

Time-Based Maintenance Models Under Uncertainty

Peter Buchholz, Iryna Dohndorf^(✉), and Dimitri Scheftelowitsch

Informatik IV, TU Dortmund, Dortmund, Germany

{peter.buchholz, iryna.dohndorf, dimitri.scheftelowitsch}@cs.tu-dortmund.de

Abstract. Model based computation of optimal maintenance strategies is one of the classical applications of Markov Decision Processes. Unfortunately, a Markov Decision Process often does not capture the behavior of a component or system of components correctly because the duration of different operational phases is not exponentially distributed and the status of component is often only partially observable during operational times. The paper presents a general model for components with partially observable states and non-exponential failure, maintenance and repair times which are modeled by phase type distributions. Optimal maintenance strategies are computed using Markov decision theory. However, since the internal state of a component is not completely known, only bounds for the parameters of a Markov decision process can be computed resulting in a bounded parameters Markov decision process. For this kind of process optimal strategies can be computed assuming best, worst or average case behavior.

Keywords: Maintenance models · Markov decision processes
Stochastic dynamic programming · Numerical methods

1 Introduction

Maintenance problems appear in various aspects of human life wherever there exists a *system* or *component* that degrades with time and has to be kept in an operational condition by means of *active maintenance*, i. e., a conscious effort to inspect and revert the effects of degradation. In general one can distinguish between *condition-* or *event-*based maintenance and *preventive-* or *time-*based maintenance [11]. In the former case, which is more often analyzed, maintenance takes places whenever a specific condition or event is observed in a system which indicates a degradation of the system. In preventive or time-based maintenance, the system is inspected at predefined time points and based on the results of inspection it is decided whether a maintenance operation is performed or not. Usually this decision is based on incomplete information about the system state because the degradation of most systems can only be partially observed or tested [16]. This implies that maintenance decisions have to be made under uncertainty about the current state and future behavior of the system.

A popular approach to model this kind of problems are Markov decision processes (MDPs) [6, 15, 19]. In it, the stages of degradation are modeled as individual *states* of the system, the maintenance options as possible *actions* of the

controller, the results of actions as a probabilistic transition mechanism between the states, and the degradation itself is modeled as a *reward function*. The optimal policy computed from the MDP describes the optimal maintenance strategy. Since failures occur at random times, they are usually described by continuous time models. Fixed inspection times introduce a time-step in the model which implies that a discrete time model is more appropriate for the computation of optimal policies. Thus, we consider discrete time MDPs (DTMDPs) [19] resulting from the embedding of a continuous time MDP (CTMDP) to compute optimal time-based maintenance strategies.

The basis of Markov models is the memoryless property which implies that the time to the next event, in our case next degradation of the component or system state, is independent of the past. This means that failure or degradation times are exponentially distributed which is usually not the case for real systems. It is known that Weibull or log-normal distributions are much more realistic models for the behavior of systems. Consequently, MDPs are only an approximate model because the future behavior depends on additional non-observable states of the system. To express this uncertainty in the model, one has to extend the model class to partially observable MDPs (POMDPs) or MDPs with uncertain parameters. The price for such an extension is a more complex model and the need to compute bounds rather than exact results.

In our work, we combine in some sense POMDPs and MDPs with uncertain parameters in order to provide a framework for maintenance models of systems with several components. Prior to the introduction of the approach, we give in Sect. 2 a brief overview of related work. Then, in Sect. 3, we model failure time distributions by embedded Markov processes to build a realistic model of component aging. In Sect. 4, we apply knowledge about uncertainty in Markov decision processes to derive a decision model and reduce the state space and make computing optimal maintenance policies tractable even for more complex models. Finally, we present simple examples in Sect. 5 and discuss the results in Sect. 6. The proofs for the main theorems used in the paper can be found online [1].

2 Related Work

An enormous number of papers on MDPs and their use in maintenance modeling exists. Therefore we mention only briefly the main references.

Markov decision processes and, briefly, their application to maintenance and dependability have been thoroughly analyzed in the textbook by Puterman [19]. A more elaborate discussion of Markov models for maintenance problems can be found in [11, 15]; for specific applications in this context, we refer to [3, 6, 12, 16]. Probability distributions of the phase type and their application to model failure and repair times have been considered in [5, 13]. More complex maintenance models based on phase-type distributions can be found in [2, 7].

The extension of Markov decision processes with interval-type transition probabilities has been discussed in [14]. Maintenance applications of a related uncertain model can be found in [18].

3 Markov Models for Maintenance

We consider dependability models that are described by Markov processes. In the following first the basic model of a component and then the composition of components is introduced.

Component Models. The behavior of a component can be described by a stochastic process with $2N + 1$ phases, which are divided in three groups: N *operational* phases, N *maintenance* phases and one *failure* phase. We use the term *phase* rather than *state* because states will later be used to model non-exponential durations of phases. Operational phases are numbered 1 through N , maintenance phases receive numbers $N + 1$ through $2N$ and the failure phase number 0.

A new component starts in operational phase N and then continuously degrades by decreasing the operational phase or by a hard failure that immediately brings the component down to the failure phase. Let T_i be the random variable describing the duration of phase i and p_i the probability of entering from operational phase i immediately failure phase 0. Consequently, $\bar{p}_i = 1 - p_i$ is the probability of entering operational phase $i - 1$ from i and $p_1 = 1$.

Eventually a component will end in phase 0, the failure phase, if no action is taken. In the failure phase, the component may be repaired or substituted by a new component, in both cases the component starts again in operational phase N . Additionally, maintenance operations may be introduced. Maintenance can be initiated at inspection times $l \cdot \Delta$ ($l \in \mathbb{N}$) when the phase of the component is determined. If a maintenance operation is initiated when the component is in operational phase i , then the phase changes to $N + i$. In phase $N + i$ ($i \in \{1, \dots, N\}$) maintenance is performed, which requires time T_{N+i} . Afterwards the phase changes with probability $h_{N+i,j}$ to operational phase $j \in \{i, \dots, N\}$, i.e., we assume that maintenance cannot deteriorate the component.

Figure 1 shows the state transition diagram of a component. Solid arcs describe transitions that occur if nothing is done and dashed arcs describe transitions that take place if a maintenance operation is initiated. The above model is similar to the classical models for maintenance [8, 19]. To analyze the behavior and to determine optimal decisions, quantitative measures in the form of rewards

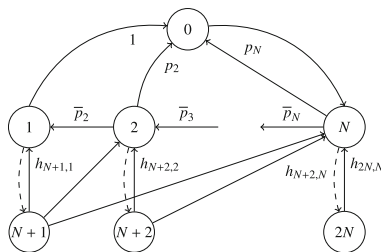


Fig. 1. Phase transition diagram of a component.

have to be associated with phases or transitions. We assume that in operational phase i , the component gains reward r_i in phase i and additionally a penalty $r_0 < 0$ has to be paid whenever the failure phase is entered.

