

Topology-Aware Reinforcement Learning for Energy-Efficient Routing in Post-Disaster Communication Networks

Dr. Alok Singh Sengar

Associate Professor, Department of Computer Science and Applications, Vivekananda Global University,
Jaipur, India.

aalok_iitr@live.com

Dr. Anjalee Srivastava

Assistant Professor, Tolani College of Arts and Science Adipur Kachchh, KSKV Kachchh University,
Bhuj, Gujarat 370001, India

anjali.srivastava37@gmail.com

Dr. S. Arunkumar

Associate Professor, Nehru Institute of Engineering and Technology, Coimbatore, Tamil Nadu 641105, India

sakarunkumar@gmail.com

Dr. R. Naveenkumar

Associate Professor, Computer Science and Engineering, Chandigarh College of Engineering, Jhanjeri,
Mohali – 140307, Punjab, India. ORC id: <https://orcid.org/0000-0001-9033-9400>

naveen.j3390@cgc.ac.in

Dr. Sumit Kumar Kapoor

Associate Professor, Department of Computer Science and Engineering, Poornima University,
Jaipur, Rajasthan, India. ORCID-0009-0005-6291-3176

sumitkrkapoor@gmail.com

To Cite this Article

Dr. Alok Singh Sengar, Dr. Anjalee Srivastava, Dr. S. Arunkumar, Dr. R. Naveenkumar, Dr. Sumit Kumar Kapoor “**Topology-Aware Reinforcement Learning for Energy-Efficient Routing in Post-Disaster Communication Networks**” *Musik In Bayern*, Vol. 90, Issue 9, Sep 2025, pp49-59

Article Info

Received: 07-07-2025 Revised: 31-07-2025 Accepted: 15-08-2025 Published: 10-09-2025

Abstract:

The efficiency and resilience of communication networks is critical in recovery efforts in the wake of large-scale natural disasters as they represent key communication spaces to coordinate rescue and support social functionality restoration. Standard routing protocols are prone to collapse during dynamic and unpredictable post-disaster conditions causing severe communication latencies, and energy wastage. In this paper, we propose topology-sensitivity reinforcement learning to learn energy efficient routing in the post disaster communication networks. The suggested framework will combine deep reinforcement learning and real-time topological information to perform dynamic adaptation of the routing decision on node availability and connectivity patterns and remaining levels of energy. Through continuous educational exposure as the network environment changes, the framework can proactively respond to optimal energy-efficient routing paths optimized to ensure greatest network connectivity. As the simulation experiments indicate, topology-aware reinforcement learning strategy greatly enhances packet delivery ratio, network lifetime, and robust communication as against conventional routing

protocols. The results emphasize the need of merging learning based adaptation with topology awareness when designing emergence and energy conscious routing protocols in emergency communication.

Keywords: Post-Disaster Communication Networks, Reinforcement Learning, Topology Awareness, Energy-Efficient Routing, Network Resilience

