Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/OJIM.2022.1234567

Combining Software-Defined and Delay-Tolerant Networking Concepts with Deep Reinforcement Learning Technology to Enhance Vehicular Networks

OLIVIA NAKAYIMA*, MOSTAFA I. SOLIMAN*, KAZUNORI UEDA‡ AND SAMIR A. ELSAGHEER MOHAMED*

[1]Department of Computer Science and Engineering, Egypt–Japan University of Science and Technology, New Borg El-Arab City, 21934, Egypt

[2]Department of Electrical Engineering, Aswan University, Aswan, 81542, Egypt

[3]Department of Computer Science and Engineering, Waseda University 3-4-1, Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

CORRESPONDING AUTHOR: OLIVIA NAKAYIMA (e-mail: olivia.nakayima@ejust.edu.eg).

ABSTRACT Ensuring reliable data transmission in all Vehicular Ad-hoc Network (VANET) segments is paramount in modern vehicular communications. Vehicular operations face unpredictable network conditions which affect routing protocol adaptiveness. Several solutions have addressed those challenges, but each has noted shortcomings. This work proposes a centralised-controller multi-agent (CCMA) algorithm based on Software-Defined Networking (SDN) and Delay-Tolerant Networking (DTN) principles, to enhance VANET performance using Reinforcement Learning (RL). This algorithm is trained and validated with a simulation environment modelling the network nodes, routing protocols and buffer schedules. It optimally deploys DTN routing protocols (Spray and Wait, Epidemic, and PRoPHETv2) and buffer schedules (Random, Defer, Earliest Deadline First, First In First Out, Large/smallest bundle first) based on network state information (that is; traffic pattern, buffer size variance, node and link uptime, bundle Time To Live (TTL), link loss and capacity). These are implemented in three environment types; Advanced Technological Regions, Limited Resource Regions and Opportunistic Communication Regions. The study assesses the performance of the multi-protocol approach using metrics: TTL, buffer management,link quality, delivery ratio, Latency and overhead scores for optimal network performance. Comparative analysis with single-protocol VANETs (simulated using the Opportunistic Network Environment (ONE)), demonstrate an improved performance of the proposed algorithm in all VANET scenarios.

INDEX TERMS Delay-Tolerant networks, Performance Analysis, Reinforcement Learning, Simulator, Software-Defined networking, Vehicular Ad-hoc Networks.

I. INTRODUCTION

IN the ever-evolving landscape of modern vehicular operations, the effectiveness of communication networks is paramount. Modern vehicular communications are organised as a wireless network that enables communication among vehicles (V2V) and infrastructure (V2I). An active Vehicular Ad-hoc Network (VANET) is one that can dynamically adjust its topology and routing based on the road traffic conditions, network scalability, vehicle mobility and network quality [1], [2], yet they are expected to have maximum efficiency in data transmission. Due to the critical nature of their communications, where any disruption could mean the difference between life and death, these networks have to be designed with fault-tolerance in mind. Most of these networks utilise traditional network protocols in transmitting messages across the entire VANET, such as the connection-oriented Transmission Control Protocol (TCP) and others utilise their known routing protocols specifically designed to operate in VANETs, like Topology based routing, Position based routing, Geo cast routing, Broadcast routing and Cluster based routing protocols [3], [4]. V2V is specifically based on transmission based protocols (unicast, multicast

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

and broadcast) and information based protocols(topology and position). These protocols can be classified as proactive, reactive or hybrid. Alternatively they are classified as being power aware and predictive mobility routing protocols. However, the functionality and performance of these protocols falls apart when they are exposed to high-latency or highly-disrupted network conditions. Since developing new VANET protocols has proven to been hard [5], work has been done to overcome these challenges with the incorporation of Delay-Tolerant Networking (DTN) [6], Software-Defined Networking (SDN) [7] and Machine Learning technologies [8]. These technologies utilise various DTN protocols like Spray and Wait [9], [10], Probability Based Spray and Wait (PBSW) [11], Epidemic [12], Probability Routing Protocol using History of Encounters and Transitivity (PRoPHET) [13], [14], PRoPHETv2 [15], [16], MaxProp [17] and Rapid Adaptive Routing Protocol for Intermittently Connected Delay Tolerant Networks (RAPID) [18] in an ad-hoc manner in order to reliably transmit information from source to destination. However, these protocols also have limitations and thus only work very well in specific use cases such as if the movement of the communicating nodes is predictable or if the network resources required to buffer the transmitted data are sufficient [19]. When it comes to Machine Learning, techniques like Supervised, Semi-supervised, Unsupervised [20] and Reinforcement Learning (RL) [21], [22], [23], have been put to great usage in solving the vehicular challenges as well.

VANETs require fast data processing, which is ideal for limited resource zones and infrastructure-limited settings. DTN technologies are useful in these cases since they perform better in environments with sporadic connectivity and disruptions than their more connection-oriented counterparts. Routing efficiency and network management of DTNs is further improved with the integration of SDN [6] which introduces centralized control, flexibility and security by separating the control plane from the forwarding plane. Optimization problems such as these can be solved and automated with the use of an RL agent, since it can adapt, in real-time, to changes in DTNs and SDNs by adjusting network parameters needed to improve performance, which is more than what can be said for other machine learning approaches.

This research proposes an integration of all these paradigms, which is an approach to optimize vehicular communication networks, thereby bolstering their resilience, efficiency, and adaptability. For this approach to be examined appropriately, simulations are required to understand whether positive results can be achieved when multiple DTN protocols are used in combination (in a heterogeneous manner), as opposed to each working alone, as in homogeneous networks currently. Unfortunately, the most prevalent simulation tools for this, such as the Opportunistic Network Environment (ONE) simulator, only allow homogeneous network simulation scenarios.

The primary objectives of this research are;

- To develop a robust Centralised Controller Multi-agent (CCMA) RL-based algorithm as the SDN controller in the SDN-DTN combined architecture to optimize the VANETs.
- To develop a simulation environment which mimics the behaviour of network nodes, communication dynamics and varying network conditions present in DTN-based VANET environments. The simulator's capability to mimic various DTN protocols, including Spray and Wait, Epidemic and PRoPHETv2, as well as buffer scheduling strategies such as Defer, Earliest Deadline First (EDF), First in First Out (FIFO), Random and Largest/Smallest Bundle First (L/SBF), provides a comprehensive evaluation platform for the proposed algorithm.
- To compare the performance of both the heterogeneous (multi-protocol) and the homogeneous (single-protocol) approaches.

To evaluate the model's performance in optimizing heterogeneous SDN-DTNs, the research considers Time To Live (TTL), link quality, buffer management, delivery ratio, latency and overhead as the metrics under study for their significant influence on VANET efficiency. This research firstly presents results on the model's learning performance, and the optimisation performance of the TTL, link quality and buffer management network metrics. The model's metric optimization performance is then compared to the performance observed when individual DTN protocols optimize homogeneous DTNs, as reported by simulation results from the famous state of art ONE, as detailed in the methodology. For this comparison, the TTL, delivery ratio, latency and overhead network metrics are considered, as they are the ones reported by both the ONE and the novel simulator in their simulation results. All these comparisons are performed in three different VANET environments:

- Limited Resource regions; those characterised by little to no resources.
- Opportunistic Communication Regions which mostly depict the DTN environments.
- Advanced Technological Regions, which are characterised by the sufficient resources.

All these environments are further discussed in the methodology. The core of the proposed algorithm lies in its ability to make informed decisions based on network state information from the simulator. The embedded parameters in these states include link availability and quality, data exchange and differences in buffer sizes. The algorithm employs RL technique to dynamically deploy DTN protocols and buffer schedules that are tailored to the prevailing network conditions. This dynamic decision-making process ensures that the communication strategy adapts to changing circumstances and optimally utilizes available resources.

The fusion of VANETs with the heterogeneous SDN-DTN architecture, operated by a CCMA RL-based model, represents a pioneering approach to enhancing the resilience, efficiency and adaptability of vehicular communication networks. By leveraging the simulation environment, the proposed algorithm, DTN routing protocols and buffer scheduling strategies, this research strives to contribute not only to the field of vehicular communication but also to the broader domain of dynamic network optimization. As subsequent sections delve deeper into the key concept descriptions of the technologies involved, related works performed, the nuances of the proposed approach, discussion and conclusion of this work, a comprehensive understanding of its potential to revolutionize vehicular communication networks unfolds.

The remaining paper is organized as follows: Section II discusses VANET and the key technologies used along side it. Section III discusses all the approaches utilised to optimise VANETs. Section IV presents the proposed algorithm and the state of art ONE. Section V presents key results and highlights obtained from this investigation. Section VI gives concluding remarks.

II. BACKGROUND

A. VANETs

VANETs present a paradigm shift in network design, advocating for data processing and decision-making to occur since they are characterised by environment with no physical infrastructure, as in Fig. 1. Within the VANETs, each vehicle is equipped with an Onboard Unit (OBU) for exchange of messages [24], which functions similarly to a vehicle computer but includes additional features to support VANET services . On the infrastructure side, Roadside Units (RSUs) are deployed along the roadside to form a network [25]. VANETs applications are greatly observed in collision avoidance, traffic systems, traffic violation, toll collection, geo location services, informatics [26]. In V2V, the OBUs communicate to each other either directly or through the RSU using several communication standards(IEEE 802.11p, Dedicated Short Range Communications (DSRC), cellular-Vehicle-to-Everything(C-V2X) and European Telecommunications Standards Institute (ETSI)) [27]. All these are key requirements for real time transmission Like in the traditional networks [1], some of the characteristics of VANETs include; mobility, intermittent connectivity, limited device resources, peer-to-peer network architecture, dynamic network topology, limited communication range and low data transfer rates.

Understanding these characteristics is essential when designing, deploying and managing VANETs to meet the specific requirements of mission critical operations. To address the challenges faced in the VANETs, dynamic routing protocols and Quality of Service (QoS) requirements for the data types are critical to sustain communication in a VANET.

