MAINTENANCE POLICY SIMULATION FOR A FACTORED PARTIALLY OBSERVABLE SYSTEM

Demet Özgür-Ünlüakın

Department of Industrial Engineering
Işık University
34980 Şile
Istanbul, Turkey
demet.unluakin@isikun.edu.tr

İpek Kıvanç

Department of Industrial Engineering
Boğaziçi University
34342 Bebek
Istanbul, Turkey
ipek.kivanc@boun.edu.tr

ABSTRACT

Taking maintenance decisions is one of the well-known stochastic sequential decision problems under uncertainty. Partially Observable Markov Decision Processes (POMDPs) are powerful tools for such problems. Nevertheless, POMDPs are rarely used for tackling maintenance problems of multi-component systems because their state spaces grow exponentially with the increasing number of components. Factored representations have been proposed for POMDPs taking advantage of the factored structure already available in the nature of the problem. Our aim in this study is to show how to formulate a factored POMDP model for the maintenance problem of a multi-component dynamic system and how to simulate and evaluate the obtained policy before implementing it in real life. The sensitivity of the methodology is analyzed under several cost values, and the methodology is compared to other predefined policies. The results show that the policies generated via the POMDP solver perform better than the predefined policies.

Keywords: maintenance policy, POMDP, DBN, factored representation, sensitivity analysis.

1 INTRODUCTION

All kinds of systems (machine, equipment, vehicle, etc.) used in the manufacturing sector have a certain life span and they need maintenance during their life cycles. As technology evolves, systems become more complex, having more components and thus causing the plan of the maintenance activities harder. That is why maintenance management is a very important and critical process for companies with multi-component systems. Since there are different maintenance procedures for each component, selecting a maintenance policy for the entire system is difficult.

Markov decision processes (MDPs) are used to make sequential decisions under uncertainty in decision problems where system states can be fully observed throughout the planning horizon (Puterman 2014). Nevertheless, MDP models are inadequate, especially for real-life problems where the system state is not fully observable. In such systems, the state of the system can be estimated via observations which are collected with some signals or measurements. However, the accuracy of these observations is also probabilistic. For this reason, MDPs have been extended to Partially Observable Markov Decision Processes (POMDP), which can take into account the cost of information, with better results in planning optimal policies under uncertainty (Papakonstantinou and Shinozuka 2014a).

POMDPs provide a rich and powerful framework for tasks that require decision making under uncertainty. Thus, POMDPs are applied in a wide range of different real-world problems. One of the most widely used areas is robotics which includes navigation and localization of robots (Pineau and Gordon 2007, Roy, Gordon, and Thrun 2005, Foka and Trahanias 2007) and adaptive sensing (Hoey and Little 2007, Spaan, Veiga, and Lima 2010). Another application area of POMDPs is health informatics. POMDPs are applied in treatment prescription of heart disease (Hauskrecht and Fraser 2000), kinds of cancer (Maillart, Ivy, Ransom, and Diehl 2008, Ayer et al. 2010), Parkinson's disease (Goulionis and Vozikis 2009) and assistance of disabled people (Mihailidis, Boger, Craig, and Hoey 2008).

POMDPs are less popular in the domain of maintenance than other probabilistic graphical approaches because of their high operational complexity. However, POMDPs offer state-based maintenance policies that minimize operating costs and maximize the machine's availability. Early theoretical works studied by White (1976) and Rosenfield (1976) are pioneering in POMDP modeling. Monahan's survey (Monahan 1982) provides an overview of the models and algorithms dealing with POMDPs. Machine replacement is also discussed within the scope of this study. After these contributions, in literature, the number of articles that provide theoretical frameworks for the management of the maintenance has increased (Madanat 1993, David, Friedman, and Sinuany-Stern 1999, Makis and Jiang 2003).

Madanat and Ben-Akiva (1994) adopted POMDPs for the maintenance decision-making of highway-pavement networks. Ellis, Jiang, and Corotis (1995) and Corotis, Hugh Ellis, and Jiang (2005) showed how to formulate a POMDP model to plan the for bridge inspection. POMDPs are widely used for infrastructure maintenance and inspection scheduling. Papakonstantinou and Shinozuka (2014a) performed an extensive literature review in the area of inspection scheduling and maintenance planning using dynamic programming and POMDPs. Byon, Ntaimo, and Ding (2010) obtained seasonal dependent situation-based maintenance policies for wind turbines with finite horizon. Papakonstantinou and Shinozuka (2014b) obtained maintenance and inspection policies for a corroding reinforced concrete structure. AlDurgam and Duffuaa (2013) used POMDP models to create policy graphs that allow the operator to choose the optimal maintenance action and speed setting.