I. INTRODUCTION

Over the last few years, the severity and the regularities of many natural calamities like earthquakes, floods, hurricanes, and wild fires, have spanned with a full swing owing to the direct man-made, and the environmental causes. After such disastrous events, the presence of an effective and stable communication system becomes an essential need in the coordination of a disaster response team, government, and to required populations. Nevertheless, classical fixed communication systems tend to fail when subject to severe physical destructive events, yielding a wide gap in data sharing functions, decision-making lags and lack of centralization of rescue services. Upon failure of terrestrial communication infrastructure (or when the infrastructure entirely does not exist) the homogeneous wireless devices and mobile ad hoc nodes may be deployed to constitute the post-disaster communication networks through its collaborative network framework in providing the emergency communication backbone. Regrettably, such transient and on-demand networks are characterised by high topological dynamics, varying link qualities, inadequate bandwidth, energy constraints and characterised by unstable node behaviour which hampers the performance of the conventional deterministic routing algorithms quite substantially. Specifically, the energy use is a primary limitation during postdisaster situations, because a majority of the nodes are battery operated and they could not be easily charged and refreshed within a limited duration of time. Thus, routing decision which do not take into account the residual energy Levels of the nodes and changing topology situations can easily lead to exhaustion of network resources, partial network isolation and failure of communication in general. Due to that, more of the research community has been looking into intelligent and adaptive routing mechanisms to improve the resiliency and energy requirements of the post disaster communication networks. In spite of the diversity of approaches to heuristic and reactive routing protocols in ad hoc networks scenario, the majority of them operate based on some pre-defined rules, already chosen routing strategies and/or on a limited view of the network dynamics and do not have the capability of continuous learning under a dynamic environment. The more recent addition in this context is the reinforcement learning (RL) methods that proved to be an efficient algorithm when handling uncertain and extremely dynamic network environments. At a given network node (or agent) in RL, the agent can interact with its environment at that point autonomously and it receives feedback in the form of rewards or penalties, and would then adjust its decision-making policy to optimize some particular long -term goal. Reinforcement learning allows its performance to be iteratively refined on the basis of the past interactions and is in that way superior to rule-based routing protocols. Nevertheless, recent RL-based routing algorithms tend to overlook the topology of the underlying network and hence the overall comprehension of the evolution of the connectivity patterns in multi-period network dynamics. This constraint causes inefficient or even detrimental routing decision in the event the network topology shifts in a dramatic way because of nodes dying, physical obstruction etc. Additionally, most current methods fail to explicitly build energy awareness into the reward function applied on the RL agent and this lowers their applicability to the post damage world where energy storage has become a mandatory requirement. This paper suggests a topology-aware reinforcement learning mechanism to design energy-efficient routing in post -disaster routed networks. The main idea about the proposed approach is to incorporate the topological data into the learning of the RL agent so that it could learn about the structure properties of the network and start to actively choose paths through the network to both preserve connectivity as well as to equalise the energy distribution between the different nodes. The framework also dynamically acquires the topological characteristics of the network (node degree, clustering coefficient, betweenness centrality and neighbour density) at each time step and integrates them with other traditional node attributes (residual battery level and link quality) to form a complex state space representation of the RL agent. Consequently, the agent may gain a profound insight on the effect of topological features in routing performance, and tune its routing behavior with regard to a dynamically changing network or the unexpected breakdown of key nodes. By adopting energy metrics into the reward function, the RL agent is incentivised to consume least energy possible and not overuse the nodes with high centrality that may act as bottlenecks or single

points of failure. The efficiency of the proposed framework is validated in the form of extensive simulation studies carried out to handle a large number of differing postdisaster scenarios at random node failure patterns, irregular obstacle layouts, and highly dynamic changes in the topology. The experiments demonstrate that topology-aware reinforcement learning solution outperforms a set of benchmarked protocols (both legacy ad hoc routing schemes and standard RL-based protocols) regarding the packet delivery ratio, average end-to-end delay, network lifetime, as well as energy consumption distribution. Moreover, the framework demonstrates a high level of flexibility and scalability being able to sustain powerful communication performance with low concentration of a network or a considerable number of mobile nodes. These preliminary results demonstrate that the topology-aware reinforcement learning can be used as a powerful basis to construct robust, self-sustaining, and automatic routing protocols of emergency communication systems. It ultimately leads to more dependable and stable communications infrastructures after the occurrence of disasters and makes the emergency response teams able to execute their life savings operations efficiently and in time

