OPTIMIZING POST-EARTHQUAKE BRIDGE NETWORK RESTORATION PLANNING USING GRAPH NEURAL NETWORKS AND DEEP REINFORCEMENT LEARNING

MUHAMMAD ALI[1], AO DU, PH.D.[2*], AND JIANNAN CAI, PH.D.[3]

[1]School of Civil and Environmental Engineering, and Construction Management,
University of Texas at San Antonio
1 UTSA Circle, San Antonio, TX 78249
muhammad.ali2@utsa.edu

[2*]School of Civil and Environmental Engineering, and Construction Management,
University of Texas at San Antonio
1 UTSA Circle, San Antonio, TX 78249
ao.du@utsa.edu (Corresponding Author)

[3]School of Civil and Environmental Engineering, and Construction Management,
University of Texas at San Antonio
1 UTSA Circle, San Antonio, TX 78249
jiannan.cai@utsa.edu

Key words: Bridge networks, post-earthquake restoration, sequential decision-making, deep reinforcement learning, graph neural networks

Abstract. *Transportation networks are essential to the regional built environment and are highly vulnerable to seismic hazard impact. Efficient and rapid restoration of bridges following an earthquake is crucial to minimizing disruption and ensuring rapid recovery of the network. However, it is a challenging spatial-temporal decision-making problem, where sequential repair decisions should be made to the spatially distributed bridges within the transportation network under various uncertainties and constraints. This study proposes an innovative approach to optimize the restoration planning of post-earthquake bridge networks, by introducing Graph Neural Networks (GNNs) into Deep Reinforcement Learning (DRL) based decision policies. The integrated restoration decision-support tool can holistically consider the bridge network's spatial and topological information and the complex temporal dynamics during the restoration process. The restoration decisions are optimized under budget constraints through a compound reward function design that aggregates multiple reward signals during the restoration process, such as indirect costs, budget exceedance, and incentives for repair actions. The proposed GNN-DRL agent can effectively coordinate and prioritize network-level bridge repairs based on the bridges' damage states, restoration status, connectivity, and importance within the network. This method is evaluated on two small-scale bridge networks, with a comparative study against a ranking based benchmark approach. The results demonstrate that the integrated GNN-DRL agent outperforms the ranking based approach, achieving a higher success rate and improved restoration performance. The findings reflect promising results in the effectiveness of combining GNN and DRL for efficient decision-making of spatial bridge networks, ensuring the restoration of high-priority bridges and minimizing the overall post-earthquake socioeconomic impact.*

1 INTRODUCTION

Transportation infrastructure is essential for economic activities, public safety, and social mobility. Among its components, bridges play a crucial role in maintaining network continuity but are vulnerable to seismic hazards. Earthquakes have historically caused significant bridge damage, as seen in the 1994 Northridge, 1995 Kobe, and 2011 Tohoku earthquakes [1–3]. Such disruptions can severely hinder emergency response and recovery efforts, amplifying societal and economic losses. Therefore, rapidly restoring bridge functionality is vital for post-earthquake recovery [4]. Yet, restoration planning is complex as it requires optimal repair sequencing under spatial, temporal, and budget constraints, compounded by uncertainties in damage, resource availability, and interdependencies within the network.

To support recovery planning, researchers have developed component-ranking approaches that prioritize bridges based on static metrics like centrality or proximity to critical services [5,6]. While efficient, these methods often overlook system dynamics and interdependencies. Optimization-based strategies, such as mixed-integer programming and genetic algorithms [7,8], offer greater flexibility but can be computationally intensive and less adaptable in real-time scenarios.

The emergence of artificial intelligence (AI), especially deep reinforcement learning (DRL), has introduced new opportunities for post-hazard infrastructure recovery. DRL has shown success in traffic control, construction robotics, and infrastructure maintenance [9-11], but often treats components independently, missing key network relationships. Graph Neural Networks (GNNs), capable of modeling graph-structured data, address this limitation by incorporating spatial and topological dependencies. Recent applications include traffic forecasting, structural health monitoring, and disaster response [12-14]. Integrating GNNs with DRL allows for learning spatially informed, system-wide recovery policies. For example, GNN-DRL models have shown promising results in traffic control and disaster logistics [14] but rarely have been applied to post-earthquake bridge restoration.

This study addresses this gap by introducing a GNN-DRL framework for optimizing the restoration of earthquake-damaged bridge networks. The model uses GNNs to represent the spatial and structural attributes of bridges and employs a DRL agent to make sequential repair decisions. A tailored reward function incentivizes actions that improve long-term network resilience while adhering to cost constraints. The framework is adaptable to different network configurations and spatial damage scenarios, offering promising performance for post-disaster recovery planning.