This approach aims to reduce latency, enhance responsiveness and enable decision-making using multiple routing

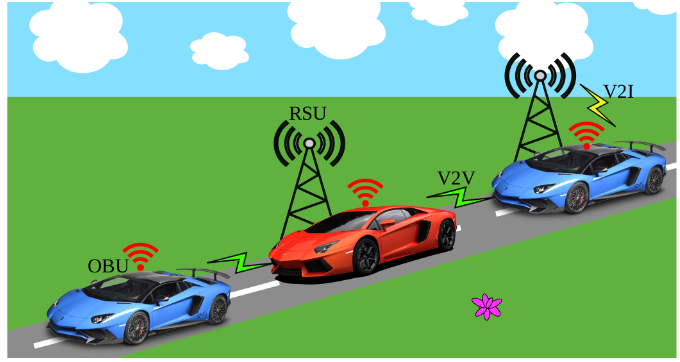

FIGURE 1. **Dynamic vehicular Ad-hoc Network.**

protocols that are suitable in different environments. In the context of vehicular communication networks, this concept gains significance due to the criticality of timely information dissemination, particularly in tactical scenarios where split-second decisions can shape outcomes.

B. DTNs

DTNs employ a store-carry-and-forward data transmission strategy to improve communication performance in settings characterized by limited network resources, frequent or prolonged connectivity disruptions, variable data delivery times, and peer-to-peer communications. These characteristics present themselves in vehicular operations, disaster-stricken zones, underwater networks, and satellite-based space networks [28]. DTNs operate on opportunistic routing protocols to select feasible paths (when available) over which to transmit network traffic. This transmission is carried out in a hop-by-hop fashion, ensuring minimal data loss due to the DTN's intermittency. Some of these opportunistic DTN routing protocols include; Spray and Wait, PBSW, Epidemic, PRoPHET, PRoPHETv2, MaxProp and RAPID. For better deployments, DTNs are firstly simulated to have the environments tested prior and some of the commonly used simulators are the ONE [29], Network Simulator-3 (ns3) [30], Contact Simulator, Epidemic Routing Simulator (EpiSim), SimBet, PROTON and DTN-NS3.

Due to the observed challenges in DTNs and their ad-hoc nature, there arises a necessity to centrally administer them despite said nature. This requirement has led to the introduction of SDN, which primarily aims to centralise control of routing processes with greater flexibility, without sacrificing mobility. Consequently, the integration of these two architectures emerges as a superior solution, as it enhances overall performance in comparison to their separate implementations.

C. SDNs

SDNs, as illustrated in Fig. 2, separates the control plane from the forwarding plane, enhancing network flexibility by centralizing control and management functions within a controller, while the forwarding devices solely focus on

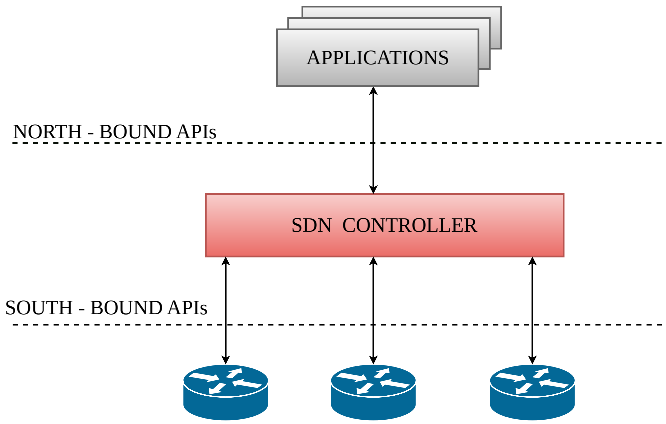

FIGURE 2. SDN Architecture

data forwarding [31], [32]. This methodology introduces security by virtue of the centralized control it introduces to the network. SDN is based on the following key aspects; control plane versus forwarding plane, decoupling control and forwarding, centralized control, flexibility and malleability, security benefits, traffic optimization and scalability [33]. In order to simulate SDN environments, researchers make use of controllers like POX, RYU, TREMA, FloodLight and Open DayLight [34] and simulators like Mininet, Estinet and PlanetLab [35], [36].

D. Reinforcement Learning

The core components of RL include the agent, environment, actions, states, rewards, and policies. The agent's objective is to find an optimal policy, a mapping from states to actions, that maximizes its long-term expected reward.

By integrating VANETs with a reinforced SDN-DTN architecture, this research extends the benefits of localized data processing to the realm of centralized control and dynamic routing.

The effectiveness of the integrated approach is evaluated through a comprehensive set of performance metrics. These metrics include TTL, which measures the time a message remains viable in the network, buffer management efficiency, which assesses the utilization of node buffer resources, and link quality, which indicates the reliability of communication links , delivery ratio [37], [38], latency and Overhead [44]. Through the systematic analysis of these metrics, the research assesses the extent to which the proposed approach optimizes vehicular communication networks.

III. RELATED WORK

In this section, a comprehensive review that delves into relevant works that revolve around the integration of SDN, DTN and RL, aiming to transform the realm of vehicular communication networks have been highlighted.

VANETs, with their limited wireless capacity and susceptibility to disruptions, raise the possibility of information replication which could enhance redundancy and accessibility, yet introduces security concerns by broadening potential attack points. Duplicating information diverts resources from other tasks and sparks information availability, security and resource allocation trade-offs. This research emphasized a need to formulate measures that handle the trade-offs to ensure that networks can securely overcome the resource-constrained scenarios. Pratima et al proposed multi-layered Vehicular-Internet of Things framework that enhanced intelligence, security, reliability and efficiency in VANET environments [39].

Fahad et al introduced an SDN based self-organizing map for the 5G based VANET to improve on its security considering the Distributed Denial of Service (DDOS) [40]. Various surveillance tools, such as Unmanned Aerial Vehicles (UAVs) have also been put to play [41]. However, these UAVs are confronted with security concerns, primarily because of their high mobility, resulting in inconsistent communication. Such challenges culminate in compromised and unreliable communication, characterized by buffer overflow and DDOS attacks.

Since VANETs are mostly affected by high vehicle mobility, researchers in [42] introduced a DTN based routing protocol which was used to minimise energy consumption especially in low density VANETs where it is practically hard to reach other vehicles, thus maximising delivery ratio and overhead. Researchers in [43] proposed a reinforcement based routing protocol that enhanced energy consumption in VANETS and also greately reduced overhead

The integration of SDN and DTN has been individually examined to address issues within VANETs. A novel approach was investigated that combines these architectures (SDN-DTN). Zacarias et al delved into melding SDN and DTN within a battlefield setting, emulating the tactical networks using UAVs, a subcategory of Vehicular communication [6]. These UAVs were used to expedite the transmission of data, especially video and audio, tackling the persistent problem of sluggish information exchange in combat zones. The VANET set up utilized SDN controllers arranged in a master-slave hierarchy to oversee and guide the UAV-gathered data to its end location. This methodology sought to refine data pathways and boost the speed of information relay in demanding battlefield situations using a traditional (Internet Protocol) IP-based DTN routing mechanism. Nonetheless, a prominent issue arose concerning the modification of video and audio resolutions in line with fluctuating network conditions. Given the frequently changing and unpredictable networking conditions on battlefields, tailoring the resolutions of audio-visual content became essential to guarantee consistent and effective data delivery.

In the intricate landscape of modern vehicular communication networks, having highlighted the relevant works undertaken like DTN based traditional routing protocols, SDN driven techniques and ML based approaches, the need for agile, adaptive, and efficient methodologies to combat the SDN traditional routing protocols, homogeneous protocol

This article has been accepted for publication in IEEE Open Journal of Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJVT.2024.3396637

architectures as exhibited in DTN architectures, has never been greater.

IV. METHODOLOGY

A. Proposed Approach

The main idea of this study is to borrow architectural and functional characteristics of SDNs to enhance VANET performance using a heterogeneous DTN-protocol deployment strategy effected by an RL model. This study employs an RL model, emulating a centralised SDN controller's functionality, which dynamically assigns DTN routing protocols and buffer schedules to each node participating in a VANET, based on the VANET's state information transmitted to it from the nodes via out-of-band back-haul communication infrastructure, such as the V2I Internet of Vehicles (IoV) infrastructure [44]. The back-haul network could be using Optical-Fibre, Ethernet, LTE or other high-capacity connection-oriented means.

The VANET state information is received when each of the VANET nodes forwards their state information (connected neighbours, link availability, link quality, node buffer utilisation, and data exchanged) to the RL agent (controller).

The RL agent in Fig 5 is a dedicated computing component, reachable via Vehicle-to-Infrastructure (V2I) connectivity between the RSUs and the VANET nodes (vehicles). This connectivity facilitates the state and action information exchange between the RL agent and the VANET nodes.

After the state information is received by the controller, as an output, it will reward each node with the expected protocol and buffer schedule as per environment. The selected protocol will help improve on the delivery ratio, TTL, Latency and Overhead performance so as to leverage the exchange of information in the VANET. When the environment states fluctuates, the new states will be sent to the controller and different protocols will be fed back to the VANET.

To train and evaluate this RL model, it is exposed to a comprehensive simulation environment that emulates the behavior of vehicular communication networks discussed in Subsection D, ensuring a controlled, yet realistic, evaluation of the integrated SDN-DTN approach. The summarised setup of the entire proposed approach is as in Fig. 3.

The proposed solution assumes constant connectivity between the RL agent and the VANET nodes. However, in the event of disconnection, the RL agent only acts on the state information reported by connected nodes and optimises the sub-VANET associated with them while others recover.

B. RL Model Design

Constructed within the environment-agent framework, as in Fig. 3 and Fig. 4, an advanced RL agent takes on the pivotal role of a decision-maker (similar to that of an SDN controller), orchestrating optimal routing protocol and buffer schedule choices based on observed network conditions.