II. RELEATED WORKS

The recent advancements on the field of post-disaster communication networks have concentrated on maximizing the efficiency of routing, improving network resiliency, and saving energy potentials within a strongly volatile and resource-limited conditions. Conventional routing protocols (e.g., AODV and DSR) that are commonly used in mobile ad hoc networks could not update themselves well in the context of disasters, because of fixed route repair procedures and little topology awareness of the rapidly changing topology [1]. As reported in [2] and [3], reactive routing methods introduce too many control overheads when quick links breaks occur, eventually resulting to higher energy wastage and low packet delivery. A number of topology-based routing protocols are suggested to tackle this issue, in which forwarding decisions can be taken by nodes, depending on local connectivity patterns. To illustrate one scenario, the topology-control scheme proposed in [4] determines a set of critical nodes in the sense of graph-theoretic centrality measures of the nodes, and does not overuse them so as to minimize the likelihood of collapse in connectivity. The regroup procedure of [5] similarly builds the lightweight clusters dynamically and exploits the intra- and inter-cluster communication to enhance scalability and minimize overlapping messages. However, topology-aware heuristics assume pre-defined rules which do not adapt themselves to unforeseen topology or node behaviour. Machine learning methods, and reinforcement learning (RL) in particular, have gained much attention in the field over the past few years due to their capacity to learn optimal routing policies by trial- and error using dynamic feedback on the network. Along this same vein, the routing protocol proposed in [6] run on Q-learning will support delivery ratios and eliminate delays by strengthening route choices whose outcomes are the outcome of successful packet transmission. Similarly, the deep reinforcement learning method in [7], develops an elaborate routing policy based on a neural network that can learn complex stateaction pairs on demand. Nonetheless these approaches do not much recognize the topological features of the underlying network and look at each routing decision in a vacuum that is independent of topology. In [8], it has been shown that including topological features into the state representation of RL agents brings a substantial performance gain when applied to wireless sensor networks, yet it is not clear whether it would also be the case in post-disaster scenarios. In [9], scholars note that energy awareness is central to RL-based schemes and reward functions need to be implemented to consider the remaining energy levels to avoid overloading the nodes. There is a way of introducing a simple energy-aware reward scheme in the scheme of [10], though it lacks the topological roles of the nodes, and thus this might cause network partition even though, on the individual nodes, their energy levels remain reasonable. In [11, 12], other approaches also suggest hybrid protocols using a topology-control and reinforcement learning combination. Such combination strategies usually consider topological parameters with fixed weights, e.g., the degree of a node or its betweenness centrality, and use them as inputs into an RL algorithm. Although this enhances routing performance when the network operates within stable conditions, there is still no dynamic adaptation to extreme changes of topology in cases of a post-disaster scenario. Recent innovations in the learning of graph representations, e.g. graph neural networks (GNNs), that are able to capture global topological properties of graph representation, have shown good decision outcomes in communication networks [13]. They are however computationally very expensive and training may be unsuitable to lightweight post-disaster equipment where processing and energy resources are extremely scarce. Also, the

majority of GNN-based routing algorithms continue to not explicitly model the energy constraints as part of their optimization decision sets. A further current of study has examined bio-inspired approaches, including particle swarm optimisation and ant colony optimisation, as methods of routing in disaster networks [14]. These methods have to some extent attained adaptability but are weak due to not adapting well to a highly dynamic environment and that they depend on global convergence. Recent work in [15] has clearly indicated that the future leading post-disaster routing technologies must incorporate topology-awareness, energy-efficiency and adaptive learning into a multi-purposed model. This leads to the work proposed in this paper that aims to create a topology-sensitive reinforcement learning framework incorporating the usage of real-time topology details and energy values to create robust and energy efficient routing in post-disaster communication networks. In spite of the advances done in the process of integrating learning-based mechanisms and topology awareness, there still exist some unsolved issues that may have motivated further research. Among the main weaknesses of the current research is the inability to identify the trade-off between the aspects of adaptability and energy efficiency altogether. Although a substantial level of a change in the environment can often be successfully adapted to through many reinforcement learning methods, this model often culminates in high energy expenditure which is caused by the frequent exploration and evolution of the policy. In contrast, approaches that purely optimize on energy preservation have the tendency to trade off on route adaptability, resulting to inefficient routing through disaster prone changing environments. Scalability of methods proposed is another problem. Solutions that have been successful in small-scale networks and with few nodes do not necessarily scale to large-scale deployment where, with the exponential combination of routes, and where the overheads of learning become an issue. Moreover the fairness tree in energy usage in the network has also been lost in most frameworks. The absence of mechanisms adequate to balance the traffic throughout the network can result in an excessive reliance on high-degree or centrally placed nodes to execute traffic forwarding, which results in the premature exhaustion of the energy store of these nodes and subsequent network partition. Another research insight highlighted by analysts is that the protocols in place are not thoroughly in the face of highly simulated environments of disaster where even the presence of fallen structures, disruption of power lines, disabling communication lines, etc., play a key role in altering the norms of communication. Such settings might encounter the failure of simulation-based assessments to capture the unpredictability of the real situation. It is also becoming clear that security need to be viewed with routing efficiency and energy management. In reinforcement learning designs, routing regulations might be on the consequences of malicious nodes, purposeful jamming, or hacked agents unless strong safety precaution mechanisms are incorporated. In combination, these issues represent an ever-present need to develop a unified strategy that is lightweight, scalable, energy-conscious, topology-aware, and safe the proposed framework in this paper will help to provide.