Sojourn times in the phases are characterized by independently distributed random variables with cumulative distribution function $F_i(t)$ and probability density $f_i(t)$ for phase i . The time-dependent rate out of a phase is then given by

$$\lambda_i(t) = \frac{f_i(t)}{1 - F_i(t)}. \quad (1)$$

For the phases 1 through N this quantity describes the rate of deterioration in the operational phase. For modeling failure times, standard distributions like Weibull, Gamma or log-normal or phase type distributions are applied [5, 13]. For repair or maintenance times less information is available in the literature but it seems that also in these cases phase type distributions are appropriate for modeling the durations [2, 7]. In this paper, we use acyclic phase type distributions (APHs), a subclass of phase type distributions, which are described by an absorbing Markov chain with m transient states, initial vector $\boldsymbol{\nu}$ and generator matrix \mathbf{D} that can be reordered to an upper triangular matrix. As shown for example in [5, 9] general phase type distributions as well as standard failure time distributions like the Weibull distribution can be approximated sufficiently accurately by APHs. One of the advantages of APHs is the availability of a canonical representation developed in [10]. In the canonical representation matrix \mathbf{D} has the form

$$\mathbf{D} = \begin{pmatrix} -\mu_1 & \mu_1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & -\mu_{m-1} & \mu_{m-1} \\ 0 & \cdots & \cdots & 0 & -\mu_m \end{pmatrix} \quad (2)$$

with $\mu_i \leq \mu_{i+1}$ for $1 \leq i < m - 1$. This implies that the distribution has only one exit state, namely the last state. The representation $(\boldsymbol{\nu}, \mathbf{D})$ characterizes the APH and

$$f(t) = \boldsymbol{\nu} e^{t\mathbf{D}} (-\mathbf{D}\mathbf{1}), \quad F(t) = 1 - \boldsymbol{\nu} e^{t\mathbf{D}} \mathbf{1}, \quad \lambda(t) = \frac{\boldsymbol{\nu} e^{t\mathbf{D}} (-\mathbf{D}\mathbf{1})}{\boldsymbol{\nu} e^{t\mathbf{D}} \mathbf{1}}, \quad (3)$$

where $\mathbf{1}$ is a column vector of 1 of appropriate length. For a time-dependent rate $\lambda(t)$, the bounds $\lambda_{t^-, t^+}^- = \min_{t^+ \geq t \geq t^-} \lambda(t)$ and $\lambda_{t^-, t^+}^+ = \max_{t^+ \geq t \geq t^-} \lambda(t)$ with $t^- \leq t^+$ can be computed. We use the notation $\lambda^- = \lambda_{0, \infty}^-$ and $\lambda^+ = \lambda_{0, \infty}^+$.

The average rate is given by $\lambda^{av} = -(\boldsymbol{\nu} \mathbf{D}^{-1} \mathbf{1})^{-1}$. If the APH represents an exponential distribution, then $\lambda^{av} = \lambda^- = \lambda^+$.

We assume that the sojourn time in phase i is characterized by an APH $(\boldsymbol{\nu}_i, \mathbf{D}_i)$ of order m_i and define $\mathbf{d}_i = -\mathbf{D}_i \mathbf{1} = (0, \dots, 0, \mu_{m_i})^T$, for $i \in \{0, \dots, 2N\}$. This implies that all times are phase type distributed according to an APH. Thus, the state of a component is defined by the observable phase of

the component and the non-observable state of the APH defining the duration of the phase. The detailed stochastic process describing a component is defined in Sect. 4.

Composition of Components. If several components are considered, they run in parallel. As long as they are independent, components can be analyzed in isolation and rewards are simply added. However, usually components are dependent. We consider here the case that rewards depend on the joint phase. E.g., a system with two components is no longer operating if both components are in maintenance or in the failure phase. Thus, a penalty has to be paid, if this situation occurs. Additionally, components may interact due to limited capacity for maintenance. In this case, a maintenance operation can only be started if capacity for maintenance is available, otherwise the component has to continue working until the next decision point for maintenance.

We restrict the description and analysis to systems built from two components to avoid an extensive and complex notation. However, the general approach can be extended to more than two components following the presented ideas.

4 Decision Models

The models described above implicitly include decisions to perform a maintenance operation or let the system run. It is now shown that these decisions can be modeled adequately by a Markov Decision Process (MDP) [19], if only the average behavior is analyzed. In the general case, it can be modeled as a Bounded Parameter Markov Decision Process (BMDP) [14]. From these processes optimal maintenance strategies are computed. We consider in the following first single components and then models composed of two components.

For notational convenience we define the following abbreviations $m_{i:j} = \sum_{h=i}^j m_h$ if $i \leq j$ and 0 otherwise, $e_{i,n}$ ($i \leq n$) for the i th unit row vector of length n , $\mathbf{0}_m$ for a zero row vector of length m , $\mathbf{0}_{n,m}$ for a zero matrix of order $n \times m$ and \mathbf{I}_m for an identity matrix of order m . If the order of vectors or matrices follows from the context, we avoid the use of prefixes. We assume that the indices of matrices and vectors start with 0 and end with $n - 1$ for a vector or matrix of dimension n .

4.1 Decision Model for a Component

For maintenance the component is observed at discrete time points $l \cdot \Delta$ ($l = 0, 1, 2, \dots$) for some time step $\Delta > 0$ and maintenance is possibly initiated at those time points. At time $l \cdot \Delta$ we can observe the phase but not the internal state which in some sense describes the internal wear of the component. Consequently, decisions have to be based on the phase which implies that we build a decision process with $2N + 1$ states. However, since the behavior of this process depends on the unknown internal state at a decision point, we consider the worst, best and average behavior. The former two define a BMDP [14], the latter

an MDP. Before the decision process can be generated, the internal behavior of the component between decision points has to be characterized. Without maintenance the component can be described by a Markov chain with $m_{0:2N}$ states and generator matrix \mathbf{Q} .

