PROGNOSTIC AND HEALTH MANAGEMENT IN OCEAN ENERGY SYSTEM: A SELF-HEALING FRAMEWORK BASED ON REINFORCEMENT LEARNING

Yu Huang, Yufei Tang*, and James VanZwieten
Florida Atlantic University
Boca Raton, Florida, USA

Feng Wu
Hohai University
Nanjing, Jiangsu, China

*Corresponding author: tangy@fau.edu

INTRODUCTION

Research efforts have been focused on harvesting electricity from renewable ocean energy in a commercially and technologically acceptable manner [1]. Since the harsh and remote working environment, one of the major issues in cost-effectively integrating renewable ocean energy into power grids is the prognostic and health management (PHM) of multiple offshore/inshore devices, which drives the need for facilitating Systems-Level Thinking in PHM system [2]. This requires developing reliable self-prognostic and self-decision-making techniques which could account for both the complexity of the asset and the uncertainties on its operational conditions, failure modes, degradation behaviors, external environment, etc.

In this paper, for minimizing the cost from the ocean generator power production by optimizing the operation and maintenance (O&M) policy over an infinite time horizon, while considering the uncertainty of the renewable sources and components failure behaviors, we develop a self-healing framework for ocean energy systems, shown in Figure 1. It consists of three major modules: data manipulation, health assessment and decision-making. Specifically, a graph theoretic approach is first proposed for ocean generator health monitoring utilizing multivariate time-series data, then, reinforcement learning (RL) based technique exploits the health states of system that provides decision support for optimal O&M management.

SELF-HEALING PHM SYSTEM

A self-healing PHM system automatically integrates the results from the well-designed sensor net all the way through to the decision-making module that provides support for optimal use of O&M resources. The core of this strategy is based on: 1) accurately forecasting the onset of imminent health conditions or failures of critical components and 2) efficiently spotting the root cause of failures once effects have been detected. From this perspective, if health conditions/failures predictions can be made, the allocation of preventive or corrective actions can be scheduled in an optimal fashion.

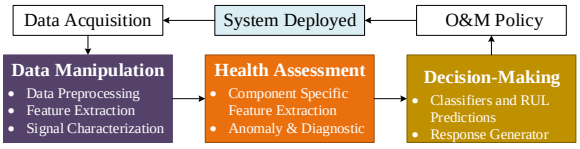

FIGURE 1. SELF-HEALING PHM SYSTEM.

DATA MANIPULATION & HEALTH ASSESSMENT

The aim of data manipulation is to represent a multivariate time series system measurement $\mathcal{X}$ as a lower-dimensional weighted and undirected network graph that contain sufficient degradation/failure signatures in order to increase the efficiency and reliability of health assessment. This approach involves the following steps:

1. Consider a segmented signal $x_i = \left[x_i^1 \ \cdots \ x_i^h\right]^T$, corresponding to known status $s_i, i = 1, \cdots, n$ with each window x_i^j a $k \times d$ matrix.
2. Transform the signal x_i into a weighted and undirected network graph $G(x_i) \equiv (\mathcal{V}, \mathcal{E}, \mathcal{W})$. The nodes $\mathcal{V}$ are the rows and columns of the symmetric similarity matrix $S^{k\times k} = \left[\omega_{pq}\right]$, where the pairwise ω_{pq} is computed by Mahala Nobis kernel Ω for each window: $\omega_{pq} = \Omega\left(x_p, x_q\right) \forall p, q \in (1,2,\cdots k)$ and the correlation between each pair of nodes is indexed by edges, i.e., connection status $\mathcal{E}$ and weights $\mathcal{W}$.

