Reinforcement Learning-Enhanced Path Planning for Mobile Cranes in Dynamic Construction Environments: A Virtual Reality Simulation Approach

Rafik Lemouchi[1], Mohamed Assaf[1], Ahmed Bouferguene[2] and Mohamed Al-hussein[1]

[1] Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
[2] Campus Saint-Jean, University of Alberta, Edmonton, Canada
lemouchi@ualberta.ca, massaf2@ualberta.ca,ahmedb@ualberta.ca, malhussein@ualberta.ca

Abstract

This paper introduces a novel approach to crane path planning on construction sites through the utilization of Reinforcement Learning (RL) and Virtual Reality (VR) simulations. The strategy includes a comprehensive simulation model that incorporates an agent, actions, states, environment, and a reward system. After undergoing extensive training across millions of episodes, the crane agent has acquired optimal path-planning techniques that enhance lifting time, manage energy consumption, and improve collision detection. The results highlight the agent's impressive growth from initial exploration to peak efficiency, represented by cumulative rewards and evolving simulation times. The findings also demonstrate the effectiveness of RL-based path planning in maneuvering dynamic construction environments and optimizing crane operations.

Keywords –
Path Planning, Simulation, Reinforcement Learning, Virtual Reality.

1 Introduction

The construction industry constantly evolves, aiming to maximize efficiency and minimize costs. Over the past few years, off-site construction has gained traction due to its ability to save time and money. This approach involves transporting prefabricated modules to the construction site for installation, making cranes an essential component.

Despite numerous attempts to improve crane operations, outdated tools and planning methodologies still need to be used. According to [1,2], current practices involve lift engineers generating CAD-based 2D and 3D simulations of various lift scenarios in a static and time-consuming manner. The planning process is often trial and error, with better alternatives not being discovered until later stages.

To address the path planning problem, many studies have explored using automated planning tools and information technology to enhance current practices. The initial works focused on using deterministic algorithms for path planning, as [3] outlined. To automate the path planning task, they employed two heuristic search methods, hill climbing and A*. Hill climbing involves an iterative approach of adjusting a solution to minimize the distance or any desired cost function. In comparison, A* uses a nodes approach, where it tries to find the shortest path between the start and end nodes using a cost function similar to hill climbing. However, these methods proved time-consuming and often stuck in local optima instead of finding the optimal solution.

Several studies have explored using metaheuristic algorithms to improve crane path planning. [4] used an ant colony to achieve collision-free path planning for mobile cranes, while [5] employed genetic algorithms to plan lifts in complex environments. However, metaheuristic algorithms are only sometimes the optimal solution and can be heavily influenced by initial conditions. At the same time, other works, such as [6], attempted to use hybrid approaches to tackle the issue, with similar results as previous works.

Recently, many researchers have turned to 3D/4D simulations and Building Information Modelling (BIM) to simulate and generate feasible solutions for lift planning. [1] integrated 4D crane simulation and BIM to manage operations on a construction site, while [7] presented a data-driven crane management system for industrial projects. Although BIM-based simulations offer detailed visualization, they may struggle to handle dynamic scenarios, making them less adaptable. [8] developed a methodology that enables automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments to address this issue. They used a GPU-based parallelization approach for discrete and continuous collision detection. However, the methodology was built using a genetic algorithm, which may generate premature solutions.

According to [9], Reinforcement Learning (RL) is a type of machine learning that relies on learning through experience without needing previous data. The learner is tasked with discovering which actions yield the highest rewards by experimenting with different approaches, as explained by [10]. RL has been applied successfully to challenging tasks, such as game-solving. AlphaGo was developed to solve the game of Go by Silver et al.2017 [11], notably self-driving and robotics. As for the implementation of RL in construction projects, there have been some works using it to enhance the overall planning process. [12]explored RL applications during the design phase for decision-making purposes. As for crane operations, [9] used a hybrid greedy and RL approach for crane mat layout optimization. The same techniques used to tackle these complex issues can also be used to address the crane path planning issue.

