Deep Reinforcement Learning Scheduling Method for Prefabricated Bridge Construction

Yuan Yao, Vivian WY Tam*, Jun Wang, Khoa N Le and Anthony Butera

Western Sydney University, School of Engineering, Design and Built Environment, Locked Bag 1797, Penrith, NSW 2751, Australia
* Corresponding Author, Email: vivianwytam@gmail.com

Abstract

Off-site construction has become a widely accepted method due to its advantages in time-saving, rapid erection, and low cost. The rapid growth in this area demands better and more refined construction scheduling methods. Construction scheduling is complicated by the nature of various constraints in different aspects, such as resources and labour. Traditional methods, such as the Critical Path Method (CPM), lack consideration of various constraints, making them less applicable in real-world projects. This study proposes a Deep Reinforcement Learning (DRL) method to generate optimal construction schedules under limited labour and resource constraints. The objective of this study is to minimize the duration of construction projects. The proposed method introduces an improved DRL framework that enables the DRL to handle the scheduling of all construction processes. A case study is conducted on a real-world prefabricated bridge with 9 spans to evaluate the proposed method's performance. The DRL method is compared with traditional methods and the Genetic Algorithm (GA). The results show that DRL outperformed other methods in generating optimal construction schedules and required less running time. Therefore, the proposed method in this study extends the construction scheduling method and can be used in real projects.

Keywords: *Reinforcement Learning, Construction Scheduling, Off-site Construction, Scheduling Optimization*

1. INTRODUCTION

Automated construction scheduling is a technique that assists project managers in improving the management of labor, equipment, time, costs, and other project elements. It is a decision-making process, and it identifies which tasks need to be completed and outlines the method and timing for their execution. As Ding et al. (2023) pointed out, automated construction scheduling is a Resource Constrained Project Scheduling Problem (RCPSP). In a standard construction timeline, the total workload is broken down into various work breakdown structures and specific tasks, which are then allocated to the appropriate subcontractors (Bai et al. 2009). Construction planning encompasses scheduling and various other planning activities, such as material management, site layout design, equipment movement coordination, and overall site logistics. An experienced scheduler with extensive prior knowledge should meticulously manage all the information from the construction site to complete a construction schedule, though this process can be quite time-consuming (Amer, Koh & Golparvar-Fard 2021).

Various methods and theories have been developed for construction scheduling. Among the most commonly used is the Critical Path Method (CPM), which divides projects into distinct work breakdown structures and then manually rearranges them for better organization. Alamode and Plaza (1994) proposed case-based reasoning (CBR) techniques, which leverage previous cases to evaluate and resolve new ones. These approaches are dependent on past experiences and lack the adaptability required for handling diverse construction projects. Alternative approaches, such as heuristic techniques and Genetic Algorithms (GA), treat the scheduling challenge as an optimization problem, often with one or more objectives constrained by specific conditions (Ahmed, Hossain & Hossain 2021; Bettemir & Sonmez 2015; Erdal & Kanit 2021; Lin et al. 2022; Yuan et al. 2021). These approaches have limited ability to adapt to various types of construction projects and fail to account for uncertainties that arise during the construction phase. In real-world construction projects, organizing tasks is challenging due to multiple constraints, including labor, resources, and construction methods that affect the order in which activities can be carried out.

Reinforcement Learning (RL), a widely used approach for decision-making problems, has been employed to tackle scheduling optimization challenges involving intricate constraints (Ratajczak-Ropel 2018). In the reinforcement learning (RL) framework, the agent discovers the best action policy, like the ideal sequence of activities, by utilizing rewards from the environment through a mechanism involving delayed feedback (Sutton & Barto 2018). In construction scheduling, this reward could be tied to objectives such as project duration or overall cost. Recent work by Soman and Molina-Solana. (2022) recently explored the use of reinforcement learning (RL) techniques in construction look-ahead scheduling (LAS). Their study demonstrated that RL-based LAS could be produced significantly faster than conventional Critical Path Method (CPM) approaches while avoiding scheduling conflicts. Kadir et al. (2022) designed a network combining reinforcement learning and graph embedding to optimize construction project planning. However, the case studies presented in their research are fairly straightforward and do not fully reflect the complexity of real-world projects. Such methods prove inefficient when used in large-scale construction projects. These projects are typically categorized into a few broad groups, which may encompass various processes. However, this broad classification lacks the necessary clarity and detail needed by schedulers and contractors to properly track activities.

