# Risk Management of Multi-state Multi-component Bridge Systems Using Partially Observable Markov Decision Processes

Abdollah Shafieezadeh
*Assistant Professor, Dept. of Civil, Env. & Geo. Engineering, Ohio State Univ., Columbus, USA*

Ehsan Fereshtehnejad
*Graduate Student, Dept. of Civil, Env. & Geo. Engineering, Ohio State Univ., Columbus, USA*

ABSTRACT: Infrastructure systems play a critical role in providing continuous services to societies. Exposure to stressors such as aging, demand loads, and environmental factors threatens the functionality and safety of infrastructure systems, highlighting the necessity for proper decision-making frameworks. Toward this goal, in the light of imperfect asset condition state evaluation, this paper presents a stochastic framework based on partially observable Markov decision process (POMDP) for the determination of optimal maintenance actions. A feature of this approach is its ability to effectively and accurately manage large scale, multi-state multi-component bridge systems. To overcome the dimensionality curse of the decision-making for such large systems without losing accuracy, the “counting process” state reduction technique is applied and conformed in a novel way. Further, to significantly reduce the computational runtime while keeping the accuracy in a high level, a randomized point-based value iteration POMDP is utilized. The proposed framework is applied to a case study bridge system with four steel girders and one concrete deck. Results of 12 random runs showed acceptable convergence in the optimized average expected long-run reward. The applied framework provides optimal policies for the concrete deck and girders in each of the possible states. It is also concluded that the combination of the POMDP decision-making framework and the “counting process” technique gives rise to an efficient and accurate approach for the optimal management of large scale systems.

## 1. INTRODUCTION

Infrastructure systems play a critical role in providing continuous services to societies to support their economic prosperity and public health and safety. The quality and safety of infrastructure systems depend considerably on their physical and functional states. Aging, demand loads, and environmental stressors are among factors that bring about various degradation processes in infrastructure systems. For instance, a bridge system consists of multiple critical components which may undergo stochastic degradation processes depending on their exposure to various stressors mentioned above. The combined effects of degradation in bridge components can lower the service quality, decrease strength of structural elements and therefore, lessen the reliability of the bridge. This highlights the need for proper infrastructure maintenance, repair and rehabilitation (MR&R) management to reduce the likelihood of degraded functionality and incurred costs, as well as, potentially detrimental consequences of severe element damage.

Toward this goal, in the presence of imperfect asset condition state evaluation, a number of probabilistic component-level decision making frameworks have been implemented in civil engineering applications. Incorporating measuring randomness in addition to forecasting uncertainty, LMDP (Latent Markov Decision Process) proposed by Madanat and Ben-Akiva, (1994) and POMDP (Partially Observable Markov Decision Process) first offered by Monahan (1982), gained attention for a single-component pavement and bridge management. While in the LMDP within the optimization process, all possible combinations of scenarios through time are simulated, POMDP in general, determines regions in the belief state with similar optimal actions (Monahan, 1982) resulting in a significant reduction in the required runtime compared to the LMDP.

In addition to the fact that inspection uncertainties are neglected in MDP-based frameworks, a pitfall of these frameworks is the poor adaptation to a portfolio of joint components (Kuhn, 2010). From a practical viewpoint, in many research papers, components of infrastructure systems are grouped into few larger-components through which the decision-making framework has reduced and manageable dimensions (Ellis et al., 1995, Scherer and Glagola, 1994, Durango-Cohen et al., 2007, Kuhn, 2009). This however, gives birth to considerable level of approximation for the assignment of element-level optimal strategies. To overcome this difficulty, in this article, a "counting process" technique is applied and conformed to reduce the number of state combinations. Via this approach, instead of considering all possible combinations of the elements condition states individually, a new condition state is introduced that represents the total *number* of elements in a specific state. The application of this technique, paves the way towards a more practical decision making framework, while keeping the accuracy unchanged.

In this paper, a POMDP-based decision making framework is applied and enhanced with the "counting process" state reduction technique. The platform is implemented on a realistic case study bridge system entailing 4 girders and a concrete deck. The rest of this paper is organized as follows: In section 2, the applied POMDP-based framework is reviewed and discussed in technical terms. Section 3, explains the "counting process" technique. Following that in section 4, the framework is implemented on the example bridge system and the numerical results are provided. Finally, as the conclusion part, the characteristics of the applied framework for large scale systems will be discussed briefly and recommendations for future research will be given accordingly.