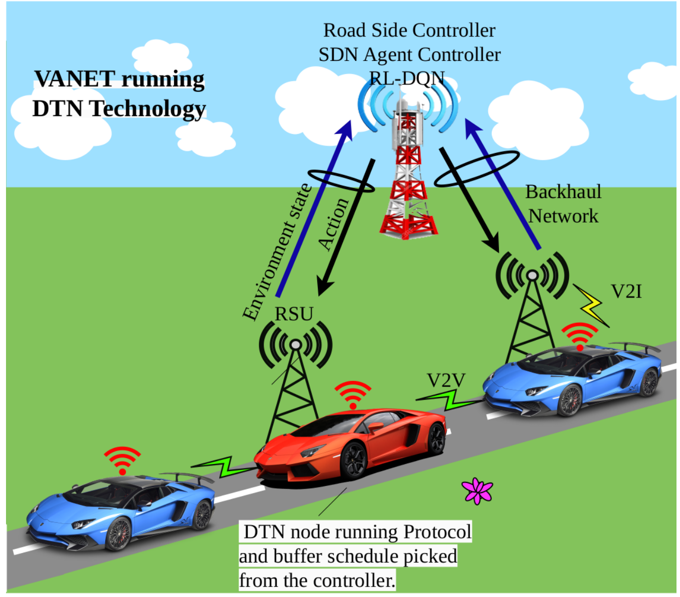

FIGURE 3. Proposed solution deployment architecture: An RL agent, acting as an SDN controller for the VANET, is deployed at the core of the VANET infrastructure. It obtains VANET state information from the VANET nodes (vehicles) and it assigns them a DTN routing protocol and buffer scheduling algorithm as a response action to optimise the VANET, both of which are transmitted over an out-of-band back-haul network. The details of the data exchanged in the Environment States and RL Actions are as shown in Fig. 3

The designed RL model adopts a deep Q-learning method supplemented by experience replay. Within this framework, a Deep Neural Network is used to ascertain Q-values from a defined batch of network states. Action selections are guided by the epsilon-greedy approach, where actions start randomly but become progressively more deterministic [45]. This form of using neural networks is known as Deep-Q Networks (DQN). The structured architecture developed in this study comprises:

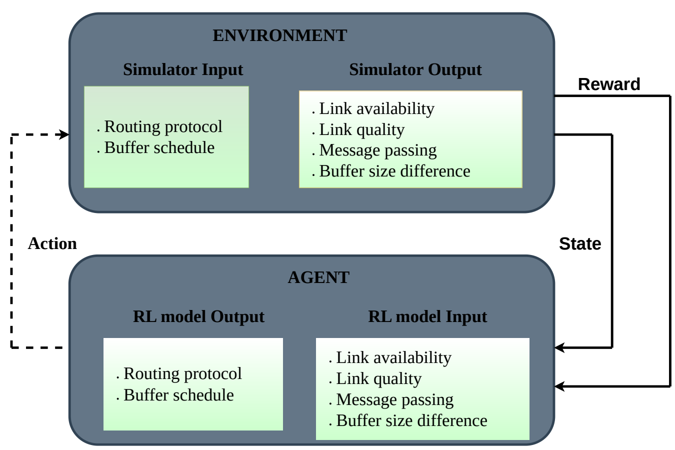

FIGURE 4. Detailed Environment-Agent data exchange during RL agent training and evaluation

1) Flattening layer: Transforms multidimensional data into a linear array of input elements.
2) Four sequential fully connected Layers each with 16 neurons and followed by a Rectified Linear Unit (ReLU) activation function to prune insignificant features.
3) Dropout layer: Functions as a regularizer by sporadically excluding features from prior layers.
4) Three Subsequent fully connected layers: Each trailed by a ReLU activation function. This structure is duplicated to produce outputs for every node's routing protocol and buffer management strategy. For instance, with a network state reflecting 8 nodes, 16 such layer clusters would exist: half for determining routing protocols for each node and the other half for buffer management strategies.
5) Softmax layers: These layers create a probability distribution for each output, facilitating the selection of the most suited routing protocol and buffer management strategy for each node.
6) Concatenation layer: Merges all output distributions into an array, which is then relayed back to the environment. This architecture is further illustrated in Fig. 5:

The agent's state space, constituting a fundamental component of the RL process, is composed of critical attributes such as the quality of communication links, link capacities, messages passed over the network and disparities in buffer sizes between connected nodes. This information is encapsulated in an $L \times N \times N$ multi-layered adjacency matrix, where N is the number of nodes in the network and L is the number of layers, each corresponding to one of the attributes mentioned previously. These multidimensional states offer a comprehensive snapshot of the network's status, serving as the foundation for the agent's decision-making process. The agent's action space is designed to enable the agent to select from the DTN routing protocols and buffer scheduling algorithms implemented in the simulator. These actions are encapsulated as an array of probability distributions for selecting the most appropriate routing protocol and buffer schedule for each node in the network. This enables the agent to adapt to varying communication scenarios with granularity and precision, by assigning each node its own routing protocol and buffer schedule to execute, to optimise the network through network heterogeneity. Crucially, the RL agent's decision-making prowess is galvanized by a well-defined reward signal. The choice of protocol and buffer schedule is tantamount to the agent's strategic move, which inherently determines the immediate reward, since these parameters have significant bearing on the efficiency of data transmission, retention, and delivery within the network. Thus, the reward signal is calculated based on parameters determined to be crucial to the efficiency of a DTN-like environment, like a VANET, and determined to be influenced by or compensated for with the (routing protocol, buffer schedule) combinations assigned to each node by the RL agent. These parameters are:

1) TTL score: TTL refers to the number of intermediary nodes that a data bundle is left to traverse within a network before it expires. TTL is a crucial mechanism in network communication to prevent data packets from circulating indefinitely in cases of network anomalies or failures and ensuring timely delivery of data, thereby optimizing network resources [46]. In the simulator, the TTL component of the reward signal is expressed as a cumulative average of the TTL values of all delivered data bundles from the start of a simulation. The simulator tracks two of its components defined by (1) and (2):

$$N_i = N_{i-1} + n \tag{1}$$

$$A_i = \left\{ \begin{array}{ll} \frac{[A_{i-1} * N_{i-1}] + \sum_{j=1}^{n} \frac{t_{ji}}{M}}{N_i}, & \forall N_i > 0 \\ A_{i-1}, & otherwise \end{array} \right\} \tag{2}$$

where; i = current simulation time step,
i – 1 = previous simulation time step,
n = number of delivered bundles in current time step; i,
A_i = cumulative average TTL of delivered bundles (as at simulation time step; i),
N_i = cumulative number of delivered bundles in the network (as at simulation time step; i),
t_{ji} = TTL value of delivered bundle (j) at current time step; i M = maximum allowed TTL value in network simulator (> 0)
Initial condition: At i = 0, N_0 = 0 and A_0= 0

Properly managing TTL ensures that data packets are not endlessly circulating in the network, thereby minimising network congestion. In a VANET scenario, it is not just about transmitting messages but ensuring they are timely, especially when decisions need to be made in real-time based on received intelligence.

2) Link quality score: This is a measure of how good a communication link is regardless of the traffic on the link. This metric gauges the degree of impairment or degradation in the performance of the link, accounting for factors such as signal interference, noise and other factors that can negatively affect a link's overall quality. This measure is valuable in assessing the overall health and reliability of a communication link, aiding in identifying and addressing issues that might compromise the efficiency of data transmission. In the simulator, the link quality component of the reward signal is calculated using (3):

$$L_i = \left\{ \begin{array}{ll} \frac{\sum_{j=1}^{n} l_{ji}}{n}, & \forall n > 0 \\ 0, & otherwise \end{array} \right\} \tag{3}$$

where;
n = number of links in the network that are up (as at

This article has been accepted for publication in IEEE Open Journal of Vehicular Technology. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/OJVT.2024.3396637

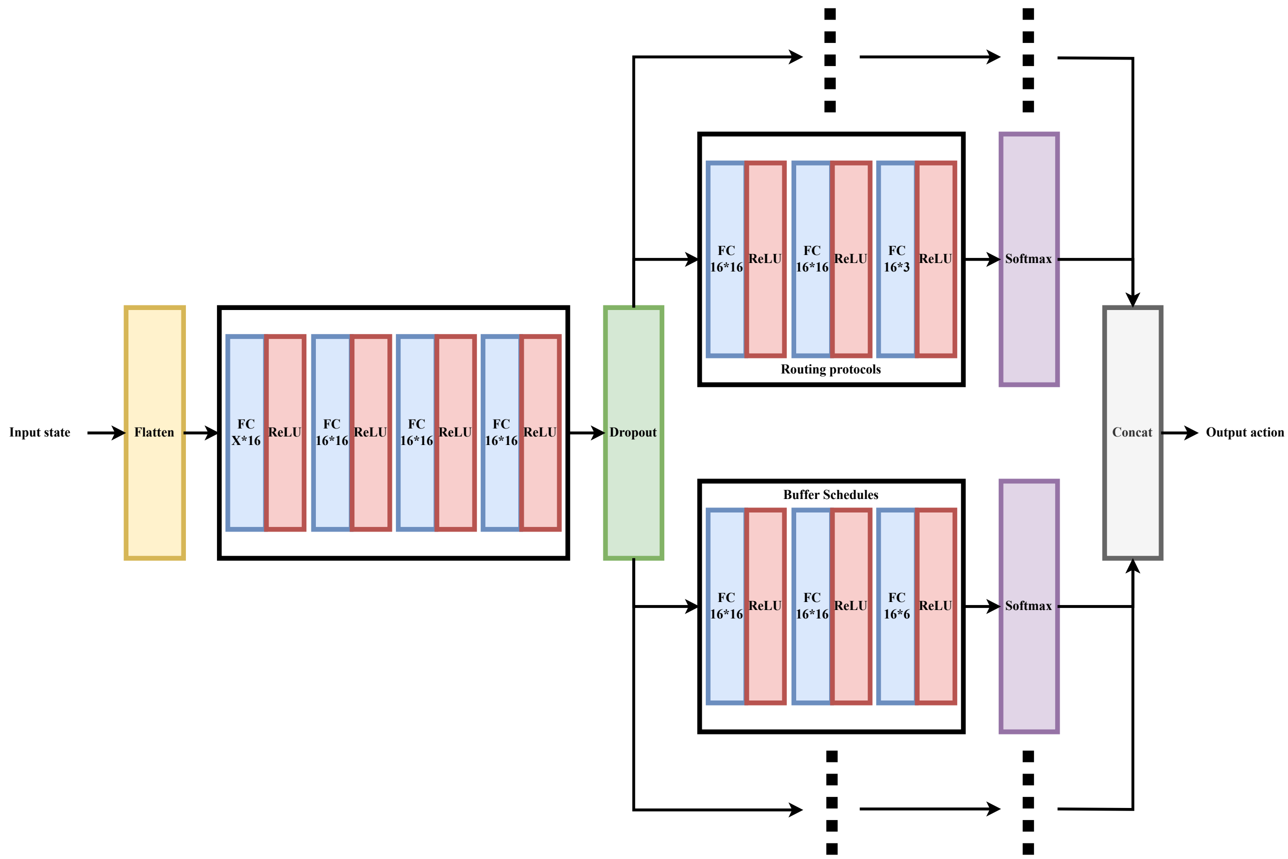

FIGURE 5. The Proposed Deep Q-Learning Model Architecture

simulation time step; i),
l_{ji} = link quality (in the range (0, 1)) of link; j (as at simulation time step; i),
L_i = average link quality of all links in the network that are up (as at simulation time step; i)

A high link quality is paramount in vehicular operations where terrain and adversarial electronic warfare tactics might affect signal strength. Ensuring consistent link quality means that units remain in touch, even in hostile environments.