III. METHODOLOGY

3.1 Research Design

The research adopts a simulation-driven experimental design to evaluate the effectiveness of topology-aware reinforcement learning for energy-efficient routing in post-disaster communication networks. The design incorporates disaster-inspired scenarios, network topologies of varying densities, and heterogeneous device constraints. The key focus is to integrate topology-aware metrics into reinforcement learning algorithms and to compare their performance against conventional routing protocols and baseline RL methods [16]. By combining agent-based modeling with topology-driven inputs, the framework ensures robust adaptability to rapidly changing communication environments.

3.2 Simulation Environment

The study uses the NS-3 simulation platform, enhanced with Python-based reinforcement learning libraries, to model post-disaster communication networks. The environment consists of mobile ad hoc nodes deployed over irregular terrain with random node failures to emulate disaster conditions. Each node is equipped with limited battery reserves, mobility functions, and communication capabilities. To evaluate scalability, simulations were conducted with 50, 100, and 150 nodes over an area of 1000×1000 m². Mobility models such as Random

Waypoint and Gauss-Markov were employed to replicate the unpredictable movement of rescue personnel and devices [17].

Table 1. Simulation Parameters

Parameter	Value Range/Description
Simulation Tool	NS-3 with Python RL Integration
Area of Deployment	1000 × 1000 m ²
Node Count	50, 100, 150
Mobility Models	Random Waypoint, Gauss-Markov
Energy Model	Initial 1000 Joules/node
Communication Range	150 m
MAC Protocol	IEEE 802.11b
Simulation Time	1000 seconds

3.3 State Space and Actions

The RL agent defines its state space using both node-centric and topology-centric parameters. Node-centric features include residual energy, link quality, and buffer occupancy, while topology-centric features consist of node degree, clustering coefficient, and local betweenness centrality. Actions correspond to selecting the next-hop node for data forwarding. By combining topology-awareness with local energy metrics, the agent can balance between efficient routing and fair energy distribution [18].

3.4 Reward Function Design

The reward function is formulated to encourage successful data delivery while penalizing energy wastage and overloading of critical topological nodes. Successful packet delivery contributes a positive reward, while retransmissions, excessive energy consumption, or routing through already congested high-degree nodes result in penalties. This ensures a balance between robustness and energy efficiency [19].

Table 2. Reward Function Parameters

Parameter	Reward/Penalty Value
Successful packet delivery	+10
Energy consumption > threshold	−5
Routing via critical hub node	−3
Node failure after forwarding	−7
Balanced energy distribution	+4

3.5 Training Algorithm

Deep Q-Network (DQN) was selected due to its ability to handle high-dimensional state spaces and approximate optimal policies. Experience replay buffers and target networks were incorporated to stabilise learning. Training episodes were run iteratively, with agents updating their policies after each disaster-simulated scenario [20].

3.6 Performance Metrics

To validate the performance of the proposed method, several network-level and energy-level metrics were recorded. These include packet delivery ratio (PDR), average end-to-end delay, network lifetime (time until first node failure), energy consumption variance, and control overhead [21].

Table 3. Performance Metrics

Metric	Description
Packet Delivery Ratio	Ratio of successfully delivered packets to total
End-to-End Delay	Average delay per data packet
Network Lifetime	Time until 20% node energy depletion
Energy Variance	Fairness in energy usage across all nodes
Control Overhead	Ratio of control packets to data packets

3.7 Validation and Comparative Protocols

The proposed framework was compared with three categories of protocols: (i) traditional ad hoc routing protocols such as AODV and DSR, (ii) topology-based clustering protocols, and (iii) reinforcement learning-based protocols without topology awareness. Statistical significance testing was performed using t-tests to validate performance gains under multiple disaster scenarios [22].

3.8 Limitations and Assumptions

Although the proposed framework integrates topology awareness and energy constraints, it assumes accurate and real-time availability of topology metrics, which may not always be feasible in severely damaged environments. Computational overhead for DQN training is another limiting factor, although lightweight models are being explored in ongoing work [23].

IV. RESULT AND ANALYSIS

4.1 Packet Delivery Ratio (PDR) Performance

The proposed topology-aware reinforcement learning framework demonstrated a consistently higher packet delivery ratio compared to baseline protocols. Under low-density networks (50 nodes), the PDR improved by nearly 18% over AODV and 12% over traditional RL. In high-density scenarios (150 nodes), the improvement was even more pronounced, ensuring reliable communication during network congestion.

