A PRACTICAL DEPENDABILITY MEASURE FOR EMBEDDED COMPUTER SYSTEMS

E. JONSSON [1)], S. ASMUSSEN [2)]

[1)] *Department of Computer Engineering, Chalmers University of Technology, Sweden*
[2)] *Institute of Electronic Systems, Aalborg University, Denmark*

Abstract. Dependability is commonly described by a number of attributes, such as reliability, availability, safety and security. Quantitative measures are found for each separate attribute e.g. reliability and availability, but are not defined for the totality of all attributes. This paper suggests a vectorized measure based on Markov processes. The measure covers reliability, safety and a modified version of the security attribute. It should be used for practical dependability trade-offs and is especially applicable to autonomous systems with embedded computers, such as aerospace vehicles and control systems. Key issues are the concepts of *degradability*, *subservice* and *service level*. The measure is based on the expected operating time on an operational service level and the total operational time before failure for failed service levels.

Keywords. Dependability, Safety, Security, Measure, Embedded Computer System

1. INTRODUCTION

The increasing complexity and criticality of systems with embedded computers, such as aeroplanes, cars and control systems, have emphasized the importance that these systems have the quality of trustworthiness, normally described by the term **dependability** (Carter 1979, Laprie 1985, Laprie 1992). Further, it must be possible to demonstrate that this quality exists. No general dependability measure, covering all dependability attributes has yet been defined. Instead, the measures used reflect one single dependability attribute at a time.

In many cases, however, it would be valuable for the system designers to calculate a measure that would describe several dependability attributes at the same time. This would help them to make the correct trade-offs and comparisons between different design alternatives. This paper presents a vectorized performance measure based on continuous-time Markov processes. For fully operational or degraded *service levels*, the numbers in the vector describe the expected amount of time the system will be operating at that specific level. For failed service levels the expected time of operation before each type of failure is used as a measure.

It is thus possible to merge attributes such as *reliability*, *privacy* and *performability* into one single quality. If two or more service levels for a failed system are introduced, the concept of *safety* can also be reflected.

A fact that is not considered is that systems can very often be repaired and upgraded. However, the measure can be applied to repairable systems for each specific operational period.

2. CLASSICAL DEPENDABILITY ATTRIBUTES

The classical viewpoint of dependability is described in (Laprie 1992) and references therein. Dependability is described by four basic attributes which are primarily related to non-degradable systems: *reliability, availability, safety* and *security*. *Performability* is a performance-related measure of dependability used for degradable systems.

The concepts of security and performability are discussed below, whereas reliability and safety should need no further comments. Availability is only applicable to repairable systems and not covered by this paper.

2.1 Security

Security has traditionally been seen as composed of three characteristics - *confidentiality, integrity* and *availability*, (Deswarte *et al.* 1991, ITSEC 1990, Pfleeger 1989). Confidentiality, which is also called *secrecy*, deals with the unauthorized disclosure of information. Integrity means that information or other assets of the system must only be changed, deleted, created etc. by authorized parties. Availability means that system assets must indeed be available to the proper system user.

Security is probably the attribute that is least well integrated into the dependability concept and there exists several other definitions of this attribute, see e.g. (Date 1990, Muftic 1989).

2.2 Performability

An item of particular interest is the modelling of the service delivered by a system. In many cases the service does not have a binary characteristic such as presence or absence, but is rather something that can be reduced gradually or *degraded*. The service will be delivered at a reduced performance level, specified by a pre-defined subset of the full specification.

Performability is the attribute used to describe systems whose service can be reduced gradually or be degraded. A number of performance measures have been suggested by various authors, see e.g. (Beaudry 1978, Laprie 1983, Meyer 1980, Smith and Trivedi 1987, Sanders and Meyer 1989).

It should be noted that performance can fall into two categories: *Quantitative performance reduction* reflects a continued delivery of the same service, but at a reduced level. However, many systems do not deliver one service only, but a number of different **subservices**. A disruption of one or several subservices means that a system is offering *functional performance reduction*.

A performance reduction, be it quantitative or functional, is due to a failure of a system component or subsystem. A proper term for this performance reduction is **system degradation**, since the failure of one component will in general only lead to a degraded system and not a failed system. A *failure* is then a special case of degradation, leading to a failed system.

3. SYSTEM MODEL

3.1 Attributes

The measure derived later in this paper is based on the traditional dependability attributes as described above, where performability is used in the sense of functional as well as quantitative performance reduction. However, for security, a different, user-oriented approach is taken.