## 2. POMDP FRAMEWORK

In a discrete MDP, the time variant behavior of a component is predicted through Markov chains. A Markov chain consists of a transition probability matrix that defines the conditional probability of the true condition state of the utility at time *t* given the state at time *t-1*

$$P(X_t|X_{t-1}, X_{t-2}, \dots, X_0) = P(X_t|X_{t-1}) \quad (1)$$

where $P(.)$ denotes the probability and $X_i$ stands for the condition state at time *i*. For each of the possible condition states at time *t-1*, a probability mass function (PMF) is provided in the Markov chain to describe the likelihood of the condition states for the next time period. A graph-based representation of a 5-state Markov chain is depicted in Figure 1 where PMF values are shown on the graph edges. The goal of the MDP is to choose among a set of actions such that an objective function is optimized. Each considered action requires a transition probability matrix describing the probabilistic impact of the action on the condition states, as well as the reward associated with that action. Defining the objective function as the expected accumulated reward at each of the decision making times, the MDP framework can be mathematically described as

$$V_n(s_n) = max_{a_n}\{r_{a_n,S} + \gamma \sum_{\acute{s}} P(\acute{s}_{n-1}|s_n, a_n).V_{n-1}(\acute{s}_{n-1})\} \quad (2)$$

where $V_n$ and $V_{n-1}$ stand for the expected accumulated values at stage *n* and *n-1*, respectively, $r_{a_n,s}$ denotes the immediate reward of taking action $a_n$ when the system is in condition $s$, and $\gamma$ represents the discount factor. Finally, $P(\acute{s}_{n-1}|s_n, a_n)$ is the Markov transition probability from state $s$ at stage $n$ to state $\acute{s}$ at stage *n-1*.

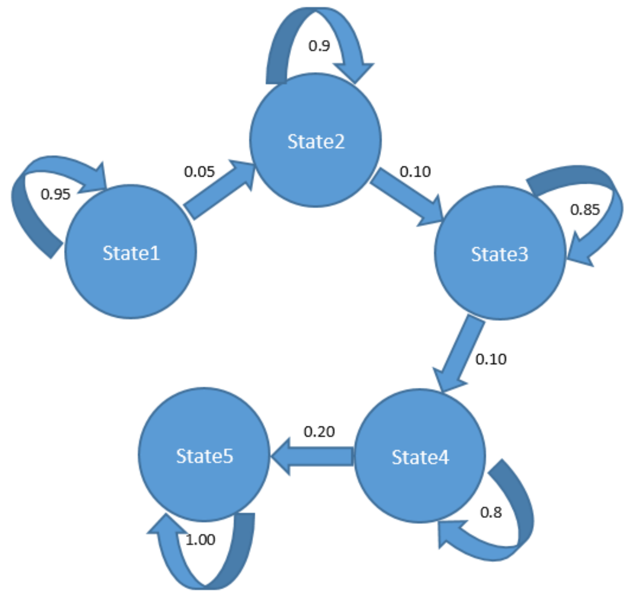


*Figure 1: Markov chain graphical representation for the deck of the case study bridge system under Do nothing action*

In the MDP framework, it is assumed that the state of the system is fully known, meaning that the system state is observed at each stage and there is no uncertainty in the observation process. However, in reality, the system condition is perceived through a set of observations that provide an estimate of the true condition state. Therefore following the observation, the true state of the system is not known with certainty; instead, it can be described probabilistically using a set of PMFs for each of the system condition state (called hereafter as a belief state) at the decision making time epochs. As a result, the belief state depends upon the entire history of observations and actions taken beforehand (Madanat and Ben-Akiva, 1994). In an MDP problem with measurement randomness considered, the optimal decisions are identical for many of the neighbor belief states, i.e. states with close probability mass functions. This property that leads to a decrease in the number of required analyses, gave birth to the POMDP framework as an efficient tool for optimal decision making. At each stage of analysis, POMDP frameworks find regions in the belief space (the entire region of the belief state possibilities) that have identical optimal value functions. The total number of optimal value functions is finite over the entire belief space, since the value functions are always piecewise linear and convex (Cassandra, 1999).

