Computers and Chemical Engineering 112 (2018) 239–252

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

A POMDP framework for integrated scheduling of infrastructure maintenance and inspection

Jong Woo Kim, Go Bong Choi, Jong Min Lee*

School of Chemical and Biological Engineering, Institute of Chemical Processes, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

ARTICLE INFO

Article history:
Received 10 October 2017
Revised 12 February 2018
Accepted 18 February 2018
Available online 19 February 2018

Keywords:
Infrastructure scheduling
Sensor scheduling
POMDP
Asset optimization
Stochastic optimization
Survival analysis

ABSTRACT

This work presents an optimization scheme for maintenance and inspection scheduling of the infrastructure system whose states are nearly impossible or prohibitively expensive to estimate or measure online. The suggested framework describes state transition under the observation uncertainty as Partially Observable Markov Decision Process (POMDP) and can integrate heterogeneous scheduling jobs including maintenance, inspection, and sensor installation within a single model. The proposed approach performs survival analysis to obtain time-variant transition probabilities. A POMDP problem is then formulated via state augmentation. The resulting large-scale POMDP is solved by an approximate point-based solver. We exploit the idea of receding horizon control to the POMDP framework as a feedback rule for the online evaluation. Water distribution pipeline is analyzed as an illustrative example, and the results indicate that the proposed POMDP framework can improve the overall cost for maintenance tasks and thus the system's sustainability.

1. Introduction

According to the report of the America's infrastructure grades in 2013 (Herrmann, 2013), the overall grade of the infrastructure was diagnosed as D+ (poor) and the required investment in infrastructure upgrades and maintenance by 2020 was estimated to be \$3.6 trillion. Infrastructure scheduling plays a critical role in ensuring safe operation and economic maintenance for chemical processes and process system peripherals. Infrastructure scheduling includes three major tasks: maintenance, inspection, and sensor-installation. Maintenance is done to improve the overall functionality of the system such as grade and failure rate. Inspection is carried out to assess the current condition and gather the information of the system. Sensor-installation allows inspection using a device rather than human senses. Proper and timely installation of sensors can enhance the quality of service (QoS) of a system by structural health monitoring (SHM) (Flammini et al., 2010; Jawhar et al., 2007).

Infrastructure has two distinguishing characteristics which hinder rational decision-making. First, it deteriorates with a low failure rate whereas serious damage to a wide area is inevitable when it fails. Second, it is nearly impossible to measure or estimate the system's state in real time due to the size and complexity of the system, the high cost of scheduling actions, and the intrinsic uncertainty of non-destructive inspection methods.

A scheduling approach based on deterioration model and optimization can be a promising tool to address these two issues, and thus can lead to economical and sustainable operation. In the case of the infrastructure network, a prioritization method can be applied to narrow the scope of the scheduling problem into the single infrastructure system (Choi et al., 2017; Memarzadeh and Pozzi, 2016). The majority of deterioration models for infrastructure system consider uncertainties due to the lack of knowledge about fundamental principles and the limited amount of available data. Discrete-time states and actions are usually appropriate for stochastic optimization models. The discrete system state is often designated as a grade or condition of the system, and its validity is discussed in Madanat et al. (1995). For this reason, the infrastructure scheduling problem can be regarded as a discrete stochastic sequential decision process. An appropriate framework for this problem is Markov Decision Process (MDP), whose objective is to calculate an optimal maintenance schedule within a control horizon (Puterman, 2014).

State transition randomness has been successfully modeled in the MDP framework via probability matrix (Golabi et al., 1982; Guignier and Madanat, 1999). However, observation uncertainty leads to suboptimal solutions to MDP. Moreover, inspection and

* Corresponding author.
E-mail address: jongmin@snu.ac.kr (J.M. Lee).

https://doi.org/10.1016/j.compchemeng.2018.02.015
0098-1354/

sensor-installation, which are other essential parts of the infrastructure scheduling, cannot be considered in the MDP framework which inherently assumes the 'fully observable state'. However, Partially Observable MDP (POMDP) formalism, which is a natural extension of MDP and involves probability of state observation, can tackle the problem of the uncertainty in observation and allow for integrating inspection scheduling and sensor-installation.

Several studies of the infrastructure scheduling using POMDP model have been conducted. Small-scale problems with the number of states less than 10 are solved with POMDP formulation (Ellis et al., 1995; Jiang et al., 2000; Madanat and Ben-Akiva, 1994). Jiang et al. (2000) suggested failure criteria by comparing minimum resistance and maximum loading effect on the system. Byon and Ding (2010) solved a finite horizon POMDP problem with season-dependent parameters.

The history-dependent and time-variant transition process improves the accuracy of the model because it can take account of maintenance records and system's age. The time-variant transition model can be obtained by survival analysis with the semi-Markov assumption (Kleiner, 2001). History-dependency can be modeled by the state-augmentation (Robelin and Madanat, 2007) or by the concept of periodic replacement (Kim et al., 2016). In Papakonstantinou and Shinozuka (2014) a large-scale POMDP problem is solved for the system having the aforementioned issues. Whereas the large-scale POMDP could describe more practical systems, the previous studies could not either consider the sensor installation or revise the resulting policy when feedback information is available.

