Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

HYBRID SIMULATION AND REINFORCEMENT LEARNING-BASED SCHEDULING FOR RESILIENT INFRASTRUCTURE NETWORKS

Pavithra Sripathanallur Murali,[1] and Shima Mohebbi[1]

[1]Department of Systems Engineering and Operations Research,
George Mason University, Fairfax, VA, USA.

ABSTRACT

Infrastructure systems are interdependent at various levels, and their collective performance is influenced by factors such as topology, budgetary decisions, resource availability, and awareness of interdependency. Traditional resource allocation models for improving resilience often assume a single decision-maker overseeing all scheduling decisions. However, critical infrastructures, characterized by a network-of-networks structure, are managed by individual entities with distinct boundaries. Moreover, the dynamic and stochastic nature of decision-making processes cannot always be captured via mathematical programming. This study develops a hybrid simulation model that merges top-down and bottom-up approaches. It captures organizational-level budgetary decision-making dynamics through system dynamics, and maintenance activities alongside evolving network performance through an agent-based model. Optimal restoration strategies maximizing network resilience are identified via deep reinforcement learning, constrained by financial allocations. This approach is applied to water distribution and mobility networks in Tampa, FL demonstrating our method's efficacy for restoring interdependent infrastructures.

1 INTRODUCTION

The United States Department of Homeland Security (DHS) has identified 16 critical infrastructures (CI) crucial for the sustained functionality of a city (DHS 2003). These systems exhibit significant inter connectivity, influenced by management practices, budgetary constraints, and geographical proximity (Rinaldi et al. 2001). The interdependency of infrastructure networks often leads to a domino effect, where failures in one system can precipitate cascading failures across others (Buldyrev et al. 2010). For example, Hurricane Irma's devastating impact on Saint-Martin in 2017, resulting in significant deaths, economic damage, and prolonged isolation due to damaged CIs, exemplifies the cascading effects and vulnerability caused by extreme events on interconnected systems (Der Sarkissian et al. 2022). Therefore, restoration activities must depend not only on the network elements but also on other elements that are physically or logically connected. Spatial interdependencies, for example, disrupt co-located roads during pipe repairs, while social interdependencies are evident when households use multiple utilities simultaneously. Ouyang (2014) and Mohebbi et al. (2020) present strong cases that acknowledgement of various interdependencies is the first step towards addressing disruptions in CIs. Furthermore, financial interdependency introduces a competitive dynamic for shared resources that complicates managerial decisions in restoration processes (Zhang and Peeta 2011). To address the effects of financial allocations on resilience, Karamouz et al. (2019) applied a multi-criteria decision-making approach to New York City's interconnected wastewater treatment plants.

Resilience — the capacity to withstand disruptions is a critical characteristic of infrastructure systems. Resilience relates to (i) the level of maximum operational capability (Pant et al. 2014), (ii) the short-term capacity to cope with disruptions (Yohe and Tol 2002), and (iii) the long-term adaptability of the system to maintain or enhance its functionality in the face of adverse changes in its operational environment (Gallopín 2006). With climate change predictions indicating rising sea levels of 0.3–1.3 m by 2100 (Hayhoe et al.

2018), increased frequency of tropical storms, and potential flooding in coastal cities (USGCRP 2017), understanding and enhancing the resilience of infrastructure systems is paramount. Swift restoration of failed components is essential for obtaining improved network resilience. Therefore, carrying out restoration activities solely based on myopic conditions such as FIFO (First in First Out), cannot ensure the ideal scheduling. Numerous studies have been carried out for developing optimal crew schedules in infrastructure systems. Mathematical modeling and machine learning are some of the methodologies utilized by researchers to solve the restoration allocation problem in interdependent infrastructure systems (Baidya and Sun 2017; Sun and Zhang 2020; Rahimi-Golkhandan et al. 2022).

Infrastructure networks, characterized as the network-of-networks (Gao et al. 2011), face significant challenges in financial resource sharing and restoration scheduling due to competing interests between agencies, dynamically changing environments and failure modes. It is challenging to represent them via mathematical optimization models because of the complex nature of interactions and scalability in city-scale applications. Traditional methodologies, including standalone mathematical modeling and conventional machine learning approaches, often lack the ability to fully capture and simulate the dynamic and stochastic nature of interdependencies within complex infrastructure systems. This limitation underscores the necessity for a more comprehensive approach that can accommodate the multifaceted interactions and adapt to changing conditions. Simulation approaches are thus instrumental in depicting dynamic system changes and interactions among decision-makers. System dynamics (SD) modeling approach has allowed for the successful modeling of complex systems (Links et al. 2018; Li et al. 2020), offering valuable insights via computational experiments. SD models have been used to examine community resilience against natural disasters (Feofilovs et al. 2020), hospital seismic response capabilities (Khanmohammadi et al. 2018), and urban resilience in the face of epidemics (Zhang and Wang 2023), enhancing our understanding of systems under stress. The application of agent-based model (ABM) to model systems as agents has also provided profound insights. Esmalian et al. (2019) employed ABM to simulate the interplay among decision-makers and their surroundings, shedding light on the effects of varying strategies on infrastructure resilience and performance. Similarly, Kandiah et al. (2019) applied ABM to explore water reuse adoption and the growth of infrastructure in sociotechnical systems.