The general decision-making framework is given in the following equation

$$V_n = max_{a_n}\left\{\overrightarrow{\pi_n^S}.\overrightarrow{r_{a_n,S}} + \gamma \sum_O P(o_{n-1}, a_n, \overrightarrow{\pi_n^S}).V_{n-1}\left(o_{n-1}, a_n, \overrightarrow{\pi_n^S}\right)\right\} \quad (3)$$

where $V_n$ and $V_{n-1}$ stand for the expected accumulated values (in terms of reward) at stages *n* and *n-1*, respectively, $\overrightarrow{\pi_n^S}$ represents PMFs of the vector of condition states at stage *n* (i.e. $\overrightarrow{\pi_n^S} = \left[\overrightarrow{\pi_n^1}, \overrightarrow{\pi_n^2}, \ldots, \overrightarrow{\pi_n^S}\right]$ where $s$ is the total number of condition states), $P(o_{n-1}, a_n, \overrightarrow{\pi_n^S})$ indicates the probability of an observation (measured value) at stage *n-1* when action $a_n$ is taken and the condition state at stage *n* has the PMF of $\overrightarrow{\pi_S^n}$. Starting from the first stage, the value function is represented as $\overrightarrow{\pi_S^0} \times \overrightarrow{v_0}^t(S)$ where $\overrightarrow{v_0}^t(S)$ is the transpose of the salvage value vector for different condition states. This forms a hyper-plane with the normal vector $\overrightarrow{v_0}^t(S)$. The components of this vector are called $\alpha_0^{\pi_{\acute{s}}^*}$ coefficients for the next stage, i.e. stage 1. In general, $V_{n-1}\left(o_{n-1}, a_n, \overrightarrow{\pi_S^n}\right)$ can be simplified to

$$V_{n-1}\left(o_{n-1}, a_n, \overrightarrow{\pi_n^S}\right) = \sum_{\acute{S}} \alpha_{n-1}^{\pi_{\acute{s}}^*}.\vec{P}(\acute{s}_{n-1}|o_{n-1}, a_n, s_n) = \sum_{\acute{S}} \alpha_{n-1}^{\pi_{\acute{s}}^*} \times \frac{P(o_{n-1}|\acute{s}_{n-1}, a_n) \times \vec{P}(\acute{s}_{n-1}|a_n, s_n).\overrightarrow{\pi_n^S}}{P(o_{n-1}, a_n, \overrightarrow{\pi_n^S})} = \overrightarrow{\pi_S^n}.\sum_{\acute{S}} \alpha_{n-1}^{\pi_{\acute{s}}^*} \times \frac{P(o_{n-1}|\acute{s}_{n-1}, a_n) \times \vec{P}(\acute{s}_{n-1}|a_n, s_n)}{P(o_{n-1}, a_n, \overrightarrow{\pi_n^S})} \quad (4)$$

Replacing Equation (4) into Equation (3), the general framework of POMDP for two successive stages is derived as follows

$$V_n = max_{a_n}\left\{\overrightarrow{\pi_n^S}.\left[\overrightarrow{r_{a_n,S}} + \gamma \sum_O \sum_{\acute{S}} \alpha_{n-1}^{\pi_{\acute{s}}^*} \times P(o_{n-1}|\acute{s}_{n-1}, a_n) \times \vec{P}(\acute{s}_{n-1}|a_n, s_n)\right]\right\} = max_{a_n}\left\{\overrightarrow{\pi_n^S}.\alpha_n^{\pi_{\acute{s}}^*}\right\} \quad (5)$$

The maximum value function derived according to Equation (5) for a vector of $\overrightarrow{\pi_n^S}$ becomes the $\alpha_n^{\pi_S^*}$ coefficient for the analysis at the next stage.

The total number of realizations at each time of the POMDP analysis, is $\mathrm{L} \times |V_{n-1}|^{\mathrm{M}}$ where $|V_{n-1}|$ is the total number of value functions at stage $n$-1 and $L$ and $M$ are the number of possible actions and observations, respectively (Spaan and Vlassis, 2005). Therefore, the total number of realizations for the entire time horizon, $TH$, for the exact POMDP becomes $\sum_{T=1}^{TH} \mathrm{L} \times |V_T|^{\mathrm{M}}$. Thus, with respect to the decision-making time, $T$, POMDP has polynomial order time complexity, whereas LMDP framework has an exponential one.