Two main drawbacks of classical POMDPs are "curse of dimensionality" and "curse of history". Size of the state space grows exponentially with the number of states, and also the number of action-observation histories grows exponentially with the planning horizon (Pineau, Gordon, Thrun, et al. 2003). Thus, classical POMDPs are limited to solve problems with large state spaces. Some approaches have been proposed in the literature to overcome these obstacles. In order to address the curse of history, factored representations have been proposed. In a classical POMDP representation, the system is represented by a single node which has multiple states. However, factored POMPDs are efficient tools for reducing the computational complexity since variables are represented compactly via data structure representations such as Decision trees (DTs) (Boutilier and Poole 1996) and Algebraic Decision Diagrams (ADDs) (Hansen and Feng 2000). ADDs have the conditional independence and context-specific independence that allow representing probability and utility tables in smaller sizes (Jesse Hoey and Robert St-Aubin and Alan J Hu and Craig Boutilier 1999). We used Symbolic Perseus (SP) which is one of the ADD-based algorithms. It includes an adapted variant of the point-based value iteration algorithm to solve the factored POMDP models (Poupart 2005).

Bayesian Networks (BNs) are directed acyclic probabilistic graphics to model the dependencies among variables (Nielsen and Jensen 2009). BNs have been used to explain the dependency between the parts of the device that has failed in decision-theoretic troubleshooting problems (Heckerman, Breese, and Rommelse 1994), especially in the late 1980s. Reliability is another area where BNs are used frequently (Langseth and Portinale 2007). Dynamic Bayesian networks (DBNs) are extended BNs including the temporal dimension (Murphy 2002). Recently DBNs are also used for dependability, risk analysis and maintenance areas (Weber et al. 2012). Since actions and costs cannot be explicitly represented in DBNs, they are generally used for prognosis purposes (Portinale, Raiteri, and Montani 2010).

Our aim in this study is to show how to formulate a factored POMDP model for the maintenance problem of a multi-component dynamic system and how to simulate and evaluate the generated policy before implementing it in real life. The organization of the paper is as follow: Section 2 gives the methodology used in the study. The empirical POMDP model is presented in Section 3. We explain the proposed simulation environment in Section 4 and give the replication results in Section 5. Finally, Section 6 concludes the study and gives future research directions.

2 METHODOLOGY

The maintenance problem is first formulated as a POMDP and a maintenance policy is obtained using Symbolic Perseus solver. The policy is then simulated using DBNs to make sensitivity analysis under several cost parameters.

2.1 POMDPs

In POMDP models, unlike MDPs, the agent cannot fully observe the state of the system. However, the agent knows the probabilities of being in a state where he/she may be after his/her action from his/her current state. These probabilities are called "Belief State Probabilities". In addition, when the agent takes an action, that is to say, he/she passes to another belief state, he/she obtains an observation. This observation affects belief state probabilities at the next step. In other words, the agent has obtained a hint for the next state.

2.1.1 POMDP Definition

A POMDP model is defined by the 6-tuple. s_t, a_t and o_t are the system state, the action state and the received observation at time t respectively. r_t represents the total reward incurred between times t and $t+1$. O and T are the observation and transition functions respectively. A typical POMDP is illustrated in Figure 1.

- **States**: System state space S is a finite set of states. A state is the definition of the environment at any point in the horizon. The relationship between states is Markovian. The fact that a process is Markovian requires that the next state depends solely on the current state, regardless of past states. $S = \{S_0, S_1, \ldots, S_N\}$ is a finite state set. In this case, the Markov property is as follows: $Pr(s_{t+1}|s_0, s_1, \ldots, s_t) = Pr(s_{t+1}|s_t)$.
- **Actions**: Action state space A is a set of all alternative actions that can be chosen. The agent's main goal is maximizing his/her reward by taking correct actions.
- **Observations** Observation state space θ is the all possible observations that the agent can observe.
- **Transition Function**: $T : SxAxS \rightarrow \Delta(S)$ is the function of the transition probabilities of the process. When the current state s and the selected action a are known, this is the function that gives all possible probability of transitions to the next state s'.
- **Observation Function**: $O : SxAx\theta \rightarrow \Delta(O)$ is a function of the state of the process and action-based observations' probabilities. It gives the probabilities of obtaining each observation at the new state where the agent passes by an action from the previous state.
- **Reward Function**: R: $SxAxSx\theta \rightarrow \Re$ is the function of calculating rewards due to state transitions, actions and observations.

