Optimal replacement policy for safety-related multi-component multi-state systems

Ming Xu¹, Tao Chen²*, Xianhui Yang¹

¹Department of Automation, Tsinghua University, Beijing 100084, China
²Division of Civil, Chemical and Environmental Engineering, University of Surrey, Guildford GU2 7XH, UK

* Corresponding author. Tel.: +44 1483 686593; Fax: +44 1483 686581.
Email: t.chen@surrey.ac.uk

Abstract
This paper investigates replacement scheduling for non-repairable safety-related systems (SRS) with multiple components and states. The aim is to determine the cost-minimizing time for replacing SRS while meeting the required safety. Traditionally, such scheduling decisions are made without considering the interaction between the SRS and the production system under protection, the interaction being essential to formulate the expected cost to be minimized. In this paper, the SRS is represented by a non-homogeneous continuous time Markov model, and its state distribution is evaluated with the aid of the universal generating function. Moreover, a structure function of SRS with recursive property is developed to evaluate the state distribution efficiently. These methods form the basis to derive an explicit expression of the expected system cost per unit time, and to determine the optimal time to replace the SRS. The proposed methodology is demonstrated through an illustrative example.

Keywords: Safety-related system; Schedule point; Safety integrity level; Universal generating function; Structure function

1. Introduction
Engineered systems suffer from gradual performance deterioration and unexpected shock damage during operation. Therefore, maintenance activities are essential to prevent system deterioration and failure, or if complete prevention is not possible, to minimize the adverse impact of their occurrences [1, 2]. If the system is non-repairable, it will be replaced upon failure, while most repairable systems can be restored to a functioning state through maintenance. Replacement can be considered as perfect maintenance [2]. Generally, maintenance activities can be classified into two categories: corrective (un-planned), and preventive (planned) [3, 4]. Corrective maintenance refers to any action that restores the failed system to a working state. In contrast, preventive maintenance refers to an action carried out when the system is still functioning, with the goal to restore the system to a specified better condition.

In this paper, we examine the replacement schedule for non-repairable safety-related systems (SRS) with multiple components and states. SRS is widely used in industry to reduce or prevent risk of the production system (PS) being protected [5, 6]. International standards like IEC 61508 [7, 8] provide a means of ensuring that safety is effectively achieved by using SRS. In order to comply with these standards, SRS is usually quantified through the “safety integrity level” (SIL). For low demand mode of operation, SILs are defined in terms of average probability of dangerous failure on demand (PFDavg) as shown in Table 1. This paper seeks the optimal schedule point when the SRS should be replaced, with the objective of minimizing expected system cost per unit time under required safety (or equivalently, SIL).
Numerous studies have been reported in the literature to improve the maintenance strategies for PS; see [9-12] for examples. However, the maintenance policy for SRS, especially those with multiple components and states, has been under-explored. Multi-component and multi-state systems are more complex than the traditional systems with a single component or binary states, and thus they require special techniques for modeling and analysis [1]. Current studies on the maintenance of multi-component and multi-state systems are primarily focused on PS and may not be directly extendable to SRS. For example, Hsieh [13] examined a multi-state deteriorating PS with standby redundancy in which components may deteriorate due to random failures. To achieve the required system availability with minimal lifecycle cost, Levitin & Lisnianski [14] formulated a joint redundancy and replacement schedule optimization problem in which the element version, redundancy level, and replacement interval were optimized simultaneously. The same authors were the first to consider element replacement in multi-component multi-state systems [15], and they also addressed the corresponding imperfect repair scheduling problem [16]. Tan & Raghavan [9] developed a predictive maintenance model for multi-component and multi-state systems. Liu & Huang [1] discussed the optimal replacement policy for multi-state systems under imperfect maintenance.

However, these aforementioned model cannot be directly applied to SRS. When dealing with SRS, one should consider not only the availability, required safety and maintenance cost for SRS, but also the interaction between the SRS and PS being protected. Usually, some hidden modes of failure in SRS may be exposed only by the occurrence of a PS risk. Conversely, SRS failure modes may affect the operation of a PS. If this interaction is not accounted for, suboptimal solutions may materialize. Nevertheless, incorporating the interaction between SRS and PS significantly complicates the design and optimization of replacement policy. For example, the SRS should be replaced immediately when some failure modes in the SRS lead to a shutdown of the PS or the risk of the PS exposes the hidden failure modes of the SRS. Otherwise, the SRS will be replaced at a scheduled point in order to achieve the required safety in its lifecycle. Thus, if the interaction between SRS and PS is not considered, this important information is ignored when designing replacement policy.

The major contribution of this paper is to find the optimal replacement policy of multi-component and multi-state SRS with consideration of interaction between SRS and PS. The SRS will be modeled by a non-homogeneous continuous time Markov model (NHCTMM). We formulate a framework to facilitate solving this model based on the method of universal generating function (UGF). Especially, the structure function of SRS with recursive property is developed, and it significantly reduces the computational burden. Furthermore, we derive an explicit expression of the expected system cost per unit time, and determine the optimal schedule point to replace the SRS using mathematical optimization technique.

