FACULTY OF LAW, ECONOMICS AND FINANCE

Department of Economics and Management
Discussion Paper
2021-01
Logistics

Department of Economics and Management
Luxembourg Centre for Logistics and Supply Chain
University of Luxembourg

A restless bandit approach for capacitated condition based maintenance scheduling

available online : https://wwwfr.uni.lu/recherche/fdef/dem/publications/discussion_papers

Ece Zeliha Demirci, TED University, TUR
Joachim Arts, LCL, University of Luxembourg
Geert-Jan Van Houtum, Eindhoven University of Technologie, NL

January, 2022

For editorial correspondence, please contact: dem@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
6, Rue Richard Coudenhove-Kalergi
L-1359 Luxembourg

A restless bandit approach for capacitated condition based maintenance scheduling

Ece Zeliha Demirci[a], Joachim Arts[b], Geert-Jan Van Houtum[c]

[a] *Department of Industrial Engineering, TED University, Turkey, ece.demirici@tedu.edu.tr*
[b] *Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg, Luxembourg, joachim.arts@uni.lu*
[c] *Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, The Netherlands, g.j.v.houtum@tue.nl*

Abstract

This paper considers the problem of optimally maintaining multiple non-identical machines deteriorating over time. The number of maintenance activities that can be carried out simultaneously is restricted by the number of maintenance workers. The main goal is to propose a heuristic with low complexity that consistently produces solutions close to the optimal strategy for problems of real size. We cast the problem as a restless bandit problem and propose an index based heuristic (Whittle's index policy) which can be computed efficiently. Another goal is to empirically compare the performance of the index heuristic with alternative policies. In addition to achieving superior performance over failure-based and threshold policies, Whittle's policy converges to the optimal solution when the number of machines is moderately high and/or maintenance workload is high.
Keywords: Maintenance, Restless bandit, Whittle's index heuristic

1. Introduction

The machine repairman problem, also called the machine interference problem, is characterized by a collection of M machines subject to random failures and R repairmen ($R \leq M$). Its classical formulation determines the optimal sequence of machines to be assigned to the repair crew (See Haque and Armstrong 2007 for a detailed review of studies in this direction). It has been extensively studied in the literature as this type of problem arises not only in maintenance operations but also in manufacturing, transportation, telecommunication, and computer systems (Desruelle and Steudel, 1996; Bunday et al., 1997; Kryvinska, 2004; Armstrong, 2002; Wang, 1994).

In this paper, we consider a machine repairman problem in which each machine gradually deteriorates over time rather than suddenly breaks down. Each machine eventually fails unless there is a maintenance intervention. The failure of a machine results in high costs related to production losses and delays, safety issues and unplanned intervention on the machine. The deterioration level of a machine can be directly measured by techniques such as vibration analysis or wear monitoring. Thus it is reasonable to schedule maintenance actions based on the degradation statuses of machines. This maintenance scheme falls into the category of condition-based maintenance (CBM). Under such a scheme, maintenance decisions are taken prior to any predicted failures using the information collected via continuous monitoring or inspections. Early work on CBM has shown that it reduces maintenance costs, improves system reliability and reduces the number of failures.

The conditions of the machines can be expressed through states based on the data collected. A machine follows a stochastic degradation process over a finite number of discrete states starting from the as good as new state and ending in the failed state. During maintenance, necessary overhaul, replacement, and repair operations are carried out to bring the machine into the 'as good as new' condition. The performance of the machines is assessed by output quality and it is a non-increasing function of the accumulated degradation. This setup is especially relevant for high-precision machining and heavy machine tooling that is utilized in many sectors such as aerospace, electronics, defense, and medical technology (Akcay et al., 2021). Degradation increases revenue losses due to decreasing output quality and also the likelihood of failure. The maintenance interventions mitigate the risk of costly machine breakdown as well as escalating revenue losses due to machine deterioration. When a factory has several machines but there is only a limited number of repairmen, it is important to effectively utilize the repairmen. Hence, the problem is to determine which machines to maintain at each decision point to ensure a high performance level at minimum cost.

A primary approach to exploit maintenance optimization and resource allocation problems is to formulate them as a Markov Decision Process (MDP). A MDP formulation can be solved exactly through standard techniques such as value iteration or policy improvement (Puterman, 2014). However, these algorithms are computationally intractable for large-scale real life problems. Naturally, the development of heuristics that can find near-optimal solutions in a time efficient manner is an area of interest. Our problem relates to the restless bandit problem (RBP) which has been introduced by Whittle (1988). The problem deals with the sequential allocation of resources to a collection of stochastic reward-generating projects. As it cannot be solved analytically except for some toy problems, Whittle (1988) develops an index based heuristic that emerges from a relaxation of the problem in which the resource capacity constraint at any decision time is replaced with its time-averaged version. The index policy has been shown to be asymptotically optimal for RBPs under certain conditions. These conditions are called *indexability property* and the heuristic is referred to as *Whittle's index policy* in literature. The heuristic relies on the computation of indices or scores for all projects and choosing the projects with the highest index values at each decision point. For Whittle's index policy to be applicable to a particular problem, one needs to establish the indexability property which is not trivial in general.

In this paper, we present two analytical models for condition-based maintenance scheduling of deteriorating machines. Both models assume that inter-arrival times of degradation processes and maintenance times are exponentially distributed. Also, the maintenance intervention decisions are given dynamically at any decision epoch. We first develop an average cost MDP formulation of the problem, which enables us to obtain the optimal policy and the corresponding optimal cost for small sized instances. Then, we cast the problem as a RBP. We show that the indexability property holds and the optimal policy for the relaxed RBP is of threshold type. Under a threshold policy, a maintenance action is initiated as soon as the degradation level of the machine exceeds a certain threshold. Next, a closed form expression for the Whittle indices is derived in terms of the problem parameters. We further propose a linear programming formulation that finds a lower bound on the expected cost of the optimal policy. Finally, we carry out a numerical performance evaluation of Whittle's index policy. Small sized settings are considered to perform a comparison with the optimal policy, and modest and large sized systems are employed for comparison with two benchmark policies as well as a lower bound. The first benchmark policy is an obvious one, the failure based policy, under which failed machines are maintained on a FCFS basis. The second benchmark policy is referred to as a naive policy as it determines the threshold wear degree levels of intervention with no consideration of capacity. This policy is applied with the first come first served (FCFS) discipline. The index policy shows superior performance compared to the benchmark policies for all instances. The cost-saving achieved by the proposed policy is more remarkable when the system size is large and the maintenance workload is high. Moreover, the performance is robust with respect to changes in the forms of maintenance cost and revenue loss. The results verify that Whittle's index policy is asymptotically optimal as the number of machines and repairmen grow in fixed proportion, especially when maintenance capacity is heavily utilized.

Our paper makes the following contributions. We propose a well performing and simple index policy to the capacitated condition-based maintenance scheduling problem based on Whittle's index theory. The resulting policy takes into account both the degradation levels and output performance of the machines. Additionally, we develop a mathematical model that finds a lower bound on its performance. Although, Whittle's policy is known to be asymptotically optimal under certain conditions, there is very limited research on how fast it converges to optimality. Our numerical experiments show that its performance converges to the lower bound with the number of machines. This explains why Whittle's policy performs well in practice even when the number of machines is moderate. Moreover, comparisons with two alternative policies show that substantial cost savings can be achieved by Whittle's policy. It is also important to notice that the framework and findings are relevant for other resource allocation problems where the assets are subject to failure and require a repair or replacement process to become operational again.

The remainder of this article is structured as follows. In Section 2, we review the related literature, focusing on condition-based scheduling and RBPs. The details of the MDP and RBP models developed are explained in Sections 3.2 and 3.3, respectively. Section 3.4 presents the linear programming model that finds a lower bound on the expected cost of the optimal policy. Next, in Section 4, we present a numerical study to investigate the performance of the proposed policy. Finally, the conclusions are stated in Section 5.

2. Literature Review

We structure the discussion of related literature as follows. We begin with discussing CBM optimization problems, the methodologies used, and the most relevant articles. Thereafter, we review RBP problems and Whittle's index theory to establish the essentials of our model. Lastly, we focus on studies which cast the CBM scheduling problem as a RBP.

2.1. CBM scheduling

Maintenance optimization problems have been extensively studied in the literature in the past several decades. Cho and Parlar (1991); Wang (2002); Alaswad and Xiang (2017); de Jonge and Scarf (2020) survey and summarize the research and practice in this field for different time periods using different classification schemes. Recent advances in sensor and ICT technology make condition monitoring data collection less costly and facilitate CBM practices and studies. CBM optimization models and the resulting strategies rely on the choice of a stochastic deterioration model. Si et al. (2011) and Ye and Xie (2015) present comprehensive overviews of stochastic process models that capture degradation dynamics of a system. If the system condition is observable, the deterioration process can be modelled either with discrete or continuous states. When discrete state deterioration is assumed, the general approach is to formulate the problem as a MDP or one of its extensions, and then to determine optimal maintenance policies with standard algorithms such as policy and value iteration. Renewal theory and regenerative processes are commonly used to model the settings with a continuous state deterioration. In this case, non-linear optimization techniques are employed to find optimal or near-optimal solutions.

We distinguish the literature on CBM based on whether it models a single component/machine or multi-component/machine system. Within the stream of single machine CBM models, the goal is to determine the threshold level beyond which it is optimal to maintain/replace the machine/component. Kolesar (1966) and Kao (1973) are early studies on a single machine system subject to discrete-state Markovian deterioration. They show that a control limit type policy is optimal, i.e. the machine is maintained whenever its wear status exceeds a certain level. As the concept of CBM has become more established, various extensions and variations of models have been proposed for the single machine setting. One extension is the joint optimization of the maintenance threshold level and a periodic inspection schedule. Grall et al. (2002) illustrate that a multi-level control limit rule performs well in this case. Another extension is the consideration of a limited number of imperfect repairs, for which Kurt and Kharoufeh (2010) prove that the optimal CBM policy exhibits control limit characteristics. Additionally, dynamic deterioration rates are adopted by studies like Fouladirad et al. (2008); Van der Weide et al. (2010); van der Weide and Pandey (2011); Fouladirad and Grall (2014, 2015).

The mathematical modeling and optimization of maintenance policies for multi-component/machine systems are more complex than those for single machine systems due to possible dependencies among machines. If all machines in the system are independent of each other, a CBM policy for the single machine setting is applicable to a multi-machine system. On the other hand, optimal decisions for one machine are not necessarily optimal for a group of machines that exhibit any kind of dependency on each other. Still, several studies base their methodology on analyzing the CBM policy per machine and employing the optimal decision of each in the multi-machine environment (Zhu et al., 2010; Tian and Liao, 2011; Tian et al., 2011).