Matrix \mathbf{Q} can be built from the phase type distributions and the different probabilities of the degradation process. Let \mathbf{F} be a $m_{1:N} \times m_{1:N}$ matrix with $N \times N$ submatrices $\mathbf{F}_{i,j}$ ($1 \leq i, j \leq N$) of size $m_i \times m_j$.

$$\mathbf{F}_{i,j} = \begin{cases} \mathbf{D}_i & \text{if } j = i, \\ \bar{p}_i \mathbf{d}_i \boldsymbol{\nu}_{i-1} & \text{if } j = i - 1, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Matrix \mathbf{F} includes all transitions between operational phases. Matrix $\mathbf{D}_{N+1:2N}$ is a $m_{N+1:2N} \times m_{N+1:2N}$ matrix with diagonal blocks \mathbf{D}_i ($i = N + 1, \dots, 2N$), it contains all transitions between maintenance phases. To describe transitions between maintenance, operational and failure phases define the following matrices

$$\mathbf{E}_0 = (\mathbf{0}_{m_0, m_{1:N-1}} \quad \mathbf{d}_0 \boldsymbol{\nu}_N), \quad \mathbf{E}_1 = \begin{pmatrix} p_1 \mathbf{d}_1 \boldsymbol{\nu}_0 \\ \vdots \\ p_N \mathbf{d}_N \boldsymbol{\nu}_0 \end{pmatrix},$$

and

$$\mathbf{E}_2 = \begin{pmatrix} h_{N+1,1} \mathbf{d}_{N+1} \boldsymbol{\nu}_1 & \cdots & \cdots & h_{N+1,N} \mathbf{d}_{N+1} \boldsymbol{\nu}_N \\ 0 & h_{N+2,2} \mathbf{d}_{N+2} \boldsymbol{\nu}_2 & \cdots & h_{N+2,N} \mathbf{d}_{N+2} \boldsymbol{\nu}_N \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{2N,N} \mathbf{d}_{2N} \boldsymbol{\nu}_N \end{pmatrix}.$$

\mathbf{E}_0 describes the start of component after a repair, \mathbf{E}_1 the immediate failure from one of the operational phases and \mathbf{E}_2 contains all transitions that transfer the component from a maintenance phase to an operational phase. Then

$$\mathbf{Q} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{E}_0 & \mathbf{0} \\ \mathbf{E}_1 & \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_2 & \mathbf{D}_{N+1:2N} \end{pmatrix} \quad (4)$$

contains all transitions between decision points. Matrix \mathbf{Q} is block structured into $(2N + 1) \times (2N + 1)$ blocks of size $m_i \times m_j$ for sub-matrix $\mathbf{Q}_{i,j}$.

Rewards are collected in a column vector \mathbf{r} of length $\sum_{i=0}^{2N} m_i$. All states belonging to the phase i with $i = 0 \vee N < i \leq 2N$ have reward r_i . The state belonging to phase i ($1 \leq i \leq N$) and state j ($1 \leq j \leq m_i$), which corresponds to entry number $\sum_{k=0}^{i-1} m_k + j - 1$ in vector \mathbf{r} , equals $r_i + p_i \mathbf{d}_i(j) r_0$. For most algorithms to compute optimal policies for MDPs or BMDPs it is advantageous to have only non-negative rewards. In finite state processes with bounded rewards

this can be achieved by adding an appropriate constant which is later removed from the solution [20]. Thus define $r_{pos} = \left\lceil \min(0, \min_j(\mathbf{r}(j))) \right\rceil$ and substitute \mathbf{r} by $\mathbf{r} + r_{pos}\mathbf{1} \geq \mathbf{0}$.

Now consider an inspection time and assume that the component is in phase $i \in \{0, \dots, 2N\}$ at time $l \cdot \Delta$. We want to compute the probability that the component is in phase $j \in \{0, \dots, 2N\}$ at time $(l+1) \cdot \Delta$. If the internal state of phase i is known, this probability can be computed exactly using matrix \mathbf{Q} . However, we only know that for the distribution of states in phase i the relation $\psi = \boldsymbol{\nu}_i e^{t^* \mathbf{D}_i} / (\boldsymbol{\nu}_i e^{t^* \mathbf{D}_i} \mathbf{1})$ for some $t^* \geq 0$ holds. To compute bounds for the probabilities we define the following matrices.

$$\hat{\mathbf{B}}_i^- = \begin{pmatrix} -\lambda_i^+ & \lambda_i^- \mathbf{b}_i \\ \mathbf{0} & \mathbf{Q} \end{pmatrix}, \quad \hat{\mathbf{B}}_i^+ = \begin{pmatrix} -\lambda_i^- & \lambda_i^+ \mathbf{b}_i \\ \mathbf{0} & \mathbf{Q} \end{pmatrix}, \quad \hat{\mathbf{B}}_i^{av} = \begin{pmatrix} -\lambda_i^{av} & \lambda_i^{av} \mathbf{b}_i \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} \quad (5)$$

where $i \in \{0, \dots, 2N\}$ and

$$\mathbf{b}_i = \begin{cases} (\mathbf{0}_{m_{1:N-1}}, \boldsymbol{\nu}_N, \mathbf{0}_{m_{N+1:2N}}) & \text{if } i = 0, \\ (\boldsymbol{\nu}_0, \mathbf{0}_{m_{1:2N}}) & \text{if } i = 1, \\ (p_i \boldsymbol{\nu}_0, \mathbf{0}_{m_{1:i-2}}, \bar{p}_i \boldsymbol{\nu}_{i-1}, \mathbf{0}_{m_{i:2N}}) & \text{if } 1 < i \leq N \\ (\mathbf{0}_{m_{0:i-N-1}}, h_{-N, i-N} \boldsymbol{\nu}_{i-N}, \dots, h_{i-N, N} \boldsymbol{\nu}_N, \mathbf{0}_{m_{N+1:2N}}) & \text{if } N < i \leq 2N. \end{cases}$$

Matrix $\hat{\mathbf{B}}_i^-$ describes the component when it leaves phase i as soon as possible and the flow into the other phases is minimal. $\hat{\mathbf{B}}_i^+$ describes the component when it leaves phase i as late as possible and the flow into the other phases is maximal. $\hat{\mathbf{B}}_i^{av}$ describes the average behavior. Observe that only $\hat{\mathbf{B}}_i^{av}$ is a generator matrix. In the sequel we use \pm for one of the elements from $\{-, +, av\}$.

Let $\phi = (1, \mathbf{0}_{m_{0:2N}})$ and compute $\phi_i^\pm = \phi e^{\Delta \hat{\mathbf{B}}_i^\pm}$. All vectors can be decomposed into $\phi_i^\pm = (\phi_{i,-1}^\pm, \phi_{i,0}^\pm, \dots, \phi_{i,2N}^\pm)$, $\phi_{i,-1}^\pm$ is a scalar which denotes the probability that state i has not been left during the whole interval and $\phi_{i,j}^\pm$ is a vector of length m_j which includes (bounds for) the probabilities that the component is in one of the states of phase j after starting in i . Let

$$\hat{\phi}_i^\pm = \left(\phi_{i,0}^\pm \mathbf{1}, \dots, \phi_{i,i-1}^\pm \mathbf{1}, \phi_{i,i}^\pm \mathbf{1} + \phi_{i,-1}^\pm, \phi_{i,i+1}^\pm \mathbf{1}, \dots, \phi_{i,2N}^\pm \mathbf{1} \right)^T. \quad (6)$$

Theorem 1. *Vector $\hat{\phi}_i^-$ is a lower bound for the conditional probability distribution over the phases of the component at time $(l+1) \cdot \Delta$ under the condition that the process starts in phase i at time $l \cdot \Delta$.*

Vector $\hat{\phi}_i^+$ is an upper bound for the conditional probability distribution over the phases of the component at time $(l+1) \cdot \Delta$ under the condition that the process starts in phase i at time $l \cdot \Delta$.

