)}_{F_i}$ is defined in Theorem 4. From [28], we have,

$$\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(t)=\widetilde{\mathbf{P}}^{(i)}_{W_iW_i}(t)+\int_0^t\int_{\tau_i}^{t-u}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(u)\mathbf{Q}^{(i)}_{W_iF_i}e^{\mathbf{Q}^{(i)}_{F_iF_i}s}\mathbf{Q}^{(i)}_{F_iW_i}\widetilde{\mathbf{P}}^{(i)}_{W_iW_i}(t-u-s)\,ds\,du,$$

$$\widetilde{\mathbf{T}}^{(i)}_{W_iF_i}(t)=\int_0^{\min(t,\tau_i)}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(t-s)\mathbf{Q}^{(i)}_{W_iF_i}e^{\mathbf{Q}^{(i)}_{F_iF_i}s}ds.$$

By taking Laplace transform, we have,

$$\begin{aligned}\widetilde{\mathbf{T}}^{(i)*}_{W_iW_i}(s) &= \widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s) + \widetilde{\mathbf{T}}^{(i)*}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)\\ &= \widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)[\mathbf{I}-\mathbf{Q}^{(i)}_{W_iF_i}e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)]^{-1}\\ &= \left\{\left[\widetilde{\mathbf{P}}^{(i)*}_{W_iW_i}(s)\right]^{-1} - \mathbf{Q}^{(i)}_{W_iF_i}e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{Q}^{(i)}_{F_iW_i}\right\}^{-1},\end{aligned}$$

$$\widetilde{\mathbf{T}}^{(i)*}_{W_iF_i}(s) = \widetilde{\mathbf{T}}^{(i)*}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1},$$

and then $\widetilde{A}^*_i(s) = \boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)*}_{W_iW_i}(s)\left\{\mathbf{u}^{(i)}_{W_i} + \mathbf{Q}^{(i)}_{W_iF_i}\left[\mathbf{I}-e^{-(s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})\tau_i}\right](s\mathbf{I}-\mathbf{Q}^{(i)}_{F_iF_i})^{-1}\mathbf{u}^{(i)}_{F_i}\right\}.$

□

Theorem 6 Interval availability of general repairable series system with neglected failures can be given as follows,

$$\widetilde{A}[a,b] = \prod_{i=1}^{n}\widetilde{A}_i[a,b].$$

Proof Denote interval availability of unit i as $\widetilde{A}_i[a,b]$ which is the probability that failures of unit i are not found during [a, b]. Then interval availability of the system $\widetilde{A}[a,b]$ can be given as $\widetilde{A}[a,b] = \prod_{i=1}^{n}\widetilde{A}_i[a,b]$, because it is the probability that the system is in working states during [a, b] which is true only when failures of all its units are not found during [a, b]. It is easy to find that, $\widetilde{A}_i[a,b] = 0$ for $a < b < \tau_i$ and $\widetilde{A}_i[a,b] = \widetilde{A}_i[\tau_i,b]$ for $a < \tau_i < b$. When $\tau_i < a < b$, we have,

$$\begin{aligned}\widetilde{A}_i[a,b] &= \boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(a)\widetilde{\mathbf{P}}^{(i)}_{W_iW_i}(b-a)\mathbf{u}^{(i)}_{W_i} + \boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(a)\widetilde{\mathbf{P}}^{(i)}_{W_iF_i}(b-a)\mathbf{u}^{(i)}_{F_i}\\ &+ \int_{a-\tau_i}^{a}\int_{a-s}^{\min(b-s,\tau_i)}\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}e^{\mathbf{Q}^{(i)}_{F_iF_i}u}\mathbf{Q}^{(i)}_{F_iW_i}\widetilde{\mathbf{P}}^{(i)}_{W_iW_i}(b-a-s-u)\mathbf{u}^{(i)}_{W_i}\,duds\\ &+ \int_{a-\tau_i}^{a}\int_{a-s,\tau_i}^{\min(b-s,\tau_i)}\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(s)\mathbf{Q}^{(i)}_{W_iF_i}e^{\mathbf{Q}^{(i)}_{F_iF_i}u}\mathbf{Q}^{(i)}_{F_iW_i}\widetilde{\mathbf{P}}^{(i)}_{W_iF_i}(b-a-s-u)\mathbf{u}^{(i)}_{F_i}\,duds,\end{aligned}$$

where the first term $\boldsymbol{\pi}^{(i)}_{W_i}\widetilde{\mathbf{T}}^{(i)}_{W_iW_i}(a)\widetilde{\mathbf{P}}^{(i)}_{W_iW_i}(b-a)\mathbf{u}^{(i)}_{W_i}$ means failures of unit i are not found during [a, b] and unit i is in working state at both time a and time b, the second term means failures of unit i are not found during [a, b] and unit i is in working state at time a and in failure state at time b, the third term means failures of unit i are not found during [a, b] and unit i is in failure state at time a and in working state at time b, and the last term means failures of unit i are not found during [a, b] and unit i is in failure state at both time a and time b.

□

Remark The results given in Theorems 1–6 can be applied to general MSS whose structure function can be given by maximum operator and minimum operator. Reliability, instantaneous availability and interval availability of these systems can be given by using *Lz*-transform method for corresponding binary state system, respectively. To explain how those reliability measures are calculated for general MSSs, a numerical example (Example 2) is given in Sect. 7.

5 Phased-Mission Repairable Series System

Theorem 7 Reliability of phased-mission repairable series system can be given as follows,

(1) When $t \in [mT, mT + T_1],\ m = 0, 1, 2, \ldots$, unit 1 is in use at time t and we have,

$$\vec{R}(t) = A_1(\prod_{k=0}^{m-1}[kT, kT + T_1], [mT, t]) \prod_{i=2}^{n} A_i(\prod_{k=0}^{m-1}[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

(2) When $t \in [mT + \sum_{j=1}^{n-1} T_j, mT + \sum_{j=1}^{n-1} T_j], m = 0, 1, 2, \ldots$, unit n is in use at time t and we have,

$$\vec{R}(t) = A_n(\prod_{k=0}^{m-1}[kT + \sum_{j=1}^{n-1} T_j, kT + \sum_{j=1}^{n} T_j], [mT + \sum_{j=1}^{n-1} T_j, t])$$
$$\times \prod_{i=1}^{n-1} A_i(\prod_{k=0}^{m}[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

(3) When $t \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ 1 < l < n,\ m = 0, 1, 2, 3 \ldots$, unit i is in use at time t and we have,

$$\vec{R}(t) = A_l(\prod_{k=0}^{m-1}[kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j], [mT + \sum_{j=1}^{l-1} T_j, t])$$
$$\prod_{i=1}^{l-1} A_i(\prod_{k=0}^{m}[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) \times \prod_{i=l+1}^{n} A_i(\prod_{k=0}^{m-1}[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j].$$

Proof General phased-mission repairable series system fails once the unit in use fails. Unit i is in use during $[mT+\sum_{j=1}^{i-1}T_j, mT+\sum_{j=1}^{i}T_j]$, $T=T_1+T_2+\cdots+T_n$, $m=0,1,\ldots,\infty$.

(1) When $t\in[mT,mT+T_1]$, $m=0,1,2,\ldots$, unit 1 is in use during $[kT,kT+T_1]$ $(k=0,1,\ldots,m-1)$ and $[mT,t]$, unit $i\,(i=2,\ldots,n)$ is in use during $[kT+\sum_{j=1}^{i-1}T_j,\ kT+\sum_{j=1}^{i}T_j]$ $(k=0,1,\ldots,m-1)$. Since the n units are independent from each other, and from [27], we have,

$$\overrightarrow{R}(t)=A_1(\prod_{k=0}^{m-1}[kT,kT+T_1],[mT,t])\prod_{i=2}^{n}A_i(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{i-1}T_j,kT+\sum_{j=1}^{i}T_j]),$$

where

$$A_1(\prod_{k=0}^{m-1}[kT,kT+T_1],[mT,t])=\boldsymbol{\pi}_{W_1}^{(1)}e^{\mathbf{Q}_{W_1W_1}^{(1)}T_1}\mathbf{E}_1^{(1)}[e^{\mathbf{Q}^{(1)}(T-T_1)}\mathbf{E}_2^{(1)}e^{\mathbf{Q}_{W_1W_1}^{(1)}T_1}\mathbf{E}_1^{(1)}]^{m-1}$$
$$e^{\mathbf{Q}^{(1)}(T-T_1)}\mathbf{E}_2^{(1)}e^{\mathbf{Q}_{W_1W_1}^{(1)}(t-mT)}\mathbf{u}_{W_1}^{(1)},$$

$$A_i(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{i-1}T_j,kT+\sum_{j=1}^{i}T_j])=\boldsymbol{\pi}_{W_i}^{(i)}\mathbf{E}_1^{(i)}e^{\mathbf{Q}^{(i)}\sum_{j=1}^{i-1}T_j}\mathbf{E}_2^{(i)}e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}$$
$$[e^{\mathbf{Q}^{(i)}(T-T_i)}\mathbf{E}_2^{(i)}e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}]^{m-1}\mathbf{u}^{(i)},$$

$$\mathbf{E}_1^{(i)}=(\mathbf{I}_{|W_i|\times|W_i|},\mathbf{0}_{|W_i|\times|F_i|}),\mathbf{E}_2^{(i)}=\begin{pmatrix}\mathbf{I}_{|W_i|\times|W_i|}\\ \mathbf{0}_{|F_i|\times|W_i|}\end{pmatrix},\mathbf{u}^{(i)}=\begin{pmatrix}\mathbf{u}_{W_i}^{(i)}\\ \mathbf{u}_{F_i}^{(i)}\end{pmatrix}.$$

(2) When $t\in[mT+\sum_{j=1}^{n-1}T_j,mT+\sum_{j=1}^{n-1}T_j]$, $m=0,1,2,\ldots$, unit n is in use during $[kT+\sum_{j=1}^{n-1}T_j,\ kT+\sum_{j=1}^{n}T_j]$ $(k=0,1,\ldots,m-1)$ and $[mT+\sum_{j=1}^{n-1}T_j,t]$, unit $i\,(i=2,\ldots,n)$ is in use during $[kT+\sum_{j=1}^{i-1}T_j,\ kT+\sum_{j=1}^{i}T_j]$ $(k=0,1,\ldots,m)$. Since the n units are independent from each other, and from [27], we have,

$$\vec{R}(t)=A_n(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{n-1}T_j,kT+\sum_{j=1}^{n}T_j],[mT+\sum_{j=1}^{n-1}T_j,t])$$
$$\prod_{i=1}^{n-1}A_i(\prod_{k=0}^{m}[kT+\sum_{j=1}^{i-1}T_j,kT+\sum_{j=1}^{i}T_j]),$$

where

$$A_n\Big(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{n-1}T_j, kT+\sum_{j=1}^{n}T_j], [mT+\sum_{j=1}^{n-1}T_j, t]\Big)$$
$$=\boldsymbol{\pi}_{W_n}^{(n)}\mathbf{E}_1^{(n)}[e^{\mathbf{Q}^{(n)}(T-T_n)}\mathbf{E}_2^{(n)}e^{\mathbf{Q}_{W_nW_n}^{(n)}T_n}\mathbf{E}_1^{(n)}]^m$$
$$e^{\mathbf{Q}^{(n)}(T-T_n)}\mathbf{E}_2^{(n)}e^{\mathbf{Q}_{W_nW_n}^{(n)}(t-mT-\sum_{j=1}^{n-1}T_j)}\mathbf{u}_{W_n}^{(n)},$$
$$A_i\Big(\prod_{k=0}^{m}[kT+\sum_{j=1}^{i-1}T_j, kT+\sum_{j=1}^{i}T_j]\Big)$$
$$=\boldsymbol{\pi}_{W_i}^{(i)}\mathbf{E}_1^{(i)}e^{\mathbf{Q}^{(i)}\sum_{j=1}^{i-1}T_j}\mathbf{E}_2^{(i)}e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}[e^{\mathbf{Q}^{(i)}(T-T_i)}\mathbf{E}_2^{(i)}e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}]^m\mathbf{u}^{(i)}.$$

(3) When $t\in[mT+\sum_{j=1}^{l-1}T_j,\ mT+\sum_{j=1}^{l-1}T_j]$, $m=0,1,2,\ldots$, unit l is in use during $[kT+\sum_{j=1}^{l-1}T_j,\ kT+\sum_{j=1}^{l}T_j]$ $(k=0,1,\ldots,m-1)$ and $[mT+\sum_{j=1}^{l-1}T_j, t]$, unit i $(i=1,\ldots,l-1)$ is in use during $[kT+\sum_{j=1}^{i-1}T_j,\ kT+\sum_{j=1}^{i}T_j]$ $(k=0,1,\ldots,m)$, unit i $(i=l+1,\ldots,n)$ is in use during $[kT+\sum_{j=1}^{i-1}T_j,\ kT+\sum_{j=1}^{i}T_j]$ $(k=0,1,\ldots,m-1)$. Since the n units are independent from each other, and from [27], we have,

$$\vec{R}(t)=A_l\Big(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{l-1}T_j, kT+\sum_{j=1}^{l}T_j], [mT+\sum_{j=1}^{l-1}T_j, t]\Big)$$
$$\times\prod_{i=1}^{l-1}A_i\Big(\prod_{k=0}^{m}[kT+\sum_{j=1}^{i-1}T_j, kT+\sum_{j=1}^{i}T_j]\Big)\times\prod_{i=l+1}^{n}A_i\Big(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{i-1}T_j, kT+\sum_{j=1}^{i}T_j],$$

where

$$A_l\Big(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{l-1}T_j, kT+\sum_{j=1}^{l}T_j], [mT+\sum_{j=1}^{l-1}T_j, t]\Big)=\pi_{W_l}^{(l)}\mathbf{E}_1^{(l)}e^{\mathbf{Q}^{(l)}\sum_{j=1}^{l-1}T_j}\mathbf{E}_2^{(l)}e^{\mathbf{Q}_{W_lW_l}^{(l)}T_l}\mathbf{E}_1^{(l)}$$
$$[e^{\mathbf{Q}^{(l)}(T-T_l)}\mathbf{E}_2^{(l)}e^{\mathbf{Q}_{W_lW_l}^{(l)}T_l}\mathbf{E}_1^{(l)}]^{m-1}$$
$$e^{\mathbf{Q}^{(l)}(T-T_l)}\mathbf{E}_2^{(l)}e^{\mathbf{Q}_{W_lW_l}^{(l)}(t-mT-\sum_{j=1}^{l-1}T_j)}\mathbf{u}_{W_l}^{(l)}.$$

□

Theorem 8 Instantaneous availability of phased-mission repairable series system can be given as follows,

$$\overrightarrow{A}(t) = A_i(t),\ t \in [mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j],\ 1 \le i \le n, m = 0, 1, \ldots, \infty.$$

Proof Unit i is in use during $[mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j]$, $T = T_1 + T_2 + \cdots + T_n$, $m = 0, 1, \ldots, \infty$, then it's easy to find that $\overrightarrow{A}(t) = \overrightarrow{A}_i(t)$ during that time. □

Theorem 9 Interval availability of phased-mission repairable series system can be given as follows,

(1) When $a, b \in [mT + \sum_{j=1}^{l-1} T_j,\ mT + \sum_{j=1}^{l} T_j], m = 0, 1, 2, \ldots$, unit l is in use during $[a,b]$ and we have $\overrightarrow{A}[a, b] = A_l[a, b]$.

(2) When $a \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ b \in [mT + \sum_{j=1}^{p-1} T_j, mT + \sum_{j=1}^{p} T_j]$, $l < p, m = 0, 1, 2, \ldots$, the system is in its $(m+1)$ th period during $[a, b]$ and we have,

$$\overrightarrow{A}[a, b] = A_l[a, mT + \sum_{j=1}^{l} T_j] A_p[mT + \sum_{j=1}^{p-1} T_j, b] \prod_{i=l+1}^{p-1} A_i[mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j].$$

(3) When $a \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ b \in [rT + \sum_{j=1}^{p-1} T_j, rT + \sum_{j=1}^{p} T_j],\ m < r$, $m = 0, 1, 2, \ldots$, the system is not in the same period during $[a,b]$ and unit l is in use at time a and unit p is in use at time b. If $l = p$, we have,

$$\begin{aligned}\vec{A}[a, b] = A_l([a, mT + \sum_{j=1}^{l} T_j], [rT + \sum_{j=1}^{l-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j]) \\ \times \prod_{i=1}^{l-1} A_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) \\ \prod_{i=l+1}^{n} A_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).\end{aligned}$$

If $l < p$, we have,

$$\vec{A}[a,b] = A_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times A_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{l-1} A_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\times \prod_{i=l+1}^{p-1} A_i(\prod_{k=m}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=p+1}^{n} A_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

If $l > p$, we have,

$$\vec{A}[a,b] = A_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times A_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{p-1} A_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\times \prod_{i=p+1}^{l-1} A_i(\prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=l+1}^{n} A_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

Proof

(1) When $a, b \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j]$, $m = 0, 1, 2, \ldots$, unit l is in use during $[a,b]$ and it's obvious that $\overrightarrow{A}[a,b] = A_l[a,b]$.

(2) When $a \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j], b \in [mT + \sum_{j=1}^{p-1} T_j,\ mT + \sum_{j=1}^{p} T_j]\ , l < p,\ m = 0, 1, 2, \ldots$, unit l is in use during $[a, mT + \sum_{j=1}^{l} T_j]$, unit p is in use during $[mT + \sum_{j=1}^{p-1} T_j, b]$, unit $i\ (i = l+1, \ldots, p-1)$ is in use during $[mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j]$, then we have,

$$\overrightarrow{A}[a,b] = A_l[a, mT + \sum_{j=1}^{l} T_j] A_p[mT + \sum_{j=1}^{p-1} T_j, b] \prod_{i=l+1}^{p-1} A_i[mT + \sum_{j=1}^{i-1} T_j, mT + \sum_{j=1}^{i} T_j].$$

(3) When $a \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ b \in [rT + \sum_{j=1}^{p-1} T_j, rT + \sum_{j=1}^{p} T_j],\ m < r,\ m = 0, 1, 2, \ldots$, there are three cases to be considered.

Case 1: If $l = p$, unit l is in use during $[a, mT + \sum_{j=1}^{l} T_j]$, $[rT + \sum_{j=1}^{l-1} T_j, b]$ and $[kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j]\ (l = m+1, \ldots, r-1)$, unit $i\ (i = 1, \ldots, l-1)$ is in use during $[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]\ (k = m+1, \ldots, r)$, unit $i\ (i = l+1, \ldots, n)$ is in use during $[kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]$ $(k = m, \ldots, r-1)$, then we have,

$$\vec{A}[a,b] = A_l([a, mT + \sum_{j=1}^{l} T_j], [rT + \sum_{j=1}^{l-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times \prod_{i=1}^{l-1} A_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=l+1}^{n} A_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]),$$

where

$$A_l([a, mT+\sum_{j=1}^{l} T_j], [rT+\sum_{j=1}^{l-1} T_j, b], \prod_{k=m+1}^{r-1} [kT+\sum_{j=1}^{l-1} T_j, kT+\sum_{j=1}^{l} T_j])$$
$$= \boldsymbol{\pi}_{W_l}^{(l)} \mathbf{E}_1^{(l)} e^{\mathbf{Q}^{(l)} a} \mathbf{E}_2^{(l)} e^{\mathbf{Q}_{W_l W_l}^{(l)} (mT+\sum_{j=1}^{l} T_j - a)} \mathbf{E}_1^{(l)} [e^{\mathbf{Q}^{(l)}(T-T_l)} \mathbf{E}_2^{(l)} e^{\mathbf{Q}_{W_l W_l}^{(l)} T_l} \mathbf{E}_1^{(l)}]^{r-m-1}$$
$$e^{\mathbf{Q}^{(l)}(T-T_l)} \mathbf{E}_2^{(l)} e^{\mathbf{Q}_{W_l W_l}^{(l)} (b-rT-\sum_{j=1}^{l-1} T_j)} \mathbf{u}_{W_l}^{(l)}.$$

Case 2: If $l<p$, unit l is in use during $[a, mT+\sum_{j=1}^{l} T_j]$ and $[kT+\sum_{j=1}^{l-1} T_j,\ kT+\sum_{j=1}^{l} T_j]$ $(l=m+1, \ldots, r)$, unit p is in use during $[rT+\sum_{j=1}^{p-1} T_j, b]$ and $[kT+\sum_{j=1}^{p-1} T_j, kT+\sum_{j=1}^{p} T_j]$ $(k=m, \ldots, r-1)$, unit i $(i=1, \ldots, l-1)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j](k=m+1, \ldots, r)$, unit i $(i=l+1, \ldots, p-1)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j](k=m, \ldots, r)$, unit i $(i=p+1, \ldots, n)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j]$ $(k=m, \ldots, r-1)$, then we have,

$$\vec{A}[a,b] = A_l([a, mT+\sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r} [kT+\sum_{j=1}^{l-1} T_j, kT+\sum_{j=1}^{l} T_j]) \times$$
$$A_p([rT+\sum_{j=1}^{p-1} T_j, b], \prod_{k=m}^{r-1} [kT+\sum_{j=1}^{p-1} T_j,\ kT+\sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{l-1} A_i(\prod_{k=m+1}^{r} [kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j])$$
$$\times \prod_{i=l+1}^{p-1} A_i(\prod_{k=m}^{r} [kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j])$$
$$\prod_{i=p+1}^{n} A_i(\prod_{k=m}^{r-1} [kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j]),$$

where

$$A_l([a, mT+\sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r}[kT+\sum_{j=1}^{l-1} T_j, kT+\sum_{j=1}^{l} T_j])$$
$$=\boldsymbol{\pi}_{W_l}^{(l)}\mathbf{E}_1^{(l)} e^{\mathbf{Q}^{(l)}a}\mathbf{E}_2^{(l)} e^{\mathbf{Q}_{W_lW_l}^{(l)}(mT+\sum_{j=1}^{l} T_j - a)}$$
$$\mathbf{E}_1^{(l)}[e^{\mathbf{Q}^{(l)}(T-T_l)}\mathbf{E}_2^{(l)} e^{\mathbf{Q}_{W_lW_l}^{(l)}T_l}\mathbf{E}_1^{(l)}]^{r-m}\mathbf{u}^{(l)},$$
$$A_p([rT+\sum_{j=1}^{p-1} T_j, b], \prod_{k=m}^{r-1}[kT+\sum_{j=1}^{p-1} T_j,\ kT+\sum_{j=1}^{p} T_j])$$
$$=\boldsymbol{\pi}_{W_p}^{(p)}\mathbf{E}_1^{(p)} e^{\mathbf{Q}^{(p)}(mT+\sum_{j=1}^{p-1} T_j)}\mathbf{E}_2^{(p)}[e^{\mathbf{Q}_{W_pW_p}^{(p)}T_p}\mathbf{E}_1^{(p)} e^{\mathbf{Q}^{(p)}(T-T_p)}\mathbf{E}_2^{(p)}]^{r-m}$$
$$e^{\mathbf{Q}_{W_pW_p}^{(p)}(b-rT-\sum_{j=1}^{p-1} T_j)}\mathbf{u}_{W_P},$$
$$A_i(\prod_{k=m}^{r}[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j])$$
$$=\boldsymbol{\pi}_{W_i}^{(i)}\mathbf{E}_1^{(i)} e^{\mathbf{Q}^{(i)}(mT+\sum_{j=1}^{i-1} T_j)}\mathbf{E}_2^{(i)} e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}[e^{\mathbf{Q}^{(i)}(T-T_i)}\mathbf{E}_2^{(i)} e^{\mathbf{Q}_{W_iW_i}^{(i)}T_i}\mathbf{E}_1^{(i)}]^{r-m}\mathbf{u}^{(i)}.$$

Case 3: If $l>p$, unit l is in use during $[a, mT+\sum_{j=1}^{l} T_j]$ and $[kT+\sum_{j=1}^{l-1} T_j, kT+\sum_{j=1}^{l} T_j]$ $(k=m+1, \ldots, r-1)$, unit p is in use during $[rT+\sum_{j=1}^{p-1} T_j, b]$ and $[kT+\sum_{j=1}^{p-1} T_j, kT+\sum_{j=1}^{p} T_j]$ $(k=m+1, \ldots, r-1)$, unit i $(i=1, \ldots, p-1)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j]$ $(k=m+1, \ldots, r)$, unit i $(i=p+1, \ldots, l-1)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j]$ $(k=m+1, \ldots, r-1)$, unit i $(i=l+1, \ldots, n)$ is in use during $[kT+\sum_{j=1}^{i-1} T_j, kT+\sum_{j=1}^{i} T_j]$ $(k=m, \ldots, r-1)$, then we have,

$$\vec{A}[a,b] = A_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times A_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{p-1} A_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\times \prod_{i=p+1}^{l-1} A_i(\prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=l+1}^{n} A_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

□

6 Phased-Mission Repairable Series System with Neglected Failures

Theorem 10 Reliability of phased-mission repairable series system with neglected failures can be given as follows,

(1) When $t \in [mT, mT + T_1],\ m = 0, 1, 2, \ldots$, unit 1 is in use at time t and we have,

$$\overrightarrow{\widetilde{R}}(t) = \widetilde{A}_1(\prod_{k=0}^{m-1} [kT, kT + T_1], [mT, t]) \prod_{i=2}^{n} \widetilde{A}_i(\prod_{k=0}^{m-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

(2) When $t \in [mT + \sum_{j=1}^{n-1} T_j, mT + \sum_{j=1}^{n-1} T_j],\ m = 0, 1, 2, \ldots$, unit n is in use at time t and we have,

$$\vec{\tilde{R}}(t) = \tilde{A}_n(\prod_{k=0}^{m-1} [kT + \sum_{j=1}^{n-1} T_j, kT + \sum_{j=1}^{n} T_j], [mT + \sum_{j=1}^{n-1} T_j, t])$$
$$\prod_{i=1}^{n-1} \tilde{A}_i(\prod_{k=0}^{m} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

(3) When $t \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ 1 < l < n,\ m = 0, 1, 2, 3 \ldots$, unit i is in use at time and we have,

$$\vec{\tilde{R}}(t) = \tilde{A}_l(\prod_{k=0}^{m-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j], [mT + \sum_{j=1}^{l-1} T_j, t]) \prod_{i=1}^{l-1} \tilde{A}_i(\prod_{k=0}^{m} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) \times \prod_{i=l+1}^{n} \tilde{A}_i(\prod_{k=0}^{m-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j].$$

Proof The proof is similar to Theorem 7. From the proof of Theorem 4, for the three cases we have,

(1) When $t \in [mT, mT + T_1],\ m = 0, 1, 2, \ldots$,

$$\tilde{A}_1(\prod_{k=0}^{m-1} [kT, kT + T_1], [mT, t]) = \boldsymbol{\pi}_{W_1}^{(1)} \mathbf{E}_1^{(1)} \tilde{\mathbf{P}}^{(1)}(T_1)[e^{\mathbf{Q}^{(1)}(T - T_1)} \tilde{\mathbf{P}}^{(1)}(T_1)]^{m-1} e^{\mathbf{Q}^{(1)}(T - T_1)} \tilde{\mathbf{P}}^{(1)}(t - mT)\mathbf{u}^{(1)},$$

where $\widetilde{\mathbf{P}}^{(1)}(T_1)$ is given in Theorem 4.

(2) When $t \in [mT + \sum_{j=1}^{n-1} T_j, mT + \sum_{j=1}^{n-1} T_j],\ m = 0, 1, 2, \ldots$,

$$\tilde{A}_n(\prod_{k=0}^{m-1} [kT + \sum_{j=1}^{n-1} T_j, kT + \sum_{j=1}^{n} T_j], [mT + \sum_{j=1}^{n-1} T_j, t]) = \boldsymbol{\pi}_{W_n}^{(n)} \mathbf{E}_1^{(n)} [e^{\mathbf{Q}^{(n)}(T - T_n)} \tilde{\mathbf{P}}^{(n)}(T_n)]^m e^{\mathbf{Q}^{(n)}(T - T_n)} \tilde{\mathbf{P}}^{(n)}(t - mT - \sum_{j=1}^{n-1} T_j)\mathbf{u}^{(n)},$$

$$\tilde{A}_i(\prod_{k=0}^{m} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) = \boldsymbol{\pi}_{W_i}^{(i)} \mathbf{E}_1^{(i)} e^{\mathbf{Q}^{(i)} \sum_{j=1}^{i-1} T_j} \tilde{\mathbf{P}}^{(i)}(T_i)[e^{\mathbf{Q}^{(i)}(T - T_i)} \mathbf{E}_2^{(i)} \tilde{\mathbf{P}}^{(i)}(T_i)]^m \mathbf{u}^{(i)}.$$

(3) When $t \in [mT + \sum_{j=1}^{l-1} T_j, mT + \sum_{j=1}^{l} T_j],\ 1 < l < n,\ m = 0, 1, 2, 3 \ldots$,

$$\tilde{A}_l(\prod_{k=0}^{m-1}[kT+\sum_{j=1}^{l-1}T_j, kT+\sum_{j=1}^{l}T_j], [mT+\sum_{j=1}^{l-1}T_j, t])$$
$$=\boldsymbol{\pi}_{W_l}^{(l)}\mathbf{E}_1^{(l)}e^{\mathbf{Q}^{(l)}\sum_{j=1}^{l-1}T_j}\tilde{\mathbf{P}}^{(l)}(T_l)[e^{\mathbf{Q}^{(l)}(T-T_l)}\tilde{\mathbf{P}}^{(l)}(T_l)]^{m-1}e^{\mathbf{Q}^{(l)}(T-T_l)}\tilde{\mathbf{P}}^{(l)}(t-mT-\sum_{j=1}^{l-1}T_j)\mathbf{u}^{(l)}.$$

□

Theorem 11 Instantaneous availability of phased-mission repairable series system with neglected failures can be given as follows,

$$\overrightarrow{\widetilde{A}}(t)=\widetilde{A}_i(t), \quad t\in[mT+\sum_{j=1}^{i-1}T_j, mT+\sum_{j=1}^{i}T_j],\ 1\le i\le n,\ m=0,1,\ldots,\infty.$$

Proof Unit i is in use during $[mT+\sum_{j=1}^{i-1}T_j, mT+\sum_{j=1}^{i}T_j]$, $T=T_1+T_2+\cdots+T_n$, $m=0,1,\ldots,\infty$, then it's easy to find that $\overrightarrow{\widetilde{A}}(t)=\overrightarrow{\widetilde{A}}_l(t)$ during that time.

□

Theorem 12 Interval availability of phased-mission repairable series system with neglected failures can be given as follows,

(1) When $a,b\in[mT+\sum_{j=1}^{l-1}T_j, mT+\sum_{j=1}^{l}T_j]$, $m=0,1,2,\ldots$, unit l is in use during $[a, b]$ and we have $\overrightarrow{\widetilde{A}}[a,b]=\widetilde{A}_l[a,b]$.

(2) When $a\in[mT+\sum_{j=1}^{l-1}T_j, mT+\sum_{j=1}^{l}T_j], b\in[mT+\sum_{j=1}^{p-1}T_j, mT+\sum_{j=1}^{p}T_j]$, $l<p, m=0,1,2,\ldots$, the system is in its $(m+1)$ th period during $[a, b]$ and we have,

$$\overrightarrow{\widetilde{A}}[a,b]=\widetilde{A}_l[a,mT+\sum_{j=1}^{l}T_j]\widetilde{A}_p[mT+\sum_{j=1}^{p-1}T_j,b]\prod_{i=l+1}^{p-1}\widetilde{A}_i[mT+\sum_{j=1}^{i-1}T_j, mT+\sum_{j=1}^{i}T_j].$$

(3) When $a\in[mT+\sum_{j=1}^{l-1}T_j, mT+\sum_{j=1}^{l}T_j]$, $b\in[rT+\sum_{j=1}^{p-1}T_j, rT+\sum_{j=1}^{p}T_j]$, $m<r$, $m=0,1,2,\ldots$, the system is not in the same period during $[a, b]$ and unit l is in use at time a and unit p is in use at time b. If $l=p$, we have,

$$\vec{\tilde{A}}[a,b] = \tilde{A}_l([a, mT + \sum_{j=1}^{l} T_j], [rT + \sum_{j=1}^{l-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times \prod_{i=1}^{l-1} \tilde{A}_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=l+1}^{n} \tilde{A}_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

If $l < p$, we have,

$$\vec{\tilde{A}}[a,b] = \tilde{A}_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times \tilde{A}_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{l-1} \tilde{A}_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) \times \prod_{i=l+1}^{p-1} \tilde{A}_i(\prod_{k=m}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=p+1}^{n} \tilde{A}_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

If $l > p$, we have,

$$\vec{\tilde{A}}[a,b] = \tilde{A}_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$\times \tilde{A}_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$\prod_{i=1}^{p-1} \tilde{A}_i(\prod_{k=m+1}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]) \times \prod_{i=p+1}^{l-1} \tilde{A}_i(\prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$\prod_{i=l+1}^{n} \tilde{A}_i(\prod_{k=m}^{r-1} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j]).$$

Proof The proof is similar to Theorem 9. From the proof of Theorem 4, we have,

$$\tilde{A}_l([a, mT + \sum_{j=1}^{l} T_j], [rT + \sum_{j=1}^{l-1} T_j, b], \prod_{k=m+1}^{r-1} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$= \boldsymbol{\pi}_{W_l}^{(l)} e^{\mathbf{Q}^{(l)} a} \tilde{\mathbf{P}}^{(l)} (mT + \sum_{j=1}^{l} T_j - a) [e^{\mathbf{Q}^{(l)} (T - T_l)} \tilde{\mathbf{P}}_l (T_l)]^{r-m-1}$$
$$e^{\mathbf{Q}^{(l)} (T - T_l)} \tilde{\mathbf{P}}^{(l)} (b - rT - \sum_{j=1}^{l-1} T_j) \mathbf{u}^{(l)},$$

$$\tilde{A}_i(\prod_{k=m}^{r} [kT + \sum_{j=1}^{i-1} T_j, kT + \sum_{j=1}^{i} T_j])$$
$$= \boldsymbol{\pi}_{W_i}^{(i)} \mathbf{E}_1^{(i)} e^{\mathbf{Q}^{(i)} (mT + \sum_{j=1}^{i-1} T_j)} \tilde{\mathbf{P}}^{(i)} (T_i) [e^{\mathbf{Q}^{(i)} (T - T_i)} \tilde{\mathbf{P}}^{(i)} (T_i)]^{r-m} \mathbf{u}^{(i)},$$

$$\tilde{A}_l([a, mT + \sum_{j=1}^{l} T_j], \prod_{k=m+1}^{r} [kT + \sum_{j=1}^{l-1} T_j, kT + \sum_{j=1}^{l} T_j])$$
$$= \boldsymbol{\pi}_{W_l}^{(l)} \mathbf{E}_1^{(l)} e^{\mathbf{Q}^{(l)} a} \tilde{\mathbf{P}}^{(l)} (mT + \sum_{j=1}^{l} T_j - a) [e^{\mathbf{Q}^{(l)} (T - T_l)} \tilde{\mathbf{P}}^{(l)} (T_l)]^{r-m} \mathbf{u}^{(l)},$$

$$\tilde{A}_p([rT + \sum_{j=1}^{p-1} T_j, b], \prod_{k=m}^{r-1} [kT + \sum_{j=1}^{p-1} T_j, kT + \sum_{j=1}^{p} T_j])$$
$$= \boldsymbol{\pi}_{W_p}^{(p)} \mathbf{E}_1^{(p)} e^{\mathbf{Q}^{(p)} (mT + \sum_{j=1}^{p-1} T_j)} \tilde{\mathbf{P}}^{(p)} (T_p) [e^{\mathbf{Q}^{(p)} (T - T_p)} \tilde{\mathbf{P}}_l (T_p)]^{r-m-1}$$
$$e^{\mathbf{Q}^{(p)} (T - T_p)} \tilde{\mathbf{P}}^{(p)} (b - rT - \sum_{j=1}^{p-1} T_j) \mathbf{u}^{(p)}.$$

□

7 Numerical Examples

Example 1 In order to illustrate the results in the previous sections, a numerical example is given as follows. Assume that the underlying repairable series system consists of three independent units which can be shown in Fig. 5.

Unit 1 follows a Markov process $\{X_1(t), t \geq 0\}$ and its state space $S_1 = \{1, 2\}$ can be divided into two subspaces $W_1 = \{1\}$ and $F_1 = \{2\}$, which means 1 is a working state and 2 is a failure state. The initial probability vector is $\boldsymbol{\pi}_{W_1} = 1$ and the infinitesimal generator is $\boldsymbol{Q}^{(1)} = \left(\begin{array}{c|c} -a_{12}^{(1)} & a_{12}^{(1)} \\ \hline a_{21}^{(1)} & -a_{21}^{(1)} \end{array}\right) = \left(\begin{array}{c|c} -1 & 1 \\ \hline 2 & -2 \end{array}\right)$.

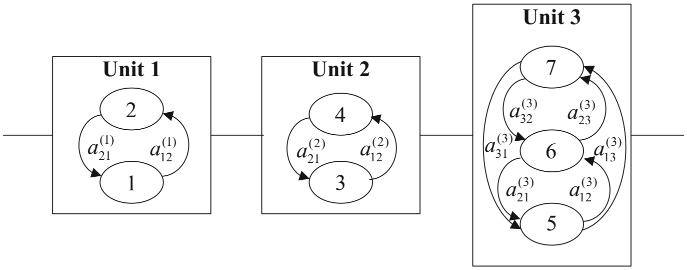


Fig. 5 Repairable series system with three independent units

Unit 2 follows a Markov process $\{X_2(t), t \geq 0\}$ and its state space $S_2 = \{3, 4\}$ can be divided into two subspaces $W_2 = \{3\}$ and $F_2 = \{4\}$, which means 3 is a working state and 4 is a failure state. The initial probability vector is $\boldsymbol{\pi}_{W_2} = 1$ and the infinitesimal generator is $\boldsymbol{Q}^{(2)} = \left(\begin{array}{c|c} -a_{12}^{(2)} & a_{12}^{(2)} \\ \hline a_{21}^{(2)} & -a_{21}^{(2)} \end{array}\right) = \left(\begin{array}{c|c} -4 & 4 \\ \hline 3 & -3 \end{array}\right)$.

Unit 3 follows a Markov process $\{X_3(t), t \geq 0\}$ and its state space $S_3 = \{5, 6, 7\}$ can be divided into two subspaces $W_3 = \{5\}$ and $F_3 = \{6, 7\}$, which means 5 is a working state and 6, 7 are failure states. The initial probability vector is $\boldsymbol{\pi}_{W_3} = 1$ and the infinitesimal generator is

$$\boldsymbol{Q}^{(3)} = \left(\begin{array}{c|cc} -a_{12}^{(3)} - a_{13}^{(3)} & a_{12}^{(3)} & a_{13}^{(3)} \\ \hline a_{21}^{(3)} & -a_{21}^{(3)} - a_{23}^{(3)} & a_{23}^{(3)} \\ a_{31}^{(3)} & a_{32}^{(3)} & -a_{31}^{(3)} - a_{32}^{(3)} \end{array}\right) = \left(\begin{array}{c|cc} -6 & 4 & 2 \\ \hline 2 & -3 & 1 \\ 1 & 3 & -4 \end{array}\right).$$

Numerical results are given for the four systems as follows.

(1) The first system: general repairable series system

(1) Reliability
The reliabilities of the three units and the system are given as follows and are shown in Fig. 6.

$$R_1(t) = e^{-a_{12}^{(1)}t} = e^{-t}, R_2(t) = e^{-a_{12}^{(1)}t} = e^{-4t}, R_3(t) = e^{-(a_{12}^{(3)} + a_{13}^{(3)})t} = e^{-6t}, R(t) = e^{-11t}.$$

(2) Instantaneous availability
The instantaneous availabilities of the three units and the system are given as follows and are shown in Fig. 6.

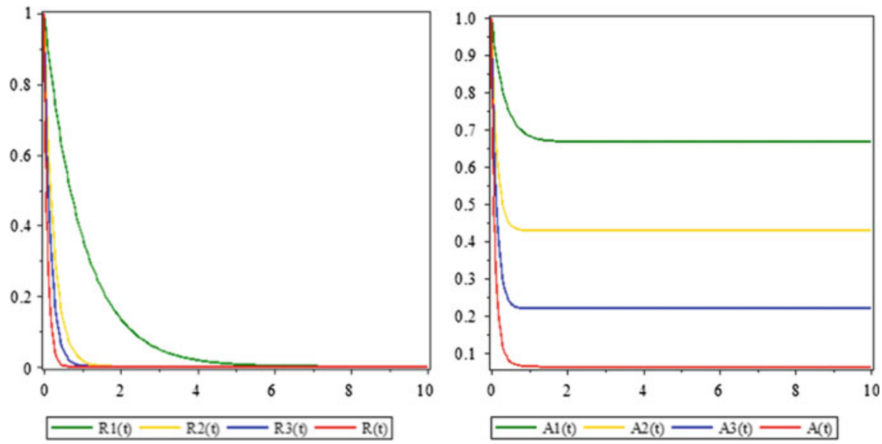


Fig. 6 Reliability and instantaneous availability for the first system

$$A_1^*(s) = [s + a_{12}^{(1)} - a_{12}^{(1)}(s + a_{21}^{(1)})^{-1}a_{21}^{(1)}]^{-1} = \frac{s+2}{s(s+3)}, A_1(t) = \frac{2+e^{-3t}}{3},$$
$$A_2^*(s) = [s + a_{12}^{(2)} - a_{12}^{(2)}(s + a_{21}^{(2)})^{-1}a_{21}^{(2)}]^{-1} = \frac{s+3}{s(s+7)}, A_2(t) = \frac{3+4e^{-7t}}{7},$$
$$A_3^*(s) = \{s + a_{12}^{(3)} + a_{13}^{(3)} - \begin{bmatrix} a_{12}^{(3)} & a_{13}^{(3)} \end{bmatrix} \begin{bmatrix} s + a_{21}^{(3)} + a_{23}^{(3)} & -a_{23}^{(3)} \\ a_{32}^{(3)} & s + a_{31}^{(3)} + a_{32}^{(3)} \end{bmatrix}^{-1} \begin{bmatrix} a_{21}^{(3)} \\ a_{31}^{(3)} \end{bmatrix}\}^{-1}$$
$$= \frac{s^2+7s+9}{s(s^2+13s+41)},$$
$$A_3(t) = \frac{9}{41} + \frac{4}{205}e^{-\frac{13}{2}t}[40\cosh(\frac{\sqrt{5}}{2}t) - 19\sqrt{5}\sinh(\frac{\sqrt{5}}{2}t)],$$
$$A(t) = \frac{1}{4305}(2+e^{-3t})(3+4e^{-7t})\{45 + 4e^{-\frac{13}{2}t}[40\cosh(\frac{\sqrt{5}}{2}t) - 19\sqrt{5}\sinh(\frac{\sqrt{5}}{2}t)]\}.$$

From Fig. 6, the steady-state availabilities are given as follows.

$$A_1^\infty = \frac{2}{3}, A_2^\infty = \frac{3}{7}, A_3^\infty = \frac{9}{41}, A^\infty = \frac{18}{287}.$$

(3) Interval availability

Interval availabilities for [0.1, 0.2] and [0, 1] are given as follows.

$$A_1[0.1, 0.2] = A_1(0.1)e^{-0.1} \approx 0.8267, A_2[0.1, 0.2] = A_2(0.1)e^{-4\times 0.1} \approx 0.4775,$$
$$A_3[0.1, 0.2] = A_3(0.1)e^{-6\times 0.1} \approx 0.3189, A[0.1, 0.2] \approx 0.1259,$$
$$A_1[0, 1] = R(1) = e^{-1} \approx 0.3679, A_2[0, 1] = R(1) = e^{-4} \approx 0.0183,$$
$$A_3[0, 1] = R(1) = e^{-6} \approx 0.0025, A[0, 1] = e^{-11} \approx 0.0000167.$$

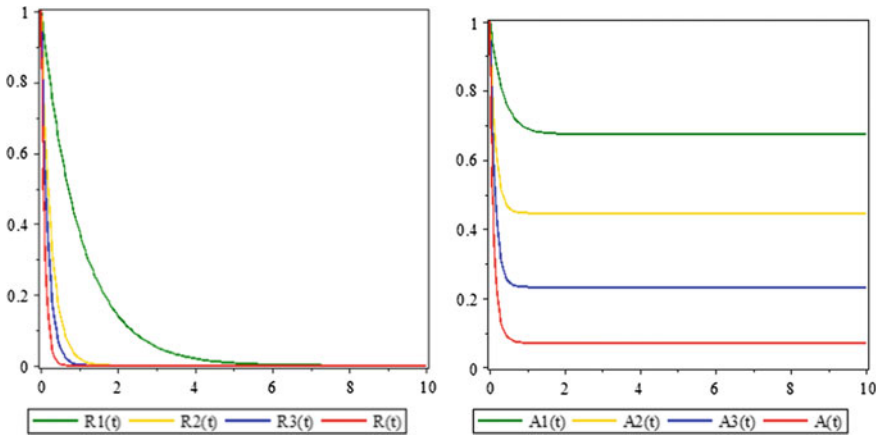


Fig. 7 Reliability and instantaneous availability for the second system

(2) The second system: general repairable series system with neglected failures $(\tau_i = 0.01, i = 1, 2, 3)$

(1) Reliability
The reliabilities of the three units and the system are shown in Fig. 7.

(2) Instantaneous availability
The instantaneous availabilities of the three units and the system are shown in Fig. 7.
From Fig. 7, the steady-state availabilities are given as follows.

$$\widetilde{A}_1^{\infty} \approx 0.6733, \widetilde{A}_2^{\infty} \approx 0.4455,\ \widetilde{A}_3^{\infty} \approx 0.2326, \widetilde{A}^{\infty} \approx 0.0698.$$

(3) Interval availability
Interval availabilities for [0.1, 0.2] and [0, 1] are given as follows.

$$\widetilde{A}_1[0.1, 0.2] \approx 0.8365, \widetilde{A}_2[0.1, 0.2] \approx 0.5020,$$
$$\widetilde{A}_3[0.1, 0.2] \approx 0.3417, \widetilde{A}[0.1, 0.2] \approx 0.1435,$$
$$\widetilde{A}_1[0, 1] = \widetilde{A}_1[0.01, 1] \approx 0.3789, \widetilde{A}_2[0, 1] = \widetilde{A}_2[0.01, 1] \approx 0.0215,$$
$$\widetilde{A}_3[0, 1] = \widetilde{A}_3[0.01, 1] \approx 0.0029, \widetilde{A}[0, 1] = \widetilde{A}[0.01, 1] \approx 0.0000224.$$

(3) The third system: phased-mission repairable series system $(T_i = 0.2, i = 1, 2, 3)$

(1) Reliability
The reliabilities of the three units and the system are given as follows and shown in Fig. 8.

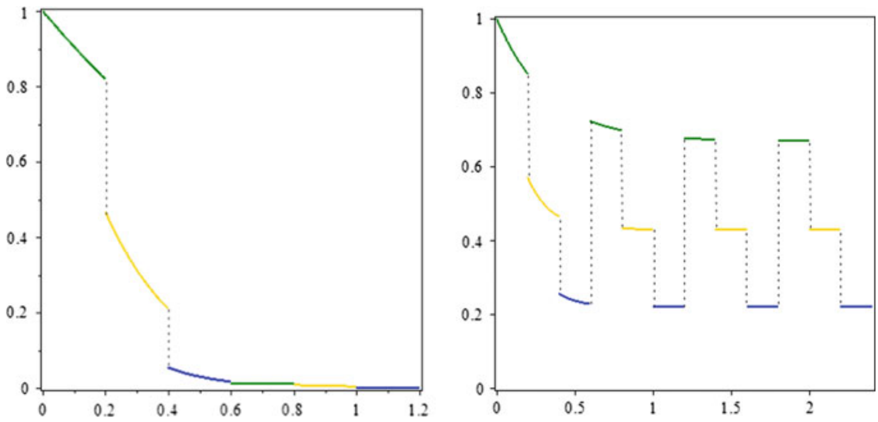


Fig. 8 Instantaneous availability for the third system

$$\vec{R}(t)=\begin{cases} A_1[0,t], & t\in(0,0.2],\\ A_1[0,0.2]A_2[0.2,t], & t\in(0.2,0.4],\\ A_1[0,0.2]A_2[0.2,0.4]A_3[0.4,t], & t\in(0.4,0.6],\\ A_1([0,0.2],[0.6,t])A_2[0.2,0.4]A_3[0.4,0.6], & t\in(0.6,0.8],\\ A_1([0,0.2],[0.6,0.8])A_2([0.2,0.4],[0.8,t])A_3[0.4,0.6], & t\in(0.8,1.0],\\ A_1([0,0.2],[0.6,0.8])A_2([0.2,0.4],[0.8,1.0])A_3([0.4,0.6],[1.0,t]), & t\in(1.0,1.2],\\ \ldots\ldots \end{cases}$$

(2) Instantaneous availability
The instantaneous availabilities of the three units and the system are given as follows and shown in Fig. 8.

$$\vec{A}(t)=\begin{cases} A_1(t), & t\in(0.2m,0.2m+0.2], & m=0,1,2,\ \ldots\\ A_2(t), & t\in(0.2m+0.2,0.2m+0.4], & m=0,1,2,\ \ldots\\ A_3(t), & t\in(0.2m+0.4,0.2m+0.6], & m=0,1,2,\ \ldots \end{cases}$$

(3) Interval availability
Interval availabilities for [0.1, 0.2] and [0, 1] are given as follows.

$$\vec{A}[0.1,0.2]=A_1[0.1,0.2]\approx 0.8267,$$
$$\vec{A}[0,1]=A_1([0,0.2],[0.6,0.8])A_2([0.2,0.4],[0.8,1.0])A_3[0.4,0.6]$$
$$\approx 0.5142\times 0.0817\times 0.0768\approx 0.0032.$$

(4) The fourth system: phased-mission repairable series system with neglected failures

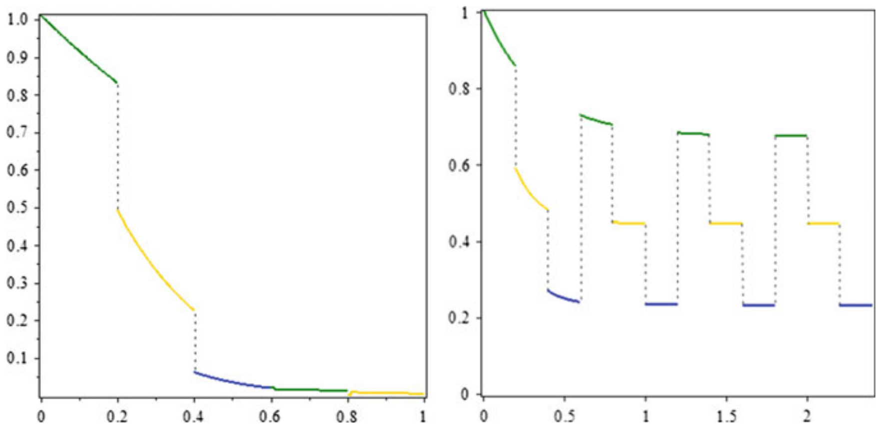


Fig. 9 Reliability and instantaneous availability for the fourth system

$$(\tau_i = 0.01, T_i = 0.2, i = 1, 2, 3)$$

(1) Reliability
The reliabilities of the three units and the system are shown in Fig. 9.
(2) Instantaneous availability
The instantaneous availabilities of the three units and the system are shown in Fig. 9.
(3) Interval availability
Interval availabilities for [0.1, 0.2] and [0, 1] are given as follows.

$$\vec{\widetilde{A}}[0.1, 0.2] = \widetilde{A}_1[0.1, 0.2] \approx 0.8365,$$

$$\vec{\widetilde{A}}[0, 1] = \widetilde{A}_1([0, 0.2], [0.6, 0.8])\widetilde{A}_2([0.2, 0.4], [0.8, 1.0])\widetilde{A}_3[0.4, 0.6]$$
$$\approx 0.5297 \times 0.0827 \times 0.0838 \approx 0.0036.$$

Example 2 In order to illustrate how theorems for the first and second system are applied to MSSs whose structure functions can be given by maximum operator and minimum operator, another numerical example is given for theorem 1 as an example. Assume that the underlying repairable system consists of three independent units which can be shown in Fig. 10 and the three units are exactly those in Example 1. Structure function of the system is given by $\phi = \min(\max(X_1, X_2), X_3)$, where X_i is performance level of unit i.

By state aggregation, state spaces of the three units are all divided into two subspaces and then multi-state system is simplified to binary system which makes the Lz-transform of the system much easier to be calculated. From Example 1, reliabilities of the three units are given as follows,

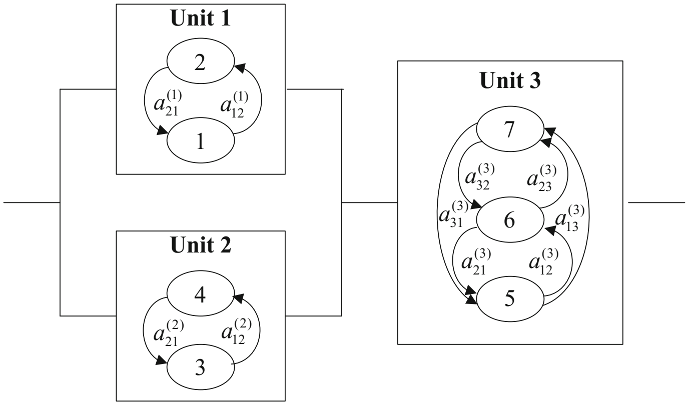


Fig. 10 Repairable series system with three independent units

$$R_1(t) = e^{-t}, R_2(t) = e^{-4t}, R_3(t) = e^{-6t}.$$

Then individual *Lz*-transform of each unit is as follows,

$$L_z\{G_i(t)\} = R_i(t)z^1 + [1 - R_i(t)]z^0,$$

where $\overline{R}_i(t) = 1 - R_i(t),\ i = 1, 2, 3.$

Using the Ushakov's UGO, we obtain the *Lz*-transform of the system as follows,

$$\begin{aligned} L_z\{G(t)\} &= \boldsymbol{\Omega_\phi}\{L_z\{G_1(t)\}, L_z\{G_2(t)\}, L_z\{G_3(t)\}\} \\ &= \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} [R_1(t)]^i[\overline{R}_1(t)]^{1-i}[R_2(t)]^j[\overline{R}_2(t)]^{1-j}[R_3(t)]^k[\overline{R}_3(t)]^{1-k}z^{\phi(i,j,k)} \\ &= \left\{R_1(t)R_2(t)R_3(t) + R_1(t)\overline{R}_2(t)R_3(t) + \overline{R}_1(t)R_2(t)R_3(t)\right\}z^1 + \left\{\overline{R}_1(t)\overline{R}_2(t)R_3(t)\right. \\ &\quad \left. + R_1(t)R_2(t)\overline{R}_3(t) + R_1(t)\overline{R}_2(t)\overline{R}_3(t) + \overline{R}_1(t)R_2(t)\overline{R}_3(t) + \overline{R}_1(t)\overline{R}_2(t)\overline{R}_3(t)\right\}z^0, \end{aligned}$$

Then the reliability of the system can be given as follows,

$$\begin{aligned} R(t) &= R_1(t)R_2(t)R_3(t) + R_1(t)\overline{R}_2(t)R_3(t) + \overline{R}_1(t)R_2(t)R_3(t) \\ &= [1 - \overline{R}_1(t)\overline{R}_2(t)]R_3(t) = e^{-7t} + e^{-10t} - e^{-11t}. \end{aligned}$$

Example 3 In order to illustrate how matrix method solves the problem of state exploration, a magnet resonance inspection (MRI) water cooling system in [5] is considered. Assume that the underlying repairable series system consists of four independent units—chillers subsystem, heat exchanger subsystem, pumps subsystem and electrical board subsystem, which can be shown in Fig. 11. Assume that the nominal performance level for the system is $24 \cdot 10^4$ BTU/h, where BTU means British thermal unit.

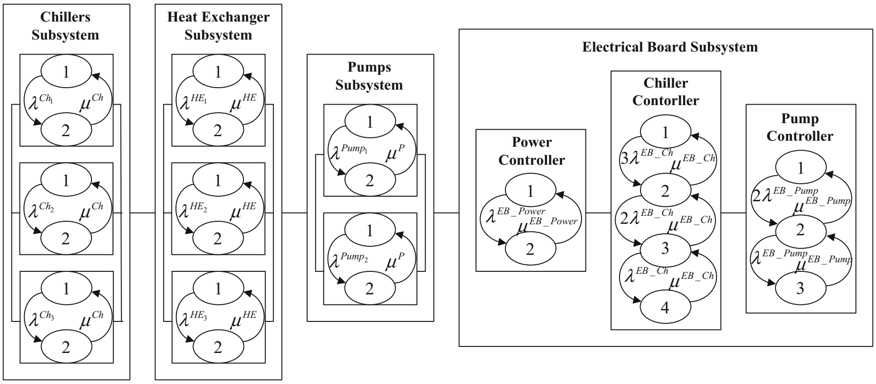


Fig. 11 Structure and state transition diagram of the water cooling system

The chillers subsystem consists of three independent identical chillers connected in parallel and each chiller has two states, state 1 for full cooling capacity $12 \cdot 10^4$ BTU/h and state 2 for zero capacity $0 \cdot 10^4$ BTU/h, then two chillers work simultaneously can cover the cooling load needed for the system. The failure rate of each chiller is $\lambda^{Ch_i} = 3\ \text{year}^{-1}$ and the repair rate of each chiller is $\mu^{Ch} = 365\ \text{year}^{-1}$.

The heat exchanger subsystem consists of three independent heat exchangers connected in parallel and each chiller has two states, state 1 for full cooling capacity and state 2 for zero capacity. Their cooling capacities are $16 \cdot 10^4$ BTU/h, $4 \cdot 10^4$ BTU/h and $4 \cdot 10^4$ BTU/h, respectively. Then three chillers work simultaneously can cover the cooling load needed for the system. Without consideration of system aging, assume that the failure rate of each heat exchanger is $\lambda^{HE_i} = 0.1\ \text{year}^{-1}$ and the repair rate of each chiller is $\mu^{HE} = 200\ \text{year}^{-1}$.

The pumps subsystem consists of two independent pumps connected in parallel and each chiller has two states, state 1 for full cooling capacity $24 \cdot 10^4$ BTU/h and state 2 for zero capacity $0 \cdot 10^4$ BTU/h. One pump in use can cover the cooling load needed for the system and two pumps work together won't make the pumps subsystem perform better. The failure rates of the two pumps are $\lambda^{Pumpi} = 3\ \text{year}^{-1}$ and the repair rate of each chiller is $\mu^{Pump} = 365\ \text{year}^{-1}$.

The electrical board system is used for power supply and control of chillers subsystem and pumps system. The power controller has two states, state 1 for fully operational state with performance level $24 \cdot 10^4$ BTU/h to the entire system and state 2 for complete failure with performance level $0 \cdot 10^4$ BTU/h to the entire system. The failure and repair rates of the power controller are $\lambda^{EB-Power} = 1\ \text{year}^{-1}$ and $\mu^{Pump_2} = 1000\ \text{year}^{-1}$, respectively.

The chiller controller has four states, state 1 and state 2 for fully operational states with performance level $24 \cdot 10^4$ BTU/h to the entire system, state 3 for partial operation with performance level $12 \cdot 10^4$ BTU/h to the entire system and state 4 for complete failure with performance level $0 \cdot 10^4$ BTU/h to the entire system. The failure and repair rates of the chiller controller are $\lambda^{EB-Ch} = 2$ year^{-1} and $\mu^{EB\_Ch} = 1000$ year^{-1}, respectively.

The pump controller has three has three states, state 1 and state 2 for fully operational states with performance level $24 \cdot 10^4$ BTU/h to the entire system and state 3 for complete failure with performance level $0 \cdot 10^4$ BTU/h to the entire system. The failure and repair rates of the pump controller are $\lambda^{EB-Pump} = 2$ year^{-1} and $\mu^{EB\_Pump} = 1000$ year^{-1}, respectively.

Assume that demand level of the system is $24 \cdot 10^4$ BTU/h and then the problem is to find reliability of the water cooling system. By *Lz*-transform method, we can get reliability of the system by the following steps [6]:

Step 1: Find *Lz*-transforms for all subsystems by using Ushakov's UGO.
Step 2: Find *Lz*-transform for the entire system by using Ushakov's UGO.
Step 3: Get underlying Markov process of the entire system by inverse *Lz*-transform.
Step 4: Investigate the Markov process to get reliability of the system.

From [5], a lot of work should be done in step 1. Take the first subsystem, chillers subsystem, for example. First, we should solve differential equations for each chiller $(i = 1, 2, 3)$: $\frac{dp_{i1}^{Ch}(t)}{dt} = -3p_{i1}^{Ch}(t) + 365p_{i2}^{Ch}(t)$, $\frac{dp_{i2}^{Ch}(t)}{dt} = 3p_{i1}^{Ch}(t) - 365p_{i2}^{Ch}(t)$ on initial condition $p_{i1}^{Ch}(0) = 1, p_{i2}^{Ch}(0) = 0$. Then we have $p_{i1}^{Ch}(t) = \frac{365}{368} + \frac{3}{368}e^{-368t}, p_{i2}^{Ch}(t) = \frac{3}{368} - \frac{3}{368}e^{-368t}$, and *Lz*-transform of each chiller is given as follows:

$$Lz\{G_i^{Ch}(t)\} = p_{i1}^{Ch}(t)z^{12} + p_{i2}^{Ch}(t)z^0.$$

Using Ushakov's UGO, we have *Lz*-transform of the chillers subsystem as follows,

$$\begin{aligned}
Lz\{G^{Ch}(t)\} &= \Omega_{fpar}\{Lz\{G_1^{Ch}(t)\}, Lz\{G_2^{Ch}(t)\}, Lz\{G_3^{Ch}(t)\}\} \\
&= \Omega_{fpar}\{p_{11}^{Ch}(t)z^{12} + p_{12}^{Ch}(t)z^0, p_{21}^{Ch}(t)z^{12} + p_{22}^{Ch}(t)z^0, p_{31}^{Ch}(t)z^{12} + p_{32}^{Ch}(t)z^0\} \\
&= p_{11}^{Ch}(t)p_{21}^{Ch}(t)p_{31}^{Ch}(t)z^{\min(12+12+12,24)} + p_{11}^{Ch}(t)p_{21}^{Ch}(t)p_{32}^{Ch}(t)z^{\min(12+12+0,24)} \\
&\quad + p_{11}^{Ch}(t)p_{22}^{Ch}(t)p_{31}^{Ch}(t)z^{\min(12+0+12,24)} + p_{11}^{Ch}(t)p_{22}^{Ch}(t)p_{32}^{Ch}(t)z^{\min(12+0+0,24)} \\
&\quad + p_{12}^{Ch}(t)p_{21}^{Ch}(t)p_{31}^{Ch}(t)z^{\min(0+12+12,24)} + p_{12}^{Ch}(t)p_{21}^{Ch}(t)p_{32}^{Ch}(t)z^{\min(0+12+0,24)} \\
&\quad + p_{12}^{Ch}(t)p_{22}^{Ch}(t)p_{31}^{Ch}(t)z^{\min(0+0+12,24)} + p_{12}^{Ch}(t)p_{22}^{Ch}(t)p_{32}^{Ch}(t)z^{\min(0+0+0,24)} \\
&= p_1^{Ch}(t)z^{24} + p_2^{Ch}(t)z^{12} + p_3^{Ch}(t)z^0,
\end{aligned}$$

where

$$\begin{aligned}
p_1^{Ch}(t) &= p_{11}^{Ch}(t)p_{21}^{Ch}(t)p_{31}^{Ch}(t) + p_{11}^{Ch}(t)p_{21}^{Ch}(t)p_{32}^{Ch}(t) \\
&\quad + p_{11}^{Ch}(t)p_{22}^{Ch}(t)p_{31}^{Ch}(t) + p_{12}^{Ch}(t)p_{21}^{Ch}(t)p_{31}^{Ch}(t), \\
p_2^{Ch}(t) &= p_{11}^{Ch}(t)p_{22}^{Ch}(t)p_{32}^{Ch}(t) + p_{12}^{Ch}(t)p_{21}^{Ch}(t)p_{32}^{Ch}(t) + p_{12}^{Ch}(t)p_{22}^{Ch}(t)p_{31}^{Ch}(t), \\
p_3^{Ch}(t) &= p_{12}^{Ch}(t)p_{22}^{Ch}(t)p_{32}^{Ch}(t).
\end{aligned}$$

As shown above, in step 1, each unit in chillers subsystem has two states, it's not hard to solve the differential functions, but for the chiller controller in electrical board subsystem who has four states, things will be harder. There are $2 \times 2 \times 2 = 8$ terms to be considered in *Lz*-transform of the chillers subsystem which is not too hard. However, we should repeat this work for the other three subsystems and there will be $2 \times 2 \times 2 = 8$ terms for heat exchanger subsystem, $2 \times 2 = 4$ terms for pumps subsystem and $2 \times 4 \times 3 = 24$ terms for the electrical board subsystem. What's more, in step 2, there will be $3 \times 4 \times 2 \times 3 = 72$ terms to be considered in *Lz*-transform of the entire system. From [6], step 3 is also not easy to be completed. To obtain infinitesimal generator of the entire system whose dimension is 5×5, genetic algorithm need to be used to generate all the 4×5 unknown coefficients in the matrix. Therefore, it's very difficult to get reliability of the system by *Lz*-transform.

However, for matrix method, it's not that hard and state exploration problem is solved to some extent by state aggregation. Reliability of the water cooling system can be given by three steps:

Step 1: Find infinitesimal generators for chillers subsystem, heat exchanger subsystem, pumps subsystem and the three controllers in electrical board subsystem as follows,

$$\boldsymbol{Q}^{Ch} = \left(\begin{array}{cc|cc} -9 & 9 & 0 & 0 \\ 365 & -371 & 6 & 0 \\ \hline 0 & 730 & -733 & 3 \\ 0 & 0 & 1095 & -1095 \end{array}\right), \boldsymbol{Q}^{HE} = \left(\begin{array}{c|ccc} -0.3 & 0.3 & 0 & 0 \\ \hline 200 & -200.2 & 0.2 & 0 \\ 0 & 400 & -400.1 & 0.1 \\ 0 & 0 & 600 & -600 \end{array}\right),$$

$$\boldsymbol{Q}^{Pump} = \left(\begin{array}{cc|c} -6 & 6 & 0 \\ 365 & -368 & 3 \\ \hline 0 & 730 & -730 \end{array}\right), \qquad \boldsymbol{Q}^{EB\_Power} = \left(\begin{array}{c|c} -1 & 1 \\ \hline 1000 & -1000 \end{array}\right),$$

$$\boldsymbol{Q}^{EB\_Ch} = \left(\begin{array}{cc|cc} -6 & 6 & 0 & 0 \\ 1000 & -1004 & 4 & 0 \\ \hline 0 & 1000 & -1002 & 2 \\ 0 & 0 & 1000 & -1000 \end{array}\right), \boldsymbol{Q}^{EB\_Pump} = \left(\begin{array}{cc|c} -4 & 4 & 0 \\ 1000 & -1002 & 2 \\ \hline 0 & 1000 & 0 \end{array}\right).$$

These infinitesimal generators are all divided into four parts according to state aggregation, which divides state space of corresponding Markov process into two subspaces, one for working states and the other one for failure states.

Step 2: Calculate reliabilities for the three subsystems and three controllers in step 1 by Theorem 1 as follows,

$$R^{Ch}(t)=\boldsymbol{\pi}_{WW}^{Ch}e^{\mathbf{Q}_{WW}^{Ch}t}\mathbf{u}_{WW}^{Ch}, R^{HE}(t)=\boldsymbol{\pi}_{WW}^{HE}e^{\mathbf{Q}_{WW}^{HE}t}\mathbf{u}_{WW}^{HE},$$
$$R^{Pump}(t)=\boldsymbol{\pi}_{WW}^{Pump}e^{\mathbf{Q}_{WW}^{Pump}t}\mathbf{u}_{WW}^{Pump}, R^{EB\_Power}(t)=\boldsymbol{\pi}_{WW}^{EB\_Power}e^{\mathbf{Q}_{WW}^{EB\_Power}t}\mathbf{u}_{WW}^{EB\_Power},$$
$$R^{EB\_Ch}(t)=\boldsymbol{\pi}_{WW}^{EB\_Ch}e^{\mathbf{Q}_{WW}^{EB\_Ch}t}\mathbf{u}_{WW}^{EB\_Ch}, R^{EB\_Pump}(t)=\boldsymbol{\pi}_{WW}^{EB\_Pump}e^{\mathbf{Q}_{WW}^{Ch}t}\mathbf{u}_{WW}^{Ch},$$

where vectors $\boldsymbol{\pi}_{WW}^{Ch}=(1 \quad 0), \boldsymbol{\pi}_{WW}^{HE}=1, \boldsymbol{\pi}_{WW}^{Pump}=(1 \quad 0), \boldsymbol{\pi}_{WW}^{EB\_Power}=1$, $\boldsymbol{\pi}_{WW}^{EB\_Ch}=(1 \quad 0)$, $\boldsymbol{\pi}_{WW}^{EB-Pump}=(1 \quad 0)$ are initial conditions, $\mathbf{u}_{WW}^{Ch}, \mathbf{u}_{WW}^{HE}, \mathbf{u}_{WW}^{Pump}, \mathbf{u}_{WW}^{EB-Power}, \mathbf{u}_{WW}^{EB-Ch}, \mathbf{u}_{WW}^{Ch}$ are vectors whose elements are 1 and

$$\mathbf{Q}_{WW}^{Ch}=\begin{pmatrix}-9 & 9\\ 365 & -371\end{pmatrix},\ \mathbf{Q}_{WW}^{HE}=-0.3,\ \mathbf{Q}_{WW}^{Pump}=\begin{pmatrix}-6 & 6\\ 365 & -368\end{pmatrix},$$
$$\mathbf{Q}_{WW}^{EB\_Power}=-1, \mathbf{Q}_{WW}^{EB\_Ch}=\begin{pmatrix}-6 & 6\\ 1000 & -1004\end{pmatrix},\ \mathbf{Q}_{WW}^{EB\_Pump}=\begin{pmatrix}-4 & 4\\ 1000 & -1002\end{pmatrix}.$$

Then by MAPLE software, we have,

$$R^{Ch}(t)\approx 1.00037438e^{-0.1421584t}-0.00037438e^{-379.8578416t}, R^{HE}(t)=e^{-0.3t},$$
$$R^{Pump}(t)\approx 1.00012874e^{-0.0481345t}-0.00012874e^{-373.9518655t}, R^{EB\_Power}(t)=e^{-t},$$
$$R^{EB\_Ch}(t)\approx 1.00002353e^{-0.0237629t}-1.00002353e^{-1009.976237t},$$
$$R^{EB\_Pump}(t)\approx 1.00000791e^{-0.0079523t}-0.00000791e^{-1005.992048t}.$$

Step 3: Get reliability of the entire system as follows,

$$\begin{aligned}R(t)&=R^{Ch}(t)R^{HE}(t)R^{Pump}(t)R^{EB\_Power}(t)R^{EB\_Ch}(t)R^{EB\_Pump}(t)\\ &\approx 1.00055025e^{-1.5378187t}-0.00012879e^{-375.4415497t}\\ &\quad -0.00037445e^{-381.2535019t}-0.00004708e^{-1011.490293t}.\end{aligned}$$

As shown above, all the three steps are not difficult and finally the result can be given in very simple form.

8 Conclusions

In this chapter, we consider repairable series systems consisting of n independent units and each of them follows a Markov process with finite state and continuous time. With state spaces of units divided into two subspaces for working states and failure states, four kinds of system are defined. The first system is general repairable system and the second system is general repairable system with neglected failures. The third system is phased-mission repairable series system and the forth system is phased-mission repairable series system with neglected failures. By aggregation of states and matrix method, problem of state explosion is solved to some extent. Reliability measures such as reliability, instantaneous availability and interval availability are discussed for the four kinds of system, respectively. The theorems for the first system and second system can be applied to MSSs whose structure functions can be given by maximum operator and minimum operator. Numerical examples are provided to illustrate results given in this chapter and present how matrix method solves the problem of state exploration. Obviously, values of reliability measures are larger with neglected failures in the numerical examples. The results in this chapter can be used in the analysis of repairable degradation systems in the field of reliability and many other fields.

Acknowledgements This work is supported by the NSF of China under grants 71371031 and 71631001.

