# A tiered modelling approach for condition based maintenance of industrial assets with load sharing interaction and fault propagation

Zhenglin Liang, Ajith Kumar Parlikad

Cambridge University Engineering Department

(Manuscript prepared for the IMA Journal of Management Mathematics)

This paper considers the problem of condition-based maintenance optimization for complex industrial assets involving load sharing interaction and fault propagation among their components using a two-tiered approach. The upper layer represents the deterioration and failure at the asset level and uses a continuous-time Markov chain to calculate the optimal inspection interval and maintenance threshold. The model is mathematics tractable. We illustrate a numerical example with closed-form solutions for a simple case. The lower layer models the load sharing interaction and fault propagation, and in turn calculates the deterioration rates and failure rates to be used in the lower layer. We derive a partitioning rule to transform the lower layer model into the layer model's format. Hence we are able to reduce the size of state space and circumvent the state space explosion problem.

# *Keywords:* Condition-based maintenance, Stochastic dependence, State space explosion, Markov aggregation

# 1. Introduction

# 1.1. Problem description

In this paper, we focus on optimizing maintenance and inspection strategies for an asset that is composed of multiple components. The components in the asset are classified into critical components and non-critical components. Critical components are those that are indispensable for an asset to perform its designed function, i.e., if a critical component fails, the asset will fail as well. The deterioration of the asset is also directly dependent on the deterioration of the critical components. On the other hand, non-critical components are those whose malfunction has an indirect impact, or in some cases, no impact on the asset. The deteriorations of components are in some cases interdependent, and we consider two types of dependencies between the components: load sharing interaction and fault propagation. Load sharing interaction is when the failure of a critical component raises the load on other components, and in turn increases the deterioration rate of those components. Fault propagation indicates the scenario when a non-critical component is damaged by an exogenous event and becomes 'malfunctioned'. In this scenario, the asset may still function for some time, possibly at a lower performance level. However, due to the malfunctioning of the component, some of the critical components may no longer be working under their rated operational environment. Therefore, malfunctioning of a non-critical component could result in the deterioration of one or more of the critical components, thus causing the asset to deteriorate faster than the normal rate.

For instance, stay cables are some of the most important elements in span bridge. It consists of multiple parallel wires (critical components) and wrapped by anti-corrosive coating (non-critical component). Bending, displacement or breakage of any wires will raise the load on the rest of wires and in turn increase their wear rates. This scenario fits in the concept of load sharing interaction. In addition to the load sharing interaction, it also experiences fault propagation, when the anti-corrosive coating is damaged, penetrated moisture will cause aggressive corrosion on wires and result in the loss of their strength.

## 1.2. Literature review

The deterioration process of an asset is often represented as a series of transitions between consecutive states. For example, the condition of a power transformer can be classified as normal, defective, faulty and failure states (Sokolov, 2000). This staged deterioration process can be adequately described by a state transition process such as Markov Chains. There are a large number of reliability models in this area e.g. (Amari and McLaughlin 2004), (Rao and Naikan 2006) and (Soro, Nourelfath, and Aït-Kadi 2010). Due to the Markovian property, these models have excellent mathematical tractability. In some cases, even closed form solutions for optimal maintenance and inspection strategies can be found based on such models (Chen and Trivedi 2002).

In some assets, deterioration of critical components can be interdependent, which is also termed in the literature as stochastic dependence. Stochastic dependence is where "the state of certain components in an asset influences the lifetime distribution of other components" (Dekker et al., 1997). Scarf and Deara (Scarf and Deara 2003) modeled this type of failure dependence in a two-component system considering the set up cost as economic dependence. Lai (Lai 2009) considered a two-unit parallel system, where the failure interaction is considered in which the failure of unit 1 increases the failure rate of unit 2, while the failure of unit 2 fails unit 1 immediately. Golmakani (Golmakani and Moakedi 2011) modeled the failure interaction in a two-component system where the failure of unit 1 is unrevealed and has no effect on the reliability of unit 2, but reduces the system's operating performance, while the failure of unit 2 is self-announcing and increases the failure rate of unit 1. The similarity of these models is that they express the stochastic dependence as failure rate interaction, which highly relies on the knowledge of failure rate. However, in reality, failure rate is often unobservable. Hence, to gain a better understanding of the underlying asset aging process and scheduling a better preventive maintenance strategy, it is necessary to model the stochastic dependence as deterioration rate interaction (Bian and Gebraeel 2013). Nonetheless, modeling such complex deterioration behavior is problematic, especially, when the number of components is large, when one often encounters the infamous state space explosion problem (Boudali and Dugan, 2006). It means the size of asset state space increases rapidly with increasing number of components (Gemikonakli, Ever, and Kocyigit 2009). This gives rise to significant difficulty in analysing the deterioration of the asset.

Most approaches to alleviate this state space explosion problem are based on aggregation techniques. Lumping of states in Markov chains is normally deemed to be one of the main techniques to mitigate the state space explosion problem. Strong lumpability allows us to aggregate the states while conserving the Markov property (Ledoux, Rubino, and Sericola 1994). A large body of research on lumpability is on discrete time Markov chain, such as (Kemeny and Snell 1976), (Rubino and Sericola 1991) and (Feinberg and Chiu 1987). The mathematical concept of lumpability for continuous time Markov chain was derived by (Burke and Rosenblatt 1958) and (Rosenblatt 1959) from the perspective of Chapman-Kolmogorov equations and transition probability functions respectively. This sets up the theoretic foundation for a multi-level solution (Leutenegger and Horton 1993). However, as showed by (Kemeny and Snell 1976) and (Feinberg and Chiu 1987), lumpability could only implemented on a very restricted class of Markov chain. In this paper, we use the concept of lumpability to investigate the applicability of aggregating component states to obtain a stochastically dependent asset deterioration model.