In the literature, four types of dependencies between machines are identified, which are economic, structural, stochastic, and resource dependencies (Olde Keizer et al., 2017; de Jonge and Scarf, 2020). In our study, we consider resource dependency which applies when multiple machines rely on a limited number of maintenance engineers. Despite its practical relevance, only a small number of studies has addressed resource dependency under a CBM regime. Liu et al. (2014) propose a dynamic CBM policy for a multi-component system maintained by a single worker. The case with multiple maintenance workers is investigated by Marseguerra et al. (2002). They calculate threshold degradation levels beyond which maintenance has to be performed based on a combination of a genetic algorithm and Monte Carlo simulation. Koochaki et al. (2013) compare the performance of CBM and age-based maintenance in the opportunistic maintenance framework for three different maintenance workforce limitation scenarios. The situations taken into account are: without worker constraint, with a single worker, and with multiple external workers subject to a certain response time. Moreover, some other studies explore resource dependency in the context of time-based maintenance with no consideration of condition information, e.g. Armstrong (2002); Camci (2015); López-Santana et al. (2016). None of the aforementioned studies have elaborated on the application of their model to instances of industrial sizes nor on the comparison with the optimal strategy. These gaps in the literature constitute an important area for research, which is addressed by our study.

2.2. The restless bandit problem: description and methodologies

RBP is as an extension of the multi-armed bandit problem (MABP) (Whittle, 1988; Gittins, 1979). In the multi-armed bandit problem (MABP), the decision maker is presented with a set of N bandits and each bandit is endowed with a finite state space. At each discrete time instant, the decision maker needs to select one of the bandits to activate so that the expected total discounted reward will be maximized. Only an active bandit earns reward and changes state. However, in the restless bandit problem the decision maker can activate a number of bandits, and inactive bandits are also allowed to change states and generate rewards (referred to as passive rewards). The restless bandit model has gained attention lately due to its applicability to many real-life problems.

Despite providing a powerful modeling framework, RBPs are PSPACE-Hard (Papadimitriou and Tsitsiklis, 1999), which makes their optimal solutions out of reach. Thus, a relaxed version of the problem is considered in the literature, where the constraint on the maximum number of active bandits at any moment is relaxed to its time average. The relaxation makes the problem analytically amenable as it allows for a decomposition to one problem per bandit. The optimal solution of the relaxed problem is defined in terms of index values per bandit depending on the state and transition rates of the bandit. The index values for the relaxed problem serve as a heuristic for the original problem, called Whittle's index policy, where the bandits with the highest Whittle index values are activated at each decision point. The Whittle index policy has been shown to be asymptotically optimal under certain conditions and performs well in practice. In spite of its practicality, application of the Whittle index involves two difficulties, (i) showing a technical property called indexability; (ii) the calculation of the index function itself. Despite being a challenging property to prove, several problems have passed the indexability test. Some examples are Ansell et al. (2003); Glazebrook et al. (2005, 2009); Archibald et al. (2009); Niño-Mora (2012); Larranaga et al. (2016); Ayesta et al. (2020); Li et al. (2020). Additionally, these studies have demonstrated the power of Whittle's index theory on multiple application areas.

More recently, Larranaga et al. (2016) and Ayesta et al. (2020) have established indexability for a family of problems and also derived closed form Whittle's index expressions. Larranaga et al. (2016) restrict the attention to restless multi-armed bandit problems under the average cost criterion, where each bandit evolves as a birth-and-death process. Birth and death state evolution implies that state transitions are only of two types which are: (i) birth transitions that increase the state of a bandit by one, and (ii) death transitions that decrease the state of a bandit by one. The paper presents a general framework for solving birth and-death restless bandits that provides a sufficient condition for the indexability property as well as a closed-form expression for the Whittle's index in terms of steady state probabilities. Ayesta et al. (2020) extend the study of Larranaga et al. (2016) for the cases where there are no upward jumps under active action and there can be an upward jump of at most one under the passive action. We base our analysis on their results for sufficient conditions of indexability and index function derivation.

2.3. Restless bandit approach to CBM scheduling

Only a few studies have mapped the CBM scheduling of deteriorating machines as a RBP. Glazebrook et al. (2005) is the first one but in a discrete time setting. They obtain the Whittle index function for the discounted cost criterion. They numerically compare Whittle's index policy with the optimal policy using small instances (i.e., $M = 4$ and $R = 2$ and $M = 5$ and $R = 3$). Ruiz-Hernández et al. (2020) apply Whittle's methodology to the case where maintenance interventions might be imperfect. They show practicality and effectiveness of the Whittle heuristic through numerical studies with problem sizes of up to 50 machines and 3 repairmen. Ayesta et al. (2020) use the machine repairmen problem in continuous time as an example to show how to retrieve Whittle's index based on their results. Their model is different from ours in terms of cost structure and considering the chance of experiencing catastrophic breakdowns. Furthermore, they have not focused on the computational implementation of the policy and its performance with respect to alternative policies in different settings.

3. Model

3.1. Problem description

A team of R repairmen is responsible for maintenance of M non-identical deteriorating machines, where $1 \leq R < M$. Each machine runs continuously while being subject to a stochastic degradation process. A machine eventually fails if no preventive maintenance is performed. As the number of repairmen is smaller than the number of machines, all machines cannot be maintained simultaneously. Thus, the decision maker needs to select the machines to be maintained at each decision epoch. The interventions are scheduled via a condition based scheme using the degradation status of the machines. The conditions of the machines are assumed to be continuously monitored.

We describe the degradation process of each machine by discrete state degradation. Specifically, a finite number of states is used to denote the condition of the machine, which starts in the new state and ends in the failed state. After every maintenance action, the machine returns to its "as good as new" condition and then gradually deteriorates to worse states. Machines evolve independently from each other. In addition to providing information about the failure probability, the degradation state also has an effect on the operational performance of the machine. The higher the degradation state, the lower the output quality of a machine. This lower quality is incorporated as revenue loss realized due to operating at a higher state wear. Both revenue loss and maintenance costs are non-decreasing with the state of the machine. Furthermore, there is a higher probability of failing at a higher degradation level for all machines. The maintenance rate is independent of

Table 1: Notation

M	number of machines
R	number of repairmen
B_m	breakdown state of machine m
n_m	degradation level of machine m, $\in \{0, 1, \cdots, B_m\}$
$\lambda_m(n_m)$	rate with which an operating machine m jumps from state n_m to $n_m + 1$
μ_m	maintenance rate of machine m
$Y_m(n_m)$	maintenance cost of machine m at state n_m
$R_m(n_m)$	revenue loss rate of operating machine m at state n_m

the wear level because maintenance operations are standardized such that they are independent of the degradation level of a machine.

The decision maker can initiate the maintenance at any moment. Whenever the machine is under maintenance, it is considered non-operational. Although any number of repairmen might be working at any time, a single repairman can conduct maintenance on one individual machine only.

3.2. MDP formulation

We first use MDP methodology in order to model the problem described above. A summary of notation used by both the MDP and Restless Bandit Problem formulation is presented in Table 3.2.

We denote the degradation state of machine m by n_m, where $m \in \{1, \cdots, M\}$ and $n_m \in \{0, 1, \cdots B_m\}$ with 0 being the as-good-as-new state and $B_m \in \mathbb{N}$ being the state where the revenue generated from the manufactured product is (close to) zero. Then, the system state is given by $\boldsymbol{n} = (n_1, n_2, \cdots, n_M)$. If a repairman is assigned to machine m, it returns to pristine state 0 with exponential repair rate μ_m while incurring a maintenance cost of $Y_m(n_m)$. Otherwise if machine m is unattended, its degradation state transitions from n_m to $n_m + 1$ with an exponential rate of $\lambda_m(n_m)$. In this case, we have a revenue loss rate of $R_m(n_m)$ due to operating at state n_m.

We assume that $\lambda_m(n_m)$ is increasing in n_m for all machines. Then the sum of the transition rates under any state $\boldsymbol{n}$ is bounded from above by $\Delta = \sum_{m=1}^{M} \lambda_m(B_m - 1) + \sum_{m=1}^{M} \mu_m$. Thus, we can formulate the system as a discrete time MDP with a time scale chosen as Δ. Under any state $\boldsymbol{n}$, the action state is defined as $\boldsymbol{a} = (a_1, \cdots, a_M)$, where action $a_m = 1$ indicates that a repairman is assigned to machine m and $a_m = 0$ indicates that machine m continues operation. Let $\mathbb{A}$ denote the set of feasible actions that satisfy the condition $\sum_{i=1}^{M} a_i \leq R$. The transition probability of going from state $\boldsymbol{n}$ to state $\boldsymbol{n'}$ given a feasible action $\boldsymbol{a}$ is denoted by $p(\boldsymbol{n}, \boldsymbol{n'}, \boldsymbol{a})$, where

$$p(\boldsymbol{n}, \boldsymbol{n'}, \boldsymbol{a}) = \begin{cases} \lambda_m(n_m)/\Delta & \text{if } \boldsymbol{n'} = \boldsymbol{n} + e_m \text{ for } m \in \{1, \cdots, M\}, \\ \mu_m/\Delta & \text{if } a_m = 1 \text{ and } \boldsymbol{n'} = \boldsymbol{n}^T e_m \text{ for } m \in \{1, \cdots, M\}, \\ 0 & \text{otherwise,} \end{cases}$$

where e_m is a vector in $\mathbf{R}^M$ with all elements 1 except m^{th} element. When no random event occurs a self transition from state $\boldsymbol{n}$ to itself takes place with transition probability

$$\sum_{m=1}^{M} \mathbb{I}\{a_m = 0\} \frac{\lambda_m(B_m - 1) - \lambda_m(n_m) + \mu_m}{\Delta} + \mathbb{I}\{a_m = 1\} \frac{\lambda_m(B_m - 1)}{\Delta},$$

where I is the indicator function. Let $C^{\boldsymbol{a}}(\boldsymbol{n})$ denote the cost for choosing action $\boldsymbol{a}$ at state $\mathbf{n}$,

$$C^{\boldsymbol{a}}(\boldsymbol{n}) = \sum_{m=1}^{M} \left[\mathbb{I}\{a_m = 0\} R_m(n_m) + \mathbb{I}\{a_m = 1\}(R_m(B_m) + \mu_m Y_m(n_m))\right]$$

The Bellman optimality equations of the MDP are given by

$$\gamma + V(\boldsymbol{n}) = \min_{\boldsymbol{a} \in \mathbb{A}} \left[\sum_{\boldsymbol{n'} \in \boldsymbol{S}} p(\boldsymbol{n'} | \boldsymbol{n}, \boldsymbol{a})(C^{\boldsymbol{a}}(\boldsymbol{n}) + V(\boldsymbol{n'})) \right], \quad \forall \boldsymbol{n} \tag{1}$$

where $V(\boldsymbol{n})$ is the value function representing the relative cost of starting at state $\boldsymbol{n}$, and γ is the optimal cost rate.