References

1. Ushakov I (1986) A universal generating function. Sov J Comput Syst Sci 24:37–49
2. Levitin G (2005) The universal generating function in reliability analysis and optimization. Springer, London
3. Lisnianski A (2012) *Lz*-transform for a discrete-state continuous-time Markov process and its applications to multi-state system reliability. In: Lisnianski A, Frenkel I (eds) Recent advances in system reliability. Signatures, multi-state systems and statistical inference. Springer, London
4. Frenkel I, Lisnianski A, Khvatskin L (2012) Availability assessment for aging refrigeration system by using *Lz*-transform. J Reliab Stat Stud 5:33–43
5. Lisnianski A, Frenkel I, Khvatskin L (2015) On Birnbaum importance assessment for aging multi-state system under minimal repair by using the *Lz*-transform method. Reliab Eng Syst Safety 142:258–266
6. Lisnianski A, Ding Y (2016) Using inverse *Lz*-transform for obtaining compact stochastic model of complex power station for short-term risk evaluation. Rel Eng Syst Safety 145:19–27
7. Colquhoun D, Hawkes A (1977) Relaxation and fluctuations of membrane currents that flow through drug-operated channel. Proc R Soc Lond B Bio Sci 199:205–235
8. Colquhoun D, Hawkes A (1982) On the stochastic properties of bursts of single ion channel openings and of clusters of bursts. Philos Trans R Soc Lond B 300:1–59
9. Ball F, Milne R, Yeo G (2002) Multivariate semi-Markov analysis of burst properties of multiconductance single ion channels. J Appl Probab 39:179–196

10. Cui L, Li H, Li J (2007) Markov repairable systems with history-dependent up and down states. Stoch Model 23:665–681
11. Hawkes A, Cui L, Zheng Z (2011) Modeling the evolution of system reliability performance under alternative environments. IIE Trans 43:761–772
12. Hawkes A, Jalali A, Colquhoun D (1990) The distributions of the apparent open times and shut times in a single channel record when brief events cannot be detected. Philos Trans R Soc Lond A 332:511–538
13. Ball F (1997) Empirical clustering of bursts of openings in Markov and semi-Markov models of single channel gating incorporating time interval omission. Adv Appl Probab 29:909–946
14. Zheng Z, Cui L, Hawkes A (2006) A study on a single-unit Markov repairable system with repair time omission. IEEE Trans Reliab 5:182–188
15. Bao X, Cui L (2010) An analysis of availability for series Markov repairable system with neglected or delayed failures. IEEE Trans Reliab 59:734–743
16. Quandt R (1958) The estimation of the parameters of a linear regression system obeying two separate regimes. J Am Stat Assoc 53:873–880
17. Goldfeld S, Quandt R (1973) A Markov model for switching regressions. J Econom 1:3–15
18. Lim T (1998) A stochastic regime switching model for the failure process of a repairable system. Rel Eng Syst Safety 59:225–238
19. Shen J, Cui L (2016) Reliability performance for dynamic systems with cycles of K regime. IIE Trans 48:389–402
20. Li Y, Cui L, Yi H (2016) Reliability of non-repairable systems with cyclic-mission switching and multimode failure components. J Comput Sci 17:126–138
21. Baxter L (1981) Availability measures for a two-state system. J Appl Probab 18:227–235
22. Sericola B (1994) Interval-availability distribution of 2-state systems with exponential failures and phase-type repairs. IEEE Trans Reliab 43:335–343
23. Goyal A, Tantawi A (1988) A measure of guaranteed availability and its numerical evaluation. IEEE Trans Comput 37:25–32
24. Goyal A, Tantawi A, Trivedi K (1985) A measure of guaranteed availability. IBM research report RC 11341
25. Csenki A (1994) Joint availability of systems modeled by finite semi-Markov processes. Appl Stoch Model Data Anal 10:279–293
26. Csenki A (2007) Joint interval reliability for Markov systems with an application in transmission line reliability. Reliab Eng Syst Safety 92:685–696
27. Cui L, Du S, Liu B (2013) Multi-point and multi-interval availabilities. IEEE Trans Reliab 62:811–820
28. Davis M (1984) Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. J R Stat Soc Serv B 46:353–388
29. Zhang Q, Cui L, Yi H (2017) A study on a single-unit repairable system with working and repair time omission under an alternative renewal process. J Risk Reliab 231:232–241

Dynamic Performance of Series Parallel Multi-state Systems with Standby Subsystems or Repairable Binary Elements

Gregory Levitin and Liudong Xing

Abstract This chapter presents a method for evaluating dynamic performance of multi-state systems with a general series parallel structure. The system components can be either repairable binary elements with given time-to-failure and repair time distributions, or 1-out-of-N warm standby configurations of heterogeneous binary elements characterized by different performances and time-to-failure distributions. The entire system needs to satisfy a random demand defined by a time-dependent distribution. Iterative algorithms are presented for determining performance stochastic processes of individual components. A universal generating function technique is implemented for evaluating the dynamic system performance indices. Examples are provided to demonstrate applications of the proposed methodology.

Keywords Multi-state system • Repair • Warm standby • Stochastic process • Instantaneous availability • Unsupplied demand

Acronyms

cdf	Cumulative distribution function
DSCTP	Discrete-state continuous-time process
MSS	Multi-state system
pdf	Probability density function
RBD	Reliability block diagram
UGF	Universal generating function (u-function)

G. Levitin (✉)
The Israel Electric Corporation, P.O. Box 10, 31000 Haifa, Israel
e-mail: levitin@iec.co.il

L. Xing
University of Massachusetts Dartmouth, 285 Old Westport Rd., Dartmouth, MA 02747, USA
e-mail: lxing@umassd.edu

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_8

Nomenclature

T	Mission time
I	Number of system components
N_i	Number of elements within component i
$G_i(t)$	Random performance of component i at time t
$g_{i,k}$	K-th possible performance level of component i
$\boldsymbol{g}_i$	Performance level vector of component i: $g_i = \{g_{i,0}, \ldots, g_{i,N_i}\}$
$p_{i,k}(t)$	Probability that component i operates at level $g_{i,k}$ at time t, *i.e.*, $\Pr(G_i(t) = g_{i,k})$
$\phi(G_1, \ldots, G_I)$	System structure function
$V(t)$	Random system performance at time t
$\boldsymbol{v}$	Vector of possible performance levels of MSS $v = \{v_0, \ldots, v_K\}$
$w_j(t)$	$\Pr(V(t) = v_j)$
D(t)	Random system demand at time t
$\boldsymbol{d}$	Vector of possible system demand levels: $\boldsymbol{d} = \{d_0, \ldots, d_L\}$
$h_l(t)$	$\Pr(D(t) = d_l)$
c(t)	Expected system performance at time t
C(T)	Expected amount of work system can complete in time T
$e(t)$	Expected instantaneous unsupplied demand at time t
E	Expected unsupplied demand over mission time T
$a(t)$	Expected instantaneous system availability at time t
A	Expected system availability over mission time T
θ	Predetermined amount of work to be completed by system,
$\tau\ (\theta)$	Expected time of completing amount of work θ
$s_i(k)$	Index determining type of element that should be activated after the $(k-1)$th element failure in component i
T_k	Random variable representing the time when the last element from sequence $s_i(1), \ldots, s_i(k)$ fails
$q_k(t)$	*pdf* of random variable T_k
$F_{i,j}(t), f_{i,j}(t)$	*cdf, pdf* of lifetime of element j within component i in the operation mode
$\omega_{i,j}$	Nominal performance of element j within standby component i
δ_{ij}	Deceleration factor of element j within component i
$\eta_{i,j}, \beta_{i,j}$	Scale, shape parameters of Weibull time-to-failure distribution for element j within component i
$\gamma_i^{\min}, \gamma_i^{\max}$	Minimum, maximum repair time of element i
J_i	Maximal number of failures of element i
π_i	Repair efficiency of element i
ω_i	Nominal performance of repairable element i
$\zeta_i(t)$	Hazard rate of element i
X_j	Random time spent by element in operation mode before the j-th failure

$Q_j(t, x)$	Joint distribution of j-th failure event parameters
Γ	Random repair time
$\Psi_i(t)$, $\psi_i(t)$	*Cdf*, *pdf* of random repair time
$r_{ij}(t)$	Probability that element i is under repair after the j-th failure at time t

1 Introduction

Many real-world systems, such as those with shared loads, performance degradation, standby sparing, imperfect coverage, or limited repair resources are multi-state systems (MSSs) [1, 2]. In MSSs, the system and/or its components can exhibit multiple different states or performance levels [3]. MSSs abound in applications including (but not limited to) medical systems [4], power systems [5, 6], computing systems [7], communication systems [8], and transportation systems [9, 10]. Due to their critical applications, the MSS modeling and analysis have attracted lots of research efforts. Diverse types of methods have been developed for MSS analysis including for example, multi-state path/cut-vector based enumerations [3, 11], simulations [12], branch-and-bound technique [13], *Lz*-transform techniques based on Markov processes [14–16], universal generating functions (UGF) [3, 9, 17], and binary or multi-valued decision diagrams [1, 8, 18, 19].

This chapter focuses on a class of MSSs with the general series parallel structure. Different from the traditional structure of multi-state series parallel systems that has been intensively studied [9, 20–23], the system considered in this chapter contains an arbitrary combination of series and parallel configurations of system components. Each system component can be either a warm standby configuration of basic binary functional elements or a repairable binary element. In contrast, the traditional structure contains subsystems connected in a purely series configuration with each subsystem being a purely parallel configuration of functional components.

An iterative algorithm is first presented for determining the performance discrete-state continuous-time process (DSCTP) of an individual component in the considered system. A UGF-based technique is then applied for evaluating system-level performance indices of expected system availability and unsupplied demand. Note that the integrated DSCTP and UGF technique has been applied to model dynamic behavior of MSSs without general standby redundancies in [4–6, 14–16]. These works are based on the Markov process model, thus being limited to exponential element time-to-failure distributions. In this chapter we extend the dynamic MSS model to considering repairable elements with arbitrary repair time distributions and to considering warm standby components (or subsystems) composed of elements with non-identical, arbitrary time-to-failure distributions.

The rest of this chapter is organized as follows. Section 2 presents the generic model and performance metrics of the MSS considered. Section 3 presents algorithms for obtaining the performance DSCTP for system components consisting of either standby elements or repairable binary elements. Section 4 gives examples of the DSCTP evaluation. Section 5 describes the UGF technique for obtaining the DSCTP of the entire system performance based on DSCTPs of its components' performances. Section 6 presents illustrative examples of obtaining the system performance DSCTPs. Section 7 concludes the chapter and outlines the optimization problems that can be solved based on the presented methodology.

2 System Model and Performance Metrics

Two types of MSSs are considered in this chapter, both of which contain I s-independent components composing a general series parallel structure.

In the first type of MSSs, each component i consists of N_i non-repairable binary elements configured as a 1-out-of-N_i warm standby structure, where one element is online and functioning with the remaining elements being kept in a warm standby mode. In the case of the online element failing, according to a pre-defined sequence a standby element is activated to take over the task of the component. If the chosen standby element is not available (fails before it should be activated), the next element in the sequence is checked etc. Elements within the same component can be non-identical, characterized by different time-to-failure distributions and nominal performance rates. Thus, depending on the element functioning at the moment, the performance $G_i(t)$ of each component i can vary dynamically, and be modeled using a DSCTP.

In the second type of MSSs, each component consists of a single binary repairable element with known time-to-failure and repair time distributions as well as nominal performance. Depending on the status of the element at the instant of time t, the performance $G_i(t)$ of component i can dynamically change from zero (failure) to nominal (operation), which constitutes a DSCTP with two discrete states.

The entire system needs to meet a random demand that can be specified by a distribution depending on weather conditions, time of day, season, etc. The demand is also modeled using a DSCTP. In many applications the demand distribution is obtained empirically for specific time periods (times of day, seasons, parts of production cycle etc.)

The considered models are based on the following assumptions.

- The time-to failure and repair time distributions of elements are independent.
- Different components are statistically independent.
- The failure detection is perfect.
- The repair/replacement starts immediately after the failure.
- Specific elements have fixed performance during their operation.
- All the system elements are available in the beginning of the mission.

2.1 Generic Model of Series Parallel MSSs

For modelling behaviour of an MSS, it is necessary to analyse characteristics of its components. In the first type of MSSs, any system component i can assume $N_i + 1$ states, corresponding to different performance levels or rates. Specifically, component i's performance rate at a time instant t can be modelled using a discrete random variable $G_i(t) \in g_i = \{g_{i,0}, g_{i,1}, \ldots, g_{i,N_i}\}$. Let $p_{i,k}(t)$ be the probability that component i operates at level $g_{i,k}$ at time t, i.e., $p_{i,k}(t) = \Pr\{G_i(t) = g_{i,k}\}$. The two vectors $p_i(t) = \{p_{i,0}(t), p_{i,1}(t), \ldots, p_{i,N_i}(t)\}$ and $g_i = \{g_{i,0}, g_{i,1}, \ldots, g_{i,N_i}\}$ can determine the performance distribution of random variable $G_j(t)$ at any time instant t. The second type of MSSs can be considered as a special case of the above model, where $p_i(t) = \{p_{i,0}(t), p_{i,1}(t)\}$ and $g_i = \{g_{i,0}, g_{i,1}\} = \{0, \omega_i\}$ determining the DSCTP $G_i(t)$ of component i consisting of a single repairable binary-state element with nominal performance ω_i.

Performance rates of its constituent components unambiguously determine the performance rate of a system; their mapping relation can be defined by a function called system structure function $\phi(G_1, \ldots, G_I)$. The system structure function and probability mass functions (*pmf*) of performances of system elements at any time instant t give the generic model of the considered MSS, as shown in (1).

$$g_i, p_i(t), 1 \le i \le I, 1 \le t \le T,$$
$$V(t) = \phi(G_1(t), \ldots, G_I(t)), \quad V(t) \in \{v_1, \ldots, v_K\} \tag{1}$$

The DSCTP $V(t)$ in (1) can be determined by *pmf* of the entire system performance at t as

$$w_k(t) = \Pr\{V(t) = v_k\}, \text{ where } 0 \le k \le K. \tag{2}$$

The system must meet a random demand $D(t)$, defined by two vectors $\boldsymbol{d} = \{d_0, \ldots, d_L\}$ and $\boldsymbol{h}(t) = \{h_0(t), \ldots, h_L(t)\}$, where $h_l(t) = \Pr\{D(t) = d_l\}$ for $l = 0, 1, \ldots, L$.

2.2 MSS Dynamic Performance Metrics

Based on the DSCTPs of $V(t)$ and $D(t)$, the following dynamic performance metrics can be defined for the considered MSSs.

- The expected system performance at time t

$$c(t) = \sum_{k=0}^{K} v_k w_k(t) \tag{3}$$

- The expected amount of work the system can complete in time T

$$C(T) = \int_0^T c(t)dt = \int_0^T \sum_{k=0}^{K} v_k w_k(t)dt; \tag{4}$$

- The expected instantaneous system availability at time t

$$a(t) = \sum_{l=0}^{L} \left(h_l(t) \sum_{k=0}^{K} w_k(t) 1(d_l \le v_k) \right); \tag{5}$$

- The expected system availability during the system mission time T

$$A = \frac{1}{T}\int_0^T a(t)dt = \frac{1}{T}\int_0^T \sum_{l=0}^{L} \left(h_l(t) \sum_{k=0}^{K} w_k(t) 1(d_l \le v_k) \right) dt; \tag{6}$$

- The expected instantaneous unsupplied demand at time t

$$e(t) = \sum_{l=0}^{L} \left(h_l(t) \sum_{k=0}^{K} w_k(t) \max(0, d_l - v_k) \right) \tag{7}$$

- The expected total unsupplied demand during mission time T

$$E = \int_0^T e(t)dt = \int_0^T \sum_{l=0}^{L} \left(h_l(t) \sum_{k=0}^{K} w_k(t) \max(0, d_l - v_k) \right) dt. \tag{8}$$

- If the system must complete a predetermined amount of work θ, the expected mission time is

$$\tau(\theta) = \arg\left(\int_0^\tau \sum_0^K v_k w_k(t) dt = \theta \right). \tag{9}$$

To evaluate system performance metrics (3)–(9), an iterative algorithm is presented in Sect. 3, which is used for obtaining the description of DSCTP characterizing components' performances in the form of $\boldsymbol{g}_i$, $\boldsymbol{p}_i(t)$, $1 \le i \le I, 1 \le t \le T$ in (1).

Then a generalized reliability block diagram (RBD) method based on the UGF technique is implemented in Sect. 5 to derive the description of the DSCTP for the entire MSS performance in the form of (2).

3 Obtaining Performance DSCTP for Individual Components

The DSCTP $G_i(t)$ for each individual component i is derived for both types of MSSs in this section.

3.1 Performance DSCTP for Warm Standby Components

To determine the DSCTP $G_i(t)$ for 1-out-of-N_i standby component i in the first type of MSS, vector $p_i(t) = \{p_{i,0}(t), p_{i,1}(t), \ldots, p_{i,N_i}(t)\}$ is derived in this subsection while vector $g_i = \{g_{i,0}, g_{i,1}, \ldots, g_{i,N_i}\}$ is given as input parameters, where $g_{i,0} = 0$, $g_{i,k} = \omega_{i,s_i(k)}$ for $k = 1, \ldots, N_i$.

Let the order $s_i(1)$, $s_i(2)$, …, $s_i(N_i)$ determine the predetermined activation sequence of elements composing component i, T_k be a random variable modeling the time when the last element from sequence $s_i(1)$, …, $s_i(k)$ fails during the operation mode, and $q_k(t)$ be the probability density function (*pdf*) of this random variable. For $k = 1$, since only one element $s_i(1)$ belongs to the sequence, $q_1(t) = f_{i,s_i(1)}(t)$, where $f_{i,s_i(1)}(t)$ is the *pdf* of element $s_i(1)$.

With $q_{k-1}(t)$ and $f_{i,s_i(k)}(t)$ $q_k(t)$ can be derived for $k = 2$, …, N_i. Specifically, there exist two scenarios that can cause failure of the last element from sequence $s_i(1)$, …, $s_i(k)$ at time t.

1. Scenario 1: $T_k = T_{k-1} = t$. The last element from sequence $s_i(1)$, …, $s_i(k-1)$ fails at time t; element $s_i(k)$ fails earlier during the standby mode. This scenario can occur with probability $F_{i,s_i(k)}(t)$, where $F_{i,s_i(k)}(t)$ is the *cdf* of element $s_i(k)$.
2. Scenario 2: $T_k = t$, $T_{k-1} = t - \tau$. The last element from sequence $s_i(1)$, …, $s_i(k-1)$ fails at certain time before t, e.g., $t - \tau$ for $0 \le \tau \le t$; element $s_i(k)$ fails after spending $(t - \tau)$ in the standby mode and then working for time τ in the operation mode.

Based on the two scenarios, *pdf* of T_k is

$$q_k(t) = q_{k-1}(t)F_{i,s_i(k)}(t) + \int_0^t q_{k-1}(t-\tau)f_{i,s_i(k)}((t-\tau)\delta_{i,s_i(k)} + \tau)d\tau. \tag{10}$$

In (10), $0 \le \delta_{i,s_i(k)} \le 1$ represents a deceleration factor of element $s_i(k)$ within component i. Such a factor is utilized to reflect lower stresses experienced by the element during the warm standby mode than during the operation mode in the commonly-used cumulative exposure model [24]. Based on (10), $q_k(t)$ can be obtained iteratively for $k = 2$, …, N_i.

Next the derivation of vector $p_i(t) = \{p_{i,0}(t), p_{i,1}(t), \ldots, p_{i,N_i}(t)\}$ is given. The probability $p_{i,1}(t)$ that component i operates with performance level $\omega_{i,s_i(1)}$ provided by element $s_i(1)$ at time t is simply the probability that element $s_i(1)$ does not fail before time t, which is given as:

$$p_{i,1}(t) = 1 - F_{i,s_i(1)}(t). \tag{11}$$

The probability $p_{i,k}(t)$ that component i works with performance level $\omega_{i,s_i(k)}$ (i.e., element $s_i(k)$ is operational) at time t can be evaluated as the probability that $T_{k-1} = t - \tau$ for any $0 \le \tau \le t$ and element $s_i(k)$ waiting for time $t - \tau$ in the warm standby mode does not fail before spending at least time τ in the operation mode:

$$p_{i,k}(t) = \int_0^t q_{k-1}(t-\tau)\left[1 - F_{i,s_i(k)}(\delta_{i,s_i(k)}(t-\tau) + \tau)\right]d\tau \tag{12}$$

The probability that component i's performance is zero (i.e., all the elements of component i fail) at time t is thus

$$p_{i,0}(t) = 1 - \sum_{k=1}^{N_i} p_{i,k}(t) \tag{13}$$

3.2 Performance DSCTP for Repairable Binary Elements

To determine the DSCTP $G_i(t)$ for component i consisting of a single repairable element in the second type of MSSs, vector $\boldsymbol{p}_i(t) = \{p_{i,0}(t), p_{i,1}(t)\}$ is derived in this subsection, while vector $g_i = \{0, \omega_i\}$ is given as input parameters meaning that element i functions with nominal performance ω_i and has performance 0 while under repair.

It is assumed that the repair starts immediately when an element fails. The repair time of element i is dependent on external factors such as availability and efficiency of repair manpower and equipment. Assume the repair time of element i is randomly distributed within interval $[\gamma_i^{\min}, \gamma_i^{\max}]$ $(0 > \gamma_i^{\min} > \gamma_i^{\max} > \infty)$ with known *cdf* $\Psi_i(t)$ ($\Psi_i(t) \equiv 0$ for $t < \gamma_i^{\min}$, $\Psi_i(t) \equiv 1$ for $t > \gamma_i^{\max}$).

The number of repairs experienced by element i during considered mission time T cannot exceed $T/\gamma_i^{\min}$. Thus, the maximal number of failures that can be experienced by this element is $J_i = 1 + \lfloor T/\gamma_i^{\min} \rfloor$.

According to the repair model in [25], a coefficient π_i can be used to model the repair efficiency of element i. Particularly, π_i can vary from 0 corresponding to perfect repair (the element after repair is as good as new) to 1 corresponding to minimal repair (the element after repair is as bad as old). Under the model of [25], for element i with hazard rate $\zeta_i(t)$ before a repair, its hazard rate after the repair is $\zeta_i(\pi_i t_0 + t)$, where t_0 and t represent operation times of element i before and after the repair, respectively. The *pdf* $f_i^*(t_0, t)$ and *cdf* $F_i^*(t_0, t)$ of the time-to-failure of element i after the repair performed at time t_0 are, respectively,

$$f_i^*(t_0,t) = f_i(\pi_i t_0 + t)/[1 - F_i(\pi_i t_0)] \tag{14}$$

and

$$F_i^*(t_0,t) = [F_i(\pi_i t_0 + t) - F_i(\pi_i t_0)]/[1 - F_i(\pi_i t_0)] \tag{15}$$

For element i, consider a random event denoted by $\langle T_j, X_j \rangle$, meaning that the j-th failure of element i occurs at time T_j from the beginning of the mission after element i has spent time $X_j \le T_j$ in the operation mode and additional time $T_j - X_j$ in repair. Each event $\langle T_j, X_j \rangle$ corresponds to the initiation of a repair action that takes random time Γ. For any realization of X_j, the time elapsed from the beginning of the mission T_j can range from $X_j + (j-1)\gamma_i^{\min}$ to $X_j + (j-1)\gamma_i^{\max}$, corresponding to the cases where the element spends minimal and maximal time in each of the $j - 1$ repairs, respectively.

Define $Q_j(t, x)$ as the joint distribution of random event parameters T_j and X_j. Because the element spends no time in repair before the first failure and $X_1 = T_1$

$$Q_1(t,x) = \begin{cases} f_i(t) & \text{if } (x = t) \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

The element in the second type of MSSs must have event transition from $\langle T_j, X_j \rangle$ to $\langle T_{j+1}, X_{j+1} \rangle$ with $T_{j+1} \ge T_j + \gamma_i^{\min}$ and $X_{j+1} \ge X_j$. Note that when element i fails immediately after the j-th repair, $X_{j+1} = X_j$ The event transition $\langle T_j, X_j \rangle \to \langle T_{j+1}, X_{j+1} \rangle$ happens when the element has functioned for time $\left(X_{j+1} - X_j\right)$ after a repair that takes time $\Gamma = \left(T_{j+1} - T_j\right) - \left(X_{j+1} - X_j\right)$. Because $\gamma_i^{\min} \le \Gamma \le \gamma_i^{\max}$, the condition (17) must hold to make the event transition possible.

$$T_{j+1} + X_j - X_{j+1} - \gamma_i^{\max} \le T_j \le T_{j+1} + X_j - X_{j+1} - \gamma_i^{\min} \tag{17}$$

With functions $Q_j(t,x)$, $\psi_i(t)$ and $f_i(t)$, $Q_{j+1}(t,x)$ can be obtained in a recursive manner for $j = 1, \ldots, J_i - 1$ as

$$\begin{aligned} Q_{j+1}(t,x) &= \int_0^x \int_{\max\left(\tilde{x} + (j-1)\gamma_i^{\min},\, t + \tilde{x} - x - \gamma_i^{\max}\right)}^{\min\left(\tilde{x} + (j-1)\gamma_i^{\max},\, t + \tilde{x} - x - \gamma_i^{\min}\right)} Q_j(\tilde{t},\tilde{x})\psi_i(t - \tilde{t} - x + \tilde{x}) f_i^*(\tilde{x}, x - \tilde{x}) d\tilde{t} d\tilde{x} \\ &= \int_0^x \int_{\max\left(\tilde{x} + (j-1)\gamma_i^{\min},\, t + \tilde{x} - x - \gamma_i^{\max}\right)}^{\min\left(\tilde{x} + (j-1)\gamma_i^{\max},\, t + \tilde{x} - x - \gamma_i^{\min}\right)} Q_j(\tilde{t},\tilde{x})\psi_i(t - \tilde{t} - x + \tilde{x}) \frac{f_i(\pi_i \tilde{x} + x - \tilde{x})}{1 - F_i(\pi_i \tilde{x})} d\tilde{t} d\tilde{x}. \end{aligned} \tag{18}$$

Note that for $t < x + j\gamma_i^{\min}$ or $t > x + j\gamma_i^{\max}$, $Q_{j+1}(t,x) = 0$.

The element can be under repair at time t since the mission beginning if the last failure occurred at time $t-\xi$ and the last repair took at least time ξ Hence, the probability that the element is under repair at time t after the occurrence of the j-th failure is

$$r_{ij}(t) = 0 \text{ for } t < (j-1)\gamma_i^{\min} \tag{19}$$

$$r_{ij}(t) = \int_0^{\min\left(t,\gamma_i^{\max}\right)} \int_{\max\left(0,t-\xi-(j-1)\gamma_i^{\max}\right)}^{t-\xi-(j-1)\gamma_i^{\min}} Q_j(t-\xi,x)(1-\Psi_i(\xi))dxd\xi \text{ for } t \geq (j-1)\gamma_i^{\min} \tag{20}$$

Observe that for any $k \neq j$ $r_{ij}(t)$ and $r_{ik}(t)$ are probabilities of mutually disjoint events corresponding to different numbers of failures occurred before time t. Therefore the overall probability that the element i undergoes repair at time t can be obtained as sum of probabilities $r_{ij}(t)$ for all the possible numbers j of failure/repair events. Because the minimal time when the j-th element failure may occur is $(j-1)\gamma_i^{\min}$, the number of failures that can occur at time not later than t cannot exceed $1+t/\gamma_i^{\min}$. The overall occurrence probability that element i is under repair at time t is thus

$$p_{i,0}(t) = \sum_{j=1}^{\lfloor 1+t/\gamma_i^{\min} \rfloor} r_{ij}(t) = \sum_{j=1}^{\lfloor 1+t/\gamma_i^{\min} \rfloor} \int_0^{\min\left(t,\gamma_i^{\max}\right)} \int_{\max\left(0,t-\xi-(j-1)\gamma_i^{\max}\right)}^{t-\xi-(j-1)\gamma_i^{\min}} Q_j(t-\xi,x)(1-\Psi_i(\xi))dxd\xi. \tag{21}$$

For the binary repairable element i, $p_{i,1}(t) = 1 - p_{i,0}(t)$ which defines the component's instantaneous availability.

4 Examples of Component Performance Evaluation

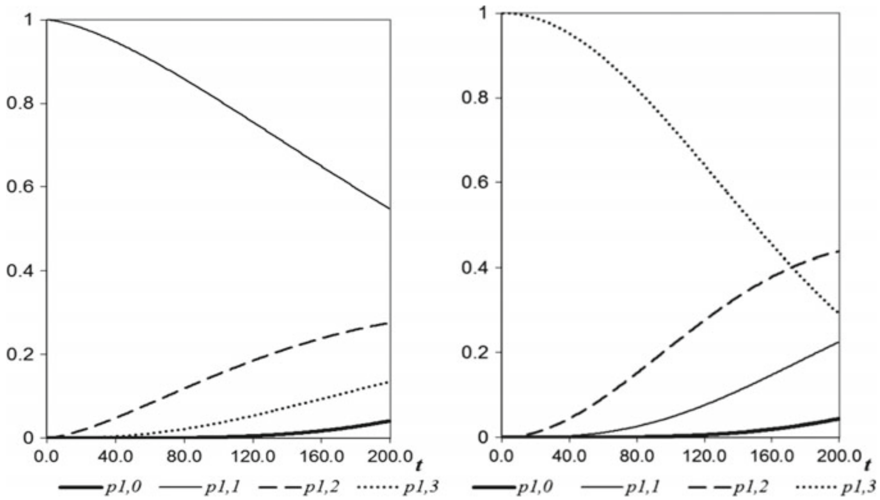
4.1 Warm Standby Components

Consider a warm standby component denoted by component 1, consisting of three elements characterized by Weibull time-to-failure distributions. The scale (η_j) and shape (β_j) parameters of the distributions, deceleration factor (δ_j) and nominal performance (ω_j) of elements are presented in Table 1.

Figure 1 illustrates the performance level probabilities $p_{1,j}(t) = \Pr\{G_1(t) = \omega_j\}$ for two different element activation sequences 1, 2, 3 and 3, 2, 1. According to Table 1, the component performance can take three different values, $G_1(t) \in$ (20, 27, 32). Thus,

Table 1 Parameters of elements composing standby component 1

Element j	η_j	β_j	δ_j	ω_j
1	280	1.5	0.2	20
2	250	1.1	0.4	27
3	180	2	0.2	32



(A) Sequence 1,2,3 (B) Sequence 3,2,1

Fig. 1 Performance level probabilities $p_{1,j}(t)$ [27]. **a** Sequence 1, 2, 3 **b** Sequence 3, 2, 1

$$\Pr(G_1(t) \geq 20) = \Pr(G_1(t) > 0) = \Pr(G_1(t) = 20) + \Pr(G_1(t) = 27) + \Pr(G_1(t) = 32),$$
$$\Pr(G_1(t) \geq 27) = \Pr(G_1(t) = 27) + \Pr(G_1(t) = 32), \Pr(G_1(t) \geq 32) = \Pr(G_1(t) = 32).$$

Figure 2 shows the cumulative performance distribution $\Pr(G_1(t) \geq x)$ of component 1 under the two activation sequences. It can be observed that the probability $\Pr(G_1(t) \geq 32)$ is always larger for sequence 3, 2, 1 where element 3 with the greatest performance is activated first; the probability $\Pr(G_1(t) \geq 20)$ is always slightly larger for sequence 1, 2, 3 where the most reliable element 1 is activated first.

4.2 Repairable Binary Element

Consider a repairable element with a Weibull time-to-failure distribution having *cdf* of $F(t) = 1 - \exp\left[(-t/20)^2\right]$. The random repair time follows a truncated normal

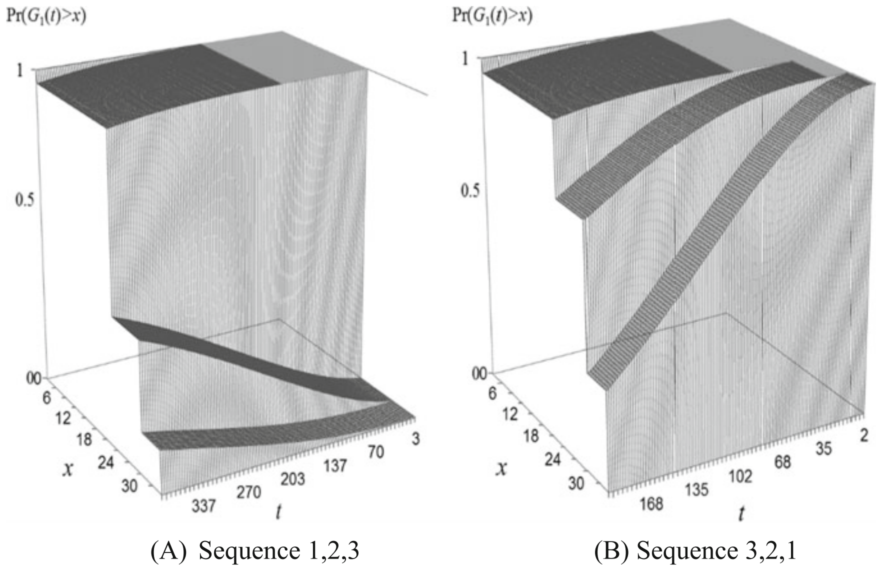


Fig. 2 Cumulative performance distributions $\Pr(G_1(t) \geq x)$ [27]. **a** Sequence 1, 2, 3 **b** Sequence 3, 2, 1

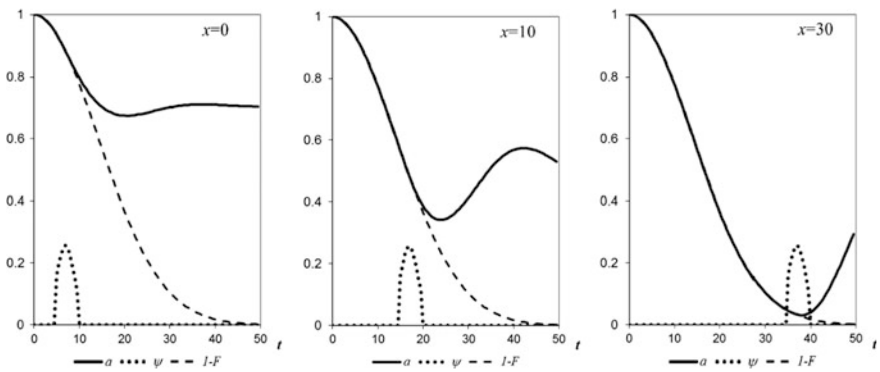


Fig. 3 Performance metrics of the example element for different repair time distributions

distribution with the following parameters: $\gamma^{\min} = x + 5, \gamma^{\max} = x + 10$, $\mu = x + 7$ (mean), $\sigma = 2$ (standard deviation). Figure 3 illustrates the reliability $1 - F(t)$, instantaneous availability $a(t)$ and repair time *pdf* $\psi(t)$ for $x = 0, x = 10$ and $x = 30$ of the element. The perfect repair is assumed, i.e., $\pi = 0$. As the value of x increases (i.e., the repair time increases), the element instantaneous availability reduces significantly.

5 Obtaining Performance DSCTP for Entire MSS5

5.1 UGF (U-Function) Technique

The polynomials in (22) give the u-function modeling the DSCTPs of random performance of *s*-independent component *i* at time *t*.

$$u_i(z,t) = \sum_{n_i=0}^{N_i} p_{i,n_i}(t) z^{g_{i,n_i}}. \tag{22}$$

The composition operator of (23) is used to obtain the *u*-function modeling the DSCTP of the system random performance $V(t)$ at time *t*.

$$\begin{aligned} U(z,t) &= \otimes_{\phi}(u_1(z,t), \ldots, u_I(z,t)) = \otimes_{\phi}\Big(\sum_{n_1=0}^{N_1} p_{1,n_1}(t) z^{g_{1,n_1}}, \ldots, \sum_{n_I=0}^{N_I} p_{I,n_I}(t) z^{g_{I,n_I}}\Big) \\ &= \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \cdots \sum_{n_n=0}^{N_I} \left(\prod_{i=1}^{I} p_{i,n_i}(t) z^{\phi(g_{1,n_1}, \ldots, g_{I,n_I})} \right) \end{aligned} \tag{23}$$

For each time instance *t*, the polynomial *U*(*z*, *t*) models all the possible disjoint combinations of realizations of *s*-independent random variables $G_1(t), \ldots, G_I(t)$ by relating the occurrence probability of each combination to the value of the structure function $\phi(G_1(t), \ldots, G_I(t))$ for this particular combination. This polynomial can eventually take the form of (24), representing performance distribution of the entire system at time *t*.

$$U(z,t) = \sum_{k=0}^{K} w_k(t) z^{v_k}, \tag{24}$$

With the MSS generic model in the form of (1), the system performance metrics (3)–(9) can be obtained through the following 3-step procedure.

1. Apply the *u*-function (22) to represent the *pmf* of random performance distribution of each component *i* at time *t*.
2. Apply the composition operator $\otimes_{\phi}$ (23) to obtain the *u*-function *U*(*z*, *t*) of the entire system random performance distribution *V*(*t*).
3. Evaluate metrics (3)–(9) based on *pmf* (2) modeled by the u-function *U*(*z*, *t*) (24).

Steps 1 and 3 are straightforward. Step 2 often involves complex computations because it can be difficult to derive the system structure functions. According to studies in [17], representing the structure function recursively can be beneficial for computation simplicity and derivation clarity. For a system with complex series parallel structure, its structure function can be represented as a composition of structure functions of the system's s-independent subsystems. Those subsystems contain only components configured in purely series or in purely parallel. During the aggregation process, the RBD method is commonly applied to distinguish recurrent subsystems and replace them with equivalent single components in a graphical representation of system structure, as detailed in Sect. 5.2.

5.2 Generalized RBD Method for Multi-state Series-Parallel System

For obtaining the u-function of a series parallel system, the composition operators is applied recursively to obtain u-functions of intermediate purely series or purely parallel subsystems using the following procedure.

1. Identify any pair of components (i and j) that are connected in parallel or in series in the considered MSS.
2. Obtain the u-function of each pair (i and j) by applying the composition operator $\otimes_{\varphi}$ over the u-functions of these two components:

$$U_{\{i,j\}}(z,t)=u_i(z,t)\otimes_{\phi}u_j(z,t)=\sum_{n_i=0}^{N_i}\sum_{n_j=0}^{N_j}p_{i,n_i}(t)p_{j,n_j}(t)z^{\phi(g_{i,n_i},g_{j,n_j})}, \tag{25}$$

The function ϕ in (25) depends on the interaction nature between the two s-independent components' performances. For example for a production system with throughput being its performance metric, if components i and j operate in parallel, the sum of throughputs of the two components gives the total throughput, as determined in (26).

$$\phi\left(G_i(t),G_j(t)\right)=G_i(t)+G_j(t) \tag{26}$$

If two components process some material consecutively (i.e., forming a series connection), the performance of the component with the minimal performance (i.e., the bottleneck) determines the entire system throughput, as shown in (27).

$$\phi\left(G_i(t),G_j(t)\right)=\min\left(G_i(t),G_j(t)\right) \tag{27}$$

3. Replace the component pair with a single component that has the u-function determined in step 2.
4. If there are more than one component remained in the system, then return to step 1.

The final u-function obtained from the above algorithm models performance distribution of the entire series parallel system.

6 Examples of System Performance Evaluation

6.1 Systems with Warm Standby Components

Figure 4 illustrates an example of MSS with each component consisting of several binary elements configured in a warm standby subsystem. All elements have Weibull time-to-failure distributions with parameters presented in Table 2.

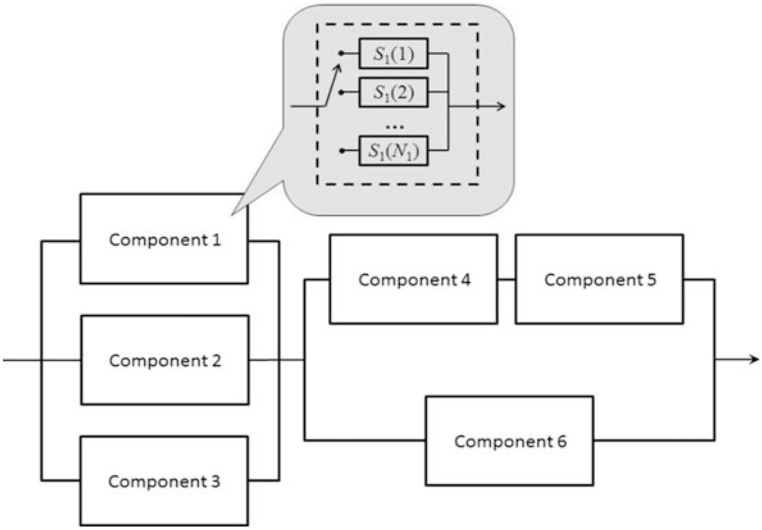


Fig. 4 Example series parallel MSS with warm standby components

Table 2 Parameters of elements composing components

Component i	Element j	$\eta_{i,j}$	$\beta_{i,j}$	$\delta_{i,j}$	$\omega_{i,j}$
1	1	280	1.5	0.2	20
	2	250	1.1	0.4	27
	3	180	2	0.2	32
2	1	300	1	0	18
	2	200	1.4	0	25
3	1	380	2.2	0.5	10
	2	360	1.8	0.7	12
	3	270	1.1	0.5	15
	4	210	1.1	0.6	17
4	1	400	1	0.2	37
	2	400	1	0.2	37
5	1	400	1.4	0.3	30
	2	540	1.2	0.1	40
6	1	380	1.1	0.2	35
	2	340	1.1	0.1	40
	3	280	1.4	0.1	45

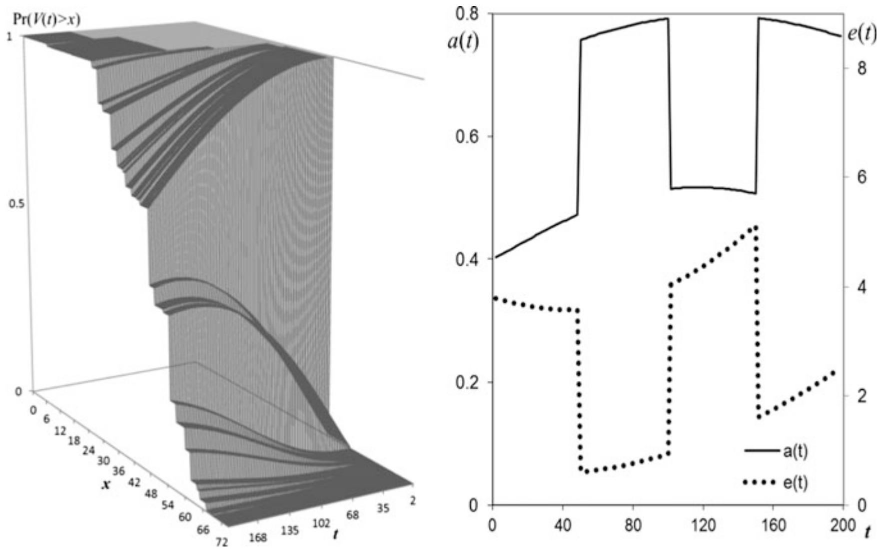


Fig. 5 Performance metrics of the example series parallel system with standby components

A random demand should be supplied by the system, which changes periodically with distribution as follows: $\boldsymbol{d} = \{63, 50, 46, 22\}$, for $t < 50$ and $100 < t < 150$ $\boldsymbol{h}(t) = \{0.2, 0.4, 0.4, 0\}$; for $50 \leq t \leq 100$ and $t \geq 150$, $\boldsymbol{h}(t) = \{0, 0.3, 0.2, 0.5\}$. The components' interaction corresponds to functions (26) and (27). The time of replacement by standby elements is negligible compared to the mission time $T = 200$ (days).

Assume the elements within each component are activated according to their numerical order. For mission time $T = 200$, the expected system availability is obtained as $A(T) = 0.629$ and the expected unsupplied demand is obtained as $E(T) = 539.9$. Figure 5 illustrates the system cumulative performance distribution $\Pr(V(t) \geq x)$, instantaneous availability $a(t)$ and unsupplied demand $e(t)$.

6.2 Systems with Repairable Binary Elements

Assume that the system presented in Fig. 4 consists of repairable binary elements with Weibull time-to-failure distribution parameters and performances presented in Table 3. Table 3 also gives repair efficiency coefficients π_j and parameters related to the truncated normal distributions of elements' repair time. The time horizon of

Table 3 Parameters of repairable elements

Element j	η_j	β_j	$\omega_{,j}$	π_j	$d_j^{\min}$	$d_j^{\max}$	μ_j	σ_j
1	60	2.0	33	0.30	15	20	17	2
2	78	1.1	22	0.70	10	40	25	100
3	90	1.0	19	0.80	28	48	32	5
4	75	1.1	48	0.00	30	40	35	2
5	60	1.0	45	0.20	5	15	10	6
6	80	2.0	33	0.00	10	15	12	100

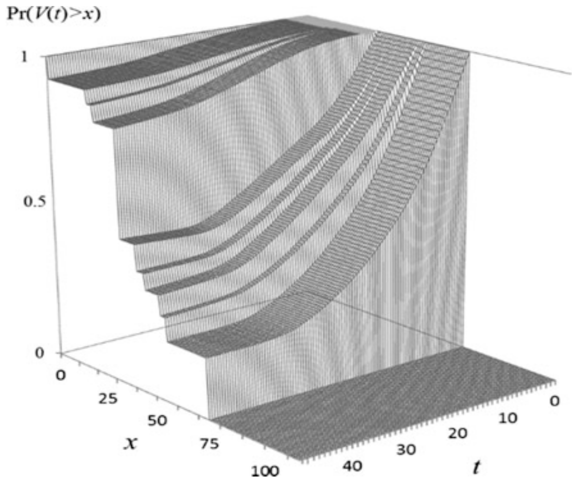


Fig. 6 Cumulative performance distribution Pr($V(t) \geq x$) of example series parallel system with repairable elements

interest is $T = 50$. The random system demand can take four different values $\boldsymbol{d} = \{70, 50, 40, 20\}$ and its distribution is $\boldsymbol{h}(t) = \{0.05, 0.25, 0.45, 0.25\}$, which does not change during T.

The performance metrics obtained for the considered system are $A = 0.6704$, $C = 2565$, $E = 267.3$ and the expected time needed to complete the amount of work $\theta = 1700$ is $\tau = 29.75$.

Figure 6 presents the system cumulative performance distribution $\Pr(V(t) \geq x)$ as a function of time. Figure 7 presents the instantaneous system performance metrics $\alpha(t), c(t)$ and $e(t)$.

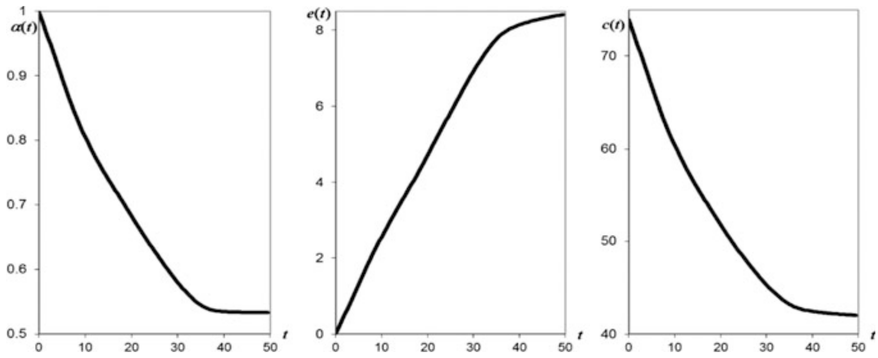


Fig. 7 System instantaneous performance metrics $\alpha(t)$, $c(t)$ and $e(t)$

7 Summary

This chapter demonstrates a methodology that extends the state-of-the-art in system modeling by considering multi-state series parallel systems with components subject to dynamic performance. Each system component can be either a heterogeneous warm standby configuration of binary elements or a repairable binary element. The entire system is considered being available if it can meet a pre-specified random system demand distribution. Iterative algorithms are described for determining dynamic stochastic performances of individual components. A generalized RBD method based on UGFs is implemented for analyzing expected system availability, performance and unsupplied demand over a specific mission time for the entire series parallel MSS.

The presented algorithms allow fast determination of system dynamic performance metrics. Based on these algorithms different optimization problems can be solved. For example, as shown in Sect. 4.1, standby element activation sequence can have significant impacts on component and thus system performance metrics. Therefore solving the following two problems can considerably improve system performance. The optimal operation problem finds the element activation sequence of each component maximizing system availability or minimizing unsupplied system demand. The optimal design problem finds both component structures and element activation sequences minimizing the total cost consisting of design and operation expenses. In addition, increased elements loading can on one hand improve the system performance; on the other hand, it can cause failures that are more frequent and, thus reduces the system availability. Hence, the element loading can be optimized to achieve a proper balance among different system performance metrics. Examples of solving the optimization problems based on the suggested methodology can be found in [26, 27].

References

1. Amari S, Xin L, Shrestha A, Akers J, Trivedi K (2010) Performability analysis of multi-state computing systems using multi-valued decision diagrams. IEEE Trans Comput 59(10):1419–1433
2. Peng R, Zhai Q, Xing L, Yang J (2014) Reliability of demand-based phased-mission systems subject to fault level coverage. Reliab Eng Syst Saf 121:18–25
3. Lisnianski A, Levitin G (2003) Multi-state system reliability. In: Assessment, optimization and applications. World Scientific, NJ
4. Frenkel I, Lisnianski A, Khvatskin L (2014) On the *Lz*-transform application for availability assessment of an aging multi-state water cooling system for medical equipment. In: Frenkel I, Lisnianski A, Karagrigoriou A, Kleiner A (eds) Applied reliability and risk analysis: probabilistic models and statistical inference. Wiley, New York, pp 60–77
5. Lisnianski A, Ben Haim H (2013) Short-term reliability evaluation for power stations by using *Lz*-transform. J Mod Pow Syst Clean Energy 1(2):110–117
6. Lisnianski A, Ding Y (2016) Using inverse *Lz*-transform for obtaining com-pact stochastic model of complex power station for short-term risk evaluation. Reliab Eng Syst Saf 145:19–27
7. Veeraraghavan M, Trivedi KS (1994) A combinatorial algorithm for performance and reliability analysis using multi-state models. IEEE Trans Comput 43:229–234
8. Zang X, Wang D, Sun H, Trivedi KS (2003) A BDD-based algorithm for analysis of multistate systems with multistate components. IEEE Trans Comput 52:1608–1618
9. Li Y, Peng R (2014) Availability modeling and optimization of dynamic multi-state series-parallel systems with random reconfiguration. Reliab Eng Syst Saf 127:47–57
10. Mo Y, Xing L, Amari S, Dugan J (2015) Efficient analysis of multi-state k-out-of-n systems. Reliab Eng Syst Saf 133:95–105
11. Ramirez-Marquez J, Coit D, Tortorella M (2006) A generalized multi-state-based path vector approach to multistate two-terminal reliability. IIE Trans 38(6):477–488
12. Zio E, Podofillini L (2003) Monte-Carlo simulation analysis of the effects on different system performance levels on the importance on multi-state components. Reliab Eng Syst Saf 82 (1):63–73
13. Levitin G, Amari S, Xing L (2013) Algorithm for reliability evaluation of non-repairable phased-mission systems consisting of gradually deteriorating multi-state elements. IEEE Trans Syst Man Cybern Syst 43(1):63–73
14. Frenkel I, Lisnianski A, Khvatskin L (2012) Availability assessment for aging refrigeration system by using *Lz*-transform. J Reliab Stat Stud 5(2):33–43
15. Lisnianski A (2012) *Lz*-transform for a discrete-state continuous-time Markov process and its applications to multi-state system reliability. In: Lisnianski A, Frenkel I (eds) Recent advances in system reliability. Signatures, multi-state systems and statistical inference. Springer, London, pp 79–96
16. Lisnianski A, Frenkel I, Khvatskin L (2015) On Birnbaum importance assessment for aging multi-state system under minimal repair by using *Lz*-transform method. Reliab Eng Syst Saf 142:258–266
17. Levitin G (2005) Universal generating function in reliability analysis and optimization. Springer, London
18. Xing L, Dai Y (2009) A new decision diagram based method for efficient analysis on multi-state systems. IEEE Trans Dependable Sec Comput 6(3):161–174
19. Xing L, Levitin G (2011) Combinatorial algorithm for reliability analysis of multi-state systems with propagated failures and failure isolation effect. IEEE Trans Syst Man Cybern Part A Syst Hum 41(6):1156–1165
20. Dao C, Zuo M, Pandey M (2014) Selective maintenance for multi-state series-parallel systems under economic dependence. Reliab Eng Syst Saf 121:240–249

21. Levitin G, Lisnianski A (1999) Joint redundancy and maintenance optimization for multistate series-parallel systems. Reliab Eng Syst Saf 64(1):33–42
22. Tian Z, Levitin G, Zuo M (2009) A joint reliability–redundancy optimization approach for multi-state series-parallel systems. Reliab Eng Syst Saf 94(10):1568–1576
23. Xiao H, Peng R (2014) Optimal allocation and maintenance of multi-state elements in series-parallel systems with common bus performance sharing. Comput Ind Eng 72:143–151
24. Levitin G, Xing L, Dai Y (2013) Optimal sequencing of warm standby elements. Comput Ind Eng 65(4):570–576
25. Lindqvist BH (2006) On the statistical modeling and analysis of repairable systems. Stat Sci 21(4):532–551
26. Levitin G, Xing L, Dai Y (2017) Optimal loading of series parallel systems with arbitrary element time-to-failure and time-to-repair distributions. Reliab Eng Syst Saf 164:34–44
27. Levitin G, Xing L, Dai Y (2016) Optimizing dynamic performance of multi-state systems with heterogeneous 1-out-of-N warm standby components". IEEE Trans Syst Man Cybern Syst (to appear). doi:10.1109/TSMC.2016.2633808

Optimal Imperfect Maintenance in a Multi-state System

Stephanie Dietrich and Waltraud Kahle

Abstract In this research, we are concerned with the modeling of optimal maintenance actions in multi-state systems. Most of the imperfect maintenance models that have been investigated in literature use either imperfect preventive maintenance actions or imperfect corrective maintenance actions. In this paper we consider a model with both imperfect preventive and imperfect corrective maintenance actions. A sequential failure limit preventive maintenance (PM) policy with infinite planning horizon and with imperfect preventive and imperfect corrective maintenance actions is used to formulate a cost optimization problem. Different cost functions for PM actions, as well as several discrete lifetime distributions are introduced. The solution of the cost optimization problem is illustrated by an example.

Keywords Multistate systems · Optimal maintenance · Discrete lifetimes · Modified Weibull distribution

1 Introduction

This paper is concerned with the statistical modeling and optimization of an imperfect maintenance model for a repairable MSS. Multi-state systems are an important area in modern reliability theory. They provide a flexible tool for modeling engineering systems in real life. This model was first introduced in [4, 7]. For a historical overview of MSS and an overview of ideas for MSS reliability theory, see for example [13]. A recent contribution on the subject is [12]. As in [10] we consider a MSS with n states and assume that in time 1 the system is in state one, in time 2 the system is in state two and so on. Using this time scale, one can model imperfect maintenance actions also for systems with discrete lifetime distribution.

S. Dietrich · W. Kahle (✉)
Otto-von-Guericke University Magdeburg, Magdeburg, Germany
e-mail: waltraud.kahle@ovgu.de

S. Dietrich
e-mail: Stephanie_Dietrich@gmx.de

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_9

Most models in reliability theory use continuous lifetime distributions. This is appropriate because in reality most of the lifetimes are continuous. But there are also several situations where discrete failure data arise. This is the case if the life length of a system is measured in cycles and the number of cycles successfully completed prior to failure are observed. The same holds if we have a multi-state system and the number of states prior to failure is observed. A recent contribution on complex systems subject to shock processes is [8].

When a failure occurs, a repairable system can be restored to an operating condition by some repair process. Therefore, it is not necessary to replace the whole system, and the failure intensity of the system depends on the history of repairs. In general, there are two kinds of maintenance actions. Preventive or planned maintenance actions and unplanned corrective maintenance actions. Corrective maintenance (CM), also called repair, is carried out after a failure of the system. The aim of CM actions is to retain the system in or restore it to an acceptable operating condition [14]. Preventive maintenance (PM) occurs when the system is operating. The aim of PM actions is to reduce the failure intensity of the system.

The two most common assumptions on the influence of maintenance actions is minimal repair or as bad as old and perfect repair or as good as new. A maintenance action is called to be minimal if the failure intensity of the system after that action is the same it had when it failed. Perfect repair means the failure intensity of the system after repair is that of a new system [3]. In reality, the state of the system after maintenance will often not be as good as new and not as bad as old, but something in between. In this case, the maintenance action is called to be imperfect. In this paper we assume that maintenance actions impacted the failure intensity of the system in such a way that they adjust the virtual age of the system. As in Kijima [11], we proposed that the state of the system just after a maintenance action can be described by its virtual age, which is smaller or equal than the real age of the system.

In [5] we have considered a system with two failure types, minor and major ones. Minor ones can be removed by minimal repair and the major ones can only be removed by replacement. The maintenance model of the present paper includes both imperfect preventive and imperfect corrective maintenance actions where the costs of maintenance actions depend on the degree of maintenance. Here a sequential failure limit PM policy (also called *age replacement* in case of perfect PM) with infinite planning horizon is used to formulate a cost optimization problem.

The paper is structured as follows. Section 2 contains essential assumptions and definitions that are needed to formulate the cost optimization problem in Sect. 3. In Sect. 4 we introduce some discrete lifetime distributions. In multi-state systems, these distributions define the probability of a failure in different states. Finally, in Sect. 5 different possible cost functions for PM and CM are introduced. In an example, the optimal maintenance strategy is computed the introduced discrete lifetime distributions and for one of the cost functions.

2 Modeling the System and Maintenance Policy

According to Kahle [10], we consider a repairable multi-state system (MSS). This system has the following properties.

1. Initially a new repairable MSS is installed. The MSS has n states in which the system can fail. A time scale is introduced so that the system at time 1 is in state one, at time 2 in state two and so on.
2. The system has only one failure type which can be removed through imperfect repair actions.
3. The repair times are negligible small.

The maintenance strategy described here is designed for an infinite time horizon. The following assumptions are made.

1. All failures that occurred after installation during the time interval $(0, v]$ are removed through minimal repair.
2. If there is a failure during the time interval $(v, v + \tau)$ a CM action is carried out. The first cycle is finished by this CM. Otherwise a PM action at time $v + \tau$ will be carried out and the first cycle is finished by a PM. Both maintenance actions reset the virtual age (state) of the system to v.
3. At the begin of a cycle the system is in state v which is the virtual age of the system. If there is no failure during the pre-defined time interval of length $\tau > 0$ after a maintenance action, a PM will be carried out. For τ it holds $\tau \in \{1, \dots, n - v\}$.
4. If a failure occurs during the time interval of the length $\tau > 0$ after a maintenance action, a CM is carried out.
5. The virtual age of the system after both PM actions and CM actions is always $v \geq 0$ and $v \in \{0, \dots, n - 1\}$. Since PM actions can be planned, they are assumed to be more cost effective than unplanned CM actions.
6. Suppose $c_1, c_2, \dots$ are the realizations of the general maintenance times. In terms of Kijima type II model the degree of the kth repair is

$$\xi_k(v, c_k, c_{k-1}) = \frac{v}{v + c_k - c_{k-1}}, \quad \forall k \geq 1. \tag{1}$$

This maintenance policy is a sequential failure limit policy (see [6, p. 765]) because an alternative formulation of Assumption 2 might be: A PM is performed when the failure intensity reaches the predetermined level $\lambda(v + \tau)$.

3 Cost Optimization Problem

Consider a technical system that is maintained with maintenance policy described in Sect. 2. The aim of this section is to formulate a cost optimization problem. The optimization criterion are the average maintenance costs per unit time. For this

purpose, the expected maintenance costs per cycle are set in relation to the mean cycle length. Here the cycle length is the time between two maintenance actions and for reasons of simplification, the time between the startup of the system and the age of v is excluded from the modeling of the cost optimization problem.
Suppose $N = (N_t)_{t \geq 0}$ is the failure counting process, i.e. N_t is the random number of failures of a repairable system with PM in the interval $[0, t]$.

The next two lemmata are obviously.

Lemma 1 (Intensity Function of $N = (N_t)_{t \geq 0}$) *Suppose $c_1, c_2, \ldots$ are realizations of the general maintenance times. The intensity function of the counting process $N = (N_t)_{t \geq 0}$ is then*

$$\lambda(t) = \begin{cases} 0, & t < 0 \\ h^{T_1}(t), & t \in [0, c_1) \\ h^{T_1}(v + t - c_k), & t \in [c_k, c_{k+1}),\ k \geq 1, \end{cases} \tag{2}$$

where $h^{T_1}(\cdot)$ is the hazard function of the time to first failure of a new system.

For the computation of the expected maintenance costs per cycle it is necessary to compute the probability that a failure occurs within τ time units after a maintenance action.

Lemma 2 (Distribution function of T^v) *Let T_1 be the discrete random time of the first failure of a repairable system without maintenance. Suppose T^v is the remaining discrete lifetime of the system after a maintenance action that reduces the virtual age of the system to v. Then T^v is a truncated discrete random variable with the following cumulative distribution function*

$$F^{T^v}(t) = P(T^v \leq t) = \frac{F^{T_1}(v + t) - F^{T_1}(v)}{1 - F^{T_1}(v)}, \ \forall t = 0, 1, 2, \ldots. \tag{3}$$

Suppose the random cycle length $L_{v,\tau}$ is the random time between two maintenance actions. Therefore, it is either the time between two PM actions (if there is no failure within the interval of length τ) or the time between PM and CM actions (if a failure occurs). It holds

$$L_{v,\tau} = \min\{T^v, \tau\} \begin{cases} < \tau \text{ with } P(T^v < \tau) = P(T_1 < v + \tau | T_1 > v) \\ = \tau \text{ with } P(T^v \geq \tau) = 1 - P(T_1 < v + \tau | T_1 > v) \end{cases} \tag{4}$$

for $\tau \in \{1, \ldots, n - v\}$. For the cumulative distribution function of $L_{v,\tau}$ it holds

$$F^{L_{v,\tau}}(t) = P(L_{v,\tau} \leq t) = \begin{cases} 0, & \text{if } t < 0 \\ F^{T_1}(v + t | T_1 > v), & \text{if } t \in \{0, \ldots, \tau - 1\} \\ 1, & \text{if } t \geq \tau \end{cases}. \tag{5}$$

Theorem 1 (Mean Cycle Length) *For the expected cycle length it holds*

$$E(L_{v,\tau}) = \sum_{j=0}^{\tau-1} j \cdot P(T_1 = v + j | T_1 > v) + \tau \cdot P(T_1 \geq v + \tau | T_1 > v) \,. \tag{6}$$

Proof The random cycle length is a positive discrete random variable with the following probability mass function

$$P(L_{v,\tau} = t) = \begin{cases} 0, & \text{if } t < 0 \\ P(T_1 = v + t | T_1 > v), & \text{if } t \in \{0, \dots, \tau - 1\} \\ P(T_1 \geq v + t | T_1 > v), & \text{if } t = \tau \\ 0, & \text{if } t > \tau \end{cases} \,. \tag{7}$$

Using this, the mean cycle length is computed as follows

$$\begin{aligned} E(L_{v,\tau}) &= \sum_{j=0}^{\infty} j \cdot P(L_{v,\tau} = j) = \sum_{j=0}^{\tau-1} j \cdot P(L_{v,\tau} = j) + \tau \cdot P(L_{v,\tau} = \tau) \\ &= \sum_{j=0}^{\tau-1} j \cdot P(T_1 = v + j | T_1 > v) + \tau \cdot P(T_1 \geq v + \tau | T_1 > v). \end{aligned} \tag{8}$$

□

The mean cycle length in the discrete case is bounded from below by one and for $v = 0$ the mean cycle length is bounded from above by $\min\{\tau, E(T_1)\}$. Note that since $P(T_1 = 0) = 0$, the expected time to first failure of a new system is always greater or equal one, i.e. $E(T_1) \geq 1$.

Problem 1 (*Cost Optimization Problem)* Let c_{CM} denotes the costs of a CM action and c_{PM} the costs of a PM action. The average maintenance costs per unit time are

$$C(v, \tau) = \frac{c_{CM} P(T^v < \tau) + c_{PM} P(T^v \geq \tau)}{E(L_{v,\tau})} \,. \tag{9}$$

The optimization problem then has the following form

$$\min_{v \in \{0,\dots,n-1\},\, \tau \in \{1,\dots,n-v\}} C(v, \tau). \tag{10}$$

Note that if $\tau = 1$, there will be no CM actions, i.e. $P(T^v < 1) = 0$, and $E(L_{v,1}) = 1$. Thus, for the average maintenance costs per unit time it holds $C(v, 1) = c_{PM}$.

In Sect. 5 an example plot of this function is given. It can be seen that it has (at least for this example) a unique minimum with respect to both τ and v.

4 Discrete Lifetime Distributions

There are different possible discrete lifetime distributions, e.g. in [10] are studied some discrete lifetime distributions like the discrete uniform distribution, the Poisson distribution and the truncated discrete Weibull distribution. In this paper we consider the discrete modified Weibull distribution (DMWD), that was introduced in [9], and the discrete version of the reduced modified Weibull distribution (RMWD), that was introduced in [2].

4.1 The Discrete Modified Weibull Distribution

The DMWD based on the modified Weibull distribution (MWD) that was introduced in [15]. This distribution generalizes some most commonly used distributions in survival analysis such as exponential, Rayleigh, linear failure rate and Weibull distribution.

In the following, the notation MWD(α, β, γ) is used to denote the modified Weibull distribution with the scale parameter α and the two shape parameters β and γ. Let X be modified Weibull distributed MWD(α, β, γ). Then the probability density function of X is

$$f^X(x; \alpha, \beta, \gamma) = (\alpha + \beta\gamma x^{\gamma-1}) \exp(-\alpha x - \beta x^\gamma), \tag{11}$$

where $x > 0$, $\gamma > 0$ and α, $\beta \geq 0$ such that $\alpha + \beta > 0$.

Let T be the random discrete lifetime of an operating unit. For the discrete version of the MWD we put the probability mass of the interval $(t-1, t]$ into the point t (see [9]), that is for $t = 1, 2, \dots$

$$\begin{aligned} P(T = t) &= \int_{t-1}^{t} f^X(s) ds \\ &= \exp(-\alpha(t-1) - \beta(t-1)^\gamma) - \exp(-\alpha t - \beta t^\gamma) \ . \end{aligned} \tag{12}$$

The corresponding reliability (survival) function is given by

$$R^T(t) = P(T > t) = \sum_{j=t+1}^{\infty} P(T = j) = \exp(-(\alpha t + \beta t^\gamma)), \quad t = 1, 2, \dots, \tag{13}$$

and $R^T(0) = 1$. Finally, the failure rate is given by

$$\begin{aligned} h^T(t) &= \frac{P(T = t)}{P(T \geq t)} = \frac{R^T(t-1) - R^T(t)}{R^T(t-1)} \\ &= 1 - \exp(\alpha(t-1) + \beta(t-1)^\gamma - (\alpha t + \beta t^\gamma)), \quad t = 1, 2, \dots . \end{aligned} \tag{14}$$

This DMWD(α, β, γ) includes other important lifetime distributions. If $\gamma = 2$ the DMWD(α, β, γ) becomes the discrete linear failure rate distribution (DLFRD) with parameters α and β. In case of $\alpha = 0$, we obtain the discrete Weibull distribution (DWD) with parameters β and γ. By setting $\alpha = 0$ and $\gamma = 2$, we get the discrete Rayleigh distribution (DRD) with parameter β. Finally, if $\beta = 0$, we obtain the discrete exponential distribution with parameter α, that is, a geometric distribution with success probability $1 - \exp(-\alpha)$.

4.2 *The Discrete Reduced Modified Weibull Distribution*

The discrete version of the RMWD was introduced in [2] and based on a three-parameter modified Weibull distribution, which was developed in [1]. In the following the notation RMWD(α, β, γ) is used to denote the reduced modified Weibull distribution with scale parameters α and β and acceleration parameter γ. Let X be reduced modified Weibull distributed RMWD(α, β, γ). Then the probability density function of X is

$$f^X(x; \alpha, \beta, \gamma) = \frac{1}{2\sqrt{x}} (\alpha + \beta(1 + 2\gamma x) \exp(\gamma x)) \times \exp\left(-\alpha\sqrt{x} - \beta\sqrt{x}\exp(\gamma x)\right), \tag{15}$$

for $x > 0$, α, $\beta > 0$ and $\gamma > 0$. Now set $q = \exp(-\alpha)$, $b = \beta/\alpha$ and $c = \exp(\gamma)$.

Let T be the random discrete lifetime of an operating unit. As before, for the discrete version of the RMWD(q, b, c) we put the probability mass of the interval $(t - 1, t]$ into the point t, i.e.

$$P(T = t) = \int_{t-1}^{t} f^X(s)ds = q^{\sqrt{t-1}(1+bc^{t-1})} - q^{\sqrt{t}(1+bc^t)}, \quad t = 1, 2, \ldots. \tag{16}$$

The reliability (survival) function of the discrete reduced modified Weibull distribution DRMWD(q, b, c) is given by

$$R^T(t) = P(T > t) = q^{\sqrt{t}(1+bc^t)}, \quad t = 1, 2, \ldots, \tag{17}$$

and $R^T(0) = 1$. The failure rate is given by

$$h^T(t) = \frac{P(T = t)}{P(T \geq t)} = 1 - q^{\sqrt{t}(1+bc^t) - \sqrt{t-1}(1+bc^{t-1})}, \quad t = 1, 2, \ldots. \tag{18}$$

As shown in [2], the failure rate of the DRMWD(q, b, c) can be increasing or has a bathtub shape.

Fig. 1 Probability mass functions

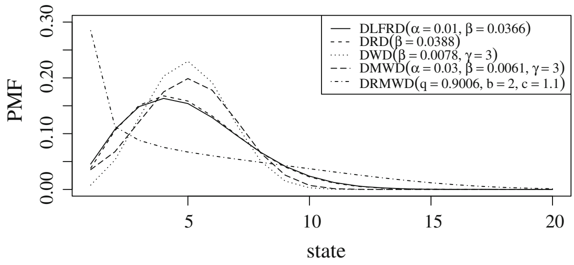


Fig. 2 Discrete failure rates

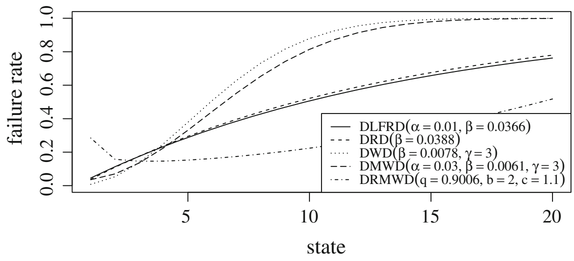


Fig. 3 Cumulative hazard functions

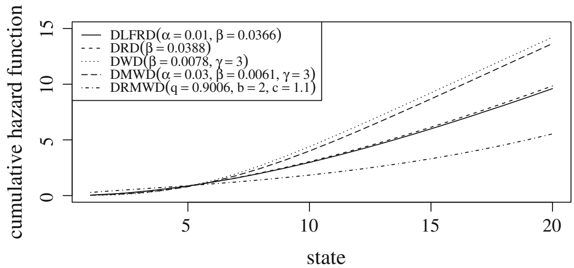


Figure 1 shows the different probability mass functions (PMF). It can be seen that the PMF of the DLFRD and the DRD have a similar shape. This is true also for the failure rate and the cumulative hazard function as seen in Figs. 2 and 3. As can be seen in Fig. 2, the chosen parameters for the DRMWD leads to a bathtub shaped failure rate.

Note that in Figs. 1, 2 and 3 the points are connected with lines for better visibility. Further, the parameters for all distributions are chosen so that the expectation is equal to 5 and the number of states in the MSS is $n = 20$.

Usually truncated distributions, i.e. $P(T = t|T \leq n)$, are used to model the lifetime of MSS. For the chosen distribution parameters of all previously defined lifetime distributions, the probability that a failure occurs at a point in time $t > 20$ is close to zero. This means there is no significant difference between the lifetime distribution and the truncated version of it. Hence, it is appropriate here to use the untruncated lifetime distribution instead of the truncated version.

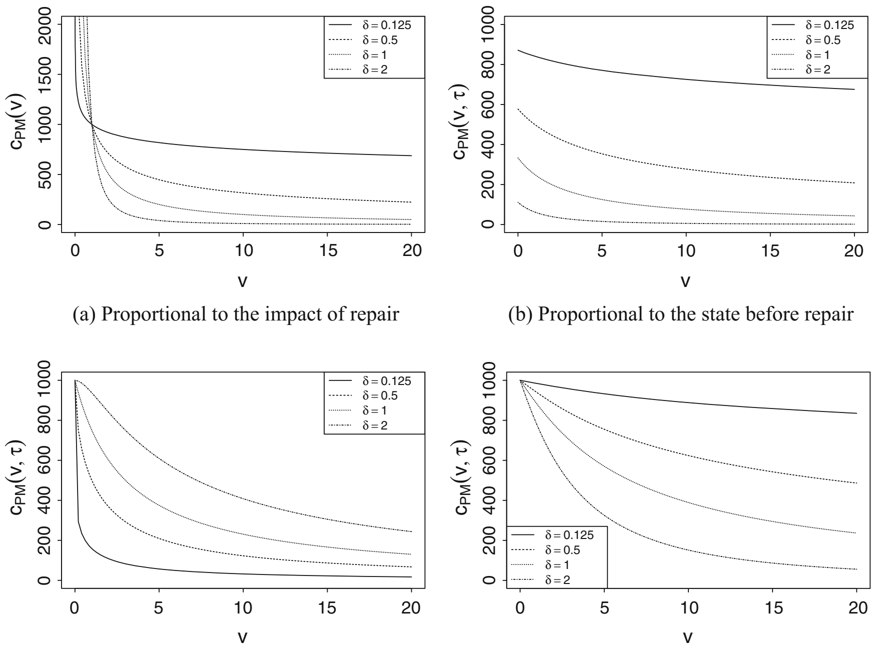
5 Example for Cost Optimal Maintenance

The costs for CM and PM actions in optimization problem 1 are yet unspecified. In this section four different special cost functions (plotted in Fig. 4) for CM and PM actions are considered and cost optimal parameter v and τ are computed for different discrete lifetime distributions using R and complete enumeration.