The proof for the theorem can be found online [1]. Vector $\hat{\phi}_i^-$ contains only non-negative elements, whereas $\hat{\phi}_i^+$ might include elements larger than 1 since the

matrix $\hat{\mathbf{B}}_i^+$ is superstochastic. Then vector elements can be substituted by the trivial upper bound 1 for probabilities. Now define $(2N+1) \times (2N+1)$ matrices

$$\hat{\mathbf{S}}^\pm = \left(\hat{\phi}_0^\pm \dots \hat{\phi}_{2N}^\pm \right)^\top \quad (7)$$

$\hat{\mathbf{S}}^-$ is substochastic, $\hat{\mathbf{S}}^{av}$ is stochastic and $\hat{\mathbf{S}}^+$ is superstochastic¹ and $\hat{\mathbf{S}}^- \leq \hat{\mathbf{S}}^{av} \leq \hat{\mathbf{S}}^+$. Matrices $\hat{\mathbf{S}}^-$ and $\hat{\mathbf{S}}^+$ bound the transition probabilities of the component between two decision periods.

A maintenance policy can now be described by a vector $\hat{\mathbf{a}} \in \{0, 1\}^N$. $\hat{\mathbf{a}}(i) = 1$ implies that maintenance is initiated if the component is in phase $i \in \{1, \dots, N\}$. Obviously maintenance can only be initiated if the component is in an operational phase. Matrix $\hat{\mathbf{T}}_{\hat{\mathbf{a}}}$ describes the transitions induced by policy $\hat{\mathbf{a}}$ and is defined as

$$\hat{\mathbf{T}}_{\hat{\mathbf{a}}}(i, j) = \begin{cases} 1 & \text{if } (1 \leq i \leq N \wedge j = N + i \wedge \hat{\mathbf{a}}(i) = 1) \vee ((i \notin \{1, \dots, N\} \vee \hat{\mathbf{a}}(i) = 0) \wedge i = j) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Then $\hat{\mathbf{P}}_{\hat{\mathbf{a}}}^- = \hat{\mathbf{T}}_{\hat{\mathbf{a}}} \hat{\mathbf{S}}^-$ and $\hat{\mathbf{P}}_{\hat{\mathbf{a}}}^+ = \hat{\mathbf{T}}_{\hat{\mathbf{a}}} \hat{\mathbf{S}}^+$ are matrices containing transition probability bounds of a BMDP. Only for the first state, the reward has to be computed, for the remaining states reward vector \mathbf{r} can be used. Define $r_i^\pm = r_i + \lambda_i^\pm r_0 + r_{pos}$ if $i \in \{1, \dots, N\}$ and $r_i^\pm = r_i + r_{pos}$ otherwise, where r_{pos} has been defined above. Furthermore $u^+ = \max_i r_i^+$ and $u^- = \min_i r_i^-$. Reward vector $\hat{\mathbf{s}}^\pm$ is defined element-wise as

$$\hat{\mathbf{s}}^\pm(i) = \mathbf{e}_{0, m_{0:2N+2}} \int_0^\Delta e^{\tau \hat{\mathbf{C}}_i} \begin{pmatrix} \hat{r}_i^\pm \\ \mathbf{r} \\ u^\pm \end{pmatrix} d\tau \text{ where } \hat{\mathbf{C}}_i = \begin{pmatrix} -\lambda_i^+ & \lambda_i^- \mathbf{b}_i & \lambda_i^+ - \lambda_i^- \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ 0 & \mathbf{0} & 0 \end{pmatrix}. \quad (9)$$

It is easy to show that $\hat{\mathbf{s}}^- \leq \hat{\mathbf{s}}^{av} \leq \hat{\mathbf{s}}^+$ holds.

Theorem 2. $\hat{\mathbf{s}}^-(i)$ is a lower bound for the reward that is gained in an interval $[l \cdot \Delta, (l+1) \cdot \Delta]$ under the condition that the process starts at time $l \cdot \Delta$ in phase i .

$\hat{\mathbf{s}}^+(i)$ is an upper bound for the reward that is gained in an interval $[l \cdot \Delta, (l+1) \cdot \Delta]$ under the condition that the process starts at time $l \cdot \Delta$ in phase i .

Again, the proof can be found online [1]. $\left(\hat{\mathbf{P}}_{\hat{\mathbf{a}}}^-, \hat{\mathbf{P}}_{\hat{\mathbf{a}}}^+ \right)_{\hat{\mathbf{a}} \in \{0, 1\}^N}$ and $\left(\hat{\mathbf{s}}^-, \hat{\mathbf{s}}^+ \right)$ define a BMDP to bound the reward of the policies. Furthermore, $\hat{\mathbf{P}}_{\hat{\mathbf{a}}}^{av} = \hat{\mathbf{T}}_{\hat{\mathbf{a}}} \hat{\mathbf{S}}^{av}$ is the stochastic transition matrix and $\left(\hat{\mathbf{P}}_{\hat{\mathbf{a}}}^{av} \right)_{\hat{\mathbf{a}} \in \{0, 1\}^N}$ and $\hat{\mathbf{s}}^{av}$ define the discrete time MDP which models the average behavior of the component.

After the optimal policy has been computed from the BMDP, the policy can be analyzed by building the CTMC described by the policy and then analyzing the policy on this process. Time complexity analysis can be found in [1].

¹ We denote a non-negative matrix \mathbf{A} as substochastic, iff a stochastic matrix \mathbf{P} with $\mathbf{A} \leq \mathbf{P}$ and $\mathbf{A} \neq \mathbf{P}$ exists. Similarly we denote \mathbf{A} as superstochastic, iff a stochastic matrix \mathbf{P} with $\mathbf{A} \geq \mathbf{P}$ and $\mathbf{A} \neq \mathbf{P}$ exists.

4.2 Decision Models for Composed Components

Modeling of composed components is more difficult than the analysis of a single component. We consider here the case of two components but the approach can be easily extended to more than two components.