3) Buffer management score: This term signifies the effectiveness of the node's internal processes in managing data bundles, encompassing the steps from their generation or arrival, through their successful transmission or forwarding, to their eventual removal from the node's memory or storage [46], [47]. This metric offers insight into the node's ability to ensure that there is still sufficient node buffer storage space left to receive and persistently store data bundles until they are successfully transmitted to their destination or to an intermediary node. In this simulator, the buffer management component of the reward signal is calculated using (4):

$$B_i = \left\{ \begin{array}{ll} \frac{\sum_{j=1}^{n} \frac{l_{ji}}{s_{ji}}}{n} \forall n, & s_{ji} > 0 \\ 0, & otherwise \end{array} \right\} \tag{4}$$

where; n = number of nodes in the network that are up (as at simulation time step; i),
B_i = average buffer management score of all n network nodes (as at simulation time step; i),
l_{ji} = buffer space left (as at simulation time step; i) on node V_J
s_{ji}= total buffer size (as at simulation time step; i) of node V_J
Effective buffer management is crucial when multiple data streams, such as video surveillance, voice communication, and data packets, compete for network resources. In tactical situations, deciding which data to prioritize can significantly impact mission outcomes.

4) Delivery ratio: This refers to the number of successfully delivered message bundles at the destination about the total bundles generated at the source. It evaluates the bundle count that is confirmatory evidence for efficient communication in DTNs. In this research, bundle delivery is defined in (5):

This article has been accepted for publication in IEEE Open Journal of Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJVT.2024.3396637

$$D_t = \left\{ \begin{array}{ll} \frac{\sum_{j=1}^{n} \frac{d_{jt}}{w_{jt}}}{n} \forall n, & w_{jt} > 0 \\ 0, & otherwise \end{array} \right\} \quad (5)$$

where;
n = DTN node count.
D_t = average bundle count of all n DTN nodes at time t,
d_{jt} = cumulative numbers of bundles that were generated by node V_j at time t and were successfully delivered to their destination
w_{jt} = cumulative bundle count generated by node V_j at time t

5) Overhead: This simply means the extra bundles transmitted beyond the actual payload where a lower overhead implies a minimal resource utilisation.
6) Latency: Time taken to propagate messages from source to destination, where a lower latency depicts a better performance

C. RL Model Training and Validation Procedure

The training and evaluation of the RL model is done using the novel simulator detailed in Sub-section D. The hyper parameters used to develop the RL model are summarised in Table 1. This linkage between decisions and rewards forms the essence of the RL agent's learning process, steering it toward optimal strategies that resonate with the overarching goal of communication optimization.

TABLE 1. Table detailing all the model training and validation Hyper Parameters.

Hyper Parameter	Value
Training Epochs	512
Training Episodes	50
Learning rate	0.02
Batch size	4
Replay memory size	10
Epsilon decay rate	0.95
Discount factor	0.99
Validation Epochs	128
Validation Episodes	50
Network size (Number of nodes)	8, 16, 32, 64

In addition, the model's training and validation is performed on a workstation with the following specifications:

- 10 Intel® Xeon(R) Gold 6230R 64-bit central processing units (CPUs) with 2.10GHz processing speed,
- 2 NVIDIA TU104GL Quadro RTX 5000 graphics processing units (GPUs), and
- 125.5 Gibibytes (GiBs) of random access memory (RAM)

D. Novel Simulation Environment Setup

The novel simulator mimics a centrally-controllable VANET, easily configurable to simulate a range of network properties (such as the frequency of bundle generation from each node, network size, link capacity, DTN routing protocols and buffer schedules employed and node buffer sizes) as depicted in Fig.6 and capable of interacting with RL models. This prompted for the utilisation of OpenAI Gymnasium as the framework to build the simulator, since it is highly regarded and utilised in training RL models in a multitude of ways [45].

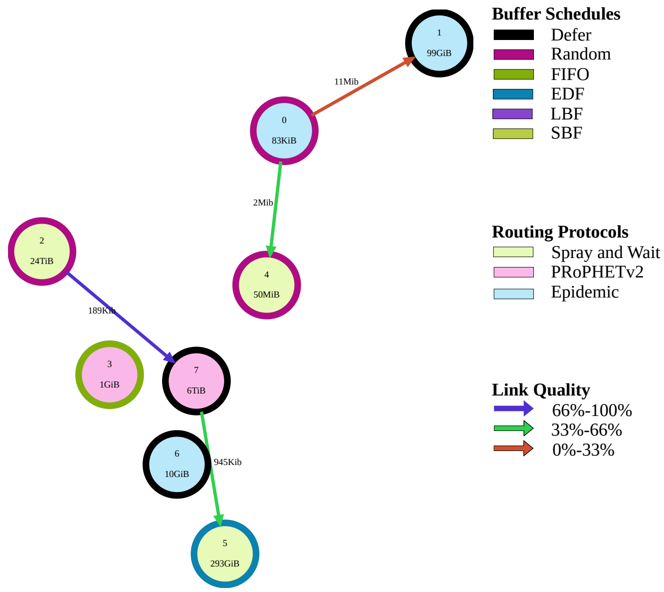

FIGURE 6. One of the proposed simulation environments depicting 8 nodes running all the three network protocols, buffer schedules and environment characteristics.

Epidemic, Spray and Wait, and PRoPHETv2 are implemented within the simulator. In addition, six buffer management scheduling algorithms are integrated to regulate the utilisation of the storage space allocated to buffering DTN bundles on each node in the simulated networks. These scheduling algorithms include [49]:

1) Position-based algorithms: Random and FIFO; These algorithms select the next bundle in the buffer to transmit basing on their position in the buffer's queue. FIFO prioritises bundles at the head of the queue for transmission while Random, selects any bundle at random for transmission. These algorithms are simple and unbiased, since they are selected for transmission without any priority or discrimination based on content or characteristics. However, they may not always be the most efficient scheduling strategies, especially in scenarios where bundles have different levels of importance or where some bundles have more stringent timing requirements. In such cases, more advanced scheduling algorithms like priority scheduling or EDF

scheduling may be employed to optimize network performance.

2) Time-based algorithms: EDF; This is a scheduling algorithm used in buffer scheduling and real-time systems. EDF refers to a mechanism where data packets are scheduled for transmission or processing based on their respective deadlines [50]. In the realm of buffer scheduling, EDF ensures that packets or tasks with imminent deadlines are processed ahead of those with later deadlines, that is, the packet or task with the earliest deadline is given priority and transmitted or executed first. In this research, the DTN bundle's TTL is used to determine the time of expiry and those nodes configured to use this scheduling algorithm expedite bundles with the smallest TTL values (earliest deadline to expire).
3) Size-based algorithms: Smallest/Largest Bundle First (SBF/LBF); These are algorithms that prioritise bundle transmission basing on the size of their payloads, where either the smallest bundles or largest bundles are transmitted first. This is advantageous in situations where node buffer storage space or link capacity are severely lacking in the DTN. SBF takes advantage of available lower-capacity links and thus increases bundle throughput despite the low capacity, while LBF ensures bundle storage space is available to receive (and buffer) more bundles by evicting those taking up the most space in a node's buffer.
4) Defer: When a packet is deferred, it is temporarily held back and not immediately sent or forwarded to its intended destination. Buffer scheduling mechanisms of VANET make decisions about when to transmit or process packets based on various criteria such as priority, available bandwidth, or the specific scheduling algorithm in use. Deferring packets can be a deliberate strategy to enhance network performance by ensuring that higher-priority or more time-sensitive packets are processed or transmitted first, while lower-priority or less time-sensitive packets are delayed preventing congestion in the DTN.

This study concentrates on the resource-based attributes within DTNs, as they are the primary determinants influencing MANETs, including VANETs, across various scenarios. Key network characteristics emphasized during the modelling phase encompass: link capacity, link loss, link uptime, node buffer size, node uptime and bundle TTL. For a comprehensive simulation, the following aspects are also integrated:

1) Traffic Pattern: This parameter reflects the quantity of bundles generated by a node at specific time intervals throughout the simulation. A consistent traffic pattern is adopted to replicate situations in VANETs where content, such as video streams from reconnaissance operations, is transmitted at a constant rate across the network.
2) Buffer Size Variance: This factor represents the degree to which the buffer size of a node fluctuates during the simulation. Random variance suggests the buffer size can either increase or decrease unpredictably, simulating conditions where VANET devices might experience inconsistent storage availability. Conversely, a stable variance implies the buffer size remains unchanged, reminiscent of advanced VANET devices with dedicated buffering capacities for network traffic.

To determine how the RL model shall optimise VANETs, the simulator is configured to mimic three network environments; Limited Resource Regions, Opportunistic Communication Regions and Advanced Technological Regions, as summarised in Table 2. These environments serve as use cases to depict the network scenarios exhibited in VANETs.

E. ONE Simulation Setup

To evaluate the behavior and performance of homogeneous VANETs, where a single protocol operates throughout the entire network, researchers use the ONE simulator. This simulator inherently supports such a configuration, allowing the utilization of only one protocol at a given time, unlike the proposed approach. Various simulation environments are established to represent the three distinct protocols. The specific simulation parameters can be found in Table 2.

To gain insights into the dynamics and effectiveness of homogeneous vehicular DTNs, this table provides a comprehensive overview of the specific simulation settings, encompassing key variables and configurations that influenced the performance evaluation. The utilization of the ONE simulator, coupled with the specified parameters, allows for a targeted examination of the homogeneous vehicular DTNs, providing valuable insights into the system's operation under the conditions of exclusive reliance on a single protocol.

The setup depicts the VANETs that adopt one routing protocol at a time for communication and the rest of the other parameters. These parameters are configured in the ONE simulator source code obtained [49], [48] compiled and run with the Eclipse Integrated Development Environment (IDE) [51], [52], to obtain results which are compared with the results obtained from the algorithm setup proposed.

V. SIMULATION RESULTS AND MODEL PERFORMANCE EVALUATION

A. Model Learning Performance

The analysis involves examining the overall learning performance trend of the model across different network densities by plotting both training and validation results.

Figure 7 shows how the model's training performance varies as the network density increases. The performance ranges from a minimum of 40% to a maximum of 84%, indicating a considerable improvement in its effectiveness with denser networks.