Another primary concern regarding the computational cost of optimization-based decision making frameworks is that, for a system of multiple elements, the size of observation and action increases exponentially with the number of components. Since the observation size appears in the power, the total number of realizations grows significantly even in the POMDP framework. To overcome this limitation for multi-component systems, Spaan and Vlassis (2005) proposed the randomized point-based value iteration POMDP procedure, called "Perseus". Perseus is a steady state stationary policy POMDP that reduces the computational demand by using the fact that a single optimal action for a belief point may improve many other belief points. Within this framework, first, a set of likely random belief points are generated through multiple epochs of random walks. Second, in the POMDP framework, optimal value functions and strategies are computed for these points.

The mathematical representation of the decision making framework is the same as Equation (5), while in Perseus, optimization is performed for the set of discrete random belief points. In addition, Spaan and Vlassis (2005) proposed a steady state algorithm with the key idea that in each optimization stage, the value and the policy of all the random points can be optimized only by a subset of these points. The convergence criteria for the algorithm is set as either a relative tolerance value for two successive value functions, $max_{\overrightarrow{\pi_n^S}}\,(V_n - V_{n-1})/V_{n-1}$, or the number of optimal policy change between two successive stages.

At the end as a result, this algorithm provides the agencies with the optimal policies to make if any of the random belief points should be met.

## 3. "COUNTING PROCESS" TECHNIQUE

In MDP-based frameworks, the explosion of the condition states has been identified as a common problem. This issue becomes even more pronounced for POMDP models since measurement randomness adds another layer of time complexity to the framework. For moderate-scale systems, Scherer et al (1994) suggested the so called "counting process" technique to reduce the size of the state space without missing any decision-making information. In this technique, instead of considering all possible combinations of the elements condition states individually, a new condition state is introduced that represents the total *number* of components in a specific state. The new set of states are called Super States in this paper. Each Super State is a vector representing the total number of components in each condition state. For instance, V=[2 0 3] is a Super State vector indicating two out of five components of a system are in condition state 1, while the three remaining components are in condition state 3. Since the grouped components have identical consequences, no approximation is introduced into the decision-making model. Considering $|S|$ as the number of states for a component type and $N$ as the number of component types, the total number of states using the "counting process" ($|SS|$, i.e. the size of Super States) can be found from Equation (6).

$$|SS| = \binom{N+|S|-1}{|S|-1} \tag{6}$$

In order to compute the transition probabilities for Super States as well as the conditional probabilities for observations, the following procedure is proposed: First, all permutations of the possible realizations of Super States going from one specific Super State to another should be determined and then be added together as mutually exclusive events. With $N$ as the total number of components and $N_i$ as the number of components in state *i*, the total number of permutations ($N_k$) for a Super State follows Equation (7).

$$N_k = \prod_{i=1}^{n} \binom{N-\sum_{j=1}^{i-1} N_j}{N_i} \quad (7)$$

Then, based on the total probability theorem, the following equation is used to compute the transition probabilities of Super State *A* evolving to Super State *B*:

$$P(SS_B|SS_A) = \sum_{i=1}^{N_A} \frac{1}{N_A} \times \sum_{j=1}^{N_B} \prod_{k=1}^{N} P\left(S_{ij,k}\right) \quad (8)$$

where $N_A$ and $N_B$ are the total number of permutations for Super States *A* and *B*, respectively, and $P\left(S_{ij,k}\right)$ is the probability of element *k* transitioning from state *i* to state *j*. It is worthy to note that since the permutations of the Super State *A* have identical likelihood of occurrence, a non-informative uniform PMF of $\frac{1}{N_A}$ should be considered as a multiplicative factor in Equation (9).

A similar procedure can be derived for computing the conditional probabilities of observation Super States.

## 4. NUMERICAL RESULTS

Application of the proposed framework is demonstrated here for MR&R decision making of a bridge system. However, it should be noted that the proposed approach could be applied for the optimal management of any multi-state multi-component system. The case study bridge system, is a single span bridge structure composing of four girders and a deck element. The structural specifications of the girders and concrete deck are taken from Al-Wazeer (2007).

By adopting the "counting process" technique and applying the proposed modification, the total number of states corresponding to girders, and the system of girders and deck, reduces to $\binom{4+5-1}{5-1} = 70$ and $70 \times 5 = 350$, transforming the problem into a feasible scope for the POMDP decision-making framework.