# 1.3. The multi-layered modeling approach

In order to develop an optimal maintenance and inspection plan for an asset in which fault propagation is prevalent, we propose a mathematical model that consists of two layers. In the upper layer, we represent the degradation process of the asset using a continuous-time Markov chain (CTMC) as two interactive multi-state processes – one indicating the normal deterioration process and the other indicating the accelerated deterioration process. The accelerated deterioration process represents the faster deterioration caused by fault propagation among components that is caused by malfunctioning of certain non-critical components. The calculation of the asset deterioration rate (accelerated and normal) is facilitated by the lower layer, which uses a deterioration rate interaction model that considers individual component deterioration rates to calculate the deterioration rate of asset considering load sharing interaction and fault propagation. The advantage of this modeling approach is threefold:

- The maintenance and inspection strategies has better performance than existing models that do not consider accelerated deterioration caused by fault propagation.
- It has good mathematical tractability. In our illustrative example, a closed-form solution could be attained.
- It mitigates the state space explosion problem by Markov aggregation.

# *1.4.* Structure of the paper

The remainder of the paper is organized as follows. In section 2, we present the upper layer of the approach, where we model the deterioration of an asset with fault propagation and optimise the maintenance and inspection strategies with an aim to maximise asset availability and/or minimise cost. In section 3 we present the lower layer of the approach, where we model deterioration and interaction between components to calculate the deterioration rates to be used in the upper layer. Section 4 concludes the paper.

# 2. Asset maintenance and inspection model: the upper layer

# 2.1. System description

The asset under consideration can exist in one of a number of discrete pre-defined states. The asset is initially in a *perfect state*. As time progresses, the asset can deteriorate to the first deterioration state, or it can go to a failed state upon random and sudden failure called *Poisson failure*. When the asset reaches the first deterioration state, it can again either deteriorate to the second deterioration state, or it can experience a Poisson failure, and so on until the asset deteriorates to an unacceptable condition, in which situation, we say the asset has experienced deterioration failure. If the asset experiences a Poisson failure, a corrective repair action is performed on the asset, which will bring its condition back to the previous state, i.e., if the asset experiences a Poisson failure when it is in the first deterioration state, the corrective repair action will bring it back to the first deterioration state. If the asset experiences a deterioration failure, it is replaced, which restores its condition back to the perfect state. By default, the deterioration of the asset progresses at a normal deterioration rate unless a non-critical component malfunctions, in which case the deterioration will progress at an accelerated deterioration rate. We will describe the effect of component malfunctions and how it affects the deterioration of the asset in section 3 where we discuss the lower layer model. The asset's deterioration rate can change from the normal rate to accelerated rate when it is at any deterioration state. Note that if the asset is undergoing accelerated deterioration, it is more vulnerable to abnormal events resulting in a higher probability of triggering Poisson failure.

The failure of the asset (sudden as well as deterioration failure) is deemed to be self-evident. However, the state of deterioration of the asset is deemed to be unrevealed. The asset needs to be inspected in order to reveal its state of deterioration, and such inspections require the asset to be stopped from performing its function. Depending on the condition of the asset as revealed in the inspection, various types of preventive maintenance can be performed on the asset to reduce the risk of failure. If the asset is in good condition, no preventive action is taken. If the asset is in poor condition, a minimal preventive maintenance will be performed to restore the asset to the previous deterioration state. If the asset is in bad condition, a major preventive maintenance will be performed to bring the asset back to the perfect state. If inspection reveals that the asset is undergoing accelerated deterioration, and that it is in good condition, a corrective replacement of the malfunctioned component is performed. As a result of this corrective action, the rate of deterioration of the asset is brought back to the normal deterioration rate. On the other hand, if inspection reveals that the asset is undergoing accelerated deterioration, and that it is in poor condition, a corrective replacement of the malfunctioned component is performed along with a minor preventive maintenance action, which takes the asset back to its previous deterioration state as well as bringing the rate of deterioration to its normal rate. Further, if the inspection reveals that the asset is undergoing accelerated deterioration, and that it is in bad condition, a corrective replacement of the malfunctioned component along with a major preventive maintenance action will bring the asset back to perfect state.

The objective of the problem here is to decide the optimal time between inspections as well as the maintenance thresholds.

## 2.2. Model formulation

Fig. 1 shows the asset-state transition diagram in the form of a Markov Chain using the following notation.



FIG. 1: UPPER-LAYER MODEL

# State space

We denote the states using  $Y_{i,j}$  with a steady state probability of the asset being in a particular state  $\pi_{i,i}$ .

- States  $Y_{i,j}$ :  $i = 0 \dots k$  represent the states that the asset goes through before deterioration failure; and *j* indicates the events and activities relating to each deterioration state.
  - j = 0: normal deterioration
  - $\circ$  *j* = 1: accelerated deterioration
  - $\circ$  *j* = 2: inspection without further action
  - j = 3: inspection that will lead to a maintenance action
  - $\circ$  *j* = 4: asset under maintenance
  - j = 5: Poisson failure state

For example, state  $Y_{3,0}$  means that the asset is in the third deterioration state and is undergoing normal deterioration.

- $i \leq g$ : good condition, i.e., g is the threshold for minimal preventive maintenance
- $\circ$  g + 1  $\leq$  i  $\leq$  b: poor condition, i.e., b is the threshold for major preventive maintenance
- $\circ$  i > b: bad condition
- State  $Y_{k+1,0}$  is the deterioration failure state.