The MDP formulation given in (1) can be solved by implementing one of the well-known approaches such as value iteration or policy improvement. However, the computational complexity grows with the number of machines and repairmen. Hence, we employ this model only to generate optimal solutions to problems of small sizes in Section 4.2.

3.3. RBP formulation

Independent evolution of the machines allows us to cast the problem as a RBP and utilize Whittle's index theory to obtain a well-performing heuristic. To formulate the problem as a RBP, we represent each individual machine with a bandit, where the state of bandit m, n_m, is the degradation level. Hence, activation of bandit m corresponds to performing maintenance on machine m (i.e., $a_m = 1$ if bandit m is activated and $a_m = 0$ otherwise). The transition dynamics of a bandit are dependent on the action chosen, but are independent of the other bandits. When bandit m is in state n_m, it makes a transition to either state 0 or $n_m + 1$ after an exponentially distributed amount of time. Let $\tau_m^a(i,j)$ represent the transition rate from state i to j for bandit m under action a, then the transition rate function can be expressed as

$$\tau_m^a(i,j) = \begin{cases} \mu_m & \text{if } a_m = 1,\ j = 0,\ 1 \le i \le B_m, \\ \lambda_m(i) & \text{if } a_m = 0,\ j = i+1,\ 0 \le i < B_m, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

For bandit m, the cost per unit of time when in state i under action a, $C_m^a(i)$, can be written as

$$C_m^a(i) = \begin{cases} R_m(B_m) + \mu_m Y_m(i) & \text{if } a_m = 1, \\ R_m(i) & \text{o.w., for } 0 \le\ i \le B_m. \end{cases} \tag{3}$$

Note that the cost of activating a bandit (i.e. maintaining a machine) has two components, which can be explained as follows: (i) $R_m(B_m)$ is the revenue loss due to not operating during maintenance (i.e., equal to the maximum revenue that can be realized), and (ii) $Y_m(i)$ is the material and workforce costs.

The decision maker is interested in finding a policy ϕ, which decides on bandits to activate such that at most R out of M bandits are active at any moment in time. Given the policy ϕ, $X_m^\phi(t)$ stands for the state of bandit m at time t and $X^\phi(t) = (X_1^\phi(t), X_2^\phi(t), \cdots X_M^\phi(t))$. $Z_m(X^\phi(t))$ takes value 1 if bandit m is made active at time t under policy ϕ and 0 otherwise. A policy ϕ is called feasible if the following constraint is satisfied.

$$\sum_{m=1}^{M} Z_m(X^\phi(t)) \le R \quad \forall t \tag{4}$$

The collection of feasible policies satisfying constraint (4) is denoted by U and $U \neq \emptyset$. The original optimization problem can be represented in the following form:

$$\min\ \limsup_{T\to\infty} \sum_{m=1}^{M} \frac{1}{T} E\left[\int_0^T C_m^{Z_m(X^\phi(t))}(X_m^\phi(t))dt\right] \tag{5}$$

$$\text{s.t.}\quad \phi \in U. \tag{6}$$

where the objective function is to minimize long-run average cost. Given the intractability of the problem, we relax it in two steps following the approach in Whittle (1988) to obtain an efficient solution. We first relax the class of policies from those which activate at most R bandits in every decision epoch into those which activate at most R bandits on average. Specifically, we replace constraint (4) with (7).

$$\limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T \sum_{m=1}^{M} Z_m(X^\phi(t))dt\right] \leq R \tag{7}$$

Then, the corresponding relaxed problem is to solve (5) under constraint (7), which turns out to be tractable. The Lagrangian relaxation of the optimization problem can be expressed as the following unconstrained minimization problem:

$$\limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T \sum_{m=1}^{M} C_m^{Z_m(X^\phi(t))}(X_m^\phi(t)) - W\left(R - \sum_{m=1}^{M} Z_m(X^\phi(t))\right) dt\right], \tag{8}$$

where W is the Lagrange multiplier. Problem (8) yields decomposition into M sub-problems, one for each bandit m, that is:

$$\min \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T C_m^{Z_m(X^\phi(t))}(X_m^\phi(t)) - W\left(1 - Z_m(X^\phi(t))\right) dt\right], \tag{9}$$

In other words, optimal policies for M bandits found by (9) operate as a solution to the relaxed problem (8). Due to the unichain nature of the problem, the sub-problem for each bandit becomes equivalent to

$$E[C_m^{Z_m(X^\phi(t))}(X_m^\phi(t))] - WE[1_{Z_m(X^\phi(t))=0}], \tag{10}$$

The optimal solutions of the relaxed problem facilitate the development of the Whittle's heuristic for the original problem in (5). The heuristic relies on establishing a technical property known as indexability. A bandit is called indexable if the number of states in which the optimal action is passive increases with the value of the Lagrange multiplier W. Given that indexability holds, Whittle's index value for bandit m at state n, $W_m(n)$, is defined as the minimum subsidy that makes the passive and active actions equally rewarding at state n for problem (10). Note that indexability is not a trivial property to prove and it is not always possible to obtain a closed-form expression for Whittle's index. In order to analyze our problem structure, we adapt the results found for Markovian restless bandits by Ayesta et al. (2020).

Lemma 1. *An optimal solution of* (10) *is of a 0-1 threshold type with threshold* t_m*. That is, when machine* m *is in a state* $n_m \leq t_m$*, the optimal decision is to continue operating the machine, otherwise the optimal decision is to maintain the machine.*

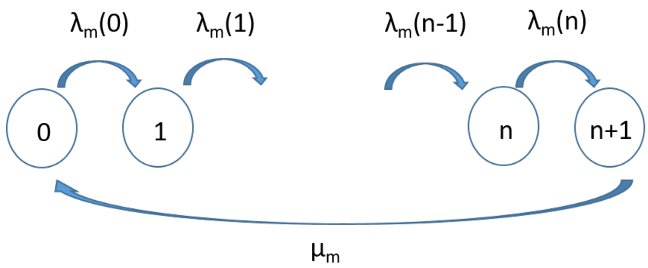

Figure 1: Transition diagram for machine m under the threshold policy n

The lemma follows from Proposition 1 in Ayesta et al. (2020). We can analyze the behavior of each machine under a threshold policy in isolation of the others. Let $\pi_m^n(.)$ denote steady state probabilities for machine m under threshold policy $t_m = n$. The transition diagram corresponding to threshold policy $t_m = n$ for machine m is presented in Figure 1. Action $a = 0$ is taken in states $0, 1, 2, \cdots, n$, whereas action $a = 1$ is taken in states $n+1, n+2, \cdots, B_m$.

The balance equations to find the stationary distribution are given by

$$\lambda_m(0)\pi_m^n(0) = \mu_m \pi_m^n(n+1)$$

$$\lambda_m(i)\pi_m^n(i) = \lambda_m(i-1)\pi_m^n(i-1), \quad \text{for} \quad i = 1, \cdots, n$$

$$\mu_m \pi_m^n(n+1) = \lambda_m(n)\pi^n(n),$$

$$\sum_{i=0}^{n+1} \pi_m^n(i) = 1.$$

From the set of equations given above, we obtain the following expressions:

$$\pi_m^n(i) = \frac{1}{\lambda_m(i)(\sum_{j=0}^n \frac{1}{\lambda_m(j)} + \frac{1}{\mu_m})}, \quad \text{for} \quad i = 0, 1, \cdots n \tag{11}$$

$$\pi_m^n(n+1) = \frac{1}{\mu_m(\sum_{j=0}^n \frac{1}{\lambda_m(j)} + \frac{1}{\mu_m})}, \tag{12}$$

$$\pi_m^n(i) = 0, \quad \text{for} \quad i = n+2, \cdots B_m \tag{13}$$

Lemma 2. *(a) Problem* (10) *is indexable, if* $E[1_{Z_m^n(X_m^n)=0}] = \sum_{j=0}^n \pi_m^n(j)$ *is non-negative and strictly increasing in* n*. (b) Let* $C_m^n(i)$ *denote the cost of bandit* m *at state* i *under threshold policy* $t_m = n$*. Whittle's index for machine* m *at state* n*,* $W_m(n)$ *is given by*

$$\frac{E[C_m^n(i)] - E[C_m^{n-1}(i)]}{\sum_{j=0}^n \pi_m^n(j) - \sum_{j=0}^{n-1} \pi_m^{n-1}(j)}, \tag{14}$$

provided that (14) *is a monotone function in* n*.*

Proof. (a) As $\pi_m^n(i) = 0$ for $i \geq n+2$, $\sum_{j=0}^{n} \pi_m^n(j)$ being strictly increasing in n is equivalent to $\pi_m^n(n+1)$ being strictly decreasing in n. We then obtain

$$\pi_m^n(n+1) - \pi_m^{n-1}(n) = \frac{1}{\mu_m(\sum_{j=0}^{n} \frac{1}{\lambda_m(j)} + \frac{1}{\mu_m})} - \frac{1}{\mu_m(\sum_{j=0}^{n-1} \frac{1}{\lambda_m(j)} + \frac{1}{\mu_m})}\ , \tag{15}$$

which is negative. So, the result follows. (b) It follows from Proposition 3 in Ayesta et al. (2020). □

Although we could not prove that $W_m(n)$ is a monotonic function of n, we observe that this always holds in the numerical experiments.

Note that the expected cost of implementing threshold policy $t_m = n$ for machine m is given by

$$E[C_m^n(i)] = \sum_{j=1}^{n} R_m(j)\pi_m^n(j) + R_m(B_m)\pi_m^n(n+1) + \mu_m Y_m(n+1)\pi_m^n(n+1) \tag{16}$$

Given index values, one can implement Whittle's index policy to determine for which machines to conduct maintenance. At any decision point, the policy decides to intervene up to R machines with the highest non-negative index values at their current states, $W_m(n)$. If all machines have negative index values, then none of them will be worked on.