2 METHODOLOGY

To apply DRL-based decision-making for post-earthquake bridge network restoration planning, two core components need to be established, including a virtual simulation environment capturing bridge network restoration dynamics, and the AI agent architecture.

At each discrete time step, the AI-based decision-making agent interacts with the simulated environment by selecting repair actions based on the current state or status of the bridge network. After executing the action, the agent receives updated observations from the environment and is provided with a reward that reflects the effectiveness of its decision, considering factors such as repair progress, budget adherence, and system-level recovery.

Through repeated interactions and feedback, the agent iteratively improves its policy, and eventually learns the optimal policy for repair decisions that maximize the long-term cumulative reward. This learning process is schematically illustrated in Figure 1, which outlines the overall architecture of the proposed GNN-DQN framework for post-earthquake bridge network restoration.

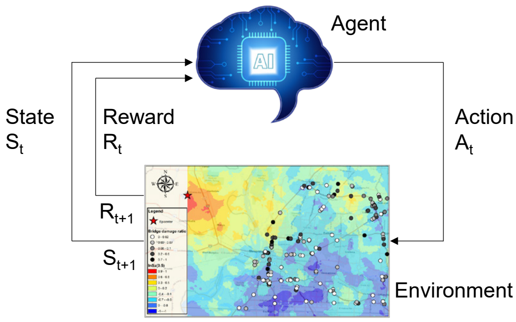

Figure 1: Agent and Environment Dynamics

2.1 Environment

The environment simulates the post-earthquake bridge network restoration process under various possible repair action combinations. Formally speaking, the environment is defined as a tuple {G,S,A,r,γ}, as illustrated subsequently:

- G: a graph defining the topology of the bridge network, where each bridge is represented as a node, and an edge exists if two nodes are directly connected. The connectivity of different node pairs is represented by the adjacency matrix. Each node has attributes that are defined in the State Space.
- S: State space of each bridge node, representing all the bridge-related parameters at each time step:
  - Damage State: One of five HAZUS [15] defined levels—None, Slight, Moderate, Extensive, Complete—randomly assigned at the start and updated to None upon repair.
  - Initial Damage State: The original damage level of each bridge recorded at the beginning of each episode to track repair progress over time.
  - Centrality Measure: Betweenness and closeness centrality metrics to indicate each bridge's importance within the network.
  - Functionality Ratio: A ratio used to track the functionality of each bridge.
  - Average Daily Traffic (ADT): The typical daily traffic volume for each bridge.
  - Deck Area: Deck area of each bridge
  - Length of Detour: The travel distance required if a bridge is non-functional.
  - Repair Cost: A function of both the bridge's damage state and its deck area.
  - Repair Duration: Indicates the time required to complete repairs.
  - Repair Flag: A binary indicator denoting whether a bridge is currently under repair.
  - Stepwise Available Budget: The remaining repair budget updated each step.
- A: Action space of each bridge node.
- r: Reward function.
- γ: Discount factor, used to weigh future rewards in the agent's decision-making process.

At each time step, the agent is allocated a specific budget (2% of the total bridges repair cost), which accumulates overtime and is used to fund the selected repair actions. Each repair action incurs a direct repair cost at the moment the repair is initiated, which is deducted from the cumulative budget. In addition, indirect costs, reflecting disruptions to transportation flow and economic activities, are continuously updated for every damaged bridge based on their restoration progress. For a given damaged bridge, after the repair is initiated, the indirect costs gradually decrease with time until zero, when the bridge is fully restored. This dynamic introduces a temporal incentive for the agent to initiate repair actions.

The agent must make repair decisions subject to budgetary constraints. If a selected action leads to a cumulative cost that exceeds the remaining budget, a substantial penalty proportional to the exceeded amount is applied to the reward. This soft constraint mechanism discourages infeasible repair sequences and promotes budget-aware decision-making. The environment continuously tracks bridge repair status, repair duration, and cost metrics, allowing the agent to learn an optimal restoration policy through repeated interactions. The goal of the agent is to maximize the expected cumulative reward over time, the instantaneous and long-term reward are calculated as below:

$$r_t = r(s_t, a_t, s_{t+1}) \tag{1}$$

$$G_t = \sum_{k=0}^{T-t} \gamma^k \, r_{t+k} \tag{2}$$

The instantaneous reward r_t is based on each time step t where s_t is the current state, a_t is the action taken at step t, s_{t+1} is the next state after action a_t, and r is the reward function. The long-term reward G_t is the cumulative reward at k steps where the discount factor is γ.