TABLE 2. Configuration Summary of Simulated Networks.

Network Scenario	**Limited Resource regions (worst-case)**	**Opportunistic Communication Regions (average-case)**	**Advanced Technological Regions (best-case)**
Link capacity (bps)	32k	16M	64 G
Link loss (% of link capacity)	75	25	5
Node buffer size (bytes)	16M	1G	32G
Bundle TTL (hops)	2-32	64	128
Node uptime (% of simulation time)	25	50	99.999
Link uptime (% of simulation time)	15	50	99.999
Buffer size variance	Random	Constant	
Traffic pattern	Constant		
Routing Protocol	Epidemic, Spray and Wait, PRoPHETv2		

Validation results, represented in Fig 8, reveal performance consistent with the model's training performance, thereby validating the model's effectiveness in new scenarios never before seen. Even in environments where performance was initially less favorable, the minimum validation score reached 55%, indicative of the model's effectiveness across different network densities.

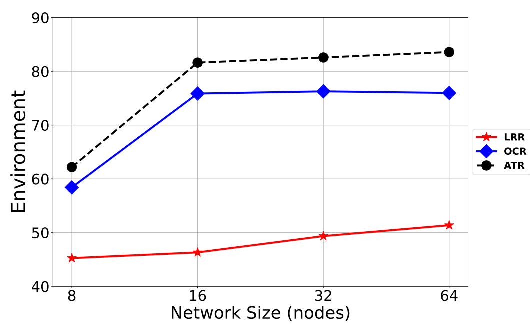

FIGURE 7. Comprehensive understanding of the model training performance under different levels of network complexity. This is crucial for assessing the model's effectiveness and robustness in diverse network environments.

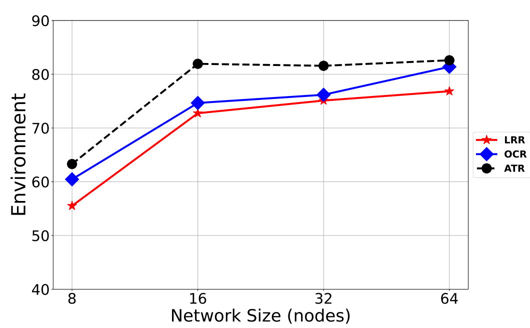

FIGURE 8. Validation results depicting the model's performance across varying network complexity. The graphs showcase the model's effectiveness diverse environmental conditions.

Where
LRR: Limited Resource Regions
OCR: Opportunistic Communication Regions
ATR: Advanced Technological Regions

To show how the model optimises the network metrics critical to the efficient performance of VANETs, multiple simulation scenario epochs, each spanning multiple point-in-time episodes (as described in the configuration in Tables 2 and 1) were executed. The outcomes are discussed in the following sections.

B. Training Phase

An exhaustive data collection and analysis process is employed to delve deeply into the performance trends of the link quality, TTL and buffer management metrics within the vehicular communication perspective.

1) Episodes

During each epoch, comprised of 50 episodes, the average trend in all metrics measured is observed to be relatively stable throughout, with the link quality score remaining mostly stable between 0.70 to 0.745, as shown in Fig. 10. Buffer management and TTL scores, in Fig. 9 and 11 respectively, show an increase between episodes 0 to 5 before stabilising at their respective levels. Obviously, the networks with more resources would perform much better, but as seen in the worst-case scenarios where networks are granted fewer resources, the proposed algorithm still manages to score above 0.7 by rapidly optimising the network resources before a steady state is reached, a phenomenon that is graphically presented in Fig. 9, 10 and 11.

2) Epochs

To get an understanding of how the proposed model improves in performance over multiple simulations, the study scrutinizes the same performance metrics over the course of 500 epochs. The outcomes in Fig. 12 show progressive improvement in buffer management score, with the worst-case performance capping out at around 0.83 in limited resource regions. In Fig. 13, it is observed that the link quality score is relatively stable, with a very significant improvement in

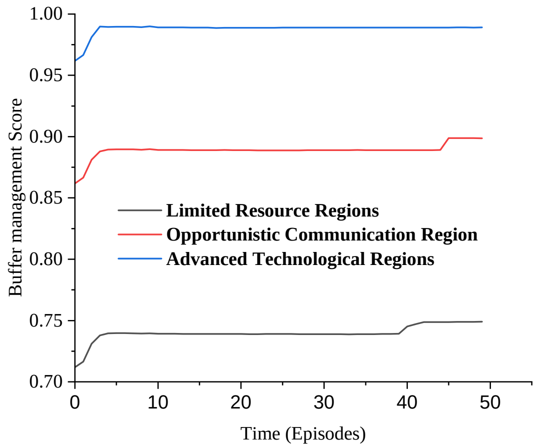

FIGURE 9. Buffer management performance score across episodes during model training

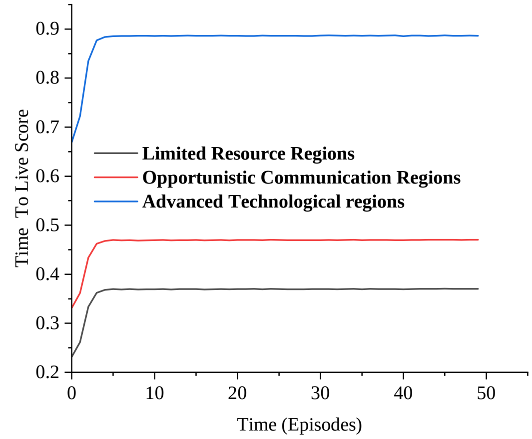

FIGURE 11. TTL performance score across episodes during model training

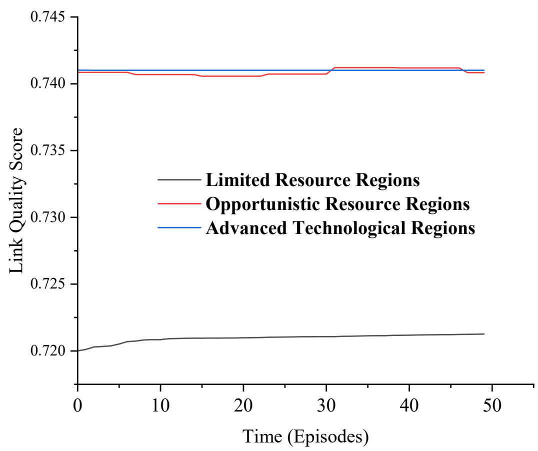

FIGURE 10. Link quality performance score across episodes during model training

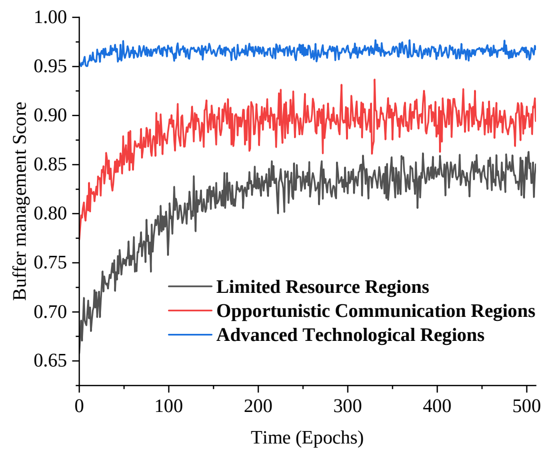

FIGURE 12. Buffer management performance score across epochs during model training

trend between epochs 0 - 50 in the worst-case scenarios, while in Fig. 14, the TTL score remains relatively stable all through, indicating no increase in performance improvement in preventing data expiry in environments with randomly-moving nodes.

C. Validation Phase

This phase, which is critical in establishing how well the model generalises its observations outside of the training set, presents results that mirror those from the training phase. Such similarity is an indicator of a well-trained and consistent model. Notably, there is minimal over fitting, meaning the model did not just memorize the training data but understood it, thereby exhibiting similar robust performance when exposed to new unseen scenarios. In the context of vehicular communications, this translates to a system that is adaptable, reliable and resilient, irrespective of randomness in the communication landscape.

Given the 120-epoch configuration, the system's performance in Fig. 15, 16 and 17 align closely with the epoch-based analysis observed during the training sessions. This is consistent in the episodic perspective as well, shown in Fig. 18, 19 and 20. As can be noted, the validation results indicate a slight reduction in performance since the model is operating on never-before-seen data. However, there is still consistency in the performance results, highlighting the robustness of the training regimen.

D. Impact of Network Density on Model Performance

The metrics, including link quality score and buffer management are of particular importance in vehicular communications. In operational environments where every millisecond

This article has been accepted for publication in IEEE Open Journal of Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJVT.2024.3396637

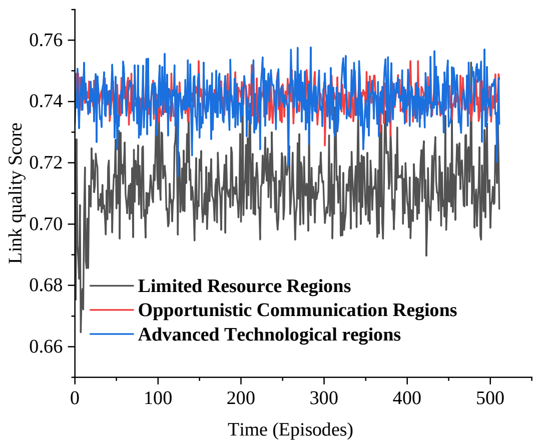

FIGURE 13. Link quality performance score across epochs during model training

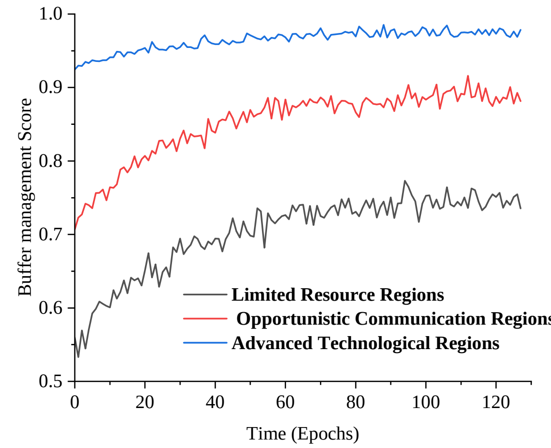

FIGURE 15. Buffer management performance score across epochs during model validation