### *4.1. System information*

Transition probabilities for the concrete deck and girders and the corresponding cost values for MR&R actions (the negative of these values should be used as the reward in the POMDP formulation) and the potential failure costs are given in Tables (1) and (2), and adapted from actual data in PONTIS and experts judgment provided by Al-Wazeer (2007). *The* total number of MR&R actions for girders in condition states 1 through 5 are 2, 3, 3, 3 and 2, respectively. For the concrete deck, the action set size is taken as 4 for all condition states. Hence, the total number of possible MR&R actions considered for the Super States is $2 \times 3 \times 3 \times 3 \times 2 \times 4 = 432$. In addition, three inspection strategies for both the steel girders and the concrete deck are considered, including "Do not observe", "Visual inspection", and "Ultrasonic test" for girders and "Do not observe", "Visual inspection", and "half-cell potential method" for the deck (see Table 3 for the associated cost values). Taking into account that visual inspections for the concrete deck and the girders are practically carried out simultaneously, the total number of possible inspection combinations will be five. Thus, the total number of Super Action combinations becomes $432 \times 5 = 2160$. Though engineering judgment, the observation transition probabilities as well as the corresponding inspection costs are adapted from Frangopol et al. (1999) and Daher (2004).

### *4.2. Results and discussion*

Random walks procedure for generating likely belief points, can result in variation in the optimal value functions. To evaluate the extent of variation in the optimal value function and the robustness of the framework, the applied POMDP framework combined with the counting process technique is carried out 12 times. Each run includes 10000 randomly generated belief points. These belief points were generated by considering uniform distribution functions for actions and observations within each stage of the random walks. The convergence track for these 12 runs in terms of the average expected discounted cost vs the CPU runtime is depicted in Figure 2. As for the stopping criteria explained in section 2, the tolerance and the maximum CPU runtime are set at 5e-3 and 21.5 hours, respectively. Out of the 12 runs, 3 of them converged, while the other 9 runs stopped by the runtime limit. The plots in Figure 2 show that in general, variation in the expected cost value for the 12 different runs is not significant. That is to say, the maximum coefficient of variation is 6% which occurs at t~20000s. This quantity becomes 2% near the stopping time, i.e. the dispersion around the average curves reduces as the number of iterations increases.

Figure 2 also shows the average long-run cost associated with an extreme scenario where the best preventive actions are applied each year. Clearly, employing the POMDP decision-making framework results in lower expected costs in the lifetime of the bridge system.

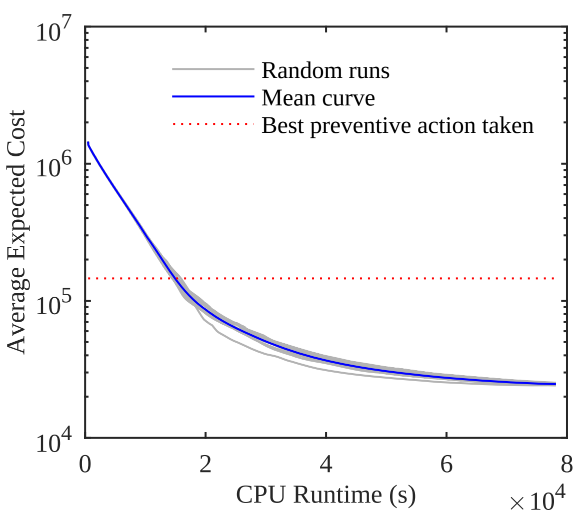


*Figure 2: The average expected discounted cost and the mean of the average expected discounted cost vs the CPU runtime for the 12 runs.*

As explained before, the main objective of the proposed decision-making framework is to determine the optimal long-run steady-state policies for the randomly generated belief points. In Figure 3, a sample run is selected and the optimal decisions are investigated. For this simulation, the framework reveals seven Super Actions that results in optimal solutions for the 10000 belief points. For each of the seven optimal policies, the PMF of a sample belief point is shown in Figure 3. The frequency values of these optimal policies as well as their descriptions are also shown in Table 4. According to Figure 3, for some belief points, many Super States contribute considerably to the PMF values (the significant Super States for each of the seven belief points are indicated in Table 5). Unlike the case in the MDP framework which considers the PMF is equal to one at the observed condition state and zero elsewhere, the plots in Figure 3 indicate that, in reality with imperfect observations, there are several condition States that can be true with different likelihoods, even if the inspection results point out to a particular state.