As shown in Equation 3, the reward function is designed to reflect the balance among repair efficiency, budget feasibility, and long-term resilience outcomes. The key components of the reward function are presented below:

$$r_t = -\sum_i w_i C_i^{indirect} + \sum_i w_i \beta f_i C_i^{indirect} + C_{comp} - \lambda \max\left(0, B_t - B_t^{avail}\right) \tag{3}$$

- **Network-level Indirect Cost ($-\sum_i w_i C_i^{indirect}$):** Negative aggregated indirect cost across all damaged bridges, weighted by the centrality measure w_i.
- **Bridge-level Repair Bonus ($\sum_i w_i \beta f_i C_i^{indirect}$)**: Incentive, controlled by β, added for bridges under repair. f_i denotes the repair flag, where $f_i = 1$ if a bridge is under repair.
- **Network-level Repair Completion Bonus** (C_{comp}): Bonus awarded if the entire bridge network is fully restored.
- **Network-level Budget Exceedance Penalty** ($-\lambda \max\left(0, B_t - B_t^{avail}\right)$): Penalty applied if the repair cost B_t exceeds the available budget B_t^{avail} in any given time step.

The reward function integrates both bridge-level and network-level objectives, promoting efficient local actions while ensuring overall network recovery and budget compliance. To further illustrate the environment dynamics and the reward formulation process, Figure 2 presents a schematic overview of the simulation framework.

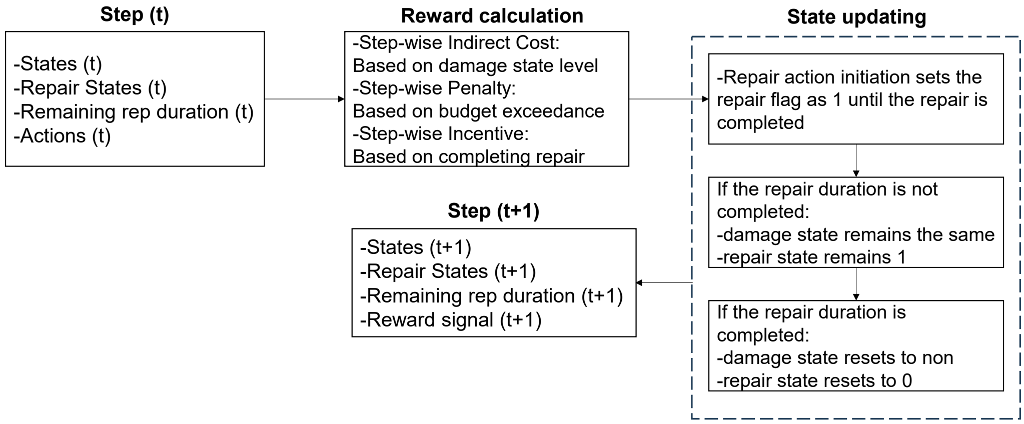

Figure 2: Single step Environment Dynamics

2.2 GNN-DQN Agent Architecture

The proposed model integrates a GNN with a Deep Q-Network (DQN) [16] to leverage the advantages of both spatial graph representation learning and deep reinforcement learning under uncertainty. DQN is selected for its ability to approximate value functions using neural networks, making it well-suited for high-dimensional, sequential decision-making tasks. The input to the agent is a graph-structured representation of the bridge network $G = (V, E)$, where each node $v \in V$ denotes a bridge and the edges $e \in E$ represent the connectivity between bridges based on road topology. The network topology is randomly generated from Simulation of Urban Mobility (SUMO) [17].

Each node is associated with a feature vector comprising the state parameters defined in the environment section. These vectors are processed through GNN layers which aggregate and propagate information from neighboring nodes. This allows the model to learn high-dimensional node embeddings that reflect both dependencies and attribute similarities within the bridge network. The encoded graph features are then passed into the DQN network layers. The output layer returns Q-values for each possible action combination, reflecting the expected cumulative reward of initiating bridge repairs at the current decision step. The agent then selects the highest Q-value action combination and deploys it in the environment.