This study identifies several areas of improvement where further work is needed, including the following:

1. Better tools must be provided for crane operators to develop practical and achievable lifting paths where less redundant and efficient lifts are desired.
2. Current path planning methodologies often fail to account for the unique complexities of mobile crane operations and their planning procedures. These complexities relate to the congested nature of construction sites and the requirements needed to perform a successful lift effectively.
3. A fully automated path planning process that explores all possible lifting approaches while considering the changing nature of construction sites needs to be developed.
4. There is a need for better-optimized procedures that can learn from the built environment and adjust to changes that occur during construction. Therefore, a framework with an adaptive learning approach is required.

This methodology combines VR simulations with RL to address previously mentioned research gaps. This work aims to address a need for developing a more comprehensive understanding of optimal solutions in complex and dynamic environments. By integrating 3D environmental elements, the methodology aims to identify precise solutions that consider aspects such as time, complex movement, and realistic scenarios that previous methods have overlooked. Furthermore, using RL, the methodology seeks to enhance exploration within construction sites, particularly in the automated path planning of mobile cranes, which can be complex and challenging to adapt to dynamic site conditions. This integration of VR simulations and RL allows a thorough evaluation of diverse alternatives, ultimately selecting the most optimal solutions. The methodology fills the gap left by the absence of fully automated path-planning methods tailored specifically for mobile cranes and dynamic construction settings. Ultimately, this methodology aims to enhance the understanding of adapting to changes in the built environment during the construction phase, leading to better solutions for complex and dynamic building projects.

2 Methodology

In this section, a brief description of the methodology followed in developing the RL-based path planning framework. Figure 1 displays the main steps followed; where after identifying the lack of a realistic solution based on a 3D environment is needed to optimize the lift planning resources, data was collected to develop a simulation model that would assess various lift alternatives, and based on the main KPI the best solution is suggested for the user.

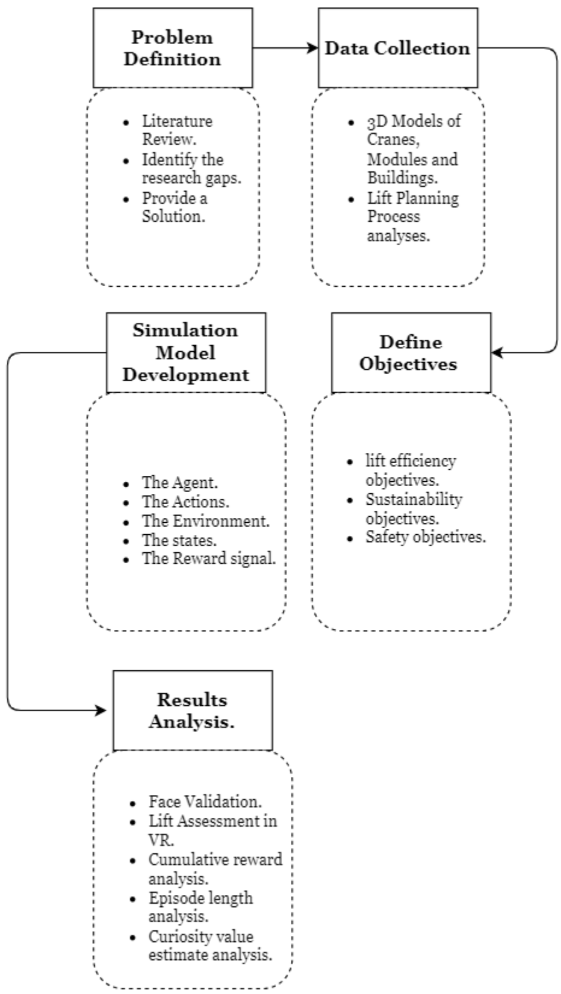


Figure 1. The overall research methodology.

2.1. Data Collection

In order to realistically replicate the lifting planning process, the authors determined that three main types of data are necessary: Crane data, Lifting Module data, and Building data.