(a) **Costs Proportional to the Impact of Repair**
We assume, that the costs for maintenance actions depends only on the virtual age after repair, i.e. the cost functions for PM and CM are

$$c_{PM}(v) = c_I \left(\frac{1}{v}\right)^{\delta}, \qquad c_{CM}(v) = c_F + c_I \left(\frac{1}{v}\right)^{\delta}, \tag{19}$$

where $v > 0$, $\delta > 0$, $c_I > 0$ is a constant cost value and $c_F > 0$ is the fixed amount, by which the costs for CM are higher than for PM. Note that in case of costs proportional to the impact of repair the extreme case of perfect repair, i.e. $v = 0$, have to be excluded from optimization problem 1.
Cost function (19) has the following properties:



(a) Proportional to the impact of repair
(b) Proportional to the state before repair
(c) Proportional to the degree of repair – 1
(d) Proportional to the degree of repair – 2

Fig. 4 Cost functions for PM depending on v with $c_R = c_I = c_S = 1000$, $c_M = 500$ and $\tau = 3$

- The costs of PM are only bounded below by zero, i.e. $c_{PM}(v) > 0$.
- Figure 4a shows the better the PM actions, i.e. the smaller v, the higher are the costs for PM. For a perfect repair the costs tend to infinity, i.e. the limit of $c_{PM}(v)$ as v approaches zero is infinity, on condition that δ does not tend towards zero.
- The worse the PM actions, i.e. the higher v, the smaller are the costs of PM. The limit of $c_{PM}(v)$ as v approaches infinity is zero, on condition that δ does not tend towards zero.
- If δ tends to zero the costs of PM converge to c_I for all values of v. Therefore, the costs of relative good PM actions with $v < 1$ decrease and the costs of less good PM actions with $v > 1$ increase.
- The higher δ, the faster the costs of PM converge to infinity as v approaches zero and the faster the costs of PM converge to zero as v approaches infinity.
- If $v = 1$ the costs of PM are always c_I independent of δ.

(b) **Costs Proportional to the State Before Repair**
Assume the cost function c_{PM} depends on the state just before PM. It holds

$$c_{PM}(v, \tau) = c_S \left(\frac{1}{v + \tau} \right)^{\delta}, \tag{20}$$

where $\delta > 0$ and $c_S > 0$ is a constant cost value. The following cost function is used for CM actions

$$c_{CM}(v, \tau) = \sum_{t=1}^{\tau-1} \left(c_F + c_S \left(\frac{1}{v + t} \right)^{\delta} \right) P(T^v = t | T^v < \tau), \tag{21}$$

where $c_F > 0$ is the fixed amount, by which the costs for CM are higher than for PM.

- The costs of PM are only bounded below by zero, i.e. $c_{PM}(v, \tau) > 0$.
- Figure 4b shows the better the PM actions and the shorter the distance between PM actions, the higher are the costs for PM. In case of perfect repair, i.e. $v = 0$, the costs of PM are $c_{PM} = c_S(1/\tau)^{\delta}$.
- The worse the PM actions and the longer the distance between PM actions, i.e. the higher v and τ, the lower are the costs of PM. The limit of $c_{PM}(v, \tau)$ as v or τ approaches infinity is zero, on condition that δ does not tends towards zero.
- If δ tends to infinity and v or τ tend to infinity the costs of PM converge to zero. The higher δ the faster the costs of PM tend towards zero.
- If δ tends to zero the costs of PM converge to c_S for all possible values of v and τ.
- If δ is increasing the costs of PM tend to infinity if $v + \tau$ is close to zero.

(c) **Costs Proportional to the Degree of Repair**
It is assumed that a PM action reduces the virtual age of a repairable system to $v \geq 0$. If the PM action is done $\tau \geq 0$ time units after the last maintenance action that reduced the virtual age of the system to v, the degree of repair in a Kijima type II manner is

$$\xi(v, \tau) = \frac{v}{v + \tau}. \tag{22}$$

If the costs of PM depend on the degree of repair, a possible cost function could be

$$c_{PM}(v, \tau) = c_R \left(1 - \left(\frac{v}{v + t}\right)^{\delta}\right), \tag{23}$$

with $\delta > 0$ and replacement costs $c_R > 0$. The cost function used for CM actions is

$$c_{CM}(v, \tau) = \sum_{t=1}^{\tau-1} \left(c_F + c_R \left(1 - \left(\frac{v}{v + t}\right)^{\delta}\right)\right) P(T^v = t | T^v < \tau), \tag{24}$$

where $c_F > 0$ is the fixed amount by which the costs of CM are higher than for PM. Figure 4c shows the shape of cost function (24) for different parameter constellations. This cost function has the following properties:

- The costs of PM actions are bounded below and above. They are greater than zero and smaller than the costs of a replacement, i.e. $0 \leq c_{PM}(v, \tau) \leq c_R$.
- The better the PM actions, i.e. the smaller v, the higher are the costs for PM. In case of perfect repair, i.e. $\xi(v, \tau) = 0$, implying that $v = 0$, the costs of PM are $c_{PM}(v, \tau) = c_R$. The same holds if v tends to zero and δ does not at the same time tends to zero.
- The worse the PM actions and the lower the distance between PM actions, i.e. the higher v and the smaller τ, the lower are the costs for PM. In case of minimal repair, i.e. $\xi(v, \tau) = 1$, involving that v tends to infinity or $\tau = 0$, there are no costs of PM, i.e. $c_{PM}(v, \tau) = 0$.
- If v and τ tend to infinity and δ does not tend to zero, the costs of PM tend to the cost of a replacement, i.e. $c_{PM}(v, \tau)$ tends to c_R.
- The lower δ the lower are the costs of PM. If $v \neq 0$ and δ tends to zero, the costs of PM tend to zero too.

(d) **Costs Proportional to the Degree of Repair-2**
Another way to model the costs of PM actions if they depend on the degree of repair is

$$c_{PM}(v, \tau) = c_R \left(1 - \left(\frac{v}{v + \tau}\right) \exp\left(\frac{v}{v + \tau} - 1\right)\right)^{\delta}, \tag{25}$$

where $\delta > 0$ and $c_R > 0$ are the costs of a replacement. For CM actions the following cost function is used

$$c_{CM}(v,\tau) = \sum_{t=1}^{\tau-1} \left(c_F + c_R \left(1 - \left(\frac{v}{v+t} \right) \exp \left(\frac{v}{v+t} - 1 \right) \right)^{\delta} \right) P(T^v = t | T^v < \tau). \tag{26}$$

where $v, \tau \geq 0$, $v + \tau \neq 0$, $\delta > 0$ and $c_R > 0$ are the costs of a replacement. This cost function was introduced e.g. in Gasmi and Mannai [9]. Figure 4d shows the shape of cost function (25) for different parameter constellations. This cost function has the following properties:

- The costs of PM actions are bounded below and above. They are greater than zero and smaller than the costs of a replacement, i.e. $0 \leq c_{PM}(v,\tau) \leq c_R$.
- The better the PM actions, i.e. the smaller v, the higher are the costs for PM. In case of perfect repair, i.e. $\xi(v,\tau) = 0$, implying that $v = 0$, the costs of PM are $c_{PM}(v,\tau) = c_R$. The same holds if v tends to zero and δ does not at the same time tends to zero.
- The worse the PM actions and the lower the distance between PM actions, i.e. the higher v and the smaller τ, the lower are the costs for PM. In case of minimal repair, i.e. $\xi(v,\tau) = 1$, involving that v tends to infinity or $\tau = 0$, there are no costs of PM, i.e. $c_{PM}(v,\tau) = 0$. The same holds if τ tends to zero and v and δ does not at the same time tend to zero.
- The lower δ, the higher are the costs of PM actions. If $v \neq 0$ and δ tends to zero, the costs of PM tend to the costs of a replacement.
- If v and τ tend to infinity and δ does not tend to zero, the costs of PM tend to the cost of a replacement, i.e. $c_{PM}(v,\tau)$ tends to c_R.

As illustrated in Fig. 4c, d, the difference between cost function (24) and (25) is that if $\delta \leq 1$ the costs of PM actions are now higher than before and if $\delta > 1$ the costs of PM actions are lower than before. In contrast to cost function (24), the costs of PM decrease with rising δ.

As a concrete example we consider case (b). The cost optimal maintenance strategies for different cost ratios c_F/c_I and different δ are given in Table 1 and lead to the following conclusions:

1. Only for $\delta = 0.125$ it is cost optimal to do very good maintenance actions with $v = 1$. For higher δ these maintenance actions are too expensive compared to worse maintenance actions. Therefore, with rising δ it is cost optimal to have higher values of v.
2. The higher δ the greater is the cost difference between good and less good maintenance actions and the faster the costs of PM tend to zero and the costs of CM tend to c_F. Therefore, for high values of δ it is cost optimal to do nonstop bad PM actions.

Table 1 Optimal values in case of costs proportional to the state before repair

	DLFRD $\alpha = 0.01$ $\beta = 0.02944$	DRD $\beta = 0.03142$	DWD $\beta = 0.0057$ $\gamma = 3$	DMWD $\alpha = 0.03$ $\beta = 0.004335$ $\gamma = 3$	DRMWD $\alpha = 0.1$ $\beta = 0.1746$ $\gamma = 0.1$
$c_F/c_S = 0.05$					
$\delta = 0.125$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 1$
	$\tau = 20$	$\tau = 20$	$\tau = 10$	$\tau = 11$	$\tau = 19$
$\delta = 0.5$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 3$
	$\tau = 20$	$\tau = 19$	$\tau = 9$	$\tau = 10$	$\tau = 17$
$\delta = 1$	$v = 19$	$v = 19$	$v = 0$	$v = 19$	$v = 5$
	$\tau = 1$	$\tau = 1$	$\tau = 8$	$\tau = 1$	$\tau = 13$
$\delta = 2$	$v = 19$	$v = 19$	$v = 19$	$v = 19$	$v = 19$
	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$
$c_F/c_R = 0.1$					
$\delta = 0.125$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 1$
	$\tau = 18$	$\tau = 17$	$\tau = 9$	$\tau = 10$	$\tau = 19$
$\delta = 0.5$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 3$
	$\tau = 14$	$\tau = 13$	$\tau = 8$	$\tau = 9$	$\tau = 17$
$\delta = 1$	$v = 19$	$v = 19$	$v = 19$	$v = 19$	$v = 4$
	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 11$
$\delta = 2$	$v = 19$	$v = 19$	$v = 19$	$v = 19$	$v = 19$
	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$
$c_F/c_R = 0.2$					
$\delta = 0.125$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 1$
	$\tau = 12$	$\tau = 12$	$\tau = 7$	$\tau = 8$	$\tau = 18$
$\delta = 0.5$	$v = 0$	$v = 0$	$v = 0$	$v = 0$	$v = 3$
	$\tau = 10$	$\tau = 9$	$\tau = 7$	$\tau = 7$	$\tau = 13$
$\delta = 1$	$v = 19$	$v = 19$	$v = 19$	$v = 19$	$v = 19$
	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$
$\delta = 2$	$v = 19$	$v = 19$	$v = 19$	$v = 19$	$v = 19$
	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$	$\tau = 1$

3. For lower values of δ the maintenance costs are relatively high and the cost difference between good and less good maintenance actions is relatively small. Therefore, the higher the amount, by which the costs of CM are higher than for PM, i.e. the higher the ratio c_F/c_I, the better are the maintenance actions, i.e. v is decreasing.
4. If the cost optimal values of v for the other distributions are at a low level, the DRMWD has higher cost optimal values of v. The reason for this lies in the high failure rate at lower states for the DRMWD.

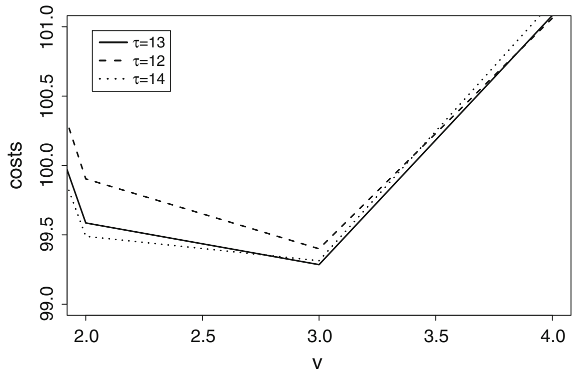


Fig. 5 Costs depending on v for different τ

In Fig. 5 the costs depending on v are plotted for the DRMWD, $c_F/c_I = 0.2$, $\delta = 0.5$. As one can see from Table 1, the optimal solution in the case is $v = 3, \tau = 13$.

6 Conclusion

In this paper an imperfect maintenance model with both imperfect preventive and imperfect corrective maintenance actions is investigated. A sequential failure limit PM policy with infinite planning horizon is used to formulate a cost optimization problem. The remaining lifetime of the system after a maintenance action is modeled as a truncated random variable. With the help of this truncated random variable the costs of corrective maintenance actions, the mean cycle length and thus the average maintenance costs per unit time can be computed. Optimal maintenance strategies are computed with R. The resulting cost optimal maintenance strategies shows that in special parameter constellations it is cost optimal to do quasi-non-stop PM actions if the costs of PM are close to zero.

References

1. Almalki S (2013) A reduced new modified Weibull distribution. arXiv:1307.3925
2. Almalki S, Nadarajah S (2014) A new discrete modified Weibull distribution. IEEE Trans Reliab 63:68–80
3. Barlow R, Proschan F (1965) Mathematical theory of reliability. Wiley, New York
4. Barlow R, Wu A (1978) Coherent systems with multi-state components. Math Oper Res 3(4):275–281
5. Dietrich S, Kahle W (2016) Optimal imperfect maintenance in a multi-state system with two failure types. In: Frenkel I, Lisnianski A (eds) Proceedings of the second international symposium on stochastic models in reliability engineering, life science and operations management (SMRLO16). IEEE CPS (2016), pp 233–243. doi:10.1109/SMRLO.2016.47

6. Doyen L, Gaudoin O (2011) Modeling and assessment of aging and efficiency of corrective and planned preventive maintenance. IEEE Trans Reliab 60(4):759–769
7. El-Neveihi E, Proschan F, Sethuraman J (1978) Multi-state coherent systems. J Appl Probab 15:675–688
8. Finkelstein M, Gertsbakh I, Vaisman R (2017) On a single discrete scale for preventive maintenance with two shock processes affecting a complex system. Appl Stoch Models Bus Ind 33(1):54–62
9. Gasmi S, Mannai N (2014) Optimization problems of costs in a multistate system with imperfect repair. J Qual Reliab Eng 10 pp. Article ID 592527
10. Kahle W (2012) Optimal incomplete maintenance in multi-state systems. In: Lisnianski A, Frenkel I (eds) Recent advances in system reliability. Springer, London, pp 209–218
11. Kijima M (1989) Some results for repairable systems with general repair. J Appl Probab 26(1):89–102
12. Lisnianski A (2012) Lz-transform for a discrete-state continuous-time Markov process and its applications to multi-state system reliability. In: Lisnianski A, Frenkel I (eds) Recent advances in system reliability. Springer, London, pp 79–95
13. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
14. Pham H, Wang H (1996) Imperfect maintenance. Eur J Oper Res 94:425–438
15. Sarhan A, Zaindin M (2009) Modified Weibull distribution. Appl Sci 11:123–136

Reliability Evaluation of Non-repairable Multi-state Systems Considering Survival-Death Markov Processes

Yan Yuan, Yi Ding, Chuanxin Guo and Yuanzhang Sun

Abstract Multi-state system (MSS) models have been extensively studied in recent years, because of their accuracy and flexibility for reliability evaluation of complex systems. One of the most important multi-state systems is the non-repairable multi-state system, which cannot be repaired during its operating time or whose repair is not economical. The "death" Markov process provides a basis for reliability analysis of the non-repairable multi-state system. It does not consider, however, the impact of start–up failures of components on system reliability. In this chapter, two models of modified "death" Markov processes considering component start–up failures are proposed. They are referred to as "survival-death" Markov processes and they differ in that the first model considers not only completely successful and failed start up but also partially successful start-up, whereas the second model only considers completely successful or failed start up. In such processes, the analytic expressions of the time-dependent transition probabilities can be obtained by using the Laplace-Stieltjes transform and the inverse Laplace-Stieltjes transform. The stochastic processes are combined with the *Lz*-transform technique for evaluating dynamic reliability of non-repairable MSS.

Y. Yuan
School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
e-mail: yymail@sjtu.edu.cn

Y. Ding (✉) · C. Guo
School of Electrical Engineering, Zhejiang University, Hangzhou, China
e-mail: yiding@zju.edu.cn

C. Guo
e-mail: guochuangxin@zju.edu.cn

Y. Sun
School of Electrical Engineering, Wuhan University, Wuhan, China
e-mail: yzsun@whu.edu.cn

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_10

1 Introduction

Compared with a conventional binary-state system model, a multi-state system (MSS) model is capable of describing systems with different levels of performance. MSS models have found wide applications in engineering and medicine [14], and have been extensively studied in the last three decades [1, 4, 6, 7, 9, 10, 13–15]. References [9, 10, 13, 14] provide insightful summaries of MSS reliability theory developments. In recent research, a general fuzzy multi-state system (FMSS) model has been proposed to extend the basic MSS model with crisp sets [2, 3].

The MSSs can be divided into two groups: non-repairable and repairable [13, 14]. To our knowledge, most existing research is more focused on the studies of repairable MSSs and usually long-term (steady state) system reliabilities are used to analyze them. However, there exist many non-repairable MSSs in real life applications. A non-repairable MSS consists of components, which either cannot be repaired during their operating time, or the repair is uneconomical, or only the life history up to the first failure is of interest [14]. Non-repairable systems and non-repairable components include both complex systems such as spacecraft and missile and most "one-shot" household devices. Instead of steady state reliabilities, short-term reliabilities (dynamic reliabilities) should be utilized to analyze the dynamic behavior of non-repairable MSSs.

Continuous-time discrete state Markov processes are widely used for the reliability analysis of MSSs because they fit very well in many practical cases. The "death" Markov process [13, 14, 16, 17] is a special case of the well-known "birth-death" Markov process [8, 13, 14, 17], where state transition only goes from state j to the adjacent state $j - 1$. The "death" Markov process provides a basis for reliability analysis of a non-repairable multi-state system with minor failures and degradation [13, 14]. The "death" Markov process is also used in ecology to evaluate the expected time of population extinction as well as the probability of population to extinct at certain time.

In the "death" process, it is usually assumed that a component has already been in the operational phase and begins from its best state K with maximal performance level. However, it does not consider failures of multi-state components during "start–up". For example, an emergency generating unit can fail to start when needed, which may have serious impact on system reliability. Though the basic Markov process of non-repairable components can be easily extended for including the start-up failures, the conceptual modeling and corresponding reliability solutions have not been comprehensively developed and evaluated based on our knowledge. This can lead to inaccurate and optimistic assessment of reliability indices of non-repairable MSSs.

In this chapter, two models of modified "death" Markov processes which account for start–up failures are presented in Sect. 2. They are named "survival-death" Markov processes. They are referred to as "survival-death" Markov processes and they differ in that the first model considers not only completely successful and failed start up but also partially successful start-up, whereas the second model only considers completely successful or failed start up. In "survival-death" Markov processes, the analytic expressions of the time-dependent state probabilities (dynamic probabilities) can be obtained by utilizing the Laplace-Stieltjes transform and the inverse Laplace-Stieltjes transform.

The universal generating function techniques (UGF) [5, 10, 13, 14, 18] have been widely used for MSS reliability evaluation. However, the Z-transform used in the UGF is only restricted for random variables. The Z-transform cannot be applied to discrete-state continuous-time stochastic process [11, 12]. Therefore, the basic UGF techniques are only suitable for steady-state reliability evaluation of repairable MSS, which cannot be used for analyzing dynamic reliability of non-repairable MSS. In some recent research [11, 12], the basic UGF techniques have been further extended and a special transform technique named as "Lz-transform" has been developed for discrete-state continuous-time stochastic process. In Sect. 3, this technique is applied for evaluating the dynamic reliability measures of non-repairable MSSs. Illustrative examples are also given.

Notation:

l	Component index
j, i	State index
$\lambda_{j,i}$	Transition rate from state j to state i
N	Number of system components
K^l	State number of component l
K^s	System state number
$\mathrm{g}^l = \{g_1^l, g_2^l, \ldots, g_K^l\}$	Performance set of component l
$\mathrm{p}^l(t) = \{p_0^l(t), p_1^l(t), \ldots, p_K^l(t)\}$	Probability set of component l at time instant t
$G^l(t)$	Performance level of component l at time instant t

2 Multi-state Models and Markov Processes for Non-repairable Components

In this section, key definitions and concepts of multi-state models and corresponding Markov processes for non-repairable components are introduced and developed. According to the generic MSS model, any system component l can have K^l different states corresponding to the performance levels, which can be represented by the set $\mathrm{g}^l = \{g_1^l, g_2^l, \ldots, g_{K^l}^l\}$ [10, 13, 14]. The performance level $G^l(t)$ of component l at time instant t takes a value from g^l: $G^l(t) \in \mathrm{g}^l$, which is a random variable. The performance level of component l forms a stochastic process during the time interval [0, T], where T is the operation period of the MSS. If the state probabilities at a future instant only depend on the current state, the stochastic process is considered as the Markov process.

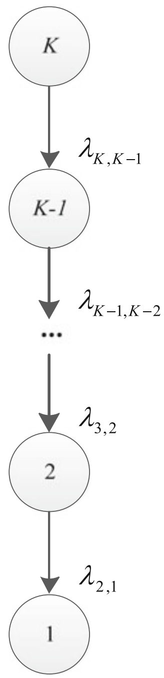


Fig. 1 State-space diagram for a non-repairable multi-state component with minor failures

In this section, the study is more focused on a single multi-state component. For simplifying expression, the component index l can be omitted. Therefore the set of performance levels is represented as $\mathrm{g} = \{g_1, g_2, \ldots, g_K\}$ and ordered so that $g_{j+1} > g_j$.

A non-repairable multi-state component transits from a higher state level to a lower state level. Firstly consider a multi-state component with only minor failures defined as failures that cause transition from state j to the adjacent state $j-1$ [13, 14]. During the deterioration process of the component, a minor failure causes minimal performance degradation. A typical state-space diagram of a multi-state component with minor failures is shown in Fig. 1 [14].

When the sojourn time in any state j is exponentially distributed with parameter $\lambda_{j,j-1}$, the process is the well known "death" Markov process [13, 14, 17]. Based on the standard Markov technique the following system of differential equations (Kolmogorov equations) can be written for the state probabilities of the component.

$$\begin{cases} \frac{dp_K(t)}{dt} = -\lambda_{K,K-1} \cdot p_K(t) \\ \frac{dp_j(t)}{dt} = \lambda_{j+1,j} \cdot p_{j+1}(t) - \lambda_{j,j-1} \cdot p_j(t), \quad for\ 1 < j < K \\ \frac{dp_1(t)}{dt} = \lambda_{2,1} \cdot p_2(t) \end{cases} \tag{1}$$

Suppose the process starts from the highest state K. The initial conditions of the state probabilities are:

$$p_K(0) = 1,\ p_{K-1}(0) = p_{K-2}(0) = \cdots = p_1(0) = 0 \quad (2)$$

The analytical solution of the system of differential Eqs. (1) under initial conditions (2) can be obtained using the Laplace–Stieltjes transform and inverse Laplace–Stieltjes transform [14]. More detailed discussions of analytical solution for the "death" Markov process can be found in [14, 17].

2.1 Model I for Non-repairable Components

The simplified model shown in Fig. 1 assumes that multi-state components have already started up and been in the operating cycles. However, the starting period is the most critical period in the system or/and component operation, where the start-up may succeed or not succeed or be partially successful. The model without considering start–up failures can lead to accuracy decreasing and optimistic assessment of reliability indices.

A more complex multi-state model that takes into consideration of the start-up process is illustrated in Fig. 2.

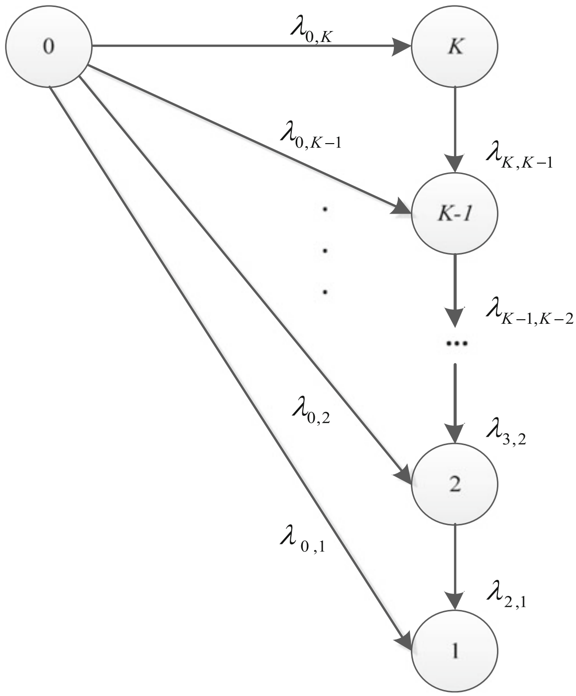


Fig. 2 State-space diagram for a non-repairable multi-state component with minor failures considering the start-up process (model I)

As shown in Fig. 2, the partially successful start-up causes the component transition from the start-up state to a performance derating state. The formal definition of Model I of a non-repairable multi-state component with minor failures considering the start-up process is given below.

Definition 1 The component may be in $K+1$ possible states: 0, 1, 2, …, K. State 0 and state 1 represent the start-up state and the completely failure state, respectively. State j $(1 < j \leq K)$ represents a functional state with a performance level $g_j > 0$. The component starts at state 0 at time $t = 0$. The component start-up can be completely successful, partially successful or failed. The completely successful or failed start-up causes the transition from state 0 to state K or state 1 with the transition rate $\lambda_{0,K}$ or $\lambda_{0,1}$, respectively. The partially successful start-up makes the transition from state 0 to a functional state j $(1 \leq j \leq K-1)$ with a derating performance level, which has the transition rate $\lambda_{0,j}$. Following a completely successful or partially successful start-up, the component degrades with minor failures with transition rate $\lambda_{j,j-1}$.

Suppose the transition time between any states is exponentially distributed. The homogeneous continuous-time Markov process describing the component evolution process is illustrated in Fig. 2. The states probabilities, $p_0(t)$, $p_K(t)$ and $p_j(t)$ for $1 < j < K$ for the Model I of the "survival-death" Markov process can be evaluated by solving the differential equations:

$$\begin{cases} \frac{dp_0(t)}{dt} = -c_0 \cdot p_0(t) \\ \frac{dp_K(t)}{dt} = \lambda_{0,K} \cdot p_0(t) - c_K \cdot p_K(t) \\ \frac{dp_j(t)}{dt} = \lambda_{0,j} \cdot p_0(t) + \lambda_{j+1,j} \cdot p_{j+1}(t) - c_j \cdot p_j(t), \, for \; 1 \leq j < K \end{cases} \tag{3}$$

where $c_j = \begin{cases} \sum_{i=1}^{K} \lambda_{0,i} & for \; j = 0 \\ \lambda_{j,j-1} & for \; 1 < j \leq K \\ 0 & for \; j = 1 \end{cases}$

The initial conditions are $p_0(0) = 1$, $p_K(0) = p_{K-1}(0) = \cdots = p_1(0) = 0$.

Under the initial conditions, the following linear algebraic equations can be obtained using the Laplace-Stieltjes transform:

$$\begin{cases} s\tilde{p}_0(s) = -c_0 \cdot \tilde{p}_0(s) \\ s\tilde{p}_K(s) = \lambda_{0,K} \cdot \tilde{p}_0(s) - c_K \cdot \tilde{p}_K(s) \\ s\tilde{p}_j(s) = \lambda_{0,j} \cdot \tilde{p}_0(s) + \lambda_{j+1,j} \cdot \tilde{p}_{j+1}(s) - c_j \cdot \tilde{p}_j(s), \; for \; 1 \leq j < K \end{cases} \tag{4}$$

Solving this system of algebraic equations (4) and apply the inverse Laplace-Stieltjes transform and obtain the probability function of state 0:

$$p_0(t) = e^{-c_0 t} \tag{5}$$

$c_i \neq c_l$ for any $i, l = 0, j, j+1, \ldots, K$ but $i \neq l$, therefore

$$p_j(t) = a_{j,0}e^{-c_0 t} + \sum_{l=j}^{K} a_{j,l}e^{-c_l t} \quad \text{for } 1 \leq j \leq K \tag{6}$$

where

$$a_{j,0} = \sum_{l=j}^{K} \frac{\lambda_{0,l}\lambda_{l,l-1}\lambda_{l-1,l-2}\cdots\lambda_{j+1,j}}{(c_l - c_0)(c_{l-1} - c_0)\cdots(c_j - c_0)} \tag{7}$$

$$a_{j,i} = \sum_{l=i}^{K} \frac{\lambda_{0,l}\prod_{m=j+1}^{l}\lambda_{m,m-1}}{(c_0 - c_i)\prod_{\substack{n=j \\ n \neq i}}^{l}(c_n - c_i)} \tag{8}$$

for $1 \leq j \leq K$ and $j \leq i \leq K$ where $\prod_{n=a}^{b}(\bullet) = 1$ when $a > b$.

2.2 Model II for Non-repairable Components

Model II is illustrated in Fig. 3, and its formal definition given below. Model II is the specific case of Model I. Model II can be used in many engineering applications e.g. power generating unit, where the start up can be simplified as completely successful or failed start up.

Definition 2 The component may be in $K+1$ possible states: 0, 1, 2, …, K. State 0 and state 1 represent the start-up state and the completely failure state, respectively. State j ($1 < j \leq K$) represents a functional state with a performance level $g_j > 0$. The component starts at state 0 at time $t = 0$. The successful or failed start-up causes the transition from state 0 to state K or state 1 with the transition rate $\lambda_{0,K}$ or $\lambda_{0,1}$, respectively. Following a successful start-up, the component degrades with minor failures with transition rate $\lambda_{j,j-1}$.

$$\begin{cases} p_0(t) = e^{-c_0 t} \\ p_j(t) = a_{j,0}e^{-c_0 t} + \sum_{l=j}^{K} a_{j,l}e^{-c_l t} \quad \textit{for } 1 \leq j \leq K \end{cases} \tag{9}$$

where

$$a_{j,0} = \left(\frac{\lambda_{0,K}}{c_K - c_0}\right) \cdot \left(\frac{\lambda_{K,K-1}}{c_{K-1} - c_0}\right) \cdot \ \ldots \ \cdot \left(\frac{\lambda_{j+1,j}}{c_j - c_0}\right) \quad \textit{for } 1 \leq j \leq K, \tag{10}$$

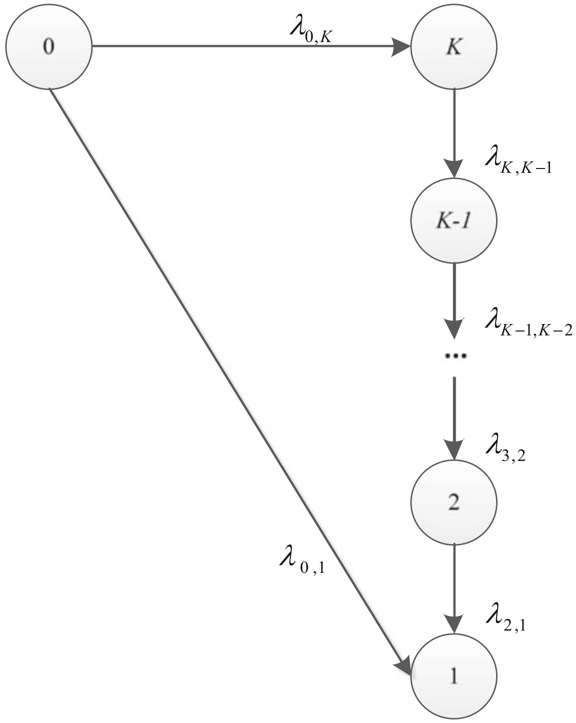


Fig. 3 State-space diagram for a non-repairable multi-state component with minor failures considering the start-up process (model II)

$$a_{j,i} = \left(\frac{\lambda_{0,K}}{c_0 - c_i}\right) \cdot \left(\frac{\lambda_{K,K-1}}{c_K - c_i}\right) \cdot \ldots \cdot \left(\frac{\lambda_{i+1,i}}{c_{i+1} - c_i}\right) \cdot \left(\frac{\lambda_{i,i-1}}{c_{i-1} - c_i}\right) \cdot \left(\frac{\lambda_{i-1,i-2}}{c_{i-2} - c_i}\right) \cdot \ldots \cdot \left(\frac{\lambda_{j+1,j}}{c_j - c_i}\right) \quad (11)$$
$$\text{for } 1 \le j \le K \text{ and } j \le i \le K$$

3 Dynamic Reliability Evaluation for Non-repairable Multi-state Systems

There are many optimization methods available and used to various reliability optimization problems [14]. The applied algorithms can be classified into two categories: heuristics and exact techniques. Most reliability optimization problems are solved by using general heuristic techniques based on artificial intelligence and stochastic techniques. Based on the classification of [14] and some recent research, the heuristics techniques include simulating annealing, ant colony, tabu searching, genetic algorithm and particle swarm optimization.

The L_Z-transform of a discrete-state continuous-time stochastic process $\{G^l(t)|t \ge 0\}$ of the component l can be represented as a function $u(z, t, \mathrm{p}(0))$ [11, 12] associated with $\{G^l(t)|t \ge 0\}$:

$$L_Z\left(G^l(t)\right) = u_l\left(z, t, \mathrm{p}^l(0)\right)$$
$$\sum_{j=0}^{K^l} p_j^l(t)\cdot z^{g_j^l} = p_0^l(t)\cdot z^{g_0^l} + p_1^l(t)\cdot z^{g_1^l} + \cdots + p_{K_l}^l(t)\cdot z^{g_{K_l}^l} \tag{12}$$

where $p_j^l(t)$ is the probability that the stochastic process is in state j at time $t \geq 0$ for given initial states probability distribution $\mathrm{p}^l(0) = \left\{p_0^l(0), p_1^l(0), p_2^l(0), \ldots, p_K^l(0)\right\}$, g_j^l is the performance level of the component l in state j, and z is a complex variable.

Equation (12) represents the performance distribution (PD) of the component l at time $t \geq 0$. The instantaneous state probabilities of $\mathrm{p}^l(\mathrm{t}) = \left\{p_0^l(t), p_1^l(t), p_2^l(t), \ldots, p_K^l(t)\right\}$ can be obtained analytically or numerically using the models proposed in Sect. 2.

The performance levels of state-up state (state 0) and completely failure state (state 1) can be zero. Therefore, these two states can be united in one state with the probability $(p_0^l(t) + p_1^l(t))$. The Lz-transform of output performance process of the component l can be represented as:

$$u_l\left(z, t, \mathrm{p}^l(0)\right) = \left(p_0^l(t) + p_1^l(t)\right)\cdot z^0 + \sum_{j=0}^{K^l} p_j^l(t)\cdot z^{g_j^l} = \sum_{\hat{j}=1}^{K^l} p_{\hat{j}}^l(t)\cdot z^{g_{\hat{j}}^l} \tag{13}$$

where $p_{\hat{1}}^l(t) = p_0^l(t) + p_1^l(t)$, $g_{\hat{1}}^l(t) = 0$, $p_{\hat{j}}^l(t) = p_j^l(t)$ and $g_{\hat{j}}^l(t) = g_j^l(t){:}2 \leq j \leq K$.

As it was proved in [11] in order to obtain Lz-transform of output performance process for entire MSS at time $t \geq 0$ with the arbitrary structure function ϕ, a general composition operator Ω_ϕ (Ushakov's universal generating operator) is applied to Lz-transforms of all individual system components over all time points t:

$$\begin{aligned}
U(z,t) &= \Omega_\phi\left\{u_1\left(z, t, \mathrm{p}^1(0)\right), \ldots, u_N\left(z, t, \mathrm{p}^N(0)\right)\right\} \\
&= \Omega_\phi\{\sum_{\hat{j}_1=1}^{K^1} p_{\hat{j}_1}^1(t)\cdot z^{g_{\hat{j}_1}^1}, \ldots \sum_{\hat{j}_N=1}^{K^N} p_{\hat{j}_N}^l(t)\cdot z^{g_{\hat{j}_N}^N}\} \\
&= \sum_{\hat{j}_1}^{K^1}\sum_{\hat{j}_1}^{K^2}\cdots\sum_{\hat{j}_N}^{K^N}(\prod_{l=1}^{N} p_{\hat{j}_l}^l(t)\cdot z^{\phi(g_{\hat{j}_1}^1, \ldots g_{\hat{j}_N}^N)}) \\
&= \sum_{j_s=1}^{K^s} p_{j_s}^s(t)\cdot z^{g_{j_s}^s}
\end{aligned} \tag{14}$$

where $U(z,t)$ is Lz-transform representation of output performance stochastic process for the entire MSS at time $t \geq 0$, K^s is the number of system states, and $p_{j_s}^s(t)$ and $g_{j_s}^s$ are probability at time $t \geq 0$ and performance level of system state j_s, respectively. The system performance is determined by the structure function $\phi(\cdot)$, which is strictly defined by the type of connection between elements in the reliability logic-diagram sense [13]. For example, consider a generating system with two units connected in parallel, we may have $\phi(G_1(t), G_2(t)) = G_1(t) + G_2(t)$.

After obtaining the Lz-transform for output performance process of the entire system, the system reliability at time instant t $(R(w,t))$ for the arbitrary constant demand w can be calculated. Generally, it is impossible to find entire system reliability based on its UGF. However, in our case where the system is non-repairable it can be done by using the operator δ_A [13, 14], which in general case (for repairable system) is used for availability computation:

$$R(w,t) = \delta_A(U(z,t), w) = \delta_A(\sum_{j_s=1}^{K_s} p_{j_s}^s(t) \cdot z^{g_{j_s}^s}, w) = \sum_{j_s=1}^{K_s} p_{j_s}^s(t) \cdot \alpha_{i_s} \quad (15)$$

where $\alpha_{i_s} = \begin{cases} 1, & (g_{j_s}^s - w) \geq 0 \\ 0, & (g_{j_s}^s - w) < 0 \end{cases}$

The expected instantaneous performance deficiency $(D(w,t))$ can be evaluated using the δ_D operator [13, 14]:

$$\begin{aligned} D(w,t) = \delta_D(U(z,t), w) &= \delta_D(\sum_{j_s=1}^{K_s} p_{j_s}^s(t) \cdot z^{g_{j_s}^s}, w) \\ &= \sum_{j_s=1}^{K_s} p_{j_s}^s(t) \cdot \max\left(w - g_{j_s}^s, 0\right) \end{aligned} \quad (16)$$

4 System Studies

In this section, we will use three application examples to illustrate the proposed models. The first and second examples correspond to multi-state generating units of model II and model I, respectively. The third example illustrates the reliability evaluation of a generation and transmission system represented by the MSS with non-repairable components. The analytic expressions of the time-dependent state probabilities of generating units of model II and model I are evaluated in the first and second examples, respectively. Based on that, the Lz-transform for the generation and transmission system is obtained and corresponding reliability indices are calculated.

4.1 Example 1

Consider a power generating unit that can be described by model II. The generating unit has 5 possible states: 0, 1, 2, 3, 4. State 0 and state 1 represent the start-up state and the completely failure state, respectively, both with zeroing performance levels, $g_0^{(1)} = g_1^{(1)} = 0$. The generating unit can either start up successfully and transit to the best performance state (state 4) or fail to start-up and goes into the completely failure state (state 1). States 2, 3, 4 represent functional states with performance levels $g_2^{(1)} = 50$ KW, $g_3^{(1)} = 80$ KW, $g_4^{(1)} = 100$ KW, respectively. The transition rates are $\lambda_{0,4}^{(1)} = 49932$ year^{-1}, $\lambda_{0,1}^{(1)} = 2628$ year^{-1}, $\lambda_{4,3}^{(1)} = 2$ year^{-1}, $\lambda_{3,2}^{(1)} = 1$ year^{-1}, $\lambda_{2,1}^{(1)} = 0.7$ year^{-1}. The initial state is state 0.

In order to find state probabilities the following system of differential equations should be solved:

$$\begin{cases} \frac{dp_0^{(1)}(t)}{dt} = -(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) \cdot p_0^{(1)}(t) \\ \frac{dp_4^{(1)}(t)}{dt} = \lambda_{0,4}^{(1)} \cdot p_0^{(1)}(t) - \lambda_{4,3}^{(1)} \cdot p_4^{(1)}(t) \\ \frac{dp_3^{(1)}(t)}{dt} = \lambda_{4,3}^{(1)} \cdot p_4^{(1)}(t) - \lambda_{3,2}^{(1)} \cdot p_3^{(1)}(t) \\ \frac{dp_2^{(1)}(t)}{dt} = \lambda_{3,2}^{(1)} \cdot p_3^{(1)}(t) - \lambda_{2,1}^{(1)} \cdot p_2^{(1)}(t) \\ \frac{dp_1^{(1)}(t)}{dt} = \lambda_{0,1}^{(1)} \cdot p_0^{(1)}(t) + \lambda_{2,1}^{(1)} \cdot p_2^{(1)}(t) \end{cases}$$

The initial conditions are $p_0^{(1)}(0) = 1, p_4^{(1)}(0) = p_3^{(1)}(0) = p_2^{(1)}(0) = p_1^{(1)}(0) = 0$.

Using the Laplace-Stieltjes transform and the inverse Laplace-Stieltjies transform in proper order, we obtain the state probabilities as functions of time t:

$$p_0^{(1)}(t) = e^{-(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)})t},$$

$$p_4^{(1)}(t) = \frac{\lambda_{0,4}^{(1)}}{\lambda_{4,3}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)})} e^{-(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)})t} + \frac{\lambda_{0,4}^{(1)}}{(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{4,3}^{(1)}} e^{-\lambda_{4,3}^{(1)}t};$$

$$\begin{aligned} p_3^{(1)}(t) = {} & \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}}{(\lambda_{4,3}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}))(\lambda_{3,2}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}))} e^{-(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)})t} \\ & + \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}}{((\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{4,3}^{(1)})(\lambda_{3,2}^{(1)} - \lambda_{4,3}^{(1)})} e^{-\lambda_{4,3}^{(1)}t} \\ & + \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}}{((\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{3,2}^{(1)})(\lambda_{4,3}^{(1)} - \lambda_{3,2}^{(1)})} e^{-\lambda_{3,2}^{(1)}t}; \end{aligned}$$

$$\begin{aligned} p_2^{(1)}(t) = {} & \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}\lambda_{3,2}^{(1)}}{(\lambda_{4,3}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}))(\lambda_{3,2}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}))(\lambda_{2,1}^{(1)} - (\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}))} e^{-(\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)})t} \\ & + \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}\lambda_{3,2}^{(1)}}{((\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{4,3}^{(1)})(\lambda_{3,2}^{(1)} - \lambda_{4,3}^{(1)})(\lambda_{2,1} - \lambda_{4,3}^{(1)})} e^{-\lambda_{4,3}^{(1)}t} \\ & + \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}\lambda_{3,2}^{(1)}}{((\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{3,2}^{(1)})(\lambda_{4,3}^{(1)} - \lambda_{3,2}^{(1)})(\lambda_{2,1}^{(1)} - \lambda_{3,2}^{(1)})} e^{-\lambda_{3,2}^{(1)}t} \\ & + \frac{\lambda_{0,4}^{(1)}\lambda_{4,3}^{(1)}\lambda_{3,2}^{(1)}}{((\lambda_{0,4}^{(1)} + \lambda_{0,1}^{(1)}) - \lambda_{2,1}^{(1)})(\lambda_{4,3}^{(1)} - \lambda_{2,1}^{(1)})(\lambda_{3,2}^{(1)} - \lambda_{2,1}^{(1)})} e^{-\lambda_{2,1}^{(1)}t}; \end{aligned}$$

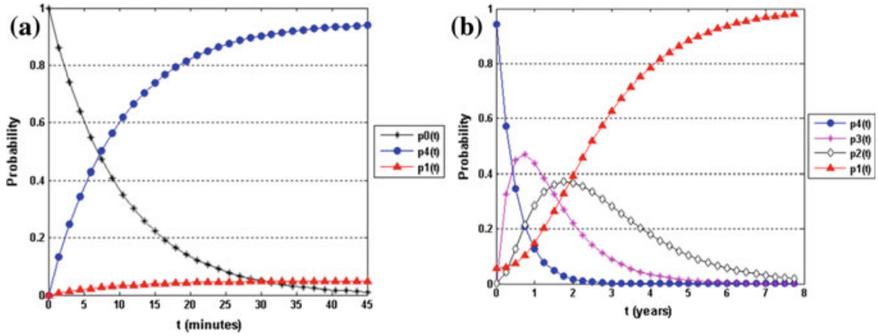


Fig. 4 State probabilities with different time scale for Example 1

$$p_1^{(1)}(t) = 1 - p_2^{(1)}(t) - p_3^{(1)}(t) - p_4^{(1)}(t) - p_0^{(1)}(t)$$

The time-dependent state probabilities with different time scales for Example 1 are presented in Fig. 4. The Fig. 4a, b illustrates the state probabilities in the short-term (45 min) and long term (8 years), respectively. It can be observed from the Fig. 4a that the probabilities in state 1 and state 4 increase with time and become relatively stable. Meanwhile the probability in state 0 gradually decreases and reaches to zero after 45 min.

The probabilities of other states are very small in the short term and are not shown in Fig. 4a. Therefore the degradation impact of the generation unit can be neglected in the short-term operation. In the long term, the generating unit will reside in the any 3 functional states (states 2–4) or state 0, as shown in Fig. 4b. The probability in state 1 (completely failure state) gradually increases and reaches to one after 8 years.

4.2 *Example 2*

Consider a power-generating unit that can be described by model I. The generating unit also has 5 possible states: 0, 1, 2, 3, 4. In this example, besides the completely successful start-up or failed start-up, the partially successful start-up can cause the component transition from the start-up state to a performance derating state. The transition rates are $\lambda_{0,4}^{(2)} = 42048\ \text{year}^{-1}$, $\lambda_{0,3}^{(2)} = 7884\ \text{year}^{-1}$, $\lambda_{0,2}^{(2)} = 2102\ \text{year}^{-1}$, $\lambda_{0,1}^{(2)} = 526\ \text{year}^{-1}$, $\lambda_{4,3}^{(2)} = 2\ \text{year}^{-1}$, $\lambda_{3,2}^{(2)} = 1\ \text{year}^{-1}$, $\lambda_{2,1}^{(2)} = 0.7\ \text{year}^{-1}$. Other conditions are the same as those in the Example 1. The initial state is state 0. In order to find state probabilities the following system of differential equations should be solved:

$$\begin{cases} \frac{dp_0^{(2)}(t)}{dt} = -c_0 \cdot p_0^{(2)}(t) \quad \text{where } c_0 = \lambda_{0,4}^{(2)} + \lambda_{0,3}^{(2)} + \lambda_{0,2}^{(2)} + \lambda_{0,1}^{(2)} \\ \frac{dp_4^{(2)}(t)}{dt} = \lambda_{0,4}^{(2)} \cdot p_0^{(2)}(t) - \lambda_{4,3}^{(2)} \cdot p_4^{(2)}(t) \\ \frac{dp_3^{(2)}(t)}{dt} = \lambda_{0,3}^{(2)} \cdot p_0^{(2)}(t) + \lambda_{4,3}^{(2)} \cdot p_4^{(2)}(t) - \lambda_{3,2}^{(2)} \cdot p_3^{(2)}(t) \\ \frac{dp_2^{(2)}(t)}{dt} = \lambda_{0,2}^{(2)} \cdot p_0^{(2)}(t) + \lambda_{3,2}^{(2)} \cdot p_3^{(2)}(t) - \lambda_{2,1}^{(2)} \cdot p_2^{(2)}(t) \\ \frac{dp_1^{(2)}(t)}{dt} = \lambda_{0,1}^{(2)} \cdot p_0^{(2)}(t) + \lambda_{2,1}^{(2)} \cdot p_2^{(2)}(t) \end{cases}$$

The initial conditions are $p_0^{(2)}(0) = 1, p_4^{(2)}(0) = p_3^{(2)}(0) = p_2^{(2)}(0) = p_1^{(2)}(0) = 0$.

$$p_0^{(2)}(t) = e^{-c_0 t};$$

$$p_4^{(2)}(t) = \frac{\lambda_{0,4}^{(2)}}{\lambda_{4,3}^{(2)} - c_0} e^{-c_0 t} + \frac{\lambda_{0,4}^{(2)}}{c_0 - \lambda_{4,3}^{(2)}} e^{-\lambda_{4,3}^{(2)} t};$$

$$p_3^{(2)}(t) = \left(\frac{\lambda_{0,3}^{(2)}}{\lambda_{3,2} - c_0} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)}}{(\lambda_{4,3}^{(2)} - c_0)(\lambda_{3,2}^{(2)} - c_0)}\right) e^{-c_0 t} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)}}{(c_0 - \lambda_{4,3}^{(2)})(\lambda_{3,2}^{(2)} - \lambda_{4,3}^{(2)})} e^{-\lambda_{4,3}^{(2)} t} + \left(\frac{\lambda_{0,3}^{(2)}}{c_0 - \lambda_{3,2}^{(2)}} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)}}{(c_0 - \lambda_{3,2}^{(2)})(\lambda_{4,3}^{(2)} - \lambda_{3,2}^{(2)})}\right) e^{-\lambda_{3,2}^{(2)} t};$$

$$p_2^{(2)}(t) = a_1 e^{-c_0 t} + a_2 e^{-\lambda_{4,3}^{(2)} t} + a_3 e^{-\lambda_{3,2}^{(2)} t} + a_4 e^{-\lambda_{2,1}^{(2)} t}$$

where

$$a_1 = \frac{\lambda_{0,2}^{(2)}}{\lambda_{2,1}^{(2)} - c_0} + \frac{\lambda_{0,3}^{(2)} \lambda_{3,2}^{(2)}}{(\lambda_{3,2}^{(2)} - c_0)(\lambda_{2,1}^{(2)} - c_0)} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)} \lambda_{3,2}^{(2)}}{(\lambda_{4,3}^{(2)} - c_0)(\lambda_{3,2}^{(2)} - c_0)(\lambda_{2,1}^{(2)} - c_0)}$$

$$a_2 = \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)} \lambda_{3,2}^{(2)}}{(c_0 - \lambda_{4,3}^{(2)})(\lambda_{3,2}^{(2)} - \lambda_{4,3}^{(2)})(\lambda_{2,1}^{(2)} - \lambda_{4,3}^{(2)})}$$

$$a_3 = \frac{\lambda_{0,3}^{(2)} \lambda_{3,2}^{(2)}}{(c_0 - \lambda_{3,2}^{(2)})(\lambda_{2,1}^{(2)} - \lambda_{3,2}^{(2)})} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)} \lambda_{3,2}^{(2)}}{(c_0 - \lambda_{3,2}^{(2)})(\lambda_{4,3}^{(2)} - \lambda_{3,2}^{(2)})(\lambda_{2,1}^{(2)} - \lambda_{3,2}^{(2)})}$$

$$a_4 = \frac{\lambda_{0,2}^{(2)}}{(c_0 - \lambda_{2,1}^{(2)})} + \frac{\lambda_{0,3}^{(2)} \lambda_{3,2}^{(2)}}{(c_0 - \lambda_{2,1}^{(2)})(\lambda_{3,2}^{(2)} - \lambda_{2,1}^{(2)})} + \frac{\lambda_{0,4}^{(2)} \lambda_{4,3}^{(2)} \lambda_{3,2}^{(2)}}{(c_0 - \lambda_{2,1}^{(2)})(\lambda_{4,3}^{(2)} - \lambda_{2,1}^{(2)})(\lambda_{3,2}^{(2)} - \lambda_{2,1}^{(2)})}$$

$$p_1^{(2)}(t) = 1 - p_2^{(2)}(t) - p_3^{(2)}(t) - p_4^{(2)}(t) - p_0^{(2)}(t)$$

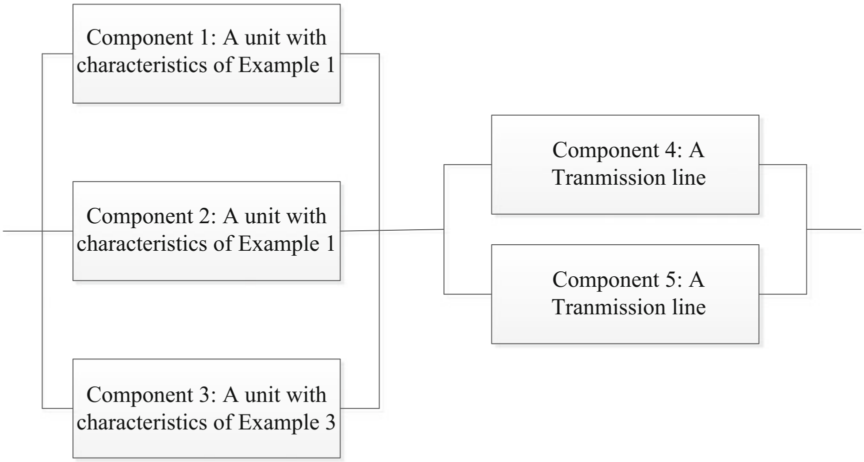


Fig. 5 Generation and transmission system

4.3 *Example 3*

An application of the proposed method is to evaluate the reliability of a generation and transmission system consisting of three generating units and two transmission lines as shown in Fig. 5.

Components 1 and 2 are two generating units with characteristics of Example 1. Component 3 is a generating unit with characteristics of Example 2. Components 1 and 2 are two identical transmission lines with the transmission capacity of 165 KW each, represented by binary state models. The failure rate of each transmission line is $\lambda_{2,1}^{(3)} = 0.51$ year^{-1}. The reliability model of the generation and transmission system can be represented as a parallel-series MSS. It is supposed that components cannot be repaired during their operating time. The initial states of generating units and transmission lines are start-up states and successful states, respectively. The system will carry a demand ($w = 100$ KW) from the time instant $t_0 = 30$ min.

The time-dependent state probabilities of transmission line in successful state $(p_2^{(3)}(t))$ and failure state $(p_1^{(3)}(t))$ can be represented as: $p_2^{(3)}(t) = e^{-\lambda_{2,1}^{(3)} \cdot t}$, $p_1^{(3)}(t) = 1 - p_2^{(3)}(t)$. The time-dependent state probabilities of generating units are obtained from Examples 1 and 2, respectively.

The Lz-transform of the components 1 and 2 can be represented as:

$$\begin{aligned} u_{l1}\left(z, t, \mathrm{p}^{l1}(0)\right) &= u_{l2}\left(z, t, \mathrm{p}^{l2}(0)\right) \\ &= (p_0^{(1)}(t) + p_1^{(1)}(t))z^{g_1^{(1)}} + p_2^{(1)}(t)z^{g_2^{(1)}} + p_3^{(1)}(t)z^{g_3^{(1)}} + p_4^{(1)}(t)z^{g_4^{(1)}} \\ &= (p_0^{(1)}(t) + p_1^{(1)}(t))z^0 + p_2^{(1)}(t)z^{50} + p_3^{(1)}(t)z^{80} + p_4^{(1)}(t)z^{100} \end{aligned}$$

The Lz-transform of the component 3

$$u_{l3}\left(z,t,\mathrm{p}^{l3}(0)\right) = (p_0^{(2)}(t)+p_1^{(2)}(t))z^{g_1^{(2)}} + p_2^{(2)}(t)z^{g_2^{(2)}} + p_3^{(2)}(t)z^{g_3^{(2)}} + p_4^{(2)}(t)z^{g_4^{(2)}}$$
$$= (p_0^{(2)}(t)+p_1^{(2)}(t))z^{0} + p_2^{(2)}(t)z^{50} + p_3^{(2)}(t)z^{80} + p_4^{(2)}(t)z^{100}$$

The three generating units are connected in parallel, which can be represented as a generation sub-system. The Lz-transform of the generation sub-system can be represented by utilizing the parallel operator:

$$\begin{aligned}U_{l,I}(z,t) &= \Omega_{\phi_P}\left(u_{l1}\left(z,t,\mathrm{p}^{l1}(0)\right), u_{l2}\left(z,t,\mathrm{p}^{l2}(0)\right), u_{l3}\left(z,t,\mathrm{p}^{l3}(0)\right)\right)\\
&= (p_0^{(1)}(t)+p_1^{(1)}(t))^2(p_0^{(2)}(t)+p_1^{(2)}(t))z^0\\
&\quad + (2p_2^{(1)}(t)(p_0^{(1)}(t)+p_1^{(1)}(t))(p_0^{(2)}(t)+p_1^{(2)}(t)) + (p_0^{(1)}(t)+p_1^{(1)}(t))^2p_2^{(2)}(t))z^{50}\\
&\quad + \cdots + ((p_4^{(1)}(t))^2p_3^{(2)}(t) + 2p_3^{(1)}(t)p_4^{(1)}(t)p_4^{(2)}(t))z^{280}\\
&\quad + (p_4^{(1)}(t))^2p_4^{(2)}(t)z^{300}\end{aligned}$$

The Lz-transform of the components 4 and 5 can be represented as:

$$\begin{aligned}u_{l4}\left(z,t,\mathrm{p}^{l4}(0)\right) &= u_{l5}\left(z,t,\mathrm{p}^{l5}(0)\right)\\
&= p_1^{(3)}(t)z^{g_1^{(3)}} + p_2^{(3)}(t)z^{g_2^{(3)}}\\
&= p_1^{(3)}(t)z^{0} + p_2^{(3)}(t)z^{165}\end{aligned}$$

The two transmission lines are connected in parallel, which can be represented as a transmission sub-system. The Lz-transform of the transmission sub-system can be represented as by utilizing the parallel operator:

$$\begin{aligned}U_{l,II}(z,t) &= \Omega_{\phi_P}\left(u_{l4}\left(z,t,\mathrm{p}^{l4}(0)\right), u_{l5}\left(z,t,\mathrm{p}^{l5}(0)\right)\right)\\
&= (p_1^{(3)}(t))^2z^{g_1^{(3)}} + 2p_1^{(3)}(t)p_2^{(3)}(t)z^{g_1^{(3)}+g_2^{(3)}} + (p_2^{(3)}(t))^2z^{2g_2^{(3)}}\\
&= (p_1^{(3)}(t))^2z^{0} + 2p_1^{(3)}(t)p_2^{(3)}(t)z^{165} + (p_2^{(3)}(t))^2z^{330}\end{aligned}$$

The generation sub-system and the transmission sub-system are connected in series. The Lz-transform of the system can be represented as by utilizing the series operator:

$$\begin{aligned}U_S(z,t) &= \Omega_{\phi_S}(U_{l,I}(z,t), U_{l,II}(z,t))\\
&= ((p_1^{(3)}(t))^2 + (p_0^{(1)}(t)+p_1^{(1)}(t))^2(p_0^{(2)}(t)+p_1^{(2)}(t))(1-(p_1^{(3)}(t))^2))z^0\\
&\quad + (2p_2^{(1)}(t)(p_0^{(1)}(t)+p_1^{(1)}(t))(p_0^{(2)}(t)+p_1^{(2)}(t)) + (p_0^{(1)}(t)+p_1^{(1)}(t))^2p_2^{(2)}(t))(1-(p_1^{(3)}(t))^2)z^{50}\\
&\quad + \cdots + ((p_4^{(1)}(t))^2p_3^{(2)}(t) + 2p_3^{(1)}(t)p_4^{(1)}(t)p_4^{(2)}(t))(p_2^{(3)}(t))^2z^{280}\\
&\quad + (p_4^{(1)}(t))^2p_4^{(2)}(t)(p_2^{(3)}(t))^2z^{300}\end{aligned}$$

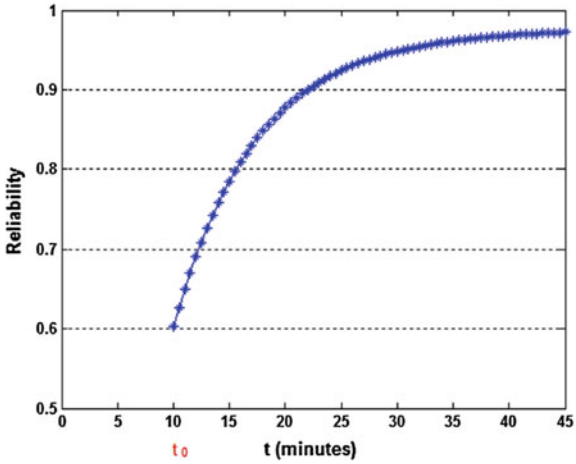


Fig. 6 Reliability of the system in the short-term

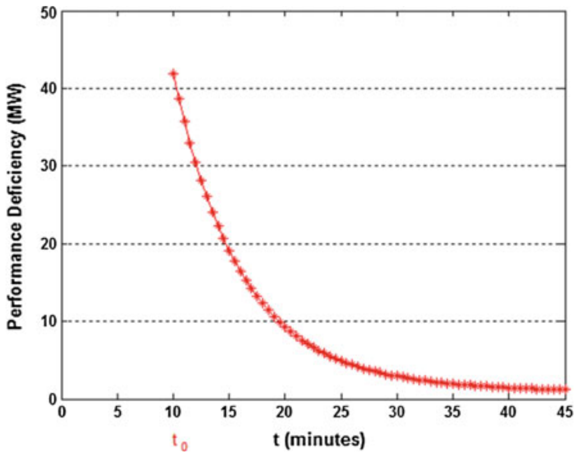


Fig. 7 Expected instantaneous performance deficiency in the short-term

After obtaining *Lz*-transform for output performance process of the entire system, the system reliability and expected instantaneous performance deficiency can be obtained, respectively. The Figs. 6 and 7 illustrate the system reliability and expected instantaneous performance deficiency in the short-term (45 min), respectively.

It can be observed from the Figs. 6 and 7 that the system reliability and expected instantaneous performance deficiency increases and decrease with the operating time in the short-term.

5 Conclusions

The non-repairable multi-state systems are important MSS models, which find many applications in practical cases. The pure "death" Markov process is an important tool for analyzing the non-repairable multi-state systems, which, however, has not considered the impact of start–up failures of components. The start–up failures may have serious impact on the reliabilities of the non-repairable multi-state systems during their mission time. Two models of modified "death" Markov processes considering component start–up failures are proposed in this chapter, which are named "survival-death" Markov processes. The Laplace-Stieltjes transform and the inverse Laplace-Stieltjes transform are used to obtain the analytic expressions of the time-dependent state probabilities of the "survival-death" Markov processes, which can provide accurate and important information for reliability evaluation. A general multi-state model for non-repairable components is also developed in this chapter. The proposed stochastic processes are combined the *Lz*-transform technique for evaluating the dynamic reliabilities of multi-state systems.

Acknowledgements The research is partially supported by the China NSFC under Grant 51537010 and 51611130197.

References

1. Barlow R, Wu A (1978) Coherent systems with multi-state elements. Math Oper Res 3(4):275–281
2. Ding Y, Lisnianski A (2008) Fuzzy universal generating functions for multi-state system reliability assessment. Fuzzy Sets Syst 159(3):307–324
3. Ding Y, Zuo M, Lisnianski A, Tian Z (2008) Fuzzy multi-state system: general definition, and performance assessment. IEEE Trans Reliab 57(4):589–594
4. El-Neweihi E, Proschan F, Sethuraman J (1978) Multistate coherent systems. J Appl Probab 15:675–688
5. Gnedenko B, Ushakov I (1995) Probabilistic reliability engineering. Wiley, NY/Chichester/Brisbane
6. Huang J, Zuo M, Wu Y (2000) Generalized multi-state k-out-of-n: G systems. IEEE Trans Reliab 49(1):105–111
7. Huang J, Zuo M, Fang Z (2003) Multi-state consecutive-k-out-of-n systems. IIE Trans 35(6):527–534
8. Kendall D (1948) On the generalized birth-and-death process. Ann Math Stat 19(1):1–15
9. Kuo W, Zuo M (2003) Optimal reliability modeling principles and applications. Wiley, New York
10. Levitin G (2005) Universal generating function and its applications. Springer
11. Lisnianski A (2012) Lz-transform for a discrete-state continuous-time Markov process and its application to multi-state system reliability. Recent advances in system reliability. signatures, multi-state systems and statistical inference. Springer, London, pp 79–95
12. Lisnianski A (2012) Reliability assessment for aging multi-state system under stochastic demand by using LZ-TRANSFORM. In: Proceedings of the 18th ISSAT international conference on reliability and quality in design, 26–28 July 2012, Boston, USA

13. Lisnianski A, Levitin G (2003) Multi-state system reliability assessment, optimization, applications. World Scientific, NJ
14. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
15. Murchland J (1975) Fundamental concepts and relations for reliability analysis of multi-state systems and fault tree analysis. In: Reliability and fault tree analysis, international atomic energy agency, INIS volume 8(04)
16. Taylor H, Karlin S (1889) An introduction to stochastic modeling. Academic Press, San Diego
17. Trivedi K (2002) Probability and statistics with reliability, queuing and computer science applications. Wiley, New York
18. Ushakov I (1986) Universal generating function. Sov J Comput Syst Sci 24(5):118–129

Reliability Assessment of Systems with Dependent Degradation Processes Based on Piecewise-Deterministic Markov Process

Yan-Hui Lin, Yan-Fu Li and Enrico Zio

Abstract This chapter presents a reliability assessment framework for multi-component systems whose degradation processes are modeled by multi-state and physics-based models. The piecewise-deterministic Markov process modeling approach is employed to treat dependencies between the degradation processes within one component or/and among components. The proposed method can handle the dependencies between physics-based models, between multi-state models and between these two types of models. A Monte Carlo simulation algorithm is designed to compute the systems reliability. A case study on one subsystem of the residual heat removal system of a nuclear power plant is illustrated as exemplification of the proposed modeling framework.

Keywords Multi-state model · Physics-based model · Dependent degradation processes · Piecewise-deterministic markov process · Monte Carlo simulation method

---

Y.-H. Lin (✉)
School of Reliability and Systems Engineering, Beihang University, Beijing, China
e-mail: linyanhui@buaa.edu.cn

Y.-F. Li
Department of Industrial Engineering, Tsinghua University, Beijing, China
e-mail: liyanfu@tsinghua.edu.cn

E. Zio
Chair on Systems Science and the Energetic Challenge, Fondation EDF, at Centrale Supelec, Paris, France
e-mail: enrico.zio@ecp.fr; enrico.zio@polimi.it

E. Zio
Politecnico di Milano, Milan, Italy

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_11

1 Introduction

Safety-critical plants, like the nuclear power plants (NPPs), are designed not to fail, i.e. with very high reliability, because of the potentially catastrophic consequences of their failures. Traditional data-based reliability analysis, based on failure data, is, then, unsuitable. On the other hand, most failure mechanisms can be traced to underlying degradation processes (e.g. wear, stress corrosion, shocks, cracking, fatigue, etc.) [30], for which models exist.

In general, the reliability of a system decreases as the degradation processes develop, eventually leading to failure [31]. In reliability engineering, degradation processes have been widely studied and different degradation models have been developed. The existing degradation models can mainly be classified into the following categories:

- statistical models of time to failure, based on degradation data (e.g. Bernstein distribution [9], Weibull distribution [21]).
- stochastic process models (e.g. Gamma processes [14], inverse Gaussian process [2]) describing the evolution of one or more degradation parameters by gradual degradation increments over time, and the failure occurs when the degradation parameter values reach predefined thresholds.
- physics-based models (PBMs), based on the knowledge of the physics of degradation, which is translated into equations to give a quantitative description (e.g. the physics functions based on critical environmental stresses, e.g. amplitude and frequency of mechanical loads, used to model the pitting and corrosion-fatigue degradation mechanisms [3]).
- multi-state models (MSMs) describing the underlying degradation process by finite degradation states (e.g. semi-Markov models for the deterioration of infrastructure systems [1]).

Among these categories, PBMs [6, 13, 27] and MSMs [18, 19, 23] can be used to describe the evolution of degradation in structures, systems and components, for which statistical degradation/failure data are insufficient, e.g. the highly reliable devices in the nuclear and aerospace industries.

In reality, components and systems are often subject to multiple competing degradation processes and any of them may cause failure [29]. The dependencies among these processes within one component (e.g. the wear of rubbing surfaces influenced by the environmental stress shock within a micro-engine [12]), or/and among different components (e.g. the degradation of the pre-filtrations stations leading to a lower performance level of the sand filter in a water treatment plant [26]) need to be considered, under certain circumstances. This renders challenging the analysis and prediction of the components and systems reliability [25]. Wang and Pham [29] applied time-varying copulas for describing the dependencies between the degradation processes modeled by statistical distributions. Straub [28] used a dynamic Bayesian network to represent the dependencies between degradation processes modeled by multi-state models.

In this chapter, we present a reliability assessment framework for multi-component systems whose degradation processes are modeled by MSMs and PBMs, capturing dependencies among the components and among multiple degradation processes within one component. The piecewise-deterministic Markov process (PDMP) modeling approach is employed. The PDMP, firstly introduced by Davis in [7, 8], and further studied by Jacobsen [11] and Cocozza-Thivent [4], is well-suited to describe degradation dependence. The remainder of this chapter is organized as follows. Section 2 presents the proposed degradation model for systems with degradation dependence. Monte Carlo simulation procedures to solve the model are presented in Sect. 3. Section 4 presents a case study on one subsystem of a residual heat removal system (RHRS) of a NPP. Section 5 concludes the work.

2 Dynamic Reliability Models for Systems with Degradation Dependence

For highly reliable systems, such as nuclear safety systems, it is relatively difficult to model their degradation and failure behaviors due to the limited amount of data available. In these cases, PBMs and MSMs are two modeling frameworks that can be used to model degradation. Systems are often subject to multiple competing degradation processes and any of them may cause failure. The dependences among these processes need to be considered under certain circumstances. In this chapter, a PDMP modeling framework is developed to treat degradation dependence in a system whose degradation processes are modeled by PBMs and MSMs.

2.1 Degradation Models

We consider a multi-component system made of Q components denoted by $O = \{O_1, O_2, \ldots, O_Q\}$. Each component may be affected by multiple degradation mechanisms or processes, possibly dependent. The degradation processes can be separated into two groups: (1) $L = \{L_1, L_2, \ldots, L_M\}$ modeled by M PBMs; (2) $K = \{K_1, K_2, \ldots, K_N\}$ modeled by N MSMs, where $L_m, m = 1, 2, \ldots, M$ and $K_n, n = 1, 2, \ldots, N$ are the indexes of the degradation processes.

2.1.1 PBMs

The following assumptions on PBMs are made:

- A degradation process $X_{L_m}(t), L_m \in L$ in the first group, has d_{L_m} time-dependent continuous variables $X_{L_m}(t) = \left(x^1_{L_m}(t), x^2_{L_m}(t), \ldots, x^{d_{Lm}}_{L_m}(t)\right) \in \mathbb{R}^{d_{Lm}}$. A system of first-order differential equations (i.e. physics equations) $\dot{X}_{L_m}(t) = f_{L_m}(X_{L_m}(t), t|\boldsymbol{\theta}_{L_m})$, are used to characterize its evolution, where $\boldsymbol{\theta}_{L_m}$ are the environmental factors influential to L_m (e.g. temperature and pressure) and the parameters used in f_{L_m}. This assumption is made in [20] and widely used in practice [5, 6]. Note that higher-order differential equations can be converted into a system of a large number of first-order differential equations by introducing extra variables [33].
- $X_{L_m}(t)$ can be divided into two groups of variables $X_{L_m}(t) = \left(X^D_{L_m}(t), X^P_{L_m}(t)\right)$: (1) $X^D_{L_m}(t)$ are the non-decreasing degradation variables describing the degradation process (e.g. leak area of the piston of the valve [6]), where D is the set of degradation variables indices; (2) $X^P_{L_m}(t)$ are the physical variables influencing $X^D_{L_m}(t)$ (e.g. velocity and force [5]), where P is the set of physical variable indices. For example, the friction-induced wear of the bearings is considered as one degradation process in [5]. It is represented by the increase in friction coefficients. The two friction coefficients associated with sliding and rolling friction are considered as the degradation variables. The rotational velocity of the pump is considered as the physical variable, since it influences the increase in the coefficients of friction. The evolution of physical variables can be characterized by physics equations. If the variables can be modeled by physics equations and influence certain degradation variables, then, they are considered as physical variables. As long as one $x^i_{L_m}(t) \in X^D_{L_m}(t)$ reaches or exceeds its corresponding failure threshold $x_{(L_m)}i^*$ the generic degradation process L_m fails. Let $\mathcal{F}_{L_m}$ denote the failure state set of L_m and $x^*_{L_m}$ denote the set of all the failure thresholds of $X^D_{L_m}(t)$. An example of L_1 is shown in Fig. 1.

2.1.2 MSMs

The following assumptions on MSMs are made:

- A degradation process, $Y_{K_n}(t), K_n \in K$ in the second group, takes values from a finite state set denoted by $S_{K_n} = \{0, 1, \ldots, d_{K_n}\}$, where '$d_{K_n}$' is the perfect functioning state and '0' is the complete failure state. The transition rates $\lambda_i(j|\boldsymbol{\theta}_{K_n}), \forall i, j \in S_{K_n}, i > j$ characterize the degradation transition probabilities from state i to state j, where $\boldsymbol{\theta}_{K_n}$ is the set of the environmental factors to K_n and

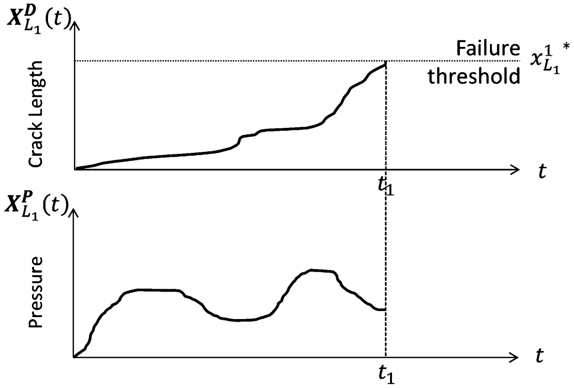


Fig. 1 An illustration of L_1

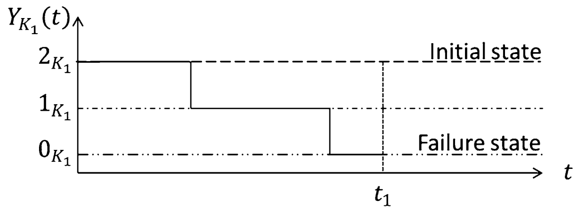


Fig. 2 An illustration of K_1

the related parameters used in λ_i. We follow the assumption of Markov property which is widely used in practice to describe components degradation processes [10]. The transition rates between different degradation states are estimated from the degradation and/or failure data from historical field collection. Let $\mathcal{F}_{K_n} = \{0\}$ denote the failure state set of K_n. An example of K_1 is shown in Fig. 2.