Several remarks should be made at this point about Whittle's index and its corresponding policy. The index value is quite intuitive and easy to compute. Given that a machine is at state n, the first term in the numerator of equation (14) represents the expected cost of continuing operating and then performing maintenance at state $n+1$; and the second term corresponds to the expected cost of performing the maintenance now. The difference between them gives us the expected cost savings realized if the maintenance is carried out immediately. The denominator of the expression calculates the difference in the fraction of time spent under threshold policy n and $(n-1)$, respectively. Thus, (14) calculates the expected cost saving per unit time due to maintaining the machine immediately. This also brings an intuitive interpretation to the Whittle policy, which is to select machines to work on that would result in higher cost savings. Another advantage of the policy is being flexible. As the index values are found independently for each machine, any change in the problem environment can be handled easily. Examples include the purchase of a new machine, removal of a machine, and changes in the availability of repairmen.

3.4. LP formulation of the threshold policy n

The dimensionality problem of the MDP formulation hinders the sub-optimality assessment of Whittle's policy for relatively large sized instances. Therefore, we now proceed to develop a performance bound that can be used to assess the strength of the policy. We develop a linear programming model of the threshold policy n to find the minimum cost achievable by it.

To formulate the problem, we introduce variable x_m^n to represent the fraction of time that maintenance of machine m is controlled by threshold deterioration state n. Recall that under the n-threshold policy, maintenance is carried on the machine when its wear state exceeds n. We consider that n can take values in $\{0, 1, \cdots, B_m\}$. $n = B_m$ corresponds to the situation when no maintenance is performed on the machine and hence it remains inoperable.

We draw on the stationary distribution and expected costs of threshold policies derived in the preceding section (i.e., equations (11)-(13) and (16)) and translate n-threshold policy into a mathematical program as follows:

$$\text{(LB)}: \quad \min \sum_{m=1}^{M} \sum_{n=0}^{B_m} x_m^n E[C_m^n(i)] \tag{17}$$

$$\text{st} \quad \sum_{n=0}^{B_m} x_m^n = 1 \quad \forall m \in \{1, \cdots, M\} \tag{18}$$

$$\sum_{m=1}^{M} \sum_{n=0}^{B_m - 1} x_m^n \pi_m^n (n+1) \leq R \tag{19}$$

$$0 \leq x_m^n \leq 1 \quad \forall m \in \{1, \cdots, M\}, \quad \forall n \in \{0, 1, \cdots, B_m - 1\}. \tag{20}$$

The formulation determines the threshold deterioration degrees that trigger an intervention decision for each machine together with the percentage of time that threshold levels are used. The objective of the model is to minimize the expected cost of exercising such a policy while ensuring that on the average at most R repairmen are working. It is important to note that this LP formulation is another way of solving Whittle's relaxed version of the problem.

It has been shown that Whittle's policy is optimal in an asymptotic regime in which M and R converge to infinity in fixed proportion (Weber and Weiss, 1990). Accordingly, Whittle's index policy achieves asymptotic optimality for problem environments satisfying the prescribed conditions in Proposition 1. Hence, LB provides a lower bound on the optimal solution of our problem.

4. Numerical experiments

In this section, we study the performance of Whittle's policy through an extensive numerical experiment. The first phase of our analysis centres on small scale problems for which it is possible to investigate the optimality gap of the policy. In the second phase, we continue with larger instances to assess the strength of the policy with problem sizes of practical interest. Given that aim, we benchmark the proposed policy against the lower bound presented in Section 3.4 and two other simpler policies, a failure based one and a naive one. In order to evaluate the average cost of implementing any policy, it is necessary to simulate the system. The set-up of the simulation study is included in Appendix B.

4.1. Instance generation

We consider non-identical machines whose characteristics will be individually generated. The degradation is modelled so as to have an increasing drift toward higher value states in which higher revenue losses and maintenance costs are incurred. We assume that maintenance cost is a linear function of the state, which is given by $Y_m(j) = \alpha_m + b_m j$, where α_m and b_m are non-negative constants. Revenue loss rates are selected so that R_m is 0 for the first two states and a linear function is considered for the later states. Specifically, $R_m(j) = 0$ for $j = 0, 1$, and $R_m(j) = f_m(j-2)$, where f_m is a positive constant. The parameters α_m, b_m, and f_m are drawn from different uniform distributions, for which the details are given in Sections 4.2 and 4.3. The deterioration rates are sampled as follows: $\lambda_m(0) \sim U[0,1]$ and $\lambda_m(j) = \lambda(j-1) + U[0,1]$ for $j = 1, \cdots B_m - 1$. Then, $\lambda_m(j)$ values are scaled so that the mean time to failure is equal to a specified value.

We are interested in the impact of different levels of maintenance workload (i.e. utilization of repairmen) on the performance of Whittle's policy. Higher maintenance workload may result in queues for repairmen and consequently maintenance delays. Thus, it is critical to identify whether the policy is robust enough in a congested setting. As it is analytically challenging to derive an exact expression for the utilization level, we can look at the workload under the failure-based policy. To that aim, we deploy an approximation using closed queueing network theory. To facilitate application of the theory, we assume that the maintenance rate is independent of the machine, $\mu_m = \mu$ for all m. We first formulate a multi-class closed queueing network model of the system operating under a failure based policy and then use the SCAT algorithm to calculate

Table 2: Parameters for the illustrative example

Parameter	Machine 1	Machine 2	Machine 3
α_m	80	50	53
b_m	15	5	5
f_m	100	45	20
ρ	0.8		
Degradation states	$\{0, 1, \cdots, 6\}$		
Mean life time	10		

the utilization for a given maintenance rate (Lavenberg and Reiser, 1980; Neuse and Chandy, 1981). μ is calibrated for each instance to achieve the target utilization level. The details of the queueing model and calibration procedure are presented in Appendix A. Maintenance workload, ρ, is assumed to range over $\{0.8, 0.85, 0.9, 0.95\}$ by controlling the maintenance rates.

As indicated in Section 3.3, the Whittle's index policy is given by (14) for the parameter settings such that (14) is non-decreasing with respect to the wear state. Thus, this precondition is checked for every randomly generated parameter set before running the simulation experiment.

4.2. Small sized systems

Whittle's policy can be easily applied to systems with a large number of machines, however it is not generally tractable to evaluate the MDP formulation in such systems. Therefore, we restrict ourselves to small sized instances to assess the optimality gap of Whittle's policy. The optimal cost is obtained by means of a standard value iteration algorithm with a tolerance limit of 10^{-5} Puterman (2014).

First of all, we investigate the behavior of the Whittle index on a small example with 3 machines and 7 degradation levels. We employ linear configurations for the revenue loss and maintenance cost functions as explained in the previous section. The specific parameter set considered is presented in Table 2. Notice that the deterioration rates are generated as described in Section 4.1. Whittle index values are calculated by equation (14) and reported in Table 3. The index values show that ordering the machines with respect to degradation level does not necessarily return the same sequence as the one found by index values. The reason is that the Whittle index also encodes information about how the degradation processes are likely to evolve in the future, as well as cost differentials between machines.

Thereafter, we explore the optimality gap of the Whittle's policy based on 20 randomly gener-

Table 3: Whittle index values for the illustrative example

State	Machine 1	Machine 2	Machine 3
1	-492.78	-231.40	-105.22
2	-236.58	-131.22	-54.13
3	47.25	-14.12	0.38
4	323.97	103.84	58.07
5	621.32	227.97	114.82
6	10^8	10^8	10^8

Table 4: Input parameters for optimality gap analysis

Input parameters	Values
Degradation states	$\{0, 1, \cdots, 6\}$
Mean life time	10
M, R	$(3, 1)$
ρ	0.8, 0.85, 0.9, 0.95
α_m	$\sim\ U[80, 110]$
b_m	$\sim U[5, 15]$
f_m	$\sim U[40, 60]$

ated instances (i.e., deterioration rates) for 3 machines and 1 repairmen. In the RBP framework, our problem translates into maintaining cost minimizing R machines at any decision point, and this corresponds to preemptive scheduling discipline. Hence, we conduct the optimality gap analysis with simulation results of Whittle's policy under both preemptive and non-preemptive scheduling rules. We use the instance generation process explained in Section 4.1. The results of the data set given in Table 4 is summarized in Table 5. In this table, we report minimum, mean, maximum percentage of optimality gap for both of the scheduling rules.

The majority of instances have a cost rate within 3% of what can be optimally achieved for our problem. The optimality gap decreases with the workload of the repairman, which is a strong endorsement of Whittle's policy for congested systems. Moreover, the performance of the policy under preemptive and non-preemptive scheduling rules are very close to each other. This shows applicability of the policy for practical situations where non-preemptive scheduling is preferable. Thus, we only consider the performance of Whittle's policy under non-preemptive scheduling in the next section.

4.3. Large sized systems

In this section, we subject the proposed policy to numerical investigation for large sized instances, which preclude the use of the value iteration algorithm. Therefore, the performance com-

Table 5: Results of optimality gap analysis

	% GAP					
	Preemptive Scheduling			Non-preemptive Scheduling		
ρ	Min	Avg	Max	Min	Avg	Max
0.8	1.74	2.33	2.96	2.50	2.94	3.23
0.85	1.05	1.94	2.41	1.59	2.24	2.61
0.9	0.52	1.34	1.94	0.83	1.53	2.03
0.95	0.36	1.06	1.59	0.44	1.11	1.58

Table 6: Input parameters for simulation study

Input parameters	Values
Degradation states	$\{0, 1, \cdots, 6\}$
Mean life time	10
M, R	$(10, 1),\ (40, 4),\ (80, 8),\ (160, 16)$
ρ	0.8, 0.85, 0.9, 0.95
α_m	$\sim U[50, 80],\ U[80, 110],\ U[150, 200]$
b_m	$\sim U[5, 15]$
f_m	$\sim U[40, 60],\ U[20, 40]$

parison of interest is conducted between Whittle's policy exercised with non-preemptive scheduling rule, the lower bound on the optimal solution (formulation LB), and two other scheduling policies that emerge in practice. These two policies are:

- Failure based policy: Only the failed machines are maintained under a first come first served (FCFS) discipline. This policy is reasonable if there is no information regarding the conditions of the machines. It acts as a basic benchmark since any decent degradation state dependent policy should perform better.

- Naive policy: The machines exceeding their deterioration thresholds are maintained on a FCFS basis. The threshold level of a machine is determined as the degradation state that minimizes the cost expression (16). Note that the thresholds are found with no consideration of maintenance capacity.