# **Transition rates**

- $\lambda_{n,i}$ : normal deterioration rate between the i<sup>th</sup> state and the (i+1)<sup>th</sup> state.
- $\lambda_{d,i}^{n,i}$ : accelerated deterioration rate between the i<sup>th</sup> state and the (i+1)<sup>th</sup> state, where  $\lambda_{d,i} > \lambda_{n,i}$ .
- $\lambda_F$ : Poisson failure rate under normal deterioration
- $\lambda_{Fd}$ : Poisson failure rate under accelerated deterioration, where  $\lambda_{Fd} > \lambda_F$
- $\lambda_{f,i}^{-1}$ : mean time between malfunctions of non-critical components at asset state *i*
- $\lambda_{in}^{-1}$ : mean time between two successive inspections
- $\mu_{in}^{-1}$ : mean duration of inspection

- $\mu_c^{-1}$ : mean duration of corrective replacement of the malfunctioned component
- $\mu_F^{-1}$ : mean duration of corrective repair for revealed Poisson failure
- $\mu_m^{-1}$ : mean duration of minimal preventive maintenance
- $\mu_{M}^{-1}$ : mean duration of major preventive maintenance
- $\mu_R^{-1}$ : mean duration of asset replacement

# 2.3. Asset performance measurement

As described before, the asset is available for operation when it is in states  $Y_{i,j}$  where i = 0, 1, ..., k and j = 0, 1. Therefore, under steady-state conditions, the availability of the asset is given by:

$$A_s = \sum_{i=0}^k \pi_{i,0} + \sum_{i=0}^k \pi_{i,i}$$

(1)

The expected cost is given by:

$$C_{s} = C_{p} \left( \sum_{i=0}^{k} \pi_{i,3} + \sum_{i=0}^{g} \pi_{i,2} + \sum_{i=0}^{k} \pi_{i,4} \right) + C_{u} \left( \sum_{i=0}^{k} \pi_{i,5} + \pi_{k+1,0} \right) + C_{in}' \left( \sum_{i=0}^{k} \mu_{in} \pi_{i,3} + \sum_{i=0}^{g} \mu_{in} \pi_{i,2} \right) + C_{c}' \left( \sum_{i=0}^{g} \mu_{c} \pi_{i,4} \right) + C_{m}' \left( \sum_{i=g+1}^{b} \mu_{m} \pi_{i,4} \right) + C_{M}' \left( \sum_{i=b+1}^{k} \mu_{M} \pi_{i,4} \right) + C_{F}' \left( \sum_{i=0}^{k} \mu_{F} \pi_{i,5} \right) + C_{R}' \mu_{R} \pi_{k+1,0}$$
(2)

# where

- $C'_{in}$ : Cost for each inspection
- $C'_c$ : Cost for repair after a component malfunctioned
- $C'_F$ : Cost for each repair after the Poisson failure
- $C'_m$ : Cost for each minimal preventive maintenance
- $C'_M$ : Cost for each major preventive maintenance
- $C'_{R}$ : Cost for each replacement activity
- $C_u$ : Down time cost per unit time of repair after failure
- $C_p$ : Scheduled down time cost per unit time
- $C_s$ : Operational cost of asset per unit time
- 2.4. Closed-form solution for k=1

We now provide a closed-form solution for the asset availability and expected cost for a special case where k=1. Although this might sound trivial, in many industrial situations, assets are often considered to exist in three states (new, degrading, failed). The simplified diagram is illustrated in Fig. 2. We are interested in computing the steady-state values for all asset state probabilities. Hence we have to solve a CTMC with 12 states. This task can be accomplished by solving a linear system of 12 equations in 12 unknowns. By recalling basic structural properties of the chain in Fig. 2, it is easy to demonstrate how closed-form equations for each of the 12 states can be obtained. The linear system of 12 equations in 12 unknowns is shown in the following:

$$\begin{aligned} \pi_{0,1} &= \frac{\lambda_{f,0}}{D} \pi_{0,0} & \pi_{0,2} &= \frac{\lambda_{in}}{\mu_{in}} \pi_{0,0} \\ \pi_{0,3} &= \frac{\lambda_{in} \lambda_{f,0}}{D\mu_{in}} \pi_{0,0} & \pi_{0,4} &= \frac{\lambda_{in} \lambda_{f,0}}{D\mu_c} \pi_{0,0} \\ \pi_{0,5} &= \left(\frac{\lambda_F}{\mu_F} + \frac{\lambda_{f,0} \lambda_{Fd}}{D\mu_F}\right) \pi_{0,0} & \pi_{1,0} &= X \pi_{0,0} \\ \pi_{1,1} &= Y \pi_{0,0} & \pi_{1,3} &= \frac{\lambda_{in} (X+Y)}{\mu_{in}} \pi_{0,0} \\ \pi_{1,4} &= \frac{\lambda_{in} (X+Y)}{\mu_M} \pi_{0,0} & \pi_{1,5} &= \left(\frac{\lambda_F}{\mu_F} X + \frac{\lambda_{Fd}}{\mu_F} Y\right) \pi_{0,0} \end{aligned}$$

$$\pi_{2,0} = \left(\frac{\lambda_{n,1}}{\mu_R}X + \frac{\lambda_{d,1}}{\mu_R}Y\right)\pi_{0,0} \qquad \sum \pi_{i,j} = 1$$

where

$$B = \lambda_{d,1} + \lambda_{in} + \lambda_{Fd}$$

$$D = \lambda_{d,0} + \lambda_{in} + \lambda_{Fd}$$

$$X = \frac{DB\lambda_{n,0} + \lambda_{d,0}\lambda_{f,0}\lambda_{Fd}}{D\{B(\lambda_{n,1} + \lambda_{in} + \lambda_{f,1}) - \lambda_{f,1}\lambda_{Fd}\}}$$

$$Y = \frac{\lambda_{d,0}\lambda_{f,0}(\lambda_{n,1} + \lambda_{in} + \lambda_{f,1}) + D\lambda_{f,1}\lambda_{n,0}}{D\{B(\lambda_{n,1} + \lambda_{in} + \lambda_{f,1}) - \lambda_{Fd}\lambda_{f,1}\}}$$

