Methods of reliability assessment of heterogeneous redundant systems

Rogova, Elena; Lodewijks, Gabri

DOI
10.1016/j.ifacol.2016.07.564

Publication date
2016

Document Version
Accepted author manuscript

Published in
IFAC-PapersOnLine - 8th IFAC Conference on Manufacturing Modelling, Management and Control (MIM 2016)

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
Methods of reliability assessment of heterogeneous redundant systems

Elena Rogova*. Gabriel Lodewijks.*

*Department of Maritime and Transport Technology,
Delft University of Technology, Mekelweg2, 2628CD Delft, The Netherlands
(Tel: +31 015 27 82889; e-mail: E.S.Rogova@tudelft.nl, G.Lodewijks@tudelft.nl)

Abstract: Reliability assessment is required for the determination of a safety integrity level (SIL) of safety systems in accordance to the functional safety approach. Functional safety standards suggest formulas for calculating PFD/PFH which numerical values are used for establishing correspondence to the SIL. However these formulas cannot be used for heterogeneous redundant systems with a combination of mechanical, electronic/electrical components and constant and non-constant failure rates. In this paper we present an overview of reliability assessment methods that are able to cope with these features of heterogeneous redundant systems, show their advantages, drawbacks and limitations in application.

Keywords: reliability analysis, redundancy, heterogeneous systems, non-constant failure rates, probability of failure on demand, safety integrity level.

1. INTRODUCTION

In the functional safety approach, safety instrumented systems (SIS) perform safety functions. Each safety function has a determined SIL (safety integrity level) from 1(minimum) till 4 (maximum). Correspondence of the SIL to the required SIL is a very important step in the design stage of a control system. Standards IEC 61508, 61511 and 62061 describe in details the procedure of reliability assessment of SIS for the determination of the corresponding SIL (IEC 61511-1, 2004; IEC 62061, 2005; IEC 61508-1, 2010). Analytical formulas for calculating PFD (Probability of failure on demand) and PFH (Dangerous Failure Frequency) for systems with M-out-of-N architecture are presented in book 6 of IEC 61508 (IEC 61508-6, 2010). However these formulas can be used only if the failure rates of a system are constant and channels are identical. For heterogeneous redundancy, that is defined as mixing of different types of components (Sharma et al. 2011) with different channels and combination of constant and non-constant failure rates, it is necessary to apply other methods.

A heterogeneous M-out-of-N redundancy architecture can be used in old mechanical safety systems when, instead of its full replacement, redundancy can be introduced by adding the required electrical/electronic components into the system. Due to the hardware diversity such redundancy significantly reduces common cause failures (CCF) and dramatically increases diagnostic coverage (DC) (IEC 62061, 2005; IEC 61508-6, 2010).

In this paper we will consider different methods that can be applied for the reliability assessment of different types of heterogeneous redundant systems. In addition we will show some possibilities to avoid excessive complexity and describe conditions when systems with non-constant failure rates can be considered as systems with constant failure rates and can be calculated by using conventional formulas presented in functional safety standards. Analytical formulas and algorithms suggested by the methods, considered in this paper, can be used in different control systems at the design stage to suit the required SIL. It is also important for the determination of a repair/maintenance policy.

The structure of the paper is as follows: in Section 2 we consider the main features of heterogeneous M-out-of-N redundant systems and the main issues in reliability assessment of such systems. Section 3 presents the difference between systems with constant and non-constant failure rates and describes conditions under which systems with degradation can be considered as systems with approximately constant failure rates. Reliability assessment methods divided into several groups are discussed in Section 4. Section 5 contains a description of the practical implementation of heterogeneous M-out-of-N redundancy architectures as a part of large engineering systems. In Section 6 we conclude.

2. HETEROGENEOUS REDUNDANT SYSTEMS

As was mentioned in Section 1, the main feature of heterogeneous redundant systems is the existence of different types of components. There are many different components that can be used in control systems from the level of sensors and detectors till the level of actuators and mechanisms. From the reliability point of view we separate these components based on three categories:

1) The first category is based on the nature of component: mechanical or electrical/electronic.

2) The second category is a sequence of the first one: constant (λ) or non-constant (λ(t)) failure rates.

3) The third category defines the difference or identity of channels in redundancy architecture:
a. different components are located in the same channel, but all channels are identical;
b. channels are also different.

The choice of constant or non-constant failure rate in the second category depends on many parameters. First of all it depends on the available information for the specific component and approximation on the basis of a chosen model. Mechanical and electrical/electronic components have different physical principals. Many mechanical components have degradation of their reliability parameters that means non-constant failure rates. Electronic/electrical components also can have degradation. However the majority of them are assumed to have approximately constant failure rates.

