Novel Multi-objective Optimisation with Deep Q-Reinforcement Learning (DQN) for Maintenance Activities of Floating Production Storage and Offloading Facilities (FPSOs)

Biju George*, Jonathan Loo**

(Ship Super UK, ** Queen Mary University of London, Mile End Road, London E1 4NS)*

{Corresponding author's email: bijuookken@hotmail.com}

Abstract - A novel work management framework for offshore floating systems has been proposed in this paper that comprises of Deep Q-Reinforcement Learning (DQN) problem formulation as a solution to multi-objective optimisation problem for maintenance activities of Floating Production Storage and Offloading Facilities (FPSOs). It has been demonstrated that DQN has the potential to be employed to develop a dynamic Work Management System that adjusts maintenance activities by achieving optimal path for carrying out activities that liquidates the risks to asset's performance depending on the changes achieved on the asset condition based on the activities completed. This potentially maximise resource utilisations, enable enhanced asset condition and lead to reduction of emissions from the asset and supplement the Regulatory oversight requirements.

Keywords: floating system; planned maintenance; optimisation; DQN; artificial intelligence; digitalisation.

1. INTRODUCTION

An investigation of the recent advancements in modelling and optimisation techniques to develop maintenance strategies for offshore floating systems have been carried out in [1], [2] and have demonstrated that the current literature does not incorporate site constraints of the asset related to offshore resource availability for the maintenance activity, which is a limitation of the existing state-of-the art maintenance frameworks. Also, there is lack of evidence to support that dynamic and autonomous resource allocations for maintenance activities take place in the offshore maintenance planning systems.

In summary, the following contributions are made in this paper:

A novel work management framework has been proposed that comprises of Deep Q-Reinforcement Learning (DQN) problem formulation as a solution to multi-objective optimisation problem for maintenance activities of Floating Production Storage and Offloading Facilities (FPSOs). The framework enables carrying out activities that have minimal site constraints, considering the design features, operating conditions, deteriorations, consequences of not doing the activities and time required to complete the activities, to get higher weighted sum of the completion times at short time as possible, whereby achieving higher resource utilisations. A greedy algorithm benchmarks the performances of DQN model and a hybrid model comprising of greedy and DQN parameters. This formulation enables achieving the optimal path for carrying out activities that liquidates the risks to the asset's performance, which would supplement the Regulatory oversight requirements of the FPSO.

2. MAINTENANCE SYSTEM FORMULATION

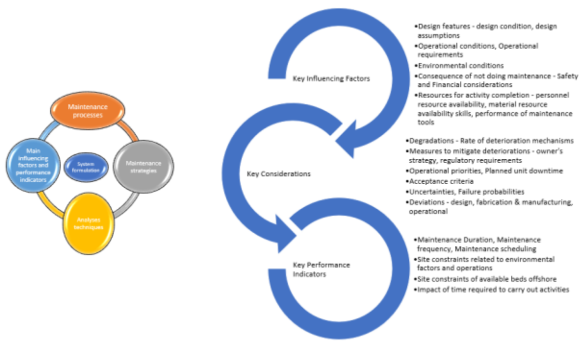


Figure 1. Maintenance system formulation

Figure 1 indicates an overview of maintenance system formulation. Maintenance processes, main influencing factors, performance indicators and analyses techniques formulate a series of maintenance strategies to achieve the desired goals, with feedback loop for continuous improvement of the maintenance program. The maintenance strategies develop tasks that could be considered to restore the desired functionality.

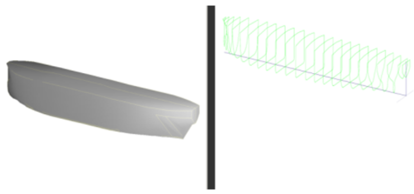
Deep Learning mathematical programming optimisation technique for aero-propulsion system has been employed by [3], whereas an unsupervised machine learning optimisation for offshore wind turbines was used by [4], a mathematical nondominated sorting genetic algorithm by [5] and a deterministic non-linear programming problem by [6].