Building on prior research, this paper explores the application of reinforcement learning (RL) to address the challenges of construction scheduling under resource constraints. This research introduces a constrained scheduling approach based on Deep Q-Networks (DQN), designed to serve as a decision-support tool in construction scheduling. It generates optimized workflows that are free from constraints, enabling more efficient project planning. The proposed study seeks to automatically create an optimized schedule that maximizes resource utilization and manpower distribution without any constraints. It introduces an action masking algorithm to improve the DQN model's ability to handle complex construction project tasks.

2. METHOD

2.1 Problem Definition

The construction process of prefabricated bridge is broken to various structure groups, namely superstructure and substructure. Structure groups consist of various segments, which include piles, pile caps, piers, pier caps, girders, and joints. Each segment requires several activities to complete. Such three-level structure represents a simplified version of the construction process (Figure 1). Since schedulers aim to finish the project ahead of schedule, the goal of the proposed optimization is to reduce the overall construction project duration. (Equation 1). The optimization is constrained by resources constraints and precedence constraints (Equation 2 to Equation 4).

Objective:

$$minimize\ D_{project} \tag{1}$$

Subject to:

$$\sum R_t(a_{i,j,k}) \leq R \tag{2}$$

$$t_{k+1} - t_k \geq d_k \tag{3}$$

$$N_c = 0 \tag{4}$$

Where $D_{project}$ denotes the total duration of the project; $R_t(a)$ denotes the resources required for activity a; R represents the resources for each operation; t_k denotes the starting time of the k^{th} activity; d_k represents the duration of k^{th} activity; N_c denotes the number of constraint violations.

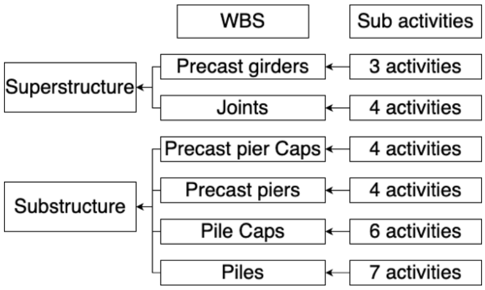

Figure 1: Construction process representation

2.2 RL development

Reinforcement learning is defined as a Markov Decision Process (MDP). In a Deep Q-Network (DQN), the agent selects actions based on a deep neural network that predicts the maximum reward value for each possible state in the environment. State and action representation is important to let the DQN agent "learns" better. In this study, state S is defined as the structure completion state (Equation 5). It indicates the overall progress of the project. The element $S_{m,n}$ denotes the number of n^{th} recipe under m^{th} structure group. It can also display the precedence constraints during the construction process. This state will guide the DRL policy to select actions that do not violate the constraints.

Action space $\hat{A}$ represents the possible construction activity combinations (Equation 6). Where m denotes the number of structure groups.

A reward is the response from the environment following an action. The way the reward function is defined shapes the optimization path for the DQN model. In this research, the reward function is composed of five components (Equation 7 to Equation 11).

$$S = \begin{bmatrix} S_{1,1} & \cdots & S_{1,n} \\ \vdots & \ddots & \vdots \\ S_{m,1} & \cdots & S_{m,n} \end{bmatrix} \tag{5}$$

$$\hat{A} = \begin{bmatrix} \overline{a_{1,1}} & \cdots & \overline{a_{1,m}} \\ \vdots & \ddots & \vdots \\ \overline{a_{2^m,1}} & \cdots & \overline{a_{2^m,m}} \end{bmatrix} \tag{6}$$

$$RW_1 = 10 \times N_{complete\ activity} + 50 \times N_{complete\ recipe} \tag{7}$$

$$RW_2 = 2 \times N_{ongoing\ activities} \tag{8}$$

$$RW_3 = -5 \tag{9}$$

$$RW_4 = 1000 \tag{10}$$

$$RW = RW_1 + RW_2 + RW_3 + RW_4 \tag{11}$$

The structure of the proposed DQN model is illustrated in Figure 2. The DQN includes one input layer, two hidden layers, and one output layer This architecture consists of an input layer, two hidden layers, and an output layer. Due to the irregular shape of the environment state S, a flattening layer is used to convert the two-dimensional state S into a one-dimensional vector S_f. The dimension of input corresponds to the length of S_f and the dimension of output is the number of actions. The dimension of first and second hidden layers are 128 and 64respectively. The agent was trained over 200 episodes with a minibatch size of 64, using a 4-step frequency for updating the network and a replay buffer size of 10,000. Training followed an ε-greedy policy, starting with an exploration rate of 0.95 and gradually decreasing to 0.01 by the final episode. The DQN model is trained on a personal computer with i7-11700 and RTX3070. The interpreter is based on Python 3.9.0, Pytorch 1.12.1, and Numpy 1.23.4.