The remainder of the paper is organized as follows. In section 2, the SRS model is presented and the structure function of SRS is developed to aid efficient computation. Section 3 describes a replacement model of SRS by jointly considering SRS and PS in order to find the cost-minimizing schedule point. The proposed model is illustrated through the study of a high-integrity pressure protection system in Section 4. Finally, Section 5 concludes the paper.
### Abbreviations and nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>schedule point to replace SRS</td>
</tr>
<tr>
<td>$a_i$</td>
<td>system idle time in case $i$ ($i = 1, 2, 3, 4$)</td>
</tr>
<tr>
<td>$C(t)$</td>
<td>cost of system idleness for time $t$</td>
</tr>
<tr>
<td>$C_{FD}$</td>
<td>penalty cost of FD</td>
</tr>
<tr>
<td>$C_{pFS}$</td>
<td>penalty cost of FS</td>
</tr>
<tr>
<td>$C_R$</td>
<td>cost of repair the SRS</td>
</tr>
<tr>
<td>DG</td>
<td>degraded state</td>
</tr>
<tr>
<td>$E[T(A)]$</td>
<td>expected system operating cycle time</td>
</tr>
<tr>
<td>$E[SC(A)]$</td>
<td>expected system cost per operating cycle</td>
</tr>
<tr>
<td>FD</td>
<td>failure-dangerous state</td>
</tr>
<tr>
<td>FP</td>
<td>full performance state</td>
</tr>
<tr>
<td>FS</td>
<td>Failure-safe state</td>
</tr>
<tr>
<td>$f_i(t)$</td>
<td>probability density of state $x$ ($x \in G_S$) of SRS at instant $t$</td>
</tr>
<tr>
<td>$G_C$</td>
<td>set of states of component: $G_C = {\text{OP, FS, FD}}$</td>
</tr>
<tr>
<td>$G_S$</td>
<td>set of states of SRS: $G_S = {\text{FP, DG, FS, FD}}$</td>
</tr>
<tr>
<td>$g_s(i)$</td>
<td>state of subsystem combined by components ${1, 2, \ldots, i}$: $g_s(i) \in G_S$</td>
</tr>
<tr>
<td>HCTMMM</td>
<td>homogeneous continuous time Markov model</td>
</tr>
<tr>
<td>$I_{FS}(i)$</td>
<td>the collection of sets which are included in $\text{subs}<em>{FS}(i)$ and belong to $MCS</em>{FS}$</td>
</tr>
<tr>
<td>$I_{FD}(i)$</td>
<td>the collection of sets which are included in $\text{subs}<em>{FD}(i)$ and belong to $MCS</em>{FD}$</td>
</tr>
<tr>
<td>NHCTMMM</td>
<td>non-homogeneous continuous time Markov model</td>
</tr>
<tr>
<td>$MCS_{FD}$</td>
<td>collection of the minimal cut sets of SRS for FD</td>
</tr>
<tr>
<td>$MCS_{FS}$</td>
<td>collection of the minimal cut sets of SRS for FS</td>
</tr>
<tr>
<td>OP</td>
<td>operation state</td>
</tr>
<tr>
<td>PFD$_{avg}$</td>
<td>average probability of dangerous failure on demand</td>
</tr>
<tr>
<td>PS</td>
<td>production system</td>
</tr>
<tr>
<td>$P'(t)$</td>
<td>state distribution of component $i$ at time $t$</td>
</tr>
<tr>
<td>$P_x(t)$</td>
<td>probability of state $x$ ($x \in G_S$) of SRS at instant $t$</td>
</tr>
<tr>
<td>$p_{SIL}$</td>
<td>upper bound under SIL $x$</td>
</tr>
<tr>
<td>$p_s(t)$</td>
<td>probability of $s(t) = x$, ($x \in G_C$) at time instant $t$</td>
</tr>
<tr>
<td>$p_k(t)$</td>
<td>probability of SRS’s individual state $k$ ($k=1,2,\ldots,M$) at instant $t$</td>
</tr>
<tr>
<td>UGF</td>
<td>universal generating function</td>
</tr>
<tr>
<td>$V$</td>
<td>a set equals to ${0, 1, 2, 3}$</td>
</tr>
<tr>
<td>$R(i)$</td>
<td>$R(i) = {g_j(i-1), \phi(S_x(x_i))}$, and $R(i) \in V \times V$</td>
</tr>
<tr>
<td>$SC(A)$</td>
<td>cost in case $i$</td>
</tr>
<tr>
<td>SIL$x$</td>
<td>safety integrity level $x$</td>
</tr>
<tr>
<td>SRS</td>
<td>safety-related system</td>
</tr>
<tr>
<td>$s_i(t)$</td>
<td>state of component $i$ at time instant $t$</td>
</tr>
<tr>
<td>$\text{subs}_{FD}(i)$</td>
<td>the combination of components ${1, 2, \ldots, i}$ which are in FD state</td>
</tr>
<tr>
<td>$\text{subs}_{FS}(i)$</td>
<td>the combination of components ${1, 2, \ldots, i}$ which are in FS state</td>
</tr>
<tr>
<td>$S_x(x_i)$</td>
<td>the system state when ${s_1(t) = \text{OP}, \ldots, s_{i_1}(t) = \text{OP}, s_i(t) = x_i, s_{i+1}(t) = \text{OP}, \ldots, s_m(t) = \text{OP}}$</td>
</tr>
<tr>
<td>$T_F$</td>
<td>occurring time of failure (SRS and PS)</td>
</tr>
<tr>
<td>$x_i$</td>
<td>realization of $s_i(t)$: $x_i \in G_C$</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>system transition intensity matrix</td>
</tr>
<tr>
<td>$\beta$</td>
<td>cost per unit time due to system idleness</td>
</tr>
<tr>
<td>$\phi_{SRS}$</td>
<td>system structure function of SRS</td>
</tr>
<tr>
<td>$\phi$</td>
<td>structure function for components/subsystems in SRS</td>
</tr>
<tr>
<td>$\phi(S_x(x_i))$</td>
<td>mapping between $S_x(x_i)$ and the working state of SRS</td>
</tr>
<tr>
<td>$\chi(A)$</td>
<td>expected system cost per unit time</td>
</tr>
</tbody>
</table>
2. Model development

This study treats SRS as a multi-component and multi-state system. In addition, SRS and PS will be jointly considered in the modeling and optimization framework so as to account for their interaction. This section presents the assumptions for the system, the component model, the model for the SRS by combining multiple components, and finally the computational procedure by exploiting the recursive property of the system and the method of UGF. The abbreviations and symbols are listed as follows.

