*13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13*
*Seoul, South Korea, May 26-30, 2019*

# Predicting the Condition Evolution of Controlled Infrastructure Components Modeled by Markov Processes

Shuo Li
*Graduate Student, Dept. of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, USA*

Matteo Pozzi
*Associate Professor, Dept. of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, USA*

ABSTRACT: When the operation and maintenance (O&M) of infrastructure components is modeled as a Markov Decision Process (MDP), the stochastic evolution following the optimal policy is completely described by a Markov transition matrix. This paper illustrates how to predict relevant features of the time evolution of these controlled components. We are interested in assessing if a critical state is reachable, in assessing the probability of reaching that state within a time period, of visiting that state before another, and in returning to that state. We present analytical methods to address these questions and discuss their computational complexity. Outcomes of these analyses can provide the decision makers with deeper understanding of the component evolution and suggest revising the control policy. We formulate the framework for MDPs and extend it to Partially Observable Markov Decision Processes (POMDPs).

## 1. Introduction

The operation and maintenance (O&M) of an infrastructure component can be modeled as a sequential decision making problem, where the decision maker infers and predicts the component's condition that evolves due to deterioration and maintenance, and takes periodic actions with the goad of minimizing a long-time maintenance costs (Frangopol et al., 2004). This stochastic decision process can be formulated as a Markov Decision Process (MDP) (Scherer and Glagola, 1994; Madanat and Ben-Akiva, 1994; Smilowitz and Madanat, 2000; Robelin and Madanat, 2007; Gao and Zhang, 2013). When the state cannot be perfectly observed, O&M can be formulated as a Partially Observable Markov Decision Processes (POMDPs) (Papakonstantinou and Shinozuka, 2014; Memarzadeh et al., 2014; Memarzadeh and Pozzi, 2016). A policy defines the O&M action as a function of the observable state (for MDPs) or of the belief (for POMDPs). The optimal policy, identified by solving the Bellman equation, is guaranteed to provide the minimum expected discounted economic cost to go. Still, a decision maker adopting that policy may wonder if critical condition states can ever be visited, and how often.

In this paper, we investigate how to predict relevant features of the condition evolution of an infrastructure component following a Markov process, and how to apply analytical methods on MDP or POMDP models under a selected policy. Outcomes of these analyses can provide the decision makers with deeper understanding of the component evolution, and suggest revising the control policy. After an introduction to sequential decision making modeling in Section 2 and some properties of Markov chains in Section 3, we formulate the general problem and propose methods for predicting specific features in MDPs in Section 4. Section 5 shows how to extend our analysis to POMDPs, and Section 6 gives examples before conclusions in Section 7.

## 2. Sequential Decision Making

In sequential decision making, an agent selects a sequence of actions, paying periodic maintenance costs and receiving possibly observations from the system she is interacting with, with the aim of minimizing the long-term expected maintenance costs.

### 2.1. *Markov chains*

A Markov chain completely describes a Markov process with discrete time and a countable state set $S = \{1, 2, \ldots, n\}$. Transition probability $p_{ij} = \mathbb{P}[s_{t+1} = j | s_t = i]$ defines the probability that a process, currently in $s_t = i$ at time $t$, moves to $s_{t+1} = j$ at the next step. These probabilities are listed in transition matrix $\mathbf{T}$, where $T(i, j) = p_{ij}$.

### 2.2. *MDP framework*

In MDPs, the system is in some state $s_t \in S$ at each time step $t$. The agent chooses an available action $a_t \in A$, and pays a cost depending on the current state and the selected action. Future costs are discounted by factor $\gamma$ per step. The long-term expected discounted cumulative cost is optimized by selecting an appropriate policy $\pi$ that can be identified by solving the Bellman equation. The evolution of the controlled system will follow a Markov process (Gardiner, 2009), completely described by a transition matrix that depends on $\pi$.

### 2.3. *POMDP framework*

In POMDPs, the state is not fully observable, and the agent, at time step $t$, only receives an observation $z_t \in Z$, which can be a noisy and incomplete measure of the current state by the probabilistic observation matrix $O_a(i, j) = \mathbb{P}[z_t = j | s_t = i, a_{t-1} = a]$. The agent's belief about the current state is represented by probability distribution $\mathbf{b}_t$, where $b_t(i) = \mathbb{P}[s_t = i | z_1, z_2, \ldots, z_t, a_0, a_1, \ldots, a_{t-1}]$. Formally, a POMDP is equivalent to a MDP in the belief state (Ibe, 2013).