Let $V^*(s)$ be the optimal value function maximizing the expected total revenue of the process in the long-run for state s, and s' denote the next process state. Dynamic programming allows determining the long-term value $V^*(s)$ for each (discrete) state s and is summarized by the Bellman Equation which is given in (1)

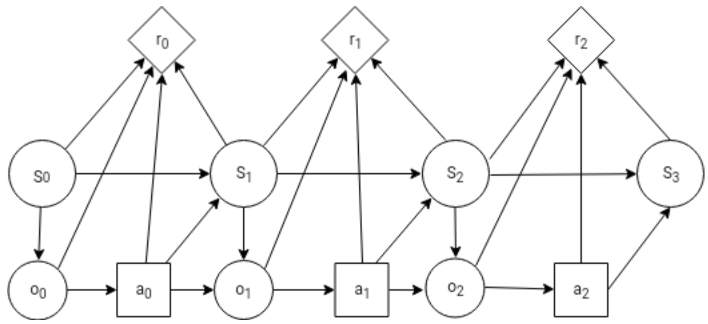

Figure 1: A typical POMDP representation.

where $0 \leq \gamma \leq 1$ is the discount factor.

$$V^*(s) = \max_{a \in A} \left\{ \sum_{s' \in S} P(s'|s,a)R(s,a,s') + \gamma \sum_{s' \in S} P(s'|s,a)V(s') \right\}, \forall s \in S \tag{1}$$

2.1.2 Belief States

In POMDP models, the agent is unable to observe the hidden state of the environment, but it is possible to estimate it with some observations. That is, the agent has a belief in the current situation and he/she has to make all decisions according to this incomplete and partial information. The agent remembers all his/her observations and actions in the past. This is called “all history” and it is difficult to process and keep it since it expands over time. Instead of keeping all history, an alternative representation has been developed that assigns probabilities to every possible state which are called “belief states". Belief state is sufficient to summarize all the history without compromising its optimality with a proper probability distribution over the state space (Aström 1965). Thus, POMDPs become belief-state MDPs, which are a special case of continuous-state MDP through a continuous space containing the probability values instead of discrete state space. Let b be this probability distribution over states. So, b is a vector of probabilities, one for each state: $b = [b(s), s \in S]$ where $b(s)$ is the probability assigned to state s. The value function in (1) can be calculated on the belief space as in (2).

$$V^*(b) = \max_{a \in A} \left\{ \sum_{s \in S} \rho(s,a)b(s) + \gamma \sum_{o \in \Theta} \sum_{s \in S} P(o|s,a)b(s)V^*(b') \right\} \tag{2}$$

Belief space is continuous; however, for a finite horizon, the optimal value function is piecewise-linear and convex. Thus, any finite-horizon solution is represented by a limited set of alpha-vectors due to this property (Sondik 1978). Alpha-vectors are a set of hyperplanes that define belief functions. For each belief point, the value function is equal to the hyperplane with the highest value. The value function is calculated on the belief space as in (3) where $\alpha(s)$ represents the element of the alpha-vector corresponding to state s.

$$V^*(b) = max_{\alpha} \left[\sum_{s} \alpha(s)b(s) \right] \tag{3}$$

2.2 DBNs

A DBN is composed of several time slots each of which includes the same BN. Let T and N be the number of time slots and random variables in a time slot respectively; X_t represent all variables each of which is denoted by X_t^i in time slot t. All parents of X_t^i and all variables in the DBN are represented by $Pa(X_t^i)$ and $X_{1:T}$ respectively. The joint probability distribution of the variables in a DBN can be calculated as in (4).

$$P(X_{1:T}) = \prod_{t=1}^{T}\prod_{i=1}^{N} P(X_t^i|Pa(X_t^i)) \tag{4}$$

3 THE EMPIRICAL POMDP MODEL

In this study, an empirical system having four components, three processes, and one observation node is constructed as a POMDP model. The relationships between components and processes are shown in Figure 2. There are three states of all components {W, D, NW} and the observation node {G, Y, R}, and all processes have two states {W, NW} where W, D, NW stand for working, degrading, nonworking respectively, and G, Y, R stand for green, yellow, red respectively. It is assumed that all components (C1, C2, C3, C4) can be changed at any time. In addition, it is not possible to observe the components and processes directly. However, the state of processes and components can be estimated by the observation node. All components deteriorate over time. The processes are defined as the result of the interaction between their parents. The main process node P3 is directly connected to the observable node used to collect information from the process node. At any time point, at most one of the components can be replaced. Hence the action node has five states which are {Do nothing, Replace C1, Replace C2, Replace C3, Replace C4}. The POMDP model for the empirical system is given in Figure 2 for two time slots. Initial states of all components are set to the working state.