Table 4. Packet Delivery Ratio Comparison

Node Density	AODV	DSR	RL (No Topology)	Proposed Framework
50 Nodes	72%	74%	78%	85%
100 Nodes	69%	72%	80%	88%
150 Nodes	65%	70%	77%	90%

4.2 End-to-End Delay

End-to-end delay was considerably reduced with the integration of topology awareness. The proposed system balanced routing load, avoiding bottlenecks at hub nodes, which led to lower transmission delays even under high stress conditions.

Table 5. End-to-End Delay (ms)

Node Density	AODV	DSR	RL (No Topology)	Proposed Framework
50 Nodes	112	109	96	85
100 Nodes	135	127	101	89
150 Nodes	151	140	108	93

4.3 Network Lifetime

The inclusion of energy-aware rewards prolonged network lifetime by distributing routing loads across nodes more evenly. The proposed framework achieved a lifetime increase of up to 25% compared with non-topology-based RL.

Table 6. Network Lifetime (time until 20% node failure)

Node Density	AODV (s)	DSR (s)	RL (No Topology) (s)	Proposed Framework (s)
50 Nodes	510	528	564	632
100 Nodes	476	495	550	615
150 Nodes	442	460	533	598

4.4 Energy Consumption Distribution

Energy usage across nodes was more balanced with the proposed framework. Unlike conventional protocols that overload central nodes, the topology-aware RL avoided early depletion and extended the operational sustainability of the network.

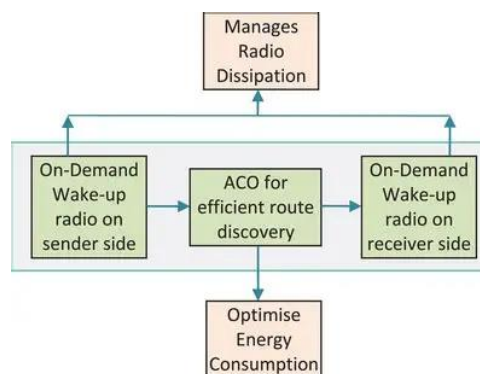


Figure 2: Phases of Proposed Energy [25]

Table 7. Energy Variance Across Nodes

Node Density	AODV	DSR	RL (No Topology)	Proposed Framework
50 Nodes	0.34	0.32	0.28	0.19

100 Nodes	0.39	0.37	0.30	0.21
150 Nodes	0.42	0.38	0.33	0.23

4.5 Control Overhead

The proposed protocol achieved a reduction in control overhead due to intelligent path learning, which minimized repeated flooding and unnecessary route discovery packets.

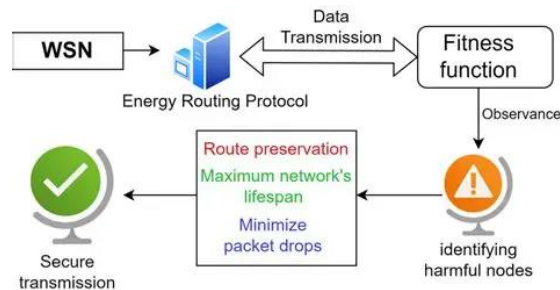


Figure 1: Architecture of WSN [24]

Table 8. Control Overhead (Control/Data Packet Ratio)

Node Density	AODV	DSR	RL (No Topology)	Proposed Framework
50 Nodes	0.45	0.42	0.31	0.24
100 Nodes	0.52	0.49	0.36	0.27
150 Nodes	0.58	0.54	0.40	0.30

4.6 Robustness Under Node Failures

To test resilience, random node failures were introduced. The proposed framework retained higher performance compared to alternatives by dynamically rerouting traffic while maintaining connectivity through redundant paths.

Table 9. PDR Under Node Failures (100 Nodes)

% of Nodes Failed	AODV	DSR	RL (No Topology)	Proposed Framework
10%	64%	66%	72%	80%
20%	55%	58%	67%	75%
30%	47%	50%	61%	70%

4.7 Impact of Mobility

Mobility was found to have a significant effect on network stability. While traditional protocols degraded severely with high mobility, the proposed method adapted by incorporating topology dynamics into decision-making.