2.1.1. Crane Model

The crane model used in the simulation is a highly accurate replica of a crawler crane widely used in industry. The model boasts a boom length of 36.5 78meters, for which the 36 meters configuration was selected, a track length of 10.0 meters and a width of 7.85 meters, and a maximum capacity of 300 tons, making it a highly versatile and reliable piece of machinery for a wide range of construction projects. The 3D rendition of the crane model is presented in Figure 2, showcasing the equipment's intricate details and precise specifications.

Figure 2. The crawler crane 3D model.

2.1.2. Modules and Buildings

Regarding the modules, a diverse array of payloads was brought into a 3D format, each possessing unique dimensions, weight, and physical attributes. Figure 3 contains a sample of a payload. A commonly used type of construction trailer is used, with dimensions of 6.01×2.34×2.69 cubic meters and a total weight of 2850kg. The trailer is transported to provide a suitable workshop area for welders and is needed in the vicinity of the construction site.

A) Construction trailer 3D model.

B) 3D representation of the lifting task.

Figure 3. The 3D model of payload and construction site.

As for the buildings were imported via a BIM format, with the models' measurements and characteristics preserved, albeit with simplified component properties and reduced detail to optimize simulation performance. A traditional construction project is selected to display the efficiency of the developed tool. The crane agent is assigned a task of lifting a payload from its pick location, highlighted in red, and delivering said payload to its set location, highlighted in green, effectively performing a single lift.

2.2. Objectives Definition

The main focus of this framework is to provide a lift plan that enhances the following components:

2.2.1. Lifting Time

Crane operators are responsible for ensuring the safe and efficient transport of payloads. To achieve this, they rely on lifting times as a crucial factor. The longer the payload spends in the air, the higher the risk of accidents, which is why keeping the lifting time to a minimum is essential. Additionally, the operator must follow the shortest and safest paths during transport to reduce the risk of damage to the payload or any surrounding structures. Maintaining the payload's elevation level throughout the lift is also preferred, as sudden changes in height can cause instability and increase the risk of accidents. According to a recent study by [13], heavy lifts are often planned to remain low until they're close to their destination, ensuring maximum safety and efficiency during transportation.

2.2.2. Energy consumption

There exist three distinct alternatives to move a load in a more sustainable manner. The first one is walking, which involves transporting the object from point A to point B while the mobile crane moves. This type of movement is deemed suboptimal, primarily because it significantly increases planning costs. The need for crane mats for crane movement is a major contributor to the high costs. To mitigate mat costs, one of the most common measures is to plan the lifts in a project such that the crane does not need to walk while performing the lifting procedure. This is according to [13].

The duration of the lifting process is critical in ensuring a safe and efficient transport of the payload when operating a crane. The primary objective of the crane operator is to move the load from one location to another using the shortest and safest route possible while minimizing lifting time. To further enhance safety measures, it is recommended to keep the payload at a low elevation until it reaches its destination. A recent study by [13] revealed that heavy lifts are often planned to maintain a low elevation during transport.

The equipment's Hook Movement feature is equipped with a primary hook that offers a singular degree of freedom in the vertical direction. This exceptional feature allows the hook to effortlessly lift or lower attached objects, making it an incredibly versatile and efficient tool suitable for a wide range of applications.

2.2.3. Collision Detection

Typically, construction sites implement numerous safety measures to prevent collisions. However, for this particular project, it was determined that five types of collisions would be of particular concern. Two of these occur before lifting the load, including collisions between the crane and humans and between the crane and the building. An additional three types of collisions are considered once the load has been lifted, including collisions between the payload and the crane, collisions between the payload and humans, and collisions between the payload and the building.

3 Simulation Model Development

In this section a brief description of the methodology used to develop the simulation model is presented.it consists of five main components which are the agent, the actions, the states, the environment, and the rewards. The interaction between the different components is represented in Figure 6.