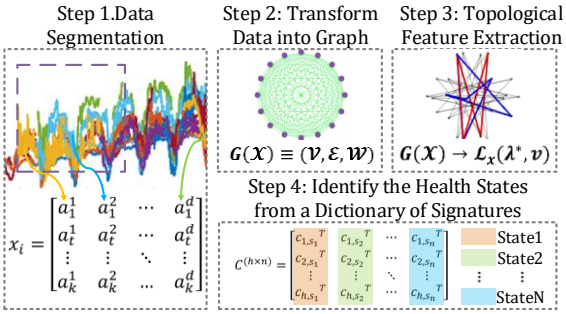

FIGURE 2. DATA MANIPULATION & HEALTH ASSESSMENT USING GRAPH THEORETIC METHOD.

3. Extract the spectral graph Laplacian matrix $\mathcal{L}_{x_i}(\lambda^*, v)$ from x_i once it transformed into a graph $G(x_i)$. The transformation from signal x_i corresponds to status s_j to the spectral graph is: $G(x_i) = \left[\mathcal{L}_{x_i^1} \cdots \mathcal{L}_{x_i^h}\right]^T$ which employed to capture the inherent dynamics of the signal.
4. Select an orthogonal subset of the graph Laplacian Eigenvectors as a basis set corresponding to health state s_i. Each x_i is decomposed by taking an inner product $x_i^T v_i$ akin to a Fourier transform into a set of coefficients c_i. Repeat this procedure for all status $s_i, i = 1,2,\cdots,n$, a dictionary of c_i can be formed as:

$$\mathbb{C}^{h\times n} = \begin{bmatrix} x_i^{1^T} V_{s_1} = c_{1,s_1}^T & \cdots & x_i^{1^T} V_{s_n} = c_{1,s_n}^T \\ \vdots & \ddots & \vdots \\ x_i^{h^T} V_{s_1} = c_{h,s_1}^T & \cdots & x_i^{h^T} V_{s_n} = c_{h,s_n}^T \end{bmatrix} \quad (1)$$

5. Given an unknown signal segment $y^{k\times d}$, obtain the candidate set by an inner product $y^T V_{s_i}$, that is $\hat{\mathbb{C}} = \left[\hat{c}_{s_1}^T \quad \cdots \quad \hat{c}_{s_n}^T\right]$. Then compare each $\hat{c}_{s_i}^T$ with associated coefficients $c_{s_i}^T$ (having the same label s_i) in the dictionary $\mathbb{C}$. The label assigned to y is the one with the minimum squared errors e, i.e. $s_i = \text{argmin}_{s_i} e_{s_i}$.

REINFORCEMENT DECISION-MAKING

Developing a reinforcement learning based decision-making module requires defining the environment and its stochastic behavior, the actions that the agent can take in every state of the environment and their corresponding effects and reward generated [3].

Environment state: Consider a system consists of elements $C = \{1, \ldots, N\}$, physically or functionally interconnected. The degrading elements $d \in D \subseteq C$ are affected by independent degradation mechanisms, obeying a Markov process that models the stochastic transitions from current state $s_i^d(t)$ to the next state $s_i^d(t+1)$, where $s_i^d(t) \in \{1, \cdots, n\}$, $\forall t, d \in D, i = 1, \cdots, n$. These degradation states are estimated by the health assessment modules. At each time t, the system state vector reads as $S_t = [s^1(t), s^2(t), \cdots, s^d(t)]$. Assume that the stochastic behavior of the environment is completely defined by transition probability matrices of each element $d = 1, \cdots, |D|$ and to each action a $\in \mathcal{A}$, that is,

$$\mathcal{P}_d^a = \begin{bmatrix} p_{1,1} & \cdots & p_{1,s^d} \\ \vdots & \ddots & \vdots \\ p_{s^d,1} & \cdots & p_{s^d,s^d} \end{bmatrix}, \sum_{j=1}^{s^d} p_{i,j} = 1 \quad (2)$$

where $p_{i,j}$ represents the probability $\mathcal{P}_d^a(s_j|a, s_i)$ of having a transition of element d from state i to state j, conditional to the action a.