## 3. Properties of Markov Chains

### 3.1. *Transient, recurrent and absorbing States*

Given a Markov chain, a state can be either transient or recurrent. A recurrent state is one that keeps returning, while a transient state is one that sooner or later will stop returning. Let $a = \mathbb{P}[\,\exists t > 0 : \ s_t = i | s_0 = i]$. We say a state $i$ is recurrent if $a = 1$ and it is transient if $a < 1$. State $i$ is an absorbing state, a special case of a recurrent one, if $p_{ii} = 1$.

### 3.2. *Ergodicity*

A Markov chain is ergodic if there exists a positive integer $k$ such that for all pairs of states $i$ and $j$, if the process starts at time $t = 0$ in state $s_0 = i$ then for all $t > k$, the probability of being in state $s_t = j$ at time $t$ is positive (Schütze et al., 2008). A numerical method to check the ergodicity of a Markov chain with $n$ states is to test if all elements of $\mathbf{T}^m$ are positive for $m = (n-1)^2 + 1$ (Meyer, 2000).

Given an ergodic Markov chain, there exists a limiting distribution $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)^\top$ that satisfies $\boldsymbol{\beta}^\top \mathbf{T} = \boldsymbol{\beta}^\top$, where $\beta_j$ represents the long-term probability of being in state $j$ and $\sum_{j=1}^n \beta_j = 1$. The limiting distribution can be defined as

$$\forall i, \ \ \beta_j = \lim_{m \to \infty} (\mathbf{T}^m)_{ij} \tag{1}$$

### 3.3. *Communication Class*

States $i$ and $j$ communicate if they are accessible from each other. The set of states can be partitioned into communication classes. To identify communication classes, we adopted the method proposed by James (2009). First, we define zero pattern transition matrix $\mathbf{D}$ as

$$D(i, j) = \begin{cases} 1 & \text{if} \ \ T(i, j) > 0 \\ 0 & \text{otherwise} \end{cases}$$

And then we define the reachability matrix $\mathbf{R}$ as

$$R(i, j) = \begin{cases} 1 & \text{if} \ \ (\mathbf{I} + \mathbf{D})^{n-1}(i, j) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{I}$ is an identify matrix and $D(i, j) = 1$ indicates state $i$ can reach state $j$. We assign states $i$ and $j$ to the same class if and only if each of these states can reach and be reached by the other. $\mathbf{Q}$ denotes the communication relationship as:

$$Q(i, j) = \begin{cases} 1 & \text{if} \ \ R(i, j) \cdot R(j, i) = 1 \\ 0 & \text{otherwise} \end{cases}$$

where states $i$ and $j$ belong to the same class if $Q(i, j) = 1$.

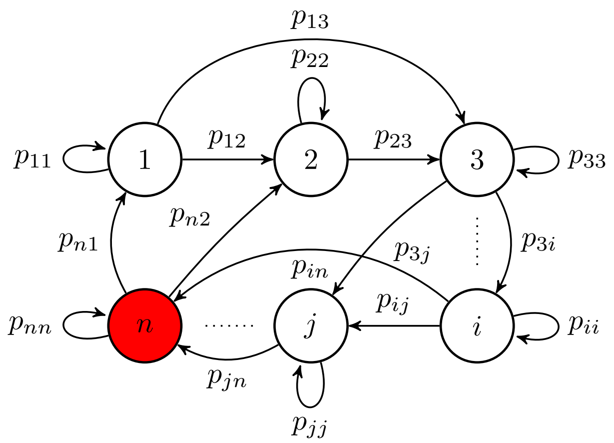


*Figure 1: A Markov transition graph.*

A communication class can be either closed or open. It is open if there is a directed connection from this class to another, otherwise it is closed.

### *3.4. Transition graphs*

As illustrated in Figure 1, a Markov chain can be represented as a directed graph $G$. Each node in $G$ corresponds to a state. $G$ contains edge $(i, j)$ if and only if $p_{ij} > 0$. A subset of $S$, of interest for our analysis, can be referred to a critical region. For example, all failure states can be grouped into that region. In Figure 1, state $n$ is the critical region, marked in red.

## 4. PREDICTING EVOLUTION

In sequential decision problems, the optimal policy finds the best trade-off between maintenance costs and frequency of critical events. To do so, a high penalty cost is usually assigned to the visit of critical states, as those related to the failure events. However, even if adopting that optimal policy, the agent cannot easily predict how the component evolves. That policy is not directly informative, for example, about whether some critical states are reachable or not, when and how frequent the process will visit those states (if reachable). Transition time from one state to another is a random variable, and we can compute some of its features, such as the expected transition time, its variance, the probability of that transition given a time period, etc. In this section, we are going to introduce analytical methods to compute these features and discuss the computational complexity.