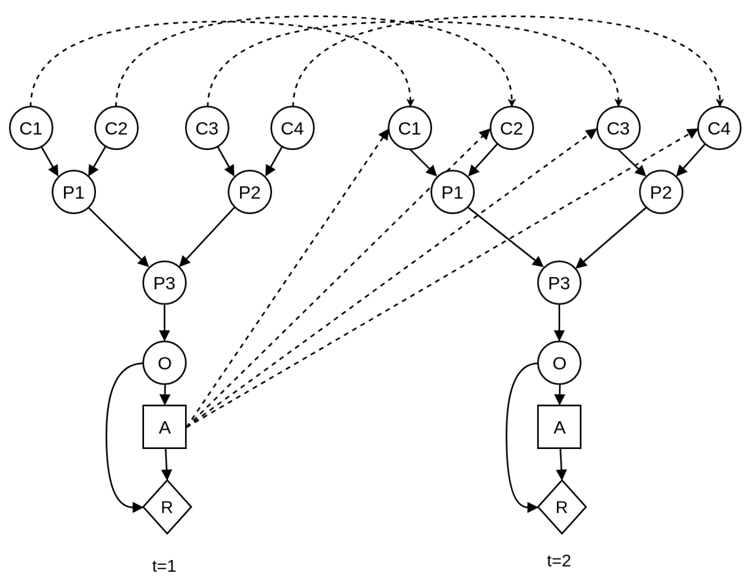

Figure 2: Empirical POMDP model for two time slots.

Rewards (costs in this study) are collected by two means: actions and observations. Maintenance costs incurred at each t depend on the observations received and the maintenance action performed. The total maintenance cost at time t consists of the cost of production loss and the replacement cost if a component is replaced at that time slot. Replacement costs of C1, C2, C3, and C4 are {100, 200, 300, 400} when a green or yellow signal is observed; and {200, 400, 600, 800} when a red signal is observed respectively. On the other side, a downtime cost of 2,500 incurs during the maintenance when a green or yellow signal is observed; and the system is assumed to have a downtime cost of 7,500 when a red signal is received. Maintenance costs of each component including both the cost of production losses depending on the observations received and also the cost of the action taken are given in Table 1.

Table 1: Maintenance costs.

	Observation		
Action	**G**	**Y**	**R**
Do nothing	0	0	7500
Replace C1	2600	2600	7700
Replace C2	2700	2700	7900
Replace C3	2800	2800	8100
Replace C4	2900	2900	8300

The BN of the empirical model at a time slot is given in Figure 3. In order to reduce the inference burden of the DBN, the non-temporal and non-replaceable intermediate nodes P1, P2 and P3 are eliminated (absorbed) from the BN, without affecting anything, with arc reversal and node elimination techniques used in graphical models. The reduced BN after absorption is depicted in Figure 4.

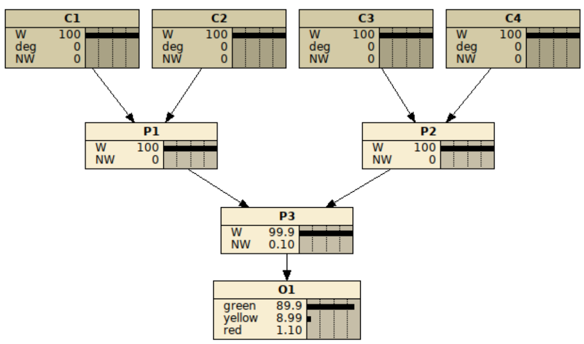

Figure 3: BN at a time slot.

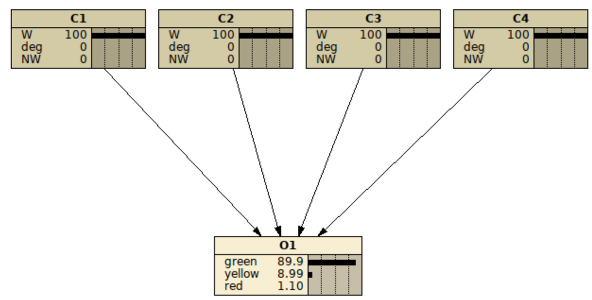

Figure 4: BN at a time slot after absorption.