Fig. 1 demonstrates different types of heterogeneous M-out-of-N architecture. Case a) is an M-out-of-N architecture with different channels and constant failure rates. The problem of reliability assessment of such architecture can be solved by using reliability block diagram (RBD) and all other methods that work with constant failure rates. Case b) looks like a homogeneous redundant system due to its identical channels. However heterogeneity of this system is in different types of components inside of each channel. For this case it is important to get a failure rate function for a channel based on failure rates of all components in a channel and use reliability assessment methods that are able to work with non-constant failure rates. Some methods (see Section 4) work only for systems with one component level redundancy and cannot be used for systems with several different components in one channel. Case d) is difficult for reliability assessment due to different channels and different non-constant failure rates. Case c) is even more difficult case because of different channels and a combination of constant and non-constant failure rates.

In general reliability assessment methods for heterogeneous redundant systems have two main issues: 1) non-identical channels and 2) non-constant failure rates. It is not difficult to find methods for each of these issues separately. But it is not easy to find a method that is able to cope with both of these issues simultaneously.

3. CONSTANT OR NON-CONSTANT FAILURE RATES

As was discussed in Section 2, many mechanical components have degradation over time that means non-constant failure rates. However sometimes it is not easy to obtain a failure rate function and to find an appropriate reliability method. In some cases non-constant failure rates can be assumed as approximately constant under specific conditions. In this section we consider these conditions and discuss when it is reasonable to calculate the failure rate function instead of a failure rate.

Alfredsson and Waak (Alfredsson and Waak, 2001) compare constant and non-constant failure rates. The authors separate constant demand rates and constant components rates. They assume constant demand rates without assuming constant component failure rates. The reason of this assumption is that ‘the demand process for a given item type at a given site is the result (in essence the superposition) of a number of component failure processes’. In this case, based on Drenick’s theorem, the demand process can be approximated by a Poisson process, that means the demand rate is approximately constant (Alfredsson and Waak, 2001). Jones (Jones, 2001) considers a failure intensity analysis for estimation of system reliability using a non-constant failure rate model. He conducts an analysis of failure intensity curve of CMOS digital integrated circuits with 1000 hour intervals. The shape of the curve obtained by Jones is ‘ample evidence that the constant failure rate assumption for this type of device is incorrect’ (Jones, 2001). It is also important to notice that Jones considers only the first part of the bath-tube curve by using an example of CMOS digital devices. For mechanical components in general we are focused on the last region of the bath-tube curve that is related to the wear-out region.

For obtaining a failure rate function it is necessary to choose an appropriate distribution that can describe a degradation process. There are different distributions that can be chosen. However, many researchers and practitioners use a Weibull distribution for the mathematical description of the wear out failure characteristics (Chudoba, 2011; Kumar and Jackson, 2009; Keller and Giblin, 1985). A failure rate function of two-parameter Weibull distribution is demonstrated in (1):

\[
\lambda(t) = \frac{\alpha t^{\alpha-1}}{\eta^\alpha}
\]

where \(\alpha\) – Weibull shape parameter; \(\eta\) – Weibull scale parameter.

Weibull shape and scale parameters can be obtained from real statistical data and also from Weibull databases where values of \(\alpha\) and \(\eta\) are presented for typical components. These databases are very helpful if real statistical data is not available. However such data from databases should be used with caution because they give very approximate average values for components. The same components produced by different manufacturers can have very different Weibull parameters.

Constant failure rates can be applied as an approximate solution for components with non-constant failure rates if the following condition is met: the difference in values of the
failure rate at the beginning and at the end of the interval is not significant. This means that the calculated PFD/PFH values of a system at the beginning and at the end of the interval should correspond to the same SIL. As a consequence of this condition, the test interval has to be chosen properly in accordance to the recommendations given by functional safety standards and Rausand and Hoyland (Rausand and Hoyland, 2004). It is important to understand that SIL-requirements for a safety system are the same for the whole test interval and if we neglect significant changes of failure rates, calculated values of PFD_{avg} and PFH may be much lower than the real values. For low-demand safety systems the proof-test interval is usually in the order of 6 months to 2-3 years (Rausand and Hoyland, 2004). Some test intervals can be too large for an approximation by a constant failure rate in case of degrading systems. Failure rates for some mechanical components obtained by using Weibull data bases and (1) are presented in Table 1.