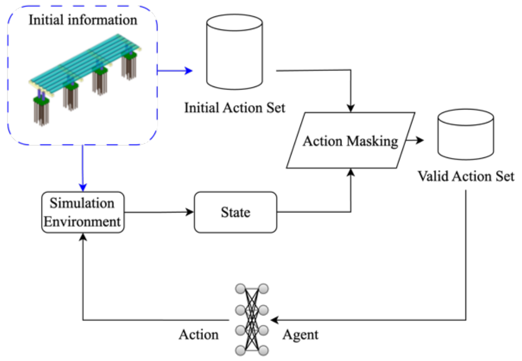

Figure 2: DRL architecture

3. CASE STUDY

The selected prefabricated bridge for case study is a 9-span simple supported bridge. The tasks will be distributed among several crews utilizing various equipment types. For example, a lifting team will be responsible for tasks such as pile lifting, pier lifting, pier cap lifting, and girder lifting, using a walking crane. In this research, the construction process is divided into 28 specific activities, involving 12 separate crews and 4 different kinds of equipment. The classification of segments considers the different bridge structures and construction techniques. The segments include: 1) pile; 2) pile cap; 3) piers; 4) piers cap; 5) girders; and 6) seams. The 9-span bridge owns 10 pile recipes, 10 pile cap recipes, 10 piers recipes, 10 piers caps recipes, 9 girder recipes, and 9 seams recipes. For each segment, there are different activities. The total quantities of sub activities are 273. The quantities of Lifting crews and Hammer & Piling crews are 4; the quantities of other crews are 2. The quantities of Mobile Crane and Hammer are 4; and the quantities of other equipment are 2.

4. RESULTS

The scheduling optimization results of proposed DQN are compared with GA. The GA algorithm adopted in this paper is proposed by Liu et al.(2020). The population size, crossover rate, mutation rate, and top rate is set to 500, 0.8, 0.2, and 0.15 accordingly. The manual schedule is translated from the original master plan which requires the project be completed with 45 days. Considering the normal work hour is 8 h/day, the total construction duration is translated to 360 h. The results are listed in Table 1. The proposed method can generate Gantt Chart for schedule evaluation (Figure 3). The chart shows no constraint violations.

Table 1: Scheduling results

Method	Manual	GA	DQN
Duration (h)	360	328	275

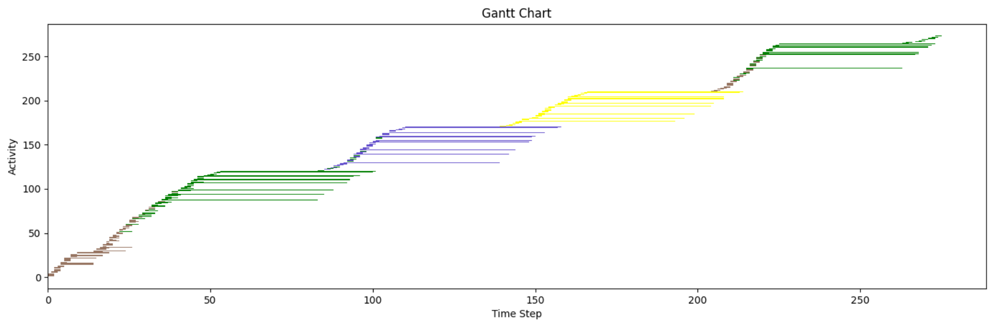

Figure 3: Gantt Chart for the 9-span bridge

5. DISCUSSION AND LIMITATION

The proposed DQN method is compared with GA and manual method. The results indicate that the performance of DQN method is better than other two methods. GA's sub-optimal performance results from its convergence to a local optimum. Its performance is highly related with the design of initial population and the algorithm configurations. However, DQN is not based on chromosome iteration to optimize construction schedules, thus the population cannot affect its performance. Action masking prevents the DRL agent from selecting invalid actions, while reward shaping directs the agent toward optimal solutions. As a result, the DRL system efficiently utilizes the environment's constraint-checking features, enabling quicker convergence and finding the global optimum.