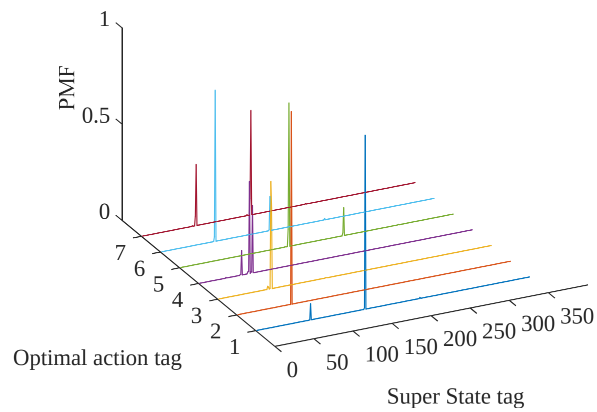


*Figure 3: The PMFs of a sample belief point for each of the seven optimal policies*

In Table 5, the optimal actions corresponding to each of the observed condition states for the girders and the optimal action for the concrete deck are provided. The optimal inspection strategy for the following decision-making year, both for girders and the deck are shown as well.

It is worth noting that the utilized steady-state POMDP framework provides optimal strategies for belief points that may be visited at any decision-making time during the lifetime of the infrastructure. Such belief points denote how probable each Super State can be the true condition of the system. An example of a simple decision-making problem can be explained as follows.

*Table 1: Cost and transition probability matrices of a girder component under various MR&R actions*

| Condition state at time "t" | | Girder components | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | State description | Action possibilities | Action cost ($/ft) | State risk cost($/ft) | Condition state at time "t+1" | | | | |
| | | | | | | 1 | 2 | 3 | 4 | 5 |
| | 1 | 1% Section loss | DN | 0 | 2.91 | 0.97 | 0.03 | 0.00 | 0.00 | 0.00 |
| | | | SC | 10 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 2 | 5% Section loss | DN | 0 | 8.14 | 0.00 | 0.94 | 0.06 | 0.00 | 0.00 |
| | | | SC | 15 | | 0.10 | 0.90 | 0.00 | 0.00 | 0.00 |
| | | | CP | 40 | | 0.95 | 0.05 | 0.00 | 0.00 | 0.00 |
| | 3 | 10% Section loss | DN | 0 | 28.15 | 0.00 | 0.00 | 0.91 | 0.09 | 0.00 |
| | | | RP | 55 | | 0.40 | 0.30 | 0.20 | 0.10 | 0.00 |
| | | | SCP | 65 | | 0.90 | 0.10 | 0.00 | 0.00 | 0.00 |
| | 4 | 15% Section loss | DN | 0 | 90.81 | 0.00 | 0.00 | 0.00 | 0.88 | 0.12 |
| | | | RP | 65 | | 0.40 | 0.20 | 0.10 | 0.30 | 0.00 |
| | | | SCP | 75 | | 0.90 | 0.05 | 0.05 | 0.00 | 0.00 |
| | 5 | 20% Section loss | DN | 0 | 266.70 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| | | | MR | 200 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Note: DN=Do Nothing, RP=Replace Paint system, MR=Major Rehabilitation, CP=Clean and Paint, SC=Surface Clean, SCP=Spot blast.

*Table 2: Cost and transition probability matrices of the concrete deck under various MR&R actions*

| Condition state at time "t" | | Deck component | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Action possibilities | Action cost ($/ft^2) | State risk cost($/ft^2) | Condition state at time "t+1" | | | | |
| | | | | | | 1 | 2 | 3 | 4 | 5 |
| | 1 | 0.01" Crack width | DN | 0 | 0.27 | 0.95 | 0.05 | 0.00 | 0.00 | 0.00 |
| | | | APS | 9 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 2 | 0.03" Crack width | DN | 0 | 0.52 | 0.00 | 0.90 | 0.10 | 0.00 | 0.00 |
| | | | RSD | 5 | | 0.90 | 0.10 | 0.00 | 0.00 | 0.00 |
| | | | APS | 10 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 3 | 0.05" Crack width | DN | 0 | 1.13 | 0.00 | 0.00 | 0.85 | 0.15 | 0.00 |
| | | | RSD | 6 | | 0.80 | 0.10 | 0.10 | 0.00 | 0.00 |
| | | | RSDAPS | 12 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 4 | 0.07" Crack width | DN | 0 | 4.63 | 0.00 | 0.00 | 0.00 | 0.80 | 0.20 |
| | | | RSD | 7 | | 0.70 | 0.10 | 0.10 | 0.10 | 0.00 |
| | | | RSDAPS | 15 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 5 | 0.10" Crack width | DN | 0 | 22.25 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| | | | RSDAPS | 20 | | 0.95 | 0.05 | 0.00 | 0.00 | 0.00 |
| | | | RD | 30 | | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Note: DN=Do Nothing, RSD=Repair Spalls and Delaminations, RSDAPS=Repair Spalls and Delaminations and Add a Protective System, RD=Replace Deck.