4 MAINTENANCE POLICY SIMULATION

The empirical maintenance problem is first formulated as a factored POMDP and it is solved for a maintenance policy using Symbolic Perseus (SP), a factored POMDP solver (Poupart 2005). The solver returns approximate policies for large space factored POMDP models. The policy is then simulated using DBNs (Özgür-Ünlüakın and Bilgiç 2017) in order to evaluate its performance before implementing it in real life and also to make sensitivity analysis under different cost parameters. We use BNT toolbox (Murphy 2001) to construct the DBN model and to compute the required inferences such as sampling the observations and inferring the belief states. Policy simulations are performed within the Matlab environment. The algorithm of the maintenance policy simulation is given in Algorithm 1 where π is the policy vector obtained from the POMDP solution of which the size is $|\pi|$ which is the same size of the set of alpha-vectors returned by the solver.

5 COMPUTATIONAL RESULTS

The planning horizon is taken as 100 days and the simulation is replicated 30 times. Results of the sensitivity analysis with respect to several downtime cost values incurred at a red signal are given in Table 2 where TCost, TRed, and TRep denote the total cost, the total number of red signals obtained and the total number of replacements respectively in 100 days of a horizon. Average and standard deviation of these measures are reported in the table. Furthermore, the average total replacements for each component is also given. According to the table, when the red signal incurs the same downtime cost as of the yellow or green signal,

Algorithm 1 Pseudocode of the Maintenance Policy Simulation.

Input π ▷ Obtained from POMDP solution
Input α vectors ▷ Obtained from POMDP solution
Set $t = 1$
Set $\varepsilon = [\,]$
Infer initial belief b_1 by computing the joint probability $P(C_{1,1}, C_{2,1}, C_{3,1}, C_{4,1}|\varepsilon)$
for t=1:T **do**
    Determine $j^* = \underset{j\in\{1..|\pi|\}}{\arg\max} \left\{\sum_s \alpha_j(s) b_t(s)\right\}$ ▷ Obtain the index j from b_t and α vectors
    Determine $a_t = \pi(j^*)$ ▷ Obtain the action from π
    Calculate $Cost_t$
    Update $\varepsilon \leftarrow \varepsilon \cup \{A_t = a_t\}$ ▷ Update evidence of the action node
    Sample observation node o_{t+1}
    Update $\varepsilon \leftarrow \varepsilon \cup \{O_{t+1} = o_{t+1}\}$ ▷ Update evidence of the observation node
    Infer next belief b_{t+1} by computing the joint probability $P(C_{1,t+1}, C_{2,t+1}, C_{3,t+1}, C_{4,t+1}|\varepsilon)$
end for

that is 2,500, SP policy behaves like a corrective maintenance strategy. In other words, replacements are done whenever the red signal is observed. Another important finding of these replication results is that as downtime cost increases, the average total number of red signals decreases whereas the average total number of replacements increases. It can be said that more replacements are performed proactively before waiting for a red signal as downtime cost increases. However, the total number of red signals reaches a steady state.

Table 2: Sensitivity analysis of SP policy under different downtime costs.

Downtime	TCost		TRed		TRep		Avg. Comp. Replacements			
Cost	Avg	Std	Avg	Std	Avg	Std	C1	C2	C3	C4
2,500	70,903	13,336	24.37	4.57	24.37	4.57	11.17	5.27	3.50	4.43
5,000	137,867	22,701	24.40	4.12	26.50	4.36	11.00	6.00	4.17	5.33
7,500	191,370	32,855	20.10	3.76	32.30	5.16	12.43	6.93	6.43	6.50
10,000	228,057	48,583	16.67	4.16	36.90	6.63	12.73	7.33	8.30	8.53
12,500	261,460	42,592	15.43	2.94	38.57	5.19	12.80	9.47	8.23	8.07
15,000	309,673	65,810	15.57	4.30	41.30	5.11	12.50	9.53	9.83	9.43
17,500	369,353	63,754	16.03	3.34	46.37	5.49	13.63	11.53	11.03	10.17
20,000	397,827	66,527	15.23	3.29	47.23	4.24	13.60	11.93	10.80	10.90

The average total number of red signals vs the average total number of replacements for different downtime cost values are depicted in Figure 5 where it is seen obviously that TRep increases as downtime cost increases. On the other side, TRed first decreases, but then is not affected by the increase in downtime cost.

In order to analyze the performance of the policy generated by the POMDP solver, we develop four predefined straightforward maintenance policies where the components are maintained randomly or in order when a red or a yellow signal (in addition to the red signal) is observed. The results are given for the downtime cost of 7,500 in Table 3 where R stands for doing corrective maintenance with a red signal and RY stands for doing maintenance with red and yellow signals. Random methods are the worst as expected. The result of the strategy belonging to the case where the components are selected in order at a time when a red or yellow signal is observed with an average total cost value of 204,100 is close to the result of the SP policy which is 191,370.