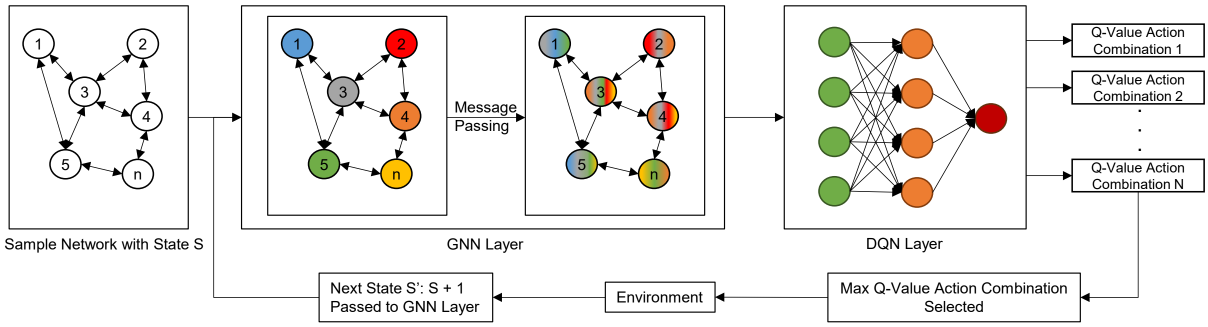

Figure 3: GNN-DQN Agent Architecture

2.3 GNN-DQN Model Training Process

The agent is trained using the DQN procedure based on the standard trial-and-error paradigm, enhanced by prioritized experience replay [18] to improve sample efficiency. An ε-greedy exploration strategy [19] is employed, where the agent initially explores the environment broadly and gradually shifts toward exploiting learned policies as training progresses. This encourages broad exploration during early training, gradually transitioning to exploitation as learning progresses. For a given bridge network, the bridge related parameters are defined and kept constant. The training process is conducted over 2000 episodes and follows a sequence of steps outlined below:

(1) Episode Initialization: A damage scenario is randomly generated for the bridge network at the beginning of each episode. (Deck Area, ADT, Length of detour, and centrality metric).

(2) State Representation: At each decision step, the environment provides the agent with a graph comprising the corresponding node features. This graph is passed through the GNN-DQN model to predict Q-values for all possible actions.

(3) Action Selection: An action is selected using an epsilon-greedy strategy, where the agent either selects the action with the highest predicted Q-value or randomly samples an action based on a decaying exploration probability. Once the repair of a bridge is initiated, the repair process continues uninterrupted for the prescribed repair duration.

(4) Environment Update and Reward Calculation: Upon executing a repair action, the environment is updated to reflect the new state of the bridge network. The immediate reward is calculated based on improvements in network-level resilience and cost-related outcomes. Following the action, the next state and all relevant bridge-specific parameters are updated and fed back into the GNN-DQN model to inform subsequent decisions. The reward signal is designed to capture both the immediate benefits of the action and the long-term impact on the overall system, as defined by:

$$Q^*(s,a) = (1-\gamma)r + \gamma \; \arg\max Q^*(s',a') \tag{4}$$

Here, $Q^*(s,a)$ represents the expected cumulative reward of taking action a with network in state s. The term r is the immediate reward reflecting the system's performance improvement resulting from the selected action. The expression $\gamma \cdot \arg\max Q^*(s',a')$ denotes the highest anticipated future reward, where s' is the next state of the system and a' is the optimal action taken in that state. The discount factor $\gamma \in [0,1]$ balances the importance of immediate versus future rewards, with γ=0 emphasizing only immediate outcomes and γ=1 prioritizes long-term benefits.

(5) Experience Replay and Model Update: The agent stores each transition tuple $\{G,S,A,r,\gamma\}$ in a prioritized experience replay buffer, where transitions are sampled based on their significance to learning. This prioritization increases the learning efficiency by allowing the agent to focus more on experiences that are expected to yield higher learning value. Mini batches are periodically sampled from this buffer to train the agent, promoting more stable and accelerated convergence compared to uniform sampling.

(6) Episode Termination: Each episode ends either when all bridges in the network have been fully repaired or when the predefined maximum number of decision steps is reached.

In this study, the time horizon is set to 48 steps, representing a 48-month total repair period. Steps (2) through (6) are repeated for 2000 episodes. After training, the model is tested over 1000 unseen damage scenarios to evaluate generalization and performance.

2.4 GNN-DQN Policy Deployment

Once trained, the GNN-DQN model serves as a pre-trained decision-making agent for post-earthquake bridge network recovery. During the deployment phase, the policy parameters are fixed and no longer updated; instead, the agent utilizes its learned knowledge to make sequential repair decisions in response to new damage scenarios. At each time step, the agent evaluates the current network state—represented as a graph encoding both structural attributes and topological relationships—and ranks all available repair actions based on their predicted Q-values. The agent selects the action with the highest predicted long-term benefit while strictly enforcing the budget constraint, ensuring that no action is taken that violates the available budget.