### *4.1. Moments of first passage time*

The first passage time is defined as the time taken to the process being at one state to first reach another. Let $\tau_{ij}$ denote the number of steps needed to first reach state $j$ from state $i$ (Ibe, 2013). $\tau_{jj}$ denotes the number of steps needed to first return to state $j$ from state $j$. For an ergodic Markov chain, the mean time between visits to state $j$ can be easily computed from the limiting distribution by $\mathbb{E}[\tau_{jj}] = 1/\beta_j$. However, we also are interested in the mean time to reaching another state, e.g. a critical state. Let $m_{ij} = \mathbb{E}[\tau_{ij}]$, then we define a recursive equation to compute $m_{ij}$ as

$$m_{ij} = 1 + \sum_{k \neq j} p_{ij} m_{kj} \qquad (2)$$

It can be rewritten in a matrix equation as:

$$(\mathbf{I} - \mathbf{T})\mathbf{M} = \mathbf{1} - \mathbf{T}\mathbf{M}_d \qquad (3)$$

where $M(i, j) = m_{ij}$, $M_d(i, j) = \delta_{ij} m_{ii}$, a diagonal matrix only with elements in the diagonal elements of $\mathbf{M}$, and $\mathbf{1}$ is a matrix with all elements equal to one.

Hunter (2008) proposed an effective way to solve Eq. (3), and showed how to compute the second moment of first passage time.

#### *4.1.1. First moment (mean time)*

If $\mathbf{G}$ is any g-inverse of $(\mathbf{I} - \mathbf{T})$, then Eq. (3) is

$$\mathbf{M} = [\mathbf{G\Pi} - \mathbf{1}(\mathbf{G}\Pi)_d + \mathbf{I} - \mathbf{G} + \mathbf{1G}_d]\mathbf{M}_d \qquad (4)$$

where $\mathbf{\Pi} = \boldsymbol{e}\boldsymbol{\beta}^\top, \boldsymbol{e} = [1, 1, \ldots, 1]^\top$, and $\boldsymbol{\beta}$ is the limiting distribution (Hunter, 2008).

#### *4.1.2. Second moment*

The second moment of first passage time $\mathbf{M}^{(2)}$,

$$\mathbf{M}^{(2)} = 2[\mathbf{GM} - \mathbf{1}(\mathbf{GM})_d] + [\mathbf{I} - \mathbf{G} + \mathbf{1G}_d][\mathbf{M}_d^{(2)} + \mathbf{M}_d] - \mathbf{M} \qquad (5)$$

where $\mathbf{M}_d^{(2)} = 2\mathbf{M}_d(\mathbf{\Pi M})_d - \mathbf{M}_d$.

The variance matrix $\mathbf{V}$ can be computed as $\mathbf{V} = \mathbf{M}^{(2)} - \mathbf{M}^2$ (Hunter, 2008). Hence, $M(i, j)$ and $V(i, j)$ are the mean and variance of first passage time from state $i$ to $j$, respectively.

### 4.2. Probability of reaching states

For an ergodic Markov chain, given any three states $i, j, c \in S$, we want to find the probability of moving from state $i$ to $j$ before hitting state $c$. This probability can be expressed as

$$\mathbb{P}(\exists t : s_t = j, s_k \neq c, s_k \neq j, \forall k \in \{1, 2, \ldots, t-1\} | s_0 = i)$$

We define $h_i$ as the probability of moving from the initial state $i$ to state $j$ before reaching the state $c$. The transition matrix $\mathbf{T}$ satisfies the following equation with the boundary conditions,

$$h_i = \sum_{k=1}^{n} T(i,k) h_k, i \neq j, c, \quad h_j = 1 \text{ and } h_c = 0 \quad (6)$$

We define $\mathbf{h} = (h_1, h_2, \ldots, h_n)^\top$, and let $\hat{\mathbf{T}}$ be $\mathbf{T}$ after deleting the $c$th and $j$th rows and columns, of the size $(n-2) \times (n-2)$. Let $\mathbf{r}$ be the $j$th column of $\mathbf{T}$ after deleting $c$th and $j$th elements, with the size of $(n-2) \times 1$. Let $\hat{\mathbf{h}}$ be $\mathbf{h}$ after deleting the $c$th and $j$th elements, with the size of $(n-2) \times 1$. Then we can solve the Eq. (6) by

$$\hat{\mathbf{h}} = (\mathbf{I} - \hat{\mathbf{T}})^{-1} \mathbf{r} \quad (7)$$

### 4.3. Probability of transition time

We now want to investigate the probability of first passage time from state $i$ to state $j$. We define $f_{i,j}(t)$ as the probability of first reaching the state $j$ within $t$ steps, starting from state $i$. We define a new transition matrix $\mathbf{T}^+$ by setting the $j$th row and column of the original transition matrix $\mathbf{T}$ to zero. Let $\mathbf{x}_j(t) = [f_{1,j}(t) \;\; f_{2,j}(t) \;\; \ldots \;\; f_{n,j}(t)]^\top$. We can derive $\mathbf{x}_j(t)$ at any time $t$ by the following recursive equation:

$$\mathbf{x}_j(t) = \mathbf{x}_j(0) + \mathbf{T}^+ \mathbf{x}_j(t-1) \quad (8)$$

where the initial condition $\mathbf{x}_j(0)$ is a vector with all element being zero except for $j$th element being one. After solving Eq. (8), we can derive the probability of first reaching state $j$ from state $i$ in exactly $t$ steps as

$$\Delta f_{i,j}(t) = f_{i,j}(t) - f_{i,j}(t-1) \quad (9)$$

*Table 1: Transition probability matrices under the optimal policy for the MDP example 1.*

$$\mathbf{T} = \begin{bmatrix} 0.8580 & 0.1358 & 0.0062 & 0 & 0 & 0 & 0 & 0 \\ 0.7625 & 0.2215 & 0.0159 & 0.0002 & 0 & 0 & 0 & 0 \\ 0.2375 & 0.5249 & 0.2215 & 0.0159 & 0.0002 & 0 & 0 & 0 \\ 0.1420 & 0.4975 & 0.3234 & 0.0364 & 0.0007 & 0 & 0 & 0 \\ 0.0766 & 0.4234 & 0.4234 & 0.0744 & 0.0021 & 0 & 0 & 0 \\ 0.2375 & 0.5249 & 0.2215 & 0.0159 & 0.0002 & 0 & 0 & 0 \\ 0.0161 & 0.2215 & 0.5249 & 0.2215 & 0.0159 & 0.0002 & 0 & 0 \\ 0.0002 & 0.0159 & 0.2215 & 0.5249 & 0.2215 & 0.0159 & 0.0002 & 0 \end{bmatrix}$$

## 5. Extension on POMDPs

POMDPs can be expressed as Markov processes in the belief state. Consider a O&M process of an infrastructure component with $n$ possible states modeled as a POMDP. In order to build a transition matrix $\mathbf{T}_\pi$ from one belief state to another following policy $\pi$ in POMDPs, we discretize the belief domain into $N$ belief states, $B = \{\mathbf{b}_1, \mathbf{h}_2, \ldots, \mathbf{b}_N\}$. For a single discrete belief state $\mathbf{b}_k \in B$, its one-step transition is defined as follows,

$$\begin{cases} \mathbf{b}_k^a = \mathbf{b}_k \mathbf{T}_a \\ v_j = \sum_{l=1}^{n} O_a(l,j) \; b_k^a(l) \\ u(j,i) = \dfrac{O_a(i,j) \; b_k^a(i)}{v_j} \end{cases} \quad (10)$$

where action $a = \pi(\mathbf{b}_k)$ is selected by the deterministic policy. $\mathbf{b}_k^a$ is an updated belief after taking action $a$, and $v_j$ represents the probability of getting $j$th observation based on the belief $\mathbf{b}_k^a$. The $j$th row of $u$ is an updated belief state $\mathbf{b}_k^{a,j}$ from $j$th observation, which is obtained by $b_k^{a,j}(i) = u(j,i)$, for $i = 1, 2, \ldots, n$. Then $\mathbf{b}_k^{a,j}$ can be approximated to $\mathbf{b}_h \in B$ by the minimum Euclidean distance. The transition probability from belief state $\mathbf{b}_k$ to $\mathbf{b}_h$ is the sum of all $v_j$ related to the observations taking from $\mathbf{b}_k$ to $\mathbf{b}_h$. Transition matrix $\mathbf{T}$, with the size $N \times N$, can be constructed by considering all discrete beliefs and all possible observations. Also, linear interpolation among beliefs can be used for approximating the inference process.

## 6. Examples

### 6.1. MDP example 1

The transition matrix of a MDP example, related to a pavement management problem (Durango and

*Table 2: Transition matrices for the MDP example 2.*

$$\mathbf{T}_1 = \begin{bmatrix} 0.9 & 0.05 & 0.02 & 0.02 & 0.01 & 0 \\ 0 & 0.8 & 0.1 & 0.06 & 0.03 & 0.01 \\ 0 & 0 & 0.75 & 0.1 & 0.1 & 0.05 \\ 0 & 0 & 0 & 0.5 & 0.3 & 0.2 \\ 0 & 0 & 0 & 0 & 0.35 & 0.65 \\ 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}$$

$$\mathbf{T}_2 = \begin{bmatrix} 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0 & 0 & 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}$$

$$\mathbf{T}_3 = \begin{bmatrix} 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.008 & 0.002 \end{bmatrix}$$

*Table 3: Transition probability matrices under the optimal policy for the MDP example 2.*

$$\mathbf{T} = \begin{bmatrix} 0.9 & 0.05 & 0.02 & 0.02 & 0.01 & 0 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.4 & 0.32 & 0.1 & 0.1 & 0.078 & 0.002 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.01 & 0 \\ 0.9 & 0.05 & 0.02 & 0.02 & 0.01 & 0 \end{bmatrix}$$

Madanat, 2002), following the optimal policy, is shown in Table 1. The pavement segment condition is discretized into 8 states, from a new pavement $(s = 1)$ to the failure state $(s = 8)$. Following the optimal policy, we can see the failure state is not reachable and states $1,2,3,4,5$ form a closed communication class by the method from Section 3.3. Therefore, states $6,7,8$ are transient.