Further comparisons are done for extreme downtime cost values which are 2,500 and 15,000 of which the results are depicted in Tables 4 and 5 respectively. When the red signal incurs a downtime cost of 2,500, the total cost of the R-Order policy is close to the cost of the SP policy. This is expected since SP policy

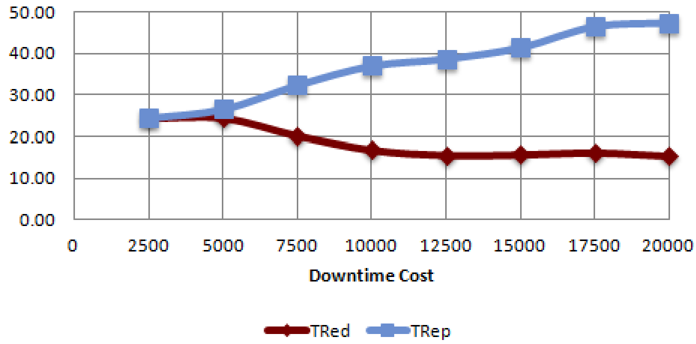

Figure 5: Tred vs TRep.

Table 3: Comparison with predefined maintenance policies when downtime cost=7,500.

	TCost		**TRed**		**TRep**		**Avg. Comp. Replacements**			
Policy	**Avg**	**Std**	**Avg**	**Std**	**Avg**	**Std**	**C1**	**C2**	**C3**	**C4**
R-Random	256,050	50,897	32.03	6.32	32.03	6.32	8.00	8.50	8.13	7.40
R-Order	215,523	39,817	26.97	4.97	26.97	4.97	7.13	6.77	6.57	6.50
RY-Random	267,500	56,199	31.23	7.30	37.67	6.49	9.77	9.20	9.30	9.40
RY-Order	204,100	32,623	23.17	4.43	30.07	4.16	7.97	7.63	7.30	7.17
SP	191,370	32,855	20.10	3.76	32.30	5.16	12.43	6.93	6.43	6.50

behaves like a corrective maintenance strategy when the downtime cost is 2,500 as understood from the equality of the values of TRed and TRep in Table 4. TRed and TRep values of the SP policy are a bit higher than that of the R-Order policy because of the difference in the distribution of the component selections for maintenance. SP policy maintains C1 more than the others, while R-Order leads a balanced selection policy. R-Random also applies a corrective maintenance policy, but since it selects the components randomly at a maintenance time, it gives a much higher cost value than that of the SP policy.

Table 4: Comparison with predefined maintenance policies when downtime cost=2,500.

	TCost		**TRed**		**TRep**		**Avg. Comp. Replacements**			
Policy	**Avg**	**Std**	**Avg**	**Std**	**Avg**	**Std**	**C1**	**C2**	**C3**	**C4**
R-Random	102,550	22,864	34.23	7.47	34.23	7.47	8.83	8.93	7.73	8.73
R-Order	79,933	15,985	26.73	5.33	26.73	5.33	7.13	6.73	6.56	6.30
RY-Random	110,017	20,838	30.57	8.01	37.03	6.85	8.20	9.10	9.20	10.53
RY-Order	97,250	14,194	26.00	5.44	33.07	4.76	8.66	8.36	8.16	7.86
SP	70,903	13,336	24.37	4.57	24.37	4.57	11.17	5.27	3.50	4.43

When the downtime cost of the red signal is increased to 15,000, of which the results are seen in Table 5, SP policy performs maintenance proactively before waiting for a red signal. None of the predefined policies result in a total cost value which is close to the total cost of the SP policy. The nearest value belongs to RY-Order policy since it behaves also proactively and selects the components in order. Although RY-Random applies a proactive maintenance policy, it gives a much higher cost value than that of the SP policy since it selects the components randomly at a maintenance time. The cost of RY-Random is also considerably higher than that of the corrective R-Order policy which indicates that selecting the component effectively is as important as deciding on the maintenance time in determining an efficient maintenance policy.

Table 5: Comparison with predefined maintenance policies when downtime cost=15,000.