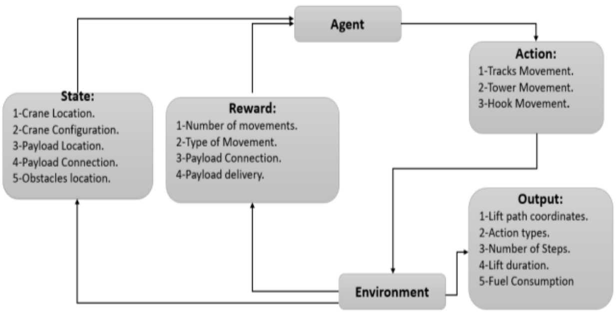


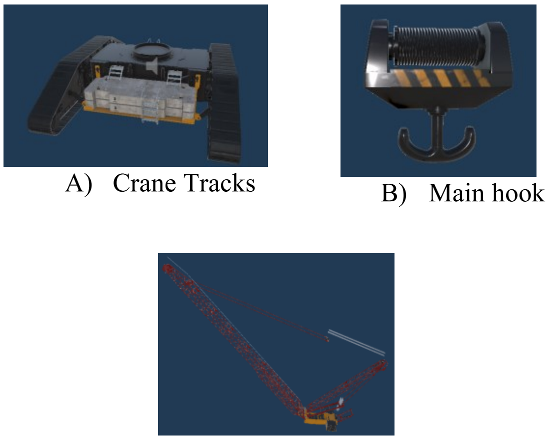
Figure 4. The reinforcement learning model.

3.1. The Agent

Initially, a crane agent is needed to take different actions to transport the payload from its original loading point to a predetermined set point. To achieve this task, a proper locomotion system is needed.

A mobile crane's locomotion system is more complex than that of a tower crane. Where the crane is made up of many parts that the crane operator needs to coordinate in order to perform a lift using the crane's full capacity. For the purposes of this work, three components' movements were modeled and are used in the training process, which is seen in Figure 5.

The main boom, where the boom is capable of rotation around the central axis of the crane. Both movements were taken into consideration, and the impact the angle has on the loading capacity was considered as well.



A) Crane Tracks

B) Main hook

C) Crane Boom

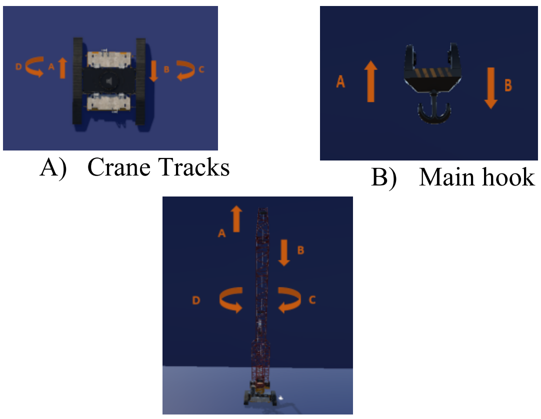
Figure 5. The main Crane Components

The second comes the crane's tracks, which enable the crane to move forwards, backwards and rotate around the it's center.

Finally, the crane's hook's movement was considered. This component has a relatively simple movement, where it moves either up or down depending on the lift phase.

3.2. The Actions

Since the agent is comprised of three main components, each component was given a degree of freedom (DOF).



A) Crane Tracks

B) Main hook

C) Crane Boom

Figure 6. The degrees of freedom of the main crane components.

The tracks are given two DOFs, one rotation around the Y axis, and one translation along the x-axis, as seen in

Figure 6A. For the agent; this constitutes four different actions that can be taken.

The main hook was given one degree of freedom along the Y axis as seen in figure 6B, which allows the agent to take two different additional actions.

Finally, the main boom was given two DOFs, two rotations, one rotation around the Y axis and one around the Z axis, as seen in Figure 6C. This constitutes four additional actions.

3.3 The environment

Once, the crane agent is instantiated in the 3D environment it starts interacting with the different environment components to explore it. In order to improve the agent's interaction, a set of sensors are attached to it. The sensors use collisions in with the different 3D models and collects data about the geometrical properties of 3D component found in the simulation environment. In the following training episodes, the agent uses the data collected through sensors to avoid colliding with the different components of the environment. for instance, when a sensor collides with a building, it can identify physical dimensions, velocity, and its tag. Where there are four main tags Human, Obstacle, and Objective, Crane.