### 6.2. *MDP example 2*

Also inspired by Durango and Madanat (2002)'s model, let us consider a component with $n = 6$ states, referring to five different level of deterioration from *intact* (s = 1) to *severely damaged* (s = 5) state and a *failure* state $(s = 6)$, considered as a critical state. The three maintenance actions include do-nothing$(a = 1)$, minor repair $(a = 2)$, replace $(a =$ 3), and the transition probabilities for each action are shown in Table 2. Time is discretized in years. Costs of minor repair, major repair and for failure are \$8K, \$20K and \$500K, respectively. The discount factor is 0.95 per year. By solving that MDP, the optimal policy $\pi = [1\ 2\ 2\ 2\ 3\ 3]^\top$ is obtained, and the expected discounted maintenance costs are \$33.5K, \$50.91K, \$50.91K, \$50.91K. \$53.53K and \$553.53K for initial states from $s_0 = 1$ to $s_0 = 6$, respectively. The corresponding transition matrix $\mathbf{T}$ is shown in Table 3. All the states are reachable and the chain is ergodic by the method from Section 3.2. Table 4 shows the mean time matrix $\mathbf{M}$ in years. The expected time to first reach the failure state is relatively much longer than that to first reach any other state from the same initial state. By computing the standard deviation matrix $\mathbf{V}_{std} = \sqrt{\mathbf{V}}$, we find that the coefficient of variation of transition time is around 1. Figure 2 illustrates the probability of transition time to the failure. As an example of the results, the probability of failure within 20 years starting from state $s_0 = 1$ is around 0.54%. Also, the probabilities of reaching failure before visiting state $s = 5$ starting from state $s_0 = 1,2,3,4$ are 1.5%, 1.67%, 1.67% and 1.67%, respectively. Hence, we predict that, with 98.5% probability, a new component will be replaced (as it will be at state 5) before its failure.

*Table 4: Mean of first passage time in years for the MDP example 2.*

$$\mathbf{M} = \begin{bmatrix} 1.22 & 15.9 & 34.9 & 34.9 & 48.2 & 3177.0 \\ 2.29 & 10.8 & 30.6 & 30.6 & 42.5 & 3166.0 \\ 2.29 & 10.8 & 30.6 & 30.6 & 42.5 & 3166.0 \\ 2.29 & 10.8 & 30.6 & 30.6 & 42.5 & 3166.0 \\ 1.22 & 15.9 & 34.9 & 34.9 & 48.2 & 3177.0 \\ 1.22 & 15.9 & 34.9 & 34.9 & 48.2 & 3177.0 \end{bmatrix}$$

### 6.3. *POMDP example 1*

Consider an infrastructure component with $n = 4$ states, *intact, minor damaged, major damaged* and *failure* states. The three maintenance actions are available: *Do-nothing* $(a = 1)$, *Minor Repair* $(a = 2)$ and *Replace* $(a = 3)$. The transition matrices are shown in Table 5. 10 different observations are available and assume the failure can be immediately