One of the problems with security is that it describes not only the system behaviour, i.e. the service that the system delivers to the environment, but also the system's ability to resist external faults, and in particular intentional faults, which is a pure fault prevention issue. The dependability measure proposed in this paper is intended to only describe the behaviour of the system, and therefore fault-prevention characteristics are ignored.

This is achieved by means of interpreting dependability as composed of three main attributes: *reliability, privacy* and *safety*. Reliability relates to the delivery-of-service to the user of the system, whereas privacy relates to the denial-of-service to unauthorized users, i.e. unauthorized users must not be able to obtain information from the system, nor be able to use the system in any other way. Degradations or failures can be due to "reliability" faults as well as "privacy" faults. Safety is the reliability or privacy with respect to the non-occurrence of catastrophic failures. For further details see (Jonsson and Olovsson 1992).

3.2 Impairments

The model used for dependability impairments is based on (Laprie 1985) with modifications as suggested in (Jonsson 1991). A fault is an event-type "undependability input" that causes an error in the system. Faults include not only component failures, but also environmental effects, deliberate system interaction with the intention to create system failures, design faults, handling faults and others. An error is defined as an undesirable system state. An error may or may not lead to a reliability or privacy degradation or failure and thereby reduce the system's dependability.

3.3 Service levels

A **service level** is defined as a group of system states, each with a user-specified degree of performance or functional accomplishment. The highest service level is denoted **service level 0 (SL0)** or **full service level**. This level must include the system state that describes the complete fulfillment of all the requirements in the specification, the fully operational state.

In the simplest case there is only one more service level, the **failed service level**, corresponding to

the system **failed state**, when no service is delivered or the service delivered is of no use to the user. A disruption of one or several subservices may lead to a performance reduction and cause a transition to a **degraded service level**. In this case the system is **degradable**.

The safety aspect of dependability is modelled by means of introducing two or more service levels for a failed system. The service level related to a catastrophic failure is the lowest one. No transitions between failed service levels are defined.

4. A DEPENDABILITY MEASURE

4.1 Introduction

This section provides a mathematical definition of the vectorized dependability measure based on the system model and covering the dependability attributes described in Section 3. The method is based on a user-defined set of service levels and a set of constant failure rates which quantify the rate of transitions between levels. In general the failure rates for each level are not constant but rather functions of time, which introduces a further complication to the procedure. This problem could be handled by means of applying phase–type assumptions, see (Neuts 1981),(Jonsson and Asmussen 1991).

4.2 Definition

It is assumed that the state of the system can be modelled as a continuous time Markov process $\{X_t\}_{t\geq 0}$ with a finite state space E, in which each service level, SLn, can be identified with a subset of states in E. Thus, E is the disjoint union $SL0 + \cdots + SL\ell$, where ℓ is the number of service levels. Further, service levels $0, \ldots, k$ correspond to operational states O, i.e. the states in which the system functions, in the sense that it delivers a full or degraded service to the user. Service levels $k+1, \ldots, \ell$ correspond to the *failed* states F, i.e. states in which the system is not functioning, meaning that it is not delivering any service of interest to the user. That is

$$\begin{aligned} E &= O + F \quad \text{where} \\ O &= SL0 + \cdots + SLk, \\ F &= SL(k+1) + \cdots + SL\ell. \end{aligned}$$

In the simplest case, corresponding to the traditional operational-failed model, O consists of just one single state o, and F consists of just one single state f. In more complex situations, the different states in O represent different full or degraded service levels, and F represents different types of failed states.

Transitions $i \to j$ have intensity λ_{ij} $(i, j \in E,\ i \neq j)$, and the initial probability $\mathbb{P}(X_0 = i)$ is denoted by π_i. In most situations, the system will always start in a fixed state i_0 so that

$$\pi_j = \begin{cases} 1 & j = i_0 \\ 0 & j \neq i_0 \end{cases} . \tag{1}$$

It is also assumed that the system starts at the highest service level, so that $i_0 \in SL0$. Transitions between operational states represent degradations, and transitions to a failed state represent failures. No transition will ever take place from a failed state, i.e. after entering a failed state, the system stops evolving. Therefore, failed states are absorbing so that $\lambda_{fj} = 0$ for $f \in F$ and all $j \in E$. For mathematical convenience we shall use the notation that the intensity for leaving state i is $\lambda_i = \sum_{j\neq i} \lambda_{ij}$, and we write $\lambda_{ii} = -\lambda_i$.

