**CAAP Quarterly Report**

**7/8/2024**

*Project Name: Pipeline Risk Management Using Artificial Intelligence-Enabled Modeling and Decision Making*

*Contract Number: 693JK32150001CAAP*

*Prime University: Rutgers University*

*Prepared by: Bingyan Cui (PhD student), Emad Farahani (PhD student), Dr. Hao Wang (PI), Dr. Qindan Huang (Co-PI)*

*Reporting Period: 4/1/2024 – 6/30/2024*

**Project Activities for Reporting Period:**

*Task 1 Literature Review (Completed)*

*Task 2 Data Collection from Industry Partners (Completed)*

*Task 3 Data-Driven Probabilistic Modeling of Pipeline Defects (Completed)*

*Task 4 Quantification of Probability of Failure (Completed)*

***Task 5 Decision Making of Inspection and Repair Strategy using Reinforcement Learning***

To ensure certain level of reliability in pipeline risk management, an AI-based reinforcement learning (RL) method will be used for pipeline maintenance planning, including maintenance timing, methods, and locations.

Fig. 1 shows the proposed framework of pipeline maintenance strategy using RL. The RL environment is the test bench of pipeline system, including the pipeline environmental parameters collection, corrosion growth model development and reliability model. The agent represents the deep Q-learning decision-maker. Actions consist of possible maintenance action, such as doing nothing, composite sleeve with different thickness and materials, and replacement. Under the influence of the action, the current state transfers and a reward can be obtained. By properly choosing the actions sequentially, the goal is to maximize the expected cumulative rewards. The constraint is to satisfy the failure probability to ensure safe pipeline operations, and the objective is to minimize the total cost of maintenance. In each simulation episode, a pipeline system is created by a stochastic environment. Then, at the end of each month, the agent will check whether the failure probability exceed the threshold. After that, the agent will choose the action from its policy until the simulation is ended. Thereafter, the program will loop more episodes until it reaches the assigned number of episodes. The following sections describe each component in detail.



**Fig. 1** Proposed framework for pipeline maintenance using reinforcement learning.

**1. Agent**

Deep Q-learning (DQL) will be employed to do the decision-making. It is a neural network as the value function to approximate the Q-value. By combining the advantages of both neural networks and Q-learning, DQL is able to manage large-scale sequential decision-making problems. The model-free algorithm chooses an action following the Action value function (Q-value) for any finite Markov decision process (MDP) and then optimize itself. The Q-value can be defined as Eq. (1).

  (1)

where, *st* is the state at time *t*; *at* is the action at time *t*; *rt* is the reward at time *t*; *γ* is the discounted factor between 0 and 1.

The objective of Q-learning is to optimize the Q-value and obtain the optimal preservation policy. DQL combines neural network and Q-learning algorithm, which is capable of making large-scale sequential decisions. The loss function is defined by the approximation function of *Q*(*s, a, ω*) = *Qπ*(*s, a*) to measure the difference between target values and predicted values according to Eq. (2).

  (2)

where, *ω* is the model parameters; *ω-* is the parameters in the target network.

*ε*-greedy policy is added to avoid the exploration-exploitation dilemma as shown in Eq. (3). Even if the agent chooses the non-optimal action, it can still learn to use it. Therefore, the agent must explore other actions for optimizing its policy. However, it should not always explore other actions as it may exploit the policy that can be optimal at any time.

  (3)

where, *ε* is 0.8 at first and then decreased to 0.01 constantly.

The procedure of using DQN algorithm can be seen in Fig. 2.

**2. RL Environment**

The Test Bench comprises three components designed to simulate the pipeline system environment and generate numerical data for interaction with the RL system. First, the relationship between soil environmental parameters and corrosion depth and length will be identified. Next, a corrosion growth model will be developed using a Bayesian neural network (BNN), based on our previous work. Finally, a reliability model will be established to calculate the failure probability of the pipeline and assess the status of the pipeline system.

*3.1 Corrosion growth model*

As established in our previous work, the corrosion growth model is based on Bayesian neural network (BNN). Unlike traditional models that learn fixed parameters, BNN model calculates the conditional distribution of weights given the training data. This distribution, known as the posterior distribution, as defined in Eq. (4).