Policy	TCost		TRed		TRep		Avg. Comp. Replacements			
	Avg	**Std**	**Avg**	**Std**	**Avg**	**Std**	**C1**	**C2**	**C3**	**C4**
R-Random	524,873	96,306	33.87	6.21	33.87	6.21	8.53	8.53	8.43	8.36
R-Order	439,953	63,133	28.40	4.07	28.40	4.07	7.53	7.16	6.90	6.80
RY-Random	518,843	107,006	32.50	7.42	38.03	6.97	9.10	10.56	9.46	8.90
RY-Order	386,863	80,810	23.77	5.19	30.57	5.70	8.06	7.66	7.50	7.33
SP	309,673	65,810	15.57	4.30	41.30	5.11	12.50	9.53	9.83	9.43

6 CONCLUSION

We formulate a factored POMDP model for the maintenance problem of a four-component partially observable dynamic system and obtain an approximate maintenance policy. We develop a maintenance policy simulator using DBNs to evaluate the policies in a finite horizon before implementing them in real life. The sensitivity of the methodology is also analyzed under several downtime cost values. We compare the performance of the proposed POMDP policy with the other predefined straightforward policies. The results show that as downtime cost of the red signal increases, proactive maintenance should be considered, and hence the obtained POMDP policy becomes more significant since the maintenance actions are determined adaptively using the information provided by the observation and action history. On the other hand, determining the right component is also important once maintenance is decided. As a result, policies generated via the POMDP solver perform better than the predefined policies. As a future study, the work can be extended for more complex real-life problems.

ACKNOWLEDGMENTS

This study is supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under grant no 117M587.

REFERENCES

AlDurgam, M. M., and S. O. Duffuaa. 2013. “Optimal joint maintenance and operation policies to maximise overall systems effectiveness”. *International Journal of Production Research* vol. 51 (5), pp. 1319–1330.

Aström, K. J. 1965. “Optimal Control of Markov Decision Processes with Incomplete State Estimation”. *J. Math. Anal. Appl.* vol. 10, pp. 174–205.

Ayer, T., J. Chhatwal, O. Alagoz, C. E. Kahn Jr, R. W. Woods, and E. S. Burnside. 2010. “Comparison of logistic regression and artificial neural network models in breast cancer risk estimation”. *Radiographics* vol. 30 (1), pp. 13–22.

Boutilier, C., and D. Poole. 1996. “Computing Optimal Policies for Partially Observable Decision Processes Using Compact Representations”. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence - Volume 2*, AAAI’96, pp. 1168–1175, AAAI Press.

Byon, E., L. Ntaimo, and Y. Ding. 2010. “Optimal maintenance strategies for wind turbine systems under stochastic weather conditions”. *IEEE Transactions on Reliability* vol. 59 (2), pp. 393–404.

Corotis, R. B., J. Hugh Ellis, and M. Jiang. 2005. “Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision processes”. *Structure and Infrastructure Engineering* vol. 1 (1), pp. 75–84.

David, I., L. Friedman, and Z. Sinuany-Stern. 1999. “A simple suboptimal algorithm for system maintenanceunder partial observability”. *Annals of Operations Research* vol. 91, pp. 25–40.

Ellis, H., M. Jiang, and R. B. Corotis. 1995. "Inspection, maintenance, and repair with partial observability". *Journal of Infrastructure Systems* vol. 1 (2), pp. 92–99.

Foka, A., and P. Trahanias. 2007. "Real-time hierarchical POMDPs for autonomous robot navigation". *Robotics and Autonomous Systems* vol. 55 (7), pp. 561–571.

Goulionis, J. E., and A. Vozikis. 2009. "Medical decision making for patients with Parkinson disease under Average Cost Criterion". *Australia and New Zealand health policy* vol. 6 (1), pp. 15.

Hansen, E. A., and Z. Feng. 2000. "Dynamic Programming for POMDPs Using a Factored State Representation". In *AIPS*.

Hauskrecht, M., and H. Fraser. 2000. "Planning treatment of ischemic heart disease with partially observable Markov decision processes". *Artificial Intelligence in Medicine* vol. 18 (3), pp. 221–244.

Heckerman, D., J. Breese, and K. Rommelse. 1994. "Troubleshooting under uncertainty". Technical report, Technical Report MSR-TR-94-07, Microsoft Research.

Hoey, J., and J. J. Little. 2007. "Value-directed human behavior analysis from video using partially observable Markov decision processes". *IEEE transactions on pattern analysis and machine intelligence* vol. 29 (7), pp. 1118–1132.

Jesse Hoey and Robert St-Aubin and Alan J Hu and Craig Boutilier 1999. "SPUDD: Stochastic Planning Using Decision Diagrams".

Langseth, H., and L. Portinale. 2007. "Bayesian networks in reliability". *Reliability Engineering & System Safety* vol. 92 (1), pp. 92–108.

Madanat, S. 1993. "Incorporating inspection decisions in pavement management". *Transportation Research Part B: Methodological* vol. 27 (6), pp. 425–438.