Example 1: In many situations, the system state is described by a finite number p of components, each of which may be functioning or failed. Thus a typical state $i \in E$ has the form $i = (b_\alpha), \alpha = 1, \ldots, p$, where the b_α are 0 or 1, $b_\alpha = 1$ indicating that component α is functioning and $b_\alpha = 0$ that it has failed. Note however, that E may be a proper subset of all such 0–1 combinations; for example in a k out of n system, it is possible to collapse all 0–1 combinations with $k+1$ or more zeros into the single state $f = 00\cdots0$. The Markovian assumption amounts to assuming that each component has an exponential lifetime, with intensity parameter ν_α, say, for the αth, or equivalently a constant failure rate ν_α. If the system starts with all components functioning, (1) holds with $i_0 = 11\cdots1$. A transition from i to j is only possible if j is obtained from i by replacing one of the 1's, say at the αth place by 0, and the intensity is then ν_α. Thus, for example, for a 2 out of 3 system or Triple Modular Redundant (TMR) system,

$$\begin{aligned} \lambda_{111,011} &= \nu_1, \quad \lambda_{111,101} = \nu_2, \quad \lambda_{111,110} = \nu_3, \\ \lambda_{011,000} &= \nu_2 + \nu_3, \quad \lambda_{101,000} = \nu_1 + \nu_3, \\ \lambda_{110,000} &= \nu_1 + \nu_2, \end{aligned}$$

and all other λ_{ij} are zero.

Instead of 0–1 combinations, a common notation is upper– and lower case Roman letters, say A meaning that the first component is functioning and a that it has failed, and then the notation ν_A is used instead of ν_1. This notation is used in the examples below. □

Assuming that O has n states and F m, the $2(n+m)$ vector

$$\boldsymbol{v} = \left((u_i)_{i\in O}, (p_i)_{i\in O+F}, (v_i)_{i\in F}\right), \tag{2}$$

is suggested as a measure of dependability of the system, where u_i is the expected time in state i or **Mean Time To Degradation** (= MTTD), p_i the probability that the system ever enters i and v_i is the expected time to absorption in any $i \in F$, weighted by the reciprocal of the probability that it actually fails in a given $i \in F$, as discussed in the following. The conditional expected time to absorption in a given i, multiplied by the probability of this absorption, is denoted w_i. This can be described in formal mathematical terms:

$$u_i = \mathbb{E}\int_0^\infty I(X_t = i)dt, \quad i \in O, \tag{3}$$

$$v_i = \frac{1}{p_i} \cdot \sum_{j \in F} w_j, \quad i \in F, \tag{4}$$

$$w_i = \mathbb{E}\int_0^\infty I(t < \tau_i < \infty)dt, \quad i \in F, \tag{5}$$

$$p_i = \mathbb{P}(\tau_i < \infty), \quad i \in E = O + F, \tag{6}$$

where $\tau_i = \inf\{t \geq 0 : X_t = i\}$ ($\tau_i = \infty$ if no t with $X_t = i$ exists) is the hitting time of i, i.e. the time of first entry, and where I is the indicator function (e.g., $I(X_t = i) = 1$ in (3) if $X_t = i$ and $I(X_t = i) = 0$ if $X_t \neq i$). For computational purposes, it is noted that

$$u_i = \frac{p_i}{\lambda_i}. \tag{7}$$

In order to illuminate the interpretation of v_i, the following discussion is pursued: Let L be the lifetime of the system. Then

$$L = \min_{i \in \text{set of failed states}} \tau_i \tag{8}$$

and it is noted that there is precisely one i such that

$L = \tau_i < \infty$ and all other $\tau_j = \infty$. Then

$$w_i = E[L; \tau_i < \infty] = E[L|\tau_i < \infty] \cdot P(\tau_i < \infty).$$

Now let $k = 1, 2, \ldots$ be different but identical systems and let $L^{(k)}$ refer to system k. Assume that system $k+1$ replaces system k upon failure, let M_i be the number of systems used before a failure due to i occurs (M_i includes the failed system), and let S_i be the time until a failure due to i occurs. Then M_i is a geometric random variable with mean $1/p_i$, and by Wald's identity, see e.g. (Grimmet and Stirzaker 1992), the expected time to failure due to i is

$$E[S_i] = E\sum_{k=1}^{M_i} L^{(k)} = E[M_i] \cdot E[L] = \frac{E[L]}{p_i},$$

which is equal to v_i. Furthermore, if t is large, approximately $t/E[S_i]$ failures due to i will be seen. For example, $1/E[S_i]$ is a rate or intensity of failure due to i, in a long–term sense; intervals between failures due to i do not have an exponential distribution (but approximately so if p_i is small).