2.1 System assumptions

The model of system is based on the following assumptions:
1) The entire system is composed of PS (for production) and SRS (for reducing or preventing PS risks) [5, 17].
2) SRS is composed of multiple components and each component can experience two categories of failures: failure-safe (FS) and failure-dangerous (FD) [18]. In the absence of failure, the component is in operation (OP) state. SRS is non-repairable when failure occurs, and a replacement schedule will take place if needed.
3) Time to failure (FS or FD) is distributed exponentially, which implies that the SRS is a Markov process [19]. In addition, system components age over time, and thus the state transition intensity varies with time.
4) The state of any composition of components is unambiguously defined by the states of these components and the nature of the interaction of the components in SRS [18]. The occurrence of FS and FD events in one or more components may result in the occurrence of FS and FD events in SRS.
5) If a risk of PS arises, the specified safety function of SRS is required to act. At this point, if no FD event has occurred in SRS, the required function of SRS can be implemented and the PS can be taken to the safe state. Otherwise if FD event occurred before the occurrence of PS risk, SRS cannot function as required and a serious accident will occur. Furthermore, we assume that FD is unrecognizable by the operator until a serious accident occurs.
6) The PS will be taken to a safe state, if FS occurred in SRS. In addition, FS is always recognizable because it results in shut-down of the plant.

2.2 Model of SRS components

According to assumption 2, each SRS component has three states: OP, FS and FD. The state of component can be represented by the set

\[ G_c = \{OP, FS, FD\} \]

The state \( s_i(t) \) of component \( i \) at any time instant \( t \geq 0 \) is a random variable taking a value from \( G_c: s_i(t) \in G_c \). The state distribution of component \( i \) at time \( t \) can be represented by the set

\[ p^i(t) = \{p_{i,0}(t), p_{i,FS}(t), p_{i,FD}(t)\} \]

where \( p_{i,x}(t) \) represents the probability that \( s_i(t) = x, (x \in G_c) \).

When describing the degradation process of individual components, Lisnianski [20] employed the homogeneous continuous time Markov model (HCTMM). Basically, it assumes the time of transition between any two states follows a negative exponential distribution. The hypothesis that the transition intensity to the next state only depends on the current state is applicable to components having no age effect. In fact, it is more realistic to consider the case that a component’s deterioration process is not only related to the current state, but also to the age of the component [1, 21]. Taking this concept into account, the NHCTMM is utilized in this paper.
to derive the stochastic behavior of individual aging components through considering the age-related increasing state transition intensity [21, 22].

For a non-repairable aging component, the state-space diagram of an aging component is illustrated in Fig. 1, where $\lambda'_{op}(t)$ is the intensity of transition from state $OP$ to state $FS$, and $\lambda'_{op}(t)$ is the intensity of transition from state $OP$ to state $FD$.

![Fig. 1 State-space diagram of component i](image)

The probability of component $i$ in each state can be expressed as:

$$p'_{op}(t) = \exp \left[-\int_{0}^{t} (\lambda'_{op}(\tau) + \lambda'_{fd}(\tau)) d\tau \right]$$  \hspace{1cm} (1)

$$p'_{fs}(t) = \int_{0}^{t} \lambda'_{fs}(\tau) \exp \left[-\int_{0}^{\tau} (\lambda'_{fs}(s) + \lambda'_{op}(s)) ds \right] d\tau$$  \hspace{1cm} (2)

$$p'_{fd}(t) = 1 - p'_{op}(t) - p'_{fs}(t)$$  \hspace{1cm} (3)

2.3 State distribution of SRS according to a Markov process

![Fig. 2 State-space diagram of SRS](image)

The states of all $m$ components composing SRS at time $t$ are represented by the state vector $(s_1(t), s_2(t), \ldots, s_m(t))$. Suppose that the SRS has $M$ states; its state-space diagram is shown in Fig. 2 and can be described as a NHCTMM. According to the structure of SRS, the $M$ states can be classified into four sub-states: 1) full performance state (FP); 2) degraded state (DG); 3) FD; and 4) FS. Thus, the set of states of SRS can be represented by

$$G_s = \{FP, DG, FD, FS\}$$

Refer to the illustration in Fig. 2, state $M$ ($M \geq 3$) is considered as FP; the combination of states $M-1$ to $K$
The formulations given in [23] are directly used here to obtain the dynamic state probability of SRS. From (2), all the states of transition intensity can be summarized in the matrix

\[
\Lambda = \begin{bmatrix}
\lambda_{M,M-1}(t) & \lambda_{M,M-2}(t) & \cdots & \lambda_{M,2}(t) & \lambda_{M,1}(t) \\
0 & \lambda_{M-1,M-2}(t) & \cdots & \lambda_{M-1,2}(t) & \lambda_{M-1,1}(t) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \lambda_{2,1}(t)
\end{bmatrix}
\]

Furthermore, define an auxiliary function as follows

\[
G(t_1,t_2,j) = \int_0^{t_1} \sum_{i=1}^{j-1} \lambda_{i,j}(\tau)d\tau
\]

Then, the probability that the SRS in state \( M \) at time \( t \) is:

\[
P_M(t) = \exp\left[G(0,t,M)\right]
\]

The probability that the SRS in state \( M-1 \) at time \( t \) is:

\[
P_{M-1}(t) = \int_0^t p_M(\tau) \lambda_{M,M-1}(\tau) \exp\left[G(\tau,t,M-1)\right]d\tau
\]

The probability that SRS in state \( M-2 \) at time \( t \) can be considered as the sum of probabilities of existence of two routes: 1) the route from \( M \) to \( M-2 \) directly; 2) the route from \( M \) to \( M-1 \) and then from \( M-1 \) to \( M-2 \):

\[
p_{M-2}(t) = p_{M-2}^1(t) + p_{M-2}^2(t)
\]

\[
p_{M-2}^1(t) = \int_0^t p_M(\tau) \lambda_{M,M-1}(\tau) \exp\left[G(\tau,t,M-2)\right]d\tau
\]

\[
p_{M-2}^2(t) = \int_0^t \left( \int_0^\tau p_M(\tau_j) \lambda_{M,M-1}(\tau_j) \exp\left[G(\tau_j,t,M-1)\right] \lambda_{M-1,M-2}(\tau_j) \exp\left[G(\tau_j,t,M-2)\right] d\tau_j \right) d\tau
\]

In addition, the probability that the system in state \( i \) (\( i=1,2,\ldots,M-3 \)) is the sum of all routes from \( M \) to \( i \), which can be obtained in a similar way [23]. The maximum number of the routes of SRS is not more than \( 2^{M+1}-2-M \), which grows exponentially with the increase of number of states. When SRS becomes more and more complex, this method is computationally inefficient or even infeasible. Therefore, a more efficient method, based on UGF, is introduced to deal with this problem.