Weighted sums approach for selective maintenance problem of multi-component systems was employed in the works of [7] and constrained optimisation mathematical programming technique for continuous and discontinuous operating systems was utilised in the works of [8]. Bayesian network with Monte Carlo simulation technique for marine pipelines has been employed by [9], and mixed integer non-linear programming based selective maintenance optimisation for engineering systems has been detailed by [10].

Sea current velocity predictions made by convolutional neural learning model employing residual learning with attention strategy has been proposed by [11], whereas marine heatwave prediction using long short-term memory recurrent neural network has been proposed by [12]. Deep learning optimisation model for wind turbine maintenance planning, combining variational mode decomposition, convolution neural network, long short-term memory network and full-connected network has been proposed by [13].

Opportunistic maintenance of offshore wind relying on market-based availability criteria rather than time/energy-based availability has been proposed by [14], whereas closed-loop maintenance strategy optimisation method has been proposed by [15].

3. MULTI-OBJECTIVE PROBLEM



Length between perpendiculars	360m
Moulded Breadth	55m
Moulded Depth	35m

Figure 2. Profile, dimensions of modelled FPSO [2]

A FPSO main deck that has been modelled in the works of [2] with the principal dimensions as indicated in Figure 2 has been employed in this work.

The decision variables considered were the design features, operating conditions, deteriorations experienced and the consequences of not doing the maintenance activities, detailed by [2]. The main objective was to maximize the maintenance personnel resource utilization and enable FPSO condition enhancement, considering the priorities with respect to design features, operating conditions, deteriorations, and the consequences of not doing the maintenance, taking into consideration the personnel resource time required for activity completion. The personnel resource utilisation directly relates to the key performance indicators of manpower costs, activity completion, cost related to activity duration and increase in efficiency, whereas the FPSO condition enhancement relates to the availability, reliability, safety and regulatory compliances of the asset.

4. DQN SOLUTION FORMULATION

The DQN problem statement has been defined as to carry out activities that have minimal site constraints, so as to get higher weighted sum of the completion times at short time as possible, which leads to higher resource utilization. The goal is to achieve the best trade-off between the turnaround time for the activities and liquidating the risks to the asset's performance, based on completion of activities in the FPSO work management system (WMS).

4.1 Implementation of problem formulation

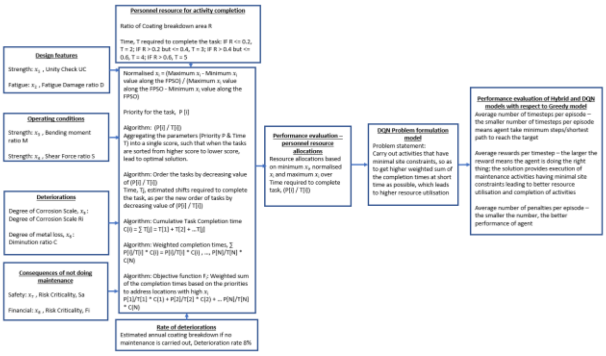


Figure 3: Multi-objective Optimisation with DQN

The Figure 3 provides an overview of formulation of multi-objective model with DQN, for FPSO main deck maintenance.

The DQN problem formulation model for FPSO main deck maintenance has been indicated in Figure 4 below.

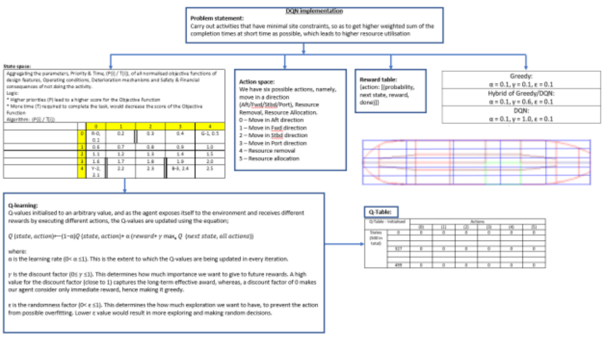


Figure 4: DQN Problem formulation model for FPSO main deck maintenance

The DQN Solution Formulation for Maintenance Activities involves:

4.1.1 State Space

FPSO Main Deck has been split into a 5 x 5 grid, which will give 25 possible locations on the Main Deck. For these grid locations the priority of the objective function over the time required to complete task (P[i] / T[i]) has been assigned from 0.1 with increments of 0.1 up to the maximum value of 2.5 that was found for the Safety & Financial Risks, from the Greedy Algorithm.