2.2 Degradation Model of the System Considering Dependence

The dependencies between degradation mechanisms or processes may exist within each group and between the two groups. The evolution trajectories of the continuous variables in the first group may be influenced by the degradation states of the second group. The transition times and transition directions of the degradation processes of the second group may depend on the degradation levels of the components in the first group [17]. PDMPs [4], which are a family of Markov processes involving deterministic evolution punctuated by random jumps, can be employed to model this type of dependence (the detailed formulations are shown in Eqs. (2) and (3)). Let $X(t)=\begin{pmatrix} X_{L_1}(t) \\ \vdots \\ X_{L_M}(t) \end{pmatrix}$ denote the degradation processes of the first group and $Y(t)=\begin{pmatrix} Y_{K_1}(t) \\ \vdots \\ Y_{K_N}(t) \end{pmatrix}$ denote the degradation processes of the second group. The overall degradation process of the system is presented as

$$Z(\mathrm{t})=\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} \in E=\mathbb{R}^{d_L}\times S \tag{1}$$

where E is a space combining $\mathbb{R}^{d_L}$ $(d_L=\sum_{m=1}^{M} d_{L_m})$ and $S=\{0,1,\ldots,d_S\}$ denotes the state set of process $Y(t)$. The evolution of $Z(t)$ has two parts: (1) the stochastic behavior of $Y(t)$ and (2) the deterministic behavior of $X(t)$ between two consecutive jumps of $Y(t)$, given $Y(t)$. The former is governed by the transition rates of $Y(t)$, which depend on the states of the degradation processes in $X(t)$ and also in $Y(t)$, as follows:

$$\begin{aligned} &\lim_{\Delta t\to 0} P\left(Y(t+\Delta t)=j|X(t),Y(t)=i,\boldsymbol{\theta}_K=\cup_{n=1}^{N}\boldsymbol{\theta}_{K_n}\right)/\Delta t \\ &=\lambda_i(j|X(t),\boldsymbol{\theta}_K),\forall t\geq 0,i,j\in S,i\neq j \end{aligned} \tag{2}$$

The latter is described by the deterministic physics, which depends on the states of the degradation processes in $Y(t)$ and also in $X(t)$, as follows:

$$\begin{aligned} \dot{X}(t)=\begin{pmatrix} \dot{X}_{L_1}(t) \\ \vdots \\ \dot{X}_{L_M}(t) \end{pmatrix} &= \begin{pmatrix} f_{L_1}^{Y(t)}(X(t),t|\boldsymbol{\theta}_{L_1}) \\ \vdots \\ f_{L_M}^{Y(t)}(X(t),t|\boldsymbol{\theta}_{L_M}) \end{pmatrix} \\ &= f_L^{Y(t)}\left(X(t),t|\boldsymbol{\theta}_L=\bigcup_{m=1}^{M}\boldsymbol{\theta}_{L_m}\right) \end{aligned} \tag{3}$$

Let $\mathcal{F}$ denote the system failure state set, which depends on the structure of the system: then, the system reliability at mission time T_{miss} can be obtained as follows:

$$R(T_{miss})=P[Z(s)\notin\mathcal{F},\forall s\leq T_{miss}] \tag{4}$$

The system failure state set is dependent on system structure. To determine this set, reliability analysis tools such as fault tree [15] can be used to identify the combination of primary failure events leading to system failure.

3 System Reliability Estimation Method

Analytically solving the PDMP is a difficult task due to the complex behavior of the system [22], which contains stochastic properties in the components modeled by MSMs and the time-dependent evolutions of the components modeled by PBMs. On the other hand, MC simulation methods are suited for the reliability estimation of the system.

Refer to the system presented in Sect. 2.2. Let $Z_k = Z(T_k) = \begin{pmatrix} X(T_k) \\ Y(T_k) \end{pmatrix} \in E, k \in \mathbb{N}$, where T_k denotes the time of the k-th transition of $Y(t)$ from the beginning. Then, $\{Z_k, T_k\}_{k \geq 0}$ is a Markov renewal process defined on the space $E \times \mathbb{R}^+$ [4], which is characterized as follows:

$$\begin{aligned} & P[Z_{k+1} \in B, T_{k+1} \in [T_k, T_k + \Delta t] | Z_k = i, \boldsymbol{\theta} = \boldsymbol{\theta}_K \cup \boldsymbol{\theta}_L] \\ & = \iint\limits_{B*[0,\Delta t]} N(i, dz, ds|\boldsymbol{\theta}), \forall k \geq 0, \Delta t \geq 0, i \in E, B \in \varepsilon \end{aligned} \tag{5}$$

where ε is a σ-algebra of E and $N(i, dz, ds|\boldsymbol{\theta})$ is a semi-Markov kernel on E, which verifies that $\int\limits_{E*[0,\Delta t]} N(i, dz, ds|\boldsymbol{\theta}) \leq 1, \forall \Delta t \geq 0, i \in E$. It can be further developed as:

$$N(i, dz, ds|\boldsymbol{\theta}) = dF_i(s|\boldsymbol{\theta})\beta(i, dz|s, \boldsymbol{\theta}) \tag{6}$$

where

$$dF_i(s|\boldsymbol{\theta}) \tag{7}$$

is the probability density function of $T_{k+1} - T_k$ given $Z_k = i$ and

$$\beta(i, dz|s, \boldsymbol{\theta}) \tag{8}$$

is the conditional probability distribution of state Z_{k+1} starting from $Z_k = i$ given $T_{k+1} - T_k = s$.

The simulation procedure consists of sampling the transition time from Eq. (7) and the arrival state from Eq. (8) for $Y(t)$, then, calculating $X(t)$ within the transition times, by using the physics equation (3) until the time of system evolution reaches a certain mission time T_{miss} or the system enters the failure space $\mathcal{F}$.

To calculate the system reliability, the procedure of the MC simulation is presented as follows:

Set N_{max} (the maximum number of replications) and $k = 0$ (index of replication)
Set $k' = 0$ (number of trials that end in the failure state)
While $k < N_{max}$
Initialize the system by setting $\boldsymbol{Z}' = \begin{pmatrix} \boldsymbol{X}(0) \\ \boldsymbol{Y}(0) \end{pmatrix}$ (initial state) and the time $T = 0$ (initial system time)
Set $t' = 0$ (state holding time)
While $T < T_{miss}$
Sample a t' by using the probability density function (11.7)
Sample an arrival state $\boldsymbol{Y}'$ for stochastic process $\boldsymbol{Y}(t)$ from all the possible states by using the conditional probability distribution (11.8)
Set $T = T + t'$
Calculate $\boldsymbol{X}(T)$ by using the physics eq. (11.3)
Set $\boldsymbol{Z}' = \begin{pmatrix} \boldsymbol{X}(T) \\ \boldsymbol{Y}' \end{pmatrix}$
If $T \le T_{miss}$
If $\boldsymbol{Z}' \in \boldsymbol{\mathcal{F}}$
Set $k' = k' + 1$
Break
End if
Else (when $T > T_{miss}$)
Calculate $\boldsymbol{Z}(T_{miss})$
If $\boldsymbol{Z}(T_{miss}) \in \boldsymbol{\mathcal{F}}$
Set $k' = k' + 1$
Break
End if
End if
End While
Set $k = k + 1$
End While

The estimated probability of occurrence of one path at time T_{miss} can be obtained by

$$\widehat{R}(T_{miss}) = 1 - k'/N_{max} \tag{9}$$

with the sample variance [16] as follows:

$$var_{\widehat{P(T_{miss})}} = \widehat{R}(T_{miss})\left(1 - \widehat{R}(T_{miss})\right)/(N_{max} - 1) \tag{10}$$

4 Case Study

The case study refers to one subsystem of the RHRS of a NPP. The system consists of a centrifugal pump and a pneumatic valve in series. Given the series configuration, the failure of anyone of the two components can lead the subsystem to failure. Dependence in the degradation processes of the two components has been indicated by the experts: the pump vibrates due to degradation [32] which, in turn, leads the valve to vibrate, aggravating its own degradation processes [24].

The pump is modeled by a MSM, modified from the one originally supplied by EDF upon discussion with the experts. It is a continuous-time homogeneous Markov chain as shown in Fig. 3.

$S_p = \{0, 1, 2, 3\}$ denotes its degradation states set, where 3 is the perfect functioning state and 0 is the complete failure state. The parameters λ_{32}, λ_{21} and λ_{10} are the transition rates between the degradation states. Due to degradation, the pump vibrates when it reaches the degradation states 2 and 1. The intensity of the vibration of the pump on states 2 and 1 is evaluated as by the experts 'smooth' and 'rough', respectively. We assume that $\lambda_{32} = \lambda_{21} = \lambda_{10} = 3e-3/s$.

The simplified scheme of the pneumatic valve is shown in Fig. 4. It is a normally closed, gas-actuated valve with a linear cylinder actuator.

By regulating the pressure of the pneumatic ports to fill or evacuate the top and bottom chambers, the position of the piston can be controlled. A return spring is linked with the piston to ensure the closure of the valve, when pressure is lost. The

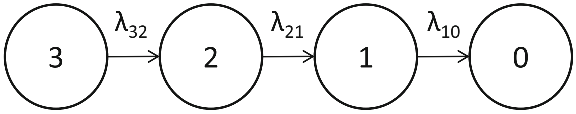


Fig. 3 Degradation process of the pump

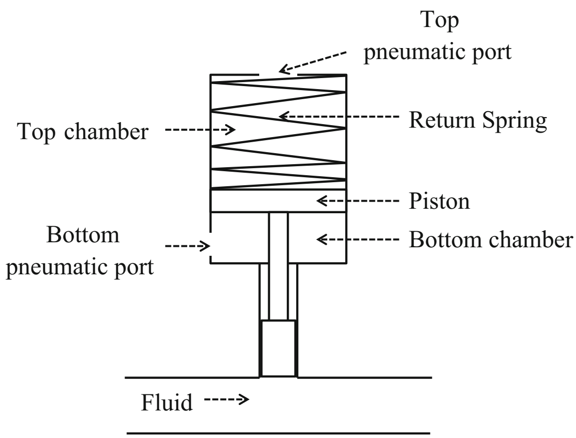


Fig. 4 Simplified scheme of the pneumatic valve [9]

external leak at the actuator connections to the bottom pneumatic port due to corrosion and other environmental factors is chosen as the degradation mechanism of the valve, which is much more significant than the other degradation mechanisms according to the results shown in [6].

Let $D_b(t)$ denote the area of the leak hole at the bottom pneumatic port at time t, the development of the leak size is described by:

$$\dot{D_b}(t) = \omega_b \left(1 + \beta_{Y_p(t)}\right) \tag{11}$$

where $\omega_b = 1e - 8m^2/s$ is the original wear coefficient and where $\beta_{Y_p(t)}$ is the relative increment of the developing rate of the external leak at the bottom pneumatic port, caused by the vibration of the pump at degradation state '2' or '1'. We assume that $\beta_2 = 10\%$ and $\beta_1 = 20\%$.

The leak will lead the valve to be more difficult to open but easier to close than in case without leak. The threshold of the area of the leak hole $D_b^* = 1.06\mathrm{e} - 5m^2$ is defined as the value above which $(D_b(t) > D_b^*)$ the valve cannot reach the fully open position from the fully closed position, within the 15 s time limit, after an opening command is executed.

The degradation processes affecting the system are modeled by PDMP as follows:

$$Z(t) = \begin{pmatrix} D_b(t) \\ Y_p(t) \end{pmatrix} \in \mathbb{R}^+ \times S_p \tag{12}$$

where $Y_p(t)$ denotes the degradation state of the pump at time t and $D_b(t)$ denotes the area of the leak hole at the bottom pneumatic port of the valve at time t. The space of the failure states of $Z(t)$ is $\mathcal{F} = [0, +\infty) \times \{'0'\} \cup \left[D_b^*, +\infty\right) \times \{1, 2, 3\}$.

The initial state of the system is assumed as follows:

$$Z_0 = \begin{pmatrix} D_b(0) \\ Y_p(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \tag{13}$$

which means that the two components are both in their perfect states. The initial probability distribution of the processes $\left(D_b(t), Y_p(t)\right)_{t \geq 0}$, $p_0(dz|\boldsymbol{\theta})$, hence, equals to $\delta_{Z_0}(dz)$, where δ is the Dirac delta function.

We perform MC simulation for the estimation of the system reliability over a time horizon of $T_{miss} = 1000$ s. The results of 10^6 trials are shown in Fig. 5. We can see from the Figure that the system reliability decreases more rapidly after around 885 s, because at that time the valve could fail, corresponding to the situation when the pump jumps to the state '1' very quickly and stays there until the valve fails.

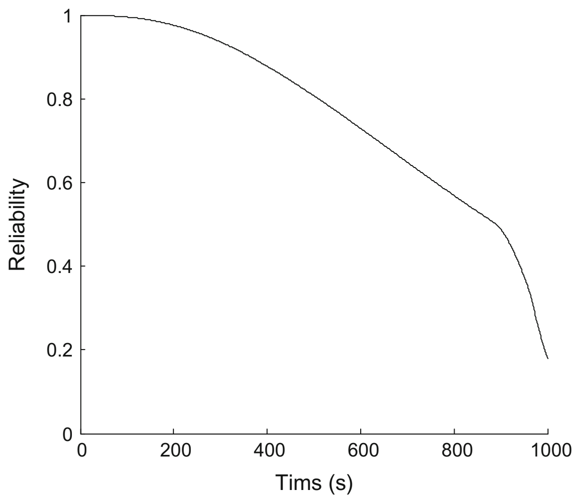


Fig. 5 Estimated system reliability

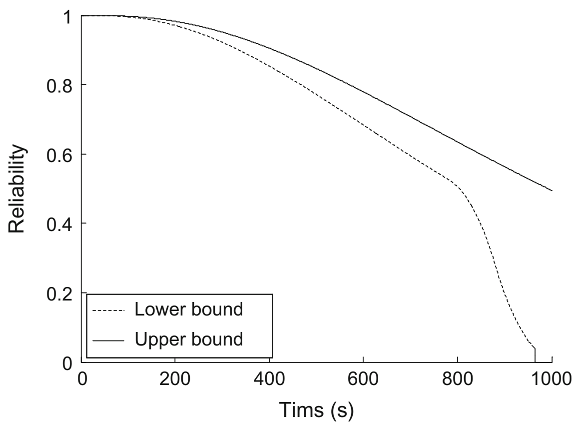


Fig. 6 Estimated system reliability in consideration of uncertainty

We further consider a relative uncertainty of $\pm10\%$ of the original parameters values. In this case study, higher parameters values lead to rapider degradation development and lower system reliability. The results of 10^6 trials are shown in Fig. 6. The lower bound of the system reliability decreases more sharply after around 790 s. It is seen that the system fails after around 964 s, because at that time the valve is completely failed. The upper bound of the system reliability does not experience a rapid decrease because the valve is mostly functioning over the time horizon.

5 Conclusion

We have illustrated a PDMP modeling approach for modeling multiple, dependent, competing degradation processes. The significance of the proposed method lies in its capability to describe the degradation dependence. A MC simulation algorithm for the system reliability assessment has been designed and an example from a real industrial system has been used to illustrate the capabilities of the modeling and simulation framework.

Limitation of the MC simulation lies in the computational burden. As future work, we plan to study acceleration techniques to improve computation efficiency, thus, enabling to extend the applications to systems of larger sizes.

References

1. Black M, Brint A, Brailsford J (2005) A semi-Markov approach for modelling asset deterioration. J Oper Res Soc 56(11):1241–1249
2. Chen N, Ye ZS, Xiang Y, Zhang L (2015) Condition-based maintenance using the inverse Gaussian degradation model. Eur J Oper Res 243(1):190–199
3. Chookah M, Nuhi M, Modarres M (2011) A probabilistic physics-of-failure model for prognostic health management of structures subject to pitting and corrosion-fatigue. Reliab Eng Syst Saf 96(12):1601–1610
4. Cocozza-Thivent C (2011) Processus de renouvellement markovien, Processus de Markov déterministes par morceaux. In: Online book. http://perso-math.univ-mlv.fr/users/cocozza.christiane/recherche-pageperso/PresentationRMetPDMP.html
5. Daigle M, Goebel K (2011) Multiple damage progression paths in model-based prognostics. In: Aerospace conference, 2011. IEEE, pp 1–10
6. Daigle M, Goebel K (2011) A model-based prognostics approach applied to pneumatic valves. Int J Prognostics Health Manage 2:84–99
7. Davis M (1984) Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. J R Stat Soc Ser B (Methodol) 46(3):353–388
8. Davis M (1993) Markov models & optimization vol 49. CRC Press
9. Gebraeel N, Elwany A, Pan J (2009) Residual life predictions in the absence of prior degradation knowledge. IEEE Trans Reliab 58(1):106–117
10. Giorgio M, Guida M, Pulcini G (2011) An age-and state-dependent Markov model for degradation processes. IIE Trans 43(9):621–632
11. Jacobsen M (2006) Point process theory and applications: marked point and piecewise deterministic processes. Birkhäuser, Boston
12. Jiang L, Feng Q, Coit D (2012) Reliability and maintenance modeling for dependent competing failure processes with shifting failure thresholds. IEEE Trans Reliab 61(4):932–948
13. Keedy E, Feng Q (2012) A physics-of-failure based reliability and maintenance modeling framework for stent deployment and operation. Reliab Eng Syst Saf 103:94–101
14. Lawless J, Crowder M (2001) 4) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10(3):213–227
15. Lee WS, Grosh D, Tillman F, Lie C (1985) Fault tree analysis, methods, and applications: a review. IEEE Trans Reliab 34(3):194–203
16. Lewis E, Böhm F (1984) Monte Carlo simulation of Markov unreliability models. Nucl Eng Des 77(1):49–62

17. Lin YH, Li Y, Zio E (2014) Dynamic reliability models for multiple dependent competing degradation processes. In: Proceedings of ESREL 2014
18. Li W, Pham H (2005) Reliability modeling of multi-state degraded systems with multi-competing failures and random shocks. IEEE Trans Reliab 54(2):297–303
19. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
20. Lorton A, Fouladirad M, Grall A (2013) A methodology for probabilistic model-based prognosis. Eur J Oper Res 225:443–454
21. Lu C, Meeker W (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
22. Marseguerra M, Zio E (1996) Monte Carlo approach to PSA for dynamic process systems. Reliab Eng Syst Saf 52:227–241
23. Moghaddass R, Zuo M (2014) Multistate degradation and supervised estimation methods for a condition-monitored device. IIE Trans 46:131–148
24. Moussou P, Cambier S, Lachene D, Longarini S, Paulhiac L, Villouvier V (2001) Vibration investigation of a French PWR power plant piping system caused by cavitating butterfly valves. ASME Publications-PVP 420:99–106
25. Peng H, Feng Q, Coit D (2010) Reliability and maintenance modeling for systems subject to multiple dependent competing failure processes. IIE Trans 43:12–22
26. Rasmekomen N, Parlikad A (2013) Maintenance optimization for asset systems with dependent performance degradation. IEEE Trans Reliab 62:362–367
27. Reggiani S, Poli S, Denison M, Gnani E, Gnudi A, Baccarani G, Pendharkar S, Wise R (2011) Physics-based analytical model for HCS degradation in STI-LDMOS transistors. IEEE Trans Electron Dev 58:3072–3080
28. Straub D (2009) Stochastic modeling of deterioration processes through dynamic Bayesian networks. J Eng Mech 135:1089–1099
29. Wang Y, Pham H (2012) Modeling the dependent competing risks with multiple degradation processes and random shock using time-varying copulas. IEEE Trans Reliab 61:13–22
30. Yang G (2002) Environmental-stress-screening using degradation measurements. IEEE Trans Reliab 51:288–293
31. Ye ZS, Xie M (2014) Stochastic modelling and analysis of degradation for highly reliable products. Appl Stoch Models Bus Ind 31:16–32
32. Zhang S, Hodkiewicz M, Ma L, Mathew J (2006) Machinery condition prognosis using multivariate analysis. In: Engineering Asset Management. Springer, pp 847–854
33. Zwillinger D (1998) Handbook of differential equations. Academic Press

Trade-Off Between Redundancy, Protection, and Imperfect False Targets in Defending Parallel Systems

Hui Xiao and Rui Peng

Abstract A substantial amount of research over the past decades has studied the reliability of different systems, but most of them are restricted to systems with only internal failures. In practice, systems may fail due to unintentional impacts or intentional attacks. In this chapter, we first provide a comprehensive review of the research on improving system reliability. The survey shows that, for systems subject to intentional attacks, providing redundant system elements, protecting genuine elements, and deploying false targets are the three important measures to increase the system survivability. The trade-off between protecting genuine elements and deployment of imperfect false targets has been studied before, however, subject to a fixed number of genuine elements in the system. This chapter studies the trade-off between building redundant genuine elements, protection of genuine elements and deploying imperfect false targets in the defense of a capacitated parallel system. Numerical examples are carried out to illustrate the applications.

Keywords Vulnerability · Attack · Defence · False target · Protection · Parallel system

H. Xiao (✉)
Department of Management Science, School of Statistics, Southwestern University of Finance and Economics, Chengdu, China
e-mail: msxh@swufe.edu.cn

R. Peng
Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing, China
e-mail: pengrui1988@ustc.edu.cn

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*, Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_12

1 Introduction

In the past two decades, reliability has received substantial attention in both industry and research [15, 17]. In order to model the complex modern systems more accurately, the traditional binary reliability theory has been extended to analyze multi-state systems that can perform their intended work at different intensities [34]. A variety of multi-state systems such as series-parallel systems [36], linear sliding window systems [18], linearly connectively connected systems [40] and multi-state network systems [48] have been studied in recent years. These reliability models have been successfully applied to analyze many real industrial systems such as power supply systems [33], maritime transportation systems [7], and high performance computing systems [50].

Most of the abovementioned research is restricted to systems with only internal failures. In practice, systems may fail due to external impacts. Examples of unintentional impacts include shocks, natural disasters, and unintentional human errors [41, 47]. Providing redundant system elements and elements protection are the two essential measures against unintentional impacts. Besides unintentional impacts, intentional attackers may circumvent the protection and choose the fragile positions to attack in order to destroy the system [42, 51]. For systems subject to intentional attacks, the defender not only need to provide redundancy and protect the genuine elements, but also can deploy some false targets in order to distract the attacker. In the literature, the false targets can be assumed to be perfect or imperfect. If the attacker cannot distinguish the false targets from the genuine elements, the false target is said to be perfect. If the false target can be detected by the attacker, the false target is imperfect. This chapter aims to analyze the trade-off among providing redundancy, investing in protection and deploying imperfect false targets. The work is closely related to some of the existing works, but differs in different aspects. Levitin and Hausken [21–23] have studied the different measures such as providing redundancy and protection, and deployment of false targets to defend parallel systems under attacks. In these works, the false targets are assumed to be perfect. In practice, false targets are usually detectable, i.e., imperfect. Thus, we study the scenario when false targets are assumed to be imperfect in this chapter. Peng et al. [38] considered the optimal number of false targets in both a parallel system and a series system assuming that the false targets are imperfect, but this paper does not consider the strategy of providing redundancy to improve the system reliability. Previous research has shown that providing redundancy is an important and useful measure to improve system reliability, therefore, this chapter proposes a model to study the optimal resource allocation in providing redundant system elements, protecting genuine system elements, and deploying false targets in the defense of capacitated parallel systems. The organization of this chapter is as follows. Section 2 conducts a comprehensive literature review on the relevant topics. Section 3 studies the defense of capacitated parallel systems with redundancy, protection, and imperfect false targets. The chapter is concluded in Sect. 4.

2 Literature Review

In the literature, approaches to minimize the system unavailability caused by internal failures include providing redundancy and finding the optimal maintenance policy [3, 44]. For example, Yeh and Fiondella [49] studied the optimal redundancy allocation for a multi-state computer network system. A correlated binomial distribution is applied to characterize the state distribution of the edges, which can exhibit multiple states. This redundancy allocation problem was solved using simulated annealing with an illustration in four practical networks. Ardakan and Hamadani [1] considered the redundancy allocation problem in a series-parallel system, where active and cold-standby strategies are simultaneously used in one subsystem. Utilizing the genetic algorithm, it determined the optimal component type, redundancy level, number of active units and cold-standby units jointly for each subsystem with the objective of maximizing the system reliability. Levitin and Amari [20] presented an approximation algorithm based on the universal generating function technique to evaluate the distribution of the time to failure for a k-out-of-n system with shared redundant elements. Besides providing redundancy, maintenance is also a frequently used measure to reduce the system internal failures. Peng et al. [40] studied the optimal preventive maintenance policy for linearly consecutively connected systems with the objective of minimizing the total maintenance cost while meeting a pre-specified system availability requirement. Xiao and Peng [46] studied the optimal element allocation and replacement interval in a series-parallel system with common bus performance sharing. Lisnianski et al. [35] considered a reliability importance evaluation for components in an aging multi-state system under minimal repair.

Besides the preventive replacement and minimal repair used in the abovementioned research, imperfect maintenance is also well studied in the literature [43]. For example, in order to model the wind turbine system in a wind farm, Ding and Tian [5] proposed an opportunistic maintenance policy that introduces different imperfect maintenance thresholds for failure turbines and working turbines. The proposed approach was shown to be effective in modeling the practical system in order to minimize the maintenance cost. Zhao et al. [52] utilized the cumulative processes theory and found the optimal imperfect maintenance policy by minimizing the expected cost rate for a used system. Pandey et al. [37] developed a selected maintenance strategy for a multi-state system with multi-state components. Different types of maintenance options such as replacement, do-nothing option, and imperfect repair were chosen to ensure that maximum system reliability could be achieved during the next mission.

Besides internal failures, systems may also fail due to external impacts. In general, the external impacts can be classified into unintentional impacts and intentional attacks. Natural disasters such as earthquakes and tsunamis are typical examples of unintentional impacts, while terrorism, warfare, intrusion, and human disruption are examples of intentional attacks [6]. In the case of external impacts, providing redundant systems elements and investing in protection are two effective measures to improve system survivability, and the two measures have been well-studied in the literature. For example, Kunreuther and Heal [16] characterized the Nash equilibrium for an interdependent security problem to analyze the incentive of firms investing in protection against intentional attacks. Bier et al. [2] proposed general models for determining the optimal defense resource allocation assuming that the attacker will maximize either the expected damage or the success probability. Hausken [9] considered the security investment of several firms in cyber wars with external intruders. Zhuang and Bier [53] considered the defender resource allocation for countering terrorism and natural disasters. Hausken [10, 11] studied the strategic defense and attack for series and parallel systems consisting of independent components against intentional attacks. Levitin [19] considered the optimal trade-off between protection and redundancy in a homogenous parallel system subject to intentional attacks. When overarching protection is provided, the attacker can only destroy the components when the outside protection layer is penetrated. Haphuriwat and Bier [8] developed a model to allocate the resource optimally between target hardening and overarching protection and analyzed the effects that influence the trade-off between target hardening and overarching protection. Hausken [12] considered the individual and overarching protection versus attack for both defenders and attackers in both series and parallel systems. The model was extended to analyze the individual and overarching protection versus attack of assets in a simultaneous game and a two-period game [13]. Levitin et al. [32] further studied the viable number of individual protection versus overarching against strategic attack. Peng et al. [41] studied the optimal individual protection versus overarching protection versus maintenance for a parallel system subject to both internal failures and unintentional impacts. Deck et al. [4] proposed a model to analyze the scenario when the contest happens between an attacker and multiple defenders, and concluded that alliance of the defenders can reduce the defense spending and result in higher profit for defenders and attackers.

Besides providing redundant system elements and investing in protection, some researchers studied the deployment of false targets in defense of systems against intentional attacks. A historical example of using false targets (decoys) can be found in WWII and the Operation Desert Storm in 1990–1991. The objective of deploying false targets is to distract the attacker and dissipate the attack resource over greater number of targets. The defense measure of deploying false targets is most effective when the attacker cannot distinguish the false targets from genuine elements. Usually, false targets are much cheaper than genuine elements, but they are not costless. Deploying more false targets results in less resource allocated to provide redundancy and protect the genuine elements. In the literature, the study on the deploying false targets against intentional attacks has been studied in a variety of ways. Firstly, based on the assumption that the false targets can be destroyed with much less effort than the genuine elements, the attacker can distribute its attack resource in two sequential attacks so that the false targets can be eliminated as many as possible in the first attack [26, 29–31]. Secondly, the defense resource is usually distributed to provide redundancy, deploy false element and protect the genuine elements. In some scenarios, the defender can also distribute its resource to strike preventively against the attacker [27, 28]. Thirdly, the quality of the false targets can be different in different scenarios. Some researchers assume that the false targets are perfect, i.e., the attacker cannot distinguish the false targets from the genuine elements [14, 21, 23, 24], while others consider the false targets as imperfect, i.e., the false target can be distinguished from the genuine elements by the attacker with certain probability [38, 39]. Lastly, some recent research has also considered the scenario that the attacker allocates part of the resource into intelligence activities to detect the false targets, and the defender allocates part of the resource into counter-intelligence activities [22, 25, 42].

In this chapter, the following notations will be used.

N The number of genuine elements in the system
H The number of false targets in the system
k The number of false targets that are detected
Q_k The number of objects the attacker tries to attack give k false targets are detected
P_k The probability that k false targets are detected
x, y The cost of a false target and a genuine element respectively
r, R The total resource of the defender and attacker respectively
a, A The unit cost of defending and attacking respectively
d The detection probability of a false target
F The system demand
g The performance of a genuine element.

3 Defense of Parallel Systems with Redundancy, Protection, and Imperfect False Targets

Consider a parallel system that consists of N identical genuine elements. The defender can deploy imperfect false targets, provide redundancy, and protect genuine elements to minimize the expected damage that may be caused by the intentional attacks. The defender builds the system and distributes its defense resource first. The attacker takes it as given when it chooses its attack strategy. Therefore, it can be modeled as a two-period min-max game of perfect information where the defender moves in the first period, and the attacker moves in the second period. The defender decides how many false targets to deploy and how many redundant genuine elements to provide in order to minimize the expected damage caused by the attacker assuming that the attacker will always use the most harmful attack strategy.

The total resource of the attacker is a fixed value R. The unit cost of attacking an object is a constant denoted by A. We assume that attacker distributes the resource evenly among all attacked objects. All elements are assumed to be mutually independent. A single attack cannot destroy more than one object. In order to improve the system reliability, the defender can perform three different measures using a fixed value of resource r. The defender can provide redundancy for the parallel system, and the cost of a genuine element is y. Therefore, the maximal number of genuine elements is $\lfloor r/y \rfloor$. The defender can deploy false targets at the cost of x each, and $x \ll y$. Similarly, the maximal number of false targets to deploy is $\lfloor (r - Ny)/x \rfloor$, where N is the number of genuine elements in this parallel system. The false target is imperfect, i.e., each false target can be detected by the attacker independently with probability d. The probability that k false targets are detected is denoted by p_k:

$$p_k = \binom{H}{k} d^k (1-d)^{H-k}, \tag{1}$$

where H is the total number of the false targets deployed.

The cost of a genuine element is much more than the cost of a false target. Additionally, the defender can protect the genuine elements using the remaining resource. The unit cost of protecting a genuine element is a, and the protection effort on each genuine element is assumed to be evenly distributed.

Given that N genuine elements and H false targets are placed in this system, the protection effort on each genuine element can be expressed as follows:

$$t = \frac{r - Ny - Hx}{Na}. \tag{2}$$

Suppose that $k\,(0 \leq k \leq H)$ false targets are detected. The attacker chooses Q_k out of $N + H - k$ objects to attack, where $1 \leq Q_k \leq N + H - k$. The attack effort on each attacked object is $T = R/(Q_k A)$. Therefore, the destruction probability of each element can be written using the contest function suggested by Tullock [45]:

$$v = \frac{T^m}{T^m + t^m} = \frac{R^m}{R^m + \frac{(r-Ny-Hx)^m (Q_k A)^m}{N^m a^m}} = \frac{R^m}{R^m + Q_k^m \varepsilon^m \frac{(r-Ny-Hx)^m}{N^m}}, \tag{3}$$

where $\varepsilon = A/a$, T and t are the efforts allocated to a single object by the attacker and the defender respectively. m is the parameter describing the intensity of the contest. If the no protection effort is provided, the element is destroyed with probability 1 when it is attacked since $v = 1$ if $t = 0$. The destruction probability will be always 0.5 if the intensity parameter m is zero. If $m \to \infty$, v is a step function where "winner takes all".

Given k and Q_k, the number of attacked genuine elements can vary from $\max(0, Q_k - H + k)$ to $\min(N, Q_k)$, where $\max(0, Q_k - H + k)$ refers to the scenario when all the false targets are attacked while $\min(N, Q_k)$ refers to the case when no false target is attacked. Let $\varphi(Q_k, i)$ denote the probability that among the Q_k attacked objects i of them are genuine elements. $\varphi(Q_k, i)$ can be derived using hyper-geometric distribution:

$$\varphi(Q_k, i) = \frac{\binom{N}{i}\binom{H-k}{Q_k - i}}{\binom{N+H-k}{Q_k}}. \tag{4}$$

Let $\theta(i,j)$ denote the probability that j elements out of i attacked genuine elements are destroyed. $\theta(i,j)$ can be determined based on the destruction probability function as follows:

$$\theta(i,j) = \binom{i}{j} v^j (1-v)^{i-j}. \tag{5}$$

In the parallel system, it is assumed that all genuine elements have the same functionality with performance rate g. The system demand is F. The system fails if at least $\lfloor N - F/g \rfloor + 1$ elements are destroyed by the attacker. When the system fails, the expected damage is proportional to the loss of demand probability. In this case, the risk can be obtained as follows:

$$D(Q_k, N, H) = F \cdot \sum_{i=\max(\lfloor N-F/g \rfloor + 1, Q_k - H + k)}^{\min(N, Q_k)} \left(\varphi(Q_k, i) \cdot \sum_{j=\lfloor N-F/g \rfloor + 1}^{i} \theta(i,j) \right). \tag{6}$$

The attacker chooses the most harmful strategy Q_k^* which maximizes $D(Q_k, N, H)$. The total risk over all possible values of k can be obtained as follows:

$$D(N,H) = \sum_{k=0}^{H} p_k \cdot D\left(Q_k^*, N, H\right) = \sum_{k=0}^{H} \left(\binom{H}{k} \cdot d_k \cdot (1-d_k) \cdot D\left(Q_k^*, N, H\right) \right). \tag{7}$$

In this two-period game, the defender moves first. Given the defender's strategy of H and N, the attacker chooses the best Q_k to maximize the damage $D(Q_k, N, H)$ when k false targets are detected. The defender knows that the attacker will maximize its damage for any value of H and N, and chooses the optimal H and N such that the expected risk $D(N,H)$ can be minimized.

For example, consider a parallel system that is made up by genuine elements to satisfy the demand of $F = 4$. The performance of each genuine element is $g = 1$. In order to protect the system, the defender can deploy false targets at the cost of $x = 0.03$ each and provide redundancy at the cost of $y = 0.1$ for a genuine element. The cost of the attacker's effort unit and defender's effort unit is assumed to be equal, i.e., $\varepsilon = 1$. The total resource of the attacker and defender are assumed to be $R = 1$ and $r = 1$ respectively.

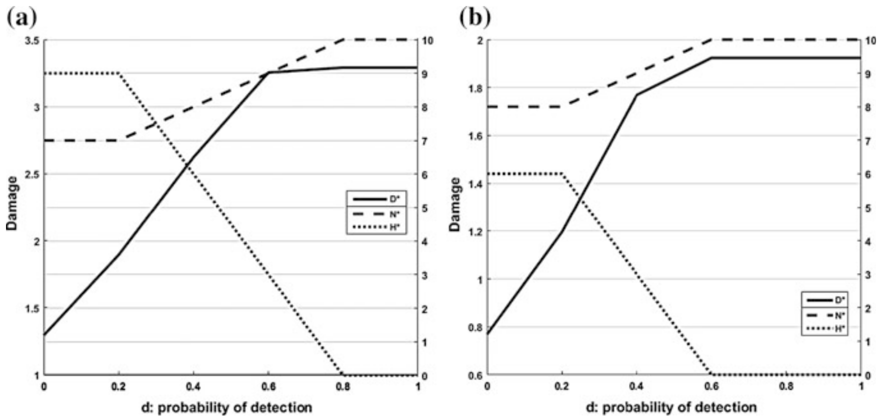


Fig. 1 Optimal number of false targets and the corresponding expected risk as a function of detection probability given intensity parameter m: (**a**) $m = 10$, (**b**) $m = 6$

Figure 1 shows the optimal redundancy, optimal number of false targets and the corresponding expected damage of this defense-attack game as a function of detection probabilities when the contest intensity parameter m is equal to 10 and 6 respectively. Figure 1 indicates that the expected damage and optimal redundancy increase when the probability of detection increases, but the optimal number of false targets decreases with increasing probability of detection. The numerical results show that it is worthy deploying false targets than providing redundancy when the detection probability is very low. When the probability of detection increases, the benefit of deploying false targets will be reduced. Furthermore, the redundancy reduces the need of false targets.

Figure 2 shows how the amount of resource owned by the attacker affects the optimal values of N, H and D. The results indicate that the expected damage increases due to the increase of resource owned by the attacker. It is also interesting to note that the optimal redundancy decreases and the optimal number of false targets increases when the attacker's resource increases. When the resource of the attacker increases, the destruction probability increases given the fixed contest intensity parameter. Therefore, the attacked GEs are more likely to be destroyed. This is why it is better to deploy more false targets so that the attack effect on each GE will be reduced.

Figure 3 shows that the expected damage is reduced when the resource of the defender increases. Naturally, increasing the resource of defense reduces the probability of destruction, therefore, the expected damage can be reduced. In Fig. 3, the optimal redundancy increases and the optimal number of false targets decreases in general when the resource of defense is increasing. It shows that increasing redundancy is better than increasing the false targets in order to minimize the expected damage if the resource of defense is unlimited.

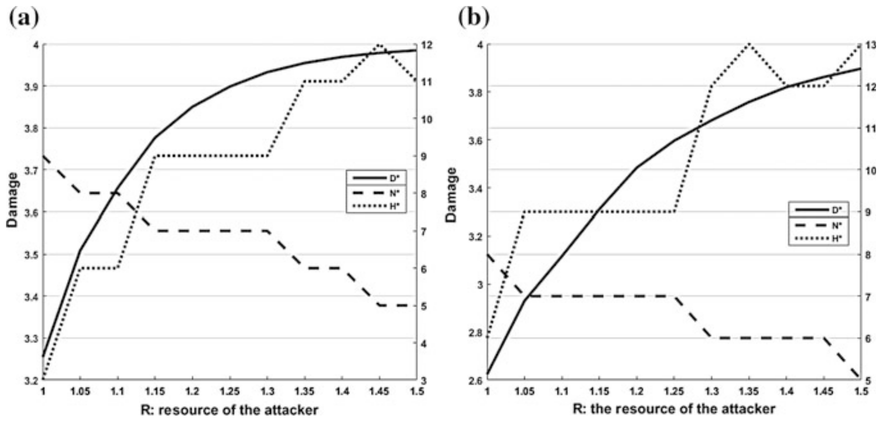


Fig. 2 Optimal number of false targets and the corresponding expected risk as a function of attacker's resource given detection probability d: (**a**) $d = 0.6$, (**b**) $d = 0.4$

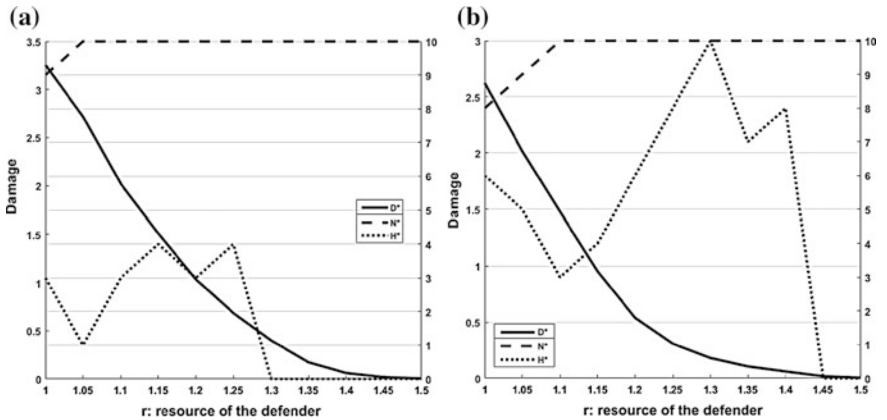


Fig. 3 Optimal number of false targets and the corresponding expected risk as a function of defender's resource when detection probability d: (**a**) $d = 0.6$, (**b**) $d = 0.4$

Figures 4 and 5 show the optimal redundancy, optimal number of false targets and the corresponding expected damage under different cost of a genuine element and different cost of a false target. Both figures indicate that the expected damage increases with increasing cost of a genuine element or a false target. When the cost of a genuine element increases, the optimal redundancy decreases and the optimal number of false targets increases. On the other hand, the optimal redundancy increases and the optimal number of false targets decreases when the cost of a false target increases.

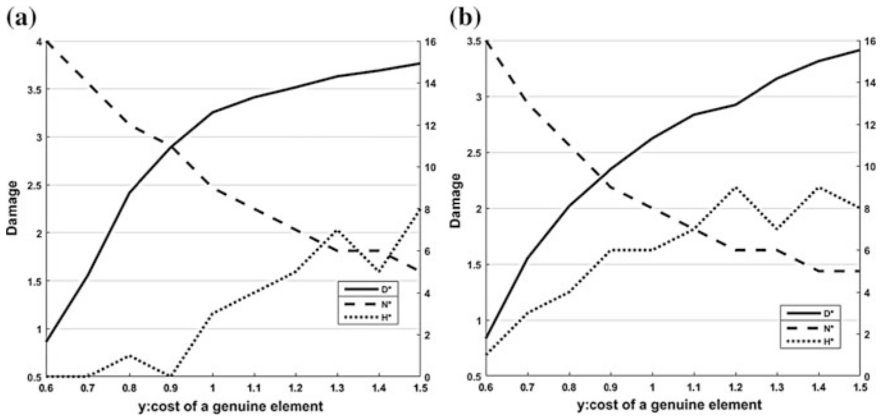


Fig. 4 Optimal number of false targets and the corresponding expected risk as a function of cost of a genuine element when detection probability d: (**a**) $d = 0.6$, (**b**) $d = 0.4$

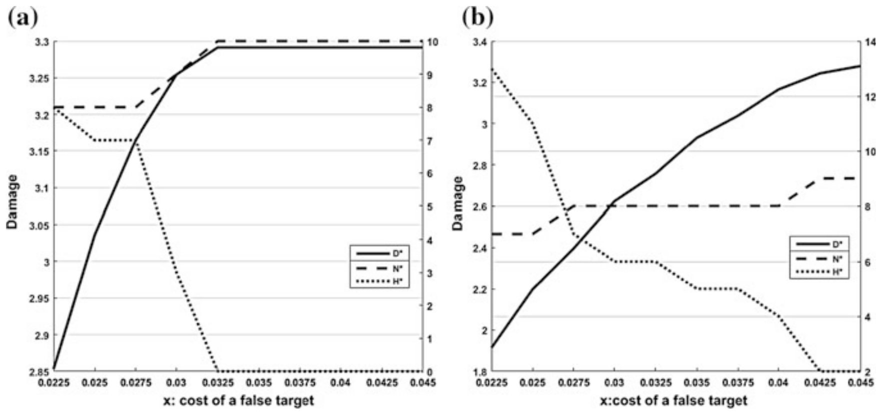


Fig. 5 Optimal number of false targets and the corresponding expected risk as a function of cost of a false target when detection probability d: (**a**) $d = 0.6$, (**b**) $d = 0.4$

4 Conclusions and Future Research

This chapter considers protecting a capacitated parallel system under intentional attacks. The defense resource can be allocated to protect the genuine element, deploy imperfect targets, and provide redundancy in order to minimize the expected damage caused by intentional attacks. Given the protection strategy of the defender, the attacker will choose the optimal number of elements to attack such that the damage can be maximized. Therefore, the defender must take consideration of the attacker's decision to optimally allocate the defense resource to minimize the total expected damage. To illustrate the attack-defense model, several numerical experiments are conducted to analyze the effect of detection probability, the amount of resource owned by the defender and the attacker, and the cost of a genuine element and a false target. The numerical experiments indicate that the optimal number of false targets decreases when the probability of detection increases. When the detection probability become larger, it is better to allocate more resources to provide redundancy rather than deploying false targets. When the resource of the attacker becomes large, it is more worthy deploying more false targets than providing redundancy. However, more redundancy should be provided rather than deploying false targets if the resource of the defender is large. Besides, as consistent as the theoretic argument, higher cost of a genuine element results in a smaller optimal value of the redundancy, while higher cost of a false targets leads to a smaller optimal number of false targets.

This chapter uses a two-period dynamic game of perfect information to model the contest between the attacker and the defender assuming each party has full information about the other. In future, the model can be extended to consider the scenario when full information is not available. Besides, it is also important to consider the scenario when preventive strikes may be used by the defender. In this case, the problem becomes more complicated since the defender will consider the resource allocation for preventive strikes, and the attacker may consider using part of the resource to defend the preventive strike. Furthermore, it is also interesting to analyze the contest model between one defender and multiple attackers in future.

Acknowledgements The research is supported by the NSFC under grant numbers 71601158, 71671016 and 71231001, and by the Fundamental Research Funds for the Central Universities of China and FRF-BR-15-001B.

References

1. Ardakan M, Hamadani A (2014) Reliability optimization of series–parallel systems with mixed redundancy strategy in subsystems. Reliab Eng Syst Saf 130:132–139
2. Bier V, Nagaraj A, Abhichandani V (2005) Protection of simple series and parallel systems with components of different values. Reliab Eng Syst Saf 87:315–323
3. Chang C (2014) Optimum preventive maintenance policies for systems subject to random working times, replacement, and minimal repair. Comput Ind Eng 67:185–194
4. Deck C, Foster J, Song H (2015) Defense against an opportunistic challenger: theory and experiments. Eur J Oper Res 242:501–513
5. Ding F, Tian Z (2012) Opportunistic maintenance for wind farms considering multi-level imperfect maintenance thresholds. Renew Energy 45:175–182
6. Enders W, Sandler T (2006) The political economy of terrorism. Cambridge University Press, New York
7. Faghih-Roohi S, Xie M, Ng K, Yam R (2014) Dynamic availability assessment and optimal component design of multi-state weighted k-out-of-n systems. Reliab Eng Syst Saf 123:57–62
8. Haphuriwat N, Bier V (2011) Trade-offs between target hardening and overarching protection. Eur J Oper Res 213:320–328

9. Hausken K (2006) Income, interdependence, and substitution effects affecting incentives for security investment. J Acc Public Policy 25:629–665
10. Hausken K (2008) Strategic defense and attack for reliability systems. Reliab Eng Syst Saf 93:1740–1750
11. Hausken K (2008) Strategic defense and attack for series and parallel reliability systems. Eur J Oper Res 186:856–881
12. Hausken K (2013) Combined series and parallel systems subject to individual versus overarching defense and attack. Asia Pac J Oper Res 30
13. Hausken K (2014) Individual versus overarching protection and attack of assets. CEJOR 22:89–112
14. Hausken K, Levitin G (2009) Protection vs. FTs in series systems. Reliab Eng Syst Saf 94:973–981
15. Kapur K (1977) Reliability in engineering design. Wiley, New York
16. Kunreuther H, Heal G (2003) Interdependent security. J Risk Uncertainty 26:231–249
17. Kuo W, Zuo M (2002) Reliability engineering: theory and practice. Wiley, New York
18. Levitin G (2003) Linear multi-state sliding-window systems. IEEE Trans Reliab 52(2): 263–269
19. Levitin G (2007) Optimal defense strategy against intentional attacks. IEEE Trans Reliab 56:148–157
20. Levitin G, Amari S (2010) Approximation algorithm for evaluating time-to-failure distribution of k-out-of-n system with shared standby elements. Reliab Eng Syst Saf 95 (4):396–401
21. Levitin G, Hausken K (2009) False targets vs. redundancy in homogeneous parallel systems. Reliab Eng Syst Saf 94(2):588–595
22. Levitin G, Hausken K (2009) Intelligence and impact contests in systems with redundancy, false targets, and partial protection. Reliab Eng Syst Saf 94(12):1927–1941
23. Levitin G, Hausken K (2009) Redundancy vs. protection vs. false targets for systems under attack. IEEE Trans Reliab 58(1):58–68
24. Levitin G, Hausken K (2009) False targets efficiency in defense strategy. Eur J Oper Res 194 (1):155–162
25. Levitin G, Hausken K (2009) Intelligence and impact contests in systems with fake targets. Defense Secur Anal 25(2):157–173
26. Levitin G, Hausken K (2011) Is it wise to protect false targets. Reliab Eng Syst Saf 96 (12):1647–1656
27. Levitin G, Hausken K (2011) Preventive strike vs. false targets and protection in defense strategy. Reliab Eng Syst Saf 96(8):912–924
28. Levitin G, Hausken K (2012) Preventive strike vs. false targets in defense strategy. Int J Performability Eng 8(4):341–354
29. Levitin G, Hausken K (2013) is it wise to leave some false targets unprotected? Reliab Eng Syst Saf 112:176–186
30. Levitin G, Hausken K, Ben Haim H (2014) Defending systems against two sequential attacks. Mil Oper Res 19(1):19–35
31. Levitin G, Hausken K, Ben Haim H (2014) False targets in defending systems against two sequential attacks. Mil Oper Res 19(1):19–35
32. Levitin G, Hausken K, Dai Y (2014) Optimal defense with variable number of overarching and individual protections. Reliab Eng Syst Saf 123:81–90
33. Li Y, Zio E (2012) A multi-state model for the reliability assessment of a distributed generation system via universal generating function. Reliab Eng Syst Saf 106:28–36
34. Lisnianski A, Levitin G (2003) Multi-state system reliability: assessment, optimization and applications. World Scientific, Singapore
35. Lisnianski A, Frenkel I, Khvatskin L (2015) On Birnbaum importance assessment for aging multi-state system under minimal repair by using the L-z-transform method. Reliab Eng Syst Saf 142:258–266

36. Nourelfath M, Yalaoui F (2012) Integrated load distribution and production planning in series-parallel multi-state systems with failure rate depending on load. Reliab Eng Syst Saf 106:38–145
37. Pandey M, Zuo M, Moghaddass R (2013) Selective maintenance modeling for a multistate system with multistate components under imperfect maintenance. IIE Trans 45(11): 1221–1234
38. Peng R, Levitin G, Xie M, Ng S (2010) Defending simple series and parallel systems with imperfect false targets. Reliab Eng Syst Saf 95:679–688
39. Peng R, Levitin G, Xie M, Ng S (2011) Optimal defense of single object with imperfect false targets. J Oper Res Soc 62(1):134–141
40. Peng R, Xie M, Ng S, Levitin G (2012) Element maintenance and allocation for linear consecutively connected systems. IIE Trans 44(11):964–973
41. Peng R, Guo L, Levitin G, Mo H, Wang W (2014) Maintenance versus individual and overarching protections for parallel systems. Qual Technol Quant Manage 11(3):353–362
42. Peng R, Zhai Q, Levitin G (2016) Defending a single object against an attacker trying to detect a subset of false targets. Reliab Eng Syst Saf 149:137–147
43. Pham H, Wang H (1996) Imperfect maintenance. Eur J Oper Res 94(3):425–438
44. Taghipour S, Kassaei M (2015) Periodic inspection optimization of a k-out-of-n load-sharing system. IEEE Trans Reliab 64(3):1116–1127
45. Tullock G (1980) Efficient rent seeking. In Buchanan JM, Tollison RD, Tullock G (eds) Toward a theory of the rent-seeking society. Texas A&M University Press, College Station, Texas
46. Xiao H, Peng R (2014) Optimal allocation and maintenance of multi-state elements in series–parallel systems with common bus performance sharing. Comput Ind Eng 72:143–151
47. Xiao H, Shi D, Ding Y, Peng R (2016) Optimal loading and protection of multi-state systems considering performance sharing mechanism. Reliab Eng Syst Saf 149:88–95
48. Yeh W (2014) Multistate network reliability evaluation under the maintenance cost constraint. Int J Prod Econ 88(1):73–83
49. Yeh C, Fiondella L (2017) Optimal redundancy allocation to maximize multi-state computer network reliability subject to correlated failures. Reliab Eng Syst Saf 166:138–150
50. Yu H, Yang J, Mo H (2014) Reliability analysis of repairable multi-state system with common bus performance sharing. Reliab Eng Syst Saf 132:90–96
51. Zhai Q, Ye Z, Peng R, Wang W (2016) Defense and attack of performance-sharing common bus systems. Eur J Oper Res 256(3):962–975
52. Zhao X, Nakagawa T, Qian C (2012) Optimal imperfect preventive maintenance policies for a used system. Int J Syst Sci 43(9):1632–1641
53. Zhuang J, Bier V (2007) Balancing terrorism and natural disasters—defensive strategy with endogenous attacker effort. Oper Res 55(5):976–981

Optimal Testing Resources Allocation for Improving Reliability Assessment of Non-repairable Multi-state Systems

Yu Liu, Tao Jiang and Peng Lin

Abstract Due to limited reliability testing resources (e.g., budget, time, and manpower etc.), the reliability of a sophisticated system may not be able to accurately estimated by insufficient reliability testing data. The book chapter explores the reliability testing resources allocation problem for multi-state systems, so as to improve the accuracy of reliability estimation of an entire system. The Bayesian reliability assessment method is used to infer the unknown parameters of multi-state components by merging subjective information and continuous/discontinuous inspection data. The performance of each candidate testing resources allocation scheme is evaluated by the proposed uncertainty quantification metrics. By introducing the surrogate model, i.e., kriging model, the computational burden in seeking the optimal testing resources allocation scheme can be greatly reduced. The effectiveness and efficiency of the proposed method are exemplified via two illustrative case.

Keywords Multi-state system · Reliability testing resources allocation · Bayesian reliability assessment · Surrogate model

1 Introduction

Multi-state is one of the characteristics of advanced manufacturing systems and complex engineered systems [1, 2]. Both systems and components may manifest multiple states ranging from perfect working, through deterioration, to completely failed over time [1, 2]. Multi-state system (MSS) reliability modeling has, therefore,

Y. Liu (✉) · T. Jiang
School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, China
e-mail: yuliu@uestc.edu.cn

P. Lin
Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*, Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_13

received considerable attentions in recent years as it is able to characterize the complicated deteriorating process of systems in a finer fashion than that of traditional binary-state models [1–3]. In engineering practice, many systems can be regarded as multi-state systems, e.g., manufacturing systems, power generating systems, flow transmission systems, etc. Various tools, such as the stochastic models, the universal generating functions, the recursive algorithms, and the simulation-based methods etc., have been developed to assess reliability and performance of MSSs in a computationally efficient manner.

Nevertheless, all the reported studies on MSS reliability assessment are based on the critical premise that the transition intensities and/or the state distributions of components and systems are exactly known in advance. Very limited attention has been placed on the parameter inference, which is a preceding task before conducting reliability assessment and enhancement of MSSs. Lisnianski et al. [4] introduced the point estimation of the transition intensities of a multi-state power-generating unit by defining a special Markov chain embedded in the observed capacity process. However, the results of point estimation are biased when data are sparse. The parameter uncertainty due to limited data and/or vague information has been taken into account in several existing works. For example, the transition intensities or the state distributions of components in an MSS were treated as fuzzy numbers [5, 6], interval values [7], and belief function [8]. Although these methods can quantify the uncertainty associated with reliability measures of interest from various angles, such uncertainty cannot be progressively reduced by collecting additional data. The Bayesian approach, which treats the unknown parameters to be inferred as random variables, enables reliability engineers to systematically synthesize the subjective information from experts and intuitive judgements with actual observed data. The estimates can be progressively updated as more data and information become available. In our earlier work, a Bayesian framework has been proposed to assess reliability and performance of MSSs [9]. Two scenarios, i.e., components are continuously and discontinuously inspected, have been discussed. The uncertainty of estimates will eventually propagate to the reliability measures of interest via a simulation method. Yet, as demonstrated in our illustrative cases, adding the same amount of additional observations to each component of an MSS may not yield an even contribution to uncertainty reduction of system reliability function. Therefore, it raises a new research question: how to allocate the additional reliability testing resources strategically if reliability engineers aim at further reducing the uncertainty associated with the reliability measures of interest?

The testing resources allocation problem has been reported in the existing literature. A methodology for allocating additional testing resources across the fault tree events with the purpose of minimizing the uncertainty of the top event probability has been investigated in Hamada et al. [10]. The events in a fault tree were binary-state in their work. Anderson-Cook et al. [11] developed an approach to assess the relative improvement in system reliability estimation for additional data from various types of aging components. The data for components could be pass/fail observations, degradation data, and lifetime data, and components can be in one of only two states, either functioning or failed. The aforementioned optimization question has been extended to a multi-objective optimization problem, in which the widths of the credible intervals of system and two subsystem reliability estimates were maximally reduced simultaneously by allocating limited testing resources [12]. Nonetheless, to date, the reliability testing resources allocation problem has rarely been studied for MSSs, and this chapter serves this purpose.

The remainder of this chapter is organized as follows: In Sect. 2, the Bayesian reliability assessment method developed in our earlier work is briefly reviewed first. It is followed by the proposed reliability testing resources allocation approach in Sect. 3. The details of evaluating performance of candidate allocation schemes, the kriging metamodel, and the optimization algorithm are elaborated. Two illustrative examples are presented in Sect. 4 to demonstrate the effectiveness of the proposed approach. A brief conclusion is given in Sect. 5.

2 Review of Bayesian Reliability Assessment for MSS

The MSS with ordered states [13] in question is assumed to be composed of M non-repairable statistically independent multi-state components, each of which has k_l ($l \in \{1, 2, \ldots, M\}$) different states distinguished by the possible performance capacities $\mathbf{g}_l = \{g_{(l,1)}, g_{(l,2)}, \ldots, g_{(l,k_l)}\}$, where $g_{(l,i)} < g_{(l,j)}$ for $i < j$. The stochastic deteriorating behaviors of components are governed by the homogenous continuous-time Markov model. In this case, the probability of component l remaining at any particular state in a future time is statistically independent of its previous state. Many engineering systems can be characterized by the aforementioned stochastic model, such as manufacturing systems [2], power systems [1], and flow transmission systems [3].

The state-space diagram of multi-state component l is given in Fig. 1, where $\lambda^l_{(i,j)}$ ($i,j \in \{1, 2, \ldots, k_l\}$, $i \neq j$) denotes the constant intensity of component l transitioning from state i to state j. Given the initial condition that component l is at state u at $t = 0$, the corresponding Kolmogorov differential equations can be formulated as:

$$\frac{dp^l_{(u,i)}(t)}{dt} = \sum_{j=i+1}^{k_l} p^l_{(u,j)}(t)\lambda^l_{(j,i)} - p^l_{(u,i)}(t)\sum_{j=1}^{i-1}\lambda^l_{(i,j)}, \tag{1}$$

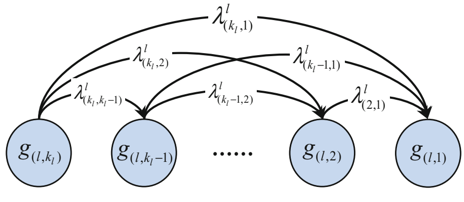


Fig. 1 The state-space diagram of non-repairable multi-state component l

where $p^l_{(u,i)}(t)$ is the probability of component l being at state i at time instant t if it is initially at state u at $t = 0$. The initial condition is set to be $p^l_{(u,u)}(0) = 1$ and $p^l_{(u,i)}(0) = 0$ $(i \neq u)$ $(i,\ u \in \{1, 2, \ldots, k_l\})$. By resolving the differential equations, one can get the state probability $p^l_{(u,i)}(t)$ as a function of $\boldsymbol{\lambda}^l$ (a vector of $\lambda^l_{(i,j)}$, $i,j \in \{1, 2, \ldots, k_l\}$, $i \neq j$) and time. Nevertheless, in this work, the transition intensities $\boldsymbol{\lambda}^l$ are unknown parameters to be estimated by observations.

In our earlier work, a Bayesian framework has been developed to infer the unknown parameters $\boldsymbol{\lambda}^l$ of multi-state components and assess reliability measures of a multi-state system [2]. It enables reliability engineers to systematically synthesize the subjective information from experts and intuitive judgements with actual observations, thereby, obtaining a balanced estimate. On the other side, the estimates from the Bayesian approach can be progressively updated as more data become available, and the uncertainty due to the limited data can also be quantified. The proposed Bayesian reliability assessment method for MSSs is composed of six steps as illustrated in Fig. 2. In Step 1, data are collected by conducting inspections. According to the data collection strategy, two types of data, i.e., continuous inspection data and discontinuous inspection data, are involved. The likelihood functions of the two types of data are constructed in Step 2. By merging the collected data with the prior knowledge of the unknown parameters, i.e., $\boldsymbol{\lambda}^l$, from Step 3, the Bayesian inference can be performed in Step 4 to yield the posterior distributions of the unknown parameters. In Step 5, by the Monte Carlo (MC) simulation, a set of samples of $\boldsymbol{\lambda}^l$ is randomly generated to get the corresponding universal generating functions (UGFs) of component l. The UGFs of the entire system, representing the system state distribution, can be derived by aggregating the UGFs of all the components. Consequently, reliability measures, such as

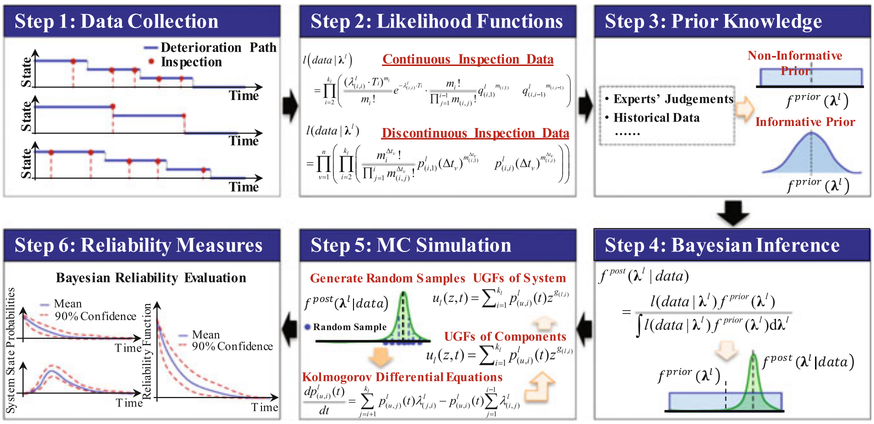


Fig. 2 The steps of the proposed Bayesian reliability assessment method for MSSs

the reliability function, the state probabilities, the instantaneous expected performance capacity, etc., can be evaluated in Step 6 via the UGFs of the entire system. The ensuing sections will review the technical details of this method.

2.1 Bayesian Parameter Inference for Multi-state Components

Following the general framework of Bayesian inference, the posterior distribution of the unknown parameters, i.e., $\boldsymbol{\lambda}^l$, of component l can be readily estimated by [9]:

$$f^{post}(\boldsymbol{\lambda}^l|data) = \frac{l(data|\boldsymbol{\lambda}^l)f^{prior}(\boldsymbol{\lambda}^l)}{\int l(data|\boldsymbol{\lambda}^l)f^{prior}(\boldsymbol{\lambda}^l)d\boldsymbol{\lambda}^l}, \quad (2)$$

where $f^{prior}(\boldsymbol{\lambda}^l)$ is the prior distribution of unknown parameters $\boldsymbol{\lambda}^l$, whereas $f^{post}(\boldsymbol{\lambda}^l|data)$ is the posterior distribution of $\boldsymbol{\lambda}^l$ given observations. $l(data|\boldsymbol{\lambda}^l)$ is the likelihood function.

In accordance to the data collection strategy, two common scenarios, i.e., continuous inspection data and discontinuous inspection data, are studied to construct the corresponding likelihood functions.

Scenario I: Continuous Inspection Data

In this scenario, components are continuously inspected all the time, and the exact times of components transitioning from one state to another can be recorded. Thus, one can evaluate the following quantities:

(1) The number of the transitions from state i to state j among all the observations, denoted as $m_{(i,j)}(i > j, i \in \{2, 3, \ldots, k_l\}, j \in \{1, 2, \ldots, k_l - 1\})$;
(2) The total time that components are remaining at state i, denoted as $T_i \ (i \in \{2, 3, \ldots, k_l\})$;
(3) The total number of the transitions from state i, denoted as $m_i(i = 2, 3, \ldots, n)$, and it can be computed by $m_i = \sum_{j=1}^{i-1} m_{(i,j)}$.

Bear in mind that the deterioration of each component is assumed to comply with a homogenous continuous-time Markov model, m_i follows a Poisson process, i.e., $m_i \sim \text{Poisson}(\sum_{j=1}^{i-1} \lambda_{(i,j)}^l \cdot T_i)$, and its probability density function is expressed as:

$$f_1\left(m_i|\sum_{j=1}^{i-1} \lambda_{(i,j)}^l \cdot T_i\right) = \frac{\left(\sum_{j=1}^{i-1} \lambda_{(i,j)}^l \cdot T_i\right)^{m_i}}{m_i!} e^{-\sum_{j=1}^{i-1} \lambda_{(i,j)}^l \cdot T_i}, \quad (3)$$

If a transition occurs at state i, the conditional probability $q^l_{(i,j)} (i > j,$ $i \in \{2, 3, \ldots, k_l\}, j \in \{1, 2, \ldots, k_l - 1\})$ that the transition is from state i to state j is given by:

$$q^l_{(i,j)} = \frac{\lambda^l_{(i,j)}}{\sum_{j=1}^{i-1} \lambda^l_{(i,j)}}. \tag{4}$$

Thereby, the set of quantities $\left(m_{(i,1)}, m_{(i,2)}, \ldots, m_{(i,i-1)}\right)$ $(i \in \{2, 3, \ldots, k_l\})$ follows the multinomial distribution with parameters m_i and $\mathbf{q}^l_i = (q^l_{(i,1)}, q^l_{(i,2)}, \ldots,$ $q^l_{(i,i-1)})$, and the corresponding probability mass function can be formulated as:

$$f_2\left(m_{(i,1)}, \ldots, m_{(i,i-1)} \middle| \mathbf{q}^l_i, m_i\right) = \frac{m_i!}{\prod_{j=1}^{i-1} m_{(i,j)}!} {q^l_{(i,1)}}^{m_{(i,1)}} \cdots {q^l_{(i,i-1)}}^{m_{(i,i-1)}}. \tag{5}$$

Hence, the likelihood function is the product of Eqs. 3 and 5, and written as [9]:

$$l\left(data \middle| \boldsymbol{\lambda}^l\right) = \prod_{i=2}^{k_l} \left(\frac{\left(\sum_{j=1}^{i-1} \lambda^l_{(i,j)} \cdot T_i\right)^{m_i}}{m_i!} e^{-\sum_{j=1}^{i-1} \lambda^l_{(i,j)} \cdot T_i} \cdot \frac{m_i!}{\prod_{j=1}^{i-1} m_{(i,j)}!} {q^l_{(i,1)}}^{m_{(i,1)}} \cdots {q^l_{(i,i-1)}}^{m_{(i,i-1)}} \right). \tag{6}$$

Scenario II: Discontinuous Inspection Data

Continuous inspections may be costly in engineering practice. Alternatively, components can be inspected periodically or non-periodically. However, in this case, only the state of components at each inspection time and the time interval between two adjacent inspections are recorded. The collected data cannot reflect the time duration that components resides in each state and the exact paths that components degrade from the best state to the worst. If components are inspected periodically with a time interval of Δt, the following quantities can be derived from discontinuous observations:

(1) The number of inspections in which components are observed at state i in the last inspection and at state j after an inspection interval Δt, denoted as $m^{\Delta t}_{(i,j)}$;

(2) The number of inspections in which the time interval between two adjacent inspections is Δt and components are observed at state i in the last inspection, denoted as $m^{\Delta t}_{(i,j)}$, and one has $m^{\Delta t}_i = \sum_{j=1}^{i} m^{\Delta t}_{(i,j)}$.

Under the assumption that a component's deterioration follows a homogenous continuous-time Markov model, all the observations with the same observed state in the last inspection and the same time interval Δt between two adjacent inspections can be regarded as repeated s-independent trials. It can be characterized by a multinomial distribution. According to Eq. 1, on the condition that component l is observed at state i in the last inspection, the probability of component l being at state $j \in \{1, 2, \ldots, i\}$ after Δt, denoted as $p^l_{(i,j)}(\Delta t)$, is a function of the inspection interval Δt and $\boldsymbol{\lambda}^l$. Hence, the set of quantities $(m^{\Delta t}_{(i,1)}, m^{\Delta t}_{(i,2)}, \ldots, m^{\Delta t}_{(i,i)})$ follows a multinomial distribution with parameters $m^{\Delta t}_i$ and $\mathbf{p}^l_i(\Delta t) = (p^l_{(i,1)}(\Delta t), p^l_{(i,2)}(\Delta t), \ldots, p^l_{(i,i)}(\Delta t))$, and it is written as:

$$f_3(m^{\Delta t}_{(i,1)}, m^{\Delta t}_{(i,2)}, \ldots, m^{\Delta t}_{(i,i)} | \mathbf{p}^l_i(\Delta t), m^{\Delta t}_i) = \frac{m^{\Delta t}_i!}{\prod_{j=1}^{i} m^{\Delta t}_{(i,j)}!} p^l_{(i,1)}(\Delta t)^{m^{\Delta t}_{(i,1)}} p^l_{(i,2)}(\Delta t)^{m^{\Delta t}_{(i,2)}} \ldots p^l_{(i,i)}(\Delta t)^{m^{\Delta t}_{(i,i)}}. \tag{7}$$

To generalize this scenario to a non-periodical case, a vector $\Delta\mathbf{t} = (\Delta t_1,\ \Delta t_2, \ldots \Delta t_n)$ $(n \in \{1, 2, \ldots\})$ with finite time intervals is used to represent the distinct time intervals between two adjacent inspections in the cases of non-periodical inspections. The quantities $m^{\Delta t_v}_i$ $(i \in \{2, 3, \ldots, k_l\})$ and $m^{\Delta t_v}_{(i,j)}$ $(j \le i,\ i \in \{2, 3, \ldots, k_l\},\ \ j \in \{1, 2, \ldots, k_l\})$ for each individual time interval $\Delta t_v (v \in \{1, 2, \ldots, n\})$ can be evaluated based on all the collected data. Thus, the likelihood function of all the observations under non-periodical inspections can be formulated as [9]:

$$l(data|\boldsymbol{\lambda}^l) = \prod_{v=1}^{n} \left(\prod_{i=2}^{k_l} \left(\frac{m^{\Delta t_v}_i!}{\prod_{j=1}^{i} m^{\Delta t_v}_{(i,j)}!} p^l_{(i,1)}(\Delta t_v)^{m^{\Delta t_v}_{(i,1)}} p^l_{(i,2)}(\Delta t_v)^{m^{\Delta t_v}_{(i,2)}} \ldots p^l_{(i,i)}(\Delta t_v)^{m^{\Delta t_v}_{(i,i)}} \right) \right). \tag{8}$$

By plugging Eq. 6 or 8 into the Bayesian formula Eq. 2, together with the prior knowledge $f^{prior}(\boldsymbol{\lambda}^l)$, one can get the posterior distributions $f^{post}(\boldsymbol{\lambda}^l|data)$ of $\boldsymbol{\lambda}^l$ via Eq. 2. Depending on the experts' knowledge or historical data, a particular distribution, such as the Gamma distribution, the Beta distribution, etc., can be chosen for the prior distribution $f^{prior}(\boldsymbol{\lambda}^l)$ [14]. Alternatively, the uniform distribution can be used as a non-informative prior if the prior knowledge is unavailable. The methods to determine the prior distribution in the Bayesian framework can be found in Wang et al. [14], Smith [15], Hamada et al. [16], and Kelly [17].

However, it should be noted that the analytical solutions to the posterior distributions of $\boldsymbol{\lambda}^l$ may not exist. In our study, the Markov Chain Monte Carlo (MCMC) method is used, as an alternative, to generate the posterior distribution via simulation. The technical details of the MCMC method are available in Hamada et al. [16] and Kelly [17].

2.2 Bayesian Reliability Assessment for Multi-state Systems

Given a particular set of the transition intensities $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$), the state distributions of components and/or systems at any time instant can be evaluated. By using the UGF, the state distribution of a multi-state component can be written as a polynomial form as $u_l(z,t) = \sum_{i=1}^{k_l} p^l_{(u,i)}(t) z^{g_{(l,i)}}$, where $p^l_{(u,i)}(t)$ derived by Eq. 1 is a function of $\boldsymbol{\lambda}^l$ and time. In the same fashion, the state distribution of an MSS can also be represented by an UGF as $U_S(z,t) = \sum_{i=1}^{N_S} p_i(t) z^{g_i}$, where $p_i(t)$ is the probability of the system staying at state i at time t and it is a function of $p^l_{(u,i)}(t)$ ($l \in \{1, 2, \ldots, M\}$). As such, $p_i(t)$ is also a function of $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$). g_i is the performance capacity of the system at state i. N_S is the number of system states. The UGF of the system can be recursively derived by the UGFs of all the components via composition operators [2, 18].