*Table 3: Cost values for inspection strategies ($)*

| Girder Components | | | Concrete Deck | | |
|---|---|---|---|---|---|
| Do not observe | Visual inspection | Ultrasonic inspection | Do not observe | Visual inspection | Half-cell potential inspection |
| 0 | 150 | 400 | 0 | 150 | 400 |

*Table 4: The frequency values and the descriptions of the optimal policies*

<table>
<tr><th rowspan="3">Optimal Policy Tag</th><th rowspan="3">Belief Point Frequency</th><th colspan="8">Policy description</th></tr>
<tr><th colspan="6">Optimal action</th><th colspan="2">Optimal inspection strategy for the following year</th></tr>
<tr><th>Girders in State 1</th><th>Girders in State 2</th><th>Girders in State 3</th><th>Girders in State 4</th><th>Girders in State 5</th><th>Concrete deck</th><th>Girders</th><th>Concrete Deck</th></tr>
<tr><td>10</td><td>21</td><td>DN</td><td>DN</td><td>RP</td><td>RP</td><td>MR</td><td>DN</td><td>DNO</td><td>DNO</td></tr>
<tr><td>41</td><td>9309</td><td>DN</td><td>CP</td><td>DN</td><td>SCP</td><td>DN</td><td>DN</td><td>DNO</td><td>DNO</td></tr>
<tr><td>107</td><td>105</td><td>SC</td><td>CP</td><td>SCP</td><td>SCP</td><td>DN</td><td>DN</td><td>DNO</td><td>DNO</td></tr>
<tr><td>108</td><td>25</td><td>SC</td><td>CP</td><td>SCP</td><td>SCP</td><td>MR</td><td>DN</td><td>DNO</td><td>DNO</td></tr>
<tr><td>322</td><td>11</td><td>SC</td><td>CP</td><td>SCP</td><td>RP</td><td>MR</td><td>RSD</td><td>DNO</td><td>DNO</td></tr>
<tr><td>918</td><td>508</td><td>DN</td><td>CP</td><td>SCP</td><td>SCP</td><td>MR</td><td>DN</td><td>VI</td><td>VI</td></tr>
<tr><td>1314</td><td>21</td><td>DN</td><td>DN</td><td>SCP</td><td>SCP</td><td>MR</td><td>DN</td><td>DNO</td><td>HP</td></tr>
</table>

Note: DN=Do Nothing, RP=Replace Paint system, MR=Major Rehabilitation, CP=Clean and Paint, SC=Surface Clean, SCP=Spot blast, Clean and Paint, RSD=Repair Spalls and Delaminations, DNO= Do Not Observe, VI= Visual Inspection, HP=Half-cell Potential inspection

*Table 5: The frequency values and the descriptions of the optimal policies*

<table>
<tr><th rowspan="2">Sample Belief point</th><th rowspan="2">Optimal Policy</th><th rowspan="2">Most Likely super state tag</th><th rowspan="2">PMF</th><th colspan="6">Super State description</th></tr>
<tr><th># Girders in State1</th><th># Girders in State2</th><th># Girders in State3</th><th># Girders in State4</th><th># Girders in State5</th><th>Deck State</th></tr>
<tr><td>1</td><td>10</td><td>140</td><td>0.90</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr>
<tr><td>2</td><td>41</td><td>70</td><td>1.00</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">107</td><td>68</td><td>0.56</td><td>3</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>69</td><td>0.41</td><td>3</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td rowspan="3">4</td><td rowspan="3">108</td><td>55</td><td>0.13</td><td>1</td><td>3</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>65</td><td>0.48</td><td>2</td><td>2</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>69</td><td>0.35</td><td>3</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2">322</td><td>140</td><td>0.74</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr>
<tr><td>210</td><td>0.14</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>3</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2">918</td><td>70</td><td>0.78</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>140</td><td>0.18</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr>
<tr><td rowspan="2">7</td><td rowspan="2">1314</td><td>70</td><td>0.32</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>140</td><td>0.54</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr>
</table>