3.4. The states

The simulation is run for a predefined number of iterations/episodes. In each episode, the location of each major element is stored alongside the element's velocity local rotation.

In the initial runs of the simulation, and prior to adding a penalty to the collision of the crane model with the payload, the agent could move the payload by pushing it towards the set point to gain the final reward. To penalize the suboptimal behaviour, an additional variable was added for the payload, which is a Boolean variable. The Boolean value represents whether the object is being lifted; once the object is connected to the hook, the Boolean value is turned to true.

Finally, the environment is reset to its initial configuration in three cases:

1. The episode will end if the maximum number of steps has been attained and the payload is yet to be delivered.
2. If the agent leaves the environment, the episode will end.
3. If the payload is delivered to the set location, the final reward is given, and the episode is terminated.

3.5. Reward signal selection

While the agent is training, it is essential to select the appropriate reward; this task is perhaps the most critical. Moreover, selecting a reward that balances penalties and rewards is necessary. For instance, if the agent is being over-penalized, the behavior resulting from the training would be suboptimal. For this work, various rewarding strategies were explored, and the final selected strategy was selected based on its merits and final training results.

The reward pseudo code.

```
Input:
  Crane initial location (Cl), Payload initial location
(PL), Crane capacity (Cc), Payload weight (Pw)
  for steps = 1 to Number of iterations
    for steps = 1 to max steps
      if PL is disconnected
        Compute D
        if D > 0.5m
          Act from possible actions
          Apply Move penalty × D.
          check for collision.
          if a collision is true
            Apply collision penalty.
        else if D < 0.5m
          Apply connection.
          Add Lift reward.
      if PL is connected
        Compute Dps.
        If Dps > 0.5m:
          Act from possible actions
          Apply Move penalty × Dps
          check for collision.
          If a collision is true:
            Apply collision penalty.
        Else if Dps < 0.5m:
          Set payload.
          Add Final reward.
    Reset environment.
```

Where:
D: distance between from crane to the object.
DPS: Distance between payload and set point.

In order to attain the previously determined objectives, which are the optimization of the lifting time, energy consumption, and collision detection, three different types of rewards were built into the training process. The first reward is related to the lifting time, which is highly affected by the number of actions taken and the overall time needed to deliver the payload to its set location. Furthermore, each action has its parameters, such as movement speed, damping ratio, and interaction with the previous movement. For instance, when the crane is moving, no other action can be taken until the crane is at a complete stop. All the previously mentioned parameters are combined to calculate the time needed to perform the lift, and the subsequent reward signal associated with the movement is developed to penalize the agent for each time spent lifting and transporting the object. This penalty ensures that the resulting lift uses the shortest lifting time. The penalty amount was selected after multiple iterations, where the initial penalties were severe enough to disable the agent from moving.

The second reward signal used is related to collisions or leaving the training area. The objective of the training is to discover all of the possible alternatives that can be used to perform a lift. It was decided that the agent would leave to explore the whole construction site. However, in some instances, the agent would leave the training area; thus, in those instances, a significant penalty is used to prevent the agent from leaving the said area. Furthermore, the episode is instantly terminated since, based on the author's experience with the model, the agent does not find its way back to the payload once it has explored the extremities of the training area.

Additionally, using the same approach, the agent sometimes collides with the surrounding obstacles and buildings despite the built-in sensors. An additional penalty is added based on the collision tag in those instances. The episode is terminated in some instances where the agent either collides with a human or a building. The other collisions are only penalized to allow the agent to train for an extended period.

The third type of reward is related to the type of action taken, where specific actions are preferred. For instance, since pick-up and walk operations are less favorable due to their increased cost, crane movements are more penalized than boom and hook movements. Next, main boom movements are expected to consume more energy than hook movements. Thus, they penalized more than hook movements. Finally, hook movements are given the lowest penalty.