3 CASE STUDY SCENARIOS

To evaluate the performance of the proposed GNN-DQN framework, we conducted two distinct case studies on two different 13-bridge network configurations. Each configuration represents a hypothetical bridge network with varying topological characteristics and earthquake damage scenarios. The agent is trained separately in each configuration to capture network-specific spatial dependencies and develop a repair policy. For comparative purposes, we consider a benchmark ranking-based restoration strategy, where bridges are ranked according to a composite importance score derived by multiplying four key indicators: Bridge functionality ratio, ADT, bridge centrality metric, and deck area. Restoration actions are executed sequentially by selecting the highest-ranked damaged bridges within the available budget at each step until all bridges are repaired or the study time horizon is completed.

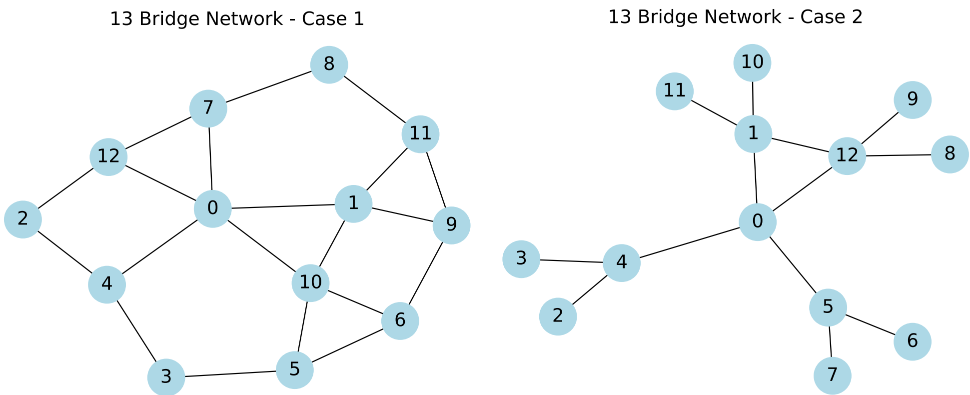

Figure 4: Case Study Network Topologies

4 RESULTS

For both case studies, the trained GNN-DQN and the ranking based models are tested under identical settings, with 1000 different initial damage state scenarios.

4.1 Comparison of Restoration Progress at the End of the Planning Horizon

The overall restoration progress at the end of the restoration planning horizon is evaluated based on key network-wide statistics, averaged across 1,000 test episodes for each case. The performance metrics include:

- **Repaired Bridges Ratio**: the percentage of bridges for which repairs were fully completed.
- **Repair Initiated but Not Completed**: the percentage of bridges where repair started but not finished due to time limitations.
- **Total Bridges with Repair Action**: the percentage of bridges where repair is initiated or completed.
- **Initial and Final Network Functionality**: computed from the average damage ratio at the beginning and the end of each episode.
- **Functionality Improvement**: the average relative gain in functionality across the episodes - the ratio difference between final and initial functionality ratio's

Table 1 summarizes the performance of the GNN-DQN model compared to the ranking-based model across two network configurations. It is shown that the GNN-DQN model is more effective than the ranking based approach, by offering higher repair rates, lower incompletion ratios, and greater improvements in the network functionality. For example, the GNN-DQN model repairs a larger portion of the network and initiates more timely actions, leading to fewer incomplete repairs. It also achieves a higher level of final network functionality—improving functionality by over 83% in both cases—compared to roughly 70% improvement under the ranking based model.

Table 1: Case Study Simulation Results Statistics

Simulation	Repaired Bridges Ratio	Repair Initiated Ratio but Not Completed	Total Bridges with Repair Action	Initial Functionality Average	Final Functionality Average	Functionality Improvement
GNN-DQN Case 1	88.33%	1.28%	89.61%	49.95%	91.77%	83.04%
Ranking Based Case 1	77.30%	4.00%	81.31%	49.83%	84.87%	69.90%
GNN-DQN Case 2	88.35%	1.20%	89.54%	49.86%	91.78%	83.13%
Ranking Based Case 2	77.19%	4.21%	81.40%	50.42%	85.01%	69.89%

4.2 Comparison of Network Performance Over Time

To further evaluate the restoration efficiency, we plotted the network functionality over time. This provides insights into how quickly the network regains functionality under different restoration strategies. Network performance is measured monthly over a 48-month simulation window, and the curves in Figure 5 represent the average performance across 1,000 test scenarios for each model. From the average restoration trends for both cases, the GNN-DQN model shows a clear advantage in accelerating and sustaining network restoration over time. While both strategies start from similar baseline performance, the GNN-DQN model consistently drives faster and makes more effective recovery throughout the planning horizon. This not only results in higher final performance levels but also achieves them more efficiently, minimizing downtime and enhancing overall system resilience. The ability to restore functionality more quickly is critical in post-disaster scenarios, where timely recovery can significantly reduce societal and economic disruptions.