We set up a test bed including instances that are obtained through the parameter values displayed in Table 6. Four possible values for the number of machines and repairmen are explored. While selecting (M, R) combinations, their ratio to each other is held constant to numerically evaluate the convergence rate for the asymptotic optimality of Whittle's policy. All of the machines have 7 degradation states (i.e. $\{0, 1, \cdots, 6\}$) and have a mean life time of 10. To vary the cost

Table 7: Results of gap analysis for $(M, R) = (10, 1)$

			% GAP_Π								
			WI			N			F		
MC	RL	ρ	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
low	high	0.8	21.46	28.22	35.00	23.67	30.99	37.47	40.04	48.16	55.64
low	high	0.85	15.07	17.94	22.18	16.08	19.75	23.46	25.46	29.08	33.32
low	high	0.9	9.05	10.61	13.00	10.32	11.52	13.38	14.33	15.60	17.39
low	high	0.95	3.96	5.64	6.69	5.64	6.10	6.46	6.19	6.63	7.05
med	high	0.8	21.21	27.36	33.73	22.64	29.45	35.40	38.12	45.59	52.44
med	high	0.85	14.79	17.48	21.42	15.62	19.04	22.66	24.52	27.93	31.98
med	high	0.9	8.97	10.46	12.70	10.05	11.28	13.07	13.92	15.16	16.90
med	high	0.95	3.93	5.58	6.62	5.56	6.03	6.39	6.04	6.53	6.96
high	high	0.8	19.50	25.16	30.66	33.49	26.50	31.22	33.49	39.61	31.22
high	high	0.85	13.66	16.42	20.53	22.24	17.59	20.89	22.24	24.97	20.89
high	high	0.9	9.54	10.37	12.09	13.80	11.21	12.07	13.80	14.52	12.07
high	high	0.95	5.29	6.14	6.70	5.97	6.65	7.93	5.97	6.99	7.93
low	low	0.8	21.88	26.74	33.33	23.16	28.72	34.63	38.57	44.59	51.90
low	low	0.85	17.04	19.78	24.91	18.07	21.01	25.48	26.69	30.07	35.01
low	low	0.9	10.12	13.56	16.59	12.79	14.30	16.52	16.68	18.35	20.50
low	low	0.95	5.48	9.12	11.26	7.56	9.42	10.66	8.18	9.97	11.20
med	low	0.8	20.73	25.30	31.03	21.65	26.73	32.16	35.58	41.04	47.51
med	low	0.85	16.24	19.04	24.05	17.17	20.11	24.58	24.93	28.31	33.04
med	low	0.9	9.93	13.19	16.03	12.18	13.87	15.92	15.97	17.57	19.75
med	low	0.95	5.38	8.98	11.17	7.47	9.24	10.34	7.98	9.75	10.91
high	low	0.8	17.72	21.59	26.64	18.86	22.96	26.92	29.02	33.03	38.13
high	low	0.85	14.80	17.05	21.42	16.17	18.13	21.27	21.79	24.16	28.17
high	low	0.9	11.89	12.89	14.96	12.46	13.52	14.31	15.48	16.41	17.52
high	low	0.95	8.04	9.52	10.71	8.04	9.90	11.77	8.38	10.22	11.89

levels of the machines, a_m, b_m, and f_m parameters are sampled as shown in Table 6. For each (M, R) scenario, we follow the data generation procedure explained in Section 4.1.

In total, 480 problem instances are randomly generated and simulated. There are 6 configurations of maintenance cost and revenue loss and 4 different ρ values. This makes 24 scenarios for each pair of (M, R) values and 5 experiments are conducted for each scenario. Those 5 experiments differ in the values of the parameters of maintenance cost and revenue loss, and degradation rates. For every simulated instance and policy, the average cost rate is recorded and subsequently the percentage gap between the lower bound is calculated as a performance metric. Specifically, we are interested in $\%GAP_\Pi = (C_\Pi - C_{LB})/C_{LB} \times 100$, where C_{LB} is the lower bound on the optimal average cost and C_Π, $\Pi = \{WI, F, N\}$ is the average cost under Whittle's index, failure based, and naive policies, respectively. In Tables 7-10, we summarize outcomes by categorizing problem instances with respect to system size, cost configurations (i.e., MC denotes maintenance cost and RL denotes revenue loss) and workload levels. We present the minimum, average, and maximum $\%GAP_\Pi$ for each policy $\Pi = \{WI, F, N\}$.

The results in Tables 7, 8, 9, and 10 confirm that the performance of the failure-based policy

Table 8: Results of gap analysis for $(M, R) = (40, 4)$

			% GAP_{Π}								
			WI			N			F		
MC	RL	ρ	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
low	high	0.8	12.95	14.88	15.72	23.92	25.38	26.13	57.89	60.70	61.92
low	high	0.85	13.25	15.07	15.73	23.86	25.70	26.86	47.73	50.59	51.70
low	high	0.9	10.68	11.99	12.68	19.04	20.67	21.94	30.11	32.46	33.29
low	high	0.95	6.38	7.08	7.86	9.96	10.77	11.29	12.39	13.07	13.68
med	high	0.8	12.34	13.84	14.42	21.32	22.65	24.06	52.62	54.90	56.11
med	high	0.85	12.71	14.21	14.99	21.23	23.03	24.30	44.03	46.49	47.42
med	high	0.9	10.28	11.54	12.07	16.74	17.95	18.49	28.42	30.54	31.37
med	high	0.95	6.27	6.94	7.69	9.24	9.86	10.25	11.89	12.61	13.24
high	high	0.8	11.08	12.21	13.24	41.70	16.03	16.64	41.70	43.12	16.64
high	high	0.85	11.56	12.88	13.75	36.22	16.84	17.68	36.22	37.89	17.68
high	high	0.9	9.32	10.68	11.43	24.21	14.04	14.79	24.21	26.13	14.79
high	high	0.95	5.79	6.18	7.29	10.43	8.31	9.23	10.43	11.12	9.23
low	low	0.8	11.81	13.20	13.79	20.06	21.28	22.08	51.15	52.97	54.21
low	low	0.85	11.97	13.44	13.98	20.14	21.44	22.39	42.40	44.73	45.56
low	low	0.9	10.00	10.90	11.44	16.74	17.34	17.77	27.87	29.58	30.12
low	low	0.95	8.75	10.06	11.17	11.77	13.36	14.02	14.89	16.26	17.12
med	low	0.8	10.91	12.15	12.81	16.73	17.33	17.88	44.44	45.64	46.63
med	low	0.85	11.27	12.61	13.12	16.91	17.69	18.54	37.79	39.45	40.13
med	low	0.9	9.43	10.30	11.02	13.76	14.63	14.97	25.55	26.95	27.43
med	low	0.95	8.31	9.65	10.70	11.00	12.52	13.11	14.06	15.39	16.06
high	low	0.8	8.78	9.43	9.88	11.10	11.83	12.23	31.56	32.12	32.70
high	low	0.85	9.33	10.33	10.84	12.20	13.06	13.42	28.42	29.27	29.72
high	low	0.9	7.93	8.99	9.77	10.44	11.69	12.33	20.16	21.46	22.29
high	low	0.95	7.75	8.45	10.18	9.66	10.44	11.96	12.32	12.99	14.48

is the weakest and Whittle's policy is consistently the strongest. Specifically, the suboptimality of the Whittle's policy is the lowest among all policies for all problem instances which shows that it is robust to changes in cost parameters. Even though it outperforms failure-based and naive policies, Whittle's policy produces rather weak results when the system size is small (i.e. $M = 10, 40$) and workload is relatively low ($\rho = 0.8, 0.85$). However, what makes it standing out is its consistency and robustness in performance relative to the lower bound for relatively larger system sizes that exist in practice. For the scenarios including 160 machines, its overall worst case performance is at most 4.9%, while the value for the other policies are 50.88%. The outperformance of Whittle's policy is due to its ability to dynamically react to deterioration levels of machines while also considering workload. Ignoring both the conditions of the machines and the maintenance capacity while scheduling maintenance can be quite costly, as gaps with the lower bound up to 50.88% are observed under the failure based policy. Although taking maintenance actions based on degradation levels of machines is beneficial, failing to consider maintenance capacity leads to gaps of up to 19.43% under naive policy. Also, Whittle's policy performs better against the lower bound as the system size increases, which is in line with the asymptotic optimality of the policy.

Table 9: Results of gap analysis for $(M, R) = (80, 8)$

			% GAP_{Π}								
			WI			N			F		
MC	RL	ρ	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
low	high	0.8	7.17	7.83	8.29	19.32	19.82	20.78	52.07	53.59	54.45
low	high	0.85	7.60	8.14	8.51	20.57	21.32	22.46	41.26	42.70	43.38
low	high	0.9	7.26	7.89	8.41	19.04	20.05	21.85	31.75	33.19	33.94
low	high	0.95	5.69	6.34	7.19	14.28	15.17	17.02	19.08	20.13	20.79
med	high	0.8	6.69	7.31	7.68	16.83	17.57	18.98	46.99	48.41	49.09
med	high	0.85	7.31	7.84	8.22	17.21	18.33	20.22	38.04	39.34	39.97
med	high	0.9	6.99	7.56	8.04	16.53	17.46	19.47	29.53	30.93	31.59
med	high	0.95	5.13	5.99	6.77	12.28	13.13	14.52	18.08	19.04	19.60
high	high	0.8	6.06	6.44	6.84	36.74	10.72	11.22	36.74	37.70	11.22
high	high	0.85	6.56	6.96	7.39	30.76	11.87	12.52	30.76	31.94	12.52
high	high	0.9	6.23	6.88	7.55	24.54	11.77	12.40	24.54	25.90	12.40
high	high	0.95	5.03	5.51	6.29	15.61	9.73	10.32	15.61	16.43	10.32
low	low	0.8	6.46	7.11	7.53	15.36	16.33	18.20	45.45	47.05	47.93
low	low	0.85	6.72	7.36	7.84	16.00	17.36	19.64	36.54	38.09	38.76
low	low	0.9	6.48	7.27	8.01	15.25	16.37	18.77	28.74	30.15	30.76
low	low	0.95	5.96	7.06	8.54	12.66	13.90	15.74	18.73	19.96	20.94
med	low	0.8	6.00	6.55	6.86	11.51	12.52	13.75	39.17	40.55	41.20
med	low	0.85	6.33	6.89	7.29	12.72	13.50	14.76	32.34	33.59	34.18
med	low	0.9	5.99	6.77	7.55	11.87	12.88	14.13	25.74	27.04	27.69
med	low	0.95	5.62	6.66	7.99	10.31	11.43	12.53	17.27	18.35	19.18
high	low	0.8	4.82	5.20	5.59	7.29	7.73	8.07	27.02	28.06	28.79
high	low	0.85	5.30	5.84	6.21	8.48	9.04	9.50	23.54	24.71	25.48
high	low	0.9	5.21	5.92	6.52	8.74	9.49	10.08	19.52	20.79	21.64
high	low	0.95	4.98	5.65	6.37	8.18	9.01	9.87	13.81	14.67	15.27