2.4 The distribution of states of SRS according to UGF

UGF is an essential tool to obtain the state distribution of the entire multi-state system. This method is based
on the z-transform and was first proposed by Ushakov [24, 25]. UGF approach is recursive so as to reduce the problem complexity and computational intensity by modularizing a system into its subsystems. Hence, this method is adopted to obtain the state distribution of SRS in this paper.

Note that the states of all \( m \) components within a SRS at time \( t \) are \( (s_1(t), s_2(t), \ldots, s_m(t)) \) and these states unambiguously determine the state of SRS (assumption 4). Therefore, the relationship between the component state vector and the SRS state variable \( S \) can be expressed by the following deterministic function (called the system structure function):

\[
S = \phi_{SRS}(s_1(t), s_2(t), \ldots, s_m(t))
\]  

(15)

In order to obtain the state distribution of the SRS one can represent the state distribution of component \( i \) as:

\[
u_i(z, t) = \sum_{s_{i,OP,FS,FD}} p_i^t(t)z^{s_{i,OP,FS,FD}} = p_i^{OP}(t)z^{OP} + p_i^{FS}(t)z^{FS} + p_i^{FD}(t)z^{FD}
\]

(16)

where Eq.(16) is called the z-transform of discrete random variable of component \( i \) [26].

To formulate the universal generating function of SRS, one needs to apply the composition operator, \( \otimes \):

\[
U(z, t) = \otimes_{\theta_SRS} \{\nu_1(z, t), \ldots, \nu_m(z, t)\}
\]

\[
= \otimes_{\theta_SRS} \left\{ \sum_{s_{i,OP,FS,FD}} p_i^t(t)z^{s_{i,OP,FS,FD}} \right\} \cdots \left\{ \sum_{s_{i,OP,FS,FD}} p_i^t(t)z^{s_{i,OP,FS,FD}} \right\}
\]

\[
= \sum_{s_{i,OP,FS,FD}} \ldots \sum_{s_{i,OP,FS,FD}} \left( \prod_{i=1}^{m} p_i^t(t)z^{s_{i,OP,FS,FD}} \right)
\]

\[
= P_F(t)z^{FP} + P_{DG}(t)z^{DG} + P_{FD}(t)z^{FD} + P_{FS}(t)z^{FS}
\]

Clearly, the total number of combinations of states of a SRS with \( m \) components is \( 3^m \) [26]. For SRS with a large number of components, this method requires an enormous number of evaluations of the structure function value. Fortunately, the structure function of SRS can be defined recursively and the states of some subsystems corresponding to SRS can be obtained. This allows to achieve considerable reduction of computation.

According to Fig. 2, the states of a SRS include FP, DG, FD and FS. Let \( \varphi(S_x(x_i)) : G^m_c \rightarrow G_2 \) be the mapping between \( S_x(x_i) \) and the working state of SRS. \( S_x(x_i) = \{s_1(t) = OP, \ldots, s_i(t) = x_i, s_{i+1}(t) = OP, \ldots, s_m(t) = OP\} \) denotes all the components are working in OP state except component \( i \). Based on assumptions 5 and 6, \( \varphi(S_x(x_i)) \) can be formulated as

\[
\varphi(S_x(x_i)) = \begin{cases} 
FP, & \text{for } x_i = OP \\
DG, & \text{for } (x_i = FD or x_i = FS) and \left( \{i\} \not\in MCS_{FD}, \{i\} \not\in MCS_{FS} \right) \\
FD, & \text{for } x_i = FD, \{i\} \in MCS_{FD} \\
FS, & \text{for } x_i = FS, \{i\} \in MCS_{FS}
\end{cases}
\]

(18)

where \( MCS_{FD} \) and \( MCS_{FS} \) are the collection of the minimal cut sets (MCSs) of SRS for FD and FS, respectively. MCS has been extensively studied in the literature [27-29], and a system’s MCSs can be obtained by many tools such as RAM commander, CAFTA and BlockSim [30]. For the same SRS, FD and FS states may have different MCSs.

We now discuss the structure function of SRS recursively. When considering the series-parallel system structure function, Levitin et al [31] used a conservative approach to obtain a simplified structure function with the following assumptions: 1) Any element in a subsystem becoming FD will always lead the related subsystem
into an FD state; 2) If the subsystem is composed of two components and is connected in a parallel way, the subsystem will be in an operational state if either component is in the operational state; 3) If the subsystem is composed of two components and is connected in series, the subsystem will be in an operational state only if both of the two components are in operational state; 4) In the remaining cases, the subsystem is in a FS state. However, this is an approximate method as the probability of FD state obtained by this method may be higher than the actual value when redundancy exists in the system. For this reason, the above assumptions are not used in this paper.

<table>
<thead>
<tr>
<th>Table 2 Structure function $\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component $\phi(S_i(x_i))$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>FP</td>
</tr>
<tr>
<td>DG</td>
</tr>
<tr>
<td>FD</td>
</tr>
<tr>
<td>FS</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>FP</td>
</tr>
<tr>
<td>DG</td>
</tr>
<tr>
<td>FD</td>
</tr>
<tr>
<td>FS</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

We encode all the $m$ components of SRS by integers from 1 to $m$ sequentially. Let $g_S(i)$, $1 \leq i \leq m$ be the working state of subsystem $\{1, 2, \ldots, i\}$ and $g_S(i) \in G_s$. Based on the assumptions 5 and 6, the structure function $\phi$ is shown in Table 2. The state of subsystem $\{1, 2, \ldots, i\}$ can be obtained by combining the subsystem $\{1, 2, \ldots, i-1\}$ and component $i$. This can be considered in five situations.