  (4)

where, *w* represents all the weights and biases; *D*train refers to the training data; *P*(*w*) symbolizes the prior distribution; *P*(*Dtrain*|*w*) represents the likelihood; *P*(*Dtrain*) is the evidence.

Then, Kullback-Leibler (KL) divergence will be used to measure the similarity between the true posterior distribution and the proposed distribution, as expressed in Eq. (5).

  (5)

where, Eq[] is the expectation of distribution *q*; ln represents natural logarithm; *q*(*w*)is the proposed distribution; *P*(*w|D*) is the actual posterior distribution.



**Fig. 2** The flowchart of using DQN for pipeline maintenance.

*3.2 Reliability model*

The reliability of pipelines will be evaluated by the probability of failure (*Pf*). Therefore, the reliability model involves calculating *Pf*, including both leakage and burst failures. The probability of failure is defined as the conditional probability of reaching or exceeding the specified limit state of the pipelines within a given condition boundary.

Small leak failure will occur when the corrosion depth exceeds the pipeline thickness. Its probability of failure can be calculated as shown in Eq. (6).

  (6)

where, *dw* is the thickness of pipeline; *d*(*t*) is the maximum corrosion depth at time *t*.

Burst failure refers to a scenario where the operation pressure of a pipeline exceeds its pressure capacity. The probability of such a failure can be quantified as presented in Eq. (7).

  (7)

where, *Pb*(*t*) is the burst pressure capacity of pipeline at time *t*; *Pp* is the operation pressure.

**3. State Space Defining**

The states in a Markov Decision Process must be defined to satisfy the Markov property. A state is a complete description of the environment at a given time, including all necessary information to predict the next state given an action. The state space consists of all possible states of the environment, representing every possible configuration the environment can be.

Given that this research focuses on pipeline corrosion, the state definition should include all essential factors to predict the pipeline future status. Therefore, the state space is discrete, including soil parameters, pipeline parameters, corrosion parameters, and the reliability index, as listed in Table 1.

**Table 1.** Features in the state space

|  |  |  |
| --- | --- | --- |
| **Category** | **Symbol**  | **Description** |
| Soil parameters | ST | soil type |
| Eh | soil redox potential |
| pH | pH value  |
| CO3 | CO32- concentration |
| HCO3 | HCO3- concentration |
| Cl | Cl- concentration |
| SO4 | SO42- concentration |
| SM | Soil moisture |
| R@1m | soil resistivity at the depth of 1m |
| R@2m | soil resistivity at the depth of 2m |
| Pipeline parameters | wt | Wall thickness |
| age | Pipeline age |
| ele | Pipeline elevation |
| Corrosion parameters | dt | Corrosion depth |
| dl | Corrosion length |
| Reliability index | Pf | Probability of failure |

**4. Action and Reward Defining**

An action is a decision or move made by an agent in response to a state of the environment. Actions determine how the agent interacts with the environment to achieve its goals. The action space is defined by the available actions that the maintenance scheduler can take in response to the system state. As presented in the maintenance actions section, the benchmark provides the scheduler with three actions. Therefore, a discrete action space of size three is considered for the agent: {Do nothing, Composite sleeve, Replacement}.

A reward is a scalar feedback signal received by the agent after taking an action in a given state. It represents the immediate benefit and cost of the action. The goal of the agent is to maximize the cumulative reward over time. The reward function defines the reward received for each action taken in each state, mapping state-action pairs to a real number. This reward function guides the agent towards desirable behaviors. In this study, the rewards will be estimated by combining the negative costs and positive benefits.

*4.1 Cost estimate*

The total cost of pipelines includes the inspection, repair and failure costs. Using the total probability rule, the expected value of total cost can be calculated as Eq. (8).

  (8)

where, *CT* is the total cost; *CT,k* is the cost due to the scenario when *k* failures occur during the lifetime considered; *Pf,k* is the probability of *k* failures occurrence during the time span considered; *E*[*CT*|*k* failures] is the expected value of total cost given *k* failures occurrence; *n* is the number of failure occurrences during the lifetime considered.