In operations research, simulation is a widely utilized tool to addressing multi-objective problems (Landa et al. 2016; Ko et al. 2006). However, simulating complex infrastructure networks presents challenges, particularly with cross-layered interactions that require different simulation methods. For instance, system dynamics is an ideal candidate for analyzing the impacts of policies and explore the mechanisms behind their effects on restoration, whereas agent-based modeling excels in capturing the behavior of the network and its intricate interactions between system actors (Ouyang 2014). An effective approach is to harness the power of integration and hybridize different modeling approaches within a single framework to leverage the strengths of different methods. Integration of the SD and ABM methods enables a feedback loop that augments model capabilities, making it particularly beneficial for addressing complex, multifaceted problems. The advantage of combining these two simulation methods is to facilitate feedback information between the models and enhance their capabilities, particularly when we deal with complex and multi-faceted problems.

Most scheduling problems turn out to be NP-hard. Reinforcement learning has been used as a viable alternative to centralized solution approaches to tackle complex scheduling in CIs (Dehghani et al. 2021). For instance, Qiu et al. (2022) introduces a multi-agent reinforcement learning strategy to enhance electric vehicle system resilience during the transition to low-carbon energy sources in power and transportation networks. Sun and Zhang (2020) is another study wherein they took an approach to model infrastructure networks through ABM along with physical interdependencies modelled by a dependency coefficient based on distance to obtain scheduling repair actions. Yang et al. (2024) developed a decision support model using function approximation with neural networks, graph theory, and an actor-critic reinforcement learning algorithm to facilitate optimal restoration policies for interdependent water, power, and transportation networks. Further, Wang et al. (2023) addressed the scheduling problem of repair crews and mobile power sources in microgrids by formulating a decentralized partially observable Markov decision process, and solving it with a hierarchical multi-agent reinforcement learning method using an actor-critic architecture. While these studies presented the problem of scheduling tasks in interdependent infrastructure networks, they do not represent different interdependencies that exist. In this study, we broaden our scope to incorporate both co-location and financial interdependencies. We also highlight the resource dynamics within the system and the corresponding impact on the network resilience evolution. With a centralized budget distributed among various CI agencies, we integrate critical financial resource dynamics through system dynamics, informing the restoration schedules within the network.

This study presents a hybrid simulation and reinforcement learning framework. We combine a system dynamics model to capture the dynamics of financial resources in infrastructure maintenance with an agent-based model to delve into the network-level evolution prompted by failures and organizational decisions. Our hybrid simulation environment facilitates cross-layered interactions, enabling an examination of the multifaceted impacts of dependencies and constraints on the restoration of infrastructure networks. A deep reinforcement learning (DRL) model is then designed to identify optimal restoration schedules under budgetary constraints. Our findings provide municipal decision-makers with actionable insights for strategic restoration scheduling during emergency response, thereby reducing recovery times and bolstering network robustness. The paper is structured as follows: Section 2 details the study site and outlines our methodology for model structure, SD & ABM development, evaluation, and strategy development. Section 3 discusses strategy testing and results for the City of Tampa case study. Section 4 concludes with future research directions and concluding remarks.

2 METHODS

In our study, we consider the restoration problem of i networks in a network-of-network setting, each representing a distinct agency. By utilizing a hybrid simulation model and deep reinforcement learning, the optimal order of repairs for failed components that maximizes each network's resilience is determined under resource constraints owing to budgetary decisions. We describe in detail our efforts to develop the hybrid simulation model by combining the SD and ABM techniques. First, we explain the approach we took to model financial resource dynamics (top-down). Next, we detail the bottom-up approach to capture the interactions of the system elements using the ABM. Finally we show the integration of SD and AB models and the application to the case study of Water and Mobility Departments in the City of Tampa, Florida.

2.1 System Dynamics Model

Here, we modeled the resource dynamics of the infrastructure network and its transition between states, influenced by changes in specific system variables. We took a layered approach to model the interaction of resource dynamics with the physical components. The financial layer manages maintenance funding, with resources triggering maintenance actions and funds flowing out based on maintenance costs. Inflows come from external sources such as federal funding, while emergency repair costs during events are estimated considering event magnitude and duration. The budget dynamics are modelled by the stock variables *Overall Financial Budget* - represents the total available budget for the utilities, *Financial Resource for Water Maintenance* - representing the budget share of Water Department and *Financial Resource for Mobility Maintenance* - representing the budget share of Mobility Department. Further, the *Cost of Road Maintenance* and the repair time was calculated based on the cost of milling per square yard, cost of mobilization and daily cost of traffic control and days to mill asphalt, days to mill base, days to restore friction course, and days to restore structural course (Lu et al. 2018). Similarly, *Cost of Pipe Maintenance* and the restoration time was calculated based on average cost of pipe repair per foot (Clark et al. 2002). The physical layer of the SD includes two stocks representing the physical components in *Good Condition* and *Poor Condition*. The components change from the state of good to poor due to random failures. We also included two important flow variables *Road Restoration*, and *Pipe Restoration* that informs about the allowable area or length that crews can restore in a given day based on the available financial resource, cost of restoration, number of available maintenance crews, and number of failed components. Further, flow variables such as failure magnitude manage the transition process. We verified the model by iteratively by comparing the financial shortfalls for restoration in each network generated by the model and the actual historic shortfalls for the Water and Mobility Departments in Tampa, FL. Upon statistical analysis we found that the difference between them is insignificant.