The reliability of the studied MSS is defined as the probability of the system's performance capacity being not less than a specified demand level W. Hence, the system reliability function can be formulated as:

$$R(t) = \sum_{j=1}^{H} q_j \sum_{i=1}^{N_S} p_i(t) \cdot 1(g_i - w_j \geq 0), \tag{9}$$

where w_j ($j \in \{1, 2, \ldots, H\}$) is the possible value of W with the associated probability mass function $\Pr\{W = w_j\} = q_j$. $1(\cdot)$ is a unity function, i.e., $1(\text{TRUE}) = 1$ and $1(\text{FALSE}) = 0$. As $p_i(t)$ is completely determined by $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$), $R(t)$ is, therefore, a function of $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$).

However, unlike most reported works in which the transition intensities $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$) of all the components are assumed to be pre-specified precise values, in this work, the transition intensities are estimated from the proposed Bayesian framework and characterized by a set of posterior distributions. Such uncertainty associated with the parameter inference will eventually propagate to the reliability measures of interests, say $R(t)$, which are functions of the estimates. Hence, at any time instant, the system reliability evaluated based on the posterior distributions of the transition intensities is a random quantity. In our earlier work, a simulation-based method was developed to approximate the posterior distributions of system reliability at any time [9]. The basic procedures of the simulation method are as follows:

(1) N_{sa}, say $N_{sa} = 1000 \sim 5000$, samples of the transition intensities $\boldsymbol{\lambda}^l$ for each component are randomly generated based on the posterior distributions.
(2) The state distribution of each component with respect to the ith sample of the transition intensities $\boldsymbol{\lambda}^l$ can be produced by solving Eq. 1.
(3) The system reliability function can be obtained by aggregating all the components' UGFs solved by the ith sample of the transition intensities $\boldsymbol{\lambda}^l$.

(4) By putting all the results from the N_{sa} samples together, the posterior distributions of the system reliability at any time instant can be approximated by fitting all the N_{sa} results with either a parametric distribution (e.g., normal, Weibull) or non-parametric distribution (e.g., the empirical distribution).

3 Optimal Testing Resources Allocation Strategy

The testing resources allocation problem concerns the sequential experiments in the reliability field, with which the best strategy to allocate the future available resources for a new data collection can be determined [11]. The role of the testing resources allocation in the progressive reliability evaluation of a product is depicted in Fig. 3. Reliability analysis is conducted based on the available initial data collected at the present Phase 1. From the analysis, the system reliability can be estimated and predicted. If the results are not credible enough, Phase 2 will be, therefore, involved to collect more data to further update the estimates and predictions. The sequential experiment process of collecting new data will continue whenever additional testing resources are available until the results of interest are satisfactory. The testing resources allocation is a decision-making action bridging the two adjacent phases, and it can provide a cost-efficient allocation strategy which yields a maximum improvement to the reliability estimates and predictions.

In our particular study, due to the limited data collected from reliability tests, the uncertainty of the estimates, i.e., the transition intensities λ^l $(l \in \{1, 2, \ldots, M\})$, and the reliability measures, i.e., reliability function, cannot be completely eliminated. These uncertainties have been quantified by the corresponding posterior distributions in the Bayesian framework as introduced in Sect. 2. On the other hand, by conducting sequential reliability tests, the newly collected data can be further used to reduce the uncertainty associated with the estimates and the reliability measures of interest. The specific objective of this study is, therefore, to determine the optimal scheme for the reliability testing resources allocation of the next phase,

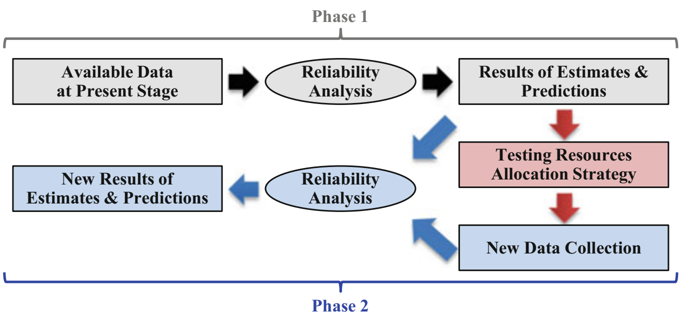


Fig. 3 The role of the testing resources allocation in the process of reliability analysis

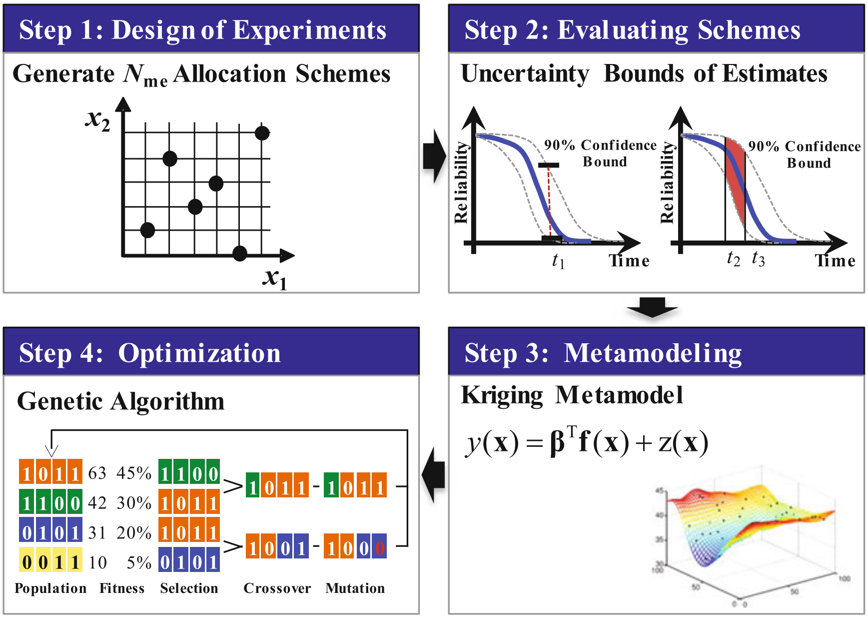


Fig. 4 The four steps of the proposed metamodel-based approach for seeking the optimal testing resources allocation scheme

so as to reduce the uncertainty of the estimated reliability measures of interest as much as possible.

As seeking the optimal testing resources allocation strategy is extremely time-consuming, a metamodel-based approach is developed in this study to alleviate the computational burden. The basic steps of the proposed metamodel-based approach to identify the optimal testing resources allocation strategy are plotted in Fig. 4.

In Step 1, a set of candidate schemes for reliability testing resources allocation is randomly generated by the design of experiment (DOE), such as the full factorial experimental design, the Latin Hypercube Design (LHD), the Improved distributed Hypercube Sampling (IHS). The candidate schemes produced by the DOE are required to evenly spread over the decision space. In Step 2, the performance of each scheme generated in Step 1 will be evaluated. In our particular study, we are concerned with the uncertainty associated with the reliability measures of interest. The improvement to the uncertainty of system reliability estimation/prediction will be quantified by a metric. As evaluating, the performance of all the candidate schemes is computationally unaffordable, the kriging model, as a surrogate model, will be constructed in Step 3 to approximate the implicit relationship between the decision variables and the performance of candidate schemes. New candidate schemes together with their performance evaluation may be added into the initial DOE to update the kriging model until the accuracy of the kriging model is acceptable. In Step 4, the optimization algorithm, such as the genetic algorithm, will be directly performed on the kriging model to seek the global optimal solution, and it corresponds to the optimal testing resources allocation strategy that decision-makers are looking for. The technical details of some steps will be elaborated in the ensuing sections.

3.1 Evaluating Performance of Candidate Allocation Schemes

To determine the optimal testing resources allocation scheme, it is necessary to define a criterion to evaluate the performance of each candidate scheme. In this study, we are concerned with the uncertainty associated with the reliability measures of interest, and the testing resources allocation scheme which can maximally reduce the uncertainty of the estimated reliability measures is preferable. As reported in Anderson-Cook et al. [11], several possible metrics can be used to quantify the uncertainty of the reliability measures of interest, such as the width of a particular $(1-\alpha)\times 100\%$ confidence bound and the entropy of the estimate. Although these metrics are all asymptotically equivalent as claimed by Wynn [19], the different metrics will lead to different relative rankings of the candidate allocation schemes. In our study, as the system reliability function is uncertain due to the uncertainty associated with the transition intensities $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$), we choose the width of the $(1-\alpha)\times 100\%$ confidence bound as a metric, denoted as $R_{(1-\alpha)\times 100\%}(t|f^{post}(\boldsymbol{\lambda}^l|data))$, to quantify the uncertainty of the system reliability at a particular time instant, as shown in Fig. 5a. If decision-makers concern with the uncertainty of the system reliability in a period of time, the integration of the width of the $(1-\alpha)\times 100\%$ confidence bound over the particular period of time, as depicted in Fig. 5b, can be used as an alternative metric. In this study, we will only focus on the former case in the illustrative examples.

The performance of a candidate allocation scheme is evaluated by examining the expected uncertainty of the reliability measures of interest after conducting the

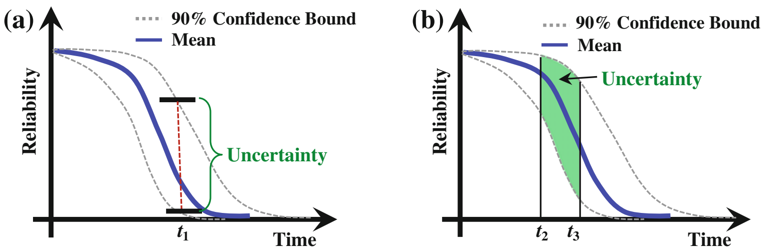


Fig. 5 The illustration of the metrics for uncertainty quantification. **a** Uncertainty at a particular time instant; **b** uncertainty of a period of time

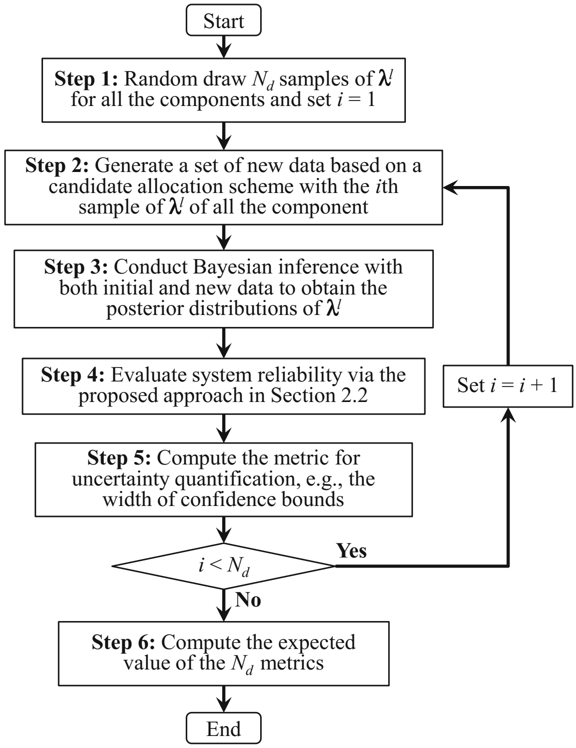


Fig. 6 The flowchart of evaluating the expected uncertainty of the reliability measures

sequential experiments. The scheme with smaller expected uncertainty is preferable. Nevertheless, the expected uncertainty of the reliability measures cannot be derived analytically for each candidate scheme. In this study, a simulation-based approach is developed to evaluate the expected uncertainty of the reliability measures after carrying out a testing resources allocation scheme. The flowchart of the simulation-based approach is shown in Fig. 6. In Step 1, N_d samples of transition intensities $\boldsymbol{\lambda}^l$ ($l \in \{1, 2, \ldots, M\}$) are randomly drawn from the posterior distributions $f^{post}(\boldsymbol{\lambda}^l|data)$. The posterior distributions are the estimates of $\boldsymbol{\lambda}^l$ based on the initial data at Phase 1 as shown in Fig. 3. And then, set the index $i = 1$. It is followed by Step 2 in which the deterioration paths of components are randomly generated based on the ith random sample of $\boldsymbol{\lambda}^l$, and a set of new artificial observations can be collected based on the candidate allocation schemes. By merging both the initial observations from Phase 1 and the new artificial observations, the Bayesian inference introduced in Sect. 2.1 can be executed to update the posterior distributions of $\boldsymbol{\lambda}^l$ in Step 3, and then, in Step 4, the reliability of the entire system can be evaluated by the proposed approach in Sect. 2.2.

The uncertainty associated with the reliability measures of interest is quantified by the metrics, such as the width of the $(1 - \alpha) \times 100\%$ confidence bound, in Step 5.

If $i < N_d$, set $i = i + 1$ and go to Step 2. Otherwise, go to Step 6 to compute the expected value of the metrics, e.g., the expected width of the $(1-\alpha) \times 100\%$ confidence bounds of the reliability measures, based on the N_d results.

It should be noted that,

(1) as the true values of $\boldsymbol{\lambda}^l$ are unknown and the new observations are artificially generated based on the posterior distributions of $\boldsymbol{\lambda}^l$, the expected uncertainty of the reliability measures from the simulation-based approach is not a true value, but a predictive value;
(2) the simulation-based approach is very time-consuming, and it is computationally unaffordable to enumerate the performance of all the candidate schemes.

3.2 Kriging Model

To mitigate the computational burden in enumerating the performance of candidate schemes, the metamodeling technique is adopted in this study to approximate the implicit relationship between the decision variables (corresponding a particular candidate allocation scheme) and the performance of schemes, i.e., the expected width of $(1-\alpha) \times 100\%$ confidence bound. Many metamodeling tools can be used here, such as the polynomial response surface, the radial basis function, the kriging, the artificial neural networks, and support vector machine, etc. It is desired that a metamodel is capable of capturing both global and local trends with a few training samples. A comparative study on the performance of various metamodels has been reported in Jin et al. [20, 21]. In this study, we choose the kriging model as a surrogate model because it has extremely widespread applications due to its flexibility and high accuracy [22–24].

In essence, the kriging model is a semi-parametric interpolation technique based on statistical theory. The kriging model is composed by a polynomial model and a stochastic model as follows:

$$y(\mathbf{x}) = \boldsymbol{\beta}^T \mathbf{f}(\mathbf{x}) + z(\mathbf{x}) = \sum_{i=1}^{P} \beta_i f(\mathbf{x}) + z(\mathbf{x}), \tag{10}$$

where P is the number of basic functions; $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_P(\mathbf{x})]^T$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_P]^T$ are polynomial function of inputs $\mathbf{x}$ and the corresponding coefficients, respectively, and they provide a global approximation. While $z(\mathbf{x})$ is the lack of fit and is represented by a realization of a random process with mean zero and non-zero covariance. The covariance of the residuals at any two sites, say $z(\mathbf{x}_i)$ and $z(\mathbf{x}_j)$, can be expressed by:

$$Cov\left[z(\mathbf{x}_i),\ z(\mathbf{x}_j)\right] = \sigma^2 R(\mathbf{x}_i,\ \mathbf{x}_j), \tag{11}$$

where σ^2 is the variance of $z(\mathbf{x})$; $R(\mathbf{x}_i, \mathbf{x}_j)$ is the spatial correlation function of any two sites, i.e., $z(\mathbf{x}_i)$ and $z(\mathbf{x}_j)$ in the sample space. The correlation function $R(\mathbf{x}_i, \mathbf{x}_j)$ plays an important role in determining the accuracy of the model. Many correlation functions can be chosen, such as linear, spherical, exponential, and Gaussian correlation functions, etc. Among these options, the Gaussian correlation function is the most popular, and it is given by:

$$R(\mathbf{x}_i,\ \mathbf{x}_j) = \exp\left(- \sum_{k=1}^{n} \theta_k \left| x_i^k - x_j^k \right|^2 \right), \tag{12}$$

where x_i^k and x_j^k are the k th elements of $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively; $\boldsymbol{\theta} = [\theta_1,\ \theta_2,\ \ldots,\theta_n]^T$ are the correlation parameters which measure how fast the correlation between $\mathbf{x}_i$ and $\mathbf{x}_j$ decays with the distance between these two sites. $\{\boldsymbol{\theta},\ \boldsymbol{\beta},\ \sigma^2\}$ are the unknown parameters of a kriging model and they can be estimated via the maximum likelihood estimations with existing training samples.

The predicted mean value of the estimated response y at any un-sampled site $\mathbf{x}$ is:

$$\hat{\mu}_y(\mathbf{x}) = f(\mathbf{x})\hat{\boldsymbol{\beta}} + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}\left(\mathbf{y}_s - \mathbf{f}\hat{\boldsymbol{\beta}}\right), \tag{13}$$

and

$$\beta = \left(\mathbf{f}^T\mathbf{R}^{-1}\mathbf{f}\right)^{-1}\mathbf{f}^T\mathbf{R}^{-1}\mathbf{y}_s. \tag{14}$$

Where the column vector $\mathbf{y}_s$ contains the response values at all sample sites; $\mathbf{f}$ are the values of the polynomial function at all of the sample sites; $\mathbf{r}(\mathbf{x})$ are the correlations between the un-sampled site $\mathbf{x}$ and all of the sample sites; and $\mathbf{R}$ is a correlation matrix of all of the sample sites. The predicted variance of the estimated response y at the un-sampled site $\mathbf{x}$ is:

$$\hat{\sigma}_y^2(\mathbf{x}) = \hat{\sigma}^2\left(1 - \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) + \left(f(\mathbf{x}) - \mathbf{f}^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})\right)^T\left(\mathbf{f}^T\mathbf{R}^{-1}\mathbf{f}\right)^{-1}\left(f(\mathbf{x}) - \mathbf{f}^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})\right)\right), \tag{15}$$

where

$$\hat{\sigma}^2 = \frac{1}{n}\left(\left(\mathbf{y}_s - \mathbf{f}\hat{\boldsymbol{\beta}}\right)^T\mathbf{R}^{-1}\left(\mathbf{y}_s - \mathbf{f}\hat{\boldsymbol{\beta}}\right)\right), \tag{16}$$

and $\hat{\sigma}_y^2(\mathbf{x})$ quantifies the interpolation uncertainty associated with the un-sampled site $\mathbf{x}$.

Given a set of training samples, the unknown parameters, i.e., $\{\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2\}$, in a kriging model can be estimated, and then the responses ast un-sampled sites can be predicted by Eq. 13 with the associated predicted variance given in Eq. 15. In this particular study, the inputs $\mathbf{x}$ correspond to a candidate testing resources allocation scheme, e.g., the number of additional specimens and/or the inspection interval for each component, and the response y is the performance of each scheme, e.g. the expected width of $(1-\alpha)\times 100\%$ confidence bound of the concerned reliability measure. The implicit relationship between the inputs and the response can be built up by the initial training samples from the DOE. The global accuracy of the kriging model can be quantified by the metrics [20], such as the R square, the Relative Average Absolute Error (RAAE), the Relative Maximum Absolute Error (RMAE), etc. If the accuracy of the kriging model is not satisfactory, additional training samples could be further generated by the sequential DOE to update the kriging model [25, 26]. The kriging model with acceptable accuracy will be used in the next step to seek the optimal testing resources allocation strategy.

3.3 Optimization Model and Algorithm

In this study, the reliability testing resources allocation for MSSs can be formulated as an optimization problem as follows:

$$\begin{array}{ll} \min & E\left[R_{(1-\alpha)\times 100\%}(t|f^{post}(\boldsymbol{\lambda}^l|data))|(\mathbf{s},\Delta\mathbf{t})\right] \\ s.t. & C(\mathbf{s},\Delta\mathbf{t})\le C_0 \\ & \quad s_l \le s_l^0 \\ & \quad \Delta t_l \ge 0 \end{array} \tag{17}$$

where $\mathbf{s}=\{s_1, s_2, \ldots, s_M\}$ and $\Delta\mathbf{t}=\{\Delta t_1, \Delta t_2, \ldots, \Delta t_M\}$ are two sets of decision variables, representing the number of additional specimens to be allocated to each component and the inspection intervals for collecting data, respectively. A setting for $\{\mathbf{s},\Delta\mathbf{t}\}$ corresponds to a candidate scheme for the reliability testing resources allocation. If Δt_l ($l\in\{1, 2, \ldots, M\}$) is set to be zero, it is the case where component l will be continuously inspected to collect observations. C_0 is the cost constraint for the testing resources; s_l^0 is the constraint for the maximum number of additional specimens of component l. $E[\cdot]$ is expectation. The objective function, i.e., $E[R_{(1-\alpha)\times 100\%}(t|f^{post}(\boldsymbol{\lambda}^l|data))|(\mathbf{s},\Delta\mathbf{t})]$, is replaced by the kriging model introduced in Sect. 3.2 to mitigate the computational burden.

It should be noted that the optimization problem in Eq. 17 involves both integer and real decision variables and the number of decision variables increases linearly with respect to the types of components in a system. An exhaustive examination of all the candidate solutions is not realistic due to the limited computational capability. In this study, the genetic algorithm (GA) is utilized to search the global optimal solution owing to its flexibility in terms of representing mixed variables in various optimization problems [27, 28].

The main procedures of implementing the GA to solve the specific optimization problem are as follows:

(1) **Population initialization**. For our specific problem, the chromosome is composed of two parts, and it can be denoted as a string $\mathbf{c} = \{s_1, s_2, \ldots, s_M, \Delta t_1, \Delta t_2, \ldots, \Delta t_M\}$. The first M elements are non-negative integers, corresponding to the additional specimens for components, whereas the last M elements are non-negative real numbers, representing the inspection intervals for data collection. A set of N_g chromosomes, as the initial population, are randomly generated in the first iteration.
(2) **Fitness evaluation**. The expected width of 90% confidence bounds of reliability measures serves as the fitness value of each chromosome, and it is predicted by the kriging model introduced in Sect. 3.2. The smaller expected width of 90% confidence bound, the higher the fitness value. The infeasible solutions which violate the constraints are handled by the penalty function approach.
(3) **New population generation**. The roulette-wheel selection strategy is used to select chromosomes based on their fitness values from the present population to form a new generation of population for the next iteration. The crossover and mutation operators are used to produce new chromosomes to explore the unsearched solution space, while maintaining the diversity of a population. As the chromosome is composed of mixed decision variables, the crossover and mutation operators are performed separately for each of the two parts to keep the digits within their allowable bounds. N_r chromosomes with the highest fitness values will be directly merge into the new generation.
(4) **Iterative process termination**. The optimization procedure terminates when the iteration count reaches N_c. Otherwise, go to Step 2 for the next iteration.

It is worth noting that many other advanced optimization algorithms, such as the Tabu search, the simulated annealing (SA), the ant colony optimization (ACO), etc., can also be used here, instead of the GA, to solve the resulting optimization problem.

4 Illustrative Examples

The illustrative example is a three-unit multi-state power generating system as shown in Fig. 7. Components 1 and 2 connected in parallel constitute a subsystem. The performance capacities and the transition intensities of all the components are

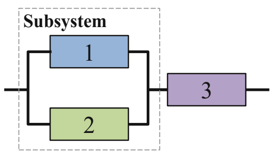


Fig. 7 The system configuration

Table 1 The performance capacities of the three components (*unit* kW)

Component ID	State 1	State 2	State 3	State 4
1	0	4	8	/
2	0	5	8	12
3	0	4	7	10

Table 2 The state transition intensities of the three components (*unit* month^{-1})

Component ID	$\lambda^l_{(4,3)}$	$\lambda^l_{(4,2)}$	$\lambda^l_{(4,1)}$	$\lambda^l_{(3,2)}$	$\lambda^l_{(3,1)}$	$\lambda^l_{(2,1)}$
1	/	/	/	0.2	0.1	0.2
2	0.3	0.2	0.1	0.3	0.2	0.3
3	0.2	0.1	0.05	0.2	0.1	0.2

Table 3 The continuous inspection data of component 1 from 50 specimens

Initial state *i* of a transition	Total sojourning time (months) T_i	Total number m_i	Destination state *j* of a transition	
			$j = 2$	$j = 2$
			$m_{(i,2)}$	$m_{(i,2)}$
$i = 3$	165.94	50	36	14
$i = 2$	167.60	36	–	36

tabulated in Tables 1 and 2, respectively. The transition intensities are assumed to be unknown to the reliability engineers, but can be inferred by the proposed Bayesian approach. The system and subsystem are viewed as failure if their performance capacities are less the required demand level. In Example 1, we only focus on allocating the limited testing resources for components 1 and 2 to improve the reliability estimation of the subsystem at a particular time, while the testing resources are allocated for all the three components in Example 2.

4.1 Example 1

In this example, the testing resources allocation is only considered for components 1 and 2 to improve the reliability estimation of the subsystem at a specific time $t = 3.0$ months. In other words, we expect to reduce the uncertainty of the subsystem reliability estimation at $t = 3.0$ months. According to the transition intensities given in Table 2, 50 deterioration paths are randomly generated for components 1 and 2, respectively. Components 1 and 2 are supposed to be continuously inspected over time. The collected data are tabulated in Tables 3 and 4, and will be used to infer the posterior distributions of transition intensities. The prior distributions of transition intensities are set to be a uniform distribution in the range of [0, 0.5] month^{-1}, together with the observations from 50 specimens, the posterior distributions of the transition intensities of the two components can be

Table 4 The continuous inspection data of component 2 from 50 specimens

Initial state i of a transition	Total sojourning time (months) T_i	Total number m_i	Destination state j of a transition		
			$j=3$	$j=2$	$j=1$
			$m_{(i,3)}$	$m_{(i,2)}$	$m_{(i,1)}$
$i=4$	73.18	50	24	19	7
$i=3$	47.05	24	–	15	9
$i=2$	122.17	34	–	–	34

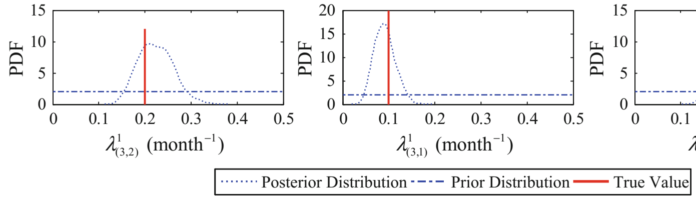


Fig. 8 The prior distributions and the posterior distributions of the transition intensities of component 1

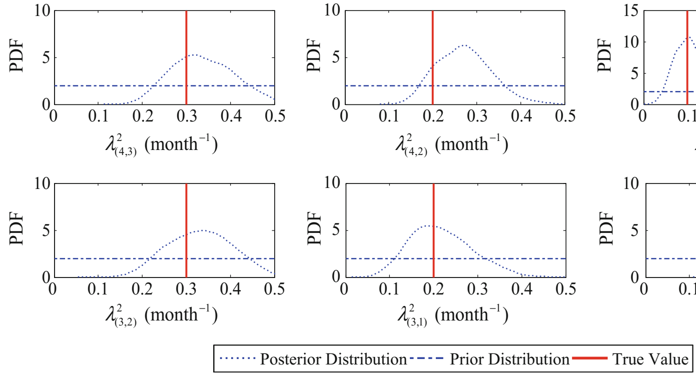


Fig. 9 The prior distributions and the posterior distributions of the transition intensities of component 2

evaluated via the proposed Bayesian approach as shown in Figs. 8 and 9. Consequently, the subsystem reliability function can be estimated, as shown in Fig. 10, via the proposed simulation method in Sect. 2.2 if the required demand level W is given as tabulated in Table 5. The width of the 90% confidence bound of the subsystem reliability is 0.1087 at $t = 3.0$ months. If the accuracy of the subsystem reliability function at $t = 3.0$ months is not satisfactory, the inference results at the present stage can be viewed as Phase 1 shown in Fig. 3 and will facilitate the reliability testing resources allocation at Phase 2.

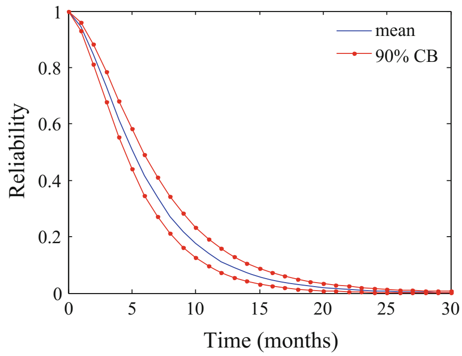


Fig. 10 The estimated reliability function of the subsystem

Table 5 The possible values of the required demand W for the subsystem (*unit* kW)

Possible required demand W	3	5	9
Probability	0.4	0.4	0.2

Suppose that the total budget for the reliability testing resources at Phase 2 is $C_0 = 29{,}500$ US dollars. The costs for conducting a test with continuous inspections for components 1 and 2 are 400 US dollars and 800 US dollars, respectively. A candidate testing resources allocation scheme, i.e., the additional number of reliability tests for each component, is denoted as $\{s_1, s_2\}$ where $s_i \in [0, 50]$ ($i \in \{1, 2\}$) is assumed as the decision space. By the Latin Hypercube Design (LHD), 15 candidate testing resources allocation schemes are randomly generated within the decision space. The performance, i.e., the expected width of 90% confidence bound of the subsystem reliability at $t = 3.0$ months, of these candidate schemes are evaluated by the proposed approach in Sect. 3.1. However, it costs around 10 min to evaluate the performance for a candidate scheme, in which the sample size N_d is set to be 50. A kriging model, as depicted in Fig. 11, is therefore constructed to approximate the relationship between candidate schemes and the predicted performance of the schemes.

As seen in Fig. 11, with the increase of the numbers of reliability tests, the expected width of 90% confidence bound of the subsystem reliability at $t = 3.0$ months declines. Additionally, adding the specimens for component 2 is more effective to reduce uncertainty than that of component 1, because the expected width of 90% confidence bound has a steeper decreasing trend along the s_2 axis.

The resulting optimization problem with a budget constraint, i.e., $400s_1 + 800s_2 \leq 29{,}500$, can be resolved by the proposed GA algorithm. The optimal allocation strategy is $\{s_1^* = 17, s_2^* = 28\}$, and the corresponding expected width of confidence bound of the subsystem reliability at $t = 3.0$ months is reduced to 0.0852. By replacing the time-consuming performance evaluation with a kriging model, the optimization algorithm takes less than 5.0 s via Matlab 2012 on a

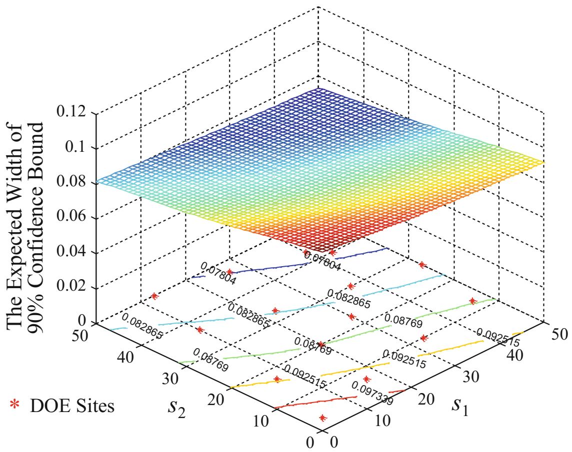


Fig. 11 The kriging model and the samples from the LHD

Table 6 The continuous inspection data of component 3 from 50 specimens

Initial state i of a transition	Total sojourning time (months) T_i	Total number m_i	Destination state j of a transition		
			$j = 3$	$j = 2$	$j = 1$
			$m_{(i,3)}$	$m_{(i,2)}$	$m_{(i,1)}$
$i = 4$	155.09	50	30	15	5
$i = 3$	94.31	30	–	20	10
$i = 2$	163.21	35	–	–	35

workstation with an Intel Xeon 2.10 GHz and 128 GB RAM when N_g and N_c are set to 50 and 1000, respectively.

4.2 *Example 2*

The proposed testing resources allocation approach is further validated in the entire system as shown in Fig. 7. The limited testing resources will be optimally distributed to the three components with the purpose of further reducing the uncertainty of the system reliability estimation at $t = 3.0$ months. In addition to the data in Tables 3 and 4, 50 specimens of component 3 are supposed to be continuously inspected at Phase 1 and the collected data are tabulated in Table 6. The posterior distributions of transition intensities of the three components are evaluated by the proposed Bayesian method and depicted in Figs. 8, 9, and 12. They serve as the inputs for the decision-making at Phase 2. The possible values and the corresponding probabilities of the required demand W are given in Table 7. Therefore,

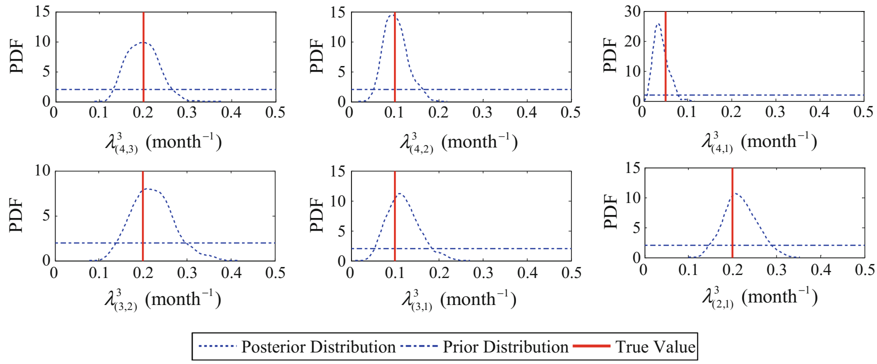
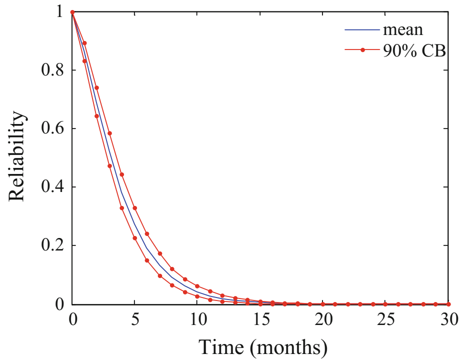


Fig. 12 The prior distribution and the posterior distributions of the transition intensities of component 3

Table 7 The possible values of the required demand W for the system (*unit* kW)

Possible required demand W	3	3	5	9
Probability	0.2	0.3	0.3	0.2

Fig. 13 The estimated system reliability function in Phase 1



the system reliability at any time instant can be evaluated via Eq. 9. The mean value and the 90% confidence bound of the estimated reliability function are shown in Fig. 13. At $t = 3.0$ months, the width of the 90% confidence bound of the system reliability is 0.1368, and such uncertainty needs to be further reduced in Phase 2.

The budget for the reliability testing resources at Phase 2 is $C_0 = 97,500$ US dollars. The components can be continuously inspected or periodically inspected, and thus, a candidate testing resources allocation scheme can be represented by $\{s_1, s_2, s_3, \Delta t_1, \Delta t_2, \Delta t_3\}$ where the number of additional specimens for each component $s_i \in [0, 50]$ $(i \in \{1, 2, 3\})$ and the inspection intervals for each component $\Delta t_i \in [0, 5]$ months $(i \in \{1, 2, 3\})$. The cost for conducting a test is associated with the inspection interval for data collection. In general, the cost of a test increases monotonically with the frequency of inspections. The following relationships are defined to link the inspection intervals with the cost for a single reliability test:

Component 1

$$c_1 = 3000 - 1000(\Delta t_1)^{0.2},$$

Component 2

$$c_2 = 400 - 100(\Delta t_2)^{0.2},$$

Component 3

$$c_3 = 800 - 200(\Delta t_3)^{0.2},$$

where $\Delta t_i = 0$ ($i \in \{1, 2, 3\}$) corresponds to the case of continuous inspections.

44 candidate allocation schemes are randomly generated via the LHD within the decision space and evaluated, and then, a kriging model is constructed to predict the performance, i.e., the expected width of 90% confidence bound of the system reliability at $t = 3.0$ months, of a candidate scheme. The genetic algorithm with mixed decision variables is used to solve the optimal allocation scheme, and it takes around 30.0 s via Matlab 2012 on a workstation with an Intel Xeon 2.10 GHz and 128 GB RAM when N_g and N_c are set to 60 and 1000, respectively. The optimal testing resources allocation scheme is $\{s_1^* = 43, s_2^* = 39, s_3^* = 22, \Delta t_1^* = 3.7, \Delta t_2^* = 1.1, \Delta t_3^* = 1.64\}$ with the expected width of 90% confidence bound equal to 0.0911 as predicted by the kriging model. By the flowchart in Fig. 6, the true performance of the optimal scheme is 0.0909 which is extremely close to the predicted value.

5 Conclusion

In this chapter, the testing resources allocation problem for MSSs is studied to optimally distribute the limited reliability testing resources to improve the accuracy of reliability estimation/prediction. The approach is on the base of the Bayesian reliability assessment method for MSSs with which both subjective information and actual continuous or discontinues inspection data can be merged to infer the unknown parameters, i.e., transition intensities λ^l. The computational burden in the performance evaluation of candidate schemes is alleviated by introducing the kriging metamodel. The genetic algorithm is utilized to resolve the resulting optimization problem with mixed decision variables. Two illustrative examples are given to demonstrate the effectiveness and efficiency of the proposed method.

As reliability tests can be conducted at various physical levels of a system, allocating the limited testing resources across multiple levels of a system [29, 30], say system-level test, component-level test, is worth exploring in our future work.

Acknowledgements The authors greatly acknowledge grant support from the National Natural Science Foundation of China under contract number 71371042 and the Fundamental Research Funds for the Central Universities under contract number ZYGX2015J082.

References

1. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system analysis and optimization for engineers and industrial managers. Springer, London
2. Lisnianski A, Levitin G (2003) Multi-state system reliability: assessment. Optimization and applications. World Scientific, Singapore
3. Kuo W, Zuo MJ (2003) Optimal reliability modeling: principles and applications. Wiley, Hoboken, NJ
4. Lisnianski A, Elmakias D, Laredo D, Haim HB (2012) A multi-state Markov model for a short-term reliability analysis of a power generating unit. Reliab Eng Sys Saf 98(1):1–6
5. Ding Y, Zuo MJ, Lisnianski A, Tian Z (2008) Fuzzy multi-state systems: general definitions, and performance assessment. IEEE Trans Reliab 57(4):589–594
6. Liu Y, Huang HZ (2010) Reliability assessment for fuzzy multi-state systems. Int J Syst Sci 41(4):365–379
7. Li CY, Chen X, Yi XS, Tao JY (2011) Interval-valued reliability analysis of multi-state systems. IEEE Trans Reliab 60(4):595–606
8. Destercke S, Sallak M (2013) An extension of universal generating function in multi-state systems considering epistemic uncertainties. IEEE Trans Reliab 62(2):504–514
9. Liu Y, Lin P, Huang HZ (2015) Bayesian reliability and performance assessment for multi-state systems. IEEE Trans Reliab 64(1):394–409
10. Hamada M, Martz HF, Reese CS, Graves T, Johnson V, Wilson AG (2004) A fully Bayesian approach for combining multilevel failure information in fault tree quantification and optimal follow-on resource allocation. Reliab Eng Sys Saf 86(3):297–305
11. Anderson-Cook CM, Graves TL, Hamada MS (2009) Reliability allocation for reliability of a complex system with aging components. Qual Reliab Eng Int 25:481–494
12. Lu L, Chapman JL, Anderson-Cook CM (2013) A case study on selecting a best allocation of new data for improving the estimation precision of system and subsystem reliability using Pareto fronts. Technometrics 55(4):473–487
13. Yu K, Koren I, Guo Y (1994) Generalized multistate monotone coherent systems. IEEE Trans Reliab 43(2):242–250
14. Wang P, Youn BD, Xi Z, Kloess A (2009) Bayesian reliability analysis with evolving, insufficient, and subjective data sets. J Mech Des 131(11):111008
15. Smith JQ (1998) Decision analysis: a Bayesian approach. Chapman and Hall, London
16. Hamada MS, Wilson AG, Reese CS, Martz HF (2008) Bayesian reliability. Springer, London
17. Kelly D, Smith C (2011) Bayesian inference for probabilistic risk assessment. Springer, London
18. Levitin G (2005) The universal generating function in reliability analysis and optimization. Springer, London
19. Wynn HP (2004) Maximum entropy sampling and general equivalence theory. Advances in model-oriented design and analysis, contributions to statistics. Physica-Verlag, Heidelberg, pp 211–218

20. Jin R, Chen W, Simpson T (2001) Comparative studies of metamodeling techniques under multiple modelling criteria. Struct Multidiscip Optim 23:1–13
21. Jin R, Du X, Chen W (2003) The use of metamodeling techniques for optimization under uncertainty. Struct Multidiscip Optim 25:99–116
22. Cremona MA, Liu B, Hu Y, Bruni S, Lewis R (2016) Predicting railway wheel wear under uncertainty of wear coefficient, using universal kriging. Reliab Eng Sys Saf 154:49–59
23. Liu Y, Shi Y, Jiang T, Liu JZ, Wang WJ (2016) Metamodel-based direction guidance system optimization for improving efficiency of aircraft emergency evacuation. Comp Ind Eng 91:302–314
24. Zhang M, Gou W, Li L, Yang F, Yue Z (2016) Multidisciplinary design and multi-objective optimization on guide fins of twin-web disk using Kriging surrogate model. Struct Multidiscip Optim 55(1):361–373
25. Jin R, Chen W, Sudjianto A (2002) On sequential sampling for global metamodeling in engineering design. In: Proceedings of ASME 2002 design engineering technical conferences and computers and information in engineering conference, Montreal, Canada, pp 539–548
26. Wang GG, Shan S (2006) Review of metamodeling techniques in support of engineering design optimization. J Mech Des 129(4):370–380
27. Gen M, Yun YS (2006) Soft computing approach for reliability optimization: state-of-the-art survey. Reliab Eng Sys Saf 91(9):1008–1026
28. Levitin G (2006) Genetic algorithms in reliability engineering. Reliab Eng Sys Saf 91(9):975–976
29. Jiang T, Liu Y (2017) Parameter inference for non-repairable multi-state system reliability models by multi-level observation sequences. Reliab Eng Sys Saf 166:3–15
30. Liu Y, Chen CJ (2017) Dynamic reliability assessment for nonrepairable multistate systems by aggregating multilevel imperfect inspection data. IEEE Trans Reliab 66(2):281–297

Topological Analysis of Multi-state Systems Based on Direct Partial Logic Derivatives

Miroslav Kvassay and Elena Zaitseva

Abstract Topological analysis deals with evaluation of influence of the system components on system operation. Such evaluation is usually performed by identification and quantification of situations in which degradation/improvement of a given system component results in system degradation/improvement. These situations can be revealed using Direct Partial Logic Derivatives (DPLDs). One of the open problems is how to compute DPLDs efficiently for large systems. In this paper, we develop a new method for their computation for systems that can be decomposed into disjoint modules. The method is based on the chain rule that is derived in this paper.

Keywords Multi-state systems · Structure function · Topological analysis · Modular decomposition · Structural importance measures · Direct partial logic derivatives

1 Introduction

Evaluation of system reliability is a complex problem. The principal task is creation of system model. As a rule, two approaches dominate in reliability analysis: Binary-State Systems (BSSs) and Multi-State Systems (MSSs).

BSSs are based on the assumption that the system and all its components can be in one of only two possible states—perfectly functioning or completely failed [1]. This approach is used in the analysis of systems in which any deviation from perfect functioning can result in a disaster, e.g. aviation systems [2], nuclear power plants [3]. However, it is not very practical for systems that operate at several performance levels, such as distribution networks [4] or complex socio-technical systems [5]. For the analysis of such systems, MSSs are more suitable because they

---

M. Kvassay (✉) · E. Zaitseva
Department of Informatics, University of Zilina, Zilina, Slovakia
e-mail: miroslav.kvassay@fri.uniza.sk

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_14

permit defining more than two states in system or component operation—from perfectly functioning through partially damaged to completely failed [6, 7].

One of the current issues in reliability analysis of MSSs is development of algorithms that can be applied to systems composed of many components. In [8], relations between MSSs and Multiple-Valued Logic (MVL) have been studied. Based on these relations, it has been shown that several tools related to the analysis of MVL functions can also be applied in investigation of MSSs. One of them is logical differential calculus, which has been used in [9] to develop a complex framework for importance analysis of MSSs.

In this paper, we extend the results from the previously mentioned papers. We primarily focus on topological analysis of MSSs and a new method for investigation of MSSs is developed using MVL. The method is based on a modular decomposition of MSSs [7]. Based on this decomposition, an efficient algorithm for quantification of reliability of system topology is proposed. Combination of this algorithm with the chain rule developed in this paper allows efficient computation of logic derivatives that can be used in importance analysis of MSSs.

2 Structure Function

Let us consider a MSS composed of n components, and let us denote the set of the components as $\mathbf{N} = \{1, 2, \ldots, n\}$. A map that defines dependency between states of the components and states of the system is known as structure function [6, 7]:

$$\phi(x_1, x_2, \ldots, x_n) = \phi(\boldsymbol{x}): \{0, 1, \ldots, m-1\}^n \rightarrow \{0, 1, \ldots, m-1\}, \tag{1}$$

where m denotes number of system/components states (state $m - 1$ agrees with perfect functioning while state 0 with complete failure), x_i is a variable defining state of the i-th component for $i \in \mathbf{N}$, and $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ is a vector of components states (state vector). Depending on properties of this function, two classes of systems can be recognized—coherent and noncoherent. In what follows, only coherent systems, i.e. systems with monotonic structure function, will be considered.

The structure function correlates with system topology, but its knowledge is not sufficient for computation of probabilistic characteristics, such as system state probability [8]:

$$\Pr\{\phi(\boldsymbol{x}) = j\}, \text{for } j \in \{0, 1, \ldots, m-1\}, \tag{2}$$

or system availability/unavailability computed with respect to system state j [6, 7]:

$$A^{\geq j} = \Pr\{\phi(\boldsymbol{x}) \geq j\}, U^{\geq j} = \Pr\{\phi(\boldsymbol{x}) < j\}, \quad \text{for } j \in \{1, 2, \ldots, m-1\}. \tag{3}$$

Computation of these and other measures requires knowledge of the state probabilities of the system components:

$$p_{i,s} = \Pr\{x_i = s\}, \quad \text{for } i = 1, 2, \ldots, n, \; s = 0, 1, \ldots, m-1. \tag{4}$$

In this paper, we primarily focus on the analysis of system topology. Because of that we will assume that only the structure function of the system is known.

2.1 Modular Decomposition

According to [7, 10], a module of a MSS is defined as a pair ($\mathbf{B}$, $\phi^{\mathbf{B}}$) such that:

$$\phi(x) = \psi\left(\phi^{\mathbf{B}}\left(x^{\mathbf{B}}\right), x^{\mathbf{B}^c}\right), \tag{5}$$

where $\mathbf{B} \cup \mathbf{B}^c = \mathbf{N}$ and $\mathbf{B} \cap \mathbf{B}^c = \{\}$, $x^{\mathbf{B}}$ is a vector of state variables x_i such that $i \in \mathbf{B}, x^{\mathbf{B}^c}$ is a vector of state variables x_i such that $i \in \mathbf{B}^c$, $\phi^{\mathbf{B}}(x^{\mathbf{B}})$ is a substructure function defined over the components from set $\mathbf{B}$, i.e.:

$$\phi^{\mathbf{B}}(x^{\mathbf{B}}) \colon \{0, 1, \ldots, m-1\}^{|\mathbf{B}|} \to \{0, 1, \ldots, m-1\}, \tag{6}$$

where $|\cdot|$ denotes size of the argument interpreted as a set, and $\psi\left(\phi^{\mathbf{B}}, x^{\mathbf{B}^c}\right)$ is a function that combines substructure function $\phi^{\mathbf{B}}(x^{\mathbf{B}})$ with the variables defining states of the components not contained in set $\mathbf{B}$ into the system structure function.

Based on the previous definition, a modular decomposition of a MSS has been introduced in [7] as a set of k disjoint modules $(\mathbf{N}_1, \phi^{\mathbf{N}_1}), (\mathbf{N}_2, \phi^{\mathbf{N}_2})$, $\ldots$, and $\left(\mathbf{N}_k, \phi^{\mathbf{N}_k}\right)$ together with function $\psi\left(\phi^{\mathbf{N}_1}, \phi^{\mathbf{N}_2}, \ldots, \phi^{\mathbf{N}_k}\right)$ such that:

$$\begin{gathered}\phi(x) = \psi\left(\phi^{\mathbf{N}_1}(x^{\mathbf{N}_1}), \phi^{\mathbf{N}_2}(x^{\mathbf{N}_2}), \ldots, \phi^{\mathbf{N}_k}(x^{\mathbf{N}_k})\right), \\ \text{where } \mathbf{N} = \bigcup_{l=1}^{k} \mathbf{N}_l \text{ and } \mathbf{N}_{l_1} \cap \mathbf{N}_{l_2} = \{\} \text{ for any } l_1, l_2 \in \{1, 2, \ldots, k\} \text{ such that } l_1 \neq l_2.\end{gathered} \tag{7}$$

Clearly, a modular decomposition can be applied not only to the entire system, but also to a module. For this purpose, the following notation will be used in the rest of the paper: ($\mathbf{N}_l, \phi^{\mathbf{N}_l}$)—the l-th module of a MSS, ($\mathbf{N}_{l_1,l_2}$, $\phi^{\mathbf{N}_{l_1,l_2}}$)—the l_2-th module of the l_1-th module of a MSS, etc.

2.2 Series and Parallel Systems

In case of MSSs, series and parallel systems can be defined in several ways. In this paper, we will adapt approach from [11, 12], and we will define a series MSS as a system with the following structure function:

$$\phi(x) = \min(x_1, x_2, \ldots, x_n), \tag{8}$$

and a parallel as a system whose structure function has the following form:

$$\phi(x) = \max(x_1, x_2, \ldots, x_n). \tag{9}$$

Structure function (or topology) of series and parallel MSSs can also be expressed using reliability block diagrams (Fig. 1), which represent one of the principal tools used in reliability analysis of BSSs [1]. The only difference is the fact that the series connections are evaluated using min-function (instead of Boolean AND) and the parallel using max-function (instead of Boolean OR).

Structure functions (8) and (9) can also be viewed as substructure functions of more complex systems that will be denoted as series-parallel systems (Fig. 2). One of the important properties of such systems is that they can be decomposed into modules, the modules into sub-modules, etc. until we obtain modules that have only series or parallel topology or that contain just one component. For example, in case of the system in Fig. 2, we can perform the following decomposition:

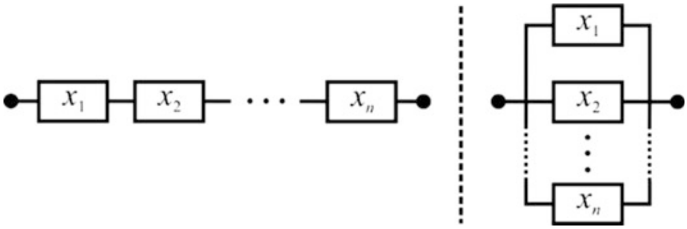
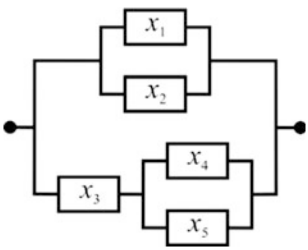


Fig. 1 Reliability block diagrams expressing topology of series and parallel MSSs

Fig. 2 Example of a series-parallel MSS



$\mathbf{N} = \{1, 2, 3, 4, 5\}$ and $\phi(x_1, x_2, x_3, x_4, x_5) = \psi(\phi^{\mathbf{N}_1}, \phi^{\mathbf{N}_2}) = \max(\phi^{\mathbf{N}_1}, \phi^{\mathbf{N}_2})$, where:

$\mathbf{N}_1 = \{1, 2\}$ and $\phi^{\mathbf{N}_1}(x_1, x_2) = \max(x_1, x_2)$
$\mathbf{N}_2 = \{3, 4, 5\}$ and $\phi^{\mathbf{N}_2}(x_3, x_4, x_5) = \psi^{\mathbf{N}_2}(\phi^{\mathbf{N}_{2,1}}, \phi^{\mathbf{N}_{2,2}}) = \min(\phi^{\mathbf{N}_{2,1}}, \phi^{\mathbf{N}_{2,2}})$, where:

$\mathbf{N}_{2,1} = \{3\}$ and $\phi^{\mathbf{N}_{2,1}}(x_3) = x_3$
$\mathbf{N}_{2,2} = \{4, 5\}$ and $\phi^{\mathbf{N}_{2,2}}(x_4, x_5) = \max(x_4, x_5)$

3 Topological Analysis of Multi-state Systems

Knowledge of the structure function without information about the state probabilities of the system components implies that we are not able to compute global characteristics, such as the system state probabilities or system availability. However, it would be nice to have some analogous measures to these characteristics that could be used to compare systems with different topologies and identify more reliable topologies. For this purpose, the relative frequencies of the system states:

$$Fr^{=j} = \mathrm{TD}(\phi(x) \leftrightarrow j), \quad \text{for } j \in \{0, 1, \ldots, m-1\}, \tag{10}$$

and the relative frequency of states greater than or equal to j:

$$Fr^{\geq j} = \mathrm{TD}(\phi(x) \geq j) = \sum_{h=j}^{m-1} Fr^{=h}, \text{ for } j \in \{1, 2, \ldots, m-1\}, \tag{11}$$

could be used. Please note that symbols $\leftrightarrow$ and $\geq$ denote logical biconditionals that are defined as follows:

$$\begin{aligned} \{\phi(x) \leftrightarrow j\} &= \begin{cases} 1 & \text{if } \phi(x) = j \\ 0 & \text{otherwise} \end{cases}, \\ \{\phi(x) \geq j\} = \vee_{h=j}^{m-1} \{\phi(x) \leftrightarrow h\} &= \begin{cases} 1 & \text{if } \phi(x) \geq j \\ 0 & \text{otherwise} \end{cases}, \end{aligned} \tag{12}$$

and TD(.) is the truth of the argument interpreted as a function with Boolean-valued output, i.e. the relative frequency of situations in which the function takes true value. It can be shown easily that measures (10) and (11) agree with the system state probabilities and the system availability respectively if the state probabilities of the system components have the following values:

Table 1 Structure functions of series and parallel MSS composed of 2 3-state components

Components states		$\phi_{\min}(x)$	$\phi_{\max}(x)$
x_1	x_2		
0	0	0	0
0	1	0	1
0	2	0	2
1	0	0	1
1	1	1	1
1	2	1	2
2	0	0	2
2	1	1	2
2	2	2	2

Table 2 The relative frequencies for the states of the MSSs defined in Table 1

System state (*j*)	$Fr_{\min}^{=j}$	$Fr_{\max}^{=j}$	$Fr_{\min}^{\geq j}$	$Fr_{\max}^{\geq j}$
0	5/9	1/9	1	1
1	3/9	3/9	4/9	8/9
2	1/9	5/9	1/9	5/9

$$p_{i,s} = 1/m, \text{for } i = 1, 2, \ldots, n, \; s = 0, 1, \ldots, m-1. \tag{13}$$

For illustration of the proposed measures, let us consider two MSSs. The first one is a series system with structure function $\phi_{\min}(\boldsymbol{x})$ = min(x_1, x_2) and the second is a parallel with structure function $\phi_{\max}(\boldsymbol{x})$ = max(x_1, x_2) where $x_1, x_2 \in \{0, 1, 2\}$. The structure functions of these systems can be expressed in form of Table 1. Using this table, the relative frequency of system state 1 can be computed as follows:

$$\begin{aligned} Fr_{\min}^{=1} &= \mathrm{TD}(\phi_{\min}(x) \leftrightarrow 1) = \frac{3}{9} \approx 0.3333, \\ Fr_{\max}^{=1} &= \mathrm{TD}(\phi_{\max}(x) \leftrightarrow 1) = \frac{3}{9} \approx 0.3333. \end{aligned} \tag{14}$$

In the similar way, we can compute the relative frequencies of system states 0 and 2. These data are in the central part of Table 2. Based on them, we can assume that the parallel system is more reliable from topological point of view because every of its states greater than state 0 (i.e. the working states) is at least such frequent as the same state in the series system with structure function $\phi_{\min}(\boldsymbol{x})$. This is more obvious using measure (11), which focuses on the occurrence of all states greater than or equal to a specific system state (the right part of Table 2). Clearly, it has no sense to compute this measure for state 0 of the system because, regardless of system topology, this measure is always equal to number 1 and, therefore, it is not useful in comparing topological reliability of different systems.

3.1 Structural Importance Measures

The measures introduced above allow us to identify system states with the greatest occurrence ($Fr^{=j}$) or systems whose topology is more reliable for satisfying demand that requires system to operate at least in state j ($Fr^{\geq j}$). However, they do not allow investigating importance of the system components. For this task other indices have to be used. These indices are known as importance measures.

One of the most fundamental importance measures is Structural Importance (SI), which focuses on system topology. In [8, 10, 13], several versions of SI have been proposed. The common characteristic of all those versions is that they investigate importance of a specific component (or its state) by counting situations when the component (or its state) is critical for system operation, i.e. situations in which component degradation (improvement) results in system degradation (improvement). In [9], some other SI measures have been introduced and have been combined with those defined in [8, 10, 13] into one complex framework that allows investigating:

- topological importance of state s of component i for system state j $\left(\mathrm{SI}_{i,s}^{j\downarrow} \text{ for } s, j \in \{1, 2, \ldots, m-1\}\right)$,
- topological importance of state s of component i for the entire system $\left(\mathrm{SI}_{i,s}^{\downarrow} \text{ for } s \in \{1, 2, \ldots, m-1\}\right)$,
- the total topological importance of component i for system state j $\left(\mathrm{SI}_{i}^{j\downarrow} \text{ for } j \in \{1, 2, \ldots, m-1\}\right)$,
- the total topological importance of component i $\left(\mathrm{SI}_{i}^{\downarrow}\right)$.

(Please note that symbol ↓ used in the notation of the SI measures indicates that they analyze consequences of component degradation.) It has been shown in [9] that the most important version of SI is $\mathrm{SI}_{i,s}^{j\downarrow}$ because all other measures can be computed from it using the following formulae:

$$\mathrm{SI}_{i,s}^{\downarrow} = \sum_{j=1}^{m-1} \mathrm{SI}_{i,s}^{j\downarrow}, \quad \mathrm{SI}_{i}^{j\downarrow} = \frac{1}{m-1} \sum_{s=1}^{m-1} \mathrm{SI}_{i,s}^{j\downarrow}, \quad \mathrm{SI}_{i}^{\downarrow} = \frac{1}{m-1} \sum_{s=1}^{m-1} \mathrm{SI}_{i,s}^{\downarrow}. \tag{15}$$

The $\mathrm{SI}_{i,s}^{j\downarrow}$ has been defined in [9] as a relative frequency of situations in which a minor degradation (degradation by one state) of state s of component i results in degradation of state j of the system. These situations can be identified using a special tool of MVL, which is known as logical differential calculus.

3.2 Logical Differential Calculus

Logical differential calculus allows analyzing dynamic properties of MVL functions [14]. The central term of this tool is logic derivative. Several types of logic derivatives exist but, for the purpose of this paper, Direct Partial Logic Derivatives (DPLDs) and Integrated DPLDs (IDPLDs) are the most important ones.

A DPLD of function $\phi(\boldsymbol{x})$ with respect to variable x_i is defined as follows [14]:

$$\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \begin{cases} 1 & \text{if } \phi(s_i,x)=j \text{ and } \phi(r_i,x)=h \\ 0 & \text{otherwise} \end{cases}, \quad (16)$$
$$\text{for } s,r,j,h\in\{0,1,\ldots,m-1\},\ s\neq r, j\neq h,$$

where $(a_i,x)=(x_1,x_2,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)$ for $a\in\{s,r\}$ This definition implies that a DPLD is a function with Boolean-valued output whose nonzero elements agree with situations in which change of variable x_i from value s to r results in change in value of function $\phi(\boldsymbol{x})$ from j to h. In reliability analysis, the nonzero elements of the DPLD agree with state vectors $(._i, \boldsymbol{x}) = (x_1, x_2,\ldots, x_{i-1}, x_{i+1},\ldots, x_n)$ at which degradation/improvement of component i from state s to r results in degradation/improvement of the system from state j to h. Since the structure function of a coherent system is monotonic, only DPLDs in which $s > r$ and $j > h$ or $s < r$ and $j < h$ are relevant in reliability analysis of such systems [8].

DPLD $\partial\phi(j\to h)/\partial x_i(s\to s-1)$ allows finding situations in which a minor degradation of state s of component i results in degradation of system from state j to h. If we want to identify all situations in which the considered degradation of the i-th system component results in degradation of system state j, then j such derivatives have to be computed (for $h = 0,1,\ldots, j-1$). Since this can be quite time-consuming, other types of logic derivatives have been introduced in [9]. These derivatives combine several DPLDs together and, therefore, they were named as IDPLDs. Depending on which DPLDs are combined, three types of IDPLDs can be recognized. In this paper, the most useful are IDPLDs of type I.

An IDPLD of type I of function $\phi(\boldsymbol{x})$ with respect to variable x_i can be defined in the following manner [9]:

$$\frac{\partial\phi(j\downarrow)}{\partial x_i(s\to r)} = \bigvee_{h=0}^{j-1}\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \begin{cases} 1 & \text{if } \phi(s_i,x)=j \text{ and } \phi(r_i,x)<j \\ 0 & \text{otherwise} \end{cases}, \quad (17)$$
$$\text{for } s,r\in\{0,1,\ldots,m-1\}, s\neq r, j\in\{1,2,\ldots,m-1\}.$$

This derivative permits detecting situations in which degradation of system component i from state s to r results in degradation of system state j. Since detection of these situations plays a key role in computation of $\mathrm{SI}_{i,s}^{j\downarrow}$, IDPLDs of type I can be used in the following way to calculate this measure [9]:

$$\mathrm{SI}_{i,s}^{j\downarrow} = \mathrm{TD}\left(\frac{\partial\phi(j\downarrow)}{\partial x_i(s \to s-1)}\right). \tag{18}$$

Specially, if a minor degradation of component i can result only in a minor degradation of the system, then the previous formula will have the following form:

$$\mathrm{SI}_{i,s}^{j\downarrow} = \mathrm{TD}\left(\frac{\partial\phi(j \to j-1)}{\partial x_i(s \to s-1)}\right). \tag{19}$$

The key question is how to compute IDPLDs $\partial\phi(j\downarrow)/\partial x_i(s \to s-1)$ or DPLDs $\partial\phi(j \to j-1)/\partial x_i(s \to s-1)$ and their truth densities efficiently.

3.2.1 Chain Rule

Let us consider a MSS in which module ($\mathbf{B}$, $\phi^{\mathbf{B}}$) can be identified. This implies that the system structure function can be expressed in the following way:

$$\phi(x) = \psi\left(\phi^{\mathbf{B}}\left(x^{\mathbf{B}}\right), x^{\mathbf{B}^c}\right), \tag{20}$$

where $\psi\left(\phi^{\mathbf{B}}, x^{\mathbf{B}^c}\right)$ denotes function that combines substructure function $\phi^{\mathbf{B}}(x^{\mathbf{B}})$ with the components not contained in module ($\mathbf{B}$, $\phi^{\mathbf{B}}$) into the system structure function.

Next, let us compute DPLD $\partial\phi(j \to h)/\partial x_i(s \to r)$ assuming that the i-th system component is in module ($\mathbf{B}$, $\phi^{\mathbf{B}}$). According to (20), we can write:

$$\begin{aligned}\frac{\partial\phi(j \to h)}{\partial x_i(s \to r)} &= \{\phi(s_i, x) \leftrightarrow j\}\{\phi(r_i, x) \leftrightarrow h\} \\ &= \left\{\psi\left(\phi^{\mathbf{B}}\left((s_i, x)^{\mathbf{B}}\right), x^{\mathbf{B}^c}\right) \leftrightarrow j\right\}\left\{\psi\left(\phi^{\mathbf{B}}\left((r_i, x)^{\mathbf{B}}\right), x^{\mathbf{B}^c}\right) \leftrightarrow h\right\}.\end{aligned} \tag{21}$$

Since this is a function with Boolean-valued output, the Shannon expansion [15] can be applied to it. This can be done in the following way:

$$\begin{aligned}\frac{\partial\phi(j \to h)}{\partial x_i(s \to r)} &= \left\{\bigvee_{v=0}^{m-1}\left\{\phi^{\mathbf{B}}\left((s_i, x)^{\mathbf{B}}\right) \leftrightarrow v\right\}\left\{\psi\left(v, x^{\mathbf{B}^c}\right) \leftrightarrow j\right\}\right\} \\ &\wedge \left\{\bigvee_{w=0}^{m-1}\left\{\phi^{\mathbf{B}}\left((r_i, x)^{\mathbf{B}}\right) \leftrightarrow w\right\}\left\{\psi\left(w, x^{\mathbf{B}^c}\right) \leftrightarrow h\right\}\right\}.\end{aligned} \tag{22}$$

This formula can be rearranged in the following way:

$$\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \bigvee_{v=0}^{m-1}\bigvee_{w=0}^{m-1}\left(\begin{array}{c}\left\{\phi^{\mathbf{B}}\left((s_i,x)^{\mathbf{B}}\right)\leftrightarrow v\right\}\left\{\psi\left(v,x^{\mathbf{B}^c}\right)\leftrightarrow j\right\}\\ \wedge\left\{\phi^{\mathbf{B}}\left((r_i,x)^{\mathbf{B}}\right)\leftrightarrow w\right\}\left\{\psi\left(w,x^{\mathbf{B}^c}\right)\leftrightarrow h\right\}\end{array}\right). \quad (23)$$

Since $j \neq h$, at least one of the expressions $\psi\left(v,x^{\mathbf{B}^c}\right)\leftrightarrow j$ and $\psi\left(w,x^{\mathbf{B}^c}\right)\leftrightarrow h$ has to be false whenever v equals w. So, we can write:

$$\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \bigvee_{v=0}^{m-1}\bigvee_{\substack{w=0\\ w\neq v}}^{m-1}\left(\begin{array}{c}\left\{\psi\left(v,x^{\mathbf{B}^c}\right)\leftrightarrow j\right\}\left\{\psi\left(w,x^{\mathbf{B}^c}\right)\leftrightarrow h\right\}\\ \wedge\left\{\phi^{\mathbf{B}}\left((s_i,x)^{\mathbf{B}}\right)\leftrightarrow v\right\}\left\{\phi^{\mathbf{B}}\left((r_i,x)^{\mathbf{B}}\right)\leftrightarrow w\right\}\end{array}\right). \quad (24)$$

One can easily recognize that the following relations are valid:

$$\begin{aligned}\left\{\psi\left(v,x^{\mathbf{B}^c}\right)\leftrightarrow j\right\}\left\{\psi\left(w,x^{\mathbf{B}^c}\right)\leftrightarrow h\right\} &= \frac{\partial\psi(j\to h)}{\partial\phi^{\mathbf{B}}(v\to w)},\\ \left\{\phi^{\mathbf{B}}\left((s_i,x)^{\mathbf{B}}\right)\leftrightarrow v\right\}\left\{\phi^{\mathbf{B}}\left((r_i,x)^{\mathbf{B}}\right)\leftrightarrow w\right\} &= \frac{\partial\phi^{\mathbf{B}}(v\to w)}{\partial x_i(s\to r)},\end{aligned} \quad (25)$$

where $v \neq w$ and, therefore, the next formula holds:

$$\begin{gathered}\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \bigvee_{v=0}^{m-1}\bigvee_{\substack{w=0\\ w\neq v}}^{m-1}\frac{\partial\psi(j\to h)}{\partial\phi^{\mathbf{B}}(v\to w)}\frac{\partial\phi^{\mathbf{B}}(v\to w)}{\partial x_i(s\to r)},\\ \text{where } \phi(x)=\psi\left(\phi^{\mathbf{B}}(x^{\mathbf{B}}),x^{\mathbf{B}^c}\right) \text{ and } i\in\mathbf{B}.\end{gathered} \quad (26)$$

Using the same procedure, the following formula can be derived for a system that can be decomposed into k disjoint modules:

$$\begin{gathered}\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)} = \bigvee_{v=0}^{m-1}\bigvee_{\substack{w=0\\ w\neq v}}^{m-1}\frac{\partial\psi(j\to h)}{\partial\phi^{\mathbf{N}_l}(v\to w)}\frac{\partial\phi^{\mathbf{N}_l}(v\to w)}{\partial x_i(s\to r)},\\ \text{where } \phi(x)=\psi\left(\phi^{\mathbf{N}_1}(x^{\mathbf{N}_1}),\phi^{\mathbf{N}_2}(x^{\mathbf{N}_2}),\ \ldots,\phi^{\mathbf{N}_k}(x^{\mathbf{N}_k})\right)\\ \text{and } l\in\{1,2,\ \ldots,k\} \text{ such that } i\in\mathbf{N}_l.\end{gathered} \quad (27)$$

It can be noticed this expression is equivalent to the chain rule, which is well known in the classical calculus. Based on (24), it is possible to show this formula allows computing the truth density of a DPLD in the following way:

$$\mathrm{TD}\left(\frac{\partial\phi(j\to h)}{\partial x_i(s\to r)}\right)=\sum_{v=0}^{m-1}\sum_{\substack{w=0\\w\neq v}}^{m-1}\mathrm{TD}\left(\frac{\partial\psi(j\to h)}{\partial\phi^{\mathbf{N}_l}(v\to w)}\right)\mathrm{TD}\left(\frac{\partial\phi^{\mathbf{N}_l}(v\to w)}{\partial x_i(s\to r)}\right). \tag{28}$$

4 Hand Calculation Example

For illustration of the formulae developed above, let us consider the series-parallel system from Fig. 2. Using the decomposition presented at the end of Sect. 2.2, its structure function can be expressed in the form of Table 3.

This table consists of 4 sub-tables that describe behavior of the modules identified in the system. For example, the first sub-table implies the system is in state 0 if modules $(\mathbf{N}_1,\phi^{\mathbf{N}_1})$ and $(\mathbf{N}_2,\phi^{\mathbf{N}_2})$ are in state 0. Situations when these modules are in state 0 can be identified in the second and third sub-tables, etc.

Table 3 can be used to compute global topological characteristics, such as $Fr^{=j}$ and $Fr^{\geq j}$, in a fast recursive way. For example, $Fr^{=0}$ can be calculated as follows:

$$\begin{aligned}Fr^{=0}&=\mathrm{TD}(\phi(x)\leftrightarrow 0)=\mathrm{TD}\big(\psi(\phi^{\mathbf{N}_1},\phi^{\mathbf{N}_2})\leftrightarrow 0\big)\\&=\mathrm{TD}\big(\{\phi^{\mathbf{N}_1}(x_1,x_2)\leftrightarrow 0\}\{\phi^{\mathbf{N}_2}(x_3,x_4,x_5)\leftrightarrow 0\}\big)\\&=\mathrm{TD}\big(\phi^{\mathbf{N}_1}(x_1,x_2)\leftrightarrow 0\big)\mathrm{TD}\big(\phi^{\mathbf{N}_2}(x_3,x_4,x_5)\leftrightarrow 0\big).\end{aligned} \tag{29}$$

Now, the truth density of function $\phi^{\mathbf{N}_1}(x_1,x_2)\leftrightarrow 0$ can be computed directly from the second sub-table as 1/9 and the truth density of function $\phi^{\mathbf{N}_2}(x_3,x_4,x_5)\leftrightarrow 0$ can be calculated using the third sub-table in the following way:

$$\begin{aligned}\mathrm{TD}\big(\phi^{\mathbf{N}_2}(x_3,x_4,x_5)\leftrightarrow 0\big)&=\mathrm{TD}\big(\psi^{\mathbf{N}_2}(\phi^{\mathbf{N}_{2,1}},\phi^{\mathbf{N}_{2,2}})\leftrightarrow 0\big)=\mathrm{TD}\big(\psi^{\mathbf{N}_2}(x_3,\phi^{\mathbf{N}_{2,2}})\leftrightarrow 0\big)\\&=\mathrm{TD}(x_3\leftrightarrow 0)+\mathrm{TD}(x_3\leftrightarrow 1)\mathrm{TD}\big(\phi^{\mathbf{N}_{2,2}}(x_4,x_5)\leftrightarrow 0\big)\\&\quad+\mathrm{TD}(x_3\leftrightarrow 2)\mathrm{TD}\big(\phi^{\mathbf{N}_{2,2}}(x_4,x_5)\leftrightarrow 0\big)\\&=\frac{1}{3}+\frac{1}{3}\mathrm{TD}\big(\phi^{\mathbf{N}_{2,2}}(x_4,x_5)\leftrightarrow 0\big)+\frac{1}{3}\mathrm{TD}\big(\phi^{\mathbf{N}_{2,2}}(x_4,x_5)\leftrightarrow 0\big).\end{aligned} \tag{30}$$

Table 3 Structure and substructure functions of modules of the system depicted in Fig. 2

$(\mathbf{N}, \phi)$			$(\mathbf{N}_1, \phi^{\mathbf{N}_1})$			$(\mathbf{N}_2, \phi^{\mathbf{N}_2})$			$(\mathbf{N}_{2,2}, \phi^{\mathbf{N}_{2,2}})$		
$\phi^{\mathbf{N}_1}$	$\phi^{\mathbf{N}_2}$	$\psi(\phi^{\mathbf{N}_1}, \phi^{\mathbf{N}_2})$	x_1	x_2	$\phi^{\mathbf{N}_1}(x^{\mathbf{N}_1})$	$\phi^{\mathbf{N}_{2,1}} = x_3$	$\phi^{\mathbf{N}_{2,2}}$	$\psi^{\mathbf{N}_2}(\phi^{\mathbf{N}_{2,1}}, \phi^{\mathbf{N}_{2,2}})$	x_4	x_5	$\phi^{\mathbf{N}_{2,2}}(x^{\mathbf{N}_{2,2}})$
0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	0	1	1	0	1	0	0	1	1
0	2	2	0	2	2	0	2	0	0	2	2
1	0	1	1	0	1	1	0	0	1	0	1
1	1	1	1	1	1	1	1	1	1	1	1
1	2	2	1	2	2	1	2	1	1	2	2
2	0	2	2	0	2	2	0	0	2	0	2
2	1	2	2	1	2	2	1	1	2	1	2
2	2	2	2	2	2	2	2	2	2	2	2

Table 4 Truth densities of states of the system in Fig. 2 and its modules

j	$Fr^{=j} = \mathrm{TD}(\phi(x) \leftrightarrow j)$	$\mathrm{TD}\left(\phi^{\mathbf{N}_1}(x^{\mathbf{N}_1}) \leftrightarrow j\right)$	$\mathrm{TD}\left(\phi^{\mathbf{N}_2}(x^{\mathbf{N}_2}) \leftrightarrow j\right)$	$\mathrm{TD}\left(\phi^{\mathbf{N}_{2,2}}(x^{\mathbf{N}_{2,2}}) \leftrightarrow j\right)$
0	11/243	1/9	11/27	1/9
1	77/243	3/9	11/27	3/9
2	155/243	5/9	5/27	5/9

Finally, the truth density of function $\phi^{\mathbf{N}_{2,2}}(x_4, x_5) \leftrightarrow 0$ can be computed directly from the fourth sub-table as 1/9. Because of that, we can write:

$$\mathrm{TD}\left(\phi^{\mathbf{N}_2}(x_3, x_4, x_5) \leftrightarrow 0\right) = \frac{1}{3} + \frac{1}{3}\frac{1}{9} + \frac{1}{3}\frac{1}{9} = \frac{11}{27} \tag{31}$$

and, based on (29), the relative frequency of system state 0 can be computed in the following manner:

$$Fr^{=0} = \frac{1}{9}\frac{11}{27} = \frac{11}{243}. \tag{32}$$

In the similar way, we can compute the relative frequencies of the other system states (Table 4) and, based on them, topological characteristic $Fr^{\geq 1}$ and $Fr^{\geq 2}$ can be calculated simply using (11). As we can see from Table 4, the relative frequencies of system states 1 and 2 are much greater than the relative frequency of state 0, what implies that the system is very reliable from topological point of view.