Table 10: Results of gap analysis for $(M, R) = (160, 16)$

			% GAP_{Π}								
			WI			N			F		
MC	RL	ρ	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
low	high	0.8	3.80	3.99	4.36	16.83	17.33	17.61	48.33	49.55	50.88
low	high	0.85	4.11	4.42	4.90	18.29	19.03	19.43	37.22	38.08	39.54
low	high	0.9	4.05	4.38	4.78	17.36	18.08	18.66	27.41	28.21	29.42
low	high	0.95	3.82	4.11	4.49	15.29	15.75	16.16	19.36	19.92	20.98
med	high	0.8	3.53	3.75	4.06	13.80	14.35	14.97	43.80	44.77	45.90
med	high	0.85	3.84	4.11	4.44	15.05	15.93	16.86	34.26	35.00	36.12
med	high	0.9	3.83	4.10	4.53	14.45	15.10	15.83	25.58	26.27	27.30
med	high	0.95	3.72	3.97	4.32	12.79	13.38	14.11	18.15	18.66	19.51
high	high	0.8	3.18	3.31	3.47	34.05	7.34	7.53	34.05	34.61	7.53
high	high	0.85	3.48	3.69	4.04	27.80	9.08	9.51	27.80	28.30	9.51
high	high	0.9	3.55	3.69	3.96	21.40	9.65	10.12	21.40	21.87	10.12
high	high	0.95	3.52	3.63	3.92	15.69	9.35	9.74	15.69	16.03	9.74
low	low	0.8	3.45	3.66	3.90	12.73	13.63	14.13	42.69	43.68	44.75
low	low	0.85	3.76	3.91	4.31	13.93	14.86	15.42	33.46	34.18	35.44
low	low	0.9	3.75	4.05	4.39	13.46	14.35	15.10	25.14	25.78	26.96
low	low	0.95	3.87	4.24	4.61	12.42	13.09	13.92	18.31	18.78	19.62
med	low	0.8	3.13	3.33	3.48	9.00	9.43	9.81	36.71	37.56	38.43
med	low	0.85	3.35	3.58	3.92	10.40	10.94	11.39	29.42	30.10	31.19
med	low	0.9	3.59	3.77	4.15	10.39	10.97	11.65	22.57	23.15	24.08
med	low	0.95	3.77	4.00	4.38	9.73	10.41	10.96	16.68	17.19	17.98
high	low	0.8	2.48	2.63	2.77	4.91	5.16	5.37	25.08	25.58	26.05
high	low	0.85	2.90	3.05	3.36	6.52	6.93	7.30	21.38	21.86	22.66
high	low	0.9	3.02	3.19	3.48	7.28	7.82	8.32	17.21	17.63	18.42
high	low	0.95	3.27	3.42	3.78	7.60	8.05	8.54	13.11	13.51	14.19

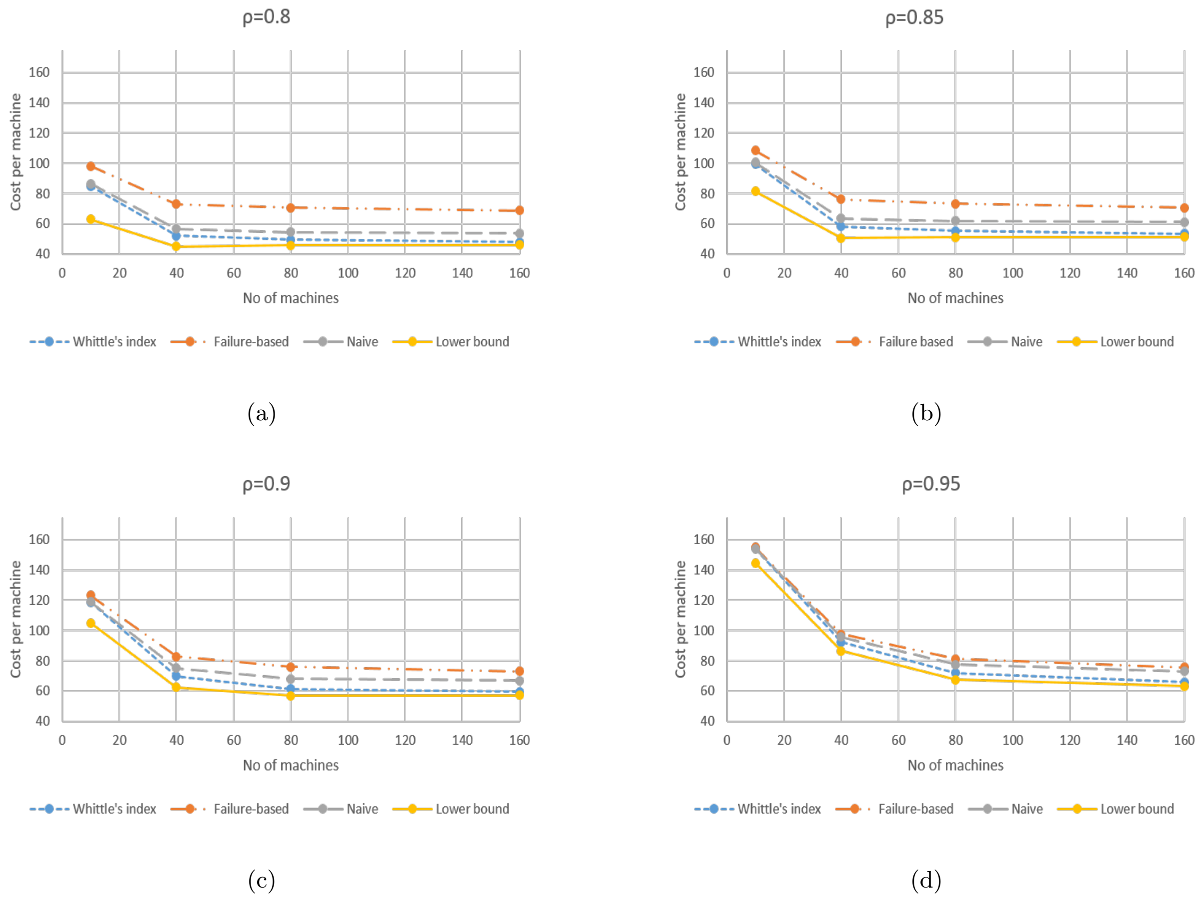

Figure 2: Average cost per machine as a function of M

Following the general discussions of the results, we shift our focus on the asymptotic optimality of Whittle's policy when the number of machines goes to infinity with a fixed proportion to the number of repairmen. This is equivalent to exploring the convergence rate of the policy to the lower bound on the optimal solution as the number of machines grows. For this purpose, we plot the average cost per machine under Whittle's policy and the lower bound as a function of the number of machines at different levels of workload. We also incorporate average cost behavior of benchmark policies to the plots for the completeness of the analysis. Even though, the plots are drawn for a single scenario, one can observe similar behaviors for other instances as well. Figure 2 shows numerical evidence for the fact that the performance of Whittle's policy converges to the lower bound quite fast with the increase in the number of machines. Although the convergence rate is high up to 80 machines, the rate of the increase slows down as the number of machines grows more as expected. The gaps up to 80 machines when $\rho = 0.8$ and 0.85 are lower than the ones when $\rho = 0.9$ and 0.95. For the same instances, we also plot the percentage gap in lower bound for all policies with respect to the number of machines (i.e. see Figure 3). Both figures justify that the performance of the Whittle's policy becomes increasingly promising for systems with heavy maintenance workload and larger number of machines.