In a two-component system, phases are described by tuples $\mathbf{i} = (i_0, i_1)$ where $i_k \in \{0, \dots, 2N^{(k)}\}$ and $k \in \{0, 1\}$ is the index of the component. We use $\cdot^{(k)}$ to denote quantities belonging to component k . Matrices and vectors belonging to the composed system will be shown without postfix. A detailed state of the composed system is given by $(\mathbf{i}, \mathbf{j}) = (i_0, i_1, j_0, j_1)$, $i_k \in \{0, \dots, 2N^{(k)}\}$ and $j_k \in \{1, \dots, m_{i_k}^{(k)}\}$.

At a decision point, the detailed state of both components is unknown. Consequently, we have to analyze all possible starting phases (i_0, i_1) using matrices $\bar{\mathbf{B}}_{i_0, i_1}^\pm$ which are an extension of the matrices $\hat{\mathbf{B}}_i^\pm$ defined in (5). However, in contrast to the case with one component we now have to distinguish four cases; both components remain in their phase for the whole interval of length Δ , the first component changes its phase and the second remains in its phase, the first component remains in its phases and the second changes its phases, and, finally, both components change their phases. This results in the definition of the matrices $\bar{\mathbf{B}}_{i_0, i_1}^\pm$ by

$$\bar{\mathbf{B}}_{i_0, i_1}^- = \begin{pmatrix} -\lambda_{i_0}^{+(0)} - \lambda_{i_1}^{+(1)} & \lambda_{i_1}^{-(1)} \mathbf{b}_{i_1}^{(1)} & \lambda_{i_0}^{-(0)} \mathbf{b}_{i_0}^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{(0)} \oplus -\lambda_{i_1}^{+(1)} & \mathbf{0} & \mathbf{I} \otimes \lambda_{i_1}^{-(1)} \mathbf{b}_{i_1}^{(1)} \\ \mathbf{0} & \mathbf{0} & -\lambda_{i_0}^{+(0)} \oplus \mathbf{Q}^{(1)} & \lambda_{i_0}^{-(0)} \mathbf{b}_{i_0}^{(0)} \otimes \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}^{(0)} \oplus \mathbf{Q}^{(1)} \end{pmatrix}.$$

The matrices $\bar{\mathbf{B}}_{i_0, i_1}^+$ and $\bar{\mathbf{B}}_{i_0, i_1}^{av}$ are defined analogously by swapping $+$ and $-$ resp. using av in the superscripts.

Then define $\bar{\Phi} = \left(\mathbf{1}, \mathbf{0}_{\binom{(0)}{0:2N^{(0)}+1} \binom{(1)}{0:2N^{(1)}+1}} \right)$ and compute

$$\bar{\Phi}_{(i_0, i_1)}^\pm = \bar{\Phi} e^{\Delta \bar{\mathbf{B}}_{i_0, i_1}^\pm}. \quad (10)$$

Vector $\bar{\Phi}_{(i_0, i_1)}^\pm$ can be decomposed into sub-vectors. The first element equals the scalar $\bar{\Phi}_{(i_0, i_1)(-1, -1)}^\pm$ including a bound or the average value for the probability that the initial phases have not been left. Then follow $2N^{(1)} + 1$ vectors $\bar{\Phi}_{(i_0, i_1)(-1, j_1)}^\pm$ of length $m_{j_1}^{(1)}$, followed by $2N^{(0)} + 1$ vectors $\bar{\Phi}_{(i_0, i_1)(j_0, -1)}^\pm$ of length $m_{j_0}^{(0)}$. Finally $N^{(0)}N^{(1)}$ vectors $\bar{\Phi}_{(i_0, i_1)(j_0, j_1)}^\pm$ of length $m_{j_0}^{(0)} m_{j_1}^{(1)}$ follow. From these sub-vectors the $(2N^{(0)} + 1)(2N^{(1)} + 1) \times (2N^{(0)} + 1)(2N^{(1)} + 1)$ matrix $\bar{\mathbf{S}}$ can be computed as shown in the following equation.

$$\bar{\mathbf{S}}((i_0, i_1)(j_0, j_1))^\pm = \begin{cases} \bar{\phi}_{(i_0, i_1)(j_0, j_1)}^\pm \mathbb{1} & \text{if } i_0 \neq j_0 \wedge i_1 \neq j_1, \\ \bar{\phi}_{(i_0, i_1)(j_0, j_1)}^\pm \mathbb{1} + \bar{\phi}_{(i_0, i_1)(-1, j_1)}^\pm \mathbb{1} & \text{if } i_0 = j_0 \wedge i_1 \neq j_1, \\ \bar{\phi}_{(i_0, i_1)(j_0, j_1)}^\pm \mathbb{1} + \bar{\phi}_{(i_0, i_1)(j_0, -1)}^\pm \mathbb{1} & \text{if } i_0 \neq j_0 \wedge i_1 = j_1, \\ \bar{\phi}_{(i_0, i_1)(j_0, j_1)}^\pm \mathbb{1} + \bar{\phi}_{(i_0, i_1)(-1, j_1)}^\pm \mathbb{1} & \\ + \bar{\phi}_{(i_0, i_1)(j_0, -1)}^\pm \mathbb{1} + \bar{\phi}_{(i_0, i_1)(-1, -1)}^\pm \mathbb{1} & \text{if } i_0 = j_0 \wedge i_1 = j_1. \end{cases} \quad (11)$$

Again the relation $\bar{\mathbf{S}}^- \leq \bar{\mathbf{S}}^{av} \leq \bar{\mathbf{S}}^+$ holds and $\bar{\mathbf{S}}^{av}$ is a stochastic matrix. In $\bar{\mathbf{S}}^+$ elements larger than 1 can be substituted by the trivial bound 1.

Maintenance decisions now can be made whenever the component is in an operational phase at a decision point. We define two vectors $\bar{\mathbf{a}}^{(k)}$ ($k = 0, 1$) of length $(2N^{(0)} + 1)(2N^{(1)} + 1)$. $\bar{\mathbf{a}}^{(k)}(i_k, i_{\bar{k}}) = 0$ if $i_k \notin \{1, \dots, N^{(k)}\}$ and for $i_k \in \{1, \dots, N^{(k)}\}$ the element can be 1, a maintenance operation for component k starts, or 0, no maintenance starts. In a state, up to 4 decisions can be made by combining the two possibilities for each component. Maintenance decisions are only possible in $N^{(k)}$ operational phases but may depend on the phase of the other component. Again, if the maintenance capacity is restricted, maintenance operations can only be started if capacity is available, i.e. the other component is not in a maintenance phase. For vectors $\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}$ the binary matrix $\bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}$ is defined by

