

Received December 3, 2021, accepted December 24, 2021, date of publication December 28, 2021, date of current version January 4, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3138903

A Novel Deep Reinforcement Learning Based Relay Selection for Broadcasting in Vehicular Ad Hoc Networks

ABIR MCHERGUI¹ AND TAREK MOULAH^{2,3}

¹Department of Computer Science, ISG, Gabes University, Gabes 6029, Tunisia

²Department of Information Technology, College of Computer, Qassim University, Buraydah 52571, Saudi Arabia

³Faculty of Science and Technology of Sidi Bouzid (FSTSB), University of Kairouan, Kairouan 3100, Tunisia

Corresponding author: Abir Mchergui (abir.mchergui1437@gmail.com)

ABSTRACT VANETs (Vehicular Ad hoc NETWORKs) are considered among the world's largest networks. These networks are providing multiple services like infotainment applications, safety services, driver assistance, and even video on demand. On one hand, VANETs are characterized by their random topology and dynamic behavior that varies in urban context, and which highly changes in highways. On the other hand, diffusing information is a fundamental task to deliver multiple services. Thus, the broadcasting task is a challenging problem and need more investigation. In fact, to achieve this task, artificial intelligence and learning based computing seem to be one of the most appropriate options that best fits the dynamic behavior of VANETs. Accordingly, in this paper we propose a novel hybrid relay selection technique to perform the broadcasting task based on a reinforcement learning method. Our proposition is initially to combine an artificial neural network-based classification applied to select forwarding nodes, and in the second phase, we apply the Viterbi algorithm as a reinforcement tool to refine the first classification. To measure the performance of our contribution, we adopt a grid map scenario with varied traffic densities. Afterwards, we analyze and compare the simulation results with