It is evident from the interpretation of $\sum w_j$ as the mean system life that

$$\sum_{j \in F} w_j = \sum_{i \in O} u_i,$$

and the calculations can be made using u_i. This can easily be checked in the following examples. However, there is an intrinsic interest of w_i as a measure referring to the individual failed states.

4.3 Examples

In this section, a few examples to illustrate the use of the dependability measure (2) are presented. The systems used in the examples are extremely simple so that the computational procedure will be as transparent as possible.

Example 2: Consider a washing machine with two functions: wash and spin-dry, A meaning that the machine can wash and a that it cannot, and B meaning that the machine can spin-dry and b that it cannot. Interpreting the failure of the washing function as more serious than that of the spin-drying function, the service levels may be interpreted as

$$\begin{aligned} SL0 &= AB \quad (\text{wash and spin} - \text{dry}) \\ SL1 &= Ab \quad (\text{wash only}) \\ SL2 &= aB + ab \quad (\text{no wash}) \end{aligned}$$

Here the finite state space E is identified by $\{SL0, SL1, SL2\} = \{AB, Ab, aB + ab\}$, where $O = \{SL0, SL1\}$ and $F = \{SL2\}$. A state diagram for the system is given in Fig. 1.

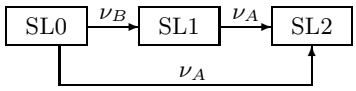

Fig. 1. State diagram for washing machine

Noting that $SL1$ or $SL2$ are entered from $SL0$ with probabilities $\nu_B/(\nu_A+\nu_B)$, $\nu_A/(\nu_A+\nu_B)$, respectively, it follows that

$$\begin{aligned} p_{SL0} &= 1, \quad u_{SL0} = \frac{1}{\nu_A + \nu_B}, \\ p_{SL1} &= \frac{\nu_B}{\nu_A + \nu_B}, \quad u_{SL1} = p_{SL1} \cdot \frac{1}{\nu_A}, \\ p_{SL2} &= 1, \quad v_{SL2} = w_{SL2} = u_{SL0} + u_{SL1}. \end{aligned}$$

□

Example 3: Consider a computerized car that has three computers: A, a background functions computer, B, a general purpose computer (ignition etc.) and C, a computer for vehicle dynamics (steering etc.). Interpreting the failure of C as catastrophic but that of A or B (or both) not, the service levels may be interpreted as

$$\begin{aligned} SL0 &= ABC \quad \text{(full service level)} \\ SL1 &= aBC \quad \text{(degraded service level)} \\ SL2 &= abC + AbC \quad \text{(failed service level)} \\ SL3 &= **c \quad \text{(catastrophically failed service level)} \end{aligned}$$

Here, E is identified with $\{SL0, SL1, SL2, SL3\}$, where $O = \{SL0, SL1\}$ and $F = \{SL2, SL3\}$. A state diagram for the system is given in Fig. 2.

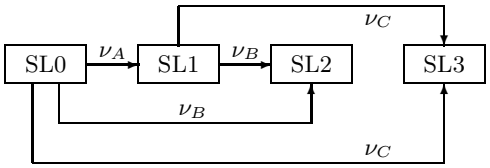

Fig. 2. State diagram for computerized car

Noting that $SL1$, $SL2$ or $SL3$ are entered from $SL0$ with probabilities

$$\frac{\nu_A}{\nu_A+\nu_B+\nu_C}, \quad \frac{\nu_B}{\nu_A+\nu_B+\nu_C}, \quad \frac{\nu_C}{\nu_A+\nu_B+\nu_C},$$

respectively, and that $SL2$ and $SL3$ are entered from $SL1$ with probabilities $\nu_B/(\nu_B+\nu_C)$, $\nu_C/(\nu_B+\nu_C)$, respectively, it follows that