$$\begin{aligned} \bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}(i_0, i_1)(j_0, j_1) &= 1 \Leftrightarrow (i_0 = j_0 \wedge i_1 = j_1 \wedge \bar{\mathbf{a}}^{(0)}(i_0, i_1) = \bar{\mathbf{a}}^{(1)}(i_0, i_1) = 0) \\ &\vee (j_0 = N^{(0)} + i_0 \wedge i_1 = j_1 \wedge \bar{\mathbf{a}}^{(0)}(i_0, i_1) = 1 \wedge \bar{\mathbf{a}}^{(1)}(i_0, i_1) = 0) \\ &\vee (i_0 = j_0 \wedge j_1 = N^{(1)} + i_1 \wedge \bar{\mathbf{a}}^{(0)}(i_0, i_1) = 0 \wedge \bar{\mathbf{a}}^{(1)}(i_0, i_1) = 1) \\ &\vee (j_0 = N^{(0)} + i_0 \wedge j_1 = N^{(1)} + i_1 \wedge \bar{\mathbf{a}}^{(0)}(i_0, i_1) = \bar{\mathbf{a}}^{(1)}(i_0, i_1) = 1) \end{aligned} \quad (12)$$

otherwise $\bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}(i_0, i_1)(j_0, j_1) = 0$. Then $\bar{\mathbf{P}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}^- = \bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}} \bar{\mathbf{S}}^-$ and $\bar{\mathbf{P}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}^+ = \bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}} \bar{\mathbf{S}}^+$ define transitions probability bounds for a BMDP. $\bar{\mathbf{P}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}}^{av} = \bar{\mathbf{T}}_{\bar{\mathbf{a}}^{(0)}, \bar{\mathbf{a}}^{(1)}} \bar{\mathbf{S}}^{av}$ is a stochastic matrix for an MDP describing the average behavior.

It remains to compute the reward vectors for the BMDP and the MDP. Reward bounds are computed similarly to the case with one component in (9).

$$\begin{aligned} \bar{\mathbf{s}}^\pm(i_0, i_1) &= e_0 \int_0^\Delta e^{\tau \bar{\mathbf{C}}_{i_0, i_1}} \begin{pmatrix} r_{i_0, i_1}^\pm \\ \bar{r}_{i_0, * }^\pm \\ \bar{r}_{*, i_1}^\pm \\ \bar{r} \\ \bar{u}^\pm \end{pmatrix} d\tau, \text{ where} \\ \bar{\mathbf{C}}_{i_0, i_1} &= \begin{pmatrix} -\lambda_{i_0}^{+(0)} - \lambda_{i_1}^{+(1)} & \lambda_{i_1}^{-(1)} \mathbf{b}_{i_1}^{(1)} & \lambda_{i_0}^{-(0)} \mathbf{b}_{i_0}^{(0)} & \mathbf{0} & \lambda_{i_0}^{+(0)} + \lambda_{i_1}^{+(1)} - \lambda_{i_0}^{-(0)} - \lambda_{i_1}^{-(1)} \\ \mathbf{0} & \mathbf{Q}^{(0)} \oplus -\lambda_{i_1}^{+(1)} & \mathbf{0} & \mathbf{I} \otimes \lambda_{i_1}^{-(1)} \mathbf{b}_{i_1}^{(1)} & (\lambda_{i_1}^{+(1)} - \lambda_{i_1}^{-(1)}) \mathbb{1} \\ \mathbf{0} & \mathbf{0} & -\lambda_{i_0}^{+(0)} \oplus \mathbf{Q}^{(1)} \lambda_{i_0}^{-(0)} \mathbf{b}_{i_0}^{(0)} \otimes \mathbf{I} & \mathbf{0} & (\lambda_{i_0}^{+(1)} - \lambda_{i_0}^{-(1)}) \mathbb{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}^{(0)} \oplus \mathbf{Q}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}. \end{aligned} \quad (13)$$

$\tilde{\mathbf{r}}$ is the reward vector of the system, $\tilde{u}^+ = \max_{i_0, i_1} r_{i_0, i_1}^+$, $\tilde{u}^- = \min_{i_0, i_1} r_{i_0, i_1}^-$ and

$$\begin{aligned} \tilde{\mathbf{r}}_{i_0, *}^{\pm} &= \left(r_{i_0, 0}^{\pm} \mathbf{1}_{m_0^{(1)}}^T, \dots, r_{i_0, 2N^{(1)}+1}^{\pm} \mathbf{1}_{m_2^{(1)} N^{(1)}+1}^T \right)^{\top} \\ \tilde{\mathbf{r}}_{*, i_1}^{\pm} &= \left(r_{0, i_1}^{\pm} \mathbf{1}_{m_0^{(0)}}^T, \dots, r_{2N^{(0)}+1, i_1}^{\pm} \mathbf{1}_{m_2^{(0)} N^{(0)}+1}^T \right)^{\top} \end{aligned} \quad (14)$$

The values r_{i_0, i_1}^{\pm} have been defined above.

The generation of the processes requires the solution of $(2N^{(0)}+1)(2N^{(1)}+1)$ systems of differential Eqs. (10) and (13) of order $(m_{0:2N^{(0)}}^{(0)}+1)(m_{0:2N^{(1)}}^{(1)}+1)$ each. The different computations can be parallelized. The resulting MDPs or BMDPs are usually of a moderate size because the number of phases is usually small.

For computing the optimal policy and average or discounted reward, standard methods for MDP analysis can be applied [19]. For BMDPs optimal policies for the worst case are usually computed. An overview of available numerical methods is given in [4].

4.3 Improved Bounds

The quality of the bounds depends on the difference between λ_i^+ and λ_i^- . Up to now we have computed $\lambda^- = \lambda_{0, \infty}^-$ and $\lambda^+ = \lambda_{0, \infty}^+$ which defines bounds without any knowledge how long the component or system is already in the current phase. However, due to regular inspection of a component, some information is available. Thus, if the component has been in a different phase at the last inspection, $\lambda^- = \lambda_{0, \Delta}^-$ and $\lambda^+ = \lambda_{0, \Delta}^+$. More general, if we know that the component has been in the same phase for the last i inspections, then $\lambda^- = \lambda_{(i-1) \cdot \Delta, i \cdot \Delta}^-$ and $\lambda^+ = \lambda_{(i-1) \cdot \Delta, i \cdot \Delta}^+$. The average rate can then be computed $\lambda^{av} = \Delta^{-1} \int_{(i-1) \cdot \Delta}^{i \cdot \Delta} \lambda(t)$.

If the number of intervals the system resides in a phase is considered, then MDPs/BMDPs have to be generated and analyzed for each combination of state and residence time which increases the effort. However, for Δ much larger than the expected residence time in a phase not much changes because it is likely that the phase has been left at the next inspection.