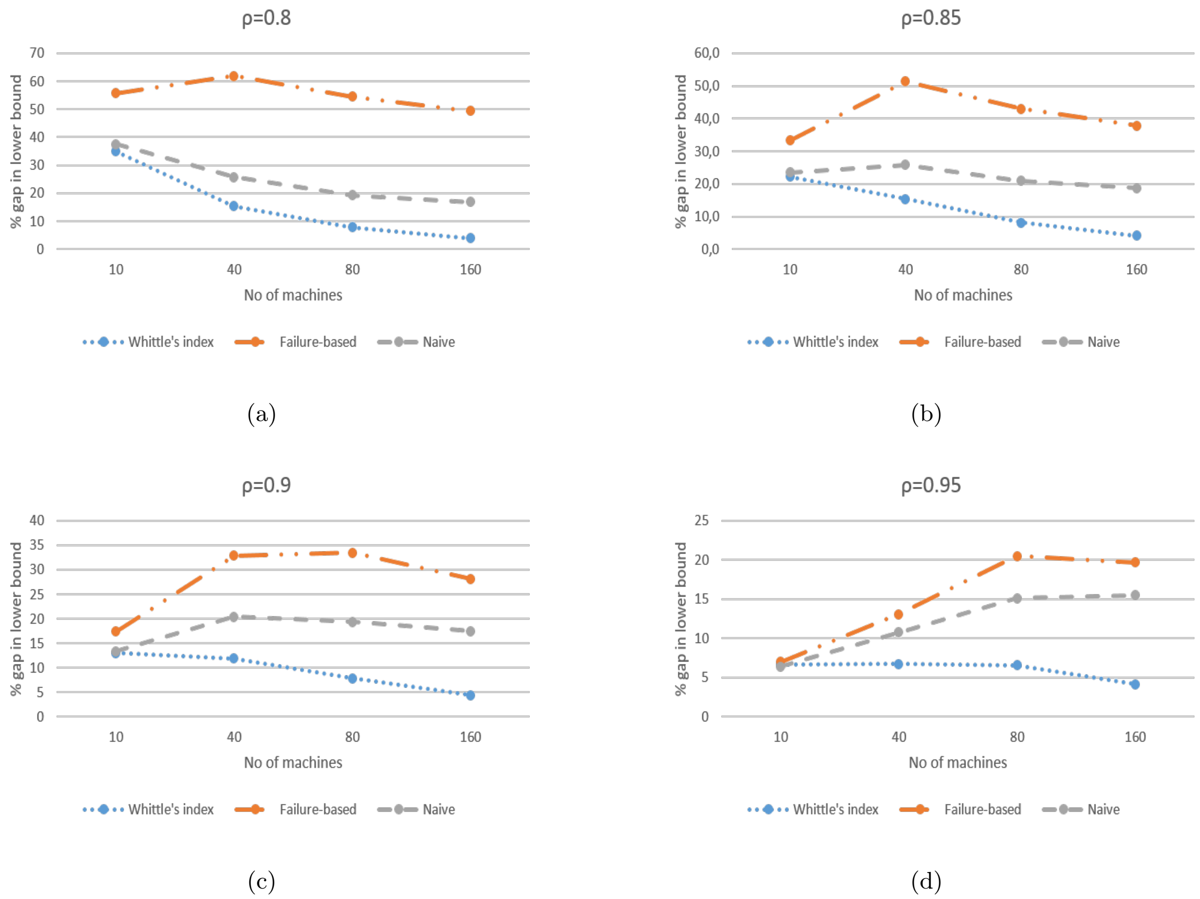

Figure 3: % gap in lower bound as a function of M

5. Conclusion

In this paper, we have considered the maintenance scheduling of deteriorating machines so that a limited number of repairmen is effectively utilized. We have first formulated the problem with a MDP, which is restrictive due to being computationally infeasible for problems of practical size. Thus, we instead used the RBP approach in deriving maintenance policies based on characteristics of the individual machines. We showed that the indexability property holds for our problem and optimal maintenance policy follows a 0-1 threshold structure. We developed an index based heuristic for our problem, which emerges from closed form expressions of the Whittle's index values of each machine. Furthermore, we developed a linear programming model to find a lower bound on the performance of the heuristic. We performed a comprehensive numerical study to show outperformance of the Whittle's heuristic compared to failure-based and naive policies. Furthermore, we have shown that Whittle's policy performs close to the optimal solution when the number of machines is high and/or maintenance workload is high. An interesting extension is to study multiple deteriorating components for each individual machine.

Appendix A. Multi-class closed queueing network model

In this section, we consider our problem under the use of the failure based policy. This scenario is modelled as a multi-class queueing network and the product form analysis of it is adapted to ensure an efficient analysis. We start with the description of the model and proceed with presenting the algorithm to find the mean system throughput.

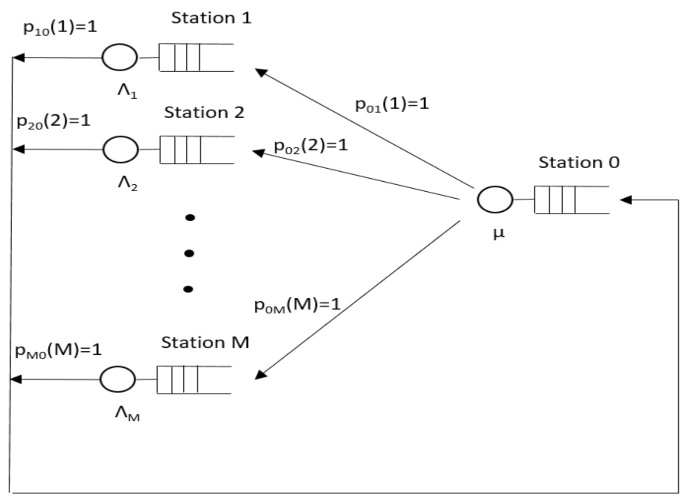

Figure A.4: Closed queueing network diagram with M customer classes and $M+1$ stations

We consider a network consisting of machine bases and a maintenance facility. Figure A.4 provides a pictorial illustration of the network. M non-identical machines are characterized as M machine bases with a single machine. Also, each machine is associated with a separate class of customers to use a multi-class queueing network framework. We represent the maintenance facility as Station 0 and machine base m as Station m, $m = 1, \cdots, M$. The operating time of base m is assumed to be exponentially distributed with mean Λ_m, the mean time to failure for machine m. In the maintenance facility, R identical repairmen are retained to service the failed machines. The service times of each repairmen follow an exponential distribution with rate μ. Once the repair of machine m is completed, it is

Table A.11: Notation used in SCAT algorithm

M	number of machine types
R	number of servers
Λ_m	mean time to failure for machine type m
μ	service rate of each repairmen at Station 0
$\boldsymbol{u}$	state vector, $(\boldsymbol{u_0}, \boldsymbol{u_1}, \cdots, \boldsymbol{u_m})$,
	where $\boldsymbol{u_i} = (u_{i1}, u_{i2}, \cdots, u_{iM})$
u_{im}	number of type m jobs at Station i
$\boldsymbol{u} - \mathbf{1_s}$	state vector when one machine is removed from type s
$e_{im}(\boldsymbol{u})$	relative frequency of visits to Station i by machine type m given state $\boldsymbol{u}$
U_m	number of type m machines (=1)
$pr(j\|\boldsymbol{u})$	queue length distribution of Station 0 given state $\boldsymbol{u}$
$t_{im}(\boldsymbol{u})$	mean time spent by machine type m in station i given state $\boldsymbol{u}$
$k_{im}(\boldsymbol{u})$	the mean number of machine type m at station i given state $\boldsymbol{u}$
$k_i(\boldsymbol{u})$	the mean number of machines at station i given state $\boldsymbol{u}$

directly sent to machine base m. Let $p_{ij}(r)$ express the probability that machine m that departs Station i will next visit Station j. Note that $p_{0m}(m) = 1$ and $p_{m0}(m) = 1 \; \forall m$; and all other transitions have zero probability.

The network is described with state $\boldsymbol{u} = (\boldsymbol{u_0}, \boldsymbol{u_1}, \cdots, \boldsymbol{u_m})$, where $\boldsymbol{u_i} = (u_{i1}, u_{i2}, \cdots, u_{iM})$ and u_{im} is the number of type m machines at Station i for $i = 0, 1, \cdots, M$ and $m = 1, \cdots, M$. Note that $U_m = \sum_{i=0}^{M} u_{im} = 1 \; \forall r$.

Suppose that e_{im} expresses the relative frequency of visits to Station i by machine type m. It can be found by using the following M sets of linear equations:

$$e_{im} = \sum_{j=0}^{M} e_{jm} p_{ji}(m), \quad i \in 0, 1, \cdots, M, \; m \in 1, \cdots, M \tag{A.1}$$

Evaluating (A.1) with the routing information gives us $(i, m \in 1, \cdots, M)$

$$e_{im} = \begin{cases} e_{0m} & \text{if } m = i, \\ 0 & \text{otherwise.} \end{cases} \tag{A.2}$$

We continue our analysis by setting e_{0m} to 1, for $m \in 1, \cdots, M$.

Mean Value Analysis is an exact technique to obtain solutions for the product-form closed queueing networks. However, its computational complexity increases very rapidly with the number of job classes. So, several algorithms have been developed over the years for computing approximate solutions. We adapt an early significant method referred to as SCAT, which allowed approximate analysis of queueing network models with multiple job classes and multiple servers for the first time (Neuse and Chandy, 1981). The method is an iterative procedure to calculate the mean number of jobs, mean residence time and average system throughput.

The notation is summarized in Table A.11 and the general form of the algorithm adapted for our problem is outlined below.

Table A.12: Core Algorithm

Inputs: M, R, U_m, $\boldsymbol{u}$, ω_{im}, e_{im} for $0 \leq i \leq M, 1 \leq m \leq M$
Initialization: $k_{im}(\boldsymbol{u}) = \frac{U_m}{M+1}$ for $0 \leq i \leq M, 1 \leq m \leq M$
$D_{ims} = 0$ for $0 \leq i \leq M, 1 \leq m, s \leq M$
I=1
Step 1: Compute estimates of $k_{im}(\boldsymbol{u} - \mathbf{1_s})$ using equations below for $0 \leq i \leq M, 1 \leq m \leq M$.
$F_{im}(\boldsymbol{u}) = \frac{k_{im}(\boldsymbol{u})}{U_m}$
$k_{im}(\boldsymbol{u} - \mathbf{1_s}) = (\boldsymbol{u} - \mathbf{1_s})_s(F_{im}(\boldsymbol{u}) + D_{ims}(\boldsymbol{u}))$
Compute estimates of $pr(j|\boldsymbol{u})$ for $0 \leq j \leq M$ along with the equations below.
$\text{floor}_{im} = \lfloor k_i(\boldsymbol{u} - \mathbf{1_m}) \rfloor$
$pr(\text{floor}_{im}|\boldsymbol{u} - \mathbf{1_m}) = \text{floor}_{im} + 1 - k_i(\boldsymbol{u} - \mathbf{1_m})$
$pr(\text{floor}_{im} + 1|\boldsymbol{u} - \mathbf{1_m}) = 1 - pr(\text{floor}_{im}|\boldsymbol{u} - \mathbf{1_m})$
$pr(j|\boldsymbol{u} - \mathbf{1_m}) = 0 \; \forall j < \text{floor}_{im}$ and $\forall j > \text{floor}_{im} + 1$
Step 2: Compute new estimates of $k_{im}(\boldsymbol{u})$ using equations below for $0 \leq i \leq M, 1 \leq m \leq M$.
$t_{0m}(\boldsymbol{u}) = \frac{1}{\mu} \sum_{j=1}^{R-1} \left(\frac{1}{j} - \frac{1}{R} \right) jpr(j-1|\boldsymbol{u} - \mathbf{1_m}) + \frac{1}{R\mu}[1 + k_i(\boldsymbol{u} - \mathbf{1_m})]$
$t_{0m}(\boldsymbol{u}) = U_m \frac{e_{im}(\boldsymbol{u}) t_{0m}(\boldsymbol{u})}{\sum_{i=1}^{M} e_{im}(\boldsymbol{u}) t_{0m}(\boldsymbol{u})}$
$t_{0m}(\boldsymbol{u}) = \frac{1}{\Lambda_m}[1 + k_i(\boldsymbol{u} - \mathbf{1_m})]$
Step 3: Termination test.
Define $Ktot = \max_{i,m} \frac{|k_{im}^I - k_{im}^{I-1}|}{U_m}$ for $I \geq 1$. I stands for iteration I estimations.
If $Ktot < (1/(4000 + 16M) <$ stop. Otherwise, go to Step 1.
Outputs: $k_{im}(\boldsymbol{u})$ for $0 \leq i \leq M, 1 \leq m, s \leq M$

Table A.13: SCAT Algorithm

Inputs: M, R, U_m, $\boldsymbol{u}$, ω_{im}, e_{im} for $0 \leq i \leq M, 1 \leq m \leq M$
Step 1: Apply the Core Algorithm for state $\boldsymbol{u}$.
Step 2: Apply the Core Algorithm for each of the states $\boldsymbol{u} - \mathbf{1_x}$ for $1 \leq x \leq M$.
Step 3: Compute estimates of $F_{im}(\boldsymbol{u})$, $F_{im}(\boldsymbol{u} - \mathbf{1_x})$, $D_{ixs}(\boldsymbol{u})$ for $0 \leq i \leq M$, $1 \leq x, s \leq M$ by
$F_{im}(\boldsymbol{u}) = \frac{k_{im}(\boldsymbol{u})}{U_m}$
$D_{ixs}(\boldsymbol{u}) = F_{im}(\boldsymbol{u} - \mathbf{1_s}) - F_{im}(\boldsymbol{u})$
Step 4: Apply the Core Algorithm for state $\boldsymbol{u}$.
For the $k_i m(\boldsymbol{u})$ inputs use the values obtained from the Core Algorithm in SCAT Step 1.
For the $D_{ims}(\boldsymbol{u})$ inputs use the values computed in Step 3.
Outputs: $k_{im}(\boldsymbol{u})$, $t_{im}(\boldsymbol{u})$ for $0 \leq i \leq M, 1 \leq m, s \leq M$

Following, we obtain mean number of machine m maintained at Station 0 with the formula $Y_m = k_{0m}/t_{0m}$, which allows us to calculate utilization of the Station 0, $\rho = \sum_{m=1}^{M} \frac{Y_m}{R\mu}$.