by solving the above 12 equations, we can get

$$\pi_{0,0} = \left\{1 + \frac{\lambda_{f,0}}{D} \left(1 + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}}{\mu_c} + \frac{\lambda_{Fd}}{\mu_F}\right) + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_F}{\mu_F} + X \left(1 + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}}{\mu_M} + \frac{\lambda_{F}}{\mu_F}\right) + Y \left(1 + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}}{\mu_M} + \frac{\lambda_{d,1}}{\mu_R} + \frac{\lambda_{Fd}}{\mu_R}\right)\right\}^{-1}$$

$$(3)$$



FIG. 2 THREE STATE ASSET MODEL

The asset is operational when it is in states  $Y_{i,0}$  and  $Y_{i,1}$ . According to equation (1) the asset availability is given by:

$$A_{s} = \pi_{0,0} \left( 1 + \frac{\lambda_{f,0}}{D} + X + Y \right)$$
(4)

The overall asset operational cost is given by:

$$C_{s} = \pi_{0,0} \times \begin{cases} C_{p} \left[ \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}\lambda_{f,0}}{D\mu_{in}} + \frac{\lambda_{in}\lambda_{f,0}}{D\mu_{c}} + (X+Y) \left( \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}}{\mu_{M}} \right) \right] + \\ C_{u} \left[ \frac{1}{\mu_{F}} \left\{ \lambda_{F} (1+X) + \lambda_{Fd} \left( 1 + \frac{\lambda_{f,0}}{D} \right) \right\} + \frac{1}{\mu_{R}} (X\lambda_{n,1} + Y\lambda_{d,1}) \right] + \\ C_{in} \left[ \lambda_{in} + \frac{\lambda_{in}\lambda_{f,0}}{D\mu_{in}} + \lambda_{in}(X+Y) \right] + \\ C_{c}'_{c} \left[ \frac{\lambda_{in}\lambda_{f,0}}{D} \right] + \\ C_{M}' \left[ \lambda_{in}(X+Y) \right] + \\ C_{F}' \left[ \lambda_{F} (1+X) + \lambda_{Fd} \left( \frac{\lambda_{f,0}}{D} + Y \right) \right] + \\ C_{F}' \left[ \lambda_{R} (\lambda_{n,1}X + \lambda_{d,1}Y) \right] \end{cases}$$
(5)

When  $\lambda_{f,0}$  and  $\lambda_{f,1}$  are equal to zero, *X* and *Y* reduces to  $\frac{\lambda_{n,0}}{\lambda_{n,1}+\lambda_{in}}$  and 0 respectively. Therefore the availability  $A_s = \left(1 + \frac{\lambda_{n,0}}{\lambda_{n,1}+\lambda_{in}}\right) / \left(1 + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_F}{\mu_F} + \frac{\lambda_{n,0}}{\lambda_{n,1}+\lambda_{in}}\left(1 + \frac{\lambda_{in}}{\mu_{in}} + \frac{\lambda_{in}}{\mu_M} + \frac{\lambda_{n,1}}{\mu_F} + \frac{\lambda_F}{\mu_F}\right)\right)$ , which is identical to the closed form solution in Chen and Trivedi (2002) when g = 0, b = 0, k = 1.

### 2.5. Performance analysis of the model

In this section, we compare the results of our model with that of Amari and McLaughlin (2004) and Chen and Trivedi (2002). This will highlight the effect of incorporating the accelerated deterioration process in the model. We also examine the impact of the failure of critical components and malfunctioning of non-critical deteriorating components on asset availability.

Amari and McLaughlin (2004) provide a way to calculate the optimal maintenance threshold. However, their model optimizes the major preventive maintenance under an availability-oriented strategy. We will now examine the result obtained for the major preventive maintenance threshold for the parameter setting in their paper (see Table 1).

TABLE 1

Parameter	Value	Parameter	Value
k	7	$\lambda_{n,i}$	0.001
$\lambda_{in}$	0.005	$\mu_R$	0.01
$\mu_M$	0.05	$\mu_{in}$	0.5

In this paper, we extend this example by assuming that one of components in the asset has a malfunction rate  $\lambda_{f,i}$ =0.005. If the component malfunctions, the asset will deteriorate 10 times faster than normal condition. If inspection reveals that the component has malfunctioned, a corrective replacement action is taken on the component with a mean duration  $\mu_c^{-1} = 4$ . The simulation result under this condition using equation (4) is illustrated in Fig. 3.



FIG. 3. AVAILABILITY VS. MAJOR MAINTENANCE THRESHOLD b

As we can see, for this modified example, the major maintenance threshold (+) as calculated using Amari's model is no longer the optimum threshold (\*) if the asset undergoes accelerated deterioration due to fault propagation.

Chen and Trivedi (2002) calculated the optimum mean time between two successive inspections (MTBI) in terms of both availability and cost for a case with the parameter settings shown in Table 2.

TABLE 2

k	4	$\lambda_{n,i}$	0.03
g	1	$C_{in}'$	0
b	3	$C'_m$	10
$\mu_{in}$	2	$C'_F$	10
$\mu_m$	0.5	$C_R'$	50
$\mu_M$	0.2	$C'_M$	50
$\mu_R$	0.025	$C_u$	50
$\mu_F$	0.03	$C_p$	10

We assume that the asset has a non-critical component with a malfunction rate  $\lambda_{f,i}$ =0.01 whose malfunction causes the asset to deteriorate 5 times faster than normal. If through inspection the component is found to be malfunctioning, a corrective replacement action is taken on the component with a mean duration of  $\mu_c^{-1} = 1$  and cost  $c'_c = 10$ . The calculated results using equation (4) and equation (5) for this problem are illustrated in Fig. 4 and Fig. 5 respectively.