**Situation 1:** If the subsystem $\{1, 2, \ldots, i-1\}$ is in FS state ($g_S(i-1) = FS$) or component $i$ makes the SRS in FS state ($\phi(S_i(x_i)) = FS$), then the subsystem $\{1, 2, \ldots, i\}$ is in FS state ($g_S(i) = FS$). This situation is based on the assumption that FS state is always recognizable (assumption 6). The states with a superscript ‘a’ in Table 2 belong to this situation.

**Situation 2:** If any of the subsystem $\{1, 2, \ldots, i-1\}$ (or component $i$) makes the SRS in FP state ($\phi(S_i(x_i)) = FP$), then the subsystem $\{1, 2, \ldots, i\}$ will be in the same state as component $i$ (or the subsystem $\{1, 2, \ldots, i-1\}$). This is based on the fact that the subsystem or component in perfect condition before combination will not affect the working state of the combined subsystem. The states with a superscript ‘b’ in Table 2 belong to this situation.

**Situation 3:** If the subsystem $\{1, 2, \ldots, i-1\}$ is in DG or FD state ($g_S(i-1) = DG$ or FD) and component $i$ makes the SRS in FD state ($\phi(S_i(x_i)) = FD$), then the subsystem $\{1, 2, \ldots, i\}$ is in FD state ($g_S(i) = FD$). Clearly, since component $i$ has already made the SRS FD, its combination with the subsystem $\{1, 2, \ldots, i-1\}$ can only worsen the situation no matter whether the subsystem is in DG or FD state. Thus, FP and DG states can be excluded for the new subsystem $\{1, 2, \ldots, i\}$. In addition, Eq. (18) indicates that when component $i$ is FD, the combined subsystem $\{1, 2, \ldots, i\}$ cannot be in FS state, leaving the only possibility of FD. The states with a superscript ‘c’ in Table 2 belong to this situation.

**Situation 4:** If the subsystem $\{1, 2, \ldots, i-1\}$ is in DG state ($g_S(i-1) = DG$) and component $i$ makes the SRS in DG state ($\phi(S_i(x_i)) = DG$), then the combined subsystem $\{1, 2, \ldots, i\}$ is in DG, FD or FS state ($g_S(i) = DG$, FD or FS), and the exact state depends on the structure of SRS. The states with a superscript ‘d’ in Table 2 belong to
this situation.  

**Situation 5:** If the subsystem \( \{1, 2, \ldots, i-1\} \) is in FD state \( (g_s(i-1) = FD) \) and component \( i \) makes the SRS in DG state \( (\phi(S_i(x_i)) = DG) \), then the subsystem \( \{1, 2, \ldots, i\} \) is in FD or FS state \( (g_s(i) = FD \text{ or } FS) \), and the exact state also depends on the structure of SRS. The states with a superscript ‘e’ in Table 2 belong to this situation.

From the analysis above, the working state of SRS cannot be determined uniquely without the structure information of SRS for a specific application (situations 4 and 5). Thus, for determining the structure function \( \phi \) uniquely, more information about the SRS is needed.

In the numerical realization of the composition operators, we encode the states FP, DG, FD and FS by integers 0, 1, 2 and 3, respectively. Table 2 can be rewritten as Table 3.

<table>
<thead>
<tr>
<th>Subsystem ( g_s(i-1) )</th>
<th>Component ( \phi(S_i(x_i)) )</th>
<th>0 (FP)</th>
<th>1 (DG)</th>
<th>2 (FD)</th>
<th>3 (FS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (FP)</td>
<td>( g_s(i-1) = 0 )</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1 (DG)</td>
<td>( g_s(i-1) = 1 )</td>
<td>1</td>
<td>1/2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2 (FD)</td>
<td>( g_s(i-1) = 2 )</td>
<td>2</td>
<td>2/3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3 (FS)</td>
<td>( g_s(i-1) = 3 )</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Given set \( V = \{0, 1, 2, 3\} \), thus \( \phi(S_i(x_i)) \), \( g_s(i-1) \) \( \in V \). Let \( R(i) \) be the ordered pair \( \{g_s(i-1), \phi(S_i(x_i))\} \), and thus \( R(i) \in V \times V \). Then, based on Table 3 and structure information of SRS the structure function \( \phi \) can be obtained by

\[
\phi(g_s(i-1), \phi(S_i(x_i))) = \begin{cases} 
\max\{g_s(i-1), \phi(S_i(x_i))\}, & \text{for } R(i) \neq \{1,1\} \text{ and } R(i) \neq \{2,1\} \\
\text{1, for } R(i) = \{1,1\}, I_{FD}(i) = \emptyset, I_{FS}(i) = \emptyset \\
\text{2, for } R(i) = \{1,1\} \text{ or } R(i) = \{2,1\}, I_{FD}(i) = \emptyset, I_{FS}(i) = \emptyset \\
\text{3, for } R(i) = \{1,1\} \text{ or } R(i) = \{2,1\}, I_{FS}(i) = \emptyset 
\end{cases}
\]

\( g_s(0) = 0 \)

\( g_s(i) = \phi(g_s(i-1), \phi(S_i(x_i))) \). \( 1 \leq i \leq m \) \hspace{1cm} (19)

\( I_{FD}(i) = \{x \mid x \in MCS_{FD}, x \subseteq subs_{FD}(i)\} \)

\( I_{FS}(i) = \{x \mid x \in MCS_{FS}, x \subseteq subs_{FS}(i)\} \)

where \( subs_{FD}(i) \) and \( subs_{FS}(i) \) denote the combined components \( \{1, 2, \ldots, i\} \) being FD and FS, respectively. \( I_{FD}(i) \) denotes the collection of sets which are included in \( subs_{FD}(i) \) and belong to \( MCS_{FD} \), and \( I_{FS}(i) \) is the collection of sets which are included in \( subs_{FS}(i) \) and belong to \( MCS_{FS} \). For example, a 2oo3 system with components 1, 2 and 3 has \( MCS_{FD} \), \( MCS_{FS} = \{\{1,2\}, \{2,3\}, \{1,3\}\} \). When we consider subsystem \( \{1, 2\} \), if components 1 and 2 are in FD state, then \( subs_{FD}(2) = \{1, 2\} \), \( subs_{FS}(2) = \emptyset \), \( I_{FD}(2) = \{\{1,2\}\} \) and \( I_{FS} = \emptyset \).