2.2 Agent-based Model

We first developed the simplified physical network of the pipeline and road networks using the road traffic and fluid libraries in AnyLogic (2024), a Java-based software with capabilities to model both ABM and SD. The networks consist of 197 pipes and 249 roads segments. Figure 1 shows the network as presented in our simulation model, where grey colored network represents the road network and the red one represents the water distribution network. Figure 2 also demonstrates the graphical user interface of the simulation model.

Figure 1: Water & transportation network.

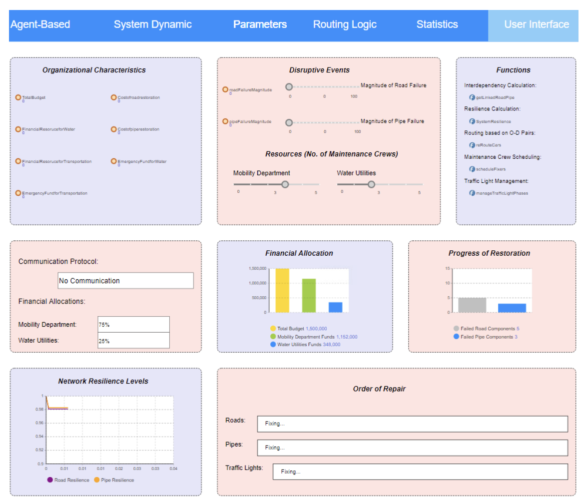

Figure 2: User interface of the hybrid simulation model.

The physical networks behave as a agent whose components can fail according to a random probability or defined conditions. The networks edges (links) that are co-located to each other within a distance threshold are considered interdependent. Therefore, a failure is transmitted to an interdependent component when its co-located one is disrupted. Crew agents belonging to the respective utilities are populated at their respective bases who travel to the nearest intersection location of failed components based on Dijkstra's shortest path algorithm to repair the failed component. Crews belonging to different utilities behave in a decentralized manner with no information exchange between them, which may simplify the coordination that often occurs in real-world scenarios. While this simplification may not fully capture the efficiency gains from explicit coordination and information sharing among crews, it allows us to focus on the individual decision-making capabilities of crews and explore the potential benefits of decentralized approaches, such as improved scalability and flexibility in decision-making. In the simulation exist cars that traverse through randomly generated origin-destination routing. Traffic light agents manage traffic through phasing rules defined in the model. Once a random failure is generated, the list of failed components are populated to the respective crew groups. Crew agents choose a shortest path to reach the failed component and restore it by spending the required repair time that changes based on an agent's random skill set. To verify the simulation model, we compared the traffic flow generated by the model with the historical average daily traffic data on various road segments in the City of Tampa, FL. Using statistical analysis of the actual vs simulated data, we ensured that the difference between them was insignificant. More detailed information can be found in Dsouza (2022).

2.3 Hybrid Simulation Model

Both the system dynamics and agent-based models we developed are standalone, simulating behavior independently without requiring external inputs. Their integration, however, markedly enhances the depiction of the interdependent infrastructure networks. Key variables from the SD model such as *Financial Resource for Water Maintenance*, *Financial Resource for Transportation Maintenance*, *Road Restoration*, and *Pipe Restoration* are incorporated into the ABM. Conversely, critical ABM variables such as the number of crews and the list of failures for each network feed back into the SD model. To capture the transition of physical components from a good to poor state in SD, we integrated the actual list of network components, modelled in the ABM, to represent the numbers of components in these states. The restoration limits, based on each network's financial resources, are applied in the ABM to realistically restrict daily repairs. To verify the restoration limits, we conducted a comparative analysis by varying the budget allocations and cost parameters within a plausible range. The simulated restoration limits were assessed for consistency across different financial allocation scenarios. The changes in the restoration limits aligned with the expected behavior and practical considerations, demonstrating the robustness of the calculations. Hence, crews can only repair failures within these financial constraints, ensuring repair efforts align with available resources. These limits are updated at every time step t, ensuring restoration activities adapt to changing resource availability. Then, a DRL algorithm was developed and integrated into the hybrid simulation model to enable agents with decentralized decision making capability. Figure 3 shows the framework that we developed. In order to measure the performance of restoration actions, we defined resilience of the network as an average of resilience of all networks, given that each network contributes equally to the resilience of the overall network. Specifically, the resilience of the interconnected network system $\mathscr{R}_t^I$ at each discrete time step t during the simulation, consisting $i \in \mathscr{I}$ distinct networks, is computed as the average resilience across all individual networks.

$$\mathscr{R}_t^{\mathscr{I}} = \frac{1}{I}\sum_{i=1}^{I}\mathscr{R}_t^I, \quad \forall i \in \mathscr{I}, \quad 0 \leq \mathscr{R}_t^{\mathscr{I}} \leq 1.$$

Finally, this hybrid model simulates the network under varying levels of failure magnitudes to understand the influence of system dependencies on the restoration of physical components.