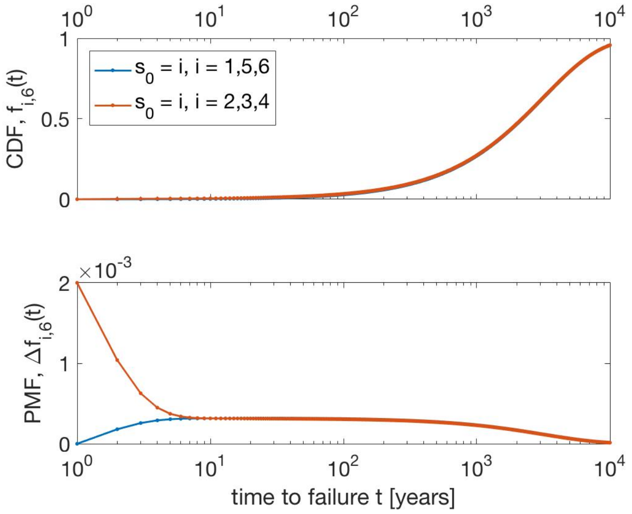


*Figure 2: Probability of transition time to failure for the MDP example 2*

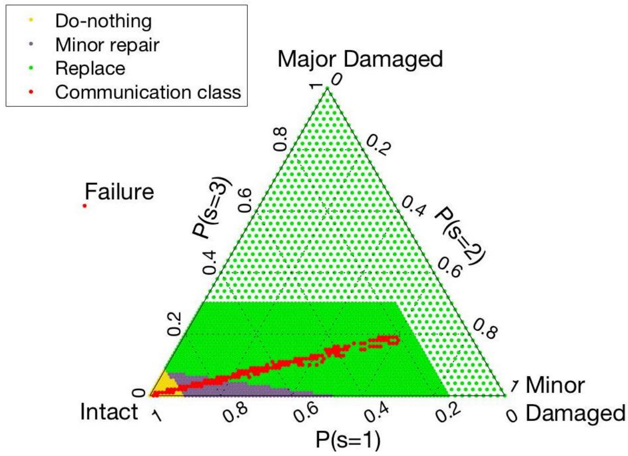


*Figure 3: Optimal policy for the discrete belief domain for the POMDP example 1*

*Table 5: Transition matrices for POMDP example 1.*

$$\mathbf{T}_1 = \begin{bmatrix} 0.9 & 0.08 & 0.02 & 0 \\ 0 & 0.7 & 0.2 & 0.1 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} \quad \mathbf{T}_2 = \begin{bmatrix} 0.9 & 0.08 & 0.02 & 0 \\ 0 & 0.85 & 0.12 & 0.03 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

$$\mathbf{T}_3 = \begin{bmatrix} 0.9 & 0.08 & 0.02 & 0 \\ 0.9 & 0.08 & 0.02 & 0 \\ 0.9 & 0.08 & 0.02 & 0 \\ 0.9 & 0.08 & 0.02 & 0 \end{bmatrix}$$

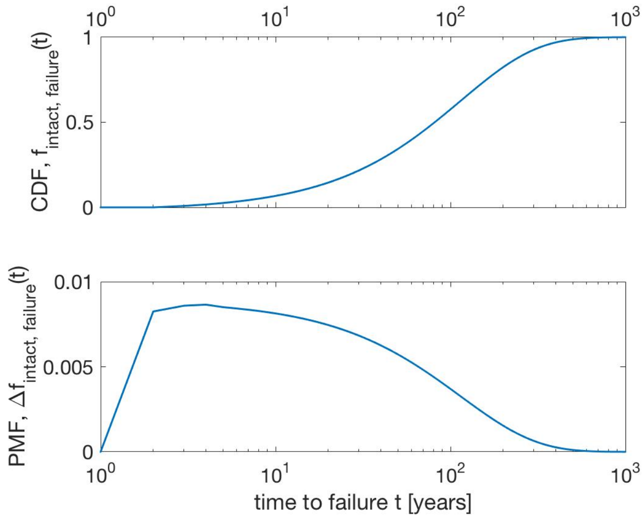


*Figure 4: Probability of transition time to failure for the POMDP example 1*

detected. Cost is assumed to be \$3.5K, \$25K for $a = 2,3$, respectively, \$500K for the *failure*, and the discount factor is 95% per year. Time is discretized in years. The belief domain with four possible states can be represented by an equilateral triangle and a single point when the failure is immediately detectable. In Figure 3, the belief domain is non-uniformly discretized with $N = 2929$ points. Each belief corresponds to an action controlled by the optimal policy that is obtained by the SARSOP solver (Kurniawati et al., 2008). We also plot the beliefs belonging to the only closed communication class, marked in red. The process will only move among those beliefs, in the long term. Figure 4 shows the cumulative distribution function (CDF) and probability mass function (PMF) of failure for an initial belief $\mathbf{b}_0 = [1\ 0\ 0\ 0]$ depending on time $t$.

*Table 6: Transition and observation matrices for the POMDP example 2.*

$$\mathbf{T}_1 = \begin{bmatrix} 0.97 & 0.02 & 0.01 & 0 \\ 0 & 0.85 & 0.1 & 0.05 \\ 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} \quad \mathbf{T}_2 = \begin{bmatrix} 0.97 & 0.02 & 0.01 & 0 \\ 0.1 & 0.8 & 0.08 & 0.02 \\ 0 & 0.2 & 0.6 & 0.2 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

$$\mathbf{T}_3 = \begin{bmatrix} 0.97 & 0.02 & 0.01 & 0 \\ 0.21 & 0.75 & 0.03 & 0.01 \\ 0 & 0.62 & 0.36 & 0.02 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} \quad \mathbf{T}_4 = \begin{bmatrix} 0.97 & 0.02 & 0.01 & 0 \\ 0.97 & 0.02 & 0.01 & 0 \\ 0.97 & 0.02 & 0.01 & 0 \\ 0.97 & 0.02 & 0.01 & 0 \end{bmatrix}$$

$$\mathbf{O}_{1-4} = \begin{bmatrix} 1-\varepsilon & \varepsilon/2 & \varepsilon/2 & 0 \\ \varepsilon/2 & 1-\varepsilon & \varepsilon/2 & 0 \\ \varepsilon/2 & \varepsilon/2 & 1-\varepsilon & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