FIG. 4. AVAILABILITY VS MTBI

It is clear from the figures that by modeling accelerated deterioration, the optimum inspection strategy is different to that determined by current approaches. We also note that if fault propagation exists, the asset will require more frequent inspections regardless of whether optimizing for cost or availability.



FIG. 5. OPERATING COST VS. MTBI

We will now compare the impact of malfunction rate of non-critical components  $(\lambda_{f,i})$  and Poisson failure rates  $(\lambda_F)$  of critical components on the optimal inspection schedule. Fig. 6 shows how asset availability changes with different MTBI when  $\lambda_{f,i}$  varies from 0.01 to 0.1. Fig. 7 indicates how asset availability changes with different MTBI when  $\lambda_F$  varies from 0.01 to 0.1. We can clearly see that malfunctions of non-critical components (resulting in accelerated deterioration) have more impact on asset availability than Poisson failures of critical components. This is because the malfunctions in the non-critical components are hidden and can only be detected by inspection causing the asset to degrade faster than normal to a worse or even faulty state between the two successive inspections. However, for Poisson failures, even though the asset fails abruptly, as it is a self-announced failure (the asset stops working) the failure is corrected immediately by corrective maintenance.



FIG. 6. AVAILABILITY VS. MTBI FOR DIFFERENT COMPONENT MALFUNCTION RATES

Therefore, if we can monitor the condition of the non-critical component and provide a timely repair, it will reduce the impact of fault propagation. In the ideal case, by applying on-line monitoring, we will be able to immediately observe and repair the malfunction of non-critical component in the same manner as self-announced Poisson failures. Hence, we are able to calculate the benefit of implementing on-line monitoring directly by comparing between figure 6 and 7. For example, if the malfunction rate of the non-critical component is 0.1, the optimal availability is 0.61 from figure 6. However, if we could perform on-line monitoring to continuously observe the condition of non-critical components, then it will be equivalent to transforming the unrevealed malfunction rate to the revealed sudden failure rate. Thus the availability can be improved to 0.8. By comparing the diagrams on both sides, we can see that an improvement about 7%-19% on availability can be attained through continuous condition monitoring.



FIG. 7. AVAILABILITY VS. MTBI FOR DIFFERENT POISSON FAILURE RATES

#### 3. Fault propagation and accelerated deterioration model: the lower layer

Information regarding asset deterioration rates (both normal as well as accelerated) is critical for determining the maintenance and inspection strategies in the upper layer. In the problem under consideration, these depend on the condition and deterioration of the components of the asset. The lower layer models the competing and interactive deterioration among components and provides a way to calculate the normal deterioration rate and accelerated deterioration rate for the asset.

# 3.1. System description

We consider an asset that contains m identical critical components and n non-critical components. The condition of the asset is dependent on the condition of the critical components. On the other hand, the condition of non-critical components will not affect the asset condition directly. However, if they malfunction, the deterioration rate of critical components will increase and cause accelerated deterioration of the asset.

Each critical component can exist in one of a number of discrete pre-defined states. Initially, all the components are in perfect state. If all the critical components are in perfect state, the asset is in a perfect state as well. The deterioration rate of a critical component consists of two parts: its inherent deterioration rate and the interactive deterioration rate. The inherent deterioration rate indicates the aging of critical component without any influence from other critical components and interactive deterioration rate caused by the aging of other influencing critical components through load sharing. The critical components deteriorate progressively until they reach a condition that is deemed to be unsuitable for operation, which we call deterioration failure state. In addition, the critical components can fail abruptly, i.e., experience a Poisson failure. The asset will experience a Poisson failure if and only if a critical component experiences a Poisson failure. The asset will reach its deterioration failure state if the sum of conditions of critical components has passed a predefined threshold. A malfunction of non-critical components will affect one critical components and situation, it can also influence other critical components though loading sharing as described above.

# 3.2. Modeling objectives

The primary aim of the lower layer model is to inform the development of the upper layer asset deterioration and maintenance/inspection optimisation model. In particular, they are as follows:

- To identify the deterioration states of the asset, which achieved in section 3.3.1 and 3.3.2
  - To identify the deterioration rates and failure rate for upper layer which are including
    - Normal deterioration rate ( $\lambda_{n,i}$ ), which is calculated in section 3.3.2 and 3.3.3.
    - Accelerated deterioration rate  $(\lambda_{d,i})$  and transition rate between normal deterioration process and accelerated deterioration process  $(\lambda_{f,i})$ , which are expressed in section 3.3.4
    - Sudden failure rate ( $\lambda_F$ ,  $\lambda_{Fd}$ ), which are derived in section 3.3.5.

# 3.3. Model formulation

# Notation of deterioration rate interaction model

*m*: Number of critical components

*N*: Deterioration failure state of critical components

 $\lambda_l(t)$ : Deterioration rate of  $l^{th}$  critical component at time t

r: Inherent deterioration rate of critical components

 $\delta_{l,h}$ : Increased deterioration rate of critical component l due to the deterioration of influencing critical component h

 $\alpha$ : Scaling factor, representing the linear mapping between increased deterioration rate of an influencing critical component and its inherent deterioration state

# Notation of Markov aggregation

A(t): Vector based asset deterioration state

 $S^{(X_l)}$ : State space of  $l^{th}$  critical component deterioration

S<sup>(A)</sup>: State space of vector-valued asset state A

S<sup>(Y)</sup>: State space of aggregated state Y

 $X_l(t)$ : Deterioration state of critical component l at time t

 $Y_{i,0}(t)$ : *i*<sup>th</sup> aggregated state in normal asset deterioration process at time t

 $q_i(t)$ : Transition rate from lumped subset  $Y_{i,0}$  to  $Y_{i+1,0}$  at time t

# Notation for fault propagation and sudden failure

 $\lambda_{f,i}^{z}$ : The malfunction rates of  $z^{th}$  non-critical component

 $\theta$ : Additional effect on deterioration rate caused by aggressive decay triggered by malfunction of non-critical component

 $\lambda_F^c$ : Sudden failure rate of a critical component

 $\rho$ : A multiplicative factor for sudden failure rate of critical components to express the increased vulnerability of critical components when asset is undergoing accelerated deterioration

3.3.1 Model formulation of interdependent deterioration of critical components

First we model the deterioration of  $l^{th}$  critical component as state transition process  $\{X_l(t)\}$ , with a countable state space  $S^{(X_l)} = \{0, 1, ..., N\}$ .  $X_l(t) = 0$  indicates the critical component l is at as good as new state.  $X_l(t) = N$  expresses the critical component is at deterioration failure state.