In Eq. (19), the first row \( \phi(g_s(i-1), \phi(S_i(x_i))) = \max\{g_s(i-1), \phi(S_i(x_i))\} \) corresponds to the states enclosed by the dotted line in Table 3. Also, the equation \( \phi(g_s(i-1), \phi(S_i(x_i))) = 1 \), 2 or 3 distinguishes the
states in bold font in Table 3 according to the structure of SRS (the minimal cut sets of SRS for FD and FS). The recursive algorithm can be derived by Eqs. (18)(19).

Furthermore, given the state distribution of component \( i \) according to the \( z \)-transform:

\[
u_i(z,t) = p_{OP}^i(t)z^{\delta(S(OP))} + p_{FS}^i(t)z^{\delta(S(FS))} + p_{FD}^i(t)z^{\delta(S(FD))},
\]

the universal generating function of SRS can be obtained recursively as

\[
U(z,t) = \otimes \left[ u_1(z,t), \ldots, u_k(z,t) \right] \\
= \otimes \left[ \otimes \left[ \otimes \left[ u_1(z,t), u_2(z,t) \right], \ldots, u_k(z,t) \right], \ldots, u_k(z,t) \right] \\
= P_{FS}(t)z^0 + P_{DG}(t)z^1 + P_{FD}(t)z^2 + P_{PS}(t)z^3
\]

Finally, the probability density function of each state of SRS is

\[
f_x(t) = \frac{dP_x(t)}{dt}, \quad x = FP, DG, FD, FS
\]

3. Maintenance model and policy

3.1 System cost

In light of SRS degradation and failure, a schedule point \( A \) is proposed with the objective of minimizing expected system cost per unit time. System cost considered in this paper comprises cost due to system idleness, replacement of the SRS, and penalty of FS and FD. System idleness arises when the SRS needs to be replaced or PS to be resumed. Let \( C_i(t) \) denote the cost of system idleness that takes a linear form in idle time \( t \):

\[C_i(t) = \beta \cdot t, \beta, t > 0. \]

\( C_k \) denotes the cost due to replacement of SRS. The penalty costs \( C_{pFS} \) and \( C_{pFD} \) are the costs incurred with FS and FD, respectively. Usually \( C_{pFS} << C_{pFD} \).

3.2 Maintenance policy

The replacement of SRS will take place at the schedule point \( A \) if the time to certain failures \( T_F \geq A \) or at the time \( T_F \) if \( T_F < A \). The time between two replacements of SRS is considered as an operating cycle. According to the failure effect and interaction between SRS and PS, the failures types of the entire system can be classified into four categories as shown in Fig. 3.

**Case 1:** FS occurs at certain time \( T_F \in (0 A) \). Replacement of SRS will take place at \( T_F \) with duration \( a_1 \), and the cost of this case is \( SC_1(A) = C_i(a_1) + C_{pFS} + C_R \).

**Case 2:** A dangerous event of PS occurs after FD at \( T_F \in (0 A) \). Replacement of SRS will take place at \( T_F \) with duration \( a_2 \), and the cost of this case is \( SC_2(A) = C_i(a_2) + C_{pFD} + C_R \).

**Case 3:** Case 1 or case 2 does not happen between \( (0 A) \). Replacement of SRS will take place at the schedule point \( A \) with duration \( a_3 \), and the cost of this case is \( SC_3(A) = C_i(a_3) + C_R \).

**Case 4:** A dangerous event of PS occurs at \( T_F \in (0 A) \) and the SRS can reduce the risk appropriately. As no FS and FD occur, replacement of SRS is not needed. After system restart, the SRS will continue to work until the occurrence of one of the four cases (1, 2, 3 or 4). However, case 4 is not independent, and it has been included in case 1, case 2 and case 3. Nevertheless, case 4 is needed in cost calculation. Assume that the duration of
restarting the entire system is \( a_4 \) and the cost of this case is \( SC_4(A) = C_s(a_4) + C_p \).

Usually, the system idle time satisfies \( a_2 > a_1 > a_3 \) and \( a_1 > a_4 \) according to the types of failure events.

Fig. 3. Four cases of system failures types

Let \( T(A) \) be the system operating cycle time. The expected system operating cycle time \( E[T(A)] \) can be obtained as

\[
E[T(A)] = \int_0^A (t + a_1) f_{FS}(t) dt + \int_0^A (t + a_2) f_{CH}(\tau) d\tau f_{FD}(t) dt + (A + a_3) \left( 1 - \int_0^A f_{FS}(t) dt - \int_0^A f_{CH}(\tau) d\tau f_{FD}(t) dt \right)
\]

(23)