The previous example illustrates a recursive method that can be used to investigate topology of systems decomposable into modules. Based on this method, the sub-modules composed only of components are evaluated firstly. Then, the modules composed of the evaluated sub-modules and other components are evaluated. This process is repeated until the truth densities of the system states are computed. These truth densities are then used to compute the global topological characteristics, such as $Fr^{=j}$ and $Fr^{\geq j}$.

In the next phase, let us focus on topological importance of system components. According to Sect. 3.1, this can be done using SI measures. The most fundamental one is $\mathrm{SI}_{i,s}^{j\downarrow}$, which quantifies topological influence of state s of component i on system state j. This measure can be computed using DPLDs or IDPLDs. Using Table 3, it is possible to show that a minor degradation of state s of any component in the system can result only in a minor degradation of system state s and, therefore, the SI measure can be computed using the next formula:

$$\mathrm{SI}_{i,s}^{j\downarrow} = \mathrm{TD}\left(\frac{\partial\phi(j \to j-1)}{\partial x_i(s \to s-1)}\right) = \begin{cases} \mathrm{TD}\left(\frac{\partial\phi(s \to s-1)}{\partial x_i(s \to s-1)}\right) & \text{if } s = j \\ 0 & \text{if } s \neq j \end{cases}. \tag{33}$$

Since the system can be decomposed into the modules, the truth density of DPLD $\partial\phi(s \to s-1)/\partial x_i(s \to s-1)$ can be computed using (28).

For illustration, let us compute $SI_{4,1}^{1\downarrow}$. Component 4 is a part of module $(\mathbf{N}_2, \phi^{\mathbf{N}_2})$ and, therefore, this measure can be computed as follows:

$$SI_{4,1}^{1\downarrow} = TD\left(\frac{\partial\phi(1\to 0)}{\partial x_4(1\to 0)}\right) = \sum_{v=0}^{2} \sum_{\substack{w=0 \\ w\neq v}}^{2} TD\left(\frac{\partial\psi(1\to 0)}{\partial\phi^{\mathbf{N}_2}(v\to w)}\right) TD\left(\frac{\partial\phi^{\mathbf{N}_2}(v\to w)}{\partial x_4(1\to 0)}\right). \quad (34)$$

With respect to Table 3, change of the system structure function from value 1 to 0 can result only from change of function $\phi^{\mathbf{N}_2}(x_3, x_4, x_5)$ from value 1 to 0. This implies the previous formula can be transformed into the following form:

$$SI_{4,1}^{1\downarrow} = TD\left(\frac{\partial\phi(1\to 0)}{\partial x_4(1\to 0)}\right) = TD\left(\frac{\partial\psi(1\to 0)}{\partial\phi^{\mathbf{N}_2}(1\to 0)}\right) TD\left(\frac{\partial\phi^{\mathbf{N}_2}(1\to 0)}{\partial x_4(1\to 0)}\right). \quad (35)$$

DPLD $\partial\psi(1\to 0)/\partial\phi^{\mathbf{N}_2}(1\to 0)$ can be calculated simply based on the first sub-table in Table 3 as $\phi^{\mathbf{N}1}(x_1, x_2) \leftrightarrow 0$, i.e. degradation of module $(\mathbf{N}_2, \phi^{\mathbf{N}_2})$ from state 1 to 0 causes degradation of the system from state 1 to 0 if module $(\mathbf{N}_1, \phi^{\mathbf{N}_1})$ is in state 0. With respect to Table 4, the truth density of this DPLD equals 1/9. The truth density of DPLD $\partial\phi^{\mathbf{N}_2}(1\to 0)/\partial x_4(1\to 0)$ can be computed using the chain rule too:

$$TD\left(\frac{\partial\phi^{\mathbf{N}_2}(1\to 0)}{\partial x_4(1\to 0)}\right) = TD\left(\frac{\partial\psi^{\mathbf{N}_2}(1\to 0)}{\partial\phi^{\mathbf{N}_{2,2}}(1\to 0)}\right) TD\left(\frac{\partial\phi^{\mathbf{N}_{2,2}}(1\to 0)}{\partial x_4(1\to 0)}\right). \quad (36)$$

All DPLDs used in computation of $SI_{4,1}^{1\downarrow}$ and their truth densities are presented in Table 5. Based on these data, we can write:

$$SI_{4,1}^{1\downarrow} = TD\left(\frac{\partial\psi(1\to 0)}{\partial\phi^{\mathbf{N}_2}(1\to 0)}\right) TD\left(\frac{\partial\psi^{\mathbf{N}_2}(1\to 0)}{\partial\phi^{\mathbf{N}_{2,2}}(1\to 0)}\right) TD\left(\frac{\partial\phi^{\mathbf{N}_{2,2}}(1\to 0)}{\partial x_4(1\to 0)}\right) = \frac{2}{81}. \quad (37)$$

Table 5 Computation of DPLD $\partial\phi(1\to 0)/\partial x_4(1\to 0)$ for the system in Fig. 2 using the chain rule

$(\mathbf{N}, \phi)$		$(\mathbf{N}_2, \phi^{\mathbf{N}_2})$		$(\mathbf{N}_{2,2}, \phi^{\mathbf{N}_{2,2}})$	
$\phi^{\mathbf{N}_1}$	$\frac{\partial\psi(1\to 0)}{\partial\varphi^{\mathbf{N}_2}(1\to 0)}$	$\phi^{\mathbf{N}_{2,1}} = x_3$	$\frac{\partial\psi^{\mathbf{N}_2}(1\to 0)}{\partial\varphi^{\mathbf{N}_{2,2}}(1\to 0)}$	x_5	$\frac{\partial\varphi^{\mathbf{N}_{2,2}}(1\to 0)}{\partial x_4(1\to 0)}$
0	1	0	0	0	1
1	0	1	1	1	0
2	0	2	1	2	0
TD	1/9	TD	2/3	TD	1/3

Table 6 SI measures computed using DPLDs for the system depicted in Fig. 2

Component (i)	1	2	3	4	5
$SI_i^{\downarrow}$	55/162	55/162	28/162	10/162	10/162

This implies that a minor degradation of state 1 of component 4 results in degradation of system state 1 in 2 out of 81 possible scenarios in which the component can degrade in the considered way.

In the similar way, we can compute $SI_{4,2}^{2\downarrow}$ measure or $SI_{i,s}^{j\downarrow}$ for the other components. Based on them, SI measures (15) can be computed and used to compare the total topological importance of the system components. These data are presented in Table 6, and they imply that components 1 and 2 have the greatest influence on system operation from topological point of view and, therefore, they should be primarily taken into account in the next phases of reliability analysis.

5 Conclusion

Reliability analysis of MSSs is a complex task. In this paper, we focused on topological analysis. Such analysis investigates only system structure, i.e. it does not take the state probabilities of the system components into account. Based on its results, we can identify the most frequent system states or compare several topologies and identify those that are most reliable for satisfying demand that requires the system to operate at least in a specific state.

Another part of topological analysis is investigation of importance of the system components or their states. Such analysis can be done using SI measures that count situations in which the component (or its state) is critical for system operation. These situations can be identified using DPLDs or IDPLDs and quantified as the truth densities of the computed derivatives [9]. In this paper, we presented the new method for these computations. The method is based on a modular decomposition of a MSS [7] and application of divide and conquer paradigm. This paradigm was applied based on the chain rule (27) derived in this paper.

The approach for topological analysis of MSSs presented in this paper can be used to investigate the structure of a MSS of any type. The main factor that determines the computational complexity of this approach is complexity of the structure function and the possibility to apply a modular decomposition to it. Typical examples of systems that can be decomposed quite easily are series-parallel systems in which the series branches are modeled by min-function and the parallel branches by max-function. For such systems, the method can be used to find closed-form expressions for computing the topological characteristics considered in this paper. The similar approach has also been used in [16] where a method for structure analysis of logic circuits has been developed. The experiments performed in [16] showed that the method was able to investigate circuits composed of hundreds of components. This indicates that the approach for topological analysis of MSSs developed in this paper could be applied to systems composed of many components.

The presented approach could also be applied in the analysis of other kinds of MSSs, i.e. not only series-parallel ones. In this case, it is important to express the structure function of the investigated system in the form that can be processed on a computer efficiently. One of the promising approaches is application of multi-valued decision diagrams [17]. Several algorithms for computation of DPLDs, which are the key tool of the approach presented in this paper, have been considered in [18]. Those algorithms could be combined with our approach to develop an efficient method for topological analysis of complex MSSs. We want to investigate this idea more deeply in further work.

Acknowledgements This work was supported by grant VEGA 1/0038/16.

References

1. Rausand M, Høyland A (2004) System reliability theory. Wiley, Hoboken, NJ
2. Nystrom B, Austrin L, Ankarback N, Nilsson E (2006) Fault tree analysis of an aircraft electric power supply system to electrical actuators. In: International Conference on Probabilistic Methods Applied to Power Systems, PMAPS, pp 1–7
3. Watanabe Y, Oikawa T, Muramatsu K (2003) Development of the DQFM method to consider the effect of correlation of component failures in seismic PSA of nuclear power plant. Reliab Eng Syst Saf 79:265–279
4. Praks P, Kopustinskas V, Masera M (2015) Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure. Reliab Eng Syst Saf 144:254–264
5. Kvassay M, Zaitseva E, Levashenko V, Kostolny J (2017) Reliability analysis of multiple-outputs logic circuits based on structure function approach. IEEE Trans Comput Aided Des Integr Circuits Syst 36:398–411
6. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
7. Natvig B (2011) Multistate systems reliability theory with applications. Wiley, Chichester
8. Zaitseva E, Levashenko V (2013) Multiple-valued logic mathematical approaches for multi-state system reliability analysis. J Appl Logic 11:350–362
9. Kvassay M, Zaitseva E, Kostolny J, Levashenko V (2015) Importance analysis of multi-state systems based on integrated direct partial logic derivatives. In: 2015 International Conference on Information and Digital Technologies (IDT), pp 183–195
10. Butler D (1979) A complete importance ranking for components of binary coherent systems, with extensions to multi-state systems. Naval Res Logist Q 26:565–578
11. Blokus-Roszkowska A, Kwiatuszewska-Sarnecka B, Soszyńska-Budny J (2016) Reliability and risk improvement of the maritime ferry technical system. In: 2016 International Conference on Information and Digital Technologies (IDT), pp 40–47
12. Kolowrocki K (2014) Reliability of large and complex systems. Elsevier, London
13. Wu S (2005) Joint importance of multistate systems. Comput Ind Eng 49:63–75
14. Yanushkevich S, Miller D, Shmerko V, Stankovic R (2005) Decision diagram techniques for micro- and nanoelectronic design handbook. CRC Press, Boca Raton, FL
15. Crama Y, Hammer P (2011) Boolean functions: theory, algorithms, and applications. Cambridge University Press, New York, NY

16. Xing L, Dai Y (2009) A new decision-diagram-based method for efficient analysis on multistate systems. IEEE Trans Depend Secure Comput 6:161–174
17. Kostolny J, Kvassay M, Zaitseva E (2014) Analysis of algorithms for computation of direct partial logic derivatives in multiple-valued decision diagrams. In: 2014 Ninth International Conference on Availability, Reliability and Security, pp 356–361
18. Zio E (2009) Reliability engineering: old problems and new challenges. Reliab Eng Syst Saf 94:125–141

Part II
Applications and Case Studies

Short-Term Reliability Analysis of Power Plants with Several Combined Cycle Units

Anatoly Lisnianski, David Laredo and Hanoch Ben Haim

Abstract This chapter presents a method for a short-term reliability analysis of power plants consisting of a number combined cycle generating units. A multi-state Markov model represents each generating unit with several states. Using a straightforward Markov method for reliability assessment is leading to explosion of number of states that should be analyzed. The chapter proposes a method for the estimation of important power system indices such as the availability for specified demand level, the loss of load probability, the expected energy not supplied to consumers etc. This method is based on *Lz*-transform of discrete-state continuous-time Markov process. The proposed approach is useful for a short-term reliability analysis of power system and operative decisions making. A numerical example is presented as an illustration of the proposed approach.

Keywords Combine cycle power generating unit · Available generating capacity · Reliability · Multi-state Markov model · *Lz*-transform · Universal generating function

Acronyms

DSCT	Discrete-state continuous-time
MSS	Multi-state system
UGF	Universal generating function
CCGT	Combined cycle gas turbine

A. Lisnianski (✉) · D. Laredo · H.B. Haim
The Israel Electric Corporation, Haifa, Israel
e-mail: anatoly-l@iec.co.il; lisnianski@bezeqint.net

D. Laredo
e-mail: uu612@iec.co.il

H.B. Haim
e-mail: hanoch@iec.co.il

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_15

Nomenclature

G_{gas}	Nominal generating capacity of gas turbine generator
G_{steam}	Nominal generating capacity of steam turbine generator
G_{cc}	Nominal generating capacity of entire combine cycle unit $G_{cc} = G_{gas} + G_{steam}$
$G_i(t)$	Stochastic generating capacity of unit i at time t
$g_{i,k}$	k-th possible generating capacity level of unit i
P_{sg}	Probability of starting failure for gas turbine
P_{ss}	Probability of starting failure for steam turbine
f	System structure function
$G(t)$	Stochastic process of entire power system at time t
w	Power demand level for entire system
D	Average in service time per occasion of demand
T	Average reserve shutdown time between periods of need
MTTF_1	Average in-service time between occasions of forced outage that shutting down at gas turbine (failure of type 1)
MTTF_2	Average in-service between occasions of forced outage that shutting down at steam turbine or steam generator (failure of type 2)
r_1	Average repair time for occasions with MTTF_1
r_2	Average repair time for occasions with MTTF_2
$EENS$	Expected energy not supplied to consumers
a_{ij}	Intensity of transition from state i to state j.
$LOLP$	Loss of load probability

1 Introduction

Combined-cycle thermal power plants are more and more widespread. The reliability and availability of such power plants depend on the perfect operation of all its subsystems (e.g. Gas Turbine GT, Heat Recovery Steam Generator HRSG, Steam Turbine ST and cooling system). The HRSG is the link between the gas turbine and steam turbine process having the function of converting the exhaust gas energy of the gas turbine into the steam. The basic model for a generating unit in power systems is a two-state representation where the unit resides (stays) either in operating state or in forced outage state [1]. In [3] it has been proven that using simple two-state models for large generating units in generating capacity adequacy assessment can yield pessimistic appraisals. It means that all types of models (block diagram, Markov models, simulation) for adequacy assessment that are based on two-state units' representation can yield pessimistic appraisals. In order to assess power system reliability more accurately many utilities now use multi-state models instead of two-state representations [3, 15].

A general Multi-state Markov model for coal fired generating unit was suggested in [11]. Based on this model such important unit's reliability indices as Forced Outage Rate (*FOR*), Equivalent Forced Outage Rate (*EFOR*) etc. were assessed for steady-state as well as for short-term unit behavior. In [11] it was shown that short-term reliability indices for coal fired units as: Loss of Load Probability (*LOLP*), Expected Energy Not Supplied (*EENS*) to consumers, etc., are substantially different from those found for a long-term reliability evaluation.

Combined cycle gas turbine (CCGT) generating unit may be used as a base load unit as well as a peaking unit. Peaking units normally operate for short periods of time. A basic four—state model that was suggested in [2] is often used for reliability investigation in this case. In [5] Fazekas and Nagy considered the case where there are a number of CCGT units in power system and each unit is represented by a multi-state model. Because of huge dimension of the problem, only steady-state (long-term) probabilities were of interest in [5]. In [4] reliability indices were introduced for combined cycle power plant by using graph theoretic approach, but representation of unit generating capacity as a stochastic process was out of the paper scope.

A combined cycle unit has three levels of available generating capacity. The first level, when both gas turbine generator and steam turbine generator are available. The second level, when only gas turbine generator is available. The third level, when both gas turbine and steam turbine are not available. Unit generating capacity at this level is zero. In order to present for CCGT unit its available generating capacity as a stochastic process four-state model that has been considered in [2], was extended in [12]. As a result, an 8-states Markov model was suggested in this paper for a two-shaft CCGT power plant, which is a suitable representation for both base load unit and intermittent operating unit. It was shown that based on this model important unit's reliability indices as; Forced Outage Rate (*FOR*), Equivalent Forced Outage Rate (*EFOR*), Loss of Load Probability (*LOLP*) etc. can be calculated both for a short and long time period.

Usually, there is a number of CCGT units that are functioning in connected power system. In order to calculate the reliability indices for a number of CCGTs (where each CCGT is represented by an 8-state Markov model) by using straight forward Markov method one should build a Markov model with m states, where $m = 8^n$, n-number of CCGT units. Even for small n, for example, if $n = 4$ then $m = 4096$ states in Markov model for entire system that should be analyzed. Then a system of 4096 differential equations should be solved in order to find the reliability indices. This explosion of states (so-called "dimension curse") is the main obstacle in reliability assessment for such power plants. In order to avoid this obstacle an approach is suggested that is based on a modern mathematical technique—*Lz*-transform. *Lz*-transform was primarily introduced in [9]. The approach will be applied to calculation of short time reliability indices for an entire power system composed by several and various CCGT units. Instead of building and solving general Markov model for entire power system with enormous number of states, by using this approach one should build and solve relatively small Markov model with

8 states for each CCGT unit. Then by using *Lz*-transform all reliability indices for the entire power system can be found by simple algebra.

A numerical example illustrates the application of the approach to short term reliability analysis of a power system and corresponding benefits.

2 *Lz*-Transform Method and Its Application to Reliability Analysis of Power Plant Consisting of Number CCGT Generating Units

The *Lz*-transform method was primarily introduced in [9] where its detailed description and corresponding mathematical proof are presented. Recently there are some successive applications of *Lz*-transform method for determining age replacement policy in multi-state systems [16] and reliability analysis for different specific multi-state systems such as refrigerating system [7] and air conditioning system [6]. In [10] the method was applied to a short-term reliability analysis of a power system consisting of number coal fired generating units and in [8] *Lz*-transform was used for reliability evaluation of smart grid. CCGT units are different from coal fired generating units, so the method suggested in [10] should be substantially corrected in order to be applied to reliability analysis of power system consisting of CCGT units.

Here we present the method application to short-term reliability assessment for power system, based on several combine cycle generating units.

In order to apply the method, each CCGT should be presented by corresponding discrete-state continuous-time (DSCT) Markov process, which describes available generating capacity of the CCGT unit.

2.1 *A Multi-state Markov Model for a Combined Cycle Generating Unit and Lz-Transform for Its Output Generating Capacity Process*

A Markov model is usually represented by a set of states in which the generating unit can reside and transitions that reflect the capability of unit to transit from one state to another in accordance with certain actions. Such state-transition diagram mimics the operating behavior of a generating unit. According to the multi-state system approach, the set of unit states should be arranged in according to the unit generating capacity. For two-shaft combined cycle unit there are three possible capacity levels. At first level, when the two turbines and generators are working, the unit capacity is the sum of nominal generating capacity of gas-turbine generator and steam turbine generator. At second level, the steam turbine is not working; therefore, the unit generating capacity is characterized only by the generating capacity of

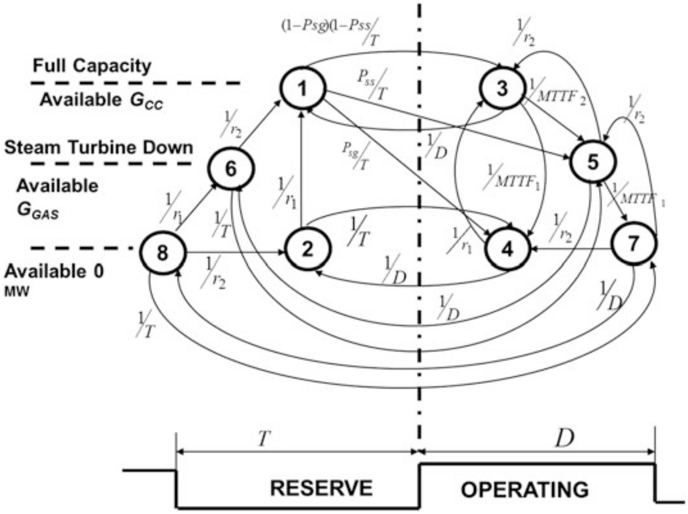


Fig. 1 Multi-state Markov model for a two-shaft CCGT unit

the gas turbine generator. When the gas turbine generator and the steam turbine generator are failed then the unit generating capacity is equal to zero. As it was said above we shall use the general 8-states Markov model that was introduced in [12] for two shaft combined cycle units. The corresponding state space diagram together with all possible transitions between states and transition intensities is presented in Fig. 1.

The model was built under the following assumptions:

- If the unit is in a reserve state (non-operating) then no failures may occur.
- The parameters Time to Failure (TTF), Time to Repair (TTR), Reserve Shutdown Time (RST) between periods of need, and average in service time per occasion of demand are exponentially distributed random variables.
- For a short-term reliability analysis only statistical data for unit's forced outages should be considered, when mean time to failures and repairs are determined.

State 1 is reserve shutdown state where the combined cycle unit doesn't operate, because it is not needed. All components of the unit are available in this state, so available generating capacity will be $G_1 = G_{cc}$. If a power demand occurs, then the unit transits from state 1 to state 3, if no starting failures occur nor in the GT or ST. State 3 is in-service state, where all the unit's components are available, therefore the available generating capacity in this state will be $G_3 = G_{cc}$ too. The intensity rate of transition from state 1 to state 3 will be as following $a_{13} = (1 - P_{ss})(1 - P_{sg})/T$.

If the unit is in state 3, it can transit back to state 1 with transition rate $a_{31} = 1/D$, when the unit being shut down ending its service time. If in the state 3 a forced outage occurred (failure) that shut down the gas turbine (GT), then the unit will transit to state 4 with transition rate $a_{34} = 1/MTTF_1$. In state 4 the unit's generating capacity is zero $G_4 = 0$. If in the state 3 a forced outage occurred (failure) that shut down the steam turbine (ST) or its steam generator (HRSG), then the unit will transit to state 5 with transition rate $a_{35} = 1/\mathrm{MTTF}_2$. In state 5 the gas turbine (GT) is working, so the generating capacity in this state is $G_5 = G_{gas}$.

From state 4 there are 2 possible transitions. The first one is the transition from state 4 back to state 3 with transition rate $a_{35} = 1/r_1$, when the unit has been repaired. The second one is the transition to state 2 with transition rate $a_{42} = 1/D$, when the unit shut down after a period of need before the repair will be completed. In state 2 available generating capacity of the unit is zero $G_2 = 0$. If the unit is repaired in state 2, it will transit to state 1.

From state 5 there are 3 possible transitions. With transition rate $a_{53} = 1/r_2$ the unit will be repaired and come back to state 3. With transition rate $a_{57} = 1/MTTR_1$ the unit will transit to state 7 after a gas turbine failure. With transition rate $a_{56} = 1/D$ the unit will transit to state 6, when the unit shut down after a period of need. In state 6 generator of gas turbine is working so the available generating capacity in this state is $\mathrm{G}_6 = G_{gas}$.

From state 7 there are three possible transitions. With transition rate $a_{75} = 1/r_1$ the unit will be repaired after a failure of type 1 and come back to state 5. With transition rate $a_{74} = 1/r_2$ the unit will be repaired after a failure of type 2 and comes to state 4. With transition rate $a_{78} = 1/D$ the unit will transit to state 8, when an operating period is finished. Available generating capacity in state 8 is zero, $G_8 = 0$, but there is no demand for the unit using in this state.

From state 8 there are three possible transitions. With transition rate $a_{87} = 1/T$ the unit will come back to state 8, when reserve shutdown period ended and the unit is needed again. With transition rate $a_{82} = 1/r_2$ the unit will be repaired after a failure of type 2 and comes to state 2. Available generating capacity in state 2 is zero, $G_2 = 0$, but there is no demand for the unit in this state. With transition rate $a_{86} = 1/\mathrm{r}_1$ the unit will be repaired after a failure of type 1 and comes to state 6. Available generating capacity in state 6, $G_6 = G_{gas}$, because gas turbine and its generator are available in this state.

From state 2 there are two possible transitions. With transition rate $a_{24} = 1/T$ the unit will transit to state 4, when reserve shutdown period ended and the unit is needed again. With transition rate $a_{21} = 1/r_1$ the unit will be repaired after a failure of type 1 and transit to state 1.

From state 6 there is only one possible transition. With transition rate $a_{61} = 1/r_2$ the unit will be repaired after a failure of type 2 and transit to state 1.

Therefore, the Multi-state Markov model has 8 states and 3 levels of available generating capacity: full capacity $G_{cc} = G_{gas} + G_{steam}$, intermediate capacity G_{gas}, when only the generator of gas turbine is available, and zero capacity.

We designate $p_i(t)$, $i = 1,\ldots, 8$, probability that at time instant t the process will be in state i, $i = 1,\ldots, 8$.

Based on the Markov processes theory [14, 17] the following system (1) of differential equations can be written for state probabilities $p_i(t)$, $i = 1,\ldots, 8$, of the Discrete-state Continuous-time (DSCT) Markov process, which state-transition diagram is depicted in Fig. 1.

Assume that at instant $t = 0$ the combine cycle generating unit is in state i, ($i = 1,\ldots, 8$). It means that system (1) should be solved under the following initial conditions

$$\mathbf{p_0} = \mathbf{p}(\mathbf{0}) = \{p_1(0) = 0, \ldots, p_i(0) = 1, \ldots, p_8(0) = 0\}.$$

After the system solving one will know all states probabilities $p_i(t)$, $i = 1,\ldots, 8$, as functions of time.

$$\begin{aligned}
\frac{dp_1(t)}{dt} &= a_{11}p_1(t) + a_{21}p_2(t) + a_{31}p_3(t) + a_{41}p_4(t) + a_{51}p_5(t) + a_{61}p_6(t) + a_{71}p_7(t) + a_{81}p_8(t)\\
\frac{dp_2(t)}{dt} &= a_{12}p_1(t) + a_{22}p_2(t) + a_{32}p_3(t) + a_{42}p_4(t) + a_{52}p_5(t) + a_{62}p_6(t) + a_{72}p_7(t) + a_{82}p_8(t)\\
\frac{dp_3(t)}{dt} &= a_{13}p_1(t) + a_{23}p_2(t) + a_{33}p_3(t) + a_{43}p_4(t) + a_{53}p_5(t) + a_{63}p_6(t) + a_{73}p_7(t) + a_{83}p_8(t)\\
\frac{dp_4(t)}{dt} &= a_{14}p_1(t) + a_{24}p_2(t) + a_{34}p_3(t) + a_{44}p_4(t) + a_{54}p_5(t) + a_{64}p_6(t) + a_{74}p_7(t) + a_{84}p_8(t)\\
\frac{dp_5(t)}{dt} &= a_{15}p_1(t) + a_{25}p_2(t) + a_{35}p_3(t) + a_{45}p_4(t) + a_{55}p_5(t) + a_{65}p_6(t) + a_{75}p_7(t) + a_{85}p_8(t)\\
\frac{dp_6(t)}{dt} &= a_{16}p_1(t) + a_{26}p_2(t) + a_{36}p_3(t) + a_{46}p_4(t) + a_{56}p_5(t) + a_{66}p_6(t) + a_{76}p_7(t) + a_{86}p_8(t)\\
\frac{dp_7(t)}{dt} &= a_{17}p_1(t) + a_{27}p_2(t) + a_{37}p_3(t) + a_{47}p_4(t) + a_{57}p_5(t) + a_{67}p_6(t) + a_{77}p_7(t) + a_{87}p_8(t)\\
\frac{dp_8(t)}{dt} &= a_{18}p_1(t) + a_{28}p_2(t) + a_{38}p_3(t) + a_{48}p_4(t) + a_{58}p_5(t) + a_{68}p_6(t) + a_{78}p_7(t) + a_{88}p_8(t)
\end{aligned} \tag{1}$$

When the unit is in states 1 or 3 its available capacity will be equal to G_{cc}. When the unit will be in state 5 or 6 its available capacity will equal G_{gas}, and when the unit is states 2, 4, 7, or 8 its available capacity will be zero. Therefore, output stochastic generating capacity process G(t) has 3 possible capacity levels: G_{cc}, G_{gas}, 0. In according to [9] Lz-transform of a DSCT Markov process $G(t) = \{g_1,\ldots,g_K\}$ is a function defined as follows

$$Lz\{G(t)\} = \sum_{k=1}^{K} p_k(t) z^{g_k}, \tag{2}$$

where $p_k(t)$, is a probability that the process is in state k with performance g_k at time instant $t \geq 0$ for a given initial states probability distribution $\mathbf{p_0}$, K is the total number of states and z in general case is a complex variable.

For CCGT unit these probabilities are finding by solving system (1) of differential equations under specified initial conditions.

Any combine cycle generating unit in entire power system can have 3 states corresponding to different generating capacity levels, represented by the set $\mathbf{g} = \{g_1, g_2, g_3\} = \left\{G_{cc}, G_{gas}, 0\right\}$.

Therefore, the corresponding *Lz*-transform of the DSCT Markov process $G(t)$, which represent available generating capacity for CCGT unit can be written as the following

$$Lz\{G(t)\} = [p_1(t) + p_3(t)]Z^{G_{cc}} + [p_5(t) + p_6(t)]z^{G_{gas}} + [p_2(t) + p_4(t) + p_7(t) + p_8(t)]z^0 \quad (3)$$

Note, that in our case, where combine cycle unit is considered, there are only three terms in polynomial form (3) (K = 3 in expression (2)), which define three different possible generating capacity levels of combine cycle unit: 0, G_{gas}, G_{cc}.

So, at this stage a Markov model for generating capacity stochastic process should be built for each CCGT generating unit in power system. Based on this model state probabilities for each generating unit are obtained as a solution of the corresponding system of differential equations (1) under given initial conditions for each unit.

Then individual *Lz*-transform (3) for each generating unit j can be obtained.

2.2 Reliability Analysis for Power System Consisting of Number Combine Cycle Generating Units

We consider the power system, which consists of n independent CCGT units, that is depicted in Fig. 2.

In this case the stochastic process $G(t)$ that represents available generating capacity of the entire power system is the sum of generating capacity processes $G_i(t)$, i = 1, … n of all n units in the system. Therefore, we have

$$G(t) = f\{G_1(t), G_2(t), \ldots, G_n(t)\} = \sum_{i=1}^{n} G_i(t). \quad (4)$$

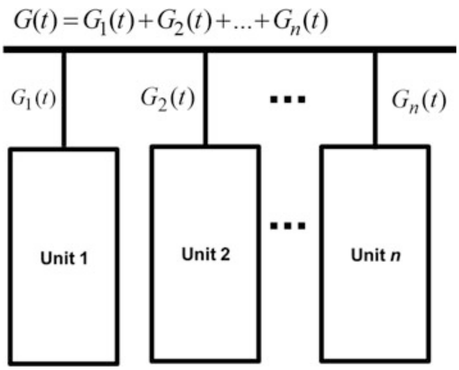


Fig. 2 Power system consisting of n CCGT units

In [9] it was shown that in order to find the Lz-transform of the resulting DSCT Markov process $G(t)$, which is the single-valued function $G(t) = f\{G_1(t), G_2(t),\ldots, G_n(t)\}$ of n independent DSCT Markov processes $G_j(t)$, $j = 1,\ldots, n$, one can apply Ushakov's Universal Generating Operator (UGO) to all individual Lz-transforms $Lz\{G_i(t)\}$ over all time points $t \geq 0$

$$Lz\{G(t)\} = \Omega_f\{Lz[G_1(t)], \ldots, Lz[G_n(t)]\}. \tag{5}$$

So, by using Ushakov's operator Ω_f over all Lz-transforms of individual generating units one can obtain the resulting Lz-transform (5) associated with output generating capacity stochastic process $G(t)$ of the entire power system.

The technique of Ushakov's operator applying is well established for many different structure functions f [13]. In our case, when f is the sum, the resulting Lz-transform of output generating capacity process $G(t)$ can be obtained by multiplying individual Lz-transforms of generating capacity processes of all generating units [13]

$$Lz[G(t)] = Lz[G_1(t)] \cdot \ldots \cdot Lz[G_n(t)] = \sum_{i=1}^{K} p_i(t) z^{g_i} \tag{6}$$

Generally each Lz-transform in polynomial form $Lz[G_i(t)]$, $i = 1, \ldots n$, has three terms (see expression (3)). So, after multiplying the resulting polynomial form in expression (6) will have $3n$ terms. But after combining similar terms (summarizing terms with same powers of z) the number K of terms in the resulting polynomial is usually less than $3n$.

Important reliability measures for the entire power system can be easily derived from the resulting Lz-transform (6). We assume that the demand level w is defined as a global level of load that should be met by the entire system (set of all combine cycle units).

The power system availability $A_w(t)$ regarding to demand level w at instant $t \geq 0$ may be obtained by

$$A_w(t) = \sum_{g_i \geq w} p_i(t). \tag{7}$$

In other words, in order to find MSS's instantaneous availability one should summarize all probabilities in Lz-transform from terms where powers of z are greater or equal to demand w.

An important reliability index that is used for power system is loss of load probability ($LOLP$)

$$LOLP(t) = \sum_{g_i = 0} p_i(t), \tag{8}$$

which is the sum of probabilities of all states where power system generating capacity is zero.

The system's Expected Generating Capacity Deficiency (ECD) is expressed by the following function:

$$ECD_w(t) = \sum_{k=1}^{K} p_k(t)(w - y_k)1(w - y_k), \tag{9}$$

where

$$1(w - y_k) = 1, \quad \textit{if } w - y_k > 0,$$
$$1(w - y_k) = 0, \quad \textit{if } w - y_k \leq 0.$$

It means that only such states of the entire system where the system generating capacity is lower than demand level ($y_k < w$) should be taken into account, when ECD is calculated.

Based on the calculated functions $ECD_w(t)$, the Expected Energy Not Supplied ($EENS_w$) to consumers during time t can be computed:

$$EENS_w(t) = \int_0^t ECD(u)du. \tag{10}$$

All reliability measures strongly depend on the initial conditions, under which the system of differential equations for each generating unit should be solved. In other words, short-term reliability measures for power system depend on initial states of all its units as well as on required demand level w.

3 Numerical Example

We consider a power system consisting of 3 independent combine cycle generating units. Each generating unit i, $i = 1, 2, 3$ is described by discrete-state continuous-time Markov process $G_i(t)$. The number of states is 8 and g_{ij} is a generating capacity of unit i in state j, $j = 1, 2, \ldots, 8$.

Available generating capacity in states 1 and 3 of each CCGT unit is $G_{cc} =$ 370 MW. So, we can write

$$g_{ij} = 370\,\text{MW}, \text{ for } i = 1, 2, 3 \text{ and } j = 1, 3. \tag{11}$$

Available generating capacity in states 5 and 6 for each CCGT unit is G_{gas} = 235 MW.

So, we have

$$g_{ij} = 235\,\text{MW}, \text{ for } i = 1, 2, 3 \text{ and } j = 5, 6. \tag{12}$$

Table 1 CCGT units parameters

j unit number	1/MTTF h^{-1}	$1/r_1$ h^{-1}	$1/MTTF_2$ h^{-1}	$1/r_2$ h^{-1}	1/D h^{-1}	1/T h^{-1}
1	0.0037	0.042	0.0014	0.042	0.0014	0.125
2	0.0045	0.039	0.0012	0.044	0.0014	0.125
3	0.0035	0.043	0.0015	0.039	0.0014	0.125

Available generating capacity in states 2, 4, 7 and 8 for each CCGT unit is *ZERO*.

So, we have

$$g_{ij} = 0\,\text{MW}, \text{ for } i = 1, 2, 3 \text{ and } j = 2, 4, 7, 8. \tag{13}$$

Units' parameters are presented in Table 1.

The problem is to calculate reliability indices $A_w(t)$, $LOLP(t)$, $ECD_w(t)$ for the entire power system.

Based on data from Table 1, all transition intensities $a_{cd}^{(j)}$ from any state c to any state $d \neq c$ of unit j, j = 1, 2, 3 may be calculated in according to Sect. 2.1 of the chapter.

Then system (1) should be written and solved for each generating unit. The solution should be performed under initial conditions that should be specified for each unit.

Here we shall use the following designations:

Initial conditions [i, j, k] means that at time instant t = 0 the unit 1 is in state i, unit 2 is in state j and unit 3 is in state k.

We shall measure generating capacity in any state and demand level w in relative units. Full (total) nominal generating capacity G_{cc} of one CCGT unit, which physically is equal to 370 MW, will be equal to 1 in our computation.

After solving all 3 systems of differential equations the Lz-transform for available generating capacity will be obtained for each unit in according to expression (2).

Then Lz-transform for the entire power system can be obtained by multiplying all 3 Lz-transforms of each unit by using expression (6).

At last stage all short-term reliability measures for the power system that we are interested in can be derived from this Lz-transform in according to expressions (7)–(10).

In Fig. 3 we present the power system availability for demand level w = 2 calculated for different initial conditions. From this figure one can see that long-term availability of the power system consisting of three CCGT units with total installed generating capacity $3G_{cc}$ = 3 for demand level w = 2 is sufficiently high—$A_{w=2}(\infty) = 0.993$.

However, if, for example, two units are in perfect state 3 at instant t = 0 and one of the units (unit 3 in our example) will be at t = 0 in state 4 with zero available capacity (because of gas turbine failure), then short-term availability will drop under level 0.986 during the first 20 h. If at t = 0 unit 3 will be in state 7 the

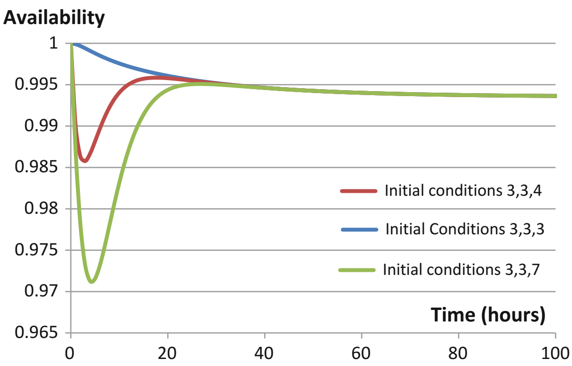


Fig. 3 Power system availability for demand $w = 2$

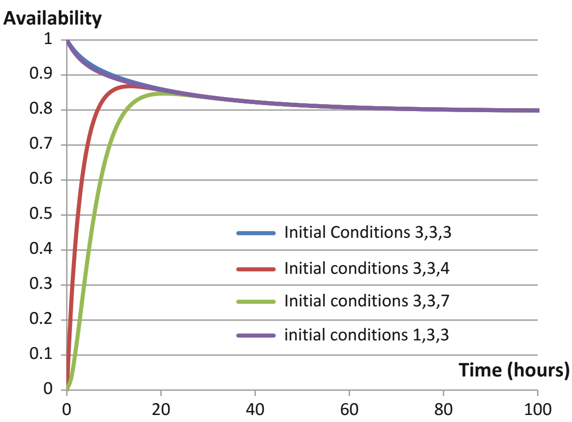


Fig. 4 Power system availability for demand $w = 3$

situation will be even worth since the availability will drop down almost to 0.97, because in state 7 gas turbine and steam turbine have failed.

Therefore, in short-term period there are substantial risks of the power system unavailability even for demand w = 2*, where only* 2 out of 3 *GGCT units installed in the power system can satisfied the demand. These risks cannot be estimated from long-term reliability analysis.*

In Fig. 4 the power system availability for demand $w = 3$ (when all 3 units are required) is shown.

It can be seen that long-term availability $A_{w=3}(\infty) = 0.79$ which is obviously much less than long-term availability in previous case. Short-term availability also strongly depends on initial conditions. If one of the units at instant $t = 0$ is in failure state 4 or 7 (available generating capacity in both these states is zero), then the power system availability at instant $t = 0$ is zero—$A_{w=3}(0) = 0$.

Curve $A_{w=3}(t)$ for initial conditions [3, 3, 7] is under the curve $A_{w=3}(t)$ for initial conditions [3, 3, 4] because complete repair CCGT unit from state 7 is more difficult (requires repair both gas turbine and steam turbine) than from state 4 (where only gas turbine should be repaired).

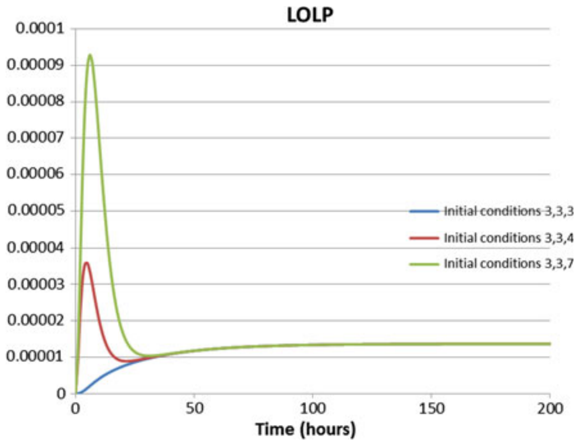


Fig. 5 LOLP for power system

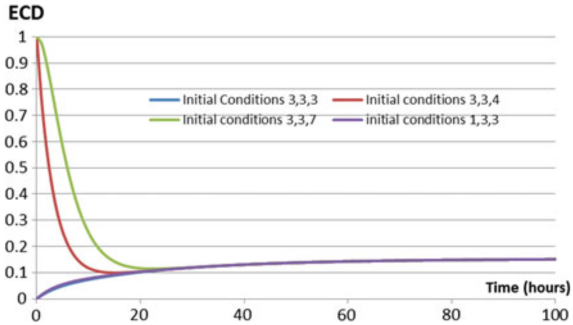


Fig. 6 Expected capacity deficiency (ECD) for demand w = 3

The curve under initial conditions [1, 3, 3] and curve under initial conditions [3, 3, 3] are almost the same. In Fig. 4 these curves are positioned so close together that the difference between them cannot be distinguished visually. This illustrates the fact that for such operating mode (when CCGT unit should start approximately one time per month), probability of starting failure has no great impact on power system availability.

In Fig. 5 we provide the power system *LOLP* for different initial conditions.

As one can see long-term $LOLP(\infty) = 0.000013$ is much less than short-term LOLP that may be $LOLP(t = 12\,\mathrm{h}) = 0.000092$ for initial conditions [3, 3, 7]. As in previous cases for availability computation, the worth (maximal) *LOLP* will be for initial conditions [3, 3, 7].

In Fig. 6 we present the power system expected capacity deficiency ECD(t) for different initial conditions of units. The required demand level is w = 3. It can be seen that long-term expected capacity deficiency is $A_{w=3}(\infty) = 0.16$. The curves under initial conditions [3, 3, 3] and [1, 3, 3] are so close together that the difference between them cannot be visible. As it was in the availability computation, it illustrates the fact that for such operating mode (when CCGT unit should start approximately one time per month), probability of starting failure has no a great impact on the power system *ECD*.

Expected energy not supplied to consumers can be obtained as a square under the corresponding curve $ECD_{w=3}$ (t) according to expression (9). Integral (10) is considered for t = 80 h, because after this time the process will be in steady-state and there is no difference between curves.

EENS (t = 80 h) = 9.01 for initial conditions [3, 3, 3]; *EENS* (t = 80 h) = 9.08 for initial conditions [1, 3, 3]; *EENS* (t = 80 h) = 12.45 for initial conditions [3, 3, 4]; *EENS* (t = 80 h) = 14.71 for initial conditions [3, 3, 7].

4 Summary

The Chapter suggests an approach for a short-term reliability analysis of a power system composed by a number of non-identical combine cycle gas turbine units. Each unit is represented by a multi-state (8-state) Markov model. The approach is based on *Lz*-transform.

Evaluation of important reliability indices as power system availability, loss of load probability, expected capacity deficiency and expected energy not supplied to consumers are considered.

The application of the proposed method decreases drastically the computation burden as compared with the straightforward Markov method.

It is shown that short-term reliability indices are essentially different from long-term (steady-state) indices. So, operative decisions for power system cannot be made correctly when based on long-term reliability indices. The short-term reliability computation method suggested in this work may be used as a base for adequate control operation.

All short-term reliability indices for a power system strongly depend on initial states of all its units as well as on required demand level w.

References

1. Billinton R, Allan R (1996) Reliability evaluation of power system. Plenum Press, NY
2. Billinton R, Ge J (2004) A comparison of four-state generating unit reliability models for peaking units. IEEE Trans Power Syst 19(2):763–768
3. Billinton R, Li Y (2004) Incorporating multi-state unit models in composite system adequacy assessment. In: Proceedings of the international conference “Probabilistic Methods Applied to Power Systems” (PMAPS)
4. Dev N, Samsher Kachhwana SS, Attri R (2014) Development of reliability index for combined cycle power plant using graph theoretic approach. Ain Shams Eng J 5:193–203
5. Fazekas A, Nagy E (2011) Modeling of combined heat and power generating power plant units using markov processes with continuous time parameter and discrete state space. Electric Power Compon Syst 39:1031–1044

6. Frenkel I, Khvatskin L, Lisnianski A (2014) *Lz*-transform application to availability assessment of the air conditioning system with rental equipment working under seasonal weather conditions. J Inf Control Manage Syst 2(2):133–140
7. Frenkel I, Lisnianski A, Khvatskin L (2012) Availability assessment for aging refrigeration system by using *Lz*-transform. J Reliab Stat Stud 5(2):33–43
8. Jia H, Jin W, Ding Y, Song Y, Yu D (2017) Multi-state time-varying reliability evaluation of smart grid with flexible demand resources utilizing *Lz*-transform. In: Proceedings of the international conference on energy engineering and environmental protection (EEEP2016), IOP conference series: Earth and Environmental Science, vol 52 012011. doi:10.1088/1755-1315/52/1/012011
9. Lisnianski A (2012) *Lz*-transform for discrete-state continuous-time Markov process and its applications to multi-state system reliability. In: Recent advances in system reliability. Signatures, multi-state systems and statistical inference. Springer, London, pp 79–95
10. Lisnianski A, BenHaim H (2013) Short-term reliability evaluation for power stations by using *Lz*-transform. J Modern Power Syst Clean Energy 1(2):110–117
11. Lisnianski A, Elmakias D, Laredo D, Ben Haim H (2012) A multi-state Markov model for a short-term reliability analysis of a power generating unit. Reliab Eng Syst Saf 98(1):1–6
12. Lisnianski A, Laredo D, BenHaim H (2016) Multi-state Markov model for reliability analysis of a combined cycle gas turbine power plant. In: Proceedings of the second international symposium on stochastic models in reliability engineering, life science and operations management, February 2016 IEEE, Israel, Beer Sheba, pp 131–135. doi:10.1109/SMRLO.2016.31
13. Lisnianski A, Levitin G (2003) Multi-state system reliability. Assessment, optimization and applications. World Scientific, NY, London, Singapore
14. Modarres M, Kaminskiy M, Krivtsov V (1999) Reliability engineering and risk analysis. A Practical Guide, Marcel Dekker Inc, NY, Basel
15. Reshid M, Abd Majid M (2011) A multi-state reliability model for a gas fueled cogenerated power plant. J Appl Sci 11(11):1945–1951
16. Sheu S-H, Zhang Z (2013) An optimal age replacement policy for multi-state systems. IEEE Trans Reliab 62(3):722–733
17. Trivedi K (2002) Probability and statistics with reliability. Queing and computer science applications. Wiley, New York

Reliability Analysis of a Modified IEEE 6BUS RBTS Multi-state System

Thomas Markopoulos and Agapios N. Platis

Abstract In this chapter, we attempt to develop a stochastic model based on a modification of a standard energy system. Aiming to achieve a high level of reliability in the system, it is necessary to implement specific modifications that are necessary to improve the structure of the system, in order to meet the demanded requirements. This improvement is actually a restructuring of an IEEE 6 BUS RBTS system by using an alternative combination of its generation units that presents the lowest possible failure rates using the same kind of generators and maintaining the level of output specifications according to the minimum reliability requirements. Using Multi-state systems and Semi-Markov modeling, the final result is a modified system that presents more flexibility and operates in less uncertainty environment, leading to a better level of reliability.

Keywords Markov chains · Semi-Markov chains · Multi-state system · Power system

1 Introduction

Reliability is a timeless problem closely related with human activity since the 18th century and has passed through stages of evolution in the course of time [35]. The systematic study of reliability has taken place at the end of the 20th century due to dramatic increase of complexity of electric and electronic systems and the cost reduction of them. In case of power generating systems, the operational parameters are the frequency of interruptions and the expected time to repair the failure [11]. The reliability as a concept is incident to two states of a system, "operation" and

T. Markopoulos · A.N. Platis (✉)
School of Science and Technology, Hellenic Open University, Patra, Greece
e-mail: platis@aegean.gr

A.N. Platis
Department of Financial and Management Engineering,
University of the Aegean, Chios, Greece

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_16

“failure”. This approach works for relative simple systems. However, a question about its effectiveness arises when we need to apply it in more complex systems, leading to the development of multi state systems analysis in the middle of 1980’s [22]. This topic is considered one of the most pioneering in the research of reliability theory [18]. Although the basic operational states of a system are two, normal operation and failure, each state of operation could consist of more than one “sub-states” (e.g. 80% or 60%, etc.) [35]. According to the literature, the term “*multi state system*” refers to a system that can be operating in a finite number of states [7]. Each complex system, consisting of a number of simple “two-state” sub-systems that have a cumulative effect on its performance can be considered a multi state system. The final performance of the whole system depends on the availability of the sub-systems and it is proportional with those that are operating [16].

In this study, we will consider failure of the system the non-acceptable level of operation due to specific requirements concerning the output level of the system, whereas some of the generators work normally and others fail. Since the failures are occurring events, and are related with independent systems (such as generators), we assume that they follow the Poisson distribution with parameter λ, whereas the time needed to repair a failed system follows the exponential distribution with a mean $1/\mu$. Referring to the IEEE 6 BUS RBTS, these parameters are depicted in Table 4. There is significant number of studies on multi state systems (MSS), because of their applicability especially in power generating and communication systems is broad [17]. Multi State Systems provide the advantage of flexibility and their representation is more accurate compared to the simple two state systems. On the other hand, their complexity holds the understanding and their performance evaluation back [33], e.g. there are systems that have hundreds or thousands of possible states. The development and the handling of such models results to depletion of the conventional Markov methods, emphasizing the need to apply innovative methods such as the EUGF [18]. The systems that are affected by the ageing of the materials with their maintenance effectiveness within limits confirm the need to combine the complexity and the flexible analysis capability [16]. Another advantage of multi-state systems is that they focus to the acceptable or non-acceptable level of operation on a specific time, instead of the “time to failure” of the simple systems, contributing to the analysis of applications closer to the real world [19], providing more accurate assessments [17] and a significant cut of the time required to develop an acceptable solution [4]. Concerning the power generating systems, all the above characteristics refer to the level of power available over a minimum acceptable level, in accordance with the requirements. An additional factor related to the reliability of power generating systems is the modeling of shocks that affect a system during its operation, perceiving these shocks in three major categories, such as the cumulative [10], the extreme [28] and the mixed [20]. All kinds of shocks mentioned above refer to a complete failure of the system. Actually, they could be an early variation of multi state systems, where the shock pushes the system from one critical state to another one, resulting to a partial failure of the system [7]. Due to their inherent complexity and the probable interaction among the existing subsystems, multi state systems present a dynamic behavior, when a sub system fails; the result on the whole system is much more severe [35]. Depending on the research requirements, the number of the states can be increase, leading to extreme complexity. This happens because it is possible to consider them as multi state systems and breaking them down in subsystems of lower level, making the need for further research intensified [18]. The mathematical form of a multi state system depicts the set of the possible states such as:

$$G_j = \{g_{j1}, g_{j2}, \ldots, g_{ji}, \ldots, g_{jk}\} \tag{1}$$

Where g_{ji} is the performance level of the subsystem j and $i \in \{1, 2, \ldots, k\}$ is the set of the possible states of the subsystems. Since we consider time of operation of the whole system, its state over time is a random variable included in a stochastic process [16]. Due to the stochastic nature of the states, there is a mean, a variance and a distribution of them as a random variable and this is precisely the importance to calculate the limits of operation of the system, in order to achieve a minimum level of reliability. One step further, the reliability function of a multi state system is:

$$R(t, w) = P\{G(t) \geq w\} \tag{2}$$

where $G(t)$ is the state of operation, at time t and w is the minimum required level of operation. The above equation leads to the separation of all states of the system in two groups. The first one is the set of acceptable states $\{d(w), d(w) + 1, \ldots, M\}$ and the second one $\{0, 1, \ldots, d(w) - 1\}$ is the set of non acceptable states [29]. The reliability function concerning the required level of operation describes the sum of the probabilities of all those acceptable states that are independent and is expressed by:

$$R(t, w) = P\{\Phi(t) \geq d(w)\} = \sum_{j=d(w)}^{M} p_j(t) \tag{3}$$

or with another expression:

$$R_{MSS}(t, W^*) = P[W(t) \geq W^*] \tag{4}$$

[14] where $W(t)$ is the level of output of the system at time t and W^* is the minimum required performance. Assuming that the states of non-acceptable operation level are equivalent to the state of failure of a simple system, the function of "state of failure" is:

$$F(t, w) = 1 - R(t, w) = \sum_{j=1}^{d(w)-1} p_j(t) \tag{5}$$

The Eq. (5) confirms the argument that there are more states of operation except for the normal operation and the complete failure [33]. Another parameter on the performance of the multi state systems is also the time the system spends in a state of operation. Assuming a transition of the system in $M + 1$ different states, where $M \in \{1, 2, 3, \ldots\}$ and $M \geq 1$. We denote that state *0* corresponds to complete failure of the system and state M reflects normal operation. Concerning the lifetime of the system in each state, we can consider the time the system lies in a state j or higher level of performance ($T \geq j$ for $j, j+1, \ldots, M$) [32]. Since $R_k(t)$ is the reliability function of the MSS (with discrete states) it is simply the probability of the system to operate in a state level higher than w at time t, this function is:

$$R_k(t) = \sum_{\phi=k}^{M} P\{\Phi(t) = \phi\}, k \in \{1, 2, 3, \ldots, M\} \tag{6}$$

with $0 \leq k \leq M$ [19]. Depending on the circumstances and because of their dynamic behavior, they proceed to states of partial operation until the state of total failure. Especially when the number of states is high, the transition is not always among consecutive states, but it is possible to omit intermediary states. The evaluation of a system is based on the assessment of reliability parameters, such as the rate the system downgrades. This parameter is related with all subsystems of the system and can be grouped in a matrix form such as

$$\Lambda = \begin{bmatrix} \lambda_{M,M-1} & \lambda_{M,M-2} & \cdots & \lambda_{M,0} \\ 0 & \lambda_{M-1,M-2} & \cdots & \lambda_{M-1,0} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_{0,0} \end{bmatrix} \tag{7}$$

With the probability the system to operate normally to be

$$P[\Phi(t) = M] = e^{[-(\lambda_{M,M-1} + \lambda_{M,M-2} + \ldots + \lambda_{M,0})t]} \tag{8}$$

the matrix of Eq. (7) can facilitate the understanding of the complexity that the researcher is possible to face, by understanding the system and the evaluation of its performance [33]. However, this complexity is useful because in many cases the two state systems often lead to erroneous and disappointing results [4]. One way to deal with the problem of the complexity is to break down the system in smaller subsystems and then to analyze them as multi state systems [15]. After all, multi state systems are a useful tool and a challenge to researchers, in order to solve different kinds of problems and research questions. It is important to understand that multi state systems are an approach in the field of probabilities modeling. Of course, there are other methodologies such as research on the optimal maintenance policy [1], or the Universal Generation function [14]. Also, Monte Carlo Simulation is a useful approach for complex systems, especially when restrictions of time exist

[23], and the recursive method, when the researcher evaluates the reliability of k out of n multi state systems [34].

Attempting to analyze and deal with the problems of power systems management, researchers are necessary to develop quantitative methods as useful tools in the decision-making. Depending on the specific needs of the problem, the system we examine, during its operation passes through certain states that could be values of random variables and the set of those states is called *state space* [16]. Using the continuous time methodology, we can study phenomena occurring in any time. The lack of memory in Markov chains implies a relationship between them and the exponential distribution which is the only one presenting this property. This property in Semi Markov models presents certain limitations concerning the time distribution that should be exponential, in case of continuous time and geometric, in case of discrete time. Especially in real world applications, these limitations could lead to erroneous conclusions [2, 13]. This means that if the system remains for certain time T in a state i, the probability to remain in this state for additional time is independent of the time T. This property is useful in the case of Semi Markov processes, where except for transient probabilities we consider the mean sojourn time for each state. This characteristic provides significant flexibility and the complexity of the calculations remains at relative low level. Semi Markov models are a generalized approach of Markov models providing an additional advantage of flexibility, concerning the distribution of the sojourn time in each state. Another characteristic of Semi Markov models is that the property of the lack of memory applies also in past states and the time the system was in those states as well. The difference comparing Semi Markov with Markov models is that time is a random variable and presumably the transition concerns only *different* states, because the probability remaining at the same state i is zero, since the system remains in this state for variable time (sojourn time). A general form of a Semi Markov Model can be represented mathematically as follows [2, 24]:

$$\begin{aligned} P(J_{n+1} = j, S_{n+1} - S_n = k | J_0, \ldots, J_n; S_0, \ldots, S_n) \\ = P(J_{n+1} = j, S_{n+1} - S_n = k | J_n) \end{aligned} \tag{9}$$

where J_n is the system state at the nth jump time and S_n is nth jump time. The embedded Markov chain associated to the jumps from one state to another of the previous Semi-Markov model is defined by its transition matrix, i.e. $p_{ij} = P(J_{n+1} = j \mid J_n = i)$.

Let $H_i(t)$ be the sojourn time distribution in state i. If $H_i(t)$ is assumed to have the exponential distribution for all i, then the previous Semi-Markov Model is simply an alternate description of a homogeneous CTMC [30].

In order now, to compute the steady state solution of the Semi-Markov Process, we first need to compute the steady state probability distribution of the previously defined embedded Markov chain by solving the following equation: $\boldsymbol{v} = \boldsymbol{v}\,\boldsymbol{P}$, with $\sum_i v_i = 1$, where v_i is the steady state probability for state i and $\boldsymbol{P}$ the transition probability matrix of the embedded Markov chain. Let additionally, define,

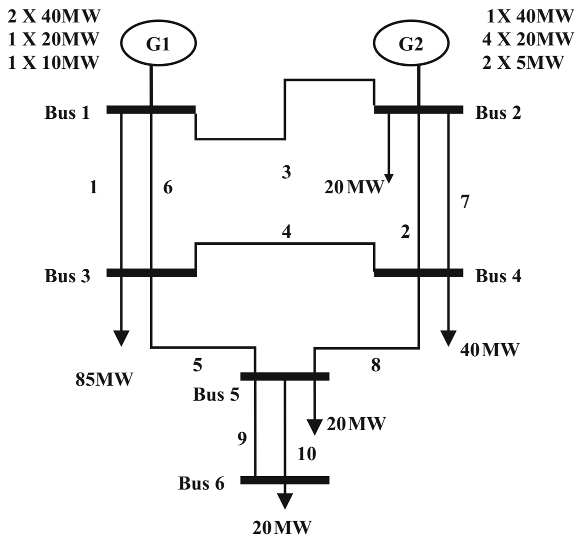


Fig. 1 IEEE 6 BUS RBTS system. *Source* Setreus [26]

the mean sojourn time in state i by: $h_i = \int_0^{\infty} [1 - H_i(t)]dt$, then the steady state probability π_i for the semi-Markov Process is given by [6]:

$$\pi_i = \frac{v_i h_i}{\sum_j v_j h_j} \tag{10}$$

The above formulation gives us a general framework to model repairs with different distributions, other than the classical exponential, given that the erroneous use of those distributions will lead to also erroneous conclusions [5]. Semi Markov models can also be applied with other distributions such as the uniform distribution [21], or combination of exponential and uniform [30], where the calculations are getting more complex, or other more "general" and flexible such as Weibul distribution [9]. When due to specific reasons of the model there is combination of two or more distributions, the sojourn time in each state is the minimum value in that state of all those distributions and/or their combinations. Of course, attempting to develop a Semi Markov model, one should consider the dramatic increase of the volume (and the complexity) of the calculations in order to balance the requirements and the available computing power. The bottom line is that the Semi Markov methodology is a useful tool especially in the analysis of multi state systems, leading to valuable inferences concerning the maintenance policy of complex systems (Fig. 1).

2 System Analysis—Original System

The IEEE 6BUS RBTS is a power generation and transmission system [3]. The main characteristics of the system are the small size, facilitating the study and the solution of different problems. Additionally, its detailed description permits the examination of new methods concerning their adequacy. The main field of the system in research is the power transmission systems, although, there are recent studies [26, 27] where it is used as a basic tool for solving problems and developing methods, concerning the power generation and power transmission as well.

The development of the IEEE 6 BUS RBTS System aims to the study of reliability, regarding the power generation and transmission; therefore, it could be a sufficient initial point of a research project. As displayed in Table 1, the capacity of the system is 240 MW, with a peak load of 185 MW and an AC Nominal voltage of 230 kV. This output is achieved by using eleven generators in two groups, #1 and #2 which consist of four and seven generators respectively. Their output power is shown in Table 2. Considering the operational conditions of each one of the generators, the system could be in different states of power output and depending on the required level of load, the possible states could be "acceptable" or "not acceptable". This fact provides an inherent uncertainty which may not have crucial extent; however, it might cause undesired consequences.

The reliability of the standard system has been studied [31] using a Markov Chain model. The interesting findings of this approach imply that the probability of the system to be in an acceptable state is inversely proportional to time. It is obvious that the probability of the system to be in a "non-acceptable" state is proportional to time. Consequently, there are inherent vulnerabilities for which there have been efforts to overcome [12]. Table 3 contains probabilities of each state.

Table 1 Basic parts of IEEE 6 BUS RBTS

Number of buses	6
Number of generators	11
Number of load points	5
Number of transmission lines	9
Number of generation buses	2
Installed generation (MW)	240
System peak load (MW)	185
AC nominal voltage (kV)	230

[27]

Table 2 Output power of the generators

Generator group	Generators—output power
#1	G1-G2 = 40 MW, G3 = 20 MW, G4 = 10 MW
#2	G5 = 40 MW, G6-G7-G8-G9 = 20 MW, G10-G11 = 5 MW

Table 3 States and operational probabilities of IEEE 6BUS RBTS

G_A	G_B	Output (MW)	State	Probability
1	1	240	P_1	0.999897
1	0	110	P_2	9.9851e−5
0	1	130	P_3	2.9998e−6
0	0	0	P_4	2.9956e−10

Table 4 Rates of failure and repair for IEEE 6 BUS RBTS

Group	Generator	MTTF (hrs)	Failure rate (Annual)	MTTR (hrs)
#1	G1	1460	6	45
	G2	1460	6	45
	G3	2190	4	45
	G4	1752	5	45
#2	G5	4380	2	45
	G6	4380	2	45
	G7	2920	3	60
	G8	3650	2.4	55
	G9	3650	2.4	55
	G10	3650	2.4	55
	G11	3650	2.4	55

Considering the reliability requirements of a system, which is able to respond at all times, the results of the study show that a system modification could cope with certain restrictions. The main purpose of this study is not to develop a new different system, but to enhance the existing one. Considering the initial layout, the system presents a relative high level of reliability. According to the literature [31] the expected probability of the complete operation of the system, is 0.99989. Table 4 shows the failure (repair) rates, and the mean time to failure (repair).

3 Modified System—Application of the Model

Indeed, the level of operability is high, but some questions arise. Is an enhancement possible? What is the effect of a change to the synthesis of the power generating system (IEEE 6 BUS RBTS)? Is an assessment concerning the performance of the system possible? Could Semi Markov modeling contribute to assess this performance? Especially, in case of a large project, where the level of operation is crucial, any kind of improvement is necessary and always welcome. Actually, there is an inherent need for further research and improvement of the original system.

More specifically, the observation of the relationship between the probability of functionality and the time of operation manifests the necessity of a further change for the better [31].

Attempting to improve the system, the authors suggest the layout of the system in a minimal way, aiming to minimize the probability of a poor state to the power transfer. The main idea is to apply a different combination of the power generators that satisfies the following criteria:

- Lower failure rates
- The same level of power per group of generators.

One of the fundamental assumptions of the authors is that not only both groups, but all generators operate independently in parallel layout. Considering the modified system, there are certain advantages by using this layout, such as that there is independence among the generators. Therefore, since a single generator fails, all the other generators continue to operate normally, until another failure to another generator takes place and/or a repair follows the existing failure (Fig. 2).

Aiming to improve the reliability characteristics of the power generation units, the change of the layout would be a sufficient start for the study. As it is shown in Fig. 2, there are two groups of generators, but with five generators in group 1 (four generators before the change) and six generators in group 2 (seven generators before the change). More analytically, and considering Table 4, the characteristics of both groups are shown in Table 5. Actually, the main idea for this change is to replace some of the generators with those that present lower failure rate.

Of course, concerning the performance of the generators, a question about the time to repair could arise. Concerning the Group #1, the replacement of the G1 and G2, by the G5 and G6 respectively, improves the performance, since there is a

Fig. 2 Modified IEEE 6 BUS RBTS suggested by authors

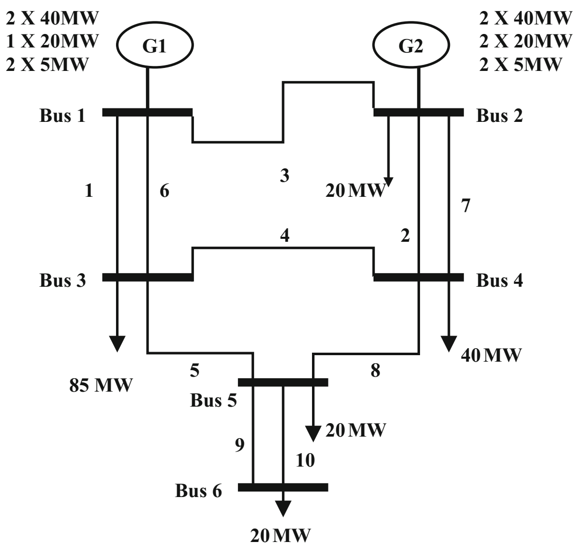


Table 5 Change of layout of group #1 and #2

Group of generators #1							
Original				Modified			
Generator	Failure rate	Power (MW)	QTY	Generator	Failure rate	Power (MW)	QTY
G1	6	40	1	G5	2	40	2
G2	6	40	1	G6	2	20	1
G3	4	20	1	G10	2.4	5	1
G4	5	10	1	G11	2.4	5	1
Group of generators #2							
Generator	Failure rate	Power (MW)	QTY	Generator	Failure rate	Power (MW)	QTY
G5	2	40	1	G3	4	20	1
G6	2	20	1	G5	4	40	2
G7	3	20	1	G6	2	20	1
G8	2.4	20	1	G10	2.4	5	2
G9	2.4	20	1				
G10	2.4	5	1				
G11	2.4	5	1				

Table 6 Brief representation of states

State	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	0
3	1	1	1	1	1	1	1	1	1	0	1
4	1	1	1	1	1	1	1	1	1	0	0
…	…	…	…	…	…	…	…	…	…	…	…
2047	0	0	0	0	0	0	0	0	0	0	1
2048	0	0	0	0	0	0	0	0	0	0	0

lowering of failure rate with the same time to repair. The other major change refers to the replacement of the G3 and G4 with the G10 and G11 respectively. Table 6 presents briefly the states of the system, where “1” is the normal operation of a generator and “0” is the status of failure. Concerning the reliability of the modified system, the states could be from full operative (240 MW—100%) to complete failure (0 MW—0%).

The probability of the system to be in an intermediate state between normal operation and complete failure proofs that it is a multi state system. After the formulation of the modified system, it is necessary to assess the new parameters and its expected behavior. After the modification, there is a lowering of the failure rate by a percentage of 40%, with an increase to the time to repair by a percentage of 22.2% (see Table 5). Consequently, this change contributes to an improvement of

Table 7 States and power output of the modified system

State #	Power output (MW)	# of states	State #	Power output (MW)	# of states	State #	Power output (MW)	# of states	State #	Power output (MW)	# of states
1	240	1	**14**	175	52	**26**	115	88	**38**	55	28
2	235	4	**15**	170	78	**27**	110	132	**39**	50	42
3	230	6	**16**	165	52	**28**	105	88	**40**	45	28
4	225	4	**17**	160	31	**29**	100	40	**41**	40	10
5	220	4	**18**	155	72	**30**	95	72	**42**	35	12
6	215	12	**19**	150	108	**31**	90	108	**43**	30	18
7	210	18	**20**	145	72	**32**	85	72	**44**	25	12
8	205	12	**21**	140	40	**33**	80	31	**45**	20	4
9	200	10	**22**	135	88	**34**	75	52	**46**	15	4
10	195	28	**23**	130	132	**35**	70	78	**47**	10	6
11	190	42	**24**	125	88	**36**	65	52	**48**	5	4
12	185	28	**25**	120	44	**37**	60	20	**49**	0	1
13	180	20									

the system as well. Considering the Group #2, there is a pure improvement of the system's performance, because of the similarity of the generators (see Table 5).

The possible states of the system are described in Table 6. There are five generators for the Group #1 and six generators for the Group #2, eleven totally. Placing both groups in parallel layout the number of states increases dramatically.

Thus, the actual number is $2^5 \cdot 2^6 = 2^{11} = 2048$. The calculation of all states aims finally to the estimation of the probability of the transition of the system to each state. These states are finally 49 *unique* states from 0 to 240 MW with an incremental step of 5 MW, due to the structure of the system (see Table 7).

Thus, the objective of the lowering of the failure rates is achieved, with only a partial increase of the time to repair. In the same time, the level of power output remains the same, achieving the same level of service with a better level of the output requirements with the original system, avoiding any major changes related with power transferring etc. Additionally, the independence of the generators mentioned above provides the flexibility that the minimum load capability is achieved with more combinations and there are forty nine output levels (as mentioned previously) compared with the four of the original system. The increase of the number of states provides more flexibility to the assessments that are necessary to manage the system and facilitate the decision-making concerning the parameters of the system. For representation purposes, due to the large size of the matrices, and in order to facilitate the understanding of the process, the authors selected to use three generators of the system in order to present the basic idea about the model. Therefore, in case of presenting systems of equations, there are eight equations (presenting a system of eight states instead of 2048). In the case of presenting matrices, there are abbreviated matrices for 2048 states. There is also another aspect concerning the issue of the reliability of the modified system. Trying to understand the meaning of reliability, it is necessary to determine the time of operation and the respective probability of an upcoming failure. First of all, in order to develop the Semi Markov Model, we should solve the system of equations in steady state conditions. The main assumption in this case is that the sojourn time in one state is exponentially distributed. Additionally, this transition is analyzed in two parts. The first one refers to the time spent in the particular state and the other one refers to the probability the system to be in another state. The jump of the system from one state to another is a result of the combination of failures and repairs of different generators. This occurs because the system consists of eleven generators and a probability to fail and/or repair one at a time or some of them simultaneously always exists. This transition will take place when the first combination of failure and/or repair comes. The transition matrix based on the failure rates and/or repairs will have the form

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,2048} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,2048} \\ \cdots & \cdots & \cdots & \cdots \\ p_{2048,1} & p_{2048,2} & \cdots & p_{2048,2048} \end{bmatrix} \quad (11)$$

Obviously, the number and the difficulty of the calculations increase dramatically with the increase of generators, confirming the findings of researchers about the complexity and the difficulty of the models. The failure and repair rates are expressed in failures and repairs per hour. These numbers are extracted by dividing the annual failure rates and repairs by 8,760 h per year. According to this analysis, the failure and repair rates are shown in Table 8. Continuing with the calculation of the steady state probabilities, we should solve the following matrix equation [30]:

$$v = vP \quad (12)$$

Table 8 Failure and repair rates

GEN	MTTF (hrs)	Failure rate (per year)	MTTR (hrs)	Failure rate (per hour)	Repair rate (per hour)
G1	1460	6	45	0.00000	0.00000
G2	1460	6	45	0.00000	0.00000
G3	2190	4	45	0.00000	0.00000
G4	1752	5	45	0.00000	0.00000
G5	4380	2	45	0.00046	0.01027
G6	4380	2	45	0.00023	0.00514
G7	2920	3	60	0.00000	0.00000
G8	3650	2.4	55	0.00000	0.00000
G9	3650	2.4	55	0.00000	0.00000
G10	3650	2.4	55	0.00027	0.00628
G11	3650	2.4	55	0.00027	0.00628

Where P is the transition probabilities Matrix and v is the vector of the discrete time Markov Chain.

$$v = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & \ldots & \ldots & v_{2047} & v_{2048} \end{bmatrix} \quad (13)$$

Of course, a unique solution of the Eq. (17) is possible only under the restriction [32].

$$\sum_{i=1}^{2048} v_i = 1 \quad (14)$$

The mean sojourn time for each state is given by the formula [30].

$$h_i = \int_0^{\infty} [1 - H_i(t)]dt \quad (15)$$

Solving the formula above, we find that the mean sojourn time has the form

$$h_i = \frac{1}{\lambda_i + \mu_i} \quad (16)$$

Once again, this expression is only indicative for representation reasons only and it is not applicable for all states.

$$V \cdot P_{semi} = U \Leftrightarrow V = U \cdot P_{semi}^{-1} \quad (17)$$

Vector U is

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & \ldots & 0_{(2047)} & 0_{(2048)} \end{bmatrix} \quad (18)$$

and V the matrix that will be combined with the mean sojourn times to calculate the final steady state probabilities. Solving the matrix Eq. (17), we have the results in Table 9. At this point, we notice that this table contains probabilities of **all** states of the system. Thus, assessing the probability of the failure related to the time of operation, these reliability parameters are shown in Table 10. An initial interest point of these results is the probability of the system to provide a level of power output.

