Adaptive Maintenance Policies for Aging Devices Using a Markov Decision Process

Saranga K. Abeygunawardane, *Student Member, IEEE*, Panida Jirutitijaroen, *Senior Member, IEEE*, and Huan Xu

***Abstract*—In competitive environments, most equipment are operated closer to or at their limits and as a result, equipment's maintenance schedules may be affected by system conditions. In this paper, we propose a Markov decision process (MDP) that allows better flexibility in conducting maintenance. The proposed MDP model is based on a state transition diagram where inspection and maintenance (I&M) delay times are explicitly incorporated. The model can be solved efficiently to determine adaptive maintenance policies. This formulation successfully combines the long term aging process with more frequently observed short term changes in equipment's condition. We demonstrate the applicability of the proposed model using I&M data of local transformers.**

***Index Terms*—Asset management, backward induction, condition monitoring (CM), maintenance, Markov decision processes, transformers.**

NOMENCLATURE

S	Set of states.
n	Total number of states.
a	Action.
t	tth decision epoch.
i	Present state of the equipment.
k	State at the decision epoch $t+1$.
$\mathrm{P}_t(k\|i,a)$	Probability of transiting from state i to any state $k \in \mathrm{S}$, upon choosing action a in state i at the tth decision epoch.
$\mathrm{r}_t(i,a)$	Immediate reward for choosing action a in state i at the decision epoch t.
$\mathrm{r_N}(i)$	Boundary value of state i.
C_i	Last known condition of the equipment.
$t_{\mathrm{I},i}$	Time to perform next inspection when the last known condition is C_i (inspection delay time in C_i).
$\mathrm{t}_{\min,i}$	Minimum time between two consecutive inspections in C_i.
$\mathrm{t}_{\max,i}$	Maximum allowable time between two consecutive inspections in C_i.
τ	Interval at which I&M decision making is performed.
N	Number of decision epochs.
$n_{\mathrm{I},i}$	Number of consecutive times that the inspection is postponed.
T	Decision horizon.
a_0	Doing nothing.
a_1	Inspection/CM.
a_2	Minor maintenance.
a_3	Major maintenance.
a_4	Replacement.
a_5	Repair.
$t_{\mathrm{M},i}$	Time spent in C_i (maintenance delay time in C_i).
$\mathrm{n}_{\max,i}$	Maximum number of decision intervals that the equipment spends in C_i.
t_i	Maximum time period spent in condition C_i.
$\mathrm{U_t}(i,a)$	Total expected reward received upon choosing action a in state i at time t.
$\mathrm{U_N^*}(i)$	Maximum total expected reward in state i, at the Nth epoch.
$\mathrm{U}_{t+1}^*(k)$	The maximum total expected reward in state k, at the decision epoch $t+1$.
$\mathrm{U}_t^*(i)$	Maximum total expected reward in state i, at the tth epoch
$a_t^*(i)$	Optimal action in state i at the decision epoch t.

Manuscript received July 18, 2012; revised November 15, 2012; accepted December 17, 2012. Date of publication January 24, 2013; date of current version July 18, 2013. This work was supported by the Singapore Ministry of Education-Academic Research Fund under Grant WBS R-263-000-691-112 and Startup Grant R-265-000-384-133. Paper no. TPWRS-00841-2012.

S. K. Abeygunawardane and P. Jirutitijaroen are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore (e-mail: g0800437@nus.edu.sg; elejp@nus.edu.sg).

H. Xu is with the Department of Mechanical Engineering, National University of Singapore, Singapore (e-mail: mpexuh@nus.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2012.2237042

I. INTRODUCTION

ASSET management is essential for reliable and economic operation of power systems. With deregulation, asset management procedures became more complicated [1]. In such environments, an asset owner can perform preventive maintenance only after the independent system operator schedules a planned outage upon the request of the asset owner. In some situations, the operator may delay certain requested outages, in order to fulfill the overall aim of serving the power consumers [1]. Such situations require equipment owners to adjust their asset management plans accordingly. This highlights the need for adaptive asset management policies which can deal with maintenance delays.

Adaptive asset management policies would also be more economical than fixed policies. When the equipment is new and in good condition, too frequent inspection or CM would not reveal any additional information about the equipment's condition, and thus, unnecessarily increases the operation cost. On the other hand, when the equipment is aged or its condition is more deteriorated, delaying inspection and maintenance (I&M) may cause huge economic losses through unexpected failures. Hence, it would be more economical to perform I&M considering the equipment's age, condition, and delay times in I&M.