4.1.2 Action Space

The agent comes across one of the 500 states and takes an action. The Action is to move in a direction along the FPSO, or to decide to remove resource and allocate resource at a location. The agent has six possible actions, namely, move in a direction -Aft/Fwd/Stbd/Port-, Resource Removal and Resource Allocation.

4.1.3 Rewards

Points considered while deciding the rewards and penalties were that the agent should receive a high positive reward for a successful resource allocation, as this action was highly desired. By trial and error, a +20 points reward was assigned for a successful resource allocation. Agent should be penalized if it tries to allocate or allocate resources at wrong locations. By trial and error, a -10 points penalty was assigned for an illegal resource allocation or removal. Agent should receive a slight negative reward for every site constraint hit and for not moving anywhere, and for not making it to the assigned location for resource removal/ allocation after every time-step. By trial and error, a -1-point penalty was assigned for these actions.

5. PERFORMANCE EVALUATION

After enough random exploration of actions, the Q-values tend to converge serving our agent as an action-value function, which it could exploit to pick the most optimal action from a given state. The Hyperparameters for the DQN model includes, α, γ, ε, whereby, α is the learning rate (0< α ≤1). This is the extent to which the Q-values are being updated in every iteration. γ is the discount factor (0≤ γ ≤1). This determines how much importance we want to give to future rewards. A high value for the discount factor (close to 1) captures the long-term effective award, whereas a discount factor of 0 makes our agent consider only immediate reward, hence making it greedy. ε the randomness factor (0< ε ≤1) determines how much exploration we want to have, to prevent the action from possible overfitting. Lower ε value would result in more exploring and making random decisions. Considering the afore-mentioned points, the hyperparameters α, γ, ε have been varied between 0.1, 0.6 and 1, to generate the Greedy, Hybrid of Greedy/DQN and DQN models.

The agent for Greedy, Hybrid of Greedy/DQN and DQN models were evaluated on the following features:

Average number of timesteps per episode – the smaller number of timesteps per episode means agent take minimum steps/shortest path to reach the target.

Average rewards per timestep – the larger the reward means the agent is doing the right thing. In this work, as both timesteps and penalties are negatively rewarded, a higher average reward would mean that the agent reaches the target as fast as possible with the least penalties. i.e. the solution provides execution of maintenance activities having minimal site constraints leading to better resource utilisation, and completion of activities.

Average number of penalties per episode – the smaller the number ideally be zero or very close to zero, the better performance of agent.

The evaluation of Greedy, Hybrid of Greedy & DQN and DQN models for up to 25,000 training episodes have been carried out.

5.1 Evaluation of agent's performance

Figure 5: Learning curves of Greedy, Hybrid and DQN models

In the simulations in Figure 5, learning curves of Greedy, Hybrid and DQN models with respect to the number of timesteps taken to reach destination and the rewards per timestep, the variation of average timesteps per episode and the variation of average rewards per timestep, for the Hybrid and DQN models with respect to Greedy model have been illustrated.

It has been noted that overall, the Hybrid model with hyperparameters of $\alpha = 0.1$, $\gamma = 0.6$, $\varepsilon = 0.1$ and the DQN model with hyperparameters of $\alpha = 0.1$, $\gamma = 1.0$, $\varepsilon = 0.1$ achieve better results when compared with the Greedy model with hyperparameters of $\alpha = 0.1$, $\gamma = 0.1$, $\varepsilon = 0.1$, as the training episode increases, towards task completion time and liquidating the risks to the asset's performance.