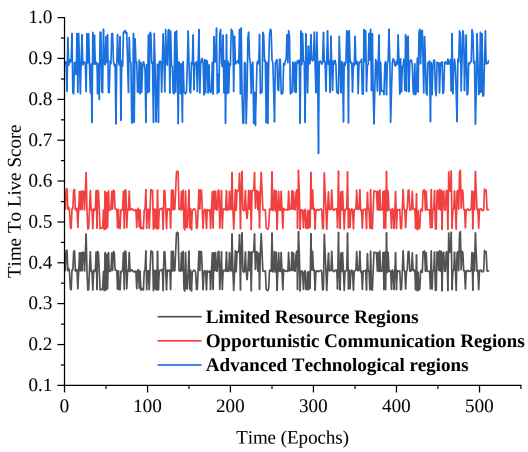

FIGURE 14. TTL performance score across epochs during model training

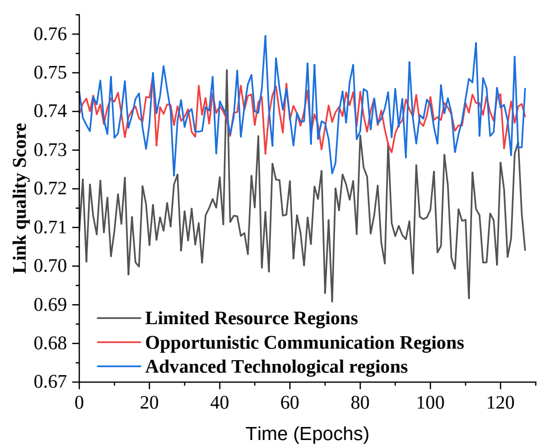

FIGURE 16. Link quality performance score across epochs during model validation

and data packet can be of strategic significance, the observed high performance of these metrics during both training and validation phases attests to the model's suitability for vehicular applications. The scalability analysis, as the network expands from 8 to 64 nodes (while keeping the world size intact) observed in Fig. 21 and 22, should be noted that there is a decline which is attributed to duplicated messages that may congest the network, thus affecting overall performance as the network gets denser. For the link quality, network congestion will degrade network links as well, thus the decline.

E. Impact of Node Buffer Size on Model Performance

In VANETs, buffer size increment is a crucial aspects related to ensuring efficient communication, handling intermittent connectivity, and managing data traffic effectively. Since vehicles often operate in dynamic environments with varying connectivity conditions. Increasing the buffer size allows vehicles to store more data packets, messages, or information when connectivity is available and transmit them opportunistically when communication links are established or improved. Fig. 23 accurately depicts a clear scenario which happens with increasing buffer sizes of 100, 500 and 1000 Mebibytes (MiB) as more bundles are effectively delivered.

F. Performance Comparison with Homogeneous VANETs

To understand how the proposed heterogeneous approach compares against the homogeneous approach, simulation results from the novel simulator and the ONE simulator were utilised respectively for this comparison. These results are presented in Table 3.

TTL, delivery ratio, latency and overhead were the metrics used to analyze and compare the performance of both the homogeneous and heterogeneous protocol approaches.

TABLE 3. Performance comparison of the homogeneous approach (running individual protocols) against the heterogeneous approach (as in the Proposed framework)

Performance metric	TTL performance				Delivery Ratio performance				Latency (s)				Overhead value			
Network Size	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64
Environment	**Limited Resource Regions**															
Epidemic	0.82	0.85	0.86	0.91	0.05	0.06	0.12	0.16	45.74	45.84	46.04	46.17	0.04	2.22	8.33	26.24
Spray and Wait	0.82	0.83	0.84	0.85	0.04	0.05	0.06	0.07	24.67	36.09	37.85	45.58	0.01	0.86	5.47	8.76
PRoPHETv2	0.81	0.82	0.85	0.87	0.01	0.15	0.20	0.27	20.34	29.52	38.53	45.75	0.03	1.1	6.49	15.90
Proposed Framework	**0.85**	**0.89**	**0.91**	**0.93**	**0.30**	**0.82**	**0.92**	**0.95**	**19.84**	**38.10**	**38.13**	**41.68**	**0.327**	**1.45**	**6.83**	**15.0**
Environment	**Opportunistic Communication Regions**															
Epidemic	0.85	0.86	0.89	0.92	0.32	0.39	0.48	0.54	22.77	37.13	37.26	39.13	0.06	2.73	11.21	42.38
Spray and Wait	0.90	0.91	0.92	0.93	0.05	0.06	0.06	0.07	22.68	22.90	22.96	33.67	0	0.12	0.23	0.44
PRoPHETv2	0.84	0.85	0.86	0.90	0.35	0.40	0.45	0.48	21.08	21.80	21.96	22.86	0	0.26	0.78	10.33
Proposed Framework	**0.92**	**0.93**	**0.94**	**0.95**	**0.50**	**0.72**	**0.90**	**0.95**	**1.191**	**3.471**	**8.150**	**17.24**	**1.02**	**1.417**	**4.17**	**15.72**
Environment	**Advanced Technological Regions**															
Epidemic	0.94	0.96	0.97	0.98	0.56	0.63	0.74	0.84	19.81	20.07	20.48	21.08	1.50	49.05	191.29	710.37
Spray and Wait	0.91	0.92	0.93	0.94	0.81	0.82	0.85	0.87	18.31	19.06	19.93	20.09	0	0.12	0.23	0.44
PRoPHETv2	0.93	0.94	0.95	0.97	0.65	0.75	0.85	0.88	11.23	13.45	13.47	13.49	0.31	45.37	89.22	123.31
Proposed Framework	**0.97**	**0.98**	**0.99**	**0.99**	**0.94**	**0.95**	**0.96**	**0.98**	**5.975**	**7.692**	**10.37**	**14.96**	**0.50**	**32.61**	**90.54**	**268.03**

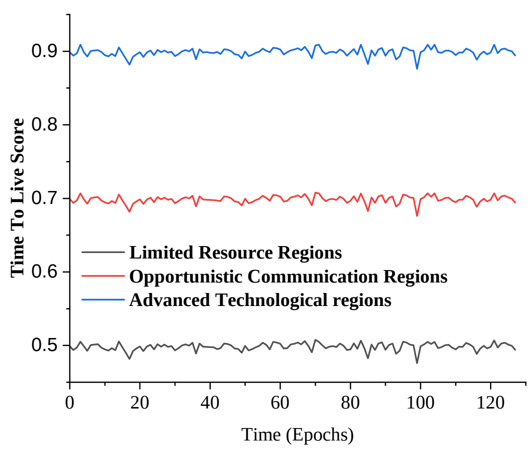

FIGURE 17. TTL performance score across epochs during model validation

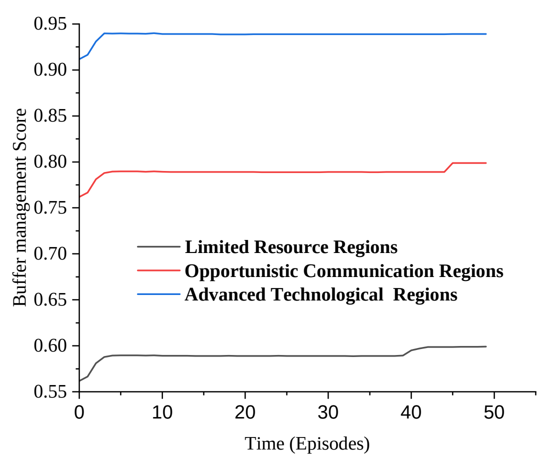

FIGURE 18. Buffer management performance score across episodes during model validation

Generally, all protocols improve in TTL and delivery ratio performance with increase in network resources and network size due to longer and more frequent contact opportunities present in denser VANETs, with more compute, storage and communication resources.

However, the latency is highest in Epidemic across all environments, due to its growing overhead as network size increases. The overhead, representing the level of data replication present in the VANET, causes congestion since data replicas in the VANET consume resources that would otherwise facilitate timely data delivery. Epidemic is ob-

This article has been accepted for publication in IEEE Open Journal of Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJVT.2024.3396637

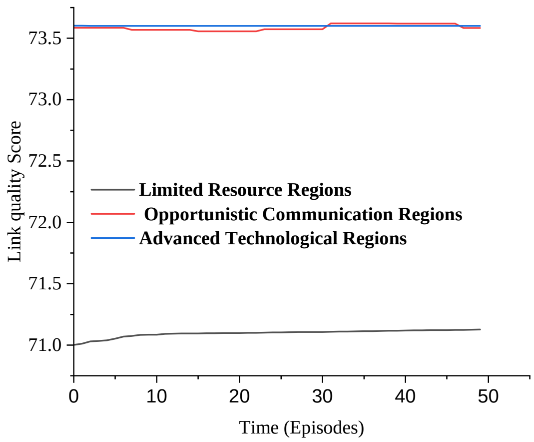

FIGURE 19. Link quality performance score across episodes during model validation

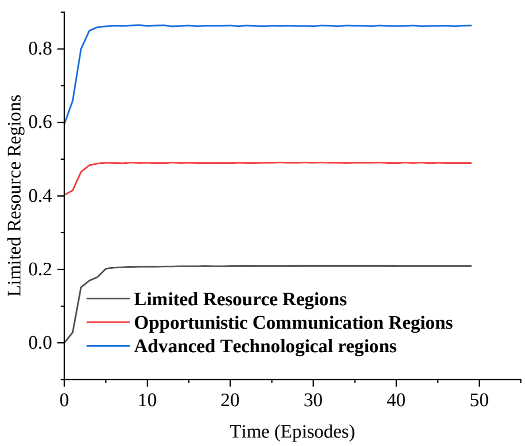

FIGURE 20. TTL performance score across episodes during model validation

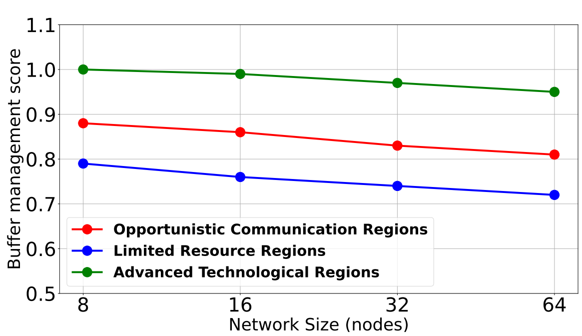

FIGURE 21. Analysis of Buffer management with increasing Network Size

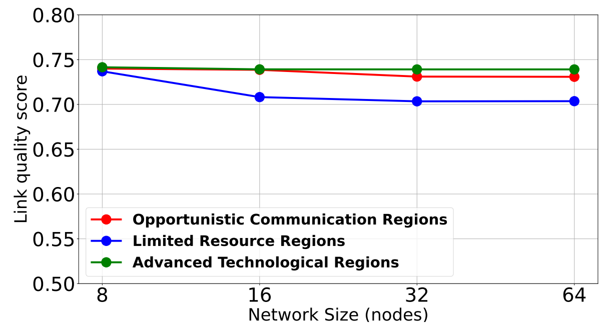

FIGURE 22. Analysis of Link quality with increasing Network Size

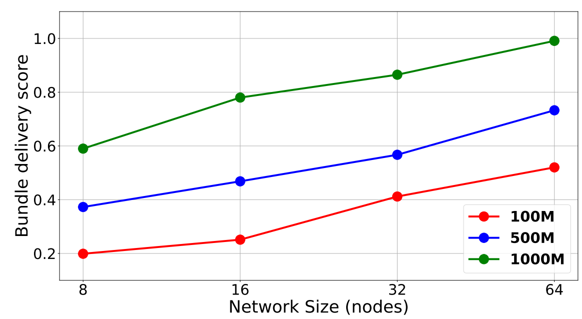

FIGURE 23. Impact of buffersize increment on the bundle delivery count with increase in network size.