Finally, the main reward signals for the crane are those related to lifting and delivering the payload. The first portion of the reward is connecting the payload to the hook; once the agent lifts the payload, it receives an enormous reward. The second portion relates to the agent setting the payload in the set place where the final reward is given and the episode is terminated. If the agent lifts the payload and fails to deliver it, the episode is terminated, and the environment is reset.

4 Results

The following section sheds light on the initial findings of our research and emphasizes the efficacy of our methodology. To quantify the impact of our methodology, three main indicators are used: cumulative rewards, episode length, and curiosity value indicators. The cumulative rewards graph displays the agent's improved training, which starts with negative outcomes (no solutions) and improves in value until it reaches a 298-point solution. To understand the cumulative rewards graph efficiently, the episode length indicator is used, which displays the number of steps needed by the agent to achieve the task; a lower number of steps indicates a more optimized solution. In order for the agent to decrease the number of episodes to achieve the lifting task, it must have a significant focus on exploring the actions, space, and the environment; this exploration task is achieved through curiosity, where the agent's focus on exploration in later episodes coincides with the decreased episode length and as a result of the latter higher cumulative rewards. To get a better understanding, you can refer to Figure 7, which displays the cumulative rewards achieved per episode and reveals some intriguing insights. In the first 500,000 episodes, the agent encountered a formidable challenge due to the task's complexity, leading to a high number of exploratory and negative rewards resulting in penalties for every action. However, the agent discovered a promising solution around the 550,000th episode, which significantly improved the outcomes. Nonetheless, further refinements were necessary to optimize the training process. In the subsequent episodes, the agent consistently improved its approach, gradually climbing towards the optimal policy. Eventually, it peaked at 298 total rewards before stabilizing at approximately 4 million episodes.

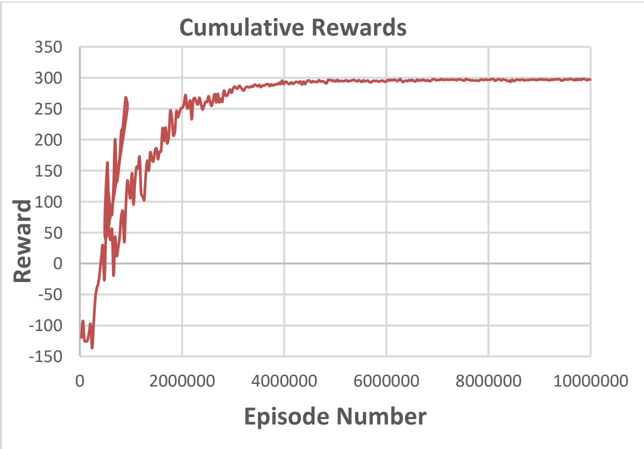


Figure 7. The cumulative reward value per episode.

The study conducted by the researchers involved a comprehensive examination of various factors that influence the efficiency of an agent in executing tasks proficiently. Amongst these factors, the duration of each episode was given special consideration, as it is a crucial parameter used to evaluate an agent's overall performance. The researchers analyzed the simulation time of the agent during its initial million episodes and found that it ranged from 1900 to 800 seconds, which is significantly longer than the desired duration. This indicates that the agent was in the exploratory phase, trying to identify the best possible path to achieve the assigned task. The researchers observed a significant decrease in the simulation time as the simulation progressed, which continued until the 3.5 millionth episode. This trend indicates that the agent had identified an optimal path and slightly improved the discovered solution. This discovery led the agent to find the most efficient path, which could be completed within 17.8 seconds in the simulation environment. The agent's continuous improvement and progress ultimately revealed the most efficient and effective way to complete the task. The episode length results are highlighted in Figure 8.