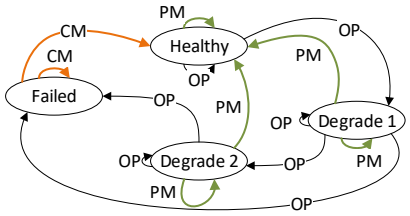

FIGURE 3. THE MARKOV DECISION PROCESS OF OCEAN TURBINE WITH ASSOCIATED STATES.

Actions: Action a_g can be performed on the system element g ∈ G ⊆ C. The action vectors a at time t is $a_t = [a_{g_1}(t), \cdots, a_{g_G}(t)] \in \mathcal{A}$. The action set $\mathcal{A}$ includes both operational actions (OM), preventive maintenance (PM) and corrective maintenance (CM) actions. CM is to fix an out-of-service faulty condition to an in-service healthy condition and PM is to improve the condition of an in-service but degraded element. Additional constraints can be defined, considering that some actions are disallow in particular states, e.g., CM is the only allowed for failed elements. Both PM and CM are assumed to restore the healthy state for each degraded element (Figure 3).

Reinforcement learning: The goal of the agent for strategy optimization is to obtain the optimal action-value function, which is the maximum sum of rewards r_t discounted by γ at each time step t, achievable by a behavior policy $\pi = P(a|s)$, after making an observation s and taking an action a:

$$Q^*(S,a) = max_\pi E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots | s_t = s, a_t = a, \pi] \quad (3)$$

and the algorithm for training Deep Q Network could be referred to [3] and shown in Algorithm 1.

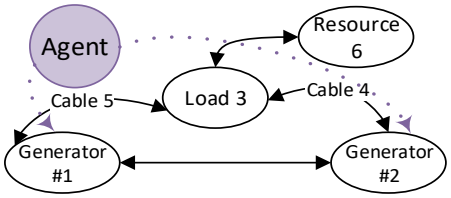

FIGURE 4. THE SIMPLIFIED OCEAN GENERATION SYSTEM WITH TWO TURBINES SUPPLYING A LOAD.

CASE STUDY

The proposed self-healing framework is applied to a scaled-down ocean power system (Figure 4). The system consists of 2 controllable generators, 1 energy source providing electricity, 1 connected load depending on random conditions and

4 transmission cable. The generators, cable 4 and 5, are under degradation and equipped with PHM capabilities to inform the decision-maker on their states. We consider 4 degradation states for generators, $s^{d=1,2} = \{1, 2, 3, 4\}$, shown in Figure 3. For the load/energy resource, we consider 3 states of rising power demand/production. For cables, 3 degrading states are defined. We assume that both generators have identical transition probability matrices (similar to [4]), and the cables degradation are described by the same Markov process. Hence, $S_t = [s^1, s^2, s^3, s^4, s^5, s^6]$ and the state space is made up of 1296 points. We defined 5 actions (3 OM, 1PM, 1CM) that can be applied to generators while keeping the system's structural and functional integrity. The action vector reads $a = [a_1, a_2]$ $a_{1,2} = \{1, \dots, 5\}$. This gives rises to 32400 state-action pairs. Each action has a specific transition probability matrix, describing the generator degradation conditioned by its operative state or maintenance action. The case-specific reward is made up of 3 contributions: the cost of demanded power from ocean generators, the cost of producing electricity by ocean generators and the cost of the performed actions, and is formulated as:

$$R_t = \sum_{i=1}^{2} [P_d - P_r] c_{el} - \sum_{i=1}^{2} P_g^i \, c_g - \sum_{i=1}^{2} c_{a-c}^i \qquad (4)$$

where c_{el} is the unit price of ocean generator produced power, P_g is the power produced by ocean generator with unit cost c_g, and c_{a-c} is the cost of actions.