$$\begin{aligned} u_{SL0} &= \frac{1}{\nu_A+\nu_B+\nu_C}, \\ p_{SL1} &= \frac{\nu_A}{\nu_A+\nu_B+\nu_C}, \quad u_{SL1} = p_{SL1}\cdot\frac{1}{\nu_B+\nu_C}, \\ p_{SL2} &= \frac{\nu_B}{\nu_B+\nu_C}, \; p_{SL3} = \frac{\nu_C}{\nu_B+\nu_C}, \\ w_{SL2} &= \cdots = \frac{\nu_B}{(\nu_B+\nu_C)^2}, \\ w_{SL3} &= \cdots = \frac{\nu_C}{(\nu_B+\nu_C)^2}, \\ v_{SL2} &= \frac{w_{SL2}+w_{SL3}}{p_{SL2}}, \; v_{SL3} = \frac{w_{SL2}+w_{SL3}}{p_{SL3}}. \end{aligned}$$

For example in the expression for p_{SL2} the first term is the contribution from the event that $SL1$ is entered after $SL0$ and the second term is the contribution from the event that $SL2$ is entered directly after $SL0$.

Note that $u_{SL0} + u_{SL1} = w_{SL2} + w_{SL3}$, as expected. Inserting numerical values $\nu_A = \nu_B = 1/1000$ failures per hour and $\nu_C = 1/10000$ failures per hour, yields $u_{SL0} = 476$hours, $u_{SL1} = 433$hours, $v_{SL2} = 1000$hours and $v_{SL3} = 10000$hours, where the last figure reflects the average (over a big number of systems) expected operation time before a catastrophic failure occurs. □

Example 4: Finally an example that includes privacy is given. Consider a process control system with two parallel, but different control lines. The process flow can take place along either line or along both lines at the same time. Each line is controlled by means of one computer, that is called A and B, respectively. Furthermore, the process control information that is stored in the system is regarded as very sensitive for competitive reasons, and is protected by a mechanism C. If either one of the computers A and B fails, the system is degraded, but still operational. If both A and B fail, the system is regarded as failed. A failure of the protection mechanism which entails a competitor obtaining the process control information is regarded as a catastrophic failure, irrespective of the system state in other respects. Thus, a catastrophic situation due to a privacy failure has occurred. The service levels are then

$$\begin{aligned} SL0 &= ABC \quad \text{(full operation and privacy)} \\ SL1 &= aBC + AbC \quad \text{(degraded operation)} \\ SL2 &= abC \quad \text{(failed operation)} \\ SL3 &= **c \quad \text{(catastrophic privacy failure)} \end{aligned}$$

Once again E is identified with $\{SL0, SL1, SL2, SL3\}$ where $O = \{SL0, SL1\}$ and $F = \{SL2, SL3\}$.
A state diagram is given in Fig. 3.

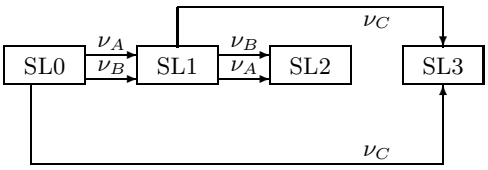

Fig. 3. State diagram for process control system

Note that the two states in $SL1$ must be handled separately. Therefore the two upper arrows to and from $SL1$ go to the state aBC and the two lower ones to the state AbC. Apart from this observation, the calculations are similar to those of example 3. For example, the expressions for u_i are:

$$\begin{aligned} u_{SL0} &= \frac{1}{\nu_A+\nu_B+\nu_C}, \quad u_{SL1} = u_{aBC} + u_{AbC}, \\ u_{aBC} &= p_{aBC}\cdot\frac{1}{\nu_B+\nu_C} = \frac{\nu_A}{\nu_A+\nu_B+\nu_C}\cdot\frac{1}{\nu_B+\nu_C}, \\ u_{AbC} &= p_{AbC}\cdot\frac{1}{\nu_A+\nu_C} = \frac{\nu_B}{\nu_A+\nu_B+\nu_C}\cdot\frac{1}{\nu_A+\nu_C}. \end{aligned}$$

In reality, it may be possible for a privacy failure to occur when the system is on service level 2. This could then be illustrated by an arrow between $SL2$ and $SL3$ in figure 3, and the calculations could be modified accordingly. □

5. SUMMARY

Even though dependability is a concept normally used in general and non-quantitative terms to de-

scribe computing systems and other systems, a step has been taken towards a quantitative understanding of it. This is done by merging attributes like reliability, performability, safety and privacy into a more general quality, and defining a measure for it. The system considered for dependability assessment is perceived in terms of service-delivery and service-denial and this service is considered as being normally degradable, i.e. it can be delivered or denied in various amounts or at different levels. A mathematical definition for a vectorized measure based on Markov processes has been given. The measure reflects the time the system is operational on a certain service level and the probability that it will reach this level, if ever. The measure is only applied to non-repairable systems, i.e. no feed-back in the Markov process is defined.