5 Examples

The following examples have all been analyzed with Matlab/Octave where the proposed approach has been implemented. Matrices are constructed from the specification of the APHs and the remaining parameters and the resulting MDP or BMDP is analyzed using value iteration. Model generation algorithms and runtime analysis are presented in the online companion [1].

We first consider the computation of maintenance policies for single components and consider afterwards the composition of two components.

Single Components: We start with a single component which has 4 operational phases and a reward of 1 in all operational phases, 0 in maintenance phases, and -10 in the failure phase. The duration of the first and the last operational phase are Weibull distributed. Phase 1 has increasing failure rate and phase 4 decreasing failure rate. The operational phases 2 and 3 are exponentially distributed. This implies that the failure rate over all operational phases is first decreasing, then constant and finally increasing which together gives the typical bathtub like curve. Figure 2a shows the time dependent failure rates of the Weibull distributions and the failure rates of the APHs that have been used to approximate the Weibull distributions. For the computation of the parameters of the APHs an EM algorithm and publicly available software has been applied [21]. It can be seen that 3 states of the underlying Markov chain in the APH model result in a reasonable and 5 states in a good approximation of the Weibull distributions.

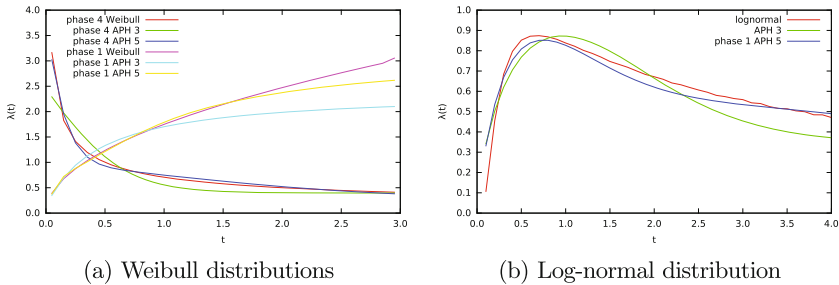


Fig. 2. Time dependent failure rates of various distributions and the respective APH approximations

For the duration of maintenance operations, log-normal distributions have shown to be an adequate model [17]. Usually the variance of the maintenance time is not too large. We generated a log-normal distribution from standard normal distribution and scaled afterwards the distribution to meet the required mean value. APHs of a low order are also suitable to approximate log-normal distributions. Figure 2b show the time dependent rate of the log-normal distribution and APHs with 3 and 5 Markov chain states which have been computed for approximating the log-normal distribution. Again 3 states provide a reasonable and 5 phases a good approximation.

The failure and replacement time is hyper-exponentially distributed with a squared coefficient of variation of 2 to indicate the uncertainty of substituting a component by a new one. We assume that the mean duration of each operational phase is 10, the mean duration of the failure phase 1 and the mean durations of the maintenance phases are 0.4, 0.6, 0.8 and 1, for maintenance starting in the operational phases 4 through 1 (i.e., maintenance becomes more expensive if started later).

The Markov models resulting from a component have 22 or 34 states depending whether we use APHs with 3 or 9 states. The MDP and BMDP have 3 states.

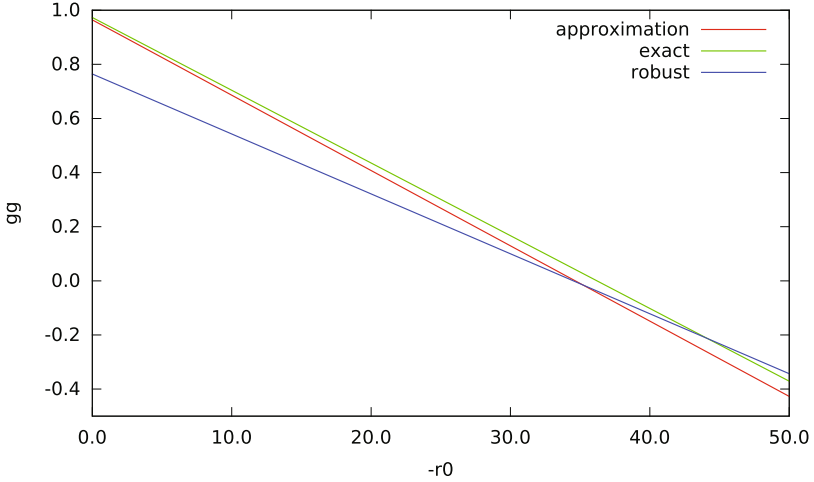


Fig. 3. Optimal gain for different values of the repair costs. (Color figure online)

We consider the configuration with $p_2 = p_3 = p_4 = 0.1$, perfect maintenance (i.e., after each maintenance operation the component behaves like a new component) and an inspection interval of length $\Delta = 1$. If we optimize the average model, no maintenance is performed and the gain is linearly increasing with the costs for a repair. If we analyze the BMDP according to the worst case behavior, a maintenance operation is performed whenever the component is in the last operational phase at an inspection interval. In Fig. 3 we compare the gain under the optimal policy without repair and the robust policy with repair in the last operational phase. The red line shows the gain from the approximation which is used to compute the policy. The green line shows the exact result for the policy. It can be seen that the approximation is for small repair costs good but becomes worse if an repair becomes more expensive. The blue line shows the exact gain of the robust policy which is for small repair cost clearly worse but becomes advantageous if repair cost are high.

Composed Components: For a composed system, we have added a second component that is identical to the first. We analyze the behavior for varying inspection interval $\Delta \in \{0.05, 0.15, \dots, 1.95, 2\}$ with a penalty of -10 . It is assumed that penalties are only paid when both components are down.

Furthermore, we consider a system with fixed $\Delta = 0.5$ and varying penalties in the range from -10 to -50 . For varying penalties, we observe in Fig. 4 that the lower bound for optimal policies in the resulting BMDP model decreases with increasing penalty value; the upper bound and the average model are not affected, as the optimal policy in both cases is very close to the upper bound, which is 2 (one reward unit for each component).

For varying time steps, in Fig. 5 one can see that more often inspection contributes to decreasing level of system uncertainty. For a small time interval Δ ,

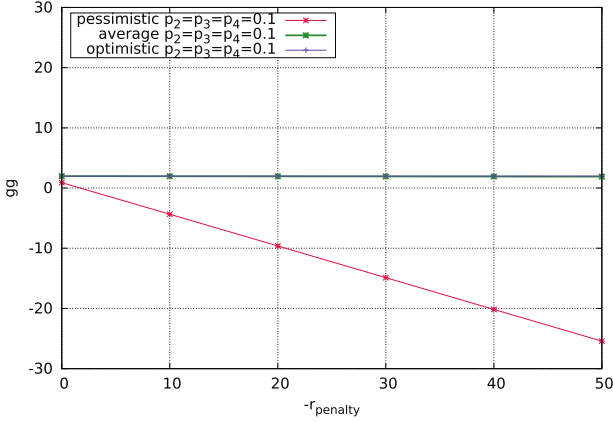


Fig. 4. Gain for the time-based composed system with varying penalties, exclusive maintenance case.