served to have the highest overhead due to its flooding-based nature and PRoPHETv2 with a lower overhead because it does selective forwarding leading to optimum resource utilisation. Spray and Wait, being a flooding-based protocol like Epidemic, exhibits a lesser overhead due to the limit on its replication while in the spraying phase, striking a balance between delivery and resource consumption. The proposed framework is noticed to have the least latency as it takes advantage of the most optimal protocols in a given scenario.

The proposed heterogeneous approach (using the RL agent) outperforms all other individually-utilised DTN protocols in TTL, delivery ratio and the latency performance, as it deploys the most optimal protocol as required to maximise the performance parameters. Despite its overhead, suffered due to its use of the Epidemic (and other similar) protocol, the approach minimises the other metrics sufficiently to maximise its performance across the board.

In the context of vehicular communications, which frequently contend with unstable network conditions, it is prudent to advocate for the deployment of networks capable of seamlessly utilising multiple DTN routing protocols concurrently. This stands in stark contrast to the conventional approach of utilizing a single protocol in isolation. The superior performance demonstrated by the proposed heterogeneous approach suggests that such integrated configurations are not only feasible but also highly advantageous, making a compelling case for the adoption of integrated network configurations, especially in contexts characterized by unpredictable and challenging communication environments such as vehicular operations.

G. Results Synthesis

- In VANETs, communication extends beyond mere data transmission; it focuses on ensuring a continuous, resilient, and real-time relay of vital information, even in challenging environments and situations. Analyzing outcomes from training and validation phases provides valuable insights into the model's potential and relevance for VANET communication.
- Alignment between training and validation outcomes and its significance in VANET scenarios: Striking a harmonious balance between training and validation results is particularly challenging in VANETs due to their unpredictable nature. The consistent performance observed suggests that the model can be relied upon in actual vehicular environments, adapting to the distinct conditions encountered.
- Adaptability across network sizes: Vehicular operations in VANETs can range widely in size. Whether it involves a specialized team of eight vehicles on a covert mission or a 64-vehicle network supporting a larger operation, effective communication remains essential. The model's proficiency across varying network sizes underscores its flexibility and scalability, both crucial attributes for a reliable VANET solution.

VI. CONCLUSION

The dynamic landscape of vehicular technology, especially within VANETs, calls for innovations that are both groundbreaking and reliable, especially in critical scenarios. The results of this study highlight the significant impact of integrating an RL agent into a system tailored to address the complexities of vehicular communication networks.

During the training phase, the consistent performance observed across various metrics confirmed the RL agent's ability to effectively navigate the specific challenges inherent to vehicular environments. The model's strong performance during training demonstrated its intrinsic capabilities, a sentiment echoed during the subsequent validation phase, emphasizing its resilience and practical utility. Such reliability is indispensable in VANETs, where errors are costly, and the ramifications of any failures can be extensive.

Our detailed examination emphasized the RL agent's potential advantages over homogeneous networks (modelled by the ONE simulator) within the VANET domain, in the Limited Resource, Opportunistic Communication and Advanced Technological Regions. The model's adaptive capacity to utilize different protocols and buffer strategies, combined with its flexibility across varying network sizes, underscores its readiness for diverse vehicular scenarios. Whether ensuring consistent communication during discrete operations or facilitating large-scale maneuvers, the RL agent's versatility shines through.

Moreover, the utilization of network state information and the strategic selection from an assortment of routing protocols and buffer strategies mark an innovative approach to vehicular communication management. Such agility and adaptability, core strengths of the RL agent, align seamlessly with the unpredictable nature of vehicular networks. However, due to the centralised nature of the controlling agent and the reliance on established back-haul infrastructure, there is need to explore offline redundancy through the use of multiple RL-enabled controller instances deployed within the VANETs themselves, to further improve on the system's reliability and robustness.

In summary, this study illuminates the potential synergies between advanced technology and the complex realm of VANETs in vehicular technology. Through its novel integration and remarkable performance, the RL agent paves the way for more resilient, efficient, and adaptable vehicular communication networks. The culmination of these insights provides a road-map for an innovative era in vehicular technology, characterized by reliability, innovation, and unparalleled flexibility. Moving forward, exploring the integration of alternative routing protocols remains a crucial avenue to ensure optimal performance across varied VANET environments.

REFERENCES

[1] K. Chandramohan, A. Manikandan, S. Ramalingam, and R. Dhanapal, “Performance evaluation of VANET using Directional Location Aided Routing (D-LAR) protocol with sleep scheduling algorithm,” *Ain Shams Engineering Journal,* vol. 15, no. 3, p. 102458, 2024.

[2] S. A. Elsagheer Mohamed, K. A. Alshalfan, M. A. Al-Hagery, and M. T. Ben Othman, “Safe driving distance and speed for collision avoidance in connected vehicles,” *Sensors,* vol. 22, no. 18, p. 7051, 2022

[3] R. Anggoro, R. Muslim, M. H. AA et al., “Performance evaluation for non-DTN and DTN geographic based protocols for VANETs,” *Journal of the Institute of Industrial Applications Engineers,* vol. 5, no. 1, pp. 47–52, 2017.

[4] N. H. Hussein, S. P. Koh, C. T. Yaw, S. K. Tiong, F. Benedict, T. Yusaf, K. Kadirgama, and T. C. Hong, “SDN-based VANET routing: A comprehensive survey on architectures, protocols, analysis, and future challenges,” *IEEE Access,* 2024.

[5] R. I. Helmi and S. M. A. H. Mizan, “Routing protocols for Vehicle-to-Vehicle (V2V) networks,” 2023.

[6] I. Zacarias, L. P. Gaspary, A. Kohl, R. Q. Fernandes, J. M. Stocchero, and E. P. de Freitas “‘Combining Software-Defined and Delay-Tolerant approaches in last-mile Tactical Edge networking,” *IEEE Communications Magazine*, vol. 55, no. 10, pp.22–29, 2017.

[7] M. A. Hoque, M. Razu, T. Islam, and A. Amin, “SDN-DTN combined architecture in post disaster scenario—a new way to start,” in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE*, 2020, pp.230–235.

[8] S. Levine, A. Kumar, G. Tucker, and J. Fu, “Offline Reinforcement Learning: Tutorial, review, and perspectives on open problems,” *arXiv preprint arXiv:2005.01643,* 2020.

[9] R. Amin, E. Rojas, A. Aqdus, S. Ramzan, D. Casillas-Perez, and J. M. Arco, “A survey on machine learning techniques for routing optimization in SDN,” *IEEE Access,* vol. 9, pp. 104 582–104 611, 2021

[10] T. E. Ali, Y.-W. Chong, and S. Manickam, “Machine learning techniques to detect a DDoS attack in SDN: A systematic review,” *Applied Sciences,* vol. 13, no. 5, p. 3183, 2023.

[11] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoel- lig, “Safe learning in robotics: From learning-based control to safe reinforcement learning,” *Annual Review of Control, Robotics, and Autonomous Systems,* vol. 5, pp. 411–444, 2022.

[12] T. M. Moerland, J. Broekens, A. Plaat, C. M. Jonker et al., “Model-based reinforcement learning: A survey,” *Foundations and Trends® in Machine Learning,* vol. 16, no. 1, pp. 1–118, 2023.

[13] J. Chen, P. Bie, J. Nie, and Z. Wei, 'HP-ECD: Heuristic PRoPHET protocol based on energy balance,cache optimization, and asynchronous dormancy', J. King Saud Univ. - *Comput. Inf. Sci.,* vol. 36, no. 1, p. 101861, Jan. 2024, doi: 10.1016/j.jksuci.2023.101861.

[14] V. Sridharan, M. Motani, B. Jalaian, and N. Suri, "Exploring performance tradeoffs in Tactical Edge Networks," *IEEE Communications Magazine,* vol. 58, no. 8, pp.28–33, 2020.

[15] S. Grasic, E. Davies, A. Lindgren, and A. Doria, "The evolution of a DTN routing protocol-PRoPHETv2," *in Proceedings of the 6th ACM Workshop on Challenged Networks,* 2011, pp. 27–30.

[16] H. Yang, G. Li, G. Sun, J. Chen, X. Meng, H. Yu, W. Xu, Q. Qu, and X. Ying, "Dispersed computing for tactical edge in future wars: vision, architecture, and challenges," *Wireless Communications and Mobile Computing,* vol. 2021, pp. 1–31,2021.

[17] A. K. Gupta, J. K. Mandal, and I. Bhattacharya, "Comparative performance analysis of DTN routing protocols in multiple post-disaster situations," *in Contemporary Advances in Innovative and Applicable Information Technology: Proceedings of ICCAIAIT 2018. Springer,* 2019, pp. 199–209

[18] A. Verma, Savita, and S. Kumar, "Routing protocols in Delay Tolerant Networks: Comparative and empirical analysis," *Wireless Personal Communications,* vol. 118, pp. 551–574, 2021

[19] E. Davies et al., "DTN—the state of the art,"*Networking for Communications Challenged Communities: Architecture, Test Beds and Innovative Alliances.,* 2009

[20] R. Amin, E. Rojas, A. Aqdus, S. Ramzan, D. Casillas-Perez, and J. M. Arco, "A survey on machine learning techniques for routing optimization in SDN," *IEEE Access*, vol. 9, pp. 104 582–104 611, 2021.