Table 1. Failure rates for mechanical components (Rogova et al. 2015)

<table>
<thead>
<tr>
<th>Failure Rate</th>
<th>Solenoid Valve</th>
<th>Gears</th>
<th>Bearings</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ(t=1h)≠const</td>
<td>2.13·10^{-4}</td>
<td>8.27·10^{-7}</td>
<td>3.86·10^{-4}</td>
</tr>
<tr>
<td>λ(t=8760h)≠const</td>
<td>5.29·10^{-4}</td>
<td>1.18·10^{-7}</td>
<td>1.00·10^{-7}</td>
</tr>
<tr>
<td>λ_{avg}(t=8760h)=const</td>
<td>3.71·10^{-4}</td>
<td>9.0·10^{-4}</td>
<td>6.93·10^{-4}</td>
</tr>
</tbody>
</table>

As Table 1 shows, the non-constant failure rate of a solenoid valve can be approximated as a constant failure rate $\lambda_{avg}$ because the difference of values at the beginning and at the end of the test interval is negligible. However difference of failure rate values for gears at the beginning and at the end of the test interval is very large and the failure rate function cannot be replaced by constant value. The difference between values of failure rates of bearings at the beginning and at the end of the test interval is larger than for solenoid valve. This change of failure rate should be considered taking into account a correspondence to the required SIL at the beginning and at the end of the test interval to take a decision about possibility to make an approximation by constant failure rate. This method of correspondence to SIL is applicable for all components (solenoid valve, gears, bearings and others) but especially useful in those cases when approximation by constant failure rate is not obvious.

It is also important to notice that non-constant failure rates allow us to make a valuable reliability prognosis of equipment. It can help in maintenance scheduling. For example if a compressor is one of the most critical components of a safety system, it is very important to follow the degradation and to build a failure rate function that can help in calculating the PFD_{avg}/PFH values and determination of the corresponding safety integrity level (SIL) of a system. The example of such measurements of vibration rate in compressor is shown in Table 2.

<p>| Table 2. Increase of vibration rate of compressor |
|-----------------|-----|-----|-----|-----|-----|</p>
<table>
<thead>
<tr>
<th>Vibration rate, mm/s</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>1</td>
<td>1.3</td>
<td>1.8</td>
<td>2.6</td>
</tr>
</tbody>
</table>

For the purpose of reliability prognosis, compressor is tested every week (see Table 2). Based on the failure rate function obtained from these measurements, it is possible to conclude that for example after N weeks of exploitation, SIL of safety system that contains compressor will not correspond to the required SIL. This means a necessity to plan maintenance before appearance of critical vibration. The similar measurements can be conducted for other mechanical equipment of heterogeneous M-out-of-N redundancy architecture where such periodical measurements (like partial stroke tests for example) are a part of diagnostics.

## 4. RELIABILITY ASSESSMENT METHODS

In this Section we consider methods of reliability assessment of heterogeneous M-out-of-N redundancy architectures. These methods are grouped in accordance to the classification introduced in Section 2. Each case (a, b, c, d) has a set of methods that are applicable for the reliability assessment of corresponding architectures (see Fig. 2).

Fig. 2. Reliability assessment for complex redundant systems.

Case a) can be easily solved by using a Reliability Block Diagram (RBD), a Fault Tree Analysis (FTA) or conventional Markov method. For example, Hildebrandt (Hildebrandt, 2007) applies a conventional Markov model for the calculation of the PFD value for a heterogeneous 1oo2 architecture. Case b) is more difficult. Here for the calculation of the PFD_{avg} value of M-out-of-N architecture we can use a direct method (Rausand and Hoyland, 2004):
\[
PFD_{avg} = 1 - \frac{1}{\tau} \int_0^\tau R_{Moore}(t) dt
\]
\[
R_{Moore} = \left( \sum_{i=1}^{N} \binom{N}{i} R^i \cdot (1 - R)^{N-i} \right) \cdot R_{CCF}
\]
where \( R_{Moore} \) is system reliability, calculated by using a Weibull distribution; \( R \) is reliability of one channel; \( \lambda_{DU}(t) \) is a failure rate function of dangerous undetected failures (DU); \( \beta \) is a CCF factor.

The direct method is simple and transparent. However exact (analytical) calculation of integral is impossible if failure rate function is described by Weibull distribution. In this case the method requires calculation of an approximate numerical solution that is not always suitable due to the accuracy of the results.