As shown in Eq. (23), if SRS is an ideal system with no failures (FS and FD) in its life cycle, then replacement will take place at the schedule point \( A \). The expected system operating cycle time can be simplified to \( A + a_3 \). In addition, the expected system cost per operating cycle is

\[
E[SC(A)] = \int_0^A SC_1(A) f_{FS}(t) dt + \int_0^A SC_2(A) f_{CH}(\tau) d\tau f_{FD}(t) dt + SC_3(A) \left( 1 - \int_0^A f_{FS}(t) dt - \int_0^A f_{CH}(\tau) d\tau f_{FD}(t) dt \right) + SC_4(A) \left( 1 - \int_0^A f_{FS}(t) dt - \int_0^A f_{CH}(\tau) d\tau f_{FD}(t) dt \right) \int_0^A f_{CH}(t) dt + \left( 1 - \int_0^A f_{FS}(t) dt \right) \int_0^A f_{CH}(\tau) d\tau f_{FD}(t) dt + \int_0^A \left( f_{FD}(t) + f_{FS}(t) \right) \int_0^A f_{CH}(\tau) d\tau dt
\]

(24)

Similarly, if no failure (FS or FD) occurs in SRS’s life cycle and replacement takes place at the schedule point \( A \), then the expected system cost per operating cycle can be simplified to \( SC_1(A) + SC_2(A) \int_0^A f_{CH}(x) dx \), which only depends on the time to replacement schedule \( A \) and the risk density \( f_{CH}(x) \) of PS. Therefore, the expected system cost per unit time is

\[
\chi(A) = \frac{E[SC(A)]}{E[T(A)]}
\]

(25)

Furthermore, the SRS is required to satisfy a certain SIL\(_x\), and \( p_{SIL_x} \) is the upper bound under SIL\(_x\) (\( x=1, 2, 3, 4 \)).
and given in Table 1. Hence the $PFD_{avg}$ should not be greater than the upper bound $p_{SILx}$. The objective of the model is to minimize the expected system cost per unit time

$$\min_A \mathcal{C}(A)$$

under the constraint of required safety

$$PFD_{avg}(A) = \frac{\int_0^{\mathcal{E}[T(A)]} \int_0^p f_{PD}(t) dt dx}{\mathcal{E}[T(A)]} \leq p_{SILx}, \quad x = 1, 2, 3, 4$$

This optimization problem can be solved by a sequential quadratic programming available in many computation software packages. In this paper, the Matlab Optimization Toolbox is used to produce the results in the following example.

The major computational cost of the proposed model is to obtain the entire SRS states distribution (FP, DG, FD and FS) by using UGF. However, the system structure $\phi$ of a general SRS cannot simply be derived from the combination of subsystems and components in some cases. To solve this problem, MCSs are used to distinguish the states of SRS. Thus, the computational cost mainly depends on the complexity of solving the MCSs of SRS, a problem that has been extensively studied in the literature [27-29]. Even in 1999, a system with 82 components, and as a result 12720 MCSs, can be solved within three minutes [32]. The state-of-the-art commercial software tools, such as RAM commander, CAFTA and BlockSim [30] are extremely powerful and capable of handling much larger problems. Furthermore, if the system is so complex as to exceed the capacity of these tools, the method of system modularization and approximation can still be used [33].

4 An illustrative example

Consider a high integrity pressure protection system that composes of one safety valve (SV) and two pressure sensors PS1 and PS2 as shown in Fig. 4 [5]. The gas flows to a pressure vessel (PV) through SV. If one or both pressure sensors issue a signal that the pressure is too high, the SV will switch off. Each component (PS1, PS2 or PV) can experience FS and FD, and if any of the components suffer FS, the SV will be closed and the entire system will shutdown. For illustration purpose, the age-related transition intensities as a function of time $t$ are given in Table 4. The PS’s time to risk is assumed to be an exponentially distributed random variable with pdf $f_{CH}(t)$ and time-varying failure rate $\lambda_{CH}(t) = 0.2 + 0.3t$. 

![Diagram](image.png)
The process of obtaining the distribution can be formulated as follows:

For component \( i \) (\( i = 1, 2 \)),

\[
u_i(z, t) = p_{op}^i(t)z^0 + p_{FD}^i(t)z^1 + p_{FS}^i(t)z^3
\]  

(28)

For component 3,

\[
u_3(z, t) = p_{op}^3(t)z^0 + p_{FD}^3(t)z^2 + p_{FS}^3(t)z^3
\]  

(29)

The process of obtaining the SRS distribution using the recursive approach is as follows:

\[
U(z, t) = u(z, t) \cup u(z, t)
\]

\[
= p_{op}^1(t)p_{op}^2(t)z^0 + \left( p_{op}^1(t)p_{FD}^2(t) + p_{FD}^2(t)p_{op}^1(t) \right)z^1 + \left( p_{op}^1(t)p_{FS}^2(t) + p_{FS}^2(t)p_{op}^1(t) \right)z^3
\]

\[
+ \left( p_{op}^2(t)p_{FD}^2(t) + p_{FD}^2(t)p_{op}^2(t) \right)z^2 + \left( p_{op}^2(t)p_{FS}^2(t) + p_{FS}^2(t)p_{op}^2(t) \right)z^3
\]

(30)

\[
U(z, t) = u(z, t) \cup u(z, t)
\]

\[
= p_{op}^1(t)p_{op}^2(t)p_{op}^3(t)z^0 + \left( p_{op}^1(t)p_{FD}^2(t)p_{op}^3(t) + p_{FD}^2(t)p_{op}^1(t)p_{op}^3(t) + p_{op}^3(t)p_{FD}^2(t)p_{op}^1(t) \right)z^1
\]

\[
+ \left( p_{op}^1(t)p_{FD}^2(t)p_{op}^3(t) + p_{op}^3(t)p_{FD}^2(t)p_{op}^1(t) \right)z^2 + \left( p_{op}^2(t)p_{FD}^2(t)p_{op}^3(t) + p_{op}^3(t)p_{FD}^2(t)p_{op}^2(t) \right)z^3
\]

\[
+ \left( p_{op}^3(t)p_{FD}^2(t)p_{op}^3(t) \right)z^3 + (1 - \eta)z^3
\]

\[
= p_{FP}(t)z^0 + P_{op}(t)z^1 + P_{FD}(t)z^2 + P_{FS}(t)z^3
\]

where \( \eta = P_{FP}(t) + P_{op}(t) + P_{FD}(t) \).