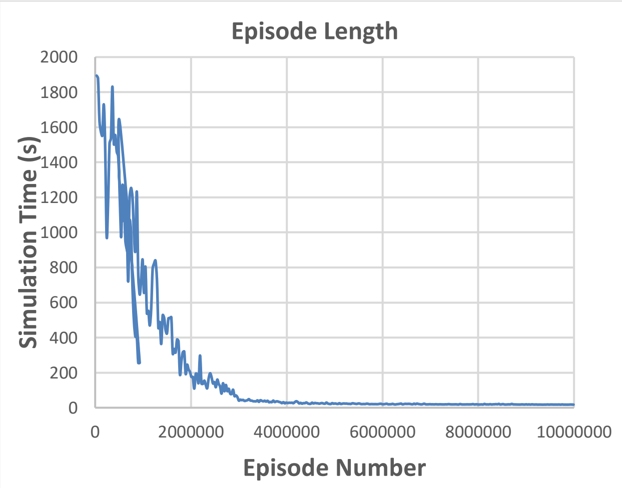


Figure 8. The episode length is in seconds per episode.

To assess an agent's learning progress, it's vital to analyze its level of curiosity and how it interacts with the environment. Curiosity value estimates provide valuable insights into the agent's learning patterns, reflecting its interest in exploring information. Higher values indicate strong engagement and a deep desire to learn, while lower values suggest a lack of enthusiasm for exploration and learning. By measuring curiosity-driven behavior, we can gain valuable insights into the agent's interest levels and learning patterns, which help us understand its progress. Figure 9 presents a graphical representation of the curiosity value estimates for behavior analysis. The results indicate a clear pattern of curiosity value progression that aligns with expected trends. At the beginning of the learning process, the agent exhibits minimal curiosity, but its curiosity level gradually increases to larger values in later episodes.

This indicates a notable surge in engagement or interest as the agent continues to learn. Statistical mean and standard deviation measures support this progression, demonstrating a distinct shift from initial disinterest to intense curiosity. Evaluating the agent's curiosity-driven behavior is crucial for understanding its learning patterns and progress. This leads to more effective learning by ensuring the agent actively engages with the environment and explores new information. After the lift was developed, the authors assessed its validation using two approaches, which are face validation and VR lift assessment.

In the first approach, the authors used the computer model in the virtual environment and analyzed the different steps used to lift and deliver the payload. However, the computer version, although a 3D approach was deemed insufficient due to the presence of certain blind spots that the users could miss, to that extent a VR, the authors explored the lift using a gamified approach, where they moved along the environment during the lift's execution and assessed the validity of the lift in the virtual environment.

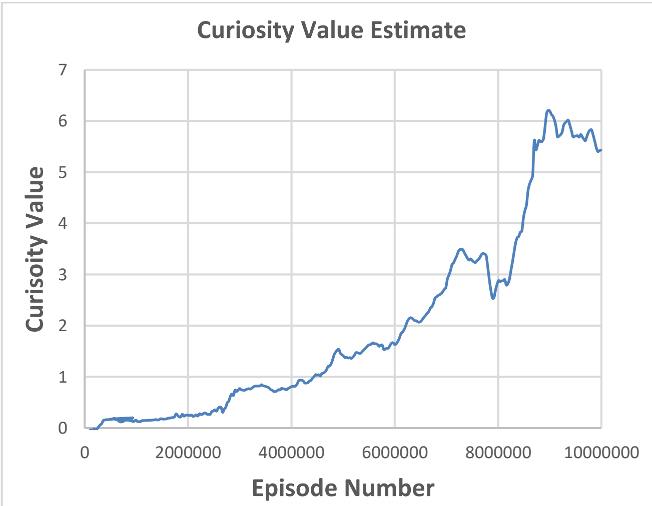


Figure 9. The curiosity value estimate per episode.

5 Conclusion and Future Directions

This work was used to develop a new tool for the construction industry that helps with crane path planning. The tool is designed to create lifting paths that are more comprehensive and take into account the changes that occur on a construction site. This leads to a smoother lifting process, which reduces planning time and produces more accurate results. The tool helps to improve collision detection, decrease lift complexity, and lower energy consumption. In addition, we have incorporated RL and VR simulations to make the planning process more resilient to unforeseen changes on the construction site. Using a virtual environment enhances the planning process, allowing for updates to be made to the crane agent training and lifting capabilities based on changes in the construction site, resulting in more reliable and efficient lifting schemes.