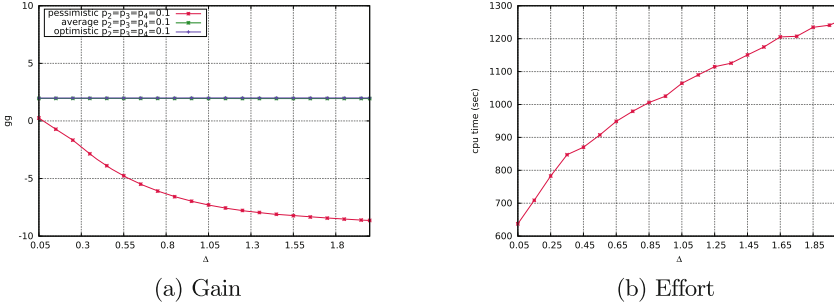


Fig. 5. Time dependent gains and CPU times for the composed system with two components, exclusive maintenance case. Computational times have been obtained from a mean of 30 independent runs.

we obtain tighter bounds of the underlying BMDP which indicates that for small inspection intervals the captured uncertainty on the internal state of a component is also small and results from uncertainty in the reward of the phase. In Fig. 5a, one can observe that a smaller time step, i. e., a smaller checking interval allows for better control of the system state and thus, for better performance of the composed system. This observation is supported by a slightly better behavior of the average MDP model. However, the price for smaller values of Δ are larger inspection costs which have to be added to the results of the optimization model.

We furthermore observe higher computation times with larger values of Δ . This includes the construction of the BMDP and MDP model as well as their optimization. The main reason for this behavior is the computational complexity of the transient analysis of reward bounds in (13).

6 Conclusion

The paper introduces a new approach to model systems of components with generally distributed failure, maintenance and repair times using Markov chains. By integrating decision points where a maintenance operation may start, a Markov decision process or a bounded parameter Markov decision process, which captures the uncertainty about the internal state of a component, can be generated and optimal maintenance policies can be computed using well established algorithms from Markov decision theory.

The approach has been described here for the composition of two components. It is not hard to extend it to more than two components. For a larger number of components the curse of dimensions resulting in the state space explosion will come up. However, the decision process which have been constructed are fairly small because they only consider the observable phase of a component and not the much larger detailed state space. For the construction of these processes larger systems of linear equations or differential equations have to be solved which can be done nowadays efficiently for relatively large state spaces.

References

1. Markov Maintenance Models Under Uncertainty Online Companion. <http://ls4-www.cs.tu-dortmund.de/cms/de/home/dohndorf/publications/>
2. Barron, Y., Frostig, E., Levikson, B.: Analysis of R out of N systems with several repairmen, exponential life times and phase type repair times: an algorithmic approach. *Eur. J. Oper. Res.* **169**(1), 202–225 (2006)
3. Bayle, T.: Preventive maintenance strategy for data centers. White Paper 124, American Power Conversation (2009)
4. Buchholz, P., Dohndorf, I., Scheftelowitsch, D.: Analysis of Markov decision processes under parameter uncertainty. In: Reinecke, P., Di Marco, A. (eds.) *EPEW 2017*. LNCS, vol. 10497, pp. 3–18. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66583-2_1
5. Buchholz, P., Kriege, J.: Markov modeling of availability and unavailability data. In: *2014 Tenth European Dependable Computing Conference*, Newcastle, United Kingdom, 13–16 May 2014, pp. 94–105. IEEE Computer Society (2014)
6. Chan, G.K., Asgarpoor, S.: Optimum maintenance policy with Markov processes. *Electr. Power Syst. Res.* **76**(67), 452–456 (2006)
7. Chen, D., Cao, Y., Trivedi, K.S., Hong, Y.: Preventive maintenance of multi-state system with phase-type failure time distribution and non-zero inspection time. *Int. J. Reliab. Qual. Saf. Eng.* **10**(3), 323–344 (2003)
8. Chen, D., Trivedi, K.S.: Optimization for condition-based maintenance with semi-Markov decision process. *Rel. Eng. Syst. Saf.* **90**(1), 25–29 (2005)
9. Crouzen, P., Pulungan, R.: Acyclic phase-type distributions in fault trees. In: *Proceedings of the International Workshop on Performability Modeling of Computer and Communication Systems* (2009)
10. Cumani, A.: On the canonical representation of homogeneous Markov processes modeling failure-time distributions. *Micorelectronics Reliab.* **22**(3), 583–602 (1982)

11. Dekker, R., Nicolai, R.P., Kallenberg, L.: Maintenance and Markov decision models. In: *Encyclopedia of Statistics in Quality and Reliability*. Wiley, Chichester (2007)
12. Endrenyi, J., et al.: The present status of maintenance strategies and the impact of maintenance on reliability. *IEEE Trans. Power Syst.* **16**(4), 638–646 (2001)
13. Faddy, M.J.: Phase-type distributions for failure times. *Math. Comput. Model.* **22**(10–12), 63–70 (1995)
14. Givan, R., Leach, S.M., Dean, T.L.: Bounded-parameter Markov decision processes. *Artif. Intell.* **122**(1–2), 71–109 (2000)
15. Hatoyama, Y.: Markov maintenance models with repair. Technical Report 175, Stanford University Department of Statistics (1976)
16. Ivy, J.S., Nembhard, H.B.: A modeling approach to maintenance decisions using statistical quality control and optimization. *Qual. Reliab. Eng. Int.* **21**(4), 355–366 (2005)
17. Kline, M.B.: Suitability of the lognormal distribution for corrective maintenance repair times. *Reliab. Eng.* **9**(2), 65–80 (1984)
18. Kuhn, K.D., Madanat, S.M.: Robust maintenance policies for Markovian systems under model uncertainty. *Comput. Aid. Civil Infrastruct. Eng.* **21**(3), 171–178 (2006)
19. Puterman, M.L.: *Markov Decision Processes*. Wiley, Boca Raton (2005)
20. Sennott, L.I.: Average reward optimization theory for denumerable state spaces. In: Feinberg, E.A., Schwartz, A. (eds.) *Handbook of Markov Decision Processes*, pp. 153–172. Kluwer, Boston (2002)
21. Thümmel, A., Buchholz, P., Telek, M.: A novel approach for phase-type fitting with the EM algorithm. *IEEE Trans. Dependable Sec. Comput.* **3**(3), 245–258 (2006)