6. Discussion and conclusion

A novel work management framework has been proposed that comprises of DQN problem formulation as a solution to multi-objective optimisation problem for maintenance activities of FPSOs. It has been demonstrated that DQN has the potential to be employed to develop a dynamic WMS that adjusts maintenance activities by achieving optimal path for carrying out activities that liquidates the risks to asset's performance depending on the changes achieved on the asset condition based on the activities completed. This potentially maximise resource utilisations, enable enhanced asset condition and lead to reduction of emissions from the asset and supplement the Regulatory oversight requirements.

References

[1] B. George, J. Loo, and W. Jie. "Novel multi-objective optimisation for maintenance activities of floating production storage and offloading facilities," *Applied Ocean Research* 130, 2023, 103440.

[2] B. George, J. Loo, and W. Jie. "Recent Advances and Future Trends on Maintenance Strategies and Optimisation Solution techniques for Offshore sector," *Ocean Engineering* 250, 2022, 110986.

[3] H. Hesabi, M. Nourelfath, and A. Hajji. "A deep learning predictive model for selective maintenance optimization," *Reliability Engineering and System Safety* 219, 2022, 108191.

[4] B. Yeter, Y. Garbatov, and C G. Soares. "Life extension classification of offshore wind assets using unsupervised machine learning," *Reliability Engineering and System Safety* 219, 2022, 108229.

[5] C. Zhang and T. Yang. "Optimal maintenance planning and resource allocation for wind farms based on non-dominated sorting genetic algorithm-II," *Renewable Energy* 164, 2021, pp. 1540-1549.

[6] B. Zhang and Z. Zhang. "A two-stage model for asynchronously scheduling offshore wind farm maintenance tasks and power productions," *Electrical Power and Energy Systems* 130, 2021, 107013.

[7] C. Diallo, A. Khatab, and U. Venkatadri. "Developing a bi-objective imperfect selective maintenance optimization model for multicomponent systems," *IFAC PapersOnLine* 52-13, 2019, pp. 1079-1084.

[8] G M. Galante, C M L. Fata, T. Lupo, and G. Passannanti. "Handling the epistemic uncertainty in the selective maintenance problem," *Computers & Industrial Engineering* 141, 2020, 106293.

[9] S. Adumene, F. Khan, S. Adedigba, and S. Zendehboudi. "Offshore system safety and reliability considering microbial influenced multiple failure modes and their interdependencies," *Reliability Engineering and System Safety* 215, 2021, 107862.

[10] T J. Ikonen, H. Mostafaei, Y. Ye, D E. Bernal, I E. Grossmann, and I. Harjunkoski. "Large-scale selective maintenance optimization using bathtub-shaped failure rates," *Computers and Chemical Engineering* 139, 2020, 106876.

[11] Z. Zhang and J. Yin. "Spatial-temporal Offshore Current Field Forecasting Using Residual-learning Based Purely CNN Methodology with Attention Mechanism," *Applied Artificial Intelligence* 38, 2024, 2323827.

[12] Q. He, Z. Zhu, D. Zhao, W. Song, and D. Huang. "An Interpretable Deep Learning Approach for Detecting Marine Heatwaves Patterns," *Applied Sciences* 14, 601, 2024, 14020601.

[13] X. Zhou, Y. Ke, J. Zhu, and W. Cui. "Sustainable Operations and Maintenance of Offshore Wind Farms Based on the Deep Wind Forecasting," *Sustainability* 16,333, 2024, 16010333.

[14] J. NcMorland, M. Collu, D. McMillan, J. Carroll, and A. Coraddu. "Opportunistic maintenance for offshore wind: A review and proposal of future framework," *Renewable and Sustainable Energy Reviews* 184, 2023, 113571.

[15] M. Li, X. Jiang, J. Carroll, and R. R. Negenborn. "A closed-loop maintenance strategy for offshore wind farms: Incorporating dynamic wind farm states and uncertainty-awareness in decision-making." *Renewable and Sustainable Energy Reviews* 184, 2023, 113535.