The authors developed a comprehensive simulation model to achieve our goal. This model included an agent, actions, states, environment, and a reward system. By training the crane agent through numerous episodes, the agent could detect optimal path-planning strategies. The agent's learning journey was tracked through cumulative rewards, which progressed from exploratory phases to achieving peak efficiency after approximately 4 million episodes. Analysis of simulation times highlighted the agent's evolving learning patterns, ultimately identifying the most efficient path within the environment. Face validation and VR-based assessment were performed to validate the developed approach's results. The results demonstrated the agent's ability to navigate complex environments, optimize lifting processes, and minimize simulation times. The insights gained through the agent's learning patterns and performance metrics validated the effectiveness of RL-based path planning in dynamic construction scenarios.

Future works will involve conducting a comparative analysis with real-life path planning scenarios by testing the developed model in live construction settings. By aligning simulated results with real-world scenarios, we aim to refine and validate our methodology for seamless integration into actual crane operations. Ultimately, our approach will contribute to safer, more efficient, and adaptable construction practices.

References

[1] A.N. Tak, H. Taghaddos, A. Mousaei, A. Bolourani, U. Hermann, BIM-based 4D mobile crane simulation and onsite operation management, Autom Constr 128 (2021) 103766.

[2] S. Hu, Y. Fang, Y. Bai, Automation and optimization in crane lift planning: A critical review, Advanced Engineering Informatics 49 (2021) 101346.

[3] P.L. Sivakumar, K. Varghese, N.R. Babu, Automated Path Planning of Cooperative Crane Lifts Using Heuristic Search, Journal of Computing in Civil Engineering 17 (2003) 197–207.

[4] X. Wang, Y.Y. Zhang, D. Wu, S. De Gao, Collision-Free Path Planning for Mobile Cranes Based on Ant Colony Algorithm, in: Materials, Mechatronics and Automation, Trans Tech Publications Ltd, 2011: pp. 1108–1115.

[5] P. Cai, I. Chandrasekaran, J. Zheng, Y. Cai, Automatic Path Planning for Dual-Crane Lifting in Complex Environments Using a Prioritized Multiobjective PGA, IEEE Trans Industr Inform ` (2018) 829–845.

[6] K. Boutouhami, A. Bouferguene, R. Lemouchi, M. Assaf, M. AL-Hussein, J. Kosa, Hybrid Approaches For Handling Mobile Crane Location Problems In Construction Sites, in: 2023 Winter Simulation Conference (WSC), 2023: pp. 2722–2733.

[7] S. Han, Z. Lei, U. (Rick) Hermann, A. Bouferguene, M. Al-Hussein, 4D-based automation of heavy lift planning in industrial construction projects, Canadian Journal of Civil Engineering 48 (2021) 1115–1129.

[8] S. Dutta, Y. Cai, L. Huang, J. Zheng, Automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments, Autom Constr 110 (2020) 102998.

[9] G.M. Ali, A. Bouferguene, M. Al-Hussein, Crane Mat Layout Optimization Based on Agent-Based Greedy and Reinforcement-Learning Approach, J Constr Eng Manag 149 (2023) 4023067.

[10] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, A Bradford Book, Cambridge, MA, USA, 2018.

[11] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge, Nature 550 (2017) 354–359.

[12] S. BuHamdan, A. Alwisy, A. Bouferguene, Explore the application of reinforced learning to support decision making during the design phase in the construction industry, in: Procedia Manuf, Elsevier B.V., 2020: pp. 181–187.

[13] N. Kayhani, H. Taghaddos, A. Mousaei, S. Behzadipour, U. Hermann, Heavy mobile crane lift path planning in congested modular industrial plants using a robotics approach, Autom Constr 122 (2021) 103508.