Madanat, S., and M. Ben-Akiva. 1994. "Optimal inspection and repair policies for infrastructure facilities". *Transportation science* vol. 28 (1), pp. 55–62.

Maillart, L. M., J. S. Ivy, S. Ransom, and K. Diehl. 2008. "Assessing dynamic breast cancer screening policies". *Operations Research* vol. 56 (6), pp. 1411–1427.

Makis, V., and X. Jiang. 2003. "Optimal replacement under partial observations". *Mathematics of Operations Research* vol. 28 (2), pp. 382–394.

Mihailidis, A., J. N. Boger, T. Craig, and J. Hoey. 2008. "The COACH prompting system to assist older adults with dementia through handwashing: An efficacy study". *BMC geriatrics* vol. 8 (1), pp. 28.

Monahan, G. E. 1982. "State of the art—a survey of partially observable Markov decision processes: theory, models, and algorithms". *Management Science* vol. 28 (1), pp. 1–16.

Kevin Murphy 2001. "Bayes net toolbox for Matlab".

Murphy, K. P. 2002. *Dynamic Bayesian Networks: Representation, Inference and Learning*. Ph. D. thesis, University of California, Berkeley. AAI3082340.

Nielsen, T. D., and F. V. Jensen. 2009. *Bayesian networks and decision graphs*. Springer Science & Business Media.

Özgür-Ünlüakın, D., and T. Bilgiç. 2017. "Performance analysis of an aggregation and disaggregation solution procedure to obtain a maintenance plan for a partially observable multi-component system". *Reliability Engineering & System Safety* vol. 167, pp. 652–662.

Papakonstantinou, K. G., and M. Shinozuka. 2014a. "Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part I: Theory". *Reliability Engineering & System Safety* vol. 130, pp. 202–213.

Papakonstantinou, K. G., and M. Shinozuka. 2014b. "Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation". *Reliability Engineering & System Safety* vol. 130, pp. 214–224.

Pineau, J., G. Gordon, S. Thrun et al. 2003. "Point-based value iteration: An anytime algorithm for POMDPs". In *IJCAI*, Volume 3, pp. 1025–1032.

Pineau, J., and G. J. Gordon. 2007. "POMDP planning for robust robot control". In *Robotics Research*, pp. 69–82. Springer.

Portinale, L., D. C. Raiteri, and S. Montani. 2010. "Supporting reliability engineers in exploiting the power of dynamic Bayesian networks". *International journal of approximate reasoning* vol. 51 (2), pp. 179–195.

Poupart, P. 2005. *Exploiting structure to efficiently solve large scale partially observable Markov decision processes*. Citeseer.

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

Rosenfield, D. 1976. "Markovian deterioration with uncertain information". *Operations Research* vol. 24 (1), pp. 141–155.

Roy, N., G. Gordon, and S. Thrun. 2005. "Finding approximate POMDP solutions through belief compression". *Journal of artificial intelligence research* vol. 23, pp. 1–40.

Sondik, E. J. 1978, April. "The Optimal Control of Partially Observable Markov Processes over the Infinite Horizon: Discounted Costs". *Oper. Res.* vol. 26 (2), pp. 282–304.

Spaan, M. T., T. S. Veiga, and P. U. Lima. 2010. "Active cooperative perception in network robot systems using POMDPs". In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4800–4805. IEEE.

Weber, P., G. Medina-Oliva, C. Simon, and B. Iung. 2012. "Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas". *Engineering Applications of Artificial Intelligence* vol. 25 (4), pp. 671–682.

White, C. C. 1976. "Procedures for the solution of a finite-horizon, partially observed, semi-Markov optimization problem". *Operations Research* vol. 24 (2), pp. 348–358.

AUTHOR BIOGRAPHIES

DEMET ÖZGÜR-ÜNLÜAKIN is an Assistant Professor in the Department of Industrial Engineering at Isik University, Istanbul, Turkey. She holds a Ph.D. degree in Industrial Engineering from Bogazici University, Istanbul, Turkey. Her research interests include maintenance, decision making under uncertainty, statistical analysis, and applied operations research. Her email address is demet.unluakin@isikun.edu.tr.

İPEK KIVANÇ is a graduate student in the Department of Industrial Engineering at Bogazici University, Istanbul, Turkey. She holds a Bachelor of Science degree in Industrial Engineering from Istanbul Technical University, Istanbul, Turkey. She currently works in Department of Industrial Engineering, Işık University as a researcher. Her research interests include maintenance, partially observable Markov decision processes, decision making under uncertainty. Her email address is ipek.kivanc@boun.edu.tr.