So, using the probabilities for each state (Table 9) and adding all probabilities with the same output, we have the final probabilities based on the level of output (see Table 10), giving the opportunity to the decision maker to shape the big picture concerning the expected level of the power output. According to Table 10, the expected level of power output which is E[X] = **237.267 MW**[1] and its variance found to be Var[X] = **81.83** $\mathbf{MW^2}$ and finally its standard deviation is at the level

[1]Calculation of expected value E[X].

Table 9 Probabilities of each state

State	Group #1					Group #2						Power (MW)	Prob
	G5	G5	G6	G10	G11	G3	G5	G5	G6	G10	G11		
1	1	1	1	1	1	1	1	1	1	1	1	240	8.681E−01
2	1	1	1	1	1	1	1	1	1	1	0	235	1.308E−02
3	1	1	1	1	1	1	1	1	1	0	1	235	1.308E−02
4	1	1	1	1	1	1	1	1	1	0	0	230	1.971E−04
5	1	1	1	1	1	1	1	1	0	1	1	220	8.918E−03
…	…	…	…	…	…	…	…	…	…	…	…	…	…
2047	0	0	0	0	0	0	0	0	0	0	1	5	1.532E−19
2048	0	0	0	0	0	0	0	0	0	0	0	0	2.541E−16

of **9.05 MW**.[2] These values assure that the system fulfills the criterion of a minimum output or peak load of **185 MW** with a probability of an almost total certain operation. Comparing the final results, the original system presents a probability for "normal operation >130 MW" equal to **0.999897**. The respective probability of the modified is **0.999995**. the probability of ">110 MW" is **0.999999** and for total failure of the original is $\mathbf{2.9956e-10}$ and for the modified is $\mathbf{2.54109E-16}$. But, the most important finding is that the minimum output limit of 185 MW is achieved in the modified system with a probability of **0.997809**. Since the peak load is 185 MW there is a difference of 5.76σ (standard deviations). Considering Chebyshev's Inequality [25] since the power output is a random variable with mean **237.267 MW** and variance $\mathbf{81.83\ MW^2}$ then for any value k > 0,

$$P\{|X-\mu|\geq k\}\leq\frac{\sigma^2}{k^2} \tag{19}$$

Therefore, if $k=\frac{237.267-185}{9.05}\Leftrightarrow k=5.76$ standard deviations, then the formula (19) gives $P\{|X-237.267|\geq 52.267\}\leq\frac{81.83}{52.267^2}\Leftrightarrow P\{|X-237.267|\geq 52.267\}\leq 0.02995$. And finally, the probability of the power output to be higher than 185 MW in any case is $P\{|X-237.267|\leq 52.267\}\geq 1-0.02995\Leftrightarrow P\{|X-237.267|\leq 52.267\}\geq 0.97$. This result agrees and confirms the findings of the analysis.

Concerning the expected time of operation in each state it is shown in Table 11. The main finding is that in annual basis, the system operates in more than 185 MW for 8,740 h, 48 min and 36 s out of 8,760 h totally. The output from 130 MW to 184 MW is 19 h, 9 min and 4 s. The operation between 110 and 130 MW is 2 min and 16 s. Finally, the operation in level lower than 110 MW is only 4 s in annual basis. All the above findings show that the suggested modification upgrades the system and its reliability and could be a starting point for further improvement.

[2]Calculation of Variance by var[X] and its standard deviation.

Table 10 Probabilities and cumulative probabilities of each output level

State (MW)	Exp hours p.a.	Prob	Cummul prob	State (MW)	Exp hours p.a.	Prob	Cummul prob
240	7,604.214	8.68060E-01	8.68060E-01	115	0.002	2.44451E-07	9.99999E-01
235	458.336	5.23214E-02	9.20382E-01	**110**	**0.000**	**5.52527E-09**	**9.99999E-01**
230	10.360	1.18260E-03	9.21565E-01	105	0.000	5.55064E-11	9.99999E-01
225	0.104	1.18800E-05	9.21576E-01	100	0.001	1.55940E-07	9.99999E-01
220	312.502	3.56737E-02	9.57250E-01	95	0.000	9.39912E-09	9.99999E-01
215	18.836	2.15019E-03	9.59400E-01	90	0.000	2.12444E-10	9.99999E-01
210	0.426	4.86003E-05	9.59449E-01	85	0.000	2.13583E-12	9.99999E-01
205	0.004	4.88222E-07	9.59449E-01	80	0.000	1.16590E-08	9.99999E-01
200	316.515	3.61318E-02	9.95581E-01	75	0.000	7.02737E-10	9.99999E-01
195	19.078	2.17781E-03	9.97759E-01	70	0.000	1.58837E-11	9.99999E-01
190	0.431	4.92245E-05	9.97808E-01	65	0.000	1.60973E-13	9.99999E-01
185	**0.004**	**4.94492E-07**	**9.97809E-01**	60	0.000	4.05634E-10	1.00000E + 00
180	12.859	1.46792E-03	9.99277E-01	55	0.000	2.44497E-11	1.00000E + 00
175	0.775	8.84777E-05	9.99365E-01	50	0.000	5.52498E-13	1.00000E + 00
170	0.018	1.99984E-06	9.99367E-01	45	0.000	5.99760E-15	1.00000E + 00
165	0.000	2.00897E-08	9.99367E-01	40	0.000	5.10464E-12	1.00000E + 00
160	4.981	5.68594E-04	9.99936E-01	35	0.000	3.07528E-13	1.00000E + 00
155	0.300	3.42714E-05	9.99970E-01	30	0.000	7.16627E-15	1.00000E + 00
150	0.007	7.74628E-07	9.99971E-01	25	0.000	1.57224E-16	1.00000E + 00
145	0.000	7.78165E-09	9.99971E-01	20	0.000	2.07729E-14	1.00000E + 00
140	0.199	2.26705E-05	9.99994E-01	15	0.000	1.22322E-15	1.00000E + 00
135	0.012	1.36644E-06	9.99995E-01	10	0.000	1.33671E-16	1.00000E + 00
130	**0.000**	**3.08853E-08**	**9.99995E-01**	5	0.000	1.96904E-17	1.00000E + 00
125	0.000	3.10264E-10	9.99995E-01	**0**	**0.000**	**2.54108E-16**	**1.00000E + 00**
120	0.036	4.05567E-06	9.99999E-01	**Expected Power (MW) E[X]=**			**237.26**

Power levels marked in bold refer to power levels (states) of initial IEEE 6BUS RBTS

Table 11 Reliability based on hrs of operation

Time (Hrs)	Prob group #1	Prob group #2	Total	Time (Hrs)	Prob group #1	Prob group #2	Total
0	1.0000E+00	1.0000E+00	1.0000E+00	4500	8.6712E−01	8.8414E-01	9.8460E-01
250	1.0000E+00	1.0000E+00	1.0000E+00	4750	8.4636E−01	8.6392E-01	9.7909E-01
500	9.9998E−01	1.0000E+00	1.0000E+00	5000	8.2456E−01	8.4245E−01	9.7236E−01
750	9.9987E−01	9.9996E−01	1.0000E+00	5250	8.0185E−01	8.1988E−01	9.6431E−01
1000	9.9951E−01	9.9982E−01	1.0000E+00	5500	7.7841E−01	7.9640E−01	9.5488E−01
1250	9.9871E−01	9.9944E−01	1.0000E+00	5750	7.5439E−01	7.7217E−01	9.4404E−01
1500	9.9723E−01	9.9863E−01	1.0000E+00	6000	7.2994E−01	7.4738E−01	9.3178E−01
1750	9.9482E−01	9.9715E−01	9.9999E−01	6750	6.5539E−01	6.7120E−01	8.8669E−01
2000	9.9123E−01	9.9475E−01	9.9995E−01	7000	6.3056E−01	6.4567E−01	8.6910E−01
2250	9.8627E−01	9.9119E−01	9.9988E−01	7250	6.0591E−01	6.2030E−01	8.5036E−01
2500	9.7977E−01	9.8623E−01	9.9972E−01	7500	5.8155E−01	5.9517E−01	8.3060E−01
2750	9.7161E−01	9.7970E−01	9.9942E−01	7750	5.5754E−01	5.7039E−01	8.0991E−01
3000	9.6171E−01	9.7144E−01	9.9891E−01	8000	5.3396E−01	5.4604E−01	7.8844E−01
3250	9.5004E−01	9.6137E−01	9.9807E−01	8250	5.1088E−01	5.2218E−01	7.6629E−01
3500	9.3664E−01	9.4945E−01	9.9680E−01	8500	4.8833E−01	4.9889E−01	7.4360E−01
3750	9.2154E−01	9.3570E−01	9.9496E−01	8750	4.6637E−01	4.7619E−01	7.2048E−01
4000	9.0484E−01	9.2016E−01	9.9240E−01	8800	4.6206E−01	4.7173E−01	7.1582E−01
4250	8.8666E−01	9.0294E−01	9.8900E−01				

4 Conclusions

In this study, we have evaluated the performance of a modified IEEE 6BUS RBTS and have shown that the recommended modification contributes to the improvement of the system performance by increasing the probability of operation within the required limits. The modification of the system aims to the reduction of failure rates maintaining the power output specifications. This objective is achieved through the restructuring of the group of generators whereas all generators are supposed to be in parallel operation. This technique led to approach the problem through the multi state systems theory. As described in previous paragraphs, the increase of the parts of the system resulted to a dramatic increase of the states of the system and consequently increased the complexity of the calculations. This problem showed in practice that probable restrictions of the computing power are possible to be overcome through programming algorithms and/or advanced software. In the case of even more complex systems, a breakdown of the system in smaller parts is a suggested complementary alternative. As multi state systems theory suggests, this model presents more flexibility than that of the original IEEE 6BUS RBTS. More specifically, there are more levels of power output and their respective probabilities. This characteristic contributes to the lowering of the uncertainty of the system, assisting decisively in the decision-making and the management of the system. Comparing the models, there is a significant improvement after the suggested modification. The reduction of the probability of "non-acceptable" output level and the increased probability of operation over the minimum required level contributes to the effective management of the system. The managerial aspect of the system's modification is that it contributes to the simplicity of the system. Since there are fewer types of generators, it simplifies the management of the system concerning the schedule of supplies and its maintenance as well. All these findings could be a starting point for further study and expansion of the methodology in other research topics. In addition, using advanced programming algorithms, the researchers could apply analytical methods in an effort to overcome the existing barriers of the computing power. This strategy can be combined with other methods such as Monte Carlo Simulation, in order to verify the effectiveness of each other method and to reduce the uncertainty of the models.

References

1. Aven T, Jensen U (1999) Stochastic models in reliability. Springer, New York
2. Barbu VS, Limnios N (2008) Semi Markov chains and hidden semi Markov models towards applications. Springer Business Media, New York
3. Billinton R, Li Y (1994) Reliability assessment of electric power systems using Monte Carlo methods. Springer Science Business Media, New York
4. Billinton R, Li Y (2007) Incorporating multi-state unit models in composite system adequacy assessment. Eur Trans Electr Power 17:375–386

5. Ciardo G, Marie RA, Sericola B, Trivedi K (1990) Performability analysis using semi-Markov reward processes 39(10):1251–1264
6. Cinlar E (1975) Introduction to stochastic processes. Prentice-Hall, Englewood Cliffs, N.J
7. Eryilmaz S (2015a) Assessment of a multi-state system under a shock model. Appl Math Comput 269:1–8
8. Eryilmaz S (2015b) Dynamic assessment of multi state systems using phase type modeling. Reliab Eng Syst Saf 140:71–77
9. Foucher Y, Mathieu E, Sait-Pierre P, Durand J-F, Daures J-P (2005) A semi markov model based on generalized Weibull distribution with an illustration for HIV disease. Biometrical J 47:1–9
10. Gut A (1990) Cumulative shock models. Adv Appl Probab 22:504–507
11. Koval et al (2007) IEEE recommended practice for the design of reliable industrial and commercial power systems. IEEE Gold Book, New York
12. Kumar S, Sankar V (2013) Enhancement of reliability analysis for a 6-bus composite power system using combination of TCSC & UPFC. International conference on recent trends in power, control & instrumentation engineering PCIE-2013. Elsevier
13. Limnios N, Oprisan G (2001) Semi-Markov processes and reliability. Springer Science Business Media, New York
14. Lisnianski A, Levitin G, Ben Heim H, Elmakias D (1999) Power system structure optimization subject to reliability constraints. Electr Power Syst Res 39:145–152
15. Lisnianski A (2007) Extended block diagram method for a multi-state system reliability assessment. Reliab Eng Syst Saf 92(12):1601–1607
16. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
17. Lisnianski A, Elmakias D, Laredo D, Haim HB (2012) A multi-state Markov model for a short-term reliability analysis of a power generating unit. Reliab Eng Syst Saf 98:1–6
18. Lisnianski A (2016) Application of extended universal generating function technique to dynamic reliability analysis of a multi state system. In: The Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management
19. Liu YW, Kapur KC (2006) Reliability measures for dynamic multistate nonrepairable systems and their applications to system performance evaluation. IIE Trans 38(6):511–520
20. Mallor F, Omey E, Santos J (2006) Asymptotic results for a run and cumulative mixed shock model. J Math Sci 138:5410–5414
21. Nabli H (2004) Performability: asymptotic distribution and moment computation. Comput Math Appl 48:1–8
22. Natvig B (2007) Multi-state reliability theory, department of Mathematics. University of Oslo Statistical Research Report No 1, ISSN 0806–3842
23. Ramirez-Marquez J, Coit D (2005) A Monte-Carlo simulation approach for approximating multi-state two-terminal reliability. Reliab Eng Syst Saf 87:253–264
24. Ross S (1995) Stochastic processes (2). Wiley, Canada
25. Ross S (2003) Introduction to probability models (8). Academic Press, Burlington
26. Setreus J (2009) On reliability methods quantifying risks to transfer capability in electric power transmission systems. Licentiate Thesis in Electrical Systems Stockholm, Sweden
27. Setreus J (2011) Identifying critical components for system reliability in power transmission systems. Doctoral Thesis in Electrical Systems, Stockholm, Sweden
28. Shanthikumar G, Sumita U (1983) General shock models associated with correlated renewal sequences. J Appl Probab 20:600–614
29. Shey-H Chang C, Chen Y, Zhang ZG (2015) Optimal preventive maintenance and repair policies for multi-state systems. Reliab Eng Syst Saf 140:78–87
30. Trivedi KS (2002) Probability and statistics with reliability, queuing and computer science applications. Wiley, New York
31. Vijayalaxmi D, Karjagi B (2015) Modeling and analysis of RBTS IEEE-6 BUS system based on Markov Chain. Int J Eng Res Gen Sci 3(2):686–696

32. Violentis J, Koutras V, Platis A, Gravvanis G (2008) Asymptotic availability of an electrical substation via a semi-markov process computed by generalized approximate inverse preconditioning. HERCMA Conference Proceedings.
33. Yingkui G, Jing L (2012) Multi-state system reliability: a new and systematic review. Procedia Eng 29:531–536
34. Zhao X, Cui L (2012) Reliability evaluation of generalised multi-state k out-of-n systems based on FMCI approach. Int J Syst Sci 41:1437–1443
35. Zio E (2009) Reliability engineering: old problems and new challenges. Reliab Eng Syst Saf 94:125–141

Lz-Transform Approach for Fault Tolerance Assessment of Various Traction Drives Topologies of Hybrid-Electric Helicopter

Ilia Frenkel, Igor Bolvashenkov, Hans-Georg Herzog and Lev Khvatskin

Abstract This chapter presents a preliminary analysis of fault tolerance, availability and performance assessment of the two promising options of the hybrid-electric traction drive version for the helicopter, which can be treated as multi-state system, where components and entire system in general case have an arbitrary finite number of states corresponding to the different performance rates. The performance rate (output nominal power) of the system at any time instant is interpreted as a discrete-state continuous-time stochastic process. In the present chapter, the *Lz*-transform is applied to a real multi-state hybrid-electric traction drive version for the helicopter system that is functioning under various stochastic demands and its availability and performance is analyzed. It is shown that *Lz*-transform application drastically simplifies the availability computation for such a system compared with the straightforward Markov method.

Keywords Hybrid-electric traction drive · Availability · Performance assessment · *Lz*-transform · Multi-state system · Discrete-state continuous-time stochastic process

1 Introduction

Nowadays, the electrification of aircraft of different types and purposes is one of the most promising directions in the development of aviation technology. According to program MEA (More Electric Aircraft) [7] developers of various specialized companies are planning the creation of electric airplanes (liner, regional, special

I. Frenkel (✉) · L. Khvatskin
Center for Reliability and Risk Management, Sami Shamoon College of Engineering, Beer Sheva, Israel
e-mail: iliaf@sce.ac.il

I. Bolvashenkov · H.-G. Herzog
Institute of Energy Conversion Technology, Technical University of Munich, Munich, Germany

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*, Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_17

purpose, manned and unmanned) as well as the helicopters (from heavy to small-size, manned and unmanned), which have a number of benefits associated with the improvement of the technical performance of the aircraft, increase their environmental performance and reducing operating costs.

The construction of electric aircraft requires a comprehensive revision of the principles of construction of a variety of devices and systems of the aircraft, which is associated with the creation of the new, having a low specific weight controlled electric drive, electric power generators, electrical energy storage (batteries, fuel cells, ultra-capacitors etc.), and electric power converters. The most important place in the solving the problem of the creation of electric aircrafts takes traction electric drive. The use of electrical technologies could lead in the near future to the changing the principles of its construction and propulsion.

The first step in the aircraft electrification is the development and implementation of hybrid-electric aircraft. This is due to the fact that today there are no electrical energy storages with high energy density and low weight and dimensions. Electric aircraft version can be implemented only with creation of new energy storage devices with the appropriate characteristics of the energy density, weight and dimensions.

This chapter presents a preliminary analysis of fault tolerance, availability and performance assessment of the two promising options of the hybrid-electric traction drive version for the conventional Airbus helicopter EC135/H135 with gas turbine engine and speed reducer [14].

Due to the system's nature, a fault in a single unit has only partial effect on the entire performance: it only reduces the system's performance. One partial failure of the traction multiphase permanent magnets synchronous motor leads to partial system failure (reduction of output nominal power), as well as the multiple consecutive multiphase motor's failures, to complete system failures. So, the hybrid-electric traction drive system can be treated as multi-state system (MSS), where components and entire system in the general case have an arbitrary finite number of states corresponding to different performance rates [1, 12, 13]. The performance rate (output nominal power) of the system at any time instant is interpreted as a discrete-state continuous-time stochastic process. Such a model is complex enough—even in relatively simple cases it has hundreds states. So, it is rather difficult to build the model and to solve the corresponding system of differential equations by using straightforward Markov method.

In recent years a specific approach called *L*z-transform, was introduced [10] for discrete-state continuous-time Markov processes. This approach is an extension of the universal generating function (UGF) technique that was introduced by Ushakov [18] and has been widely applied to MSS reliability analysis [12]. *L*z-transform was successfully applied to availability analysis of real-world multi-state systems under constant and variable demand [5, 6, 8, 11, 16] and its efficiency was demonstrated. In practice there are aging multi-state systems that are functioning under variable stochastic demand. In the present chapter, the *L*z-transform is applied to a real MSS hybrid-electric traction drive version for the conventional helicopter system that is functioning under variable stochastic demand and its availability and performance is analyzed. It is shown that *Lz*-transform application drastically simplifies the availability computation for such a system compared with the straightforward Markov method.

2 Comparative Analysis of Two Traction Drive Topologies of Hybrid-Electric Helicopter

2.1 Common Description

In this chapter for the comparative analysis were selected two traction drive topologies of hybrid-electric helicopter: a serial electric-hybrid propulsion system with a gas turbine and electrical generator (Fig. 1a) and a combined electric-hybrid propulsion system with a gas turbine, generator and the speed reducer (Fig. 1b).

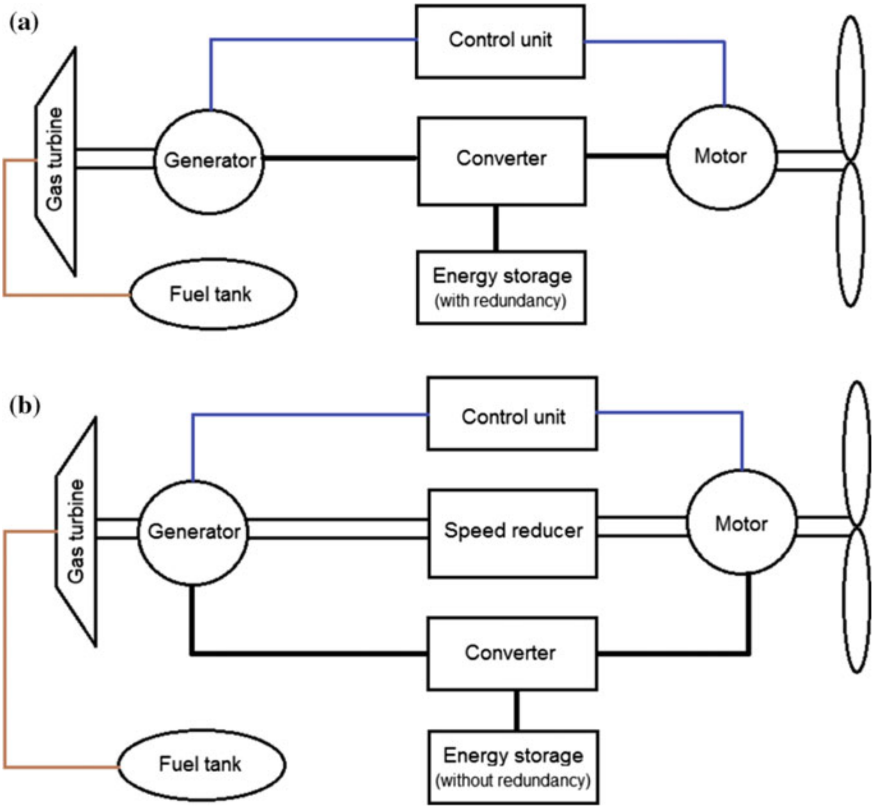


Fig. 1 Propulsion system topologies of the hybrid-electric helicopter, serial (**a**) and combined (**b**)

In these schemes the following equipment is used as follows:

- the GTE Pratt & Whitney of type PW206B2 [15] is used as the gas turbine engine,
- synchronous generator with permanent magnets [9] is used as an electric generator,
- the multiphase permanent magnets synchronous motor [4]—as the traction motor,
- multilevel cascaded H-bridge (CHB) inverter [3] is used as a converter and
- battery electric energy storage (BEES), based on the lithium-ion battery cells [2] —as energy storage.

As shown in [4], the optimal traction electric motor from the point of view of the helicopter's fault tolerance requirements is a 9-phase synchronous motor with permanent magnets and galvanically isolated phases. The states of degradation and the corresponding performance of the multiphase electric motors with the nominal power of 540 kW, which occur during the consecutive critical phase failures, is as follows: one phase failure reduces power to 480 kW, failure of two phases—to 420 kW, three phases' failure—to 360 kW. In this work, the power reduction less than 420 kW means full failure of the traction motor.

As shown in [3], the 17-level cascaded H-bridge inverter is the optimal type of electric converter in terms of helicopter's fault tolerance requirements.

The required power (demand) that must be implemented by the traction drive system of hybrid-electric helicopter for a one helicopter's flight is as follows: 540, 460 and 360 kW.

2.2 *Operational Scenarios for Various Traction Drive Topologies*

2.2.1 Serial Topology

In case of the serial topology (Fig. 1a) in normal failure-free operation, the gas turbine electric generator feeds the traction multiphase electric motor through the multilevel converter. The control unit controls the entire traction drive. The state of charge of battery cells is continually monitored, recharged from the electric generator and maintained at optimal maximum level. It is possible to install an additional amount of electric battery cells, which increase the fault tolerance indices of the battery electric energy storage of serial connection topology of traction drive units.

In failure case of the gas turbine engine or/and electric generator, depending on the value of electric energy storage capacity, it is possible, either the continuation of the flight to complete task performance, or a flight to a safe landing place. After landing the running repairs or complete replacement of the failed units are implemented.

2.2.2 Combined Topology

In case of the combined topology (Fig. 1b) in the normal failure-free operation there are two possible embodiment of the flight. In the first "electric" version of flight, the gas turbine electric generator feeds the multi-phase electric traction motor through multilevel inverter. In the second, the "mechanical" version of flight, the gas turbine engine directly rotates the propeller of the helicopter through a mechanical transmission and speed reducer. Control the components of the traction drive is carried out by the control unit. Electric battery storage is charged by the electrical generator only in the "electric" flight mode of helicopter.

In the failure case of electric generator, or/and an electric inverter, or/and the traction electric motor, the further flight is provided by "mechanical" embodiment of the traction drive scheme realization.

In the failure case of the speed reducer, the further flight is provided by "electrical" embodiment of the traction drive scheme realization.

In failure case of the gas turbine engine, a further short time flight to the safe landing of helicopter is carried out by electric energy of battery storage. After landing the running repairs or complete replacement of the failed units are implemented.

3 Brief Description of the *Lz*-Transform Method

We consider a multi-state system consisting of n multi-state components. Any j-component can have k_j different states, corresponding to different performances g_{ji}, represented by the set $\mathbf{g}_j = \{g_{j1}, \ldots, g_{jk_j}\}$, $j = \{1, \ldots, n\}$; $i = \{1, 2, \ldots, k_j\}$. The performance stochastic processes $G_j(t) \in \mathbf{g}_j$ and the system structure function $G(t) = f(G_1(t), \ldots, G_n(t))$ that produces the stochastic process corresponding to the output performance of the entire MSS, fully define the MSS model.

The MSS model definitions can be divided into the following steps. For each multi-state component we will build a model of stochastic process. Markov performance stochastic process for each component j can be represented by the expression $G_j(t) = \{\mathbf{g}_j, \mathbf{A}_j, \mathbf{p}_{j0}\}$, where $\mathbf{g}_j$ is the set of possible component's states, defined above, $\mathbf{A}_j = \left(a_{lm}^{(j)}(t)\right)$, $l, m = 1, \ldots, k; j = 1, \ldots n$ is the transition intensities matrix and $\mathbf{p}_{j0} = \left[p_{10}^{(j)} = \Pr\{G_j(0) = g_{10}\}, \ldots, p_{k_j0}^{(j)} = \Pr\{G_j(0) = g_{k_j0}\}\right]$ is the initial states probability distribution.

For each component j the system of Kolmogorov forward differential equations [17] can be written for determination of the state probabilities $p_{ji}(t) = \Pr\{G_j(t) = g_{ji}\}$, $i = 1, \ldots k_j, j = 1, \ldots, n$ under initial conditions $\mathbf{p}_{j0}$. Now Lz-transform of a discrete-state continuous-time (DSCT) Markov process $G_j(t)$ for each component j can be written as follows:

$$L_Z\{G_j(t)\} = \sum_{i=1}^{k_j} p_{ji}(t) z^{g_{ji}} \tag{1}$$

The next step, in order to find the Lz-transform of the entire MSS's output performance Markov process $G(t)$, the Ushakov's Universal Generating Operator [18] can be applied to all individual Lz-transforms $L_Z\{G_j(t)\}$ over all time points $t \geq 0$.

$$L_Z\{G(t)\} = \Omega_f\{L_Z[G_1(t)], \ldots, L_Z[G_n(t)]\} = \sum_{i=1}^{K} p_i(t) z^{g_i} \tag{2}$$

In this expression K is the number of states in the entire MSS, p_i and g_i are probabilities and performances of the entire MSS.

The technique of Ushakov's operator application is well established for many different structure functions [12].

Using the resulting Lz-transform MSS mean instantaneous availability for constant demand level w can be derived as sum all probabilities in Lz-transform from terms where powers of z are not negative:

$$A(t) = \sum_{g_i \geq w} p_i(t) \tag{3}$$

MSS's mean instantaneous performance may be calculated as sum all probabilities multiplied to performance in Lz-transform from terms where powers of z are positive:

$$E(t) = \sum_{g_i > 0} p_i(t) g_i \tag{4}$$

4 Multi-state Modeling of the Multi Power Source Traction Drive

4.1 Systems' Description

According to Lz-transform method, we build the Reliability Bloc Diagrams for presented on the Fig. 1 propulsion system topologies of the hybrid-electric helicopter. On Fig. 2 one can find the Reliability Block Diagram of the Serial Topology System and on Fig. 3—Reliability Block Diagram for Combined Topology System.

The main part of the Serial Topology System (STS) (Fig. 2) consists of two subsystems SS_{1s} and SS_{2s} connected in series. Subsystem SS_{1s} consists of 3 elements, Fuel Tank, Gas Turbine and Generator, connected in series. Subsystem SS_{2s} consists of 3 elements, Converter, Motor and Control Unit, connected in series. In

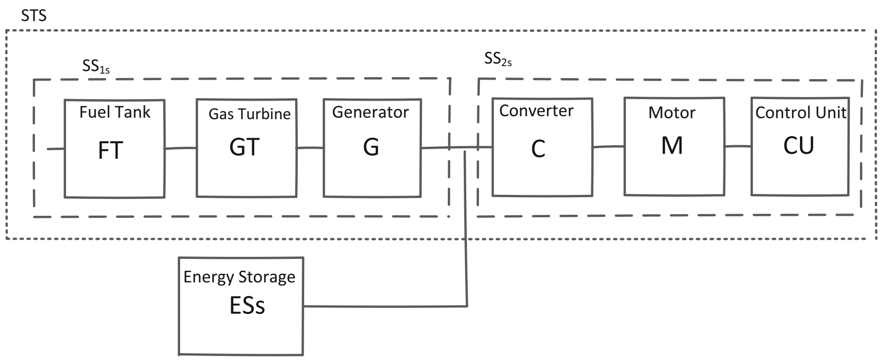


Fig. 2 Reliability Block Diagram of the Serial Topology System

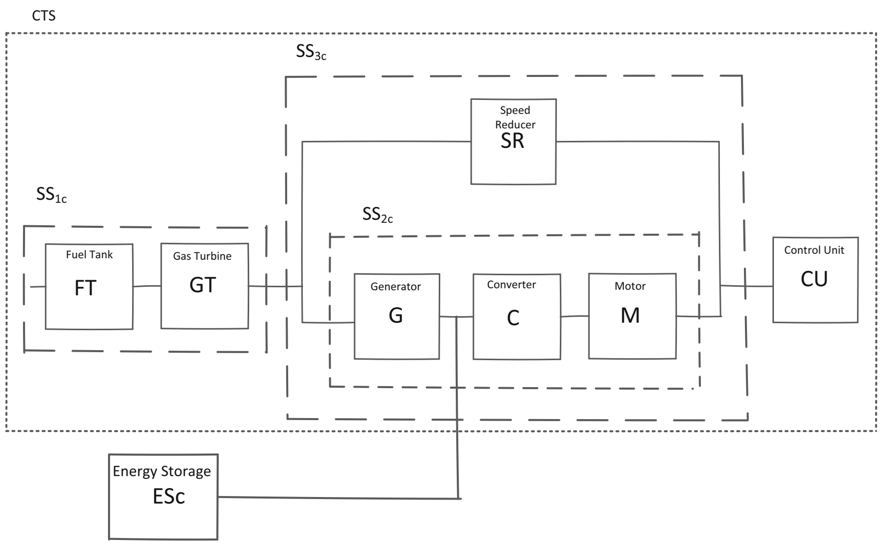


Fig. 3 Reliability Block Diagram of the Combined Topology System

case of emergency, instead main STS activates reserved system STS-ESs, where failed subsystem SS_{1s} replaced by multi-state Energy storage.

The main part of the Combined Topology System (CTS) (Fig. 3) consists of two subsystems SS_{1c} and SS_{3c} and Control Unit, connected in series. Subsystem SS_{1c} consists of 2 elements, Fuel Tank and Gas Turbine, connected in series. Subsystem SS_{3c} consists of Subsystem SS_{2c} and Speed Reducer, connected in parallel. Subsystem SS_{2c} consists of 3 elements, Generator, Converter and Motor, connected in series. In case of emergency, instead main CTS activates reserved system CTS-ESs, consists of connected in series Energy storage, Converter, Motor and Control Unit.

Elements' descriptions are presented in the next section.

4.2 Elements' Description

4.2.1 Elements with 2 States

For system's elements, which have 2 states (fully working and fully failed) in order to calculate probabilities of each state we build the state space diagram (Fig. 4) and the following system of differential equations:

$$\begin{cases} \frac{dp_{i1}(t)}{dt} = -\lambda_i p_{i1}(t) + \mu_i p_{i2}(t), \\ \frac{dp_{i2}(t)}{dt} = \lambda_i p_{i1}(t) - \mu_i p_{i2}(t). \end{cases}, i = \text{FT, GT, G, ESc, C, SR, CU}$$

Initial conditions are: $p_{i1}(0); p_{i2}(0) = 0.$

We used MATLAB® for numerical solution of these systems of DE to obtain probabilities $p_{i1}(t), p_{i2}(t)$ (i = FT, GT, G, ESc, C, SR, CU). Therefore, for such system's element the output performance stochastic processes can be obtained as follows:

$$\begin{cases} \mathbf{g}_i = \{g_{i1}, g_{i1}\} = \{540, 0\}, \\ \mathbf{p}_i(t) = \{p_{i1}(t), p_{i1}(t)\} \end{cases}.$$

Sets $\mathbf{g}_i, \mathbf{p}_i(t)$ i = FT, GT, G, ESC, C, SR, CU define Lz-transforms for each element as follows:

Fuel Tank: $L_z\{g^{FT}(t)\} = p_1^{FT}(t)z^{g_1^{FT}} + p_2^{FT}(t)z^{g_2^{FT}} = p_1^{FT}(t)z^{540} + p_2^{FT}(t)z^0.$

Gas Turbine: $L_z\{g^{GT}(t)\} = p_1^{GT}(t)z^{g_1^{GT}} + p_2^{GT}(t)z^{g_2^{GT}} = p_1^{GT}(t)z^{540} + p_2^{GT}(t)z^0.$

Generator: $L_z\{g^{G}(t)\} = p_1^{G}(t)z^{g_1^{G}} + p_2^{G}(t)z^{g_2^{G}} = p_1^{G}(t)z^{540} + p_2^{G}(t)z^0.$

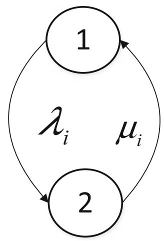


Fig. 4 State space diagram

Energy Storage

Combine: $L_z\{g^{ESc}(t)\} = p_1^{ESc}(t)z^{g_{i1}^{ESc}} + p_2^{ESc}(t)z^{g_{i2}^{ESc}} = p_1^{ESc}(t)z^{540} + p_2^{ESc}(t)z^0.$

Converter: $L_z\{g^C(t)\} = p_1^C(t)z^{g_1^C} + p_2^C(t)z^{g_2^C} = p_1^C(t)z^{540} + p_2^C(t)z^0.$

Speed Reducer: $L_z\{g^{SR}(t)\} = p_1^{SR}(t)z^{g_1^{SR}} + p_2^{SR}(t)z^{g_2^{SR}} = p_1^{SR}(t)z^{540} + p_2^{SR}(t)z^0.$

Control Unit: $L_z\{g^{CU}(t)\} = p_1^{CU}(t)z^{g_1^{CU}} + p_2^{CU}(t)z^{g_2^{CU}} = p_1^{CU}(t)z^{540} + p_2^{CU}(t)z^0.$

4.2.2 Element with 3 States

The system's element, Energy Storage Serial (ESs), has 3 states: fully working state with performance 540 KW, partial failure state with performances 440 and fully failure. The state-space diagram is presented on Fig. 5. To calculate probabilities of each state we build the following system of differential equations:

$$\begin{cases} \frac{dp_1^{ESs}(t)}{dt} = -\lambda_{12}^{ESs}p_1^{ESs}(t) + \mu_{21}^{ESs}p_2^{ESs}(t) + \mu_{31}^{ESs}p_3^{ESs}(t), \\ \frac{dp_2^{ESs}(t)}{dt} = \lambda_{12}^{ESs}p_1^{ESs}(t) - (\lambda_{23}^{ESs} + \mu_{21}^{ESs})p_2^{ESs}(t) \\ \frac{dp_3^{ESs}(t)}{dt} = \lambda_{23}^{ESs}p_2^{ESs}(t) - \mu_{31}^{ESs}p_3^{ESs}(t) \end{cases}$$

Initial conditions are

$$p_1^{ESs}(0) = 1;\ p_2^{ESs}(0) = 0;\ p_3^{ESs}(0) = 0.$$

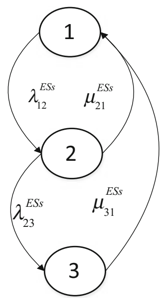
We used MATLAB® for numerical solution of this system of DE to obtain probabilities $p_1^{ESs}(t)$, $p_2^{ESs}(t)$, $p_3^{ESs}(t)$. Therefore, for such system's element the output performance stochastic processes can be obtained as follows:

$$\begin{cases} \mathbf{g}^{ESs} = \{g_1^{ESs}, g_2^{ESs}, g_3^{ESs}\} = \{540, 440, 0\}, \\ \mathbf{p}^{ESs}(t) = \{p_1^{ESs}(t), p_2^{ESs}(t), p_3^{ESs}(t)\}. \end{cases}$$

Sets $\mathbf{g}^{ESs}, \mathbf{p}^{ESs}(t)$ define *Lz*-transforms for Energy Storage Serial as follows:

$$\begin{aligned} L_z\{g^{ESs}(t)\} &= p_1^{ESs}(t)z^{g_1^{ESs}} + p_2^{ESs}(t)z^{g_2^{ESs}} + p_{32}^{ESs}(t)z^{g_3^{ESs}} \\ &= p_1^{ESs}(t)z^{540} + p_2^{ESs}(t)z^{440} + p_3^{ESs}(t)z^0. \end{aligned}$$

Fig. 5 Energy Storage Serial



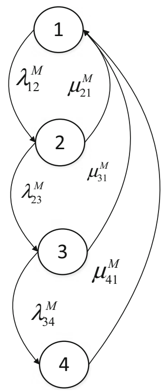


Fig. 6 State space diagram for Motor

4.2.3 Element with 4 States

The system's element, Motor (M), has 4 states: fully working state with performance 540 KW, two partial failure states with performances 480 and 420 KW and fully failure state. The state-space diagram is presented on Fig. 6. To calculate probabilities of each state we build the following system of differential equations:

$$\begin{cases} \frac{dp_1^M(t)}{dt} = -\lambda_{12}^M p_1^M(t) + \mu_{21}^M p_2^M(t) + \mu_{31}^M p_3^M(t) + \mu_{41}^M p_4^M(t), \\ \frac{dp_2^M(t)}{dt} = \lambda_{12}^M p_1^M(t) - (\lambda_{23}^M + \mu_{21}^M) p_2^M(t) \\ \frac{dp_3^M(t)}{dt} = \lambda_{23}^M p_2^M(t) - (\lambda_{34}^M + \mu_{31}^M) p_3^M(t) \\ \frac{dp_4^M(t)}{dt} = \lambda_{34}^M p_3^M(t) - \mu_{41}^M p_4^M(t) \end{cases}$$

Initial conditions are

$$p_1^M(0) = 1;\; p_2^M(0) = 0;\; p_3^M(0) = 0;\; p_4^M(0) = 0.$$

We used MATLAB® for numerical solution of this system of DE to obtain probabilities $p_1^M(t)$, $p_2^M(t)$, $p_3^M(t)$, $p_4^M(t)$. Therefore, for such system's element the output performance stochastic processes can be obtained as follows:

$$\begin{cases} \mathbf{g}^M = \{g_1^M, g_2^M, g_3^M, g_4^M\} = \{540, 480, 420, 0\}, \\ \mathbf{p}^M(t) = \{p_1^M(t), p_2^M(t), p_3^M(t), p_4^M(t)\}. \end{cases}$$

Sets $\mathbf{g}^M, \mathbf{p}^M(t)$ define Lz-transforms for each element as follows:

$$\begin{aligned} L_z\{g^M(t)\} &= p_1^M(t) z^{g_1^M} + p_2^M(t) z^{g_2^M} + p_3^M(t) z^{g_3^M} + p_4^M(t) z^{g_4^M} \\ &= p_1^M(t) z^{540} + p_2^M(t) z^{480} + p_3^M(t) z^{420} + p_4^M(t) z^0. \end{aligned}$$

4.3 Lz-Transform for Serial Topology System

4.3.1 Sub-System 1 (SS_{1s})

According to Fig. 2, Sub-system 1 consists of Fuel Tank, Gas Turbine and Generator, connected in series. Using the composition operator Ω_{fser}, we obtain the Lz-transform $L_z\{G^{SS_{1s}}(t)\}$ for the Sub-system 1, where the powers of z are found as minimum of powers of corresponding terms:

$$\begin{aligned} L_z\{G^{SS_{1s}}(t)\} &= \Omega_{f_{ser}}\left(L_z\{g^{FT}(t)\}, L_z\{g^{GT}(t)\}, L_z\{g^{G}(t)\}\right) \\ &= \Omega_{f_{ser}}\left(p_1^{FT}(t)z^{540} + p_2^{FT}(t)z^0, p_1^{GT}(t)z^{540} + p_2^{GT}(t)z^0, p_1^{G}(t)z^{540} + p_2^{G}(t)z^0\right) \\ &= p_1^{FT}(t)p_1^{GT}(t)p_1^{G}(t)z^{540} + \left(1 - p_1^{FT}(t)p_1^{GT}(t)p_1^{G}(t)\right)z^0. \end{aligned}$$

Using the following notations

$$\begin{aligned} P_1^{SS_{1s}}(t) &= p_1^{FT}(t)p_1^{GT}(t)p_1^{G}(t); \\ P_2^{SS_{1s}}(t) &= 1 - p_1^{FT}(t)p_1^{GT}(t)p_1^{G}(t); \end{aligned}$$

we obtain the resulting Lz-transform for the Sub-system 1 in the following form

$$L_z\{G^{SS_{1s}}(t)\} = P_1^{SS_{1s}}(t)z^{540} + P_2^{SS_{1s}}(t)z^0. \quad (5)$$

4.3.2 Sub-System 2 (SS_{2s})

Sub-system 2 consists of Converter, Motor and Control Unit, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, as before we obtain the Lz-transform $L_z\{G^{SS_{2s}}(t)\}$ for Sub-system 2 as follows:

$$\begin{aligned} L_z\{G^{SS_{2s}}(t)\} &= \Omega_{f_{ser}}\left(L_z\{g^{C}(t)\}, L_z\{g^{M}(t)\}, L_z\{g^{CU}(t)\}\right) \\ &= \Omega_{f_{par}}\left(p_1^{C}(t)z^{540} + p_2^{C}(t)z^0, p_1^{M}(t)z^{540} + p_2^{M}(t)z^{480} + p_3^{M}(t)z^{420} + p_4^{M}(t)z^0,\right. \\ &\quad \left. p_1^{CU}(t)z^{540} + p_2^{CU}(t)z^0\right). \end{aligned}$$

Using notations

$$\begin{aligned} P_1^{SS_{2s}}(t) &= p_1^{C}(t)p_1^{M}(t)p_1^{CU}(t); \\ P_2^{SS_{2s}}(t) &= p_1^{C}(t)p_2^{M}(t)p_1^{CU}(t); \\ P_3^{SS_{2s}}(t) &= p_1^{C}(t)p_3^{M}(t)p_1^{CU}(t); \\ P_4^{SS_{2s}}(t) &= p_1^{C}(t)p_4^{M}(t)p_1^{CU}(t) + p_1^{C}(t)p_2^{CU}(t) + p_2^{C}(t); \end{aligned}$$

we obtain the resulting *Lz*-transform for the Sub-system 2 in the following form

$$L_z\left\{G^{SS_{2s}}(t)\right\} = P_1^{SS_{2s}}(t)z^{540} + P_2^{SS_{2s}}(t)z^{480} + P_3^{SS_{2s}}(t)z^{420} + P_4^{SS_{2s}}(t)z^{0}.$$

4.3.3 Serial Topology System (STS)

Serial Topology System consists of Sub-system 1 and Sub-system 2, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, in the same way as before we obtain the *Lz*-transform $L_z\left\{G^{STS}(t)\right\}$ for Serial Topology System as follows:

$$\begin{aligned} L_z\left\{G^{STS}(t)\right\} &= \Omega_{f_{ser}}\left(L_z\left\{G^{SS_{1s}}(t)\right\}, L_z\left\{G^{SS_{2s}}(t)\right\}\right) \\ &= \Omega_{f_{ser}}\left(P_1^{SS_{1s}}(t)z^{540} + P_2^{SS_{1s}}(t)z^{0},\right. \\ &\quad \left. P_1^{SS_{2s}}(t)z^{540} + P_2^{SS_{2s}}(t)z^{480} + P_3^{SS_{2s}}(t)z^{420} + P_4^{SS_{2s}}(t)z^{0}\right) \end{aligned}$$

Using notations

$$\begin{aligned} P_1^{STS}(t) &= P_1^{SS_{1s}}(t)P_1^{SS_{2s}}(t); \\ P_2^{STS}(t) &= P_1^{SS_{1s}}(t)P_2^{SS_{2s}}(t); \\ P_3^{STS}(t) &= P_1^{SS_{1s}}(t)P_3^{SS_{2s}}(t); \\ P_4^{STS}(t) &= P_1^{SS_{1s}}(t)P_4^{SS_{2s}}(t) + P_2^{SS_{1s}}(t); \end{aligned}$$

we obtain the resulting *Lz*-transform for the Serial Topology System in the following form

$$L_z\left\{G^{STS}(t)\right\} = P_1^{STS}(t)z^{540} + P_2^{STS}(t)z^{480} + P_3^{STS}(t)z^{420} + P_4^{STS}(t)z^{0}. \tag{6}$$

4.3.4 Serial Topology System with Energy Storage (STS-ESs)

Serial Topology System with Energy Storage consists of Energy Storage and Sub-system 2, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, as in previous instances we obtain the *Lz*-transform $L_z\left\{G^{STS-ESs}(t)\right\}$ for Serial Topology System with Energy Storage as follows:

$$\begin{aligned} L_z\left\{G^{STS-ESs}(t)\right\} &= \Omega_{f_{ser}}\left(L_z\left\{g^{ESs}(t)\right\}, L_z\left\{G^{SS_{1s}}(t)\right\}\right) \\ &= \Omega_{f_{par}}\left(p_1^{ESs}(t)z^{540} + p_2^{ESs}(t)z^{440} + p_3^{ESs}(t)z^0,\right. \\ &\quad \left. P_1^{SS_{1s}}(t)z^{540} + P_2^{SS_{1s}}(t)z^{480} + P_3^{SS_{1s}}(t)z^{420} + P_4^{SS_{1s}}(t)z^0\right). \end{aligned}$$

Using notations

$$\begin{aligned} &P_1^{STS-ESs}(t) = p_1^{ESs}(t)P_1^{SS_{1s}}(t); \\ &P_2^{STS-ESs}(t) = p_1^{ESs}(t)P_2^{SS_{1s}}(t); \\ &P_3^{STS-ESs}(t) = p_2^{ESs}(t)\left(P_1^{SS_{1s}}(t) + P_2^{SS_{1s}}(t)\right); \\ &P_4^{STS-ESs}(t) = \left(p_1^{ESs}(t) + p_2^{ESs}(t)\right)P_3^{SS_{1s}}(t); \\ &P_5^{STS-ESs}(t) = \left(p_1^{ESs}(t) + p_2^{ESs}(t)\right)P_4^{SS_{1s}}(t) + P_3^{SS_{1s}}(t); \end{aligned}$$

we obtain the resulting *Lz*-transform for the Serial Topology System with Energy Storage in the following form

$$\begin{aligned} L_z\left\{G^{STS-ESs}(t)\right\} &= P_1^{STS-ESs}(t)z^{540} + P_2^{STS-ESs}(t)z^{480} \\ &\quad + P_3^{STS-ESs}(t)z^{440} + P_4^{STS-ESs}(t)z^{420} + P_5^{STS-ESs}(t)z^0. \end{aligned} \tag{7}$$

4.4 L*z-Transform for Combined Topology System*

4.4.1 Sub-System 1 (SS_{1c})

Following Fig. 3. Sub-system 1 consists of Fuel Tank and Gas Turbine, connected in series. Using the composition operator $\Omega_{f_{ser}}$, we obtain the *Lz*-transform $L_z\left\{G^{SS_{1c}}(t)\right\}$ for the Sub-system 1, where the powers of z are found as minimum of powers of corresponding terms:

$$\begin{aligned} L_z\left\{G^{SS_{1c}}(t)\right\} &= \Omega_{f_{ser}}\left(L_z\left\{g^{FT}(t)\right\}, L_z\left\{g^{GT}(t)\right\}\right) \\ &= \Omega_{f_{ser}}\left(p_1^{FT}(t)z^{540} + p_2^{FT}(t)z^0, p_1^{GT}(t)z^{540} + p_2^{GT}(t)z^0\right) \\ &= p_1^{FT}(t)p_1^{GT}(t)z^{540} + \left(p_2^{FT}(t)p_1^{GT}(t) + p_2^{GT}(t)\right)z^0. \end{aligned}$$

Using the following notations

$$\begin{aligned} &P_1^{SS_{1c}}(t) = p_1^{FT}(t)p_1^{GT}(t); \\ &P_2^{SS_{1c}}(t) = p_2^{FT}(t)p_1^{GT}(t) + p_2^{GT}(t); \end{aligned}$$

we obtain the resulting *Lz*-transform for the Sub-system 1 in the following form

$$L_z\{G^{SS_{1c}}(t)\} = P_1^{SS_{1c}}(t)z^{540} + P_2^{SS_{1c}}(t)z^0.$$

4.4.2 Sub-System 2 (SS$_2$)

Sub-system 2 consists of Generator, Converter and Motor, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, as before we obtain the *Lz*-transform $L_z\{G^{SS_{2c}}(t)\}$ for Sub-system 2 as follows:

$$\begin{aligned} L_z\{G^{SS_{2c}}(t)\} &= \Omega_{f_{ser}}\left(L_z\{G^G(t)\}, L_z\{g^C(t)\}, L_z\{g^M(t)\}\right) \\ &= \Omega_{f_{ser}}\left(p_1^G(t)z^{540} + p_2^G(t)z^0, p_1^C(t)z^{540} + p_2^C(t)z^0,\right. \\ &\quad \left. p_1^M(t)z^{540} + p_2^M(t)z^{480} + p_3^M(t)z^{420} + p_4^M(t)z^0\right) \end{aligned}$$

Using notations

$$\begin{aligned} P_1^{SS_{2c}}(t) &= p_1^G(t)p_1^C(t)p_1^M(t); \\ P_2^{SS_{2c}}(t) &= p_1^G(t)p_1^C(t)p_2^M(t); \\ P_3^{SS_{2c}}(t) &= p_1^G(t)p_1^C(t)p_3^M(t); \\ P_4^{SS_{2c}}(t) &= p_1^G(t)p_1^C(t)p_4^M(t) + p_1^G(t)p_2^C(t) + p_2^G(t); \end{aligned}$$

we obtain the resulting *Lz*-transform for the Sub-system 2 in the following form

$$L_z\{G^{SS_{2c}}(t)\} = P_1^{SS_{2c}}(t)z^{540} + P_2^{SS_{2c}}(t)z^{480} + P_3^{SS_{2c}}(t)z^{420} + P_4^{SS_{2c}}(t)z^0.$$

4.4.3 Sub-System 3 (SS$_3$)

Sub-system 3 consists of Sub-system 2 and Speed Reducer, connected in parallel, where Speed Reducer is used as backup element. Using the composition operator Ω_{fpar}, where powers of z are calculated as maximum values of powers of corresponding terms, in the same way as before we obtain the *Lz*-transform $L_z\{G^{SS_{3c}}(t)\}$ for Sub-system 3 as follows:

$$\begin{aligned} L_z\{G^{SS_{3c}}(t)\} &= \Omega_{f_{ser}}\left(L_z\{G^{SS_{2c}}(t)\}, L_z\{g^{SR}(t)\}\right) \\ &= \Omega_{f_{par}}\left(P_1^{SS_{2c}}(t)z^{540} + P_2^{SS_{2c}}(t)z^{480} + P_3^{SS_{2c}}(t)z^{420} + P_4^{SS_{2c}}(t)z^0,\right. \\ &\quad \left. p_1^{SR}(t)z^{540} + p_2^{SR}(t)z^0\right). \end{aligned}$$

Using notations

$$P_1^{SS_{3c}}(t) = p_1^{SR}(t) + P_1^{SS_{2c}}(t)p_2^{SR}(t);$$
$$P_2^{SS_{3c}}(t) = P_2^{SS_{2c}}(t)p_2^{SR}(t);$$
$$P_3^{SS_{3c}}(t) = P_3^{SS_{2c}}(t)p_2^{SR}(t);$$
$$P_4^{SS_{3c}}(t) = P_4^{SS_{2c}}(t)p_2^{SR}(t);$$

we obtain the resulting *Lz*-transform for the Sub-system 3 in the following form

$$L_z\left\{G^{SS_{3c}}(t)\right\} = P_1^{SS_{3c}}(t)z^{540} + P_2^{SS_{3c}}(t)z^{480} + P_3^{SS_{3c}}(t)z^{420} + P_4^{SS_{3c}}(t)z^{0}.$$

4.4.4 Combined Topology System (CTS)

Combined Topology System consists of Sub-System 1, Subsystem 3 and Control Unit, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, as in previous instances we obtain the *Lz*-transform $L_z\left\{G^{CTS}(t)\right\}$ for Combined Topology System as follows:

$$\begin{aligned} L_z\left\{G^{CTS}(t)\right\} &= \Omega_{f_{ser}}\left(L_z\left\{G^{SS_1}(t)\right\}, L_z\left\{G^{SS_{3c}}(t)\right\}, L_z\left\{g^{CU}(t)\right\}\right) \\ &= \Omega_{f_{ser}}\left(P_1^{SS_{1c}}(t)z^{540} + P_2^{SS_{1c}}(t)z^{0},\right. \\ &\quad \left. P_1^{SS_{3c}}(t)z^{540} + P_2^{SS_{3c}}(t)z^{480} + P_3^{SS_{3c}}(t)z^{420} + P_4^{SS_{3c}}(t)z^{0}, p_1^{CU}(t)z^{540} + p_2^{CU}(t)z^{0}\right). \end{aligned}$$

Using notations

$$P_1^{CTS}(t) = P_1^{SS_{1c}}(t)P_1^{SS_{3c}}(t)p_1^{CU}(t);$$
$$P_2^{CTS}(t) = P_1^{SS_{1c}}(t)P_2^{SS_{3c}}(t)p_1^{CU}(t);$$
$$P_3^{CTS}(t) = P_1^{SS_{1c}}(t)P_3^{SS_{3c}}(t)p_1^{CU}(t);$$
$$P_4^{CTS}(t) = P_1^{SS_{1c}}(t)\left(1 - P_4^{SS_{3c}}(t)\right)p_2^{CU}(t) + P_1^{SS_{1c}}(t)P_4^{SS_{3c}}(t) + P_2^{SS_{1c}}(t);$$

we obtain the resulting *Lz*-transform for the Combined Topology System in the following form

$$L_z\left\{G^{CTS}(t)\right\} = P_1^{CTS}(t)z^{540} + P_2^{CTS}(t)z^{480} + P_3^{CTS}(t)z^{420} + P_4^{CTS}(t)z^{0}. \tag{8}$$

4.4.5 Combined Topology System with Energy Storage (CTS-ESc)

Combined Topology System with Energy Storage consists of Combined Energy Storage, Converter, Motor and Control Unit, connected in series. Using the composition operator Ω_{fser}, where powers of z are calculated as minimum values of powers of corresponding terms, in the same way as before, we obtain the Lz-transform $L_z\{G^{CTS-ESc}(t)\}$ for Combined Topology System with Energy Storage as follows:

$$\begin{aligned} L_z\{G^{CTS-ESc}(t)\} &= \Omega_{f_{ser}}\left(L_z\{g^{ESc}(t)\}, L_z\{g^{C}(t)\}, L_z\{g^{M}(t)\}, L_z\{g^{CU}(t)\}\right) \\ &= \Omega_{f_{par}}\left(p_1^{ESc}(t)z^{540} + p_2^{ESc}(t)z^{0}, p_1^{C}(t)z^{540} + p_2^{C}(t)z^{0},\right. \\ &\quad \left. p_1^{M}(t)z^{540} + p_2^{M}(t)z^{480} + p_3^{M}(t)z^{420} + p_4^{M}(t)z^{0},\ p_1^{CU}(t)z^{540} + p_2^{CU}(t)z^{0}\right) \end{aligned}$$

Using simple algebra calculations of the powers of z as minimum values of powers of corresponding terms and the following notations

$$\begin{aligned} &P_1^{CTS-ESc}(t) = p_1^{ESc}(t)p_1^{C}(t)p_1^{M}(t)p_1^{CU}(t); \\ &P_2^{CTS-ESc}(t) = p_1^{ESc}(t)p_1^{C}(t)p_2^{M}(t)p_1^{CU}(t); \\ &P_3^{CTS-ESc}(t) = p_1^{ESc}(t)p_1^{C}(t)p_3^{M}(t)p_1^{CU}(t); \\ &P_4^{CTS-ESc}(t) = p_1^{ESc}(t)p_1^{C}(t)\left(p_2^{CU}(t) + p_4^{M}(t)p_1^{CU}(t)\right) + p_1^{ESc}(t)p_2^{C}(t) + p_2^{ESc}(t); \end{aligned}$$

the whole system's Lz-transform expression is as follows:

$$\begin{aligned} L_z\{G^{CTS-ESc}(t)\} = P_1^{CTS-ESc}(t)z^{540} + P_2^{CTS-ESc}(t)z^{480} \\ + P_3^{CTS-ESc}(t)z^{420} + P_4^{CTS-ESc}(t)z^{0}. \end{aligned} \tag{9}$$

5 Availability and Mean Power Performance Calculation

Using expression (3), the MSS mean instantaneous availability for constant demand level w may be presented as follows:

- For w = 540 KW demand level

$$\begin{aligned} A^{STS}_{w=540KW}(t) &= \sum_{g_k \geq 540} P_k^{STS}(t) = P_1^{STS}(t) \\ A^{CTS}_{w=540KW}(t) &= \sum_{g_k \geq 540} P_k^{CTS}(t) = P_1^{CTS}(t) \end{aligned} \tag{10}$$

- For w = 460 KW demand level

$$\begin{aligned} A^{STS}_{w=460KW}(t) &= \sum_{g_k \geq 460} P_k^{STS}(t) = P_1^{STS}(t) + P_2^{STS}(t) \\ A^{CTS}_{w=460KW}(t) &= \sum_{g_k \geq 460} P_k^{CTS}(t) = P_1^{CTS}(t) + P_2^{CTS}(t) \end{aligned} \tag{11}$$

- For $w = 360$ KW demand level

$$\begin{aligned} A_{w=360KW}^{STS}(t) &= \sum_{g_k \geq 360} P_k^{STS}(t) = \sum_{i=1}^{4} P_i^{STS}(t) \\ A_{w=360KW}^{CTS}(t) &= \sum_{g_k \geq 360} P_k^{CTS}(t) = \sum_{i=1}^{3} P_i^{CTS}(t) \end{aligned} \quad (12)$$

Using expression (4), the MSS instantaneous mean power performance for the Serial and Combined Topology Systems can be obtained as follows:

$$\begin{aligned} E^{STS}(t) &= \sum_{g_s > 0} g_i^{STS} P_i^{STS}(t) = \sum_{i=1}^{4} g_i^{STS} P_i^{STS}(t) \\ &= 540P_1^{STS}(t) + 480P_2^{STS}(t) + 440P_3^{STS}(t) + 420P_4^{STS}(t) \\ E^{CTS}(t) &= \sum_{g_s > 0} g_i^{CTS} P_i^{CTS}(t) = \sum_{i=1}^{3} g_i^{CTS} P_i^{CTS}(t) \\ &= 540P_1^{CTS}(t) + 480P_2^{CTS}(t) + 420P_3^{CTS}(t) \end{aligned} \quad (13)$$

The failure and repair rates (in year^{-1}) of each system's elements are presented in the Table 1.

MSS mean instantaneous availability for different constant demand levels is presented in Figs. 7 and 8. The obtained results shows that difference in instantaneous availability for Serial Topology System for different demand levels is small enough (Fig. 7), for Combined Topology System there is no distinction for different demand levels (Fig. 8).

Comparison MSS mean instantaneous availability levels for different topology shows that Combined Topology is better than Serial Topology (Figs. 9 and 10). In case of emergency and usage of Energy Storage the instantaneous availability of Serial system is greater than Combined Topology System (Fig. 11).

Table 1 Failure and repair rates of each system's elements

	Failure rates (year^{-1})	Repair rates (year^{-1})
Fuel tank (FT)	0.0584	219
Gas turbine (GT)	0.876	159
Generator (G)	0.0145	175.2
Energy Storage Series (ESs)	0.438	438
Energy Storage Combined (ESc)	0.438	250
Electric inverter (C)	0.0584	584
Motor (M)	0.0145	87.6
Speed reducer (SR)	0.0876	116.8
Control unit (CU)	0.01752	730

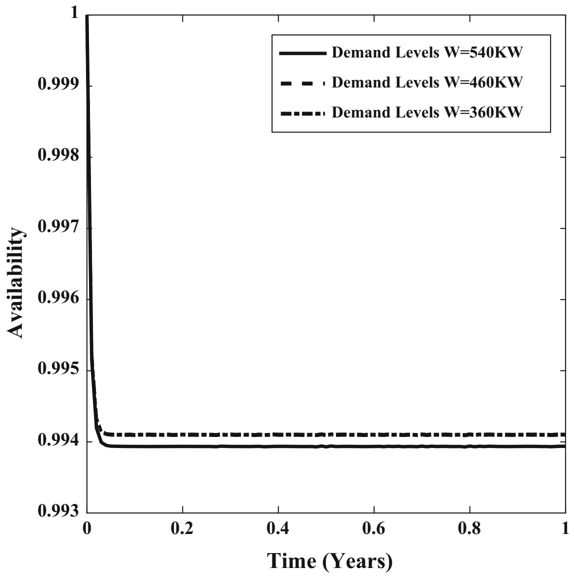


Fig. 7 MSS mean instantaneous availability for Serial Topology System for different demand levels

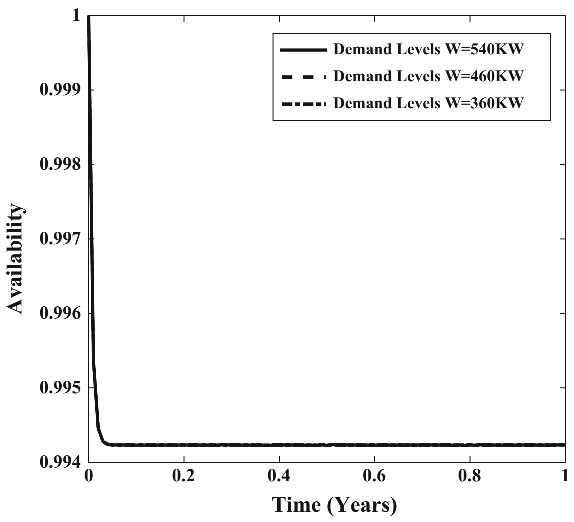


Fig. 8 MSS mean instantaneous availability for Combined Topology System for different demand levels

Fig. 9 MSS mean instantaneous availability for Serial and Combined Topology System for w = 540 KW

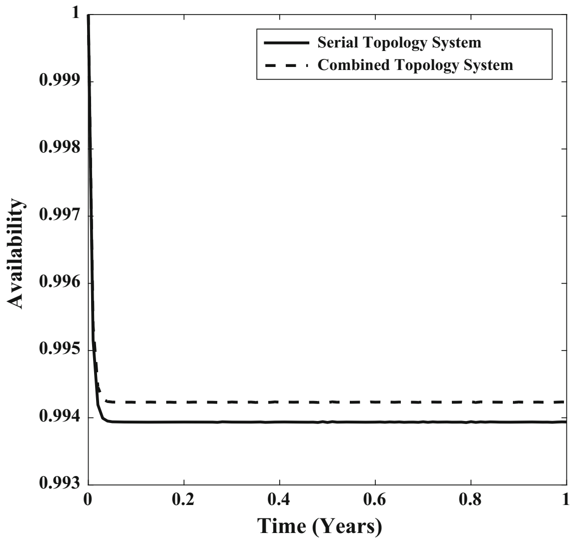
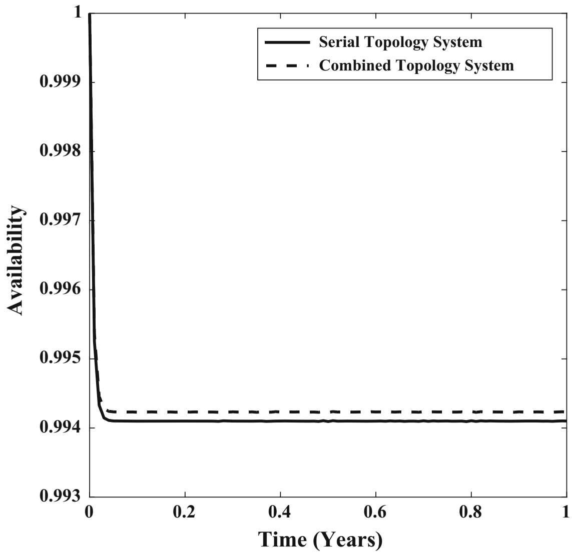


Fig. 10 MSS mean instantaneous availability for Serial and Combined Topology System for w = 460 KW



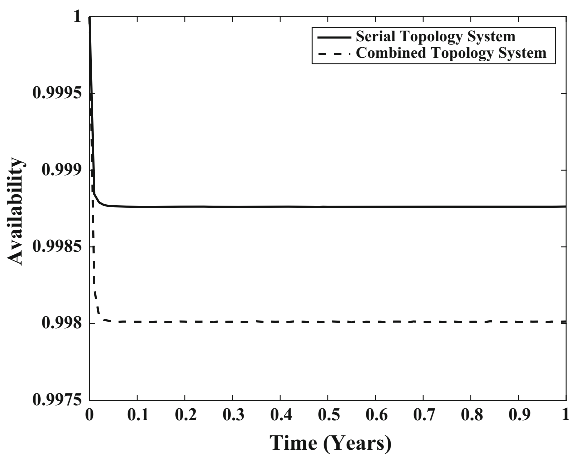


Fig. 11 MSS mean instantaneous availability for Serial and Combined Topology System with energy storage

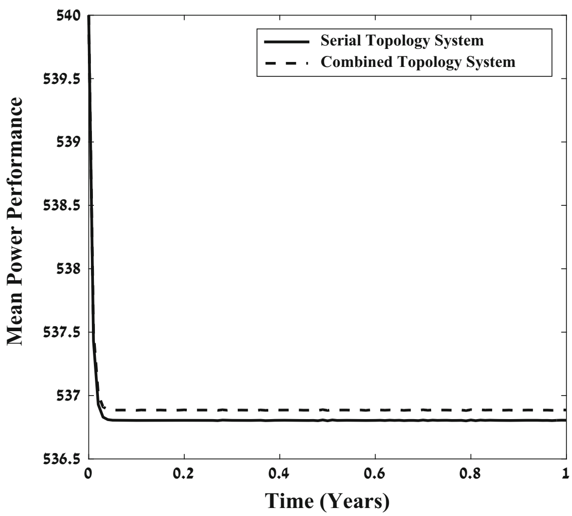


Fig. 12 MSS instantaneous mean power performance of the Serial and Combined Topology Systems

Calculated MSS instantaneous mean power performance of the Serial and Combined Topology System is presented in Fig. 12.

6 Conclusion

In this chapter, the *Lz*-transform method was used for evaluation of two important parameters—availability and performance of various topologies traction drives of hybrid-electric helicopter.

Lz-transform approach extremely simplifies the solution, which in comparison with straightforward Markov method would have required building and solving the model with 128 states for Serial and 512 states for Combined Topology Systems.

The obtained results have showed that in terms of reliability the compared variants of traction drive topologies of hybrid-electric helicopter are close enough and meet the requirements of the project. Considering this, to select the optimal variant, it is necessary to carry out a comparative analysis of their weight and size characteristics.

Nine-phase design of the traction permanent magnets synchronous motor is the best compromise solution between the required level of its fault tolerance and the complexity of manufacturing.

References

1. Bolvashenkov I, Herzog H-G (2016) Use of stochastic models for operational efficiency analysis of multi power source traction drives. In: Frenkel I, Lisnianski A (eds) Proceedings of the second international symposium on stochastic models in reliability engineering, life science and operations management, SMRLO'16, Beer Sheva, Israel, 15–18 Feb 2016. IEEE CPS, pp 124–130. doi:10.1109/SMRLO.2016.30
2. Bolvashenkov I, Frenkel I, Kammermann J, Herzog H-G (2017) The choice of an optimal structure and parameters of energy storage for an electrical helicopter traction drive. In: Proceedings of IEEE twelfth international conference on ecological vehicles and renewable energies (EVER'17), 11–13 Apr 2017, Monaco, pp 1–6
3. Bolvashenkov I, Kammermann J, Lahlou T, Herzog H-G (2016) Comparison and choice of a fault tolerant inverter topology for the traction drive of an electrical helicopter. In: 3rd international conference on electrical systems for aircraft, railway, ship propulsion, and road vehicles (ESARS'16), 02–04 Nov 2016, Toulouse, France, pp 1–6
4. Bolvashenkov I, Kammermann J, Willerich S, Herzog H-G (2016) Comparative study of reliability and fault tolerance of multi-phase permanent magnet synchronous motors for safety-critical drive trains. In: Proceedings of international conference on renewable energies and power quality (ICREPQ'16), 4–6 May Madrid, Spain, pp 1–6
5. Frenkel I, Bolvashenkov I, Herzog H-G, Khvatskin L (2016) Performance availability assessment of combined multi power source traction drive considering real operational conditions. Transp Telecommun 17(3):179–191. doi:10.1515/ttj-2016-0016
6. Frenkel I, Bolvashenkov I, Herzog H-G, Khvatskin L (2017) Operational sustainability assessment of multipower source traction drive. In: Mathematics applied to engineering. Elsevier Ltd. pp 191–203. doi:10.1016/B978-0-12-810998-4.00010-7
7. Huang H (2015) Challenges in more electric aircraft (MEA). IEEE Transp Electr Newslett, July/August, 2015
8. Jia H, Jin W, Ding Y, Song Y, Yu D (2017) Multi-state time-varying reliability evaluation of smart grid with flexible demand resources utilizing *Lz*-transform. In: Proceeding of the international conference on energy engineering and environmental protection (EEEP2016), IOP Publishing, IOP conference series: Earth and Environmental Science, vol 52, p 012011. doi:10.1088/1755-1315/52/1/012011
9. Lahne HC, Staton D, Chong YC, Gerling D (2016) Design of a 50000 rpm high-speed high-power six-phase PMSM for use in aircraft applications. In: Proceedings of IEEE eleventh international conference on ecological vehicles and renewable energies (EVER'16), 06–08 Apr 2016, Monaco, pp 1–11

10. Lisnianski A (2012) *LZ*-transform for a discrete-state continuous-time markov process and its application to multi-state system reliability. In: Lisnianski A, Frenkel I (eds) Recent advances in system reliability. Signatures, multi-state systems and statistical inference. Springer, London, pp 79–95
11. Lisnianski A, Ben Haim H (2013) Short-term reliability evaluation for power stations by using Lz-transform. J Mod Power Syst Clean Energy 1(2):110–117. doi:10.1007/s40565-013-0021-3
12. Lisnianski A, Frenkel I, Ding Y (2010) Multi-state system reliability analysis and optimization for engineers and industrial managers. Springer, London
13. Natvig B (2011) Multistate systems reliability. Theory with applications. Wiley, New York
14. Rieger H. Hubschraubertyp Airbus Helicopters EC 135. http://www.rth.info/typen/typen.php?show=ec135
15. Rick H (2013) Gasturbinen und Flugantriebe: Grundlagen, Betriebsverhalten and Simulation. Springer, Berlin
16. Sheu S-H, Zang ZC (2013) An optimal age replacement policy for multi-state systems. IEEE Trans Reliab 62(3):722–733
17. Trivedi K (2002) Probability and statistics with reliability, queuing and computer science applications. Wiley, New York
18. Ushakov I (1986) A universal generating function. Sov J Comput Syst Sci 24:37–49

Patient Diagnostic State Evolution During Hospitalization: Developing a Model for Measuring Clinical Diagnostic Dynamics

Yariv N. Marmor and Emil Bashkansky

Abstract Patient health is represented by a set of diagnoses, which determines personal health status. Each set corresponds to a certain health state and so, can be treated as an individual performance in this state and individual health can be considered as a corresponding multi-state system. Appropriate metrics for measuring patient's state diagnosis changes during hospitalization are proposed. The first metric determines the dissimilarity between two single diagnoses, each of which is based on internationally recognized classification scheme. The second metric is aimed to compare between two sets of diagnoses with respect to the same patient and is based on the first metric, but uses additional, recently proposed, ideas of measuring heterogeneity/segregation between sets of categorical data. A numerical example and a real world illustration of the above measures are provided. The ultimate goal is the analysis of multistate health status data in order to improve the accuracy and quality of medical diagnostics.