[21] T. E. Ali, Y.-W. Chong, and S. Manickam, "Machine learning techniques to detect a DDOS attack in SDN: A systematic review," *Applied Sciences,* vol. 13, no. 5, p. 3183,2023.

[22] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe Reinforcement Learning," *Annual Review of Control, Robotics, and Autonomous Systems,* vol. 5, pp. 411–444, 2022.

[23] T. M. Moerland, J. Broekens, A. Plaat, C. M. Jonker et al., "Model-based reinforcement learning: A survey," *Foundations and Trends® in Machine Learning,* vol. 16, no. 1, pp. 1–118, 2023.

[24] M. A. Al-Shareeda and S. Manickam, "A systematic literature review on security of Vehicular Ad-hoc Network (VANET) based on veins framework," *IEEE Access,* 2023.

[25] S. A. Elsagheer Mohamed and K. A. AlShalfan, "Intelligent traffic management system based on the Internet of Vehicles (IoV)," *Journal of advanced transportation, vol. 2021,* pp. 1–23, 2021.

[26] S. A. Elsagheer Mohamed, "Automatic traffic violation recording and reporting system to limit traffic accidents: based on Vehicular Adhoc Networks (VANET)," in *Proceedings of the 2019 International Conference on Innovative Trends in Computer Engineering (ITCE),* pp. 254–259, Aswan, Egypt, 2019.

[27] S. Sabeeh and K. Wesolowski, "Congestion control in autonomous resource selection of cellular-v2x," *IEEE Access,* vol. 11, pp. 7450–7460, 2023

[28] A. Verma, Savita, and S. Kumar, "Routing protocols in delay tolerant networks: Comparative and empirical analysis," Wireless Personal Communications, vol. 118, pp. 551–574, 2021.

[29] A. Keränen, J. Ott, and T. Kärkkäinen, "The one simulator for DTN protocol evaluation," in Proceedings of the *2nd international conference on simulation tools and techniques,* 2009, pp. 1–10.

[30] Z. Wang and H. Wu, "A simulation module of ns-3 for DTN with time-space model based routing," 2017

[31] A. Dhadhania, J. Bhatia, R. Mehta, S. Tanwar, R. Sharma, and A. Verma, "Un- leashing the power of SDN and GNN for network anomaly detection: State-of-the-art, challenges, and future directions," *Security and Privacy,* vol. 7, no. 1, p. e337, 2024

[32] Y. Raghavendra and U. Mahadevaswamy, "Sbldar: A link score based delay aware routing for WSNs," *Wireless Personal Communications,* vol. 132, no. 1, pp. 629–650, 2023

[33] A. K. Gupta, J. K. Mandal, and I. Bhattacharya, "Comparative performance anal- ysis of DTN routing protocols in multiple post-disaster situations," *in Contemporary Advances in Innovative and Applicable Information Technology: Proceedings of ICCAIAIT 2018. Springer,* 2019, pp. 199–209.

[34] S. Kaur, J. Singh, and N. S. Ghumman, "Network programmability using POX controller," in *ICCCS International conference on communication, computing & systems, IEEE,* vol. 138. sn, 2014, p. 70

[35] S.-Y. Wang, "Comparison of SDN openflow network simulator and emulators: Estinet vs. mininet," in *2014 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2014,* pp. 1–6.

[36] S.-Y. Wang, C.-L. Chou, and C.-M. Yang, "Estinet openflow network simulator and emulator," *IEEE Communications Magazine,* vol. 51, no. 9, pp. 110–117, 2013

[37] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, "Link quality and energy utilization based preferable next hop selection routing for wireless body area networks," *Computer Communications,* vol. 149, pp. 382–392, 2020.

[38] Y. Raghavendra and U. Mahadevaswamy, "Sbldar: A link score based delay aware routing for WSNs," *Wireless Personal Communications,* vol. 132, no. 1, pp. 629–650, 2023

[39] P. Upadhyay, S. J. Goyal, V. Marriboyina, and S. Kumar, "Securing Vehicular Internet of Things (V-IoT) communication in smart VANET infrastructure using multi-layered com- munication framework and novel threat detection algorithm," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 12, no. 6s, pp. 789–803, 2024.

[40] F. G. Abdulkadhim, Z. Yi, C. Tang, A. N. Onaizah, and B. Ahmed, "Design and development of a hybrid (SDN+ SOM) approach for enhancing security in VANET," *Applied Nanoscience,* vol. 13, no. 1, pp. 799–810, 2023.

[41] B. Deebak and S. O. Hwang, "Intelligent drone-assisted robust lightweight multifactor authentication for vehicular zone surveillance in the 6G era," *Computer Networks,* vol. 225, p. 109664, 2023.

[42] L. Mohaisen and L. Joiner, "Towards Delay Tolerant Networking for connectivity aware routing protocol for VANET-WSN communications," *Applied Sciences,* vol. 13, no. 6, p.4008, 2023.

[43] B. A. Vishwanathrao and P. A. Vikhar, "Reinforcement machine learning-based improved protocol for energy efficiency on Mobile Ad-hoc Networks," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 12, no. 8s, pp. 654–670, 2024.

[44] B. Hildebrand, M. Baza, T. Salman, S. Tabassum, B. Konatham, F. Amsaad, and A. Razaque, "A comprehensive review on blockchains for Internet of Vehicles: Chal- lenges and directions," *Computer Science Review,* vol. 48, p. 100547, 2023.

[45] A. Zai and B. Brown, Deep Reinforcement Learning in action. Manning Publications, 2020.

[46] R. Nagao, H. Masaki, and H. Harada, "A routing method using packet buffer management in wi-sun fan network," *IEICE Technical Report; IEICE Tech. Rep.,* vol. 123, no. 75, pp. 1–6, 2023.

[47] C. Szwarcfiter, Y. T. Herer, and A. Shtub, "Shortening the project schedule: solving multimode chance-constrained critical chain buffer management using Reinforcement Learning," *Annals of Operations Research,* pp. 1–28, 2023.

[48] M. Day, M. R. Penumala, and J. Gonzalez-Sanchez, "Annete: An intelligent tutoring companion embedded into the Eclipse IDE," in 2019 *IEEE First International Conference on Cognitive Machine Intelligence (CogMI).* IEEE, 2019, pp. 71–80.

[49] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE simulator for DTN protocol evaluation," in Proceedings of the *2nd international conference on simulation tools and techniques,* 2009, pp. 1–10.

[50] A. Kumar and B. Alam, "Energy harvesting Earliest Deadline First scheduling algorithm for increasing lifetime of real time systems," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 1, p. 539, 2019.

[51] V. Raju, A. S. Teja, D. Ragnesh, and M. A. Sagar, "Concrete mix design using Java Eclipse IDE," in E3S *Web of Conferences,* vol. 391. EDP Sciences, 2023, p. 01207.

[52] J. desRivieres and J. Wiegand, "Eclipse: A platform for integrating development tools," *IBM Systems Journal,* vol. 43, no. 2, pp. 371–383, 2004.

OLIVIA NAKAYIMA received her B.Sc. Degree in Computer Engineering from Makerere University, Uganda in 2019. She is currently a M.Sc. Student at Egypt-Japan University of Science and Technology, Alexandria, Egypt. Her research interests span across Computer Networks, Machine learning, Wireless Networks, Mobile AdHoc Networks, Vehicular Ad-hoc Networks, Delay Tolerant Networks.

MOSTAFA I. SOLIMAN received the B.Sc. and M.Sc. degrees in computer science and engineering from the University of Assiut, Egypt, in 1994 and 1998, respectively, and the Ph.D. degree in computer science and engineering from the University of Aizu, Japan, in 2004. He is currently a Full Professor at Aswan University, Egypt. He is also the General Director of Computer Science and Information Technology (CSIT) Programs at the Egypt–Japan University of Science and Technology. His research interests include computer architecture, parallel processing, vector/matrix processing, performance evaluation, parallel algorithms, FPGA, and SystemC implementations.

KAZUNORI UEDA received the M.Eng. And Dr. Eng. degrees from The University of Tokyo, in 1980 and 1986, respectively. He joined NEC, in 1983, and from 1985 to 1992, he was with the Institute for New Generation Computer Technology (ICOT) on loan. He joined Waseda University, in 1993, and has been a Professor, since 1997. He has been a Visiting Professor with the Egypt-Japan University of Science and Technology, since 2010. His research interests include design and implementation of programming languages, concurrency and parallelism, high-performance verification and hybrid systems. His research interests are in Cybersecurity, Network Security, Security on the Internet of Vehicles, Anti-phishing techniques, Wireless Ad hoc Networks, and Wireless Sensor Networks, Modern Technologies for the Holy Quran Services, Anti-Spam fighting and filtering, Traffic safety and optimization, Street Lights optimization and control, Vehicular Ad hoc Networks (VANET), bioinformatics, audio and video quality assessment in packet networks, neural networks training, traffic prediction, and video coding. He has got 13 grants for research projects in different areas and from different grantors.

SAMIR A. ELSAGHEER MOHAMED obtained his B.Sc. degree from the Faculty of Engineering at the University of Assuit, Egypt, in May 1994. He obtained M.Sc. and Ph.D. degrees in Computer Science and Engineering from the University of Rennes I, France, in 1998 and 2003, respectively. He worked as R&D Expert Engineer in INRA/IRISA, France after graduation. Then he joined the Faculty of Engineering, Aswan University as Assistant Professor. After that, he joined the College of Computer, Qassim University as Assistant Professor from 2006 to 2013. He was promoted to Associate professor and continue working with Qassim University till Aug. 2019. Finally, he joined the Egypt-Japan University for Science and Technology in Sept. 2019. He was in charge of the IT center and the Data Center as well as teaching and research till Jan. 2022. His research interests are in Cybersecurity, Network Security, Security on the Internet of Vehicles, Anti-phishing techniques, Wireless Ad hoc Networks, and Wireless Sensor Networks, Modern Technologies for the Holy Quran Services, Anti-Spam fighting and filtering, Traffic safety and optimization, Street Lights optimization and control, Vehicular Ad hoc Networks (VANET), bioinformatics, audio and video quality assessment in packet networks, neural networks training, traffic prediction, and video coding. He has got 13 grants for research projects in different areas and from different granters.