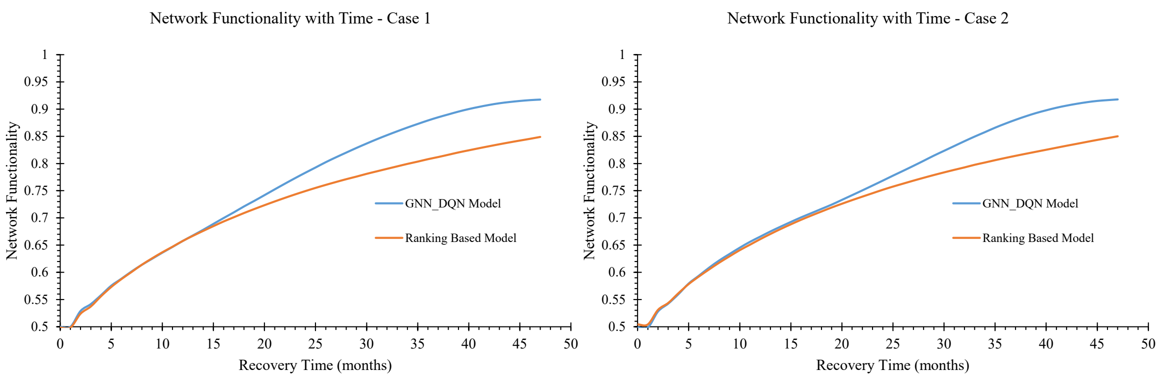

Figure 5: Comparison of Case Study Network Functionalities

4.3 Comparison of Spatial-Temporal Repair Progress Evolution

To gain spatial and temporal insights into the restoration progress, heatmaps of the network were generated at selected time steps for both the GNN-DQN and ranking-based methods for both case studies (Figures 6-9). The heatmaps visualize the average performance level at each bridge, where blue indicates higher performance (i.e., lower alpha residuals) and red indicates poor performance.

In both case studies, the GNN-DQN approach demonstrates a more consistent and spatially optimized recovery progression compared to the ranking based method, with central bridges improving earlier in the timeline. These findings highlight GNN-DQN's capability to make spatially informed decisions that enhance not only short-term restoration efficiency but also contribute to improved long-term network functionality and resilience.