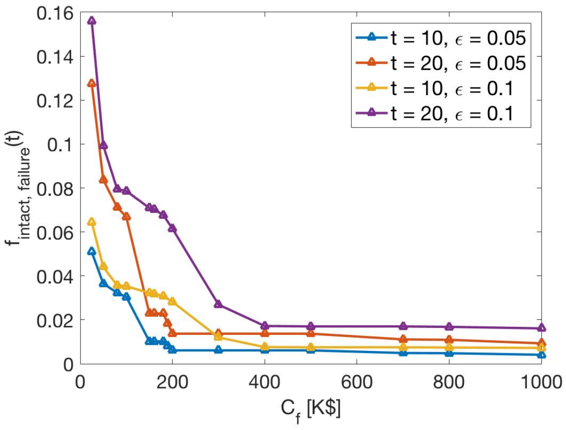


*Figure 5: Probability of transition time to failure for the POMDP example 2*

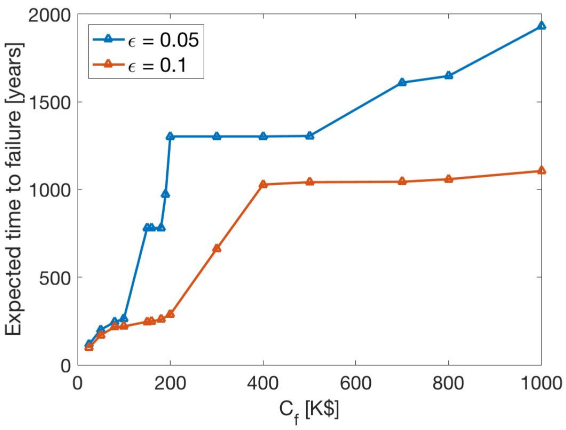


*Figure 6: Expected time to failure for the POMDP example 2*

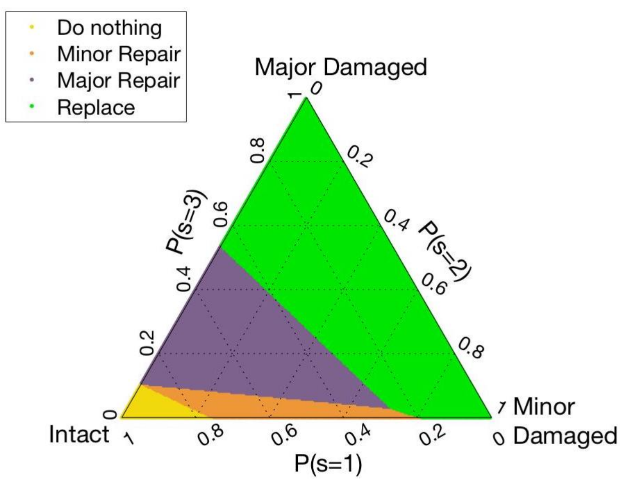


*Figure 7: Optimal policy for the POMDP example 2*

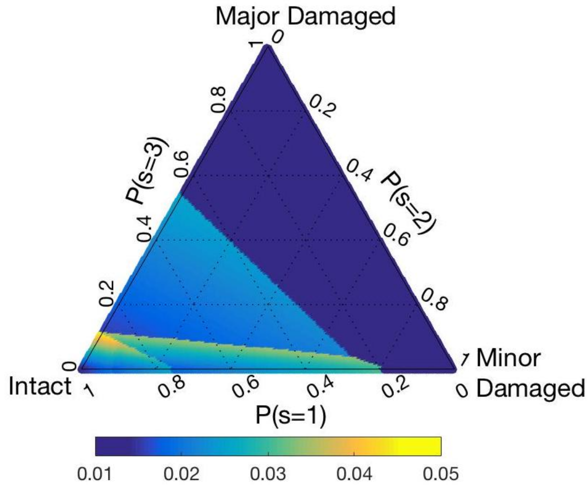


*Figure 8: Probability of failure within 20 years when* $C_f = \$300K, \varepsilon = 0.05$.