2.4 Restoration Planning

In this work, as previously mentioned, we analyze the effect of financial and physical interdependencies on the network restoration after disruptions. When a failure occurs in the ABM, the allocated budget from SD is used as input to determine the sequence of repairs. We formulated this scheduling problem as a finite-horizon Decentralized - Markov Decision Process, encompassing the components of state space, action space, reward function, and state transition dynamics. At every time step t, agent's experience is represented by a 3-tuple: state s_t, action a_t, reward r_t. The transition of the agent's states are governed by the underlying SD-ABM simulation. The sources of uncertainty in our model arrives from the random failures and restoration times. Once the financial allocation is obtained from the SD model, the next important task is to determine an optimal order of repair to repair the failed components. DRL provides the agents within our simulation the ability to carry out tasks on their own through trial and error. DRL and the hybrid simulation are connected in such a manner that the agent's experience generated as a result of their exploration within the environment is utilized to train a neural network, called Deep Q-Network (DQN) that takes in the state, action and reward values. Once trained, the DQN will be capable of providing the

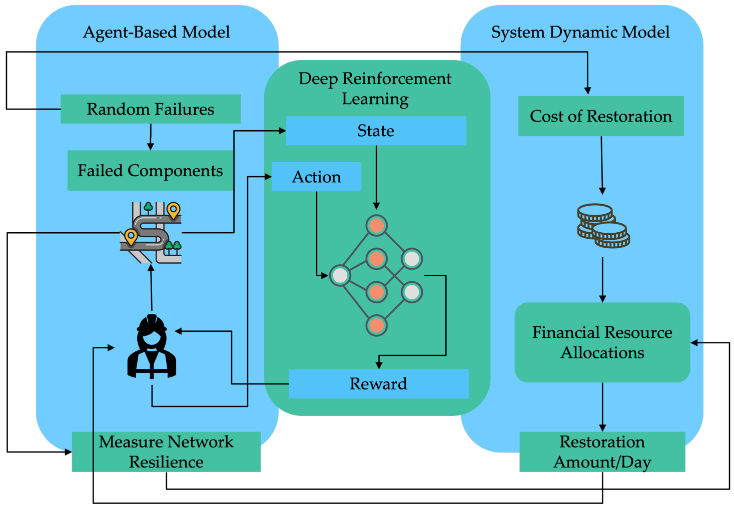

Figure 3: Hybrid simulation and restoration scheduling via deep reinforcement learning.

actions for agents that maximizes the reward earned in any state. Based on this formulation, let $f^i \in F$ represent the lists of failed components in each network i. Then,

- The state of the system at any time t is defined by the tuple $\mathscr{S} : \{f_t^1, f_t^2 .....f_t^i\}$, where f_t^i represent the lists of failed components in the i^{th} network.
- The action space $\mathscr{A}(S_t) = \{a_t^1, a_t^2 .....a_t^i\}$ comprises the order of repair actions for each network i in which the repairs are executed by the crews.
- The reward function $R : S \times A \times S \rightarrow \mathbb{R}$ of an agent is the immediate reward defined as a weighted sum of the unit length or area of the failed components f^i to directly reflect the priority of restoration for longer/larger failed components, i.e., $r_t^i = \sum_{i \in f^i}(w^n) * (q_t^n)^i$. $(q_t^n)^i$ denotes the length/area of the failed component n belonging to the network i. $w_n = \frac{f-j+1}{\frac{f \cdot (f+1)}{2}}$ is the weight, where f is the number of failed components in f^i, and j is the position in the repair array sequence.

Utilizing a repair sequence of length f–representing the number of failed components– and the position j within the order of repair array $\mathbf{a}_t^i$, we assign linearly decreasing weights, w^n, to the rewards (see Shi et al. (2021) for more details). This weighting approach prioritizes the restoration of components based on their sequence in the array, with earlier positions denoting higher importance for prompt restoration. The system transitions to the next state are governed by the underlying hybrid simulation model. The objective is to maximize the expected cumulative reward over the decision horizon, formalized through the state value function as below.

$$Q^i(s^i, a^i) = \mathbb{E}_{(s^i, a^i \rightarrow s'^i, r^i)} \left[r^i(s_t^i, a_t^i, s_{t+1}^i) + \gamma \cdot \max_{a'^i} Q(s_{t+1}^i, a'^i) \right] \tag{1}$$

where γ represents the discount factor for future rewards. Agent selects an action a^i from the current observation s^i that leads to the next observation s'^i, collecting rewards in the process. These rewards comprise not only the immediate payoff r^i but also the anticipated future rewards, discounted by a factor γ. The term a'^i denotes the prospective actions from the subsequent observation s_{t+1}^i, chosen to optimize future rewards. Within a defined planning horizon, the agents adhere to a policy $\pi(s_t^i, a_t^i)$ which dictates the viable actions a_t^i contingent on the current state s_t^i. The aim is to ascertain the optimal policy π_i^* that maximizes the state-action value function $Q^{i*}(s,a)$, mirroring the expected value formulation presented in equation (2) as follows.

$$\pi^{i*} = \arg\max_{a} Q^{i*}(s^i, a^i) \tag{2}$$

As the size of the state and action in our large-scale CI network is very huge, we utilize a neural network to approximate the state-action value function presented in the equation (1). The $Q^i(s^i, a^i)$ is approximated with weights θ^n using a neural network and its policy can be defined as follows.

$$\pi^{i*} = \arg\max_{a} Q^{i*}(s^i, a^i, \theta^i)$$

The DQN can now be trained to minimize their temporal difference loss function $L(\theta^i)$ (see equation 3). It should be noted that we calculate the target Q value of the current state by the target network $Q^i(s^i, a^i, \theta^{i'})$. We keep updating $Q^i(s^i, a^i, \theta^i)$ during the training while only copying it to $Q^i(s^i, a^i, \theta^{i'})$ after several steps to stabilize the learning procedure.