In literature, several mathematical models have been proposed to determine better maintenance policies for aging assets [2]–[23]. However, there are some issues which are still not addressed in these previously proposed maintenance models. First, time delays in I&M are not included in most previous models [2]–[21]. Since optimal I&M actions may depend on I&M delay times, these delays should be considered when determining optimal policies.

Second, time-based maintenance models represent equipment's deterioration by the overall condition based on the age [2]–[15], while condition-based maintenance models represent the deterioration of the equipment by some observable measurements [16]–[23]. However, the deterioration of the equipment's measurable conditions may accelerate with the aging. Thus, it is more accurate if models can integrate the deterioration of equipment's measurable conditions with effects of aging on deterioration. If a model can address the two aforementioned issues, such a model would be able to provide more adaptive I&M policies.

In this paper, we propose a new maintenance optimization model for I&M of aging equipment. The proposed decision-making model additionally considers delay times in performing I&M. Moreover, this model represents the deterioration of equipment using a quantifiable condition, while allowing the parameters of the deterioration process to vary with the operational age. Due to the above features of the model, it can provide more adaptive I&M policies which allow the asset owners to choose the optimal action, based on the knowledge about the equipment's *condition*, the operational *age*, and *time delays* in performing I&M.

The structure of this paper is as follows. In Section II, we provide the background theories and information. In Section III, the formulation of the maintenance optimization model is presented. In Section IV, the solution procedure is explained. A case study is presented in Section V to demonstrate a model application. Finally, conclusion is given in Section VI.

II. Background

Here, we first review the framework of a finite horizon discrete-time Markov decision process (MDP), with reference to [24]. Then, we discuss the decision-making process regarding I&M of equipment. Finally, we describe how this I&M decision-making process is modeled in the framework of a finite horizon discrete-time MDP.

A. Framework of an MDP

An MDP is a sequential decision-making model which considers uncertainties in outcomes of current and future decision-making opportunities. At each decision-making time, the system/equipment occupies a state. Based on this state, a decision is made by choosing an action from the set of actions associated with this state. Upon choosing an action, a reward is received and a state transition occurs from the present state to a new state, which is determined by a transition probability distribution. Since the process holds the Markov property, both transition probabilities and rewards only depend on the present state and the action chosen in the present state. As the process evolves, the decision maker receives a sequence of rewards. When choosing actions, the decision maker intends to maximize the total expected reward received over the total decision making period. If the total decision making period of an MDP is finite and the decisions are made in discrete time, the MDP is called a finite horizon discrete time MDP.

In standard practice, the decisions regarding I&M of equipment are made in discrete time. In addition, no equipment can be used over an infinitely long period, and, therefore, decisions regarding I&M of equipment are made over a finite time horizon. Due to these reasons, we represent the decision-making process of equipment's I&M using a finite horizon discrete-time MDP.

The five basic components of a finite horizon discrete-time MDP are as follows.

1) Decision epochs: Decision epochs are the point of times at which decisions are made. In a discrete-time MDP, the total decision making period (decision horizon) is divided into intervals which are called decision intervals, and at the beginning of each decision interval, a decision epoch occurs. The set of decision epochs is given by $\mathrm{D} = \{1, 2, 3, \ldots, \mathrm{N}\}$. In a finite horizon MDP, N is finite, and, according to the convention, decisions are not made at the Nth decision epoch. Decision horizon, decision intervals, and decision epochs of a discrete-time finite horizon MDP are shown in Fig. 1.
2) States: Different statuses of a system/equipment are modeled using a finite number of states.
3) Actions: Each state is connected with a finite number of actions.
4) Transition probabilities: As a result of choosing any action a connected with state i at the tth decision epoch, a state transition occurs. The new state at the decision epoch $t+1$ is determined by the probabilities of transiting from state i to possible states in the state space S. The probability of transiting from state i to any state $k \in \mathrm{S}$, upon choosing action a in state i at the tth decision epoch is denoted by $\mathrm{P}_t(k|i,a)$. It should be noted that $\sum_{k=1}^{\mathrm{n}} \mathrm{P}_t(k|i,a) = 1$.
5) Rewards: At each decision epoch $t < \mathrm{N}$, the decision maker receives a reward, as a result of choosing an action. The reward received upon choosing action a in state i at the tth decision epoch is denoted by $\mathrm{r}_t(i,a)$. The reward received at the Nth decision epoch is assigned based on the state that the equipment is being found at the Nth decision