The main contribution of this paper is to suggest a POMDP framework for the infrastructure scheduling with the following features: First, sensor-installation is considered as a part of the integrated scheduling, as an improvement to the maintenance and inspection-only scheduling. This can be achieved by introducing a binary state that indicates the status of sensor-installation. Second, the resulting policy derived from the POMDP is implemented in a receding horizon control (RHC) fashion so that newly obtained observations can be incorporated into the model. A step-by-step description of the formulation and solution procedure is also presented.

The remainder of this paper is organized as follows: Section 2 introduces the preliminaries of POMDP including definitions, nomenclature and solution algorithms. Section 3 presents the formulation of infrastructure scheduling problem as a POMDP framework. Section 4 illustrates the water distribution pipe system as an example and discusses the results of infinite horizon POMDP and the calculation of the offline policy. Finally, discussion and concluding remarks are provided in Section 5.

2. POMDP formulation and solution algorithm

2.1. Preliminaries

POMDP describes a stochastic sequential decision process. It is an extension of MDP to situations where observation uncertainty exists. Mathematically, POMDP is described by a tuple $< S, A, O, T, R, \Omega, \gamma >$. The nomenclature follows the most general one in the field of POMDP (Shani et al., 2013). $S = \{1, \ldots, |S|\}$ is a set of discrete internal states. $A = \{1, \ldots, |A|\}$ is a set of actions that agent can take. $\Omega = \{1, \ldots, |\Omega|\}$ is a set of discrete observations which are revealed to the agent. T: $S \times A \times S \rightarrow [0, 1]$ is a stochastic state transition function, and $T(s, a, s') = p(s'|s, a)$ means the probability of the successor state being in s' when the current state is s and the action a is taken. $R: S \times A \times S \rightarrow \mathbb{R}$ is a single stage reward function, when the current state is s and the successor state is s' with action a taken. O: $S \times A \times \Omega \rightarrow [0, 1]$ is an observation probability function. $O(a, s', o) = p(o|s', a)$ means the probability of observing o when successor state is s' and action a is taken. $\gamma \in [0, 1)$ is the discount factor to decrease the utility of later rewards.

To compare the solution of POMDP, the well-known MDP solution is first introduced. MDP is a special case of POMDP where the internal state is assumed to be fully observable. In the infinite horizon MDP problem, the solution is expressed as an optimal policy which is a mapping of each state to the corresponding optimal action: $\pi: S \rightarrow A$. The objective function or the value function is defined as the expected summation of the reward function R where any control policy π can achieve starting from state s_0.

$$V_\pi(s_0) = \mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t)) | \pi, s_0\right] \tag{1}$$

Eq. (2) is referred to as the Bellman equation for the infinite horizon discounted MDP, where $R(s, a)$ is a single stage expected reward.

$$V^*(s) = \max_{a \in A}\left[R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') V^*(s')\right] \tag{2}$$

$$R(s, a) = \sum_{s' \in S} T(s, a, s') R(s, a, s')$$

The optimal value function $V^*(s)$ is the maximal expected total discounted reward. We note that in the infinite horizon formulation, V^* is time-invariant and its uniqueness and existence are proven when $0 \le \gamma < 1$, and $R(s, a)$ is bounded. Once V^* is available, an optimal policy π^* can be obtained by solving for a given s using Eq. (2). Value iteration and policy iteration are two well-known classical methods to solve the Bellman equation (Puterman, 2014).

In the POMDP framework, the internal state S cannot be observed deterministically. Instead, a belief space B is defined as a probability distribution over S. The complete system history tuple from initial time $\eta =< S, \Omega, A >$ should be known to determine b_t. Thus a belief is expressed as:

$$b = p(s|h), h \in \eta \tag{3}$$

Similar to MDP, the solution of the infinite horizon POMDP is expressed as an optimal policy which maps beliefs b to optimal actions: $\pi: B \rightarrow A$. The objective function is the expected summation of the reward function R where the policy π starts at belief b_0:

$$V_\pi(b_0) = \mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t R(b_t, \pi(b_t)) | \pi, b_0\right] \tag{4}$$

When action a and observation o are received, the belief can be updated to $b_{a,o}$ by Bayes' rule,

$$b_{a,o}(s') = \frac{O(a, s', o)}{p(o|b, a)} \sum_{s \in S} T(s, a, s') b(s) \tag{5}$$

where the normalization constant $p(o|b, a)$ is defined as

$$p(o|b, a) = \sum_{s' \in S} O(a, s', o) \sum_{s \in S} T(s, a, s') b(s) \tag{6}$$

Since beliefs actually provide sufficient statistics for history h (Smallwood and Sondik, 1973), the value function for POMDP can be stated as a function of b only. Moreover, since the belief $b_{a,o}$ only depends on the previous belief b using Eq. (5), POMDP can be considered as an augmented MDP with continuous state. The Bellman equation for POMDP can be derived by taking the additional expectation of Eq. (2) with respect to the belief $b(s)$. We omitted s' index of b in Eq. (7) since the state dependency of b is trivial.