$$L(\theta^i) = \mathbb{E}_{(s^i, a^i, \theta^i)}\left[\left(Q^i(s^i, a^i, \theta^i) - Q^i(s^i, a^i; \theta^{i'})\right)^2\right] \tag{3}$$

3 COMPUTATIONAL STUDY

3.1 Financial Allocation

At various overall budgets for the CI networks, we carried out multiple simulations to collect data at different failure rates in the network (see Table 1). The financial allocation ratios provided the financial resource to each of utilities in the CI network, which in turn constraints the per day restoration efforts for crews. We further computationally tested these financial ratios (Water: Mobility) between the Water and Mobility departments to infer their respective resilience improvements in the network. We identified that the ratios of 27:73 and 23:77 enhance network resilience more effectively for 5% and 10% failure rates, respectively. Figures 4 and 5 show the Water network resilience over time. Notably, with a budget of \$5M and a 23% financial allocation, Water network exhibited a better resilience improvement. At a \$1.5M budget, the resilience reached a plateau at 0.99 and 0.97 for 5% and 10% scenarios. In contrast, with a \$5M financial budget, Water networks resilience rose to 0.996 and 0.987 for 5% and 10% failure rates, respectively. Similarly, Figures 6 and 7 show the Mobility network resilience improvements at different financial budget and allocations. It can be observed that a \$1.5M financial allocation resulted in resilience plateaus at 0.975 and 0.90 for 5% and 10% failure rates, while a \$5M budget yielded resilience improvements to 0.985 and 0.94 for 5% and 10% failure rates, respectively. Although improvements in resilience may seem marginal, they are significant given the extensive scale of the network, where even smaller improvements in resilience translates into significant enhancements in network performance.

Table 1: Resource allocation and crew distribution based on failure rates.

Failure	Budget (million dollars)	Finance (%): Water	Finance (%): Mobility	# Crews: Water	# Crews: Mobility
5%	1.5	0.20	0.80	1	4
	3	0.30	0.70	1	4
	5	0.27	0.73	1	4
10%	1.5	0.23	0.77	2	4
	3	0.27	0.73	2	4
	5	0.23	0.77	2	4

3.2 Restoration Scheduling

Financial allocations for the utilities set the stage for the subsequent critical step: devising an optimal repair sequence for the damaged components. Here, the DQN model emerges as a pivotal tool. The neural networks for each utility was trained with the respective agent group's experience. We collected the agent's

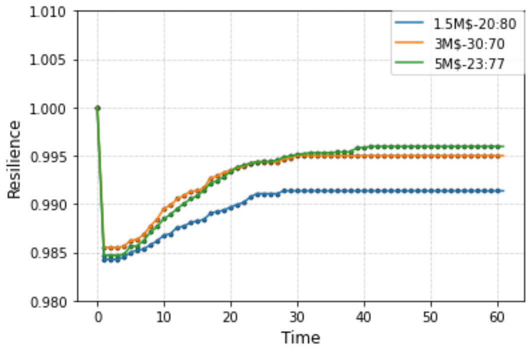

Figure 4: Water network resilience - under 5% failures & different financial allocations.

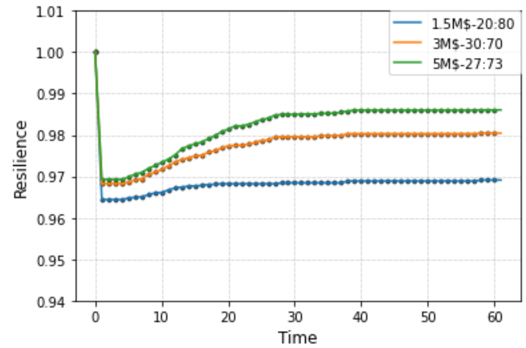

Figure 5: Water network resilience - under 10% failures & different financial allocations.

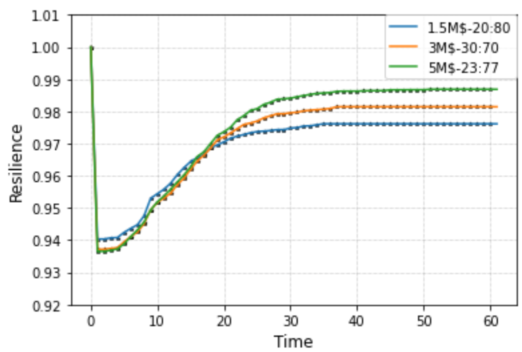

Figure 6: Mobility network resilience - under 5% failures & different financial allocations.

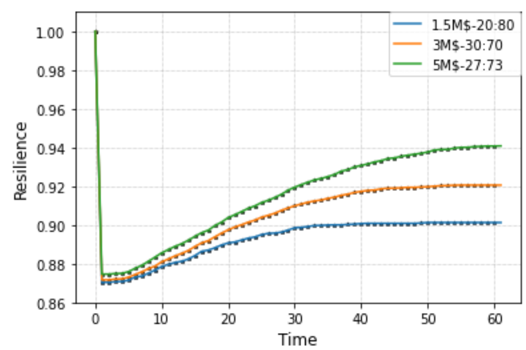

Figure 7: Mobility network resilience - under 10% failures & different financial allocations.

experience by simulating the network with 200 iterations. The loss curves, shown in Figures 8 and 9, demonstrate a significant reduction in the Loss Function value within the initial 100 epochs of training. The loss then exhibited fluctuations from epoch 200 onwards and ultimately stabilized, reaching a significance level of less than 0.01, which indicated the point to halt training. It is important to note that the extensive state space necessitated meticulous hyperparameter tuning and the adoption of optimization algorithms to ensure convergence. Despite the initial rapid decrease, the fluctuations observed between epoch 200 and the convergence point highlight the challenge of navigating the complex dynamics of the loss function in such high-dimensional spaces. The neural network's hyper-parameter settings are shown in the Table. 2.