When *k*=0, the expected total cost only consists of inspection and repair cost. Therefore, Eq. (8) can be simplified to Eq. (9).

  (9)

where, *ts* is the total time considered; E[Cin|0 failure] and E[CT|0 failure] are the expected cost of inspection and repair given no failure, respectively; P*f,0* is the probability of no failure occurrence during *ts*.

Considering the discount rate, the expected value of costs of inspection and repair can be calculated as Eq. (10) and (11).

  (10)

  (11)

where, *nin* is the total number of inspections during *ts*; *nb* is the total number of branches prior to inspection time *ti*; CI and CR is the unit costs of inspections and repairs, respectively; P*f, 0*(*ti*)=1-P*f*(*ti*) is the probability of no failure occurrence before inspection time *ti*; P*f, 0, j* (*ti*) is the probability of failure occurrence in each branch j before inspection time *ti*; P(*Ri, j*) is the probability of performing a repair action at inspection time *ti* based on the detected defect size in branch *j*.

When *k*=1, indicating there is one and only one failure occurrence during [0, *ts*]. One can divide [0, *ts*] into two periods [0, *Tf1*) and (*Tf1*, *ts*] given that the 1st failure occurs at time *Tf1*. With the assumption that a replacement will take place right after the failure and the pipeline is reset to be the initial state, the second period can be considered (0, *ts-Tf1*]. Since only one failure occurs, the costs for these two periods can be calculated based on Eq. (9) that is when there are no failure occurrences, which are CNF(*Tf1*) and CNF(*ts*-*Tf1*), respectively. Thus, the expected value of total cost when there is one and only one failure, E[CT,1], is the summation of the failure cost, CNF(*Tf1*), and CNF(*ts-Tf1*). Considering *Tf1* as a random variable, E[CT,1] is calculated by Eq. (12).

  (12)

where, *Tf1* is the occurrence time of 1st failure; P*f,1* is probability of one and only one failure occurrence during the time considered; *f*(*tf1*) is the probability density function of *Tf1*.

When *k*=2, one can divide [0, *ts*] into three periods [0, *Tf1*), (*Tf1*, *Tf2*] and (*Tf2*, *ts*]. Therefore, the total costs can be calculated as Eq. (13).

 (13)

where, P*f,2* is the probability of two and only two failure occurrences during the time considered.

*4.2 Benefit estimate*

The major benefits of pipeline maintenance include the life-extension reward, failure-reduction reward, and operation-efficiency reward.

Life-extension reward represents the economic value of the extended service life of pipelines, as calculated in Eq. (14).

  (14)

where, *te* is the extended life in years; *ca* is the annual maintenance cost; *r* is the discount rate.

Improved conditions can reduce the frequency of pipeline failures, thereby lowering the associated repair costs and environmental penalties. The calculation of failure-reduction reward can be seen in Eq. (15).

  (15)

where, *nr*is the number of failure reductions; *cf* is the average cost per failure; *t* is the evaluation years.

Proper maintenance can prevent unexpected downtime, leading to continuous operation and higher productivity. The benefits of operation-efficiency reward can be calculated as Eq. (16).

  (16)

where, *ce* is the total saved energy cost; *cd* is the saved downtime costs.

**Project Activities with Cost Share Partners:**

Cost share is provided by Rutgers University and Marquette University during this quarterly period as budgeted in the proposal.

**Project Activities with External Partners:**

N/A

**Potential Project Risks:**

Due to additional time and effort spent on Task 2 for data collection and Task 3 for model development and refinement, one-year no-cost extension is planned to extend the project date to 9/30/2025.

**Future Project Work:**

Work will be continued on Task 5 on decision making of inspection timing and repair strategy.

**Potential Impacts to Pipeline Safety:**

The AI-enabled modeling and analysis of pipeline inspection data will be used to develop probabilistic growth models of corrosion defects and make cost-effective repair or replacement decisions to minimize pipeline failure risk.