Lastly, we present steps of calibration procedure for μ to achieve a certain utilization of repairmen.

Table A.14: Calibration of μ

Inputs: M, R, λ_m for $1 \leq m \leq M$, target utilization level ρ^T.
Initialization: Set I=0 and $\mu^0 = \frac{M}{\max \Lambda_m R}$.
Step 1: Apply SCAT algorithm to find ρ^I. Step 2: If $|\rho^T - rho^I| < 0.001$, stop.
Otherwise, set I=I+1. If $\rho^T > rho^I$, $\mu^{I+1} = \mu_I - 0.01$, else $\mu^{I+1} = \mu_I + 0.01$.
Go to Step 1.
Outputs: $\mu = \mu^I$.

Appendix B. Details of the simulation study

In this part, we explain how we set up our discrete event simulations. We use the non-overlapping batch means approach for which an extensive description is provided in Steiger and Wilson (2001). After analyzing sample scenarios, the number of batches and the batch size are determined. Accordingly, each simulation run is divided in 201 batches each with 10,000 maintenance completions. The simulation is started in the state with all machines working and the first batch is disregarded to eliminate the initial bias. The average cost values are recorded for the remaining 200 batches and then the corresponding confidence interval is constructed. For each data instance, we confirm that 95% confidence interval half-widths are less than 1% of the average estimate.

Acknowledgement

Acknowledgments The authors gratefully acknowledge the support of the Netherlands Organisation for Scientific Research under grant number is 407-12-001.

References

Akcay, A., Topan, E., and van Houtum, G.-J. (2021). Machine tools with hidden defects: Optimal usage for maximum lifetime value. *IISE Transactions*, 53(1):74–87.

Alaswad, S. and Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157:54–63.

Ansell, P., Glazebrook, K. D., Nino-Mora, J., and O'Keeffe, M. (2003). Whittle's index policy for a multi-class queueing system with convex holding costs. *Mathematical Methods of Operations Research*, 57(1):21–39.

Archibald, T. W., Black, D., and Glazebrook, K. D. (2009). Indexability and index heuristics for a simple class of inventory routing problems. *Operations Research*, 57(2):314–326.

Armstrong, M. J. (2002). Age repair policies for the machine repair problem. *European Journal of Operational Research*, 138(1):127–141.

Ayesta, U., Gupta, M. K., and Verloop, I. M. (2020). On the computation of whittle's index for markovian restless bandits. *Mathematical Methods of Operations Research*, pages 1–30.

Bunday, B. D., Bokhari, H. M., and Khorram, E. (1997). The efficiency of two groups of heterogeneous stations cared for by r operatives with a priority group. *Applied Mathematical Modelling*, 21(1):42–47.

Camci, F. (2015). Maintenance scheduling of geographically distributed assets with prognostics information. *European Journal of Operational Research*, 245(2):506–516.

Cho, D. I. and Parlar, M. (1991). A survey of maintenance models for multi-unit systems. *European journal of operational research*, 51(1):1–23.

de Jonge, B. and Scarf, P. A. (2020). A review on maintenance optimization. *European Journal of Operational Research*, 285(3):805–824.

Desruelle, P. and Steudel, H. J. (1996). A queuing network model of a single-operator manufacturing workcell with machine/operator interference. *Management Science*, 42(4):576–590.

Fouladirad, M. and Grall, A. (2014). On-line change detection and condition-based maintenance for systems with unknown deterioration parameters. *IMA Journal of Management Mathematics*, 25(2):139–158.

Fouladirad, M. and Grall, A. (2015). Monitoring and condition-based maintenance with abrupt change in a system's deterioration rate. *International Journal of Systems Science*, 46(12):2183–2194.

Fouladirad, M., Grall, A., and Dieulle, L. (2008). On the use of on-line detection for maintenance of gradually deteriorating systems. *Reliability Engineering & System Safety*, 93(12):1814–1820.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2):148–164.

Glazebrook, K. D., Kirkbride, C., and Ouenniche, J. (2009). Index policies for the admission control and routing of impatient customers to heterogeneous service stations. *Operations Research*, 57(4):975–989.

Glazebrook, K. D., Mitchell, H., and Ansell, P. (2005). Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research*, 165(1):267–284.

Grall, A., Bérenguer, C., and Dieulle, L. (2002). A condition-based maintenance policy for stochastically deteriorating systems. *Reliability Engineering & System Safety*, 76(2):167–180.

Haque, L. and Armstrong, M. J. (2007). A survey of the machine interference problem. *European Journal of Operational Research*, 179(2):469–482.

Kao, E. P. (1973). Optimal replacement rules when changes of state are semi-markovian. *Operations Research*, 21(6):1231–1249.

Kolesar, P. (1966). Minimum cost replacement under markovian deterioration. *Management Science*, 12(9):694–706.

Koochaki, J., Bokhorst, J. A., Wortmann, H., and Klingenberg, W. (2013). The influence of condition-based maintenance on workforce planning and maintenance scheduling. *International Journal of Production Research*, 51(8):2339–2351.

Kryvinska, N. (2004). Intelligent network analysis by closed queuing models. *Telecommunication Systems*, 27(1):85–98.

Kurt, M. and Kharoufeh, J. P. (2010). Optimally maintaining a markovian deteriorating system with limited imperfect repairs. *European Journal of Operational Research*, 205(2):368–380.

Larranaga, M., Ayesta, U., and Verloop, I. M. (2016). Dynamic control of birth-and-death restless bandits: Application to resource-allocation problems. *IEEE/ACM Transactions on Networking*, 24(6):3812–3825.

Lavenberg, S. S. and Reiser, M. (1980). Stationary state probabilities at arrival instants for closed queueing networks with multiple types of customers. *Journal of Applied Probability*, pages 1048–1061.

Li, D., Ding, L., and Connor, S. (2020). When to switch? index policies for resource scheduling in emergency response. *Production and Operations Management*, 29(2):241–262.

Liu, B., Xu, Z., Xie, M., and Kuo, W. (2014). A value-based preventive maintenance policy for multi-component system with continuously degrading components. *Reliability Engineering & System Safety*, 132:83–89.

López-Santana, E., Akhavan-Tabatabaei, R., Dieulle, L., Labadie, N., and Medaglia, A. L. (2016). On the combined maintenance and routing optimization problem. *Reliability Engineering & System Safety*, 145:199–214.

Marseguerra, M., Zio, E., and Podofillini, L. (2002). Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation. *Reliability Engineering & System Safety*, 77(2):151–165.

Neuse, D. and Chandy, K. (1981). Scat: A heuristic algorithm for queueing network models of computing systems. In *Proceedings of the 1981 ACM SIGMETRICS conference on Measurement and modeling of computer systems*, pages 59–79.

Niño-Mora, J. (2012). Towards minimum loss job routing to parallel heterogeneous multiserver queues via index policies. *European Journal of Operational Research*, 220(3):705–715.

Olde Keizer, M. C., Flapper, S. D. P., and Teunter, R. H. (2017). Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, 261(2):405–420.

Papadimitriou, C. H. and Tsitsiklis, J. N. (1999). The complexity of optimal queuing network control. *Mathematics of Operations Research*, 24(2):293–305.

Puterman, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Ruiz-Hernández, D., Pinar-Pérez, J. M., and Delgado-Gómez, D. (2020). Multi-machine preventive maintenance scheduling with imperfect interventions: A restless bandit approach. *Computers & Operations Research*, 119:104927.

Si, X.-S., Wang, W., Hu, C.-H., and Zhou, D.-H. (2011). Remaining useful life estimation–a review on the statistical data driven approaches. *European Journal of Operational Research*, 213(1):1–14.

Steiger, N. M. and Wilson, J. R. (2001). Convergence properties of the batch means method for simulation output analysis. *INFORMS Journal on Computing*, 13(4):277–293.

Tian, Z., Jin, T., Wu, B., and Ding, F. (2011). Condition based maintenance optimization for wind power generation systems under continuous monitoring. *Renewable Energy*, 36(5):1502–1509.

Tian, Z. and Liao, H. (2011). Condition based maintenance optimization for multi-component systems using proportional hazards model. *Reliability Engineering & System Safety*, 96(5):581–589.

van der Weide, J. A. and Pandey, M. D. (2011). Stochastic analysis of shock process and modeling of condition-based maintenance. *Reliability Engineering & System Safety*, 96(6):619–626.

Van der Weide, J. A., Pandey, M. D., and van Noortwijk, J. M. (2010). Discounted cost model for condition-based maintenance optimization. *Reliability Engineering & System Safety*, 95(3):236–246.

Wang, H. (2002). A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3):469–489.

Wang, K.-H. (1994). Comparative analysis for the $\mathbf{M/K_k/1}$ machine repair problem with spares. *Computers & Industrial Engineering*, 26(4):765–774.

Weber, R. R. and Weiss, G. (1990). On an index policy for restless bandits. *Journal of Applied Probability*, 27(3):637–648.

Whittle, P. (1988). Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25(A):287–298.

Ye, Z.-S. and Xie, M. (2015). Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry*, 31(1):16–32.

Zhu, Y., Elsayed, E. A., Liao, H., and Chan, L.-Y. (2010). Availability optimization of systems subject to competing risk. *European Journal of Operational Research*, 202(3):781–788.