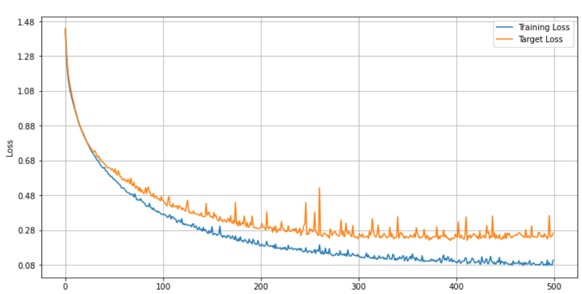

Figure 8: Loss curve - water network.

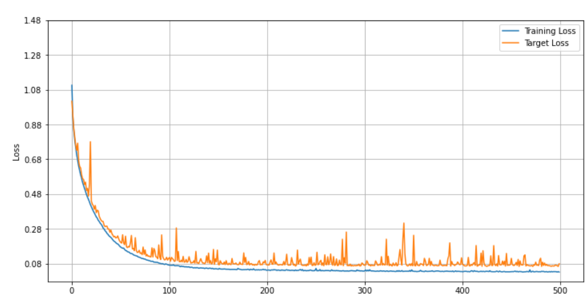

Figure 9: Loss curve - mobility network.

The neural network architecture of the DRL framework was designed to balance complexity and efficiency, with four hidden layers for hierarchical feature extraction and abstraction, ReLU activation for non-linearity and sparsity, and L2 regularization to prevent overfitting. The final set of hyperparameters, including learning rate, number of units in the hidden layers, and dropout rate, was obtained through extensive hyperparameter optimization using a grid search approach. The best combination of hyperparameters was selected based on the lowest mean squared error (MSE) achieved on the validation set, ensuring optimal model performance and generalization. The DQN model offers adaptive decision-making in complex environments, capturing intricate dependencies and trade-offs between restoration strategies and network performance metrics. However, it faces challenges such as computational complexity, extensive hyperparameter tuning, and dependence on diverse training data, which can be time-consuming and resource-intensive when applied to large-scale applications.

To demonstrate the effectiveness of the DRL algorithm, we simulated a scenario with failures comprising of 12 road and 2 pipe components with 1 water and 3 mobility crews. The pipe components had co-located roads that were transmitted as failures after they are restored. Figure 10 shows the resilience of the network under the scenario in which restoration is carried out based on the FIFO policy and Figure 11 shows the network resilience under the strategy provided by DRL. Under no learning and FIFO strategy, the CI network was fully restored by day 36, with initial slow progress in road restoration during the first 10 days. On the other hand, the actions provided by the DRL resulted in complete network restoration at day 30. The restoration strategy optimized by the DRL prioritizes restoring components with larger or longer unit sizes, thereby ensuring a more effective increase in resilience.

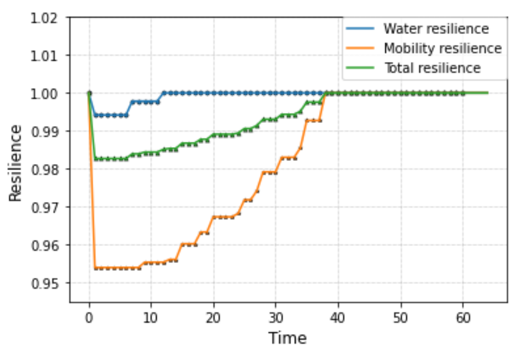

Figure 10: Network resilience over time - FIFO order.

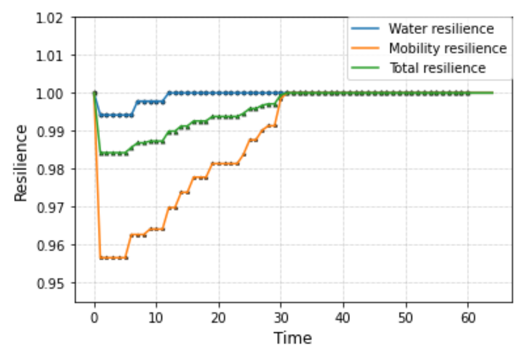

Figure 11: Network resilience over time - order provided by RL.

Table 2: Neural network hyperparameters.