The method ‘Ratio between CDFs’ also can be used for reliability analysis of an M-out-of-N redundancy architecture with identical channels and non-constant failure rates (case b). Therefore the PFD\(_{avg}\) for the first test interval \( k_i \) (Rogova et al. 2015):
\[
PFD_{avg,k_i} \approx \left( \sum_{i=1}^{N} \binom{N}{i} \right) \frac{A_k}{a+1} \frac{(1-\beta)\lambda_{DU}(\tau t)}{\alpha} \frac{k^a}{2} + \frac{\beta\lambda_{DU}(\tau t)}{2}
\]
\[
A_k = \left( \sum_{i=1}^{k} \binom{N}{i} \right) \left( \frac{(1-\beta)\lambda_{DU}(\tau)}{\alpha} \right) \frac{1}{1+\gamma(i+1)}
\]
where \( k = N-M+1; \tau \) – test interval; \( A_k \) is a ‘multiplier’ which depends only on \( k \) and the Weibull shape parameter. The method also suggests a formula for PFD\(_{avg}\) prognosis:
\[
PFD_{avg,k_i} \approx \left( \sum_{i=1}^{N} \binom{N}{i} \right) \frac{A_k}{a+1} \frac{(1-\beta)\lambda_{DU}(\tau t)}{\alpha} \frac{k^a}{2} + \frac{\beta\lambda_{DU}(\tau)}{2}
\]

where \( i \) – is a number of test interval \( \tau \).

The main limitation of the method ‘Ratio between CDFs’ is the component level redundancy. For example if there are several components in one channel, this method cannot be applied: the method uses Weibull shape and scale parameters of a component in one channel. However, this method can be used if Weibull parameters were estimated for the whole channel in general, but not for each component of a channel separately.

Cases c) and d) are the most difficult ones because they combine two main issues: non-constant failure rates and non-identical channels. All earlier mentioned above methods are not applicable to these cases. However there are a few possible solutions. The first possible way is to use Markov-methods. The conventional Markov method is not applicable because conclusions about exponential distribution of corresponding time intervals for systems with non-constant failure rates are unjustified (Harlamov, 2008). However semi-Markov methods are able to cope with this problem. The main advantage of semi-Markov processes is to allow non-exponential distributions for transitions between states and to generalize several kinds of stochastic processes. Since in most real cases the lifetime and repair time are not exponential, this is very important’ (Limnios, 2001). For example, Kumar et al. (Kumar et al., 2013) consider a steady-state semi-Markov method for calculation of availability of repairable mechanical systems. A steady-state semi-Markov method suggests a solution by using an assumption that state probabilities are not changing. This assumption is not always applicable. That is the reason why the method can be accepted with caution. A steady-state Markov method starts from the state-diagram where CDFs (Cumulative Distribution Functions) are assigned for each transition instead of failure rates. Based on known CDFs it is possible to build a kernel matrix \( Q(t)^{[P_X]} \) (\( P - \) is a number of states), which elements together with sojourn times are used for calculating state probabilities. At the final stage of this method, PFD\(_{avg}\) and PFH values can be easily calculated based on values of steady-state probabilities. The steady-state method is time-consuming and does not give exact results but it can be used as an additional method for comparison of obtained results.

In the case of complex semi-Markov models, calculating the exact probability distribution of the first passage time to the subset of states is usually very difficult. Therefore, the only way is to find an approximate probability distribution of that random variable. This is possible by using the results from the theory of semi-Markov processes (SMP) perturbations. The perturbed SMPs are defined in different way by different authors (Grabski, 2014). There are significant results presented by Korolyuk and Turbin (Korolyuk and Turbin, 1976), Gertsbakh (Gertsbakh, 1984), Pavlov and Ushakov (Pavlov and Ushakov, 1978) and others. The difference of this method in comparison to conventional and semi-Markov method is clear from the beginning: at the stage of definition of system states. The space of \( K \) states should be divided into two subspaces: subspace \( A^\uparrow = \{0,\ldots,j\} \) when the system is ‘up’ and subspace \( A^\downarrow = \{j,\ldots,K\} \) when the system is ‘down’. As a result this method allows to obtain an approximate reliability function \( R(t) \) (Grabski, 2014). However solving complex matrix equations and other time-consuming calculations make this method difficult for application in practice. In addition, this method is applicable only if all conditions of the corresponding theorems are met. Obtained results are approximate and require comparison with other methods.