Inspired by (Bian and Gebraeel 2013), we model the load sharing interaction among critical components as a deterioration rate interaction model. The deterioration rate of  $l^{th}$  critical component  $\lambda_l$  is constructed by two additive parts as a format of

$$\lambda_l(t) = r + \sum_{h \in \{1,2,\dots,m\} \setminus l} \delta_{l,h}(t)$$

Where *r* indicates the inherent deterioration rate of critical components and  $\delta_{l,h}$  denotes the increased deterioration rate on critical component *l* due to the deterioration on influencing critical component *h*. We assume  $\delta_{l,h}(t)$  is a function of condition of influencing component *h* at time *t*.

$$\delta_{l,h}(t) = f(X_h(t))$$

We limit our development to the case where  $\delta_{l,h}(t)$  is a linear function of  $X_h(t)$ , which can be expressed as

$$\delta_{l,h} = \alpha X_h(t)$$

Where  $\alpha$  is a scaling factor, which represents the linear mapping between the increased deterioration rate and the deteriorating condition of influencing critical component *h*. We assume the *m* critical component deteriorate in a competing manner as described by (Finkelstein 2008) with rate  $\lambda_l$ . Without loss of generality, the asset could be described by an irreducible vector-valued CTMC  $\{A(t)\} = \{(X_1(t), X_2(t), \dots, X_m(t))\}$  on the state space S<sup>(A)</sup>. As described in upper layer model, when asset reaches deterioration failure, the asset will be replaced, which means at that time all components will be restored back to as good as new states. The duration of the replacement is exponentially distributed with a mean time  $\mu_R^{-1}$ .

# 3.3.2 State explosion problem and aggregation of continue time Markov Chain

 $S^{(A)}$  would then consist of  $(N + 1)^m$  states, arising from the *m*-fold Cartesian product of the critical component state space. This implies that when the number of critical components and their states are large, we will encounter the infamous state space explosion problem. It is normally deemed to be impossible and mathematically intractable to find the optimal solution on such large state space (Derisavi, Hermanns, and Sanders 2003).

To overcome the state space dilemma, we implement Markov aggregation technique. It has functions to:

- Reduce the size of state space by aggregating the original state space into a smaller and more
- 11

compact state space.

 Conserve the Markov property, so that the aggregated process is stochastically equivalent to the original one.

In probability theory, this method is known as lumpability. In our case, lumpability ensures the aggregated process  $\{Y_{0,0}, Y_{1,0}, ..., Y_{k,0}, Y_{k+1,0}\}$  with state space  $S^{(Y)}$  is a CTMC (note that this state space is the same as the normal deterioration process in the upper layer model). It opens up a successful way to bridge the lower layer and upper layer. Hence, the deterioration rate interaction model (in section 3.3.1) could be transformed to a mathematical tractable CTMC. Nonetheless, most of existing partition rules are designed for Markov chain with specific topological structure, for example (Courtois 1975) provided analysis on nearly decomposable Markov chains which has the characteristic of weakly linked subsystems. To the best of authors' knowledge, there is still no deterministic partitioning rule to lump a vector-valued CTMC with interdependence among transition rates of entries. Here, we derive a theorem to solve the problem with such characteristics.

**Theorem 1:** If a vector-valued CTMC  $\{A(t)\} = \{(X_1(t), X_2(t), ..., X_m(t))\}$  satisfies the following conditions

• For any entries of  $X_l(t)$   $1 \le l \le m$ , it monotonically evolves in state space  $S^{(X_l)} = \{0, 1, ..., N\}$  with a transition rate  $\lambda_l(t)$ 

$$\lambda_l(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(X_l(t + \Delta t) = X_l(t) + 1 | X_l(t) < N)}{\Delta t} = r_l + \sum_{h \in \{1, 2, \dots, m\} \setminus l} \alpha X_h(t)$$

 All entries evolve in a competing manner. With probability 0, the transition of two or more entries happens simultaneously

then  $\{A(t)\}$  is strongly lumpable, with respect to partition

$$L = \left\{ Y_{i,0}(t) = \left\{ A(t) \middle| \sum_{X_l(t) \in A(t)} X_l(t) = i \right\}, 0 \le i \le k+1, i \in \mathbb{N} \right\}$$

and the transition rate between two adjacent lumped subsets,  $Y_{i,0}$  and  $Y_{i+1,0}$  is

$$q_i(t) = \sum_{l=1}^m r_l + \alpha i(m-1)$$

#### Proof:

For a particular  $A(t) = (X_1(t), X_2(t), ..., X_m(t))$ , because all its entries  $X_l(t)$  evolve in a competing pattern with a rate  $\lambda_l(t)$ , then according to (Finkelstein 2008) the rate for any one of entries evolve to its next state is

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}\left(\sum_{X_l(t+\Delta t) \in A(t+\Delta t)} X_l(t+\Delta t) = \sum_{X_l(t) \in A(t)} X_l(t) + 1 \big| A(t)\right)}{\Delta t} = \sum_{l=1}^m \lambda_l(t)$$