Hyperparameter	**Value**
First Hidden Layer Units	124
Second Hidden Layer Units	64
Third Hidden Layer Units (L2 Regularization)	64
Fourth Hidden Layer Units	32
Output Layer Units	1
Activation Function	ReLU (Hidden), Linear (Output)
Optimizer	SGD
Learning Rate	0.01
Loss Function	Mean Squared Error
Regularization Rate (L2)	0.01
Synchronization Frequency	50
Epochs	500
Batch Size	64

4 CONCLUSIONS

In light of the increased prevalence of intense weather-related events and aging infrastructures, it becomes imperative to establish optimal scheduling and restoration strategies to enhance the resilience and reliability of our critical infrastructure systems. We developed a hybrid simulation model to represent an interdependent infrastructure network and utilized deep reinforcement learning to obtain optimal order of restoration. We defined the resource dynamics of financial sharing between water and mobility utilities in the network. Crew agents in the network travel to the locations of failed components to restore them. At the identified ideal financial ratio obtained from computational experiments, we tested the trained the DQN algorithm to provide restoration sequences. From the results, it was evident that with identified financial ratios and repair sequences, the interdependent network's restoration occurred six days earlier. Thus, providing optimal sequence of restoration significantly improves the performance of the interdependent network.

The hybrid simulation model integrates system dynamics, agent-based modeling, and deep reinforcement learning to capture the interactions between financial decisions and infrastructure network performance. In this study, we demonstrated the ability of the DRL algorithm to provide ideal restoration strategy for the maintenance crews in an interdependent infrastructure network. However, the current model faces scalability challenges when dealing with scheduling larger number of network components. The exponential growth of the state space leads to the curse of dimensionality, slowing down the learning process and hindering convergence to optimal strategies. To address these scalability issues, future research should focus on developing advanced algorithms that can efficiently handle the complexity of large-scale networks. Techniques such as state space reduction, hierarchical learning, and transfer learning could be explored to mitigate the curse of dimensionality and improve learning efficiency. In addition, more advanced strategies that include learning from shared experiences of the maintenance crews should be evaluated to ensure more realistic illustration of decentralized decision-making paradigm.

ACKNOWLEDGMENTS

This work was partially supported by the US National Science Foundation under Grant Number 1638301. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

AnyLogic 2024. "AnyLogic: Simulation Software". https://www.anylogic.com, Version 8.7.5, accessed August 29th, 2024.

Baidya, P. M. and W. Sun. 2017. "Effective Restoration Strategies of Interdependent Power System and Communication Network". *The Journal of Engineering* 2017(13):1760–1764.

Buldyrev, S. V., R. Parshani, G. Paul, H. E. Stanley and S. Havlin. 2010. "Catastrophic Cascade of Failures in Interdependent Networks". *Nature* 464(7291):1025–1028.

Clark, R. M., M. Sivaganesan, A. Selvakumar, and V. Sethi. 2002. "Cost Models for Water Supply Distribution Systems". *Journal of Water Resources Planning and Management* 128(5):312–321.

Dehghani, N. L., A. B. Jeddi, and A. Shafieezadeh. 2021. "Intelligent Hurricane Resilience Enhancement of Power Distribution Systems via Deep Reinforcement Learning". *Applied Energy* 285:116355.

Der Sarkissian, R., J.-M. Cariolet, Y. Diab, and M. Vuillet. 2022. "Investigating the Importance of Critical Infrastructures' Interdependencies During Recovery; Lessons from Hurricane Irma in Saint-Martin's island". *International Journal of Disaster Risk Reduction* 67:102675.

DHS 2003. "Critical Infrastructure Identification, Prioritization, and Protection". https://www.cisa.gov/critical-infrastructure-sectors, accessed 25th March 2024.

Dsouza, M. H. 2022. *Designing Agent-based Simulation to Assess the Impact of Coordination Schemes on Infrastructure Networks Resilience*. Master's thesis, George Mason University, Fairfax, Virginia. https://mars.gmu.edu/items/6aac977b-aae9-4c14-bb61-446d7a03b4f0, accessed 4th September 2024.

Esmalian, A., M. Ramaswamy, K. Rasoulkhani, and A. Mostafavi. 2019. "Agent-based Modeling Framework for Simulation of Societal Impacts of Infrastructure Service Disruptions During Disasters". In *ASCE International Conference on Computing in Civil Engineering 2019*, 16–23. June 17th-19th, Reston, VA.

Feofilovs, M., F. Romagnoli, C. K. Gotangco, J. C. Josol, J. M. P. Jardeleza, J. E. Litam, *et al*. 2020. "Assessing Resilience Against Floods with a System Dynamics Approach: A Comparative Study of Two Models". *International Journal of Disaster Resilience in the Built Environment* 11(5):615–629.

Gallopín, G. C. 2006. "Linkages Between Vulnerability, Resilience, and Adaptive Capacity". *Global Environmental Change* 16(3):293–303.

Gao, J., S. V. Buldyrev, S. Havlin, and H. E. Stanley. 2011. "Robustness of a Network of Networks". *Physical Review Letters* 107(19):195701.

Hayhoe, K., D. J. Wuebbles, D. R. Easterling, D. W. Fahey, S. Doherty, J. P. Kossin, , *et al*. 2018. "Our Changing Climate. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II". https://nca2018.globalchange.gov/chapter/2/, accessed 30$^{\text{th}}$ March 2024.

Kandiah, V. K., E. Z. Berglund, and A. R. Binder. 2019. "An Agent-based Modeling Approach to Project Adoption of Water Reuse and Evaluate Expansion Plans within a Sociotechnical Water Infrastructure System". *Sustainable Cities and Society* 46:101412.

Karamouz, M., M. Taheri, P. Khalili, and X. Chen. 2019. "Building Infrastructure Resilience in Coastal Flood Risk Management". *Journal of Water Resources Planning and Management* 145(4):04019004.