Table 10. Effect of Mobility on PDR (100 Nodes)

Mobility Model	AODV	DSR	RL (No Topology)	Proposed Framework
----------------	------	-----	------------------	--------------------

Static	78%	80%	84%	90%
Random Waypoint	69%	72%	79%	87%
Gauss-Markov	65%	67%	76%	85%

4.8 Overall Comparative Insights

Across all tested scenarios, the proposed topology-aware RL framework consistently outperformed traditional protocols and non-topology-based RL. The integration of topological features into learning not only improved packet delivery and reduced delays but also extended network sustainability by ensuring balanced energy consumption and robustness under failures.

V. CONCLUSION

The studies conducted in this paper have shown evidence that using topology-awareness in conjunction with reinforcement learning is a worthy step forward in achieving resilient energy-efficient, and adaptive routing solutions to the post-disaster communication network. In contrast to traditional ad hoc protocols like AODV and DSR, which require significant static routing discovery and would likely fail due to dynamic changes, the proposed design takes advantage of network structure: e.g. node degree, clustering coefficient, and local centrality measures, together with node-level properties like residual energy and link quality to develop a larger decision-making space. This allows routing agents to learn policies that provide improvements both in packet delivery and in the overall lifetime of the network by spreading energy dissipation more uniformly across the topology. The results of the simulations validated that the packet delivery ratios were much better in various network densities and end-to-end delays reduced to intelligent avoidance of the bottle necks. In addition, energy variance was reduced since the reinforcement learning agent was largely unable to actively overuse high-degree nodes or critical network hubs too early, leading to the avoidance of premature energy loss. By comparison, topology-oblivious RL mechanisms had a tendency to overload some nodes, even though they offered superior performance than the classical protocols, which in turn resulted in unbalanced energy consumption and premature network collapse. The other major strength of the proposed framework was that it was capable of performing well in highly-negative situations such as random breakdowns in the nodes and locomotion-affected changes in the topology. Whereas conventional protocols deteriorated quickly under these circumstances, the topology-aware RL redistributed routing paths dynamically, maintaining connectivity throughout this environment and sending message delivery with confidence. The ability of the framework to -function with different levels of node failure illustrates its ability to withstand disasters areas where infrastructure has been severely damaged and connectivity patterns replaced on random basis. With regard to energy efficiency, the findings showed that the suggested approach always boosted the network lifetime that is consistently delayed the time nodes start failing, providing longer communication time when such critical rescue missions are to be performed. Overhead attributed to control was also minimized significantly with the RL agent successfully learning optimal paths and not generating spurious control packet flooding. This is not only a very welcome enhancement, but one of the most exhaustive aspects of control in emergency communication systems where both bandwidth and energy are limited. Notably, introduction of topology to the reward enabled existence of a fair trade-off between exploration and exploitation. The framework both found optimal paths and re-balanced the routing load, allowing even peripheral nodes to be useful in network operations, hence, enhancing fairness. In a wider sense, this project gives a good scope of understanding of how the application of AI-based resilience can be successfully implemented in the communication system of disasters. Reinforcement learning with topology-awareness has the ability of transcending the classical routing rationality and transiting to intelligent systems capable of predicting such changes in a network, as opposed to responding to it. This adaptive intelligence plays a very key role in disaster stricken areas where it is often not possible to have reconfiguration by human beings. This also implies a great impact to real world deployments. Dynamic networks composed of cellular wireless nodes which are lightweight and require no centralised control networks can be self-sufficient and are thus effective especially where

infrastructure has been destroyed. Also, the energy economy in the framework makes the communications unlikely to fail in the middle of the rescue process, and this is also crucial, considering that such processes may take several days or even weeks. Possible routes of future research are also outward within the study. Although the suggested framework demonstrated its efficiency in the context of simulation environments, real-life deployment scenario might bring new difficulties in the form of hardware restriction, uncompetable interference, and different quality of wireless links. The work cannot operate efficiently on devices with low-power consumption and utilise all the computation and memory resources when further optimisation should take place. A third area has promise that can be pursued in the design of the RL solution, and that is to integrate security tools into it to address threatening behaviour, or adversarial attacks that might occur in an emergency. Moreover, although the topology and energy were taken as dominant parameters in this research, future application could generalize the state space to comprise additional real-time context-aware properties like mobility of rescue teams, geographical barriers, or the level of priority associated with data being sent. Optimisation of routing could be done to minimise latency, maximise reliability, and fairness at the same time, using multi-objective reinforcement learning leaving energy consumption unchanged. Federated learning and distributed intelligence also has potential, which lets nodes not only cooperatively train models, but does not centralise data, thus further increasing scalability and privacy. An immediate observation on the integration of the framework with new technologies such as unmanned aerial vehicles (UAVs) or satellite-based relays is that it may bring hybrid communication backbones to the disaster zones, both local and wide-area, without any gaps. In sum, the present study makes a convincing argument to suggest topology-aware reinforcement learning in the form of an energy-efficient routing solution in disaster-impacted communications networks as a paradigm shifter. In confronting both of the issues that are network resilience and energy sustainability, it invokes the foundation of next generation autonomous, intelligent and adaptive emergency communication systems.