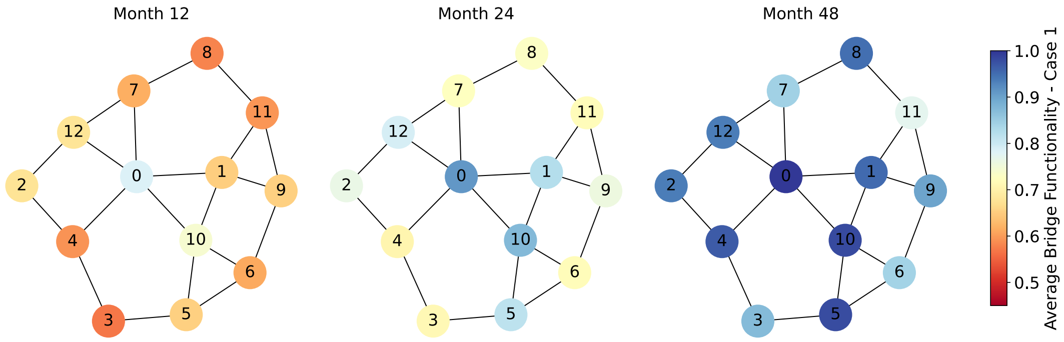

Figure 6: Heatmap for GNN-DQN - Case Study 1

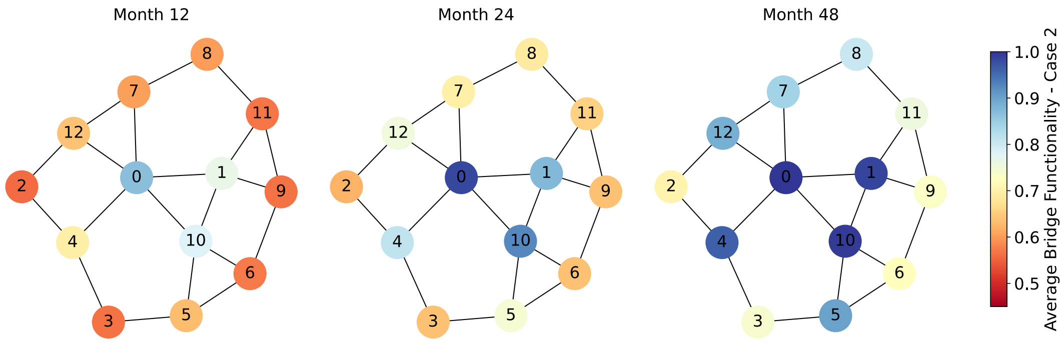

Figure 7: Heatmap for Ranking Based - Case Study 1

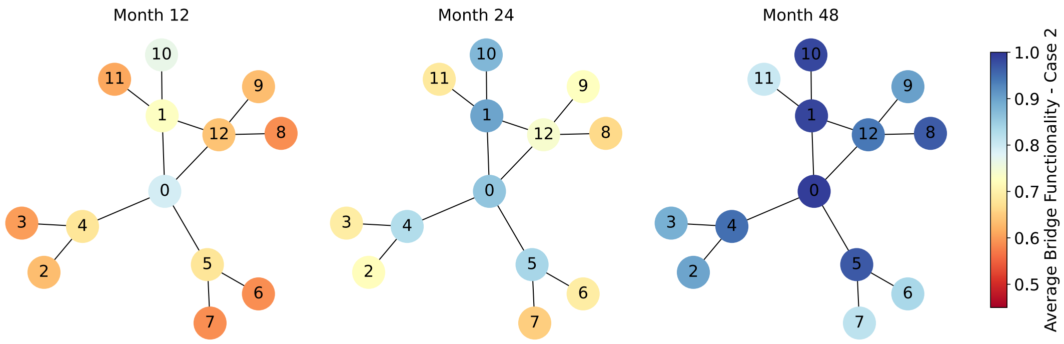

Figure 8: Heatmap for GNN-DQN - Case Study 2

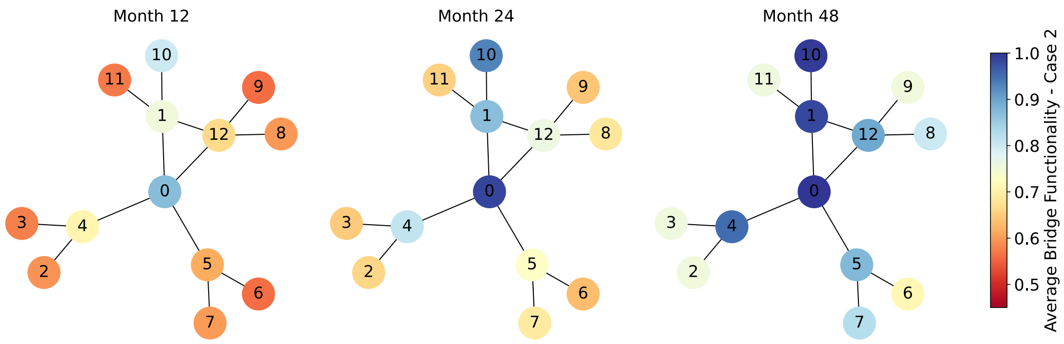

Figure 9: Heatmap for Ranking Based - Case Study 2

5 CONCLUSION

Post-earthquake recovery is critical for reducing long-term infrastructure damage and restoring functionality. This study presents an AI-based decision-making framework for optimizing bridge repair scheduling by incorporating bridges network topology with deep reinforcement learning techniques. By combining graph neural networks (GNN) with deep Q network (DQN), the proposed framework can better prioritize sequential bridge repairs based on their structural conditions and importance within the network. The framework also considers budget constraints, direct and indirect repair costs, and time-dependent functionality, enabling more strategic and resilient post-disaster restoration planning. The proposed model was compared against a ranking-based approach and consistently delivered better outcomes such as higher repair completion rates, more effective action execution, and improved final network performance. Notably, the GNN-DQN agent enabled faster recovery timelines and produced more spatially balanced repair strategies, as evidenced by functionality growth curves and heatmap visualizations.