By using this equation to express  $q_i(t)$ , we have

$$q_{i}(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(A(t + \Delta t) \in Y_{i+1,0}(t + \Delta t) \middle| A(t) \in Y_{i,0}(t)\right)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w) \mathbb{P}\left(A(t + \Delta t) \in Y_{i+1,0}(t + \Delta t) \middle| A(t) = w\right)}{\Delta t \sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w)}$$

$$= \frac{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w) \sum_{l=1}^{m} \lim_{\Delta t \to 0} \Delta t^{-1} \mathbb{P}(X_l(t + \Delta t) = X_l(t) + 1 | X_l(t) \in w)}{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w)}$$
$$= \frac{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w) \sum_{l=1}^{m} \lambda_l(t)}{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w)}$$

$$= \frac{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w) \left[ \sum_{l=1}^{m} r_l + \alpha \sum_{l=1}^{m} \sum_{\{h \in \{1, 2, \dots, m\}, h \neq l\}} X_h(t) \right]}{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w)}$$
$$= \frac{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w) \left[ \sum_{l=1}^{m} r_l + \alpha \sum_{l=1}^{m} (i - X_l(t)) \right]}{\sum_{w \in Y_{i,0}(t)} \mathbb{P}(A(t) = w)}$$
$$= \sum_{l=1}^{m} r_l + \alpha i(m-1)$$

As we can see, the  $q_i(t)$  only relies on index of subset  $Y_{i,0}$  and independent of time and condition of vector A(t), hence,  $\{A(t)\}$  is lumpable and the lumped process  $\{Y_{i,0}(t)\}$  is a time homogeneous Markovian process, see (Kemeny and Snell 1976).

Q.E.D

(6)

It is obvious that the deterioration rate interaction model in section 3.3.1 is a special case when  $r_l$  are identical. Therefore, according to theorem 1, the model is able to be aggregated into a more compact continuous-time Markovian model, which can be characterized by normal deterioration process in upper layer. Thus the asset normal deterioration rate is identical to the transition rate between lumped subsets and can be calculated by equation (6).

$$\lambda_{n,i} = q_i(t) = mr + (m-1)\alpha i$$

By using this mechanism to bridge the lower layer model and upper layer models, we could successfully reduce the number of states from  $\overline{S^{(A)}}$  to k, which is significantly smaller and irrelevant with number of critical components m, such that we alleviate the state space explosion problem.

# 3.3.3 Numerical example of aggregating deterioration rate interaction model

We give an example to demonstrate and clarify the Theorem above for an asset with m = 2, N = 4. To mirror the upper layer model, we also consider a replacement activity with a mean duration  $\mu_R^{-1}$  that takes place when the sum of conditions of the two critical components goes beyond 3. The complete state transition diagram is illustrated as follows:



FIG. 8. States transition diagram for an asset with m = 2, N = 4

By aggregating the states according to theorem 1, we can achieve an aggregated process  $\{Y_{0,0}(t), Y_{1,0}(t), Y_{2,0}(t), Y_{3,0}(t), Y_{4,0}(t)\}$  as shown in FIG 9.



FIG. 9. Partition and aggregated states transition diagram

The transition rate between each aggregated states can be calculated as:

٦.

$$\lambda_{n,0} \mathbf{1} = Q_{0,1} \mathbf{1} = \begin{bmatrix} r & r \\ 0 & r + \alpha & 0 \\ 0 & r + \alpha & r \end{bmatrix} \mathbf{1} \Rightarrow \lambda_{n,1} = 2r + \alpha$$

$$\lambda_{n,1} \mathbf{1} = Q_{1,2} \mathbf{1} = \begin{bmatrix} r & r + 2\alpha & 0 & 0 \\ 0 & r + \alpha & r + \alpha & 0 \\ 0 & 0 & r + 2\alpha & r \end{bmatrix} \mathbf{1} \Rightarrow \lambda_{n,2} = 2r + 2\alpha$$

$$\lambda_{n,3} \mathbf{1} = Q_{3,4} \mathbf{1} = \begin{bmatrix} r & r + 3\alpha & 0 & 0 & 0 \\ 0 & 0 & r + 2\alpha & r \end{bmatrix} \mathbf{1} \Rightarrow \lambda_{n,2} = 2r + 2\alpha$$

$$\mu_R \mathbf{1} = Q_{4,0} \mathbf{1} = \begin{bmatrix} r & r + 3\alpha & 0 & 0 & 0 \\ 0 & 0 & r + 2\alpha & r + \alpha & 0 \\ 0 & 0 & r & r + 3\alpha & r \end{bmatrix} \mathbf{1} \Rightarrow \lambda_{n,3} = 2r + 3\alpha$$

$$\mu_R \mathbf{1} = Q_{4,0} \mathbf{1} = \begin{bmatrix} \mu_R \\ \mu_R \\ \vdots \\ \mu_R \end{bmatrix} \mathbf{1}$$

This finding corresponds to that by (Ball and Yeo 1993) and can also be verified by the theorem they derived. This case indicates that with the designed partition rule, we are able to transform a deterioration rate interaction model with a tree structure pattern to a more compact structure with a reduced state space size.

3.3.4 Modeling of malfunction of non-critical components and resulting fault propagation

We consider *n* non-critical components with binary states (operational and malfunctioned). A malfunction of non-critical component will affect one critical component in a way of stimulating its aggressive decay, thus, the critical component is more likely to degrade to worse condition state and influencing other critical components through loading sharing interaction. We assume only one of *n* non-critical components may become malfunctioned at a time. The malfunction rates of  $z^{th}$  non-critical components are independent and time invariant as denoted by  $\lambda_{f,i}^z$ . We model the additional effect on deterioration rate caused by aggressive decay by  $\theta$ . The transition rate of asset from normal deterioration to accelerated deterioration is a competing process among all *n* non-critical components, which is expressed as:

$$\lambda_{f,i} = \sum_{z=1}^n \lambda_{f,i}^z$$

It is obvious from equation (6) that when the inherent deterioration rate of any critical component is increased by  $\theta$ , the deterioration rate of aggregated process is also increased by  $\theta$ . This implies that, without loss of generality, the detrimental impact of fault propagation could be modelled by an accelerated deterioration process within aggregated states and the accelerated deterioration can be calculated by

 $\lambda_{d,i} = \lambda_{n,i} + \theta = mr + (m-1)\alpha i + \theta$ 

It is worthwhile to highlight the inequality between  $\lambda_{n,i}$  and  $\lambda_{d,i}$  which indicates the necessity of modelling accelerated deterioration as a separate process from normal deterioration.