Thus, the state distributions of SRS are

\[
\begin{align*}
P_{op}(t) &= p_{op}^i(t)p_{op}^i(t)p_{op}^i(t) \\
P_{op}(t) &= p_{op}^3(t)p_{op}^3(t)p_{op}^3(t) \\
P_{FD}(t) &= p_{op}^i(t)p_{FD}^i(t)p_{op}^i(t) + p_{FD}^i(t)p_{op}^i(t)p_{op}^i(t) \\
P_{FD}(t) &= p_{op}^i(t)p_{FD}^i(t)p_{op}^i(t) + p_{op}^i(t)p_{FD}^i(t)p_{op}^i(t) \\
P_{FD}(t) &= p_{op}^i(t)p_{FD}^i(t)p_{op}^i(t) + p_{op}^i(t)p_{FD}^i(t)p_{op}^i(t) \\
P_{FS}(t) &= 1 - \eta
\end{align*}
\]  

(32)

Fig. 5 clearly illustrates the time-varying characteristic of the probability functions for FS and FD states.
of the SRS. According to the model presented in section 2, the SRS will eventually fall into either FD or FS state, that is $\lim_{t \to \infty} (P_{FD}(t) + P_{FS}(t)) = 1$. This phenomenon can be clearly verified from Fig. 5.

\[
\lim_{t \to \infty} (P_{FD}(t) + P_{FS}(t)) = 1
\]

Fig. 5 The time-varying probability functions for FD and FS states of the SRS.

4.2 Optimal schedule point for replacement

Assume that this SRS requires SIL1, thus $p_{SIL1} = 10^{-1}$ [7, 8]. The associated parameters are tabulated in Table 5.

<table>
<thead>
<tr>
<th>$p_{SIL1}$</th>
<th>$C_R$</th>
<th>$C_{PFS}$</th>
<th>$C_{PFD}$</th>
<th>$\beta$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-1}$</td>
<td>500</td>
<td>1000</td>
<td>$5.0 \times 10^4$</td>
<td>$5.0 \times 10^5$</td>
<td>0.02</td>
<td>0.2</td>
<td>0.006</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Fig. 6 (a) plots the expected system cost per unit time as a function of schedule point, which shows that the optimal schedule point is $A^* = 1.372$ (years) giving the minimum cost $\chi(A^*) = 6175$. The corresponding average probability of failure on demand, displayed in Fig. 6(b), is $PFD_{avg}(A^*) = 0.0299$. Fig. 6 also indicate that if we only consider the required SIL1, all schedule points before 3.132 years is feasible. Should replacement take place at $A=3.132$ years without the guidance of the proposed model, the expected cost would become 7269 per unit time, which is 18% higher than the optimal policy.

Fig. 6 (a) Expected system cost per unit time as a function of schedule point $A$
(b) Expected system cost per unit time as a function of $PFD_{avg}$
4.3 Impact of parameters

Table 6 The parametric values used in analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta)</td>
<td>(4.0\times10^3, 4.5\times10^3, 5\times10^3, 5.5\times10^3, 6.0\times10^3)</td>
</tr>
<tr>
<td>(C_R)</td>
<td>400, 450, 500, 550, 600</td>
</tr>
<tr>
<td>(C_{pFS})</td>
<td>800, 900, 1000, 1100, 1200</td>
</tr>
<tr>
<td>(C_{pFD})</td>
<td>(4.0\times10^4, 4.5\times10^4, 5.0\times10^4, 5.5\times10^4, 6.0\times10^4)</td>
</tr>
</tbody>
</table>

In this section, we investigate the effect of cost-related parameters on replacement scheduling and the corresponding cost through sensitivity analysis. The impact of individual parameters is quantified by varying that parameter according to Table 6, while keeping the remaining parameters fixed to the nominal value given in Table 5.

Fig. 7 (a) Optimal schedule point, corresponding to scaled cost-related parameters
(b) Expected system cost per unit time, corresponding to scaled cost-related parameters

Fig. 7 illustrates the effects of cost-related parameters on optimal schedule point and expected system cost per unit time. Clearly, a large value of \(\beta\) results in higher system cost per unit time due to idleness, as shown in Fig. 7 (b), and it requires a longer optimal schedule point \(A^*\) to maintain the cost (Fig. 7 (a)). The impact of cost \(C_R\) has the same trend as, but smaller magnitude than, that of \(\beta\). In addition, the effect of cost, \(C_{pFS}\) and \(C_{pFD}\), due to FS and FD, has similar trend: they both have negative impact on optimal schedule point \(A^*\) and positive impact on the expected system cost, though the effect of \(C_{pFD}\) appears to be more prominent. In summary, the analysis reveals that in practice, special attention should be paid to the parameters that have large impact to achieve minimum expected system cost per unit time.

5. Conclusion

This paper presents a replacement model for a non-repairable SRS. The state distribution of SRS are analyzed based on the structure of SRS, its components’ state distribution, the aging behavior of the components, and the relationship between SRS and PS. The state distribution is obtained by using a NHCTMM. The solution of such NHCTMM can be dramatically simplified using the UGF method and the recursive property of the structure function. The interaction between SRS and PS is accounted for when devising the objective function, i.e.,
expected system cost per unit time, to be minimized. An illustrative example is studied to demonstrate that an optimal schedule point $A^*$ can be obtained to minimize the system cost per unit time while meeting the required safety level.

Nevertheless, the proposed model has certain restrictions. First, the SRS is non-repairable, while in reality, it may be repaired and re-used. Considering a repairable SRS will significantly complicate the objective function. In addition, the system idle time $a_i (i=1,2,3,4)$ due to replacing SRS or restoring PS may not be constant, but random variables in practice. Both issues are currently being investigated.

Acknowledgements

This work was partially supported by Ministry of Railways of China (Scientific Research and Development Program No. 2008X003-C), and National Natural Science Foundation of China (Grant No. 60674064).

References