There are methods that are based on heuristic algorithms. For example Boddru and Xing consider the reliability of an M-out-of-N redundancy architecture with mixed spare types for different redundancy modes: hot, cold, mixed (Boddru and Xing, 2012). Li and Ding presented research about optimal allocation policy of active redundancies to M-out-of-N systems with heterogeneous components (Li and Ding, 2010). The question of reliability estimation of heterogeneous multi-state series-parallel systems was considered by Sharma et al. (Sharma et al., 2011) and Wang and Li (Wang and Li, 2012). However, these papers are mainly focused on existing heuristic algorithms and some difficulties related to optimization problems and do not aim at a practical
calculation of system reliability in the concept of functional safety.

GSNP (Generalized Stochastic Petri Nets) also can be used for calculation reliability in cases c)-d). This method was described for instance by Santos et al. (Santos et al., 2014). The authors use the GSNP model for an estimation of the system age and a Weibull failure rate function for the failure rate function. Dersin et al. (Dersin et al., 2008) use a Petri-nets approach for maintenance modelling. GSNP is one of the most complex and time-consuming methods for reliability assessment of architectures c)-d). It requires a high level of special knowledge and it is not easy to build a model. Markov methods are easier in the stage of model building. Often GSNP is used in a combination with Monte-Carlo simulation (MCS).

Monte-Carlo simulation is not shown on Fig. 2. However MCS is used as a part of many methods very often. It is also used for comparison and verification of the results obtained by using other methods. The main algorithm of MCS is in the discretization of the problem of calculation of state probabilities: the test interval \([0;\tau]\) should be splitted into intervals with duration \(h\). Thus the reformulated problem is the problem of defining of state probabilities at discrete moments of time: \(P_i((j-1)h, jh)\). The main principles of MCS with application in reliability theory are described by E. Zio (Zio, 2013).

5. IN PRACTICE

The purpose of applying redundancy is increase of reliability. ‘The capabilities of M-out-of-N redundancy make it an important tool for failure prevention. Sometimes components are deliberately subdivided in order to permit M-out-of-N redundancy to be applied’ (Hecht, 2004). Practical implementation of M-out-of-N heterogeneous redundancy architectures, which types are demonstrated in Section 2, is very wide. Case a) (see Fig. 1) is frequent in control systems: very often the same type of components are not totally identical and produced by different manufacturers. Moreover, as was discussed in the Introduction, this non-identity is recommended by the standards (IEC 61508-6, 2010; IEC 62061, 2005) to decrease CCF. Cases b)-d) are devoted to mechanical components in channels. Architecture b) can be not very reliable because of identical mechanical equipment in channels (even with different components inside of the channel). This type of redundancy gives us a very high probability of common cause failures and less probability to diagnose possible dangerous failures. Case d) is much better because here we use mechanical hardware diversity. A practical example of case c) is existence of different types of relays in different channels: electro-mechanical relays with degradation and electronic solid-state relay (SSR) with constant failure rates. The type of heterogeneous redundancy demonstrated in case c) is very interesting for application in control systems because the existence of components with different physical principals allow us to reach a very high reliability.

As discussed in Section 4 complex methods of reliability assessment are applied to the parts of large engineering systems because many of these methods are not able to calculate the reliability of a system with thousands of components and hundreds of subsystems. These methods are basically applied to some critical subsystems, and the obtained results are used in further work for investigation the reliability/availability of a system in general.

To start a reliability analysis of complex systems with heterogeneous subsystems (including M-out-of-N redundancy architectures), it is necessary to start from a general investigation of the system. If reliability analysis of such large systems is performed on the stage of exploitation (but not at the design stage), it is useful to focus on existing statistics of failures. In this case it is possible to use a qualitative FTA (fault tree analysis) (Rausand, 2014) or FMECA (Failure mode, effects and criticality analysis) (Rausand and Hoyland, 2004) for example. These tools will help in understanding the weakest points of a safety system from the reliability point of view. This understanding will allow to focus on specific subsystems for a detailed analysis by using methods described in Section 4. In accordance to the functional safety approach, the main purpose of a reliability assessment of critical degrading subsystems is checking of correspondence of SIL of safety system to the required SIL of safety function that is performed by safety system (IEC 61508-1, 2010).

6. CONCLUSIONS

In this paper we described main features and types of heterogeneous redundant systems and showed reliability analysis methods that are used for the reliability assessment of such systems. These methods were grouped in accordance to the suggested classification of the heterogeneous M-out-of-N redundancy architecture with corresponding comments concerning possibilities and limitations in application. This paper could serve as an input for further research and recommendations for reliability analysis of safety systems with heterogeneous M-out-of-N redundancy architectures in different engineering applications by using a functional safety approach.

ACKNOWLEDGMENTS

This research is financially supported by Lifinstituut (P90863).

REFERENCES