### 6.4. *POMDP example 2*

Consider an infrastructure component with four states ($n = 4$). Four maintenance actions including *do-nothing* ($a = 1$), *minor repair* ($a = 2$), *major repair* ($a = 3$) and *replace* ($a = 4$), are available. Costs of $a = 2,3,4$ are \$3.5K, \$9K and \$25K, respectively. The discount factor is 0.95 per year. The transition and observation matrices are shown in Table 6. By varying cost of *failure* ($C_f$) from \$25K to \$1000K, the optimal policy changes. Figure 5 illustrates how the cost of failure influences the probability of failure depending on time $t$ from an intact state under two different inaccuracy of observations $\varepsilon = 0.05, 0.1$. Figure 6 shows the expected time from an intact state to the failure when $\varepsilon = 0.05, 0.1$. The expected time to failure is monotonically increasing with $C_f$. When $C_f = \$300K$ and $\varepsilon = 0.05$, we plot the optimal policy in Figure 7 and the corresponding probability of failure within 20 years for all discrete beliefs in Figure 8.

## 7. CONCLUSIONS

This paper has illustrated how to predict the future evolution of an infrastructure component modeled by a MDP or POMDP under a selected policy. That evolution can also be investigated by analyzing the outcomes of Monte Carlo (MC) simulations. However, MC methods require high computational cost for simulating rare events, as the component failure, and they are not efficient to achieve results as those in Figure 8.

In POMDPs, the computational complexity is highly dependent on the number of discrete beliefs, and the accuracy of the approximation increases with the number of discrete beliefs. When the number of states increases, the accuracy suffers from the "curse of dimensionality". However, approaches not based on the grid discretization of Section 5, can also be investigated.

In this work, we have assumed that the control optimal is pre-assigned. The task of identifying the optimal policy is generally harder than that of predicting the evolution. An interesting line of research is to couple analytical methods for finding the optimal policy and for predicting the evolution, to improve the process achieving both goals.

## 8. ACKNOWLEDGEMENTS

We acknowledge the support of NSF project CMMI #1663479, titled "From Future Learning to Current Action: Long-Term Sequential Infrastructure Planning under Uncertainty".

## 9. REFERENCES

Durango, P. L. and Madanat, S. M. (2002). "Optimal maintenance and repair policies in infrastructure management under uncertain facility deterioration rates: an adaptive control approach." *Transportation Research Part A: Policy and Practice*, 36(9), 763–778.

Frangopol, D. M., Kallen, M.-J., and Van Noortwijk, J. M. (2004). "Probabilistic models for life-cycle performance of deteriorating structures: review and future directions." *Progress in Structural Engineering and Materials*, 6(4), 197–212.

Gao, H. and Zhang, X. (2013). "A markov-based road maintenance optimization model considering user costs." *Computer-Aided Civil and Infrastructure Engineering*, 28(6), 451–464.

Gardiner, C. (2009). *Stochastic methods*, Vol. 4. springer Berlin.

Hunter, J. J. (2008). "Variances of first passage times in a markov chain with applications to mixing times." *Linear Algebra and its Applications*, 429(5-6), 1135–1162.

Ibe, O. (2013). *Markov processes for stochastic modeling*. Newnes.

James, M. (2009). "Communication classes, <https://www.ssc.wisc.edu/ jmontgom/376textbook.htm>.

Kurniawati, H., Hsu, D., and Lee, W. S. (2008). "Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces.." *Robotics: Science and systems*, Vol. 2008, Zurich, Switzerland.

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

Memarzadeh, M. and Pozzi, M. (2016). "Value of information in sequential decision making: Component inspection, permanent monitoring and system-level scheduling." *Reliability Engineering & System Safety*, 154, 137–151.

Memarzadeh, M., Pozzi, M., and Zico Kolter, J. (2014). "Optimal planning and learning in uncertain environments for the management of wind farms." *Journal of Computing in Civil Engineering*, 29(5), 04014076.

Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*, Vol. 71. Siam.

Papakonstantinou, K. G. and Shinozuka, M. (2014). "Planning structural inspection and maintenance policies via dynamic programming and markov processes. part ii: Pomdp implementation." *Reliability Engineering & System Safety*, 130, 214–224.

Robelin, C.-A. and Madanat, S. M. (2007). "History-dependent bridge deck maintenance and replacement optimization with markov decision processes." *Journal of Infrastructure Systems*, 13(3), 195–201.

Scherer, W. T. and Glagola, D. M. (1994). "Markovian models for bridge maintenance management." *Journal of Transportation Engineering*, 120(1), 37–51.

Schütze, H., Manning, C. D., and Raghavan, P. (2008). *Introduction to information retrieval*, Vol. 39. Cambridge University Press.

Smilowitz, K. and Madanat, S. (2000). "Optimal inspection and maintenance policies for infrastructure networks." *Computer-Aided Civil and Infrastructure Engineering*, 15(1), 5–13.