These results underscore the potential of AI-based and topology-aware policies to outperform heuristic ranking-based methods, particularly under constrained post-disaster conditions. Future work will test the framework under more realistic spatial seismic damage realizations, and further improve the scalability of the framework toward larger scale bridge networks

6 ACKNOWLEDGMENT

This work was supported by the NSF/NIST Disaster Resilience Research Grants (DRRG) program under contract #70NANB24H012.

7 REFERENCES

[1] Chang, S. E., & Nojima, N. (2001). Measuring post-disaster transportation system performance: the 1995 Kobe earthquake in comparative perspective. *Transportation research part A: policy and practice*, *35*(6), 475-494.

[2] Kawashima, K. (2000). Seismic performance of RC bridge piers in Japan: an evaluation after the 1995 Hyogo-ken nanbu earthquake. *Progress in Structural Engineering and Materials*, *2*(1), 82-91..

[3] Ozawa, S., Nishimura, T., Munekane, H., Suito, H., Kobayashi, T., Tobita, M., & Imakiire, T. (2012). Preceding, coseismic, and postseismic slips of the 2011 Tohoku earthquake, Japan. *Journal of Geophysical Research: Solid Earth*, *117*(B7).

[4] Du, A., Wang, X., Xie, Y., & Dong, Y. (2023). Regional seismic risk and resilience assessment: Methodological development, applicability, and future research needs–An earthquake engineering perspective. *Reliability Engineering & System Safety*, *233*, 109104.

[5] Tang, J., Wan, L., Nochta, T., Schooling, J., & Yang, T. (2020). Exploring resilient observability in traffic-monitoring sensor networks: A study of spatial–temporal vehicle patterns. *ISPRS International Journal of Geo-Information*, *9*(4), 247.

[6] Aydin, N. Y., Duzgun, H. S., Heinimann, H. R., Wenzel, F., & Gnyawali, K. R. (2018). Framework for improving the resilience and recovery of transportation networks under geohazard risks. *International journal of disaster risk reduction*, *31*, 832-843.

[7] Zhang, W., Wang, N., & Nicholson, C. (2017). Resilience-based post-disaster recovery strategies for road-bridge networks. *Structure and Infrastructure Engineering*, *13*(11), 1404-1413.
[8] Nozhati, S., Sarkale, Y., Ellingwood, B., Chong, E. K., & Mahmoud, H. (2019). Near-optimal planning using approximate dynamic programming to enhance post-hazard community resilience management. *Reliability Engineering & System Safety*, *181*, 116-126.
[9] Li, L., Lv, Y., & Wang, F. Y. (2016). Traffic signal timing via deep reinforcement learning. *IEEE/CAA Journal of Automatica Sinica*, *3*(3), 247-254.
[10] Cai, J., Du, A., Liang, X., & Li, S. (2023). Prediction-based path planning for safe and efficient human–robot collaboration in construction via deep reinforcement learning. *Journal of computing in civil engineering*, *37*(1), 04022046.
[11] Du, A., & Ghavidel, A. (2022). Parameterized deep reinforcement learning-enabled maintenance decision-support and life-cycle risk assessment for highway bridge portfolios. *Structural Safety*, *97*, 102221.
[12] Tran, T., He, D., Kim, J., & Hickman, M. (2023). MSGNN: A Multi-structured Graph Neural Network model for real-time incident prediction in large traffic networks. *Transportation research part C: emerging technologies*, *156*, 104354.
[13] Li, T., Wu, S., Huang, Y., & Li, H. Structural Damage Detection Using Graph Neural Network-Based Probabilistic Graphical Models. *Available at SSRN 5143083*.
[14] Saadi, A., Abghour, N., Chiba, Z., Moussaid, K., & Ali, S. (2025). A survey of reinforcement and deep reinforcement learning for coordination in intelligent traffic light control. *Journal of Big Data*, *12*(1), 1-20.
[15] Federal Emergency Management Agency. (2015). *Appendix 2: Damage State Descriptions for Seismic Fragility Curves*. In *HAZUS-MH 2.1 Earthquake Model Technical Manual*. Earthquake Engineering Research Institute. https://www.eeri.org/images/sesi/Appendix-2-HAZUS-Damage-States.pdf
[16] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.
[17] Krajzewicz, D., Hertkorn, G., Rössel, C., & Wagner, P. (2002). SUMO (Simulation of Urban MObility)-an open-source traffic simulation. In *Proceedings of the 4th middle East Symposium on Simulation and Modelling (MESM20002)* (pp. 183-187).
[18] Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2015). Prioritized experience replay. *arXiv preprint arXiv:1511.05952*.
[19] Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.