Algorithm 1: Deep Q-learning

Initialize memory D to capacity N, action-value function Q with random weights θ, Initialize target action-value function $\hat{Q}$ with weights $\theta^- = \theta$

For episode =1, ..., M **do**

  Initialize sequence $s_1 = \{x_i\}$ and preprocessed sequence $\emptyset_1 = \emptyset(s_1)$

  For t=1, ..., T **do**

    With probability ε select random action a_t, otherwise $a_t = argmax_a Q(\emptyset(s_t), a; \theta)$;

    Execute a_t, observe reward r_t and x_{t+1};

    Set $s_{t+1} = s_t, a_t, x_{t+1}$ and $\emptyset_{t+1} = \emptyset(s_{t+1})$;

    Store $(\emptyset_t, a_t, r_t, \emptyset_{t+1})$ in D;

    Sample random batch of $(\emptyset_j, a_j, r_j, \emptyset_{j+1})$

$$y_j = \begin{cases} r_j & if\ episode\ terminates\ at\ j+1 \\ r_j + \gamma max_{a'} \hat{Q}(\emptyset_{j+1}, a'; \theta^-) & otherwise \end{cases}$$

    Perform a gradient descent step on $(y_j - Q(\emptyset_j, a_j; \theta))^2$ with respect to θ;

    Every C step reset $\hat{Q} = Q$

  End for

End for

RESULTS

The ocean generator's condition identification results with high accuracy (Table 1) verified that the proposed graph theoretic method is a reliable health assessment method. The RL results are summarized in Figure 5, by visualizing the distribution of $Q_{\pi^*}(S, a)$ over the states for all the combination of action $a = [a_1, a_2]$. According to the empirical CDF, we can identify three clusters: the states set (1 curve) for which both generators are under CM, the states set (8 curves) for which only one of the generators is under CM and the states set (16 curves) includes only PM and operational actions. CM is a costly action and leads to negative expectation of reward, whereas PM and operational action leads to higher positive expectation of reward.

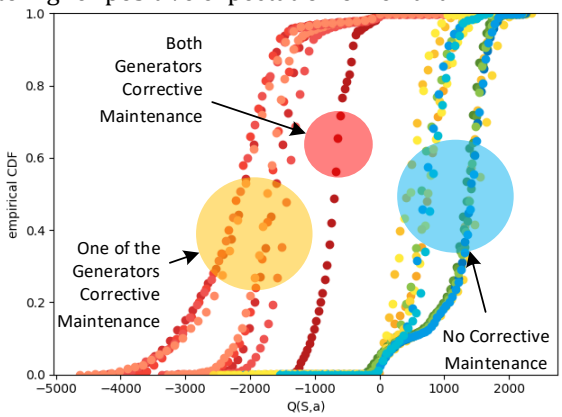

FIGURE 5. THE STATE-ACTION VALUE VISUALIZATION USING ECDFS WITH 3 CLUSTERS.

TABLE 1. HEALTH ASSESSMENT RESULT.

F-score=0.961		Predicted Condition Types				FNR
		Healthy	State1	State2	Fail	
Actual	Healthy	94	4	2	0	0.06
	State1	2	98	0	0	0.02
	State2	4	2	94	0	0.06
	Fail	0	0	0	100	0.00
FPR		0.06	0.05	0.02	0.00	Acc=0.965

CONCLUSIONS

The framework is experimented on a scaled-down ocean generator powering a load case, showing that the proposed method can effectively identify the system's operational condition and produce efficient solutions to O&M management.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation under grant ECCS-1809164 and by an Early-Career Research Fellowship Award from the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine.

REFERENCES

[1] Reikard, Gordon. "Integrating wave energy into the power grid: Simulation and forecasting." Ocean Engineering 73 (2013): 168-178.

[2] Crowder, et al. "Artificial neural diagnostics and prognostics: Self-soothing in cognitive systems." Artificial Psychology. Springer, Cham, 2020. 87-98.

[3] Mnih, et al. "Human-level control through deep reinforcement learning." Nature 518.7540 (2015): 529.

[4] Rocchetta, R., et al. "A reinforcement learning framework for optimal operation and maintenance of power grids." Applied energy 241 (2019): 291-301.