According to the history of the observations gained and actions taken on the bridge system, if the belief state matches the random belief state 322, the agency is recommended to do a "surface clean", "clean and paint", "spot blast, clean and paint", "replace paint system", and "a major rehabilitation" for the girders observed in states 1 to 5, respectively. In addition, for the concrete deck, "repair spalls and delamination" is suggested. As an interpretation, the most likely true condition of the four girders is state 1, while for the deck, it is state 2.

## 5. CONCLUSION AND DISCUSSION

A Partially Observable Markov Decision Process combined with a "counting process" technique is proposed for the optimal decision-making of infrastructure systems. The framework is implemented on a bridge system consisting of multiple girders and a concrete deck. The failure probability and its consequences in terms of failure risk cost are considered in monetary units. Hence, as a compromise between MR&R costs, and user, agency and failure risk costs, the applied decision-making framework introduces the least costly strategies for a long-run planning horizon. The optimal strategies entail MR&R actions to be taken at the current year, as well as the inspection technology to be applied for the next year.

- The proposed framework enables element-level decision-making of large-scale multi-state multi-component infrastructure systems through:
- Applying a randomized point-based value iteration POMDP, in which optimal strategies are derived for the belief states that are most likely to be visited in the lifetime of the system under consideration.
- Employing the "counting process" state reduction technique. Through this approach, for a system consisting of multiple elements, the quantity of components in each of the states is considered instead of exploring all the possible combinations of elements' condition states.

The above procedure was applied to a case study bridge system with four steel girders and one concrete deck. Results of 12 random runs showed acceptable convergence in the optimized average expected long-run reward, with less than 22 hours of runtime. The applied framework provides optimal policies for the concrete deck and girders in each of the possible states.

As an outcome of the framework, the list of optimal policies for the likely randomly generated belief points is provided. Corresponding to each of these optimal policies, a sample belief point is opted and elaborated.

The proposed framework provides optimal policies for infrastructure systems at the element-level, while forecasting uncertainty and measurement randomness are incorporated in a time-efficient manner. Thus, the proposed method can give rise to an efficient and accurate approach for the optimal management of large scale systems.

## 6. ACKNOWLEDGEMENTS

Authors would like to thank Dr. Matthijs Spaan for his assistance with the decision-making algorithm. This work was supported in part by an allocation of computing time from the Ohio Supercomputer Center. This support is greatly appreciated.

## 7. REFERENCES

AASHTO. (2005). Pontis Bridge Management release 4.4 User manual. Washington, D.C.: AASHTO.

Al-Wazeer, Adel Abdel-Rahman. "Risk-based bridge maintenance strategies." PhD Dissertation. University of Maryland. (2007). Internet resource.

Cassandra, A.R., (1999). Tony`s POMDP page. (Accessed 2014). Available from: http://cs.brown.edu/research/ai/pomdp/

Daher, B.W., "Use of sensors in monitoring civil structures". Master's Thesis. Massachusetts Institute of Technology. Dept. of Civil and Environmental Engineering. (2004).

Ellis, H., Jiang, M., Corotis, R., (1995). "Inspection, maintenance, and repair with partial observability". *Transportation Engineering* 1(2), 92–99.

Frangopol, D.M., Estes, A.C., (1999). "Optimum lifetime planning of bridge inspection and repair programs". *Structural Engineering International* 9(3), 219-223.

Kuhn, K., (2010). "Network-Level Infrastructure Management Using Approximate Dynamic Programming". *Infrastructure Systems* 16(2), 103-111.

Madanat, S., Ben-Akiva, M., (1994). "Optimal Inspection and Repair Policies for Infrastructure Facilities". *Transportation Science* 28(1), 55.

Monahan, G.E., (1982). "State of the Art. A Survey of Partially Observable Markov Decision Processes: Theory, Models, and Algorithms". *Management Science* 28(1), 1-16.

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

Spaan, M.T.J., Vlassis, N., (2005). "Perseus: Randomized Point-based Value Iteration for POMDPs". *Artificial Intelligence Research* 24, 195-220.