This study exists some limitations. First, the proposed method only considers shortening the duration as the only objective. All the algorithm features are designed for single objective purpose. Secondly, the proposed action masking still has a lot of room for improvement. The current approach relies on the DQN environment to make its own judgments and give legal action options based on the current state and constraints. Third, proposed DQN architecture is very simple, which results in the vibration of the reward curve. Improved DQN methods should be introduced to solve construction scheduling problems.

6. CONCLUSION

Proposed construction scheduling method uses DRL to generate optimized construction schedules under labour and equipment constraint. The objective of the study aims to shortening the total construction duration. To reduce the impact of the scale of construction project, this study proposes a three-level representation that divides a construction to structure groups, segments, and activities. Action masking is adopted for efficient DQN training, which reduces the dimension of action space. Compared with GA and manual method, proposed DQN method generates optimal schedules with no constraint violation.

The future direction should focus on proposing a general and flexible automated scheduling method with high generalization abilities. Additionally, it needs to have a certain degree of flexibility, meaning it can quickly make appropriate reschedules in situations of uncertainty. To have these abilities, a more comprehensive construction model should be developed by transferring human knowledge. Building Information Model (BIM) could be integrated with proposed DQN method, because BIM could directly be utilized as a simulation environment and provide various constraints. In the future, it will be important to establish solid benchmarks for construction projects. Additionally, there is a need to create construction benchmarks with optimized timelines for different scenarios, which could support the advancement of AI-driven scheduling methods.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support from the Australian Research Council (ARC), Australian Government (No: DP200100057, IH200100010; FT220100017).

REFERENCES

Aamodt, A & Plaza, E 1994, 'Case-based reasoning: Foundational issues, methodological variations, and system approaches', *AI Communications*, vol. 7, pp. 39-59.

Ahmed, T, Hossain, SM & Hossain, MA 2021, 'Reducing completion time and optimizing resource use of resource-constrained construction operation by means of simulation modeling', *International Journal of Construction Management*, vol. 21, no. 4, pp. 404-15.

Amer, F, Koh, HY & Golparvar-Fard, M 2021, 'Automated methods and systems for construction planning and scheduling: Critical review of three decades of research', *Journal of Construction Engineering and Management*, vol. 147, no. 7.

Bai, Y, Zhao, Y, Chen, Y & Chen, L 2009, 'Designing domain work breakdown structure (DWBS) using neural networks', vol. 5553, pp. 1146-53.

Bettemir, OH & Sonmez, R 2015, 'Hybrid genetic algorithm with simulated annealing for resource-constrained project scheduling', *Journal of Management in Engineering*, vol. 31, no. 5.

Ding, HY, Zhuang, CB & Liu, JH 2023, 'Extensions of the resource-constrained project scheduling problem', *Automation in Construction*, vol. 153.

Erdal, M & Kanit, R 2021, 'Scheduling of construction projects under resource-constrained conditions with a specifically developed software using genetic algorithms', *Tehnicki Vjesnik-Technical Gazette*, vol. 28, no. 4, pp. 1362-70.

Kedir, NS, Somi, S, Fayek, AR & Nguyen, PHD 2022, 'Hybridization of reinforcement learning and agent-based modeling to optimize construction planning and scheduling', *Automation in Construction*, vol. 142.

Lin, JCW, Lv, Q, Yu, DH, Srivastava, G & Chen, CH 2022, 'Optimized scheduling of resource-constraints in projects for smart construction', *Information Processing & Management*, vol. 59, no. 5.

Liu, J, Liu, Y, Shi, Y & Li, J 2020, 'Solving resource-constrained project scheduling problem via genetic algorithm', *Journal of Computing in Civil Engineering*, vol. 34, no. 2, p. 04019055.

Ratajczak-Ropel, E 2018, 'Experimental evaluation of agent-based approaches to solving multi-mode resource-constrained project scheduling problem', *Cybernetics and Systems*, vol. 49, no. 5-6, pp. 296-316.

Soman, RK & Molina-Solana, M 2022, 'Automating look-ahead schedule generation for construction using linked-data based constraint checking and reinforcement learning', *Automation in Construction*, vol. 134, p. 16.

Sutton, RS & Barto, AG 2018, *Reinforcement learning: An introduction*, MIT press, Cambridge, Massachusetts.

Yuan, YS, Ye, SD, Lin, L & Gen, MS 2021, 'Multi-objective multi-mode resource-constrained project scheduling with fuzzy activity durations in prefabricated building construction', *Computers & Industrial Engineering*, vol. 158.