ACKNOWLEDGEMENTS

We wish to thank Mikael Andersson at the Department of Mathematics, Chalmers University of Technology, for valuable comments on our paper.

REFERENCES

Beaudry, M. (1978). Performance-related reliability measures for computing systems.. In 'IEEE Transactions on Computers'. Vol. C-27.

Carter, W. (1979). 'Fault-detection and recovery algorithms for fault-tolerant systems'. *Proc. EURO IFIP79* pp. 725–734.

Date, C. (1990). *An introduction to database systems*. Vol. 1 of *ISBN 0-201-51381-1*. 5th edn. Addison-Wesley. p. 429ff.

Deswarte, Y., L. Blain, J.C. Fabre and J.M. Pons (1991). *Security in Delta-4: A generic architecture for dependable distributed computing*. ISBN 3-540-54985-4. Springer Verlag. pp. 329–350.

Fortes, J. and C.S. Raghavendra (1982). Dynamically reconfigurable fault-tolerant array processors. In 'Proc. 12th IEEE Int. Symposium'. pp. 386–392.

Grimmet, G. and D.R. Stirzaker (1992). *Probability and Random Processes*. ISBN 0-19-853666-6. Clarendon Press. p. 396ff.

Heimann, D., N. Mittal and K.S. Trivedi (1991). Dependability modelling for computer systems. In 'Proceedings of the annual Reliability and Maintainality Symposium'. pp. 120–127.

ITSEC (1990). *Information Security Evaluation Criteria: Harmonized criteria of France, Germany, the Netherlands and the United Kingdom*. Kollen-Druck, Bonn.

Jonsson, E. (1991). A system-based model for dependable computers. Technical Report 116. Department of Computer Engineering. Chalmers, Gothenburg, Sweden.

Jonsson, E. and S. Asmussen (1991). A dependability measure for degradable computing systems. Technical report 117. Department of Computer Engineering. Chalmers University of Technology, Gothenburg, Sweden.

Jonsson, E. and T. Olovsson (1992). On the integration of security and dependability in computer systems. In 'Proceedings of the IASTED International Conference: Reliability, Control and Risk assessment'. ISBN 0-88986-171-4. IASTED. pp. 93–97.

Laprie, J. (1983). Trustable evaluation of computer systems dependability. In 'International Workshop on Applied Mathmatics and Performance Reliability Models of Computer Communication Systems,'. University of Pisa.

Laprie, J. (1985). Dependable computing and fault tolerance: concepts and terminology. In 'Proc. 15th IEEE International Symposium on Fault-tolerant Computing, FTCS15'. Ann Arbor, Michigan.

Laprie, J. (Ed.) (1992). *Dependability: Basic concepts and terminology*. ISBN 3-211-82296-8. Springer Verlag.

Meyer, J. (1980). On evaluating the performability of degradable computing systems. In 'IEEE Transactions on Computers'. Vol. C-29. pp. 720–731.

Muftic, S. (1989). *Security Mechanisms for Computer Networks*. ISBN 0-7458-0613-9. Ellis Horwood Ltd, England.

Neuts, M. (1981). *Matrix-Geometric Solutions in Stochastic Models*. John Hopkins University Press, Baltimore.

Pfleeger, C. (1989). *Security in Computing*. ISBN 0-13-799016-2. Prentice-Hall International.

Sanders, W. and J.F. Meyer (1989). A unified appraoch for specifying measures of performance, dependability and performability. In '1st IFIP Conference on Dependable Computers'.

Smith, R. and K.S. Trivedi (1987). A performability analysis of of two multi-processor systems. In 'Proc. 17th IEEE International Symposium on Fault-tolerant Computing, FTCS17'. Pittsburg, Pennsylvania. pp. 224–229.

Trivedi, K. (1982). *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. ISBN 0-13-711564-4. Prentice-Hall, New Jersey.