## 3.3.5 Modeling of the sudden failure of critical components

Apart from deterioration failure, critical components also subject to sudden failure. Different from deterioration failure, which is caused by the aging and wearing of critical components, sudden failure represent a type of failure which is triggered mostly by external event, such that the rate of sudden failure does not rely on the condition of critical components. We model the sudden failure rate of a critical component as  $\lambda_F^c$ . The asset will fail under the condition of sudden failure of any critical components.

By assuming the sudden failure rates of critical components are independent and competing we have:

 $\lambda_F = m \lambda_F^c$ 

We consider, under the condition of fault propagation, the strength for critical components withstand on external events may reduce to a certain level below rated condition. We model the increase of vulnerability of critical components when asset is undergoing fault propagation as a multiplicative factor  $\rho$ . Thereby, we have:

 $\lambda_{Fd} = \rho m \lambda_F^c$ 

In summary, we have extracted all the information we need from the lower layer. Using this information, the optimal maintenance and inspection strategies can be found directly using equation (1) and (2).

# 4. Conclusion

Motivated by industrial cases such as maintenance of stay cables, we propose a 2-tiered approach to model and optimize the condition based maintenance strategy for an asset with m critical components and n non-critical components that undergoes fault propagation and load sharing between components. In the upper layer, a CTMC with multiple deterioration processes is designed to model normal and accelerated deterioration. Exploiting the power of Markovian property, a closed form solution for this model was derived. In lower layer, our model considers two types of stochastic dependences: load-sharing interaction among critical components and fault propagation caused by malfunction of non-critical component. We first expressed the load sharing interaction as a deterioration rate interaction model. Then we proved that this model can be transformed to a stochastically equivalent aggregated CTMC. This allowed us to link the upper and lower layer models, and by doing so we mitigated the state explosion problem.

Further, we numerically demonstrated the improvement on maintenance strategy that can be obtained by modelling accelerated deterioration caused by fault propagation in the upper layer. Moreover, a sufficient partition rule to aggregate a particular type of vector-valued CTMC is deduced, which can be adopted in certain cases to circumvent the state explosion problem.

# REFERENCES

Amari S. V. and McLaughlin, L. (2004) "Optimal design of a condition-based maintenance model", in *Reliability and Maintainability, 2004 Annual Symposium-RAMS*, pp 528–533.

Ball.Frank, and Geoffrey F. Yeo. 1993. "Lumpability and Marginalisability for Continuous Time Markov Chains." *Journal of Applied Probability*, 518-28.

Bian, Linkan, and Nagi Gebraeel. 2013. "Stochastic Modeling and Real-Time Prognostics for Multi-Component Systems with Degradation-Rate-Interactions." IIE Transactions.

Boudali H. and Dugan, J. B. (2006) "A continuous-time Bayesian network reliability modeling, and analysis framework", *IEEE Transactions on Reliability*, vol 55, no 1, pp 86 – 97.

Burke, C. J., and M. Rosenblatt. 1958. "A Markovian Function of a Markov Chain." The Annals of Mathematical Statistics 29 (4): 1112–22.

Chen D. and Trivedi, K. S. (2002) "Closed-form analytical results for condition-based maintenance", *Reliability Engineering & System Safety*, vol 76, no 1, pp 43–51.

Chryssaphinou, O., Limnios, N. and Malefaki, S. (2011) "Multi-State Reliability Systems Under Discrete Time Semi-Markovian Hypothesis", *IEEE Transactions on Reliability*, vol 60, no 1, pp 80–87.

Courtois, Pierre-Jacques. 1975. "Error Analysis in Nearly-Completely Decomposable Stochastic Systems." Econometrica: Journal of the Econometric Society, 691–709.

Dekker, R., Wildeman, R. E. and van der Duyn Schouten, F. A. (1997) "A review of multi-component maintenance models with economic dependence", *Mathematical Methods of Operations Research*, vol 45, no 3, pp 411–435.

Finkelstein, Maxim. 2008. Failure Rate Modelling for Reliability and Risk. Springer.

Feinberg, Brion N., and Samuel S. Chiu. 1987. "A Method to Calculate Steady-State Distributions of Large Markov Chains by Aggregating States." Operations Research 35 (2): 282–90.

Gemikonakli, Orhan, Enver Ever, and Altan Kocyigit. 2009. "Approximate Solution for Two Stage Open Networks with Markov-Modulated Queues Minimizing the State Space Explosion Problem." Journal of Computational and Applied Mathematics 223 (1): 519–33

Lai M. T. and Chen, Y. C. (2006) "Optimal periodic replacement policy for a two-unit system with failure rate interaction", *The International Journal of Advanced Manufacturing Technology*, vol 29, no 3, pp 367–371.

Ledoux, James, Gerardo Rubino, and Bruno Sericola. 1994. "Exact Aggregation of Absorbing Markov Processes Using Quasi-Stationary Distribution." Journal of Applied Probability 31 (3): 626–34.

Kemeny, John G., and James Laurie Snell. 1976. Finite Markov Chains. Vol. 28. Springer New York.

van Noortwijk, J. M.(2009) "A survey of the application of gamma processes in maintenance", *Reliability Engineering & System Safety*, vol 94, no 1, pp 2–21.