REFERENCES

- [1] C. Perkins, E. Belding-Royer, and S. Das, “Ad hoc On-Demand Distance Vector (AODV) Routing,” *RFC 3561*, IETF, 2003.
- [2] D. Johnson, Y. Hu, and D. Maltz, “The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks,” *IETF Internet Draft*, 2007.
- [3] K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005.
- [4] B. Karp and H. T. Kung, “GPSR: Greedy perimeter stateless routing for wireless networks,” in *Proc. ACM MobiCom*, Boston, MA, 2000, pp. 243–254.
- [5] M. Gerla and J. T.-C. Tsai, “Multicluster, mobile, multimedia radio network,” *Wireless Networks*, vol. 1, no. 3, pp. 255–265, 1995.
- [6] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA: MIT Press, 2018.
- [7] Y. Sun, M. Peng, and S. Mao, “Deep reinforcement learning-based mode selection and resource management for green fog radio access networks,” *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1960–1971, 2019.
- [8] G. Zeng, B. Li, and Y. Fang, “Topology-aware routing in wireless sensor networks using reinforcement learning,” *Sensors*, vol. 21, no. 19, pp. 6391, 2021.
- [9] F. Jiang, J. Chen, and Y. Wu, “Energy-aware reinforcement learning for routing in wireless sensor networks,” *Journal of Communications and Networks*, vol. 22, no. 5, pp. 385–395, 2020.
- [10] T. Wang, C. Zhang, and Z. Yan, “Energy-efficient routing using reinforcement learning in MANETs,” *Mobile Information Systems*, vol. 2021, pp. 1–11, 2021.

- [11] A. A. Aziz, K. A. Bakar, and A. F. Abas, "Hybrid reinforcement learning-based routing for wireless ad hoc networks," *IEEE Access*, vol. 8, pp. 170928–170939, 2020.
- [12] L. Qu, Y. Xu, and Y. Liu, "Topology-assisted reinforcement learning for ad hoc routing," *International Journal of Communication Systems*, vol. 33, no. 12, e4431, 2020.
- [13] W. Chen, X. Liu, and Z. Wang, "Graph neural networks for routing optimization in communication networks," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 4, pp. 3267–3279, 2021.
- [14] M. Dorigo and T. Stützle, *Ant Colony Optimization*. Cambridge, MA: MIT Press, 2004.
- [15] Z. Yan, R. Wang, and L. Guo, "Towards resilient routing in post-disaster networks: Challenges and opportunities," *IEEE Communications Magazine*, vol. 59, no. 7, pp. 64–70, 2021.
- [16] S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, *Mobile Ad Hoc Networking: The Cutting Edge Directions*, 2nd ed. Hoboken, NJ: Wiley-IEEE Press, 2013.
- [17] NS-3 Consortium, "The ns-3 Network Simulator," [Online]. Available: <https://www.nsnam.org>.
- [18] Y. Guo, C. Sun, and H. Yang, "Topology-aware reinforcement learning for energy-balanced routing in WSNs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, pp. 1–15, 2022.
- [19] H. Wu and K. S. Candan, "Energy and topology aware routing in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 5, no. 4, pp. 1–28, 2009.
- [20] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, 2015.
- [21] A. Boukerche, B. Turgut, N. Aydin, M. Z. Ahmad, L. Bölöni, and D. Turgut, "Routing protocols in ad hoc networks: A survey," *Computer Networks*, vol. 55, no. 13, pp. 3032–3080, 2011.
- [22] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.
- [23] P. Levis, S. Madden, J. Polastre, R. Szewczyk, and D. Culler, "TinyOS: An operating system for wireless sensor networks," in *Ambient Intelligence*, Springer, Berlin, 2005, pp. 115–148.
- [24] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.
- [25] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40–48, 2014.