Keywords Medical diagnosis · Accuracy · Misclassification · Dissimilarity · Healthcare quality

1 Introduction

In order to assess patients' health status and its evolution, vital signs, diagnostics, error analysis in the diagnosis process and many more, appropriate metrics must be developed and monitored. According to World Health Organization (WHO), "health is a state of complete physical, mental and social well-being… ". Setting aside the social component, it should be noted that the medical diagnosis involves

Y.N. Marmor · E. Bashkansky (✉)
Industrial Engineering and Management Department,
ORT Braude College of Engineering, Karmiel, Israel
e-mail: ebashkan@braude.ac.il

Y.N. Marmor
e-mail: myariv@braude.ac.il

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_18

thousands of physical and mental conditions or deceases—the presence or absence of which, determines the individual state of health. Specific state of individual health can be determined by a set of his or her diagnosis [2], which, by the way, could change over time: one can get sick with new disease or recover from an old one [5]. Chronical diseases may affect person's ability to perform certain tasks. That ability might gradually decrease with time (age) and thus can correlate to the life evolution of any other system operating in an external environment. Generally, patient health is represented by a set of diagnoses, which determines personal health status. Each set corresponds to a certain health state and so, can be treated as an individual performance in this state and individual health can be considered as a corresponding multi-state system. Each health state can be characterized according to the International Classification of Diseases (ICD). The last edition of ICD (the 10th) contains codes for thousands of different diseases. ICD has a hierarchical tree structure. By help of ICD medical diagnosis is assigned an appropriate code number, which essentially means selecting a suitable node—one or more—of the ICD tree.

For example, the first level of the 2015 ICD-9-CM (the 9th ICD edition that is currently being used) splits diagnosis codes for diseases and related health problems to 19 categories. The first category with codes 001–139 relates to *Infectious and parasitic diseases*, codes 280–289 relate to category of *Diseases of blood and blood-forming organs* (see Fig. 1 for partial overlook) and so on. In turns, *Diseases of blood* splits into 3 sub categories—*Anemia* (280–285), *Coagulation/hemorrhagic* (286–287) and *Other* (288–289). *Anemia* splits again into 6 subcategories such as *Iron deficiency anemias* (280), *Other deficiency anemias* (281) and so on. *Coagulation/hemorrhagic* (286–287), on the other hand, splits only into 2 subcategories—*Coagulation defects* (286) and *Purpura and other hemorrhagic conditions*

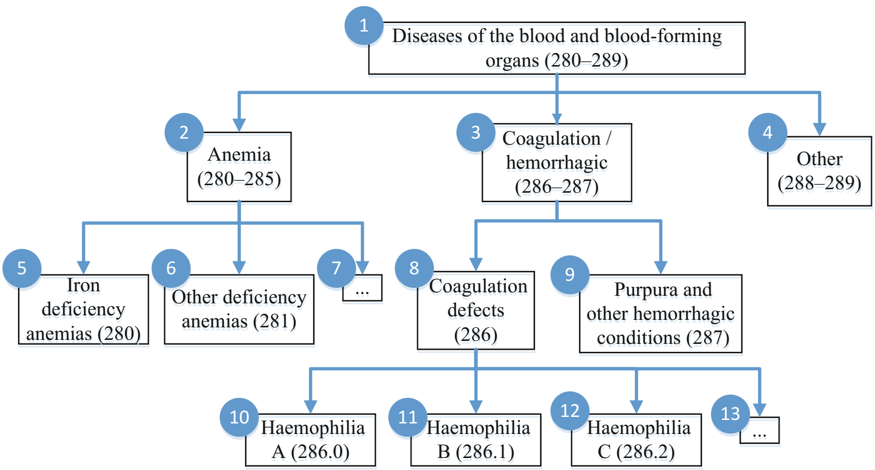


Fig. 1 Diseases of blood and blood-forming organs ICD-9 branch

(287). *Coagulation defects* (286) splits further into *Haemophilia A* (286.0), *Haemophilia B* (286.1) and so on. *Haemophilia A* (286.0), for example, is a final diagnosis—no additional split is available.

2 Some Basic Definitions

To develop an appropriate measure for diagnosis change, such as the one expressed in the mismatch between final and initial diagnoses, let us establish some basic terminology conceptions in relation to tree structure—a set of *nodes* (vertices) and *edges* (connections between two nearest linked nodes) as presented in Fig. 1 consisting of 13 nodes and 12 edges.

- Siblings: A group of nodes with the same "parent", for example nodes 2, 3 and 4 with node 1 as a "parent".
- Descendant: A node reachable by repeated proceeding downwards from the node, for example, node 10 is a descendant of node 3.
- Ancestor: A node reachable by repeated proceeding upwards from the node, for example, node 3 is an ancestor of node 10.
- Edge: Unordered connection between one node and a nearest linked another: node 8 and node 11, for example.
- Path: A sequence of nodes and edges connecting a descendant node with an ancestor node, e.g., if a descendant node is 10 and an ancestor is node 3, then their path is 10 → 8 → 3.
- Path length: The number of the path edges, for example 10 → 8 → 3 → 1 path length equals three. Two siblings have the same path length.
- Depth: The depth of a certain node is the number of edges from the certain tree's node to the root node (node 1 in our example). For example, the depth of node 10 is three.

The purpose of this Chapter is to develop metrics allowing to assess the extent of dissimilarity or distance between two diagnoses in relation to the same patient. We expect that the proposed metrics can be useful for different purposes, for example, when comparing diagnoses before and after some medical examination or while monitoring patient's health state evolution, etc.

3 The Distance Between Two Diagnoses as a Measure of Their Dissimilarity

Park et al. [6] defines the distance between two diagnoses, represented by nodes *A* and *B*, as the sum of the path length from *node A* and *node B* to the closest common ancestor *node C*, divided by that sum plus double of the depth of *node C* +1:

$$L(A,B) = \frac{l_{A,C} + l_{B,C}}{(l_{A,C} + l_{B,C}) + 2 \cdot (d_C + 1)} \quad (1)$$

where:

C is the lowest common ancestor regarding A and B
$l_{A,C}$ is a path length from A to C
$l_{B,C}$ is a path length from B to C
d_C is a depth of node C

Distance defined by Eq. (1) presents the, so-called, *semantic distance* and fulfills the following properties [7]:

- Symmetry—The distance from A to B is equal the distance from B to A: $L(A, B) = L(B, A)$.
- Normalization: The distance from A to B is zero when the two nodes are identical and is never greater than one. The larger the value is, the further the two nodes are: $0 \leq L(A, B) < 1$.
- For each 3 nodes A, B, and C, the distance between A and B is smaller than the sum of the distances from A to C and C to B (triangle rule): $L(A, B) < L(A, C) + L(B, C)$.

To demonstrate the rules, we used the tree in Fig. 1 and found that the closer the nodes are, the smaller is the distance:

- $L(node\,2, node\,5) = \frac{1+0}{1+0+2 \cdot (1+1)} \approx 0.33$
- $L(node\,2, node\,3) = \frac{1+1}{1+1+2 \cdot (0+1)} = 0.5$
- $L(node\,5, node\,8) = \frac{2+2}{2+2+2 \cdot (0+1)} \approx 0.67$
- $L(node\,5, node\,10) = \frac{2+3}{2+3+2 \cdot (0+1)} \approx 0.71$

Consequently, observe that the triangle rule is applied:

$$L(node\,5, node\,3) = \frac{2+1}{2+1+2\cdot(0+1)} = 0.6 < [L(node\,2, node\,3) = 0.5] + [L(node2, node5) \approx 0.33]$$

4 Some Facts Concerning the Similarities and Differences Between Two Consecutives Diagnoses in Hospital

Patients flow in the hospital usually starts in the emergency department (ED), where patients get initial diagnosis and if the treatment is satisfied, they go home. If not, they are hospitalized for further medical care [1]. We would like to compare the diagnosis made in the ED within few hours of care to the one that is done in the hospital wards after few days of treatment and additional diagnoses.

In a hospital we examined, we found that about 78.9% of the major diagnoses given in the ED are identical to the ones given in the ward. 6% of the major diagnoses are totally different and the rest are somewhat in between (see Fig. 2).

Mostly, the major diagnoses are singular, but we wanted to examine the overall diagnoses, considering minor ones as well. In Fig. 3 we see the CDF of the number of diagnoses each patient is evaluated with during his flow in the hospital (based on 90,250 patients)—33.6% of the set of diagnoses consist of only one diagnosis, 21.1% consist of two diagnoses and the rest evaluated with more than 3 diagnoses each visit. To deal with multi diagnoses, we would have to develop a measure of dissimilarity between two groups of diagnoses—initial state, which is evaluated in the ED and the final group of diagnoses, or state, observed in the hospital ward.

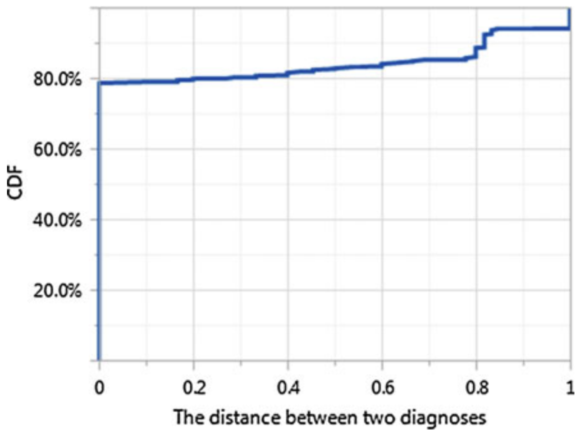
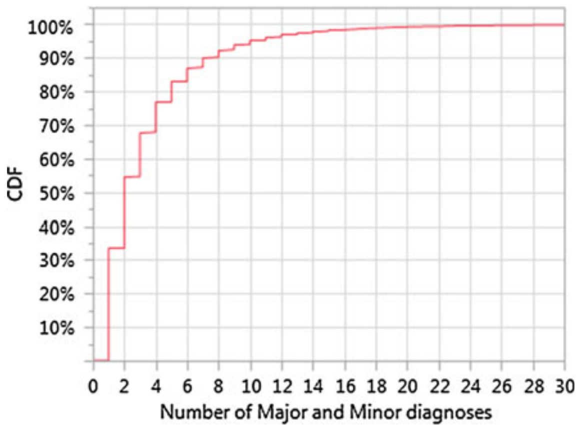


Fig. 2 CDF of the distance between each two major diagnoses in the hospital

Fig. 3 CDF of number of major and minor diagnoses for each patient per ward



5 How to Measure Divergence Between Two Sets of Diagnoses

Our suggestion and following description is based on the ideas expressed in [4]. In this publication, the concept of variation was generalized to any measurement scales, including nominal, as in our case. For the scale with K categories, designated by codes c_k, $1 \leq k \leq K$, for which the two-argument, non-negative, symmetric dissimilarity (distance) function $L(c_i, c_j)$ is defined on the codes, such that $L(c_k, c_k) = 0$, this variation is determined as:

$$V = \sum_{i=1}^{K} \sum_{j=1}^{K} L(c_i,\ c_j) p_i p_j \tag{2}$$

where p_k means the proportion of the data belonging to the k-th category.

The main result of the above-mentioned publication is the total-variation decomposition theorem, i.e., the possibility of splitting the total variation to the *intra* (within sets of data) and the *inter* (between sets of data) components, so that:

$$V_{TOTAL} = V_{WITHIN} + V_{BETWEEN} \tag{3}$$

where V_{WITHIN} is the weighted average of group/set variations, with a weight reflecting the proportion of data in the group (the number of data in the group divided to the total number of data). For our purpose, the most interesting is the second component $V_{BETWEEN}$, presenting dissimilarity between sets of data (two, in our case). For its calculation details, we refer the reader to [4]; in the next section we bring its simplified version for the case when there are only two data sets, or two lists of deceases, marked as the initial and the final diagnoses. Variation $V_{BETWEEN}$ itself may not be an appropriate measure of the difference between the two groups, as it depends on the dimension of scale (which, in our case, may vary from patient to patient). For this reason, the above-mentioned publication suggested as a suitable difference measure the so-called segregation index—SP:

$$SP = \frac{V_{BETWEEN}/df_{BETWEEN}}{V_{TOTAL}/df_{TOTAL}} \tag{4}$$

where the symbol *df* is used to indicate the corresponding degrees of freedom. In our case $df_{BETWEEN} = 1$ and df_{TOTAL} is the total number of diagnoses in both lists (initial and final) minus one. In the following section, we provide an example of segregation index calculations.

6 Example of Segregation Index Calculations

Let us illustrate the proposed measure with a simple example. Assume, for simplicity, that the range of possible diagnoses of diseases consists of $K = 5$ diseases conventionally referred to as A, B, C, D, and E.

- Initial set # I consists of 3 elements {A, C, E} and does not include 2 elements {B, D}.
- Final set # II consists of four elements {A, B, C, D} and does not include element {E}.
- Total combined set I&II consists of seven elements {2A, B, 2C, D, E} dispersed on the same five categorical scale.

We can assign to the above sets, defined on the predetermined five categorical scale, the following vectors of numbers:

- Initial set # I: $n_I = \{1, 0, 1, 0, 1\}$ with sum of modules of all its elements n_I equal 3.
- Final set # II: $n_{II} = \{1, 1, 1, 1, 0\}$ with sum of modules of all its elements n_{II} equal 4.
- Total combined set I&II: $n_T = \{2, 1, 2, 1, 1\}$ dispersed on the same five categorical scale with sum of modules of all its elements $n_T = n_I + n_{II}$ equal 7 ($df_{TOTAL} = 7–1 = 6$).

For further purposes it is more convenient to work with vectors of proportions—***p***, normalized to one:

- Initial set # I: $p_I = \{\frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3}\}$ with sum of modules of all its elements equals one.
- Final set # II: $p_{II} = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0\}$ with sum of modules of all its elements equals one.
- Total combined set I&II: $p_T = \{\frac{2}{7}, \frac{1}{7}, \frac{2}{7}, \frac{1}{7}, \frac{1}{7}\}$ dispersed on the same five categorical scale with sum of modules of all its elements equals one.

Following [4], we define the *characteristic kernel of the between-sets variation* $\theta_{i,j}$ *as* 5×5 *symmetric square matrix:*

$$\theta_{i,j} = (-1) \cdot [\frac{n_I}{n_T}(p_{i,I} - p_{i,T}) \cdot (p_{j,I} - p_{j,T}) + \frac{n_{II}}{n_T}(p_{i,II} - p_{i,T}) \cdot (p_{j,II} - p_{j,T})] \quad (5)$$

where *i, j* = 1,2,3,4,5 denote the five previously ordered deceases and

$$p_I = \{p_{1,I}, p_{2,I}, \ldots, p_{5,I}\} \quad p_{II} = \{p_{1,II}, p_{2,II}, \ldots, p_{5,II}\}$$
$$p_T = \{p_{1,T}, p_{2,T}, \ldots, p_{5,T}\}$$

For example:

$$\theta_{1,2} = (-1) \cdot [\frac{3}{7} \cdot (\frac{1}{3} - \frac{2}{7}) \cdot (0 - \frac{1}{7}) + \frac{4}{7} \cdot (\frac{1}{4} - \frac{2}{7}) \cdot (\frac{1}{4} - \frac{1}{7})] = 0.005102$$

Finally, we define the variation between the two sets of diagnoses as:

$$V_{BETWEEN} = \sum_{i=1}^{5} \sum_{j=1}^{5} L(i,j) \cdot \theta_{i,j} \quad (6)$$

or generally by:

$$V_{BETWEEN} = \sum_{i=1}^{K} \sum_{j=1}^{K} L(i,j) \cdot \theta_{i,j} \quad (7)$$

where $L(i,j)$ denotes the distance between *i* and *j* deceases as defined by (1).

If we assume that in the above example, the distance between any two mismatched deceases is the same (siblings, for example) and equals to 0.5, then $V_{BETWEEN} = 0.030612$.

Now, let's calculate the total variation following the same assumption. According to (2):

$$V_{TOTAL} = (0.5) \cdot \sum_{i=1}^{K} \sum_{j \neq i}^{K} p_{i,T} \cdot p_{j,T} = (0.5) \cdot \{[\sum_{k=1}^{K} p_{k,T}]^2 - \sum_{k=1}^{K} p_{k,T}^2\} = (0.5) \cdot (1 - \sum_{k=1}^{K} p_{k,T}^2) \quad (8)$$

Given that $p_T = \{\frac{2}{7}, \frac{1}{7}, \frac{2}{7}, \frac{1}{7}, \frac{1}{7}\}$ and substituting it in (8) we get $V_{TOTAL} = 0.3887755$, so the divergence between two diagnoses, measured by means of segregation index *SP* equals:

$$SP = \frac{0.030612}{0.387755/6} \approx 0.474$$

Table 1 Examples for different combinations of initial (I) and final (II) diagnoses sets

Set 1					Set 2									
A	B	C	D	E	A	B	C	D	E	V^a_WI	V^a_WII	V^a_T	V_B	SP
1	0	0	0	0	0	0	0	0	1	0.00	0.00	0.25	0.25	1.00
1	1	0	0	0	0	0	0	0	1	0.25	0.00	0.33	0.17	1.00
1	1	0	0	0	0	0	0	1	1	0.25	0.25	0.38	0.13	1.00
1	1	1	0	0	0	0	0	1	1	0.33	0.25	0.40	0.10	1.00
1	1	1	0	0	0	0	1	1	1	0.33	0.33	0.39	0.06	0.71
1	1	1	1	0	0	0	1	1	1	0.38	0.33	0.39	0.03	0.47
1	1	1	1	0	0	1	1	1	1	0.38	0.38	0.39	0.02	0.28
1	1	1	1	1	0	1	1	1	1	0.40	0.38	0.40	0.01	0.12
1	1	1	1	1	1	1	1	1	1	0.40	0.40	0.40	0.00	0.00

[a]V_W – V_{WITHIN}; V_B – $V_{BETWEEN}$; V_T – V_{TOTAL}

In Table 1 we bring examples for different combinations of initial (I) and final (II) diagnoses. It can be seen, that the more noteworthy the divergence between the initial and final diagnoses is, the larger the segregation index *SP* is.

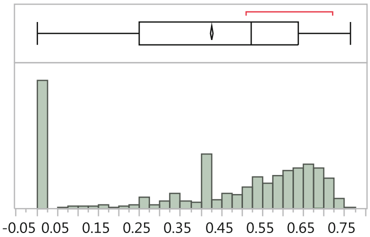
Two extremal cases must be considered separately:

- The first occurs when one of the sets is, for any reason, empty (for example, the data has been lost, or the patient deceased, or vice versa, completely recovered). The possible option is to assign to this case the maximum possible *SP*. This, however, is a special case.
- The second, more frequent, occurs when in both sets/diagnoses only one and the **same** matching disease is noted. Formally, this situation is expressed by uncertainty arising from the division of zero by zero. In this case, we assign $SP = 0$ ultimately.

7 Applying Proposed Measures to Patients Which Passed Only One Ward

To demonstrate the use of segregated dissimilarity measures, we considered only patients that went through exactly one ward after hospitalization, meaning they got to the ED and then went to one of the hospital wards and then left the hospital. In order to evaluate multiple locations, we have had to develop relevant measure as well.

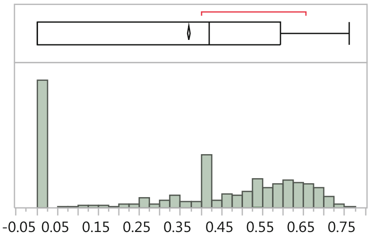
As we see in Figs. 4, 5, 6, 7 and 8, the total variation, the variation within each set and the overall variation within groups have a similar distribution. All stretch from 0 to 0.76, when the most common value is 0. Consequently, the variation



Quantiles		
100.0%	maximum	0.7641974242
99.5%		0.737470112
97.5%		0.7192957265
90.0%		0.6887550005
75.0%	quartile	0.6368484848
50.0%	median	0.5228282828
25.0%	quartile	0.25
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.4260223
Std Dev	0.2564783
Std Err Mean	0.0017491
Upper 95% Mean	0.4294508
Lower 95% Mean	0.4225939
N	21501

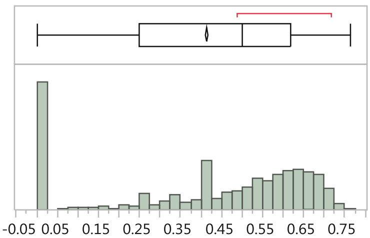
Fig. 4 V_{WITHIN} I distribution



Quantiles		
100.0%	maximum	0.7614971558
99.5%		0.7273361971
97.5%		0.7029115654
90.0%		0.6603769893
75.0%	quartile	0.5931568138
50.0%	median	0.4202020202
25.0%	quartile	0
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.3702654
Std Dev	0.2584488
Std Err Mean	0.0017619
Upper 95% Mean	0.3737189
Lower 95% Mean	0.366812
N	21517

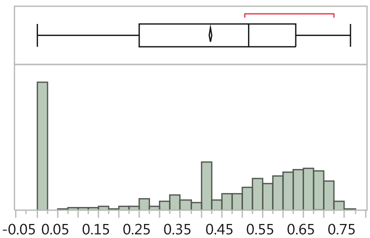
Fig. 5 V_{WITHIN} II distribution



Quantiles		
100.0%	maximum	0.763027
99.5%		0.73163957
97.5%		0.7106009
90.0%		0.6763156
75.0%	quartile	0.6198325
50.0%	median	0.500926
25.0%	quartile	0.25
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.4140815
Std Dev	0.2509548
Std Err Mean	0.0017115
Upper 95% Mean	0.4174361
Lower 95% Mean	0.4107269
N	21501

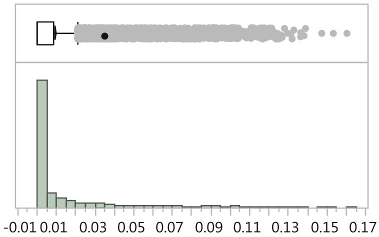
Fig. 6 V_{WITHIN} distribution



Quantiles		
100.0%	maximum	0.764691739
99.5%		0.7364997346
97.5%		0.7178749887
90.0%		0.6857653525
75.0%	quartile	0.6316253965
50.0%	median	0.5150766929
25.0%	quartile	0.25
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.423125
Std Dev	0.2548563
Std Err Mean	0.0017374
Upper 95% Mean	0.4265304
Lower 95% Mean	0.4197195
N	21517

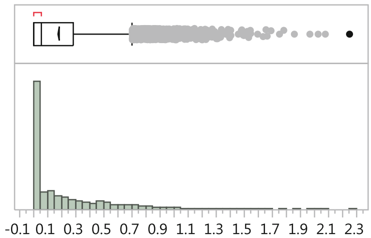
Fig. 7 V_{TOTAL} distribution



Quantiles		
100.0%	maximum	0.160363
99.5%		0.103536
97.5%		0.08464245
90.0%		0.031766
75.0%	quartile	0.008361
50.0%	median	0
25.0%	quartile	0
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.0093537
Std Dev	0.0201917
Std Err Mean	0.0001377
Upper 95% Mean	0.0096236
Lower 95% Mean	0.0090838
N	21501

Fig. 8 $V_{BETWEEN}$ distribution



Quantiles		
100.0%	maximum	2.253737619
99.5%		1.1328345716
97.5%		0.8669317638
90.0%		0.5623876477
75.0%	quartile	0.281335011
50.0%	median	0.0516841519
25.0%	quartile	0
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics	
Mean	0.1792285
Std Dev	0.2570492
Std Err Mean	0.0019863
Upper 95% Mean	0.1831219
Lower 95% Mean	0.1753352
N	16747

Fig. 9 SP distribution

between is very small, more than half with zeroes (Fig. 8). On the other hand, SP (Fig. 9) is not limited by 1, therefore we see values above 2 in our data, although again the most common value is zero.

8 Brief Summary and Future Research

In this work, we suggested a way to measure the dissimilarity between two deceases on an ICD hierarchical tree structure. Then, we add a divergence measure between two sets of diagnoses defined for the same patient on the same structure based on the idea expressed in [4]. We illustrate the proposed measures on a dataset of patients' diagnoses that passed only through one ward.

The drive to this work come from the need to evaluate the clinical quality that is given in EDs and later on to associate it with factors that can be controlled, such as load and staff expertise or seniority.

Despite the advancement in technology, the number of misclassifications is not going down, see for example the case of appendicitis in [3]. This means that further work is needed in order to find the cause of misclassifications. We believe that examining the variability between sets of patient diagnoses might give a better understanding of the cause than only focusing on errors in diagnosis (presence or absence of a disease).

References

1. Armony M, Israelit S, Mandelbaum A, Marmor YN, Tseytlin Y, Yom-Tov GB (2015) On patient flow in hospitals: a data-based queueing-science perspective. Stoch Syst 5(1):146–194
2. Deyo RA, Cherkin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 45(6):613–619
3. Flum DR, Morris A, Koepsell T, Dellinger EP (2001) Has misdiagnosis of appendicitis decreased over time?: a population-based analysis. JAMA 286(14):1748–1753
4. Gadrich T, Bashkansky E, Zitikis R (2015) Assessing variation: a unifying approach for all scales of measurement. Qual Quant 49(3):1145–1167
5. o'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM (2005) Measuring diagnoses: ICD code accuracy. Health Serv Res 40(5p2):1620–1639
6. Park SSH, Song JJ, Lee JJH, Lee W, Ree S (2015) How to measure similarity for multiple categorical data sets? Multimedia Tools Appl 74(10):3489–3505
7. Wu Z, Palmer M (1994) Verbs semantics and lexical selection. In: Proceedings of the 32nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, Las Cruces, New Mexico pp 133–138

Automated Development of the Markovian Chains to Assess the Availability and Performance of Multi-state Multiprocessor System

Bogdan Volochiy, Oleksandr Mulyak and Vyacheslav Kharchenko

Abstract Reliability design, availability and performance assessment of multi-state multiprocessor system with structural redundancy involves solving number of issues. This paper outlines a cutting-age technology of the analytical modelling of the discrete-continuous stochastic systems for automated development the Markovian chains to assess the availability and Performance of multi-state multiprocessor system, which shows the algorithm for reliability behaviour. For various configurations of the multi-state multiprocessor system, the use of the proposed model and problem-oriented software, ASNA represents the ability to automate constructed the Markovian chains after developed the structural-automated model. This model includes a number of settings: numbers of processor in the main sub-system; numbers of processor in the diverse sub-system; number of processor in hot standby; number of processor in cold standby; failure rate of the processor; mean time of sub-system repair; the structure of the system's and reliability behaviours. The proposed structural-automated model for the automated development the Markovian chains are subject to structure adaptation of the multi-state multiprocessor system and/or the algorithms of reliability behaviour. This allows us to obtain a new model and the feasibility to automate development of the Markovian chains.

Keywords Discrete-continuous stochastic system • Reliability behaviour • Structural-automated model • Markovian chains • Multi-state system

B. Volochiy · O. Mulyak (✉)
National University Lviv Polytechnic, Lviv, Ukraine
e-mail: mulyak.oleksandr@gmail.com

V. Kharchenko
National Aerospace University "KhAI", Kharkiv, Ukraine

A. Lisnianski et al. (eds.), *Recent Advances in Multi-state Systems Reliability*,
Springer Series in Reliability Engineering, DOI 10.1007/978-3-319-63423-4_19

1 Introduction

1.1 Motivation and Approach

Nowadays the developments of fault-tolerant computer-based systems are a part of weaponry components, space, aviation, energy and other critical systems. One of the main tasks is to provide requirements of reliability, availability, functional safety and performance. Thus the two types of possible risks related to the assessment of risk, and to ensuring their safety and security.

Reliability (dependability) related design (RRD) [1–6] is a main part of development of complex fault-tolerant systems based on computers. The goal of RRD is to develop the structure of fault-tolerant systems based tolerating physical failure and designs faults and assure required values of reliability, availability and other dependability attributes. The fault-tolerant systems based on computers can be in numbers of performances levels, what is the one of tasks for structure design, strategy of maintenances plan development and availability and performance assessment.

Insufficient level of adequacy of the availability models of complex multi-state fault-tolerant systems based on computers leads either to additional costs (while underestimating of the indexes), or to the risk of total failure (when inflating their values), namely accidents, material damage and even loss of life. Reliability and safety are assured by using (selection and development) fault-tolerant structures at RRD of the complex multi-state fault-tolerant systems based on computers, and identifying and implementing strategies for maintenance. Adoption of wrong decisions at this stage leads to similar risks.

Hence, we suggest systematic and formalized approach and tool-based technique of developing Markovian chains for complex fault-tolerant systems considering MSS issue.

1.2 Related Works Analysis

Most models are primarily developed to identify the impact of one the above-listed factors on reliability indexes. The rest of the factors are overlooked. Papers [1, 2] describe the reliability model of complex fault-tolerant systems based on computers which illustrates separate HW and SW failures. Paper [3] offer reliability model of a fault-tolerant system, in which HW and SW failures are differentiated and after corrections in the program code the software failure rate is accounted for. Paper [7] describes the reliability model of the complex fault-tolerant systems based on computers, which accounts for the software updates. In paper [8] the author outlines the relevance of the estimation of the reliability indexes of complex fault-tolerant systems based on computers considering the failure of SW and recommends a method for their determination. Such reliability models of the complex fault-tolerant systems based on computers produce an analysis of its conditions under the failure of SW. This research suggests that $MTTF_{system} = MTTF_{software}$. Thus, it is possible to conclude that the author considers the HW of the complex fault-tolerant systems based on computers as absolutely reliable. Such condition reduces the credibility of the result, especially when the reliability of the HW is commensurable to the reliability of the SW. Paper [9] presents the assessment of reliability parameters of complex fault-tolerant systems based on computers through modelling behaviour using Markovian chains, which account for multiple software updates. Nevertheless, there was no evidence of the quantitative assessments of the reliability measures of presented FTCS.

In paper [10], the authors propose a model of complex fault-tolerant systems based on computers using Macro-Markovian chains, where the software failure rate, duration of software verification, failure rate and repair rate of HW are accounted for. The presented method of Macro-Markovian chains modelling [9, 10] is based on logical analysis and cannot be used for profound configurations of complex fault-tolerant systems based on computers due to their complexity and high probability of the occurrence of mistakes. In addition, there is a discussion around the definition of requirements for operational verification of software of the space system, together with the research model of the object for availability evaluation and scenarios preference. It is noted that over the last ten years out of 27% of space devices failures, which were fatal or such that restricted their use, 6% were associated with HW failure and 21% with SW failure.

Mathematical methods of multi-state system availability and performance assessment are discussed in the number of research papers [8, 11–14]. The growth of MSS theories is based on distinct structural functions, Markov models and combined methods [8, 13]. The obtained results in this research area for multi-level MSS with majority voting [15]. The Markov analysis in combination with multi-level degradation diagrams was used for MSS with maintenances [11].

The primary aim of this chapter is to provide a unique technology to develop the Markovian chain for complex MSS based on computers with different structural redundancy, using the proposed formal procedure and tool. This model is proven to be flexible in configuration of the researched object, accounts for its specific features and prepares the availability and performance assessment. The main idea of technology is decreasing the risks of errors during classical methods of Markovian chain development for systems with a large (hundreds and thousands) number of stages. We propose a special notation, which allows supporting development of the Markovian chain step by step and designing the final Markovian chain using software tool.

2 Description of the Multi-state Multiprocessor System

In this Chapter, we present the comprehensive description of the multi-state multiprocessor radio-electronic data receiving computer-based system “RP-140” [4]. Generalised structure of such Multi-State multiprocessor System (MSS) with structure redundancy is shown in Fig. 1. To ensure the minimal Multi-State System (MSS) downtime, overall hot/cold standby and maintenances of the system is used.

The MSS includes: a main sub-system which consist of n—processors; diverse sub-system consist of k—processors; for two sub-systems, the common sliding standby of processors are available, the first (or m_h) processor is in hot standby and other processors m_c are in cold standby; a diagnostics control system determines the state of processors and manages the redundancy; and a switch is connected the processors to the main and diverse sub-systems from redundancy.

Based on the algorithm of MSS operation after ($\lceil n/2 \rceil + 1$) failures of processors in main sub-system or ($\lceil k/2 \rceil + 1$) failures of processors in diverse sub-system the reconfiguration is performed. If ($\lceil n/2 \rceil - 1$) or ($\lceil k/2 \rceil - 1$) processors are available in main or diverse sub-system all non-failed processors are automatically connected o the operating system. Also after ($\lceil n/2 \rceil + 1$) failures in main sub-system or ($\lceil k/2 \rceil + 1$) failures in diverse sub-system the repair procedures are performed. The system’s failure occurs when number of working processors is lower than ($\lceil n + k \rceil/2$).

The MSS with structural redundancy an stay in numbers level with different performance in particular we can divide for three levels: working when numbers of processors are equal $\lceil n + k \rceil$; partly working when numbers of processors are lower or equal ($\lceil n + k \rceil - 1$); fault when numbers of processors are lower (($\lceil n + k \rceil/2) - 1$). But for complex system this breakup are not enough, because system can stay in many partly working state with different performance coefficient and could

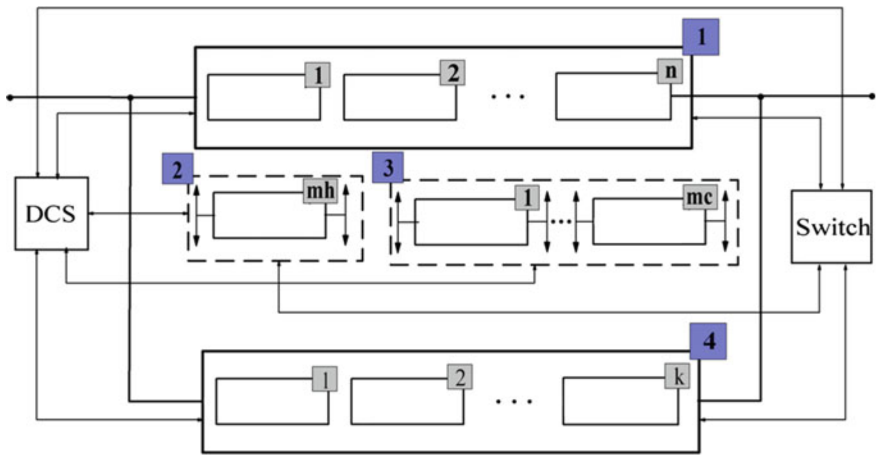


Fig. 1 Multi-state multiprocessor system (*1* main sub-system, *2* hot standby processors, *3* cold standby processors, *4* diverse sub-system, *DCS* diagnostics control system)

Table 1 Multi-State system configuration and performance level

n	k	$K_{ofPerformance}$ *of MSS, %*
6	6	100
5	6	92
6	5	92
5	5	83
4	5	75
5	4	75
4	4	67
3	4	58
4	3	58
3	3	50

be divided according to number of processors in researched MSS with coefficient of performance calculation (1). In Table 1 one of possible configuration of MSS ($n = 6$, $k = 6$, $m_h = 0$, $m_c = 0$) and performance level are shown.

$$K_{ofPerformance} = \frac{(n + k) - M_{failure}}{n + k} \cdot 100\%, \quad (1)$$

where $M_{failure}$—number of failure processors in main and diverse sub-system.

3 Technique for Automated Development of the Markov Model of the Multi-state Multiprocessor System

The detailed description of the technology of analytical modelling of the discrete-continuous stochastic systems is described in the monograph [6]. It involves a formalized representation of the study object as a “structural-automated model”. To develop structural-automated model for MSS one needs to complete the following tasks:

(1) develop a verbal description of the research object (Fig. 1);
(2) define the basic events; define the components of vector states, which can be described as a state of random time;
(3) define the parameters for the object of research, which should be in the model;
(4) shape the tree of the modification of the rules and component of the states vector.

3.1 Procedures for Behaviour Description of the Multi-state Multiprocessor System

The *MSS* behaviour is described by the following procedures: ***Procedure 1***. Failure detection in the multiprocessor system. Failure can occur in the main sub-system (MS) and diverse sub-system (DS); ***Procedure 2***. Detection of processor failure in the hot standby; ***Procedure 3***. Detection of diagnostics control system failure; ***Procedure 4***. Detection of switch failure; ***Procedure 5***. Connection of the processor from hot standby to faulty sub-systems; ***Procedure 6***. Connection of the processor from hot standby to cold standby; ***Procedure 7***. System reconfiguration is performed when number of failure processor $M_{failure} > \lceil n/2 \rceil$ for main sub-system; $M_{failure} > \lceil k/2 \rceil$ for diverse sub-system; ***Procedure 8***. Repair of the processor in main sub-system; ***Procedure 9***. Repair of the processor in diverse sub-system.

3.2 Basic Events

According to described procedures, which determine the behaviour of multi-state multiprocessor system, a list of events is composed. Events are present the end of time intervals that describe the procedure completion.

As a result of analysis, eight basic events in particular were determined: ***Event 1***—Processor failure in the main sub-system; ***Event 2***—Processor failure in the diverse sub-system; ***Event 3***—Failure of the processor in hot standby; ***Event 4***—Completing the processor switching procedure from hot standby to subsystems with failure processor; ***Event 5***—Completing the processor switching procedure from cold standby to sub-system with failure processor; ***Event 6***—Completing of the system reconfiguration procedures; ***Event 7***—Completing the maintenances procedure of the main sub-system; ***Event 8***—Completing the maintenances procedure of the diverse sub-system.

3.3 Components of Vector States

Components of the vector state that can also be described as a state of random time. To describe the state of the system, five components are used:

V1—displays the operational processors in the MS (the initial value of components V1 equal to n);
V2—displays the operational processors in the DS (the initial value of components V2 equal to k);
V3—displays the operational processors in hot standby (the initial value of V3 equal to m_h);

V4—displays the current number of processors in cold standby (the initial value of components V4 equal to m_c);
V5—displays the number of operational processors in system (the value of component V5 calculated as $V5 = (n + k) - M_{failure}$).

3.4 Parameters of the Markov Model

Developing Markov model of the FTCS, its composition and separate components should be set to relevant parameters in particular: **n**—number of processors that in the main sub-system; **k**—number of processors that in the diverse sub-system; $\mathbf{m_h}$—number of processors in the hot standby; $\mathbf{m_c}$—number of processors in the cold standby; $\boldsymbol{\lambda_{ms}}$—the failure rate of the processor in main sub-system; $\boldsymbol{\lambda_{ds}}$—the failure rate of the processor in diverse sub-system; $\boldsymbol{\lambda_{hs}}$—the failure rate of the processor in hot standby; $\mathbf{T_{rec}}$—mean time of the system reconfigurations; $\mathbf{T_{switch}}$—mean time of the processor connections from standby; $\mathbf{T_{rep}}$—mean time of hardware repair.

3.5 Structural-Automated Model for Multi-state Multiprocessor System

According to the technology of a modelling the discrete-continuous stochastic systems [6] based on certain events using the component vector state and the parameters that describe MSS, and Structural-Automated model of the MSS for automated development of the Markovian chains are presented on the Table 1.

Below is describes the procedures of structural-automated model creation:

Components of Event 1
Event 1—"Processor failure in the main sub-system" can occur when in MS are operating processor. Therefore, a failure of the processor in the MS may occur under these two conditions:

First circumstance for the Event 1—the number of operation processor is more or equal ⌈n/2⌉ (V1 >= ⌈n/2⌉); the hot standby processors are available (V3 > 0). The logical expression for first circumstance is:of the processor in the DS

$$(\mathbf{V1} >= \lceil \mathbf{n/2} \rceil)\ \mathbf{AND}\ (\mathbf{V3} > \mathbf{0}).$$

In the first circumstance the faulty processor is removed from the MS and the processor from hot standby switched to MS (*this procedures is combined with Event 4*). Failure rate of Event 1 in this circumstance is defined as: failure rate of processor in MS λ_{ms} and the number of processor in MS V1. The formula of calculating the failure rate of transition (FFRT) has the following form: $\mathbf{V1} \cdot \boldsymbol{\lambda_{ms}}$.

The consequences of the Event 1 under the first circumstance:

The processor is connected from the hot standby to the MS. As a result a number of processors in MS are not changed V1 := n but from the hot standby one processor is subtracted V3 := V3 − 1. Thus, the rule of the modification component of the state vector (MCSV) for the first circumstance is submitted as follows:

$$\mathbf{V1} := \mathbf{n};\ \mathbf{V3} := \mathbf{V3} - \mathbf{1}.$$

Second circumstance for the Event 1—the number of operation processor is more or equal ⌈n/2⌉ (V1 >= ⌈n/2⌉); the hot standby processors are not available (V3 = 0). The logical expression for second circumstance is:

$$(\mathbf{V1} >= \lceil \mathbf{n/2} \rceil)\ \mathbf{AND}\ (\mathbf{V3} = \mathbf{0}).$$

In the second circumstance the faulty processor is removed from the MS and Performance of system decreases. Failure rate of Event 1 in this circumstance is defined as: failure rate of processor in MS λ_{ms} and the number of processor in MS V1. The formula of calculating the failure rate of transition (FFRT) has the following form: $\mathbf{V1} \cdot \boldsymbol{\lambda}_{\mathbf{ms}}$.

The consequences of the Event 1 under the second circumstance:

The processor is removed from the MS. As a result a number of processors in MS are changed V1 := V1 − 1, and Performance of system is decreased V5 : = V5 − 1. Thus, the rule of the modification component of the state vector (MCSV) for the first circumstance is submitted as follows:

$$\mathbf{V1} := \mathbf{V1} - \mathbf{1};\ \mathbf{V5} := \mathbf{V5} - \mathbf{1}.$$

Components of Event 2

Event 2—"Processor failure in the diverse sub-system" can occur when in DS are operating processor. Therefore, a failure of the processor in the DS may occur under these two conditions:

Firs circumstance for the Event 2—the number of operation processor is more or equal ⌈k/2⌉ (V2 >= ⌈k/2⌉); the hot standby processors are available (V3 > 0). The logical expression for first circumstance is:

$$(\mathbf{V2} >= \lceil \mathbf{k/2} \rceil)\ \mathbf{AND}\ (\mathbf{V3} > \mathbf{0}).$$

In the first circumstance the faulty processor is removed from the DS and the processor from hot standby switched to DS (*this procedures is combined with Event 4*). Failure rate of Event 2 in this circumstance is defined as: failure rate of processor in DS λ_{ds} and the number of processor V2 in DS. The formula of calculating the failure rate of transition (FFRT) has the following form: $\mathbf{V2} \cdot \boldsymbol{\lambda}_{\mathbf{ds}}$.

The consequences of the Event 2 under the first circumstance:

The processor is connected from the hot standby to the DS. As a result a number of processors in DS are not changed V2 := k but from the hot standby one processor is subtracted V3 := V3 − 1. Thus, the rule of the modification component of the state vector (MCSV) for the first circumstance is submitted as follows:

$$\mathbf{V2}\!:=\mathbf{k};\ \mathbf{V3}\!:=\mathbf{V3}-\mathbf{1}.$$

Second circumstance for the Event 2—the number of operation processor is more or equal ⌈k/2⌉ (V2 >= ⌈k/2⌉); the hot standby processors are not available (V3 = 0). The logical expression for first circumstance is:

$$(\mathbf{V2}>=\lceil\mathbf{k/2}\rceil)\ \mathbf{AND}\ (\mathbf{V3}=\mathbf{0}).$$

In the second circumstance the faulty processor is removed from the DS and Performance of system decreases. Failure rate of Event 2 in this circumstance is defined as: failure rate of processor in DS λ_{ds} and the number of processor V2 in DS. The formula of calculating the failure rate of transition (FFRT) has the following form: $\mathbf{V2}\cdot\boldsymbol{\lambda}_{\mathbf{ds}}$.

The consequences of the Event 2 under the second circumstance:

The processor is removed from the DS. As a result a number of processors in DS are changed V2 := V2 − 1, and Performance of system is decreased V5 : = V5 − 1. Thus, the rule of the modification component of the state vector (MCSV) for the second circumstance is submitted as follows:

$$\mathbf{V2}\!:=\mathbf{V2}-\mathbf{1};\ \mathbf{V5}\!:=\mathbf{V5}-\mathbf{1}.$$

Components of Event 3

Event 3—"Failure of the processor in hot standby" can occur when processors in hot standby is operational (V3 > 0) and the main or diverse sub-system are operated (V1 >= ⌈n/2⌉ OR V2 >= ⌈k/2⌉). The logical expression for this circumstance is:

$$(\mathbf{V3}>\mathbf{0})\ \mathbf{AND}\ (\mathbf{V1}>=\lceil\mathbf{n/2}\rceil\ \mathbf{OR}\ \mathbf{V2}>=\lceil\mathbf{k/2}\rceil).$$

In this circumstance the no operational processor is removed from the hot standby. Failure rate of Event 3 in this circumstance is defined as: failure rate of processors λ_{hs} and the number of processor V3. The formula of calculating the FFRT has the following form:

$$\mathbf{V3}\cdot\boldsymbol{\lambda}_{\mathbf{hs}}.$$

The consequences of the Event 3 under this circumstance:

The processor from hot standby is subtracted V3: = V3 − 1. Thus, the rule of the MCSV for the first circumstance is submitted as follows:

$$\mathbf{V3}\!:=\mathbf{V3}-\mathbf{1}.$$

Components of Event 4

Duration the switching processor procedures from hot standby to main/diverse sub-system is very low. Taking into account this circumstance the Event 4 was combined with Event 1 and Event 2.

Components of Event 5

Event 5—"Completing of the processor switching procedure from cold standby to sub-system with failure processor" can occur when operational processors are in cold standby. Therefore, switching procedure from cold standby to sub-system may occur under these two conditions:

Firs circumstance for the Event 5—the cold standby has operational processor ($V4 > 0$), hot standby has no operational processors ($V3 = 0$) and main sub-system has failures processor (V1 <=(n − 1)). The logical expression for first circumstance is:

$$(\mathbf{V1} <= (\mathbf{n} - \mathbf{1}))\ \mathbf{AND}\ (\mathbf{V3} = \mathbf{0})\ \mathbf{AND}\ (\mathbf{V4} > \mathbf{0}).$$

In this circumstance the processor from cold standby is switching to main sub-system. The rate of Event 5 is defined as: mean time of processors switching $1/T_{switch}$. The formula of calculating the FFRT has the following form: $\mathbf{1/T_{switch}}$.

The consequences of the Event 5 under this circumstance:

The processor is switching from cold standby to main sub-system V1: = V1 + 1. The processor from cold standby is subtracted V4: = V4 − 1 and Performance of system is increased V5: = V5 + 1. Thus, the rule of the MCSV for the first circumstance is submitted as follows:

$$\mathbf{V1} := \mathbf{V1} + \mathbf{1};\ \mathbf{V4} := \mathbf{V4} - \mathbf{1};\ \mathbf{V5} := \mathbf{V5} + \mathbf{1}.$$

Second circumstance for the Event 5—the cold standby has operational processor ($V4 > 0$), hot standby has no operational processors ($V3 = 0$) and diverse sub-system has failures processor (V2 <=(k-1)). The logical expression for first circumstance is:

$$(\mathbf{V2} <= (\mathbf{k} - \mathbf{1}))\ \mathbf{AND}\ (\mathbf{V3} = \mathbf{0})\ \mathbf{AND}\ (\mathbf{V4} > \mathbf{0}).$$

In this circumstance the processor from cold standby is switching to diverse sub-system. The rate of Event 5 is defined as: mean time of processors switching $1/T_{switch}$. The formula of calculating the FFRT has the following form: $\mathbf{1/T_{switch}}$.

The consequences of the Event 5 under this circumstance:

The processor is switching from cold standby to diverse sub-system V2: = V2 + 1. The processor from cold standby is subtracted V4: = V4 − 1 and Performance of system is increased V5: = V5 + 1. Thus, the rule of the MCSV for the first circumstance is submitted as follows:

$$\mathbf{V2}:=\mathbf{V2}+\mathbf{1};\ \mathbf{V4}:=\mathbf{V4}-\mathbf{1};\ \mathbf{V5}:=\mathbf{V5}+\mathbf{1}.$$

Components of Event 6

Event 6—"Completing of the system reconfiguration procedures" can occur when main or diverse sub-systems are in no operational conditions. Therefore, reconfiguration procedure may occur under these two conditions:

First circumstance for the Event 5—the diverse sub-system is in no operational conditions (V2 < ⌈k/2⌉) and number of operational processors in main sub-system is equal ⌈n/2⌉ (V1 = ⌈n/2⌉) The logical expression for first circumstance is:

$$(\mathbf{V1}=\lceil \mathbf{n}/\mathbf{2}\rceil)\ \mathbf{AND}\ (\mathbf{V2}<\lceil \mathbf{k}/\mathbf{2}\rceil).$$

In this circumstance the all operational processors switching to main sub-system. The rate of Event 6 is defined as: mean time of reconfigurations $1/T_{rec}$. The formula of calculating the FFRT has the following form:

$$\mathbf{1}/\mathbf{T}_{\mathbf{rec}}.$$

The consequences of the Event 6 under this circumstance:

The processor is switching from diverse sub-system to main sub-system V1: = V1 + V2. The processor from diverse sub-system is subtracted V2: = V2 – V2 and Performance of system is increased V5: = V5 + V2. Thus, the rule of the MCSV for the first circumstance is submitted as follows:

$$\mathbf{V1}:=\mathbf{V1}+\mathbf{V2};\ \mathbf{V5}:=\mathbf{V5}+\mathbf{V2};\ \mathbf{V2}:=\mathbf{V2}-\mathbf{V2}.$$

Second circumstance for the Event 6—the main sub-system is in no operational conditions (V1 < ⌈n/2⌉) and number of operational processors in diverse sub-system are equal ⌈k/2⌉ (V2 = ⌈k/2⌉). The logical expression for first circumstance is:

$$(\mathbf{V1}<\lceil \mathbf{n}/\mathbf{2}\rceil)\ \mathbf{AND}\ (\mathbf{V2}=\lceil \mathbf{k}/\mathbf{2}\rceil).$$

In this circumstance the all operational processors switching to diverse sub-system. The rate of Event 6 is defined as: mean time of reconfigurations $1/T_{rec}$. The formula of calculating the FFRT has the following form: $\mathbf{1/T_{rec}}$.

The consequences of the Event 6 under this circumstance:

The processor is switching from main sub-system to diverse sub-system V2: = V2 + V1. The processor from main sub-system is subtracted V1: = V1 – V1 and Performance of system is increased V5: = V5 + V1. Thus, the rule of the MCSV for the first circumstance is submitted as follows:

$$\mathbf{V2}:=\mathbf{V2}+\mathbf{V1};\ \mathbf{V5}:=\mathbf{V5}+\mathbf{V1};\ \mathbf{V1}:=\mathbf{V1}-\mathbf{V1}.$$

Components of Event 7

Event 7—"Completing the maintenances procedure of the main sub-system" can occur when main sub-system has non operational processor (V1 < ⌈n/2⌉) and number of processors in hot/cold standby are equal 0 (V3 = 0) AND (V4 = 0) The logical expression for first circumstance is:

$$(\mathbf{V1} < \lceil \mathbf{n/2} \rceil)\ \mathbf{AND}\ (\mathbf{V3} = \mathbf{0})\ \mathbf{AND}\ (\mathbf{V4} = \mathbf{0}).$$

In this circumstance the new processors connect to main sub-system. The rate of Event 7 is defined as: mean time of repair $1/T_{rep}$. The formula of calculating the FFRT has the following form: $\mathbf{1/T_{rep}}$.

The consequences of the Event 7 under this circumstance:

The new processors connect to main sub-system V1: = n and the Performance of system, numbers of processors in hot/cold standby are returns to initial condition: V3: = m_h, V4: = m_c, V5: = n + k. Thus, the rule of the MCSV for this circumstance is submitted as follows:

$$\mathbf{V1}: = \mathbf{n};\ \mathbf{V3}: = \mathbf{m_h};\ \mathbf{V4}: = \mathbf{m_c};\ \mathbf{V5}: = \mathbf{n} + \mathbf{k}.$$

Components of Event 8

Event 8—"Completing the maintenances procedure of the diverse sub-system" can occur when diverse sub-system has non operational processor (V2 < ⌈k/2⌉) and numbers of processors in hot/cold standby are equal 0 (V3 = 0) AND (V4 = 0). The logical expression for first circumstance is:

$$(\mathbf{V2} < \lceil \mathbf{k/2} \rceil)\ \mathbf{AND}\ (\mathbf{V3} = \mathbf{0})\ \mathbf{AND}\ (\mathbf{V4} = \mathbf{0}).$$

In this circumstance the new processors connect to diverse sub-system. The rate of Event 8 is defined as: mean time of repair $1/T_{rep}$. The formula of calculating the FFRT has the following form: $\mathbf{1/T_{rep}}$.

The consequences of the Event 8 under this circumstance:

The new processors connect to diverse sub-system V2: = k and the Performance of system, numbers of processors in hot/cold standby are returns to initial condition: V3: = m_h, V4: = m_c, V5: = n + k. Thus, the rule of the MCSV for this circumstance is submitted as follows:

$$\mathbf{V1}: = \mathbf{n};\ \mathbf{V3}: = \mathbf{m_h};\ \mathbf{V4}: = \mathbf{m_c};\ \mathbf{V5}: = \mathbf{n} + \mathbf{k}.$$

4 Availability and Performance Analysis. Markov Model Development

The Structural-Automated Model for MSS provides the possibilities according to technology [6] for automated developed of the Markovian chains by tool ASNA [16]. Output data from ASNA tool is the transition rate matrix. In classical technology the above matrix is composed following the construction of the Markovian chain. In this technology the Markovian chain visualisation is based on the transition rate matrix. The system of differential equations are solved by Runge–Kutta methods in MathCad by using the standard State Space ODE Solver.

In the Table 3 are shown the parameters of Markovian chains with different configuration of researched MSS that was developed by ASNA based on prepared Structural-Automated Model (Table 2).

Table 2 Structural-automated model of the multi-state multiprocessor system

Terms and conditions of event	Formula for calculating the rate of events	Rule of modification the state vectors component
Event 1. Processor failure in the main sub-system		
(V1 >= ⌈n/2⌉) AND (V3 > 0)	$V1 \cdot \lambda_{ms}$	V3 := V3 − 1
(V1 >= ⌈n/2⌉) AND (V3 = 0)	$V1 \cdot \lambda_{ms}$	V1 := V1 − 1; V5 := V5 − 1
Event 2. Processor failure in the diverse sub-system		
(V2 >= ⌈k/2⌉) AND (V3 > 0)	$V1 \cdot \lambda_{ds}$	V3 := V3 − 1
(V2 >= ⌈k/2⌉) AND (V3 = 0)	$V1 \cdot \lambda_{ds}$	V2 := V2 − 1; V5 := V5 − 1
Event 3. Failure of the processor in hot standby		
(V3 > 0) AND ((V1 >= ⌈n/2⌉ OR V2 >= ⌈k/2⌉)	$V1 \cdot \lambda_{hs}$	V3 := V3 − 1
Event 4. Completing of the processor switching procedure from cold standby to sub-system with failure processor		
(V1 <=(n-1)) AND (V3 = 0) AND (V4 > 0)	$1/T_{switch}$	V1: = V1 + 1; V4: = V4 − 1; V5: = V5 + 1
(V2 <=(k-1)) AND (V3 = 0) AND (V4 > 0)	$1/T_{switch}$	V2: = V2 + 1; V4: = V4 − 1; V5: = V5 + 1
Event 5. Completing of the system reconfiguration procedures		
(V1 = ⌈n/2⌉) AND (V2 < ⌈k/2⌉)	$1/T_{rec}$	V1: = V1 + V2; V5: = V5 + V2; V2: = V2 − V2
(V2 = ⌈k/2⌉) AND (V1 < ⌈n/2⌉)	$1/T_{rec}$	V2: = V1 + V2; V5: = V5 + V1; V1: = V1 − V1
Event 6. Completing the maintenances procedure of the main sub-system		
(V1 < ⌈n/2⌉) AND (V3 = 0) AND (V4 = 0)	$1/T_{rep}$	V1: = n; V3: = mh; V4: = mc; V5: = n + k
Event 7. Completing the maintenances procedure of the diverse sub-system		
(V2 < ⌈k/2⌉) AND (V3 = 0) AND (V4 = 0)	$1/T_{rep}$	V2: = k; V3: = mh; V4: = mc; V5: = n + k

The Markovian chain for model N 7 (Table 3) is presented in Fig. 2. Information is available on the status of each software processor ASNA we have on file "vector. vs", which is written in the form:

Stage 1: V1 = 6; V2 = 6; V3 = 1; V4 = 1; V5 = 12
Stage 2: V1 = 6; V2 = 6; V3 = 0; V4 = 1; V5 = 12
Stage 3: V1 = 5; V2 = 6; V3 = 0; V4 = 1; V5 = 11
Stage 4: V1 = 4; V2 = 6; V3 = 0; V4 = 1; V5 = 10
Stage 5: V1 = 3; V2 = 6; V3 = 0; V4 = 1; V5 = 9
Stage 6: V1 = 2; V2 = 6; V3 = 0; V4 = 1; V5 = 8
……………………………………………………………..
Stage 76: V1 = 0; V2 = 3; V3 = 0; V4 = 0; V5 = 3
Stage 77: V1 = 0; V2 = 2; V3 = 0; V4 = 0; V5 = 2

The proposed model of MSS can be easily transformed for other features of the object of study. It is enough to: add/remove basic event; attach/remove components of the state vector; and include/remove parameters that describe the studied system. Based on information about the work of MSS an appropriate change in the model could be made (Fig. 1).

Based on the Markovian chains (Fig. 2) a system of differential equations (2) was formed. Its solution allows us to estimate the function availability value of researched MSS.

$$\left.\begin{array}{l}
\frac{dP_1(t)}{dt} = \frac{1}{T_{rep}}\left(P_{36}(t) + P_{40}(t) + P_{44}(t) + \sum_{48}^{56} P_i(t) + \sum_{61}^{67} P_i(t) + \sum_{71}^{77} P_i(t)\right) \\
\qquad - 12 \cdot \lambda_{hw} \cdot P_1(t) \\
\frac{dP_2(t)}{dt} = 12 \cdot \lambda_{hw} \cdot P_1(t) - 6 \cdot \lambda_{hw} P_2(t) - 6 \cdot \lambda_{hw} P_2(t) \\
\frac{dP_3(t)}{dt} = 6 \cdot \lambda_{hw} P_2(t) - 6 \cdot \lambda_{hw} P_3(t) - \frac{1}{T_{switch}} \cdot P_3(t) - 5 \cdot \lambda_{hw} P_3(t) \\
\frac{dP_4(t)}{dt} = 5 \cdot \lambda_{hw} P_3(t) - 6 \cdot \lambda_{hw} P_4(t) - \frac{1}{T_{switch}} \cdot P_4(t) - 4 \cdot \lambda_{hw} P_4(t) \\
\vdots \\
\frac{dP_{77}(t)}{dt} = 3 \cdot \lambda_{hw} P_{76}(t) - \frac{1}{T_{rep}} \cdot P_{77}(t)
\end{array}\right\} \quad (2)$$

Table 3 Parameters of Markovian chains with different configuration of researched MSS

N of model	Parameters of researched MSS				Parameters of Markovian chains	
	n	k	m_h	m_c	Number of stages	Number of transitions
1	4	4	1	1	47	114
2	4	4	1	2	71	184
3	4	4	1	3	95	254
4	6	6	1	1	77	190
5	6	6	1	2	122	326
6	6	6	1	3	167	462
7	12	10	1	5	609	1810

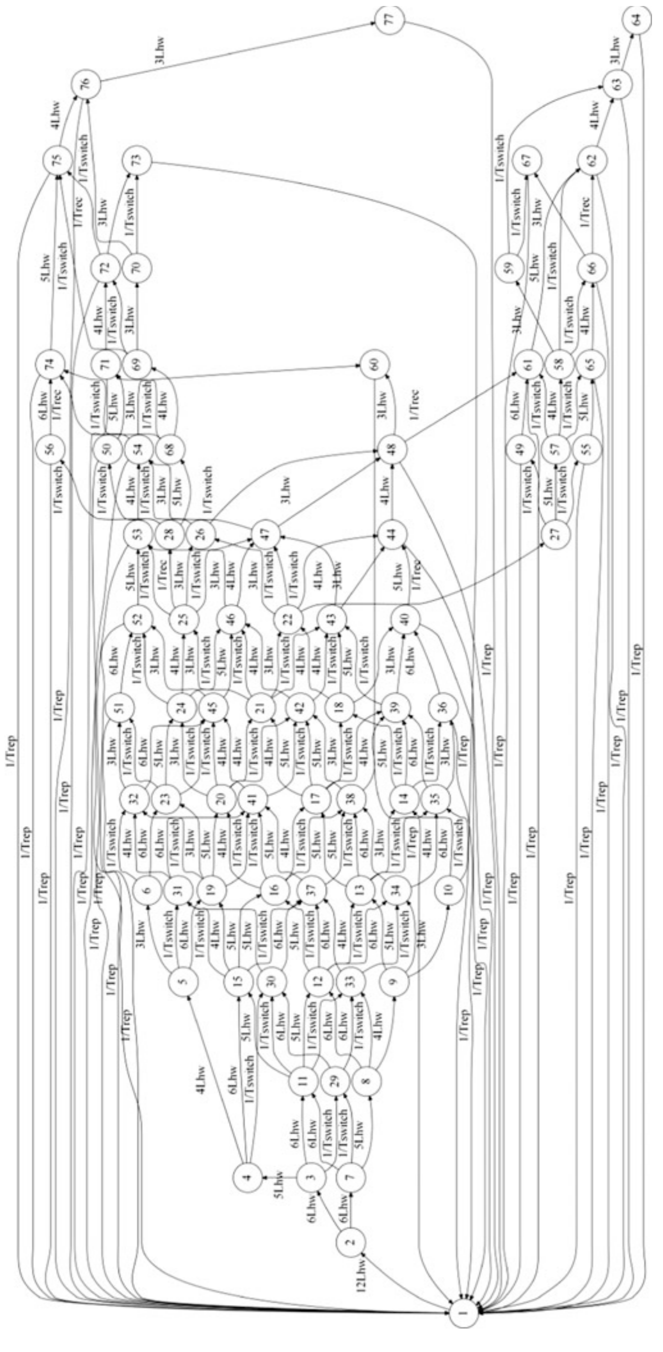


Fig. 2 The Markovian chains of the researched MSS ($n = 6$, $k = 6$, $m_h = 1$, $m_c = 1$)

Initial conditions for the system (2) for initial time t = 0, are $P_1(t) = 1$; $P_2(t) \ldots P_{77}(t) = 0$.

4.1 Simulation Results

Basing on the Markovian chains (Fig. 2) formulas for availability assessment of MSS with different performances level are presented. Availability function of MSS with different performances level is calculated as the sum of the probability functions staying in operable stages with defined the vector V5.

For 100% system performance the availability function is us follow:

$$A_{100\%}(t) = P_1(t) + P_2(t) + P_{29}(t)$$

For 92% system performance the availability function is us follow:

$$A_{92\%}(t) = \sum_{i=1}^{3} P_i(t) + P_7(t) + P_{29}(t) + P_{30}(t) + P_{33}(t)$$

For 83% system performance the availability function is us follow:

$$A_{83\%}(t) = \sum_{i=1}^{4} P_i(t) + \sum_{i=7}^{8} P_i(t) + P_{11}(t) + \sum_{i=29}^{31} P_i(t) + \sum_{i=33}^{34} P_i(t) + P_{37}(t)$$

For 75% system performance the availability function is us follow:

$$A_{75\%}(t) = \sum_{i=1}^{5} P_i(t) + \sum_{i=7}^{9} P_i(t) + \sum_{i=11}^{12} P_i(t) + P_{15}(t) + \sum_{i=29}^{35} P_i(t) + \sum_{i=37}^{38} P_i(t) + P_{41}(t)$$

For 67% system performance the availability function is us follow:

$$A_{67\%}(t) = \sum_{i=1}^{6} P_i(t) + \sum_{i=7}^{10} P_i(t) + \sum_{i=11}^{13} P_i(t) + \sum_{i=15}^{16} P_i(t) + \sum_{i=29}^{39} P_i(t) + \sum_{i=41}^{42} P_i(t) + P_{45}(t) + P_{51}(t)$$

For 58% system performance the availability function is us follow:

$$A_{58\%}(t) = \sum_{1}^{6} P_i(t) + \sum_{7}^{10} P_i(t) + \sum_{11}^{14} P_i(t) + \sum_{15}^{17} P_i(t) + P_{20}(t) + P_{23}(t) + \sum_{29}^{43} P_i(t) + \sum_{45}^{46} P_i(t) + \sum_{51}^{52} P_i(t)$$

For 50% system performance the availability function is us follow:

$$A_{50\%}(t) = \sum_{1}^{6} P_i(t) + \sum_{7}^{10} P_i(t) + \sum_{11}^{14} P_i(t) + \sum_{15}^{18} P_i(t) + \sum_{20}^{21} P_i(t) + \sum_{23}^{24} P_i(t) + \sum_{29}^{47} P_i(t) + \sum_{49}^{53} P_i(t) + \sum_{55}^{56} P_i(t)$$

Availability calculating was performed using the failure rate: $\lambda_{ms} = \lambda_{ds} = _{hs} = 1 \times 10\ h^{-1}$, mean time of switching processors $T_{rec} = 6$ min, mean time of system reconfiguration $T_{switch} = 6$ min and mean repair time $T_{rep} = 24$ h.

Availability functions of Multi-State Multiprocessor System for five performance levels (100, 92, 83, 75 and 67%) are present on Fig. 3, a availability functions for 58% performance level are presented on Fig. 4 and for 50% performance level presented on Fig. 4.

As shown in Fig. 3 the availability function decreases up to 5000 operating hours, following the availability function increase, due to the system reconfiguration procedures. After 15,000 h the availability function stabilises by repair support. The same paten shown in Figs. 3 and 4 of availability function for different performance levels (Fig. 5).

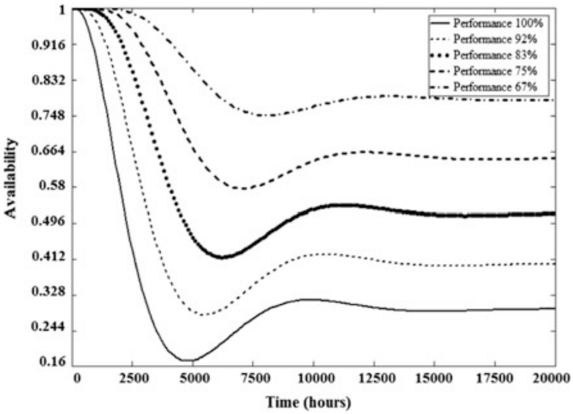


Fig. 3 System's availability for different performance level (100, 92, 83, 75, 67%)

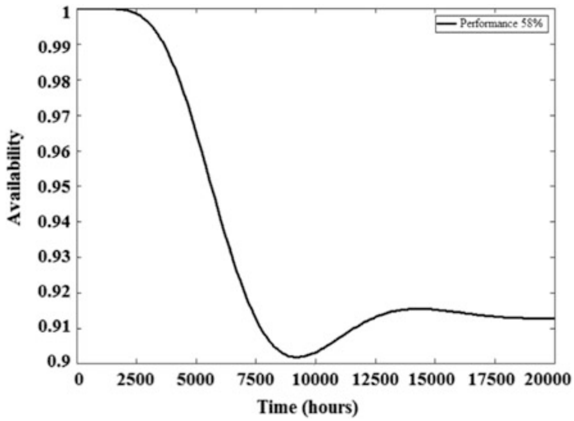


Fig. 4 System's availability for 58% performance level

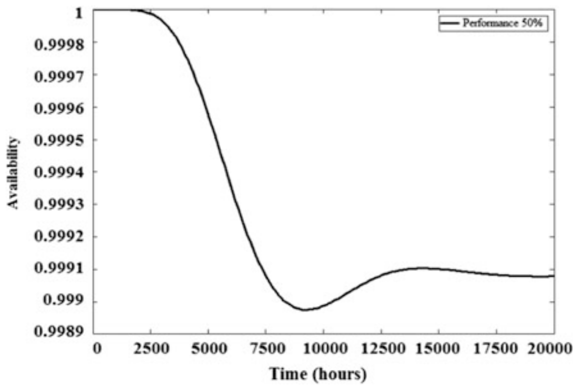


Fig. 5 System's availability for 50% performance level

5 Conclusion

This chapter presents availability and performance assessment of multi-state multiprocessor system with structural redundancy. We illustrate automated development of Markovian chains using a cutting-age technology and tool ASNA.

The presented model can be easily adapted to different configurations of multi-state multiprocessor system such us: k-out-of-n system with structural redundancy and reconfigurations. In fact, this model can be adopted for MSS with software updates and automatic restart.

Future research has the potential to supplement this model with further factors:

- Erlang distribution for durations of repair [16];
- Influence of software reliability on dependability of systems.

References

1. Amerijckx C, Legat J-D (2008) A low-power multiprocessor architecture for embedded reconfigurable systems. In: Legat JD (ed) Power and timing modeling, optimization and simulation, international workshop, 2008
2. Zhu C, Gu Z, Dick R, Shang L (2007) Reliable multiprocessor system-on-chip synthesis. In: Proceedings of the 5th IEEE/ACM international conference on hardware/software codesign and system synthesis, Salzburg, Austria, 30 Sept–03 Oct 2007
3. Korotun T (2007) Models and methods for testing software systems. Progr Probl 2:76–84
4. Mulyak O (2014) Reliability models of computer-based systems with version-structural redundancy. Dissertation, National University Lviv Polytechnic
5. Rajesh S (2008) Fault tolerance in multicore processors with reconfigurable hardware unit. In: Kumar V, Srivatsan R, Harini S, Shanthi A (eds) 15th international conference on high performance computing. Bangalore, India, 2008
6. Volochii B (2004) Technology of modeling the information systems. Lviv Polytechnic Press, Lviv
7. Xiong L, Tan Q, Xu J (2011) Effects of soft error to system reliability. In: Tan Q, Xu J (eds) Workshops of 25th international conference on advanced information networking and applications, Singapore, 22–25 Mar 2011
8. Levitin G (2003) Reliability of multi-state systems: a historical overview. Mathematical and statistical methods in reliability. In: Levitin G, Lisnianski A, Ushakov I (eds) Ser. On quality, reliability and engineering statistics, vol 7, pp 123–138
9. Gostelow K (2011) The design of a fault-tolerant, real time, multi-core computer system. In: In IEEE aerospace conference. doi:10.1109/AERO.2011.5747449
10. Kharchenko V (2016) Resilience assurance for software-based space systems with online patching. In: Kharchenko V, Ponochovnyi Y, Boyarchuk A, Brezhnev E (eds) Eleventh international conference on dependability and complex systems DepCoS-RELCOMEX, 27 June–1 July, vol 470. Springer, Brunów, Poland, pp 267–278
11. Soro I, Nourelfath M, Ait-Kadi MD (2010) Performance evaluation of multi-state degraded systems with minimal repairs and imperfect preventive maintenance. Reliab Eng Syst Saf 95 (2):65–69
12. Levitin G, Lisnianski A, Pham H (2003) Multi-state system reliability analysis and optimization (universal generating function and genetic algorithm approach). In: Handbook of reliability engineering. Springer, Berlin, pp 61–90
13. Lisnianski A (2007) Extended block diagram method for a multi-state system reliability assessment. Reliab Eng Syst Saf 92(12):1601–1607
14. Yingkui G, Jing L (2012) Multi-state system reliability: a new and systematic Review. Procedia Eng J 29:531–536
15. Zaitseva E, Levashenko V (2007) investigation multi-state system reliability by structure functions.In: Proceedings of the DepCoS conference, Poland, 2007
16. Bobalo J, Volochiy B, Lozynskyi O, Mandzii B, Ozirkovskyi L, Fedasuk D, Shcherbovskyh S, Jakovyna V (2013) Mathematical models and methods for reliability analysis of electronic, electrical and software systems. Lviv Polytechnic Press, Lviv
17. Kharchenko V (2014) Availability assessment of information and control systems with online software update and verification. In: Ermolayev V, Mayr HC, Nikitchenko M, Spivakovsky A, Zholtkevych G (eds) Information and communication technologies in education, research and industrial applications, vol 469. Springer, Heidelberg, pp 300–324
18. Mudry P (2007) A reconfigurable hardware platform for prototyping cellular architectures. In: Vannel F, Tempesti G, Mange D (eds) RAW'07. international parallel and distributed processing symposium. IEEE international, 2007
19. Viktorov O (2007) Reconfigurable Multiprocessor System Reliability Estimation. Asian J Inf Technol 6(9):958–960