Khanmohammadi, S., H. Farahmand, and H. Kashani. 2018. "A System Dynamics Approach to the Seismic Resilience Enhancement of Hospitals". *International journal of Disaster Risk Reduction* 31:220–233.

Ko, H. J., C. S. Ko, and T. Kim. 2006. "A Hybrid Optimization/Simulation Approach for a Distribution Network Design of 3PLS". *Computers & Industrial Engineering* 50(4):440–449.

Landa, P., R. Aringhieri, P. Soriano, E. Tànfani and A. Testi. 2016. "A Hybrid Optimization Algorithm for Surgeries Scheduling". *Operations Research for Health Care* 8:103–114.

Li, G., C. Kou, Y. Wang, and H. Yang. 2020. "System Dynamics Modelling for Improving Urban Resilience in Beijing, China". *Resources, Conservation and Recycling* 161:104954.

Links, J. M., B. S. Schwartz, S. Lin, N. Kanarek, J. Mitrani-Reiser, Sell *et al*. 2018. "COPEWELL: A Conceptual Framework and System Dynamics Model for Predicting Community Functioning and Resilience After Disasters". *Disaster Medicine and Public Health Preparedness* 12(1):127–137.

Lu, Q., M. Li, M. Gunaratne, C. Xin, M. Hoque, M. Rajalingola *et al*. 2018. "Best Practices for Construction and Repair of Bridge Approaches and Departures". Florida. Dept. of Transportation.

Mohebbi, S., Q. Zhang, E. C. Wells, T. Zhao, H. Nguyen, M. Li *et al*. 2020. "Cyber-Physical-Social Interdependencies and Organizational Resilience: A Review of Water, Transportation, and Cyber Infrastructure Systems and Processes". *Sustainable Cities and Society* 62:102327.

Ouyang, M. 2014. "Review on Modeling and Simulation of Interdependent Critical Infrastructure Systems". *Reliability Engineering & System Safety* 121:43–60.

Pant, R., K. Barker, and C. W. Zobel. 2014. "Static and Dynamic Metrics of Economic Resilience for Interdependent Infrastructure and Industry Sectors". *Reliability Engineering & System Safety* 125:92–102.

Qiu, D., Y. Wang, T. Zhang, M. Sun and G. Strbac. 2022. "Hybrid Multiagent Reinforcement Learning for Electric Vehicle Resilience Control towards a Low-Carbon Transition". *IEEE Transactions on Industrial Informatics* 18(11):8258–8269.

Rahimi-Golkhandan, A., B. Aslani, and S. Mohebbi. 2022. "Predictive Resilience of Interdependent Water and Transportation Infrastructures: A Sociotechnical Approach". *Socio-Economic Planning Sciences* 80:101166.

Rinaldi, S. M., J. P. Peerenboom, and T. K. Kelly. 2001. "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies". *IEEE Control Systems Magazine* 21(6):11–25.

Shi, Y., Y. Qi, L. Lv, and D. Liang. 2021. "A Particle Swarm Optimisation with Linearly Decreasing Weight for Real-time Traffic Signal Control". *Machines* 9(11):280.

Sun, J. and Z. Zhang. 2020. "A Post-disaster Resource Allocation Framework for Improving Resilience of Interdependent Infrastructure Networks". *Transportation Research Part D: Transport and Environment* 85:102455.

USGCRP 2017. "Climate Science Special Report: Fourth National Climate Assessment, Volume I". Technical report, U.S. Global Change Research Program., Washington, D.C.

Wang, Y., D. Qiu, F. Teng, and G. Strbac. 2023. "Towards Microgrid Resilience Enhancement via Mobile Power Sources and Repair Crews: A Multi-Agent Reinforcement Learning Approach". *IEEE Transactions on Power Systems*.

Yang, S., Y. Zhang, X. Lu, W. Guo and H. Miao. 2024. "Multi-Agent Deep Reinforcement Learning Based Decision Support Model for Resilient Community Post-Hazard Recovery". *Reliability Engineering & System Safety* 242:109754.

Yohe, G. and R. S. Tol. 2002. "Indicators for Social and Economic Coping Capacity—Moving Toward a Working Definition of Adaptive Capacity". *Global Environmental Change* 12(1):25–40.

Zhang, J. and T. Wang. 2023. "Urban Resilience under the COVID-19 Pandemic: A Quantitative Assessment Framework Based on System Dynamics". *Cities* 136:104265.

Zhang, P. and S. Peeta. 2011. "A Generalized Modeling Framework to Analyze Interdependencies among Infrastructure Systems". *Transportation Research Part B: Methodological* 45(3):553–579.

AUTHOR BIOGRAPHIES

PAVITHRA SRIPATHANALLUR MURALI is a PhD student in the Department of Systems Engineering and Operations Research at George Mason University. Her research interests include simulation and reinforcement learning with applications in networked systems such as infrastructure systems, supply chain and logistics. Her email address is psripath@gmu.edu

SHIMA MOHEBBI is an assistant professor in the Department of Systems Engineering and Operations Research and affiliate faculty of Computer Science at George Mason University. Her research interests include game theory, network optimization, simulation, and machine learning with applications in resilient infrastructure systems, sustainable water systems, and smart and connected communities. Her website is https://mohebbilab.cec.gmu.edu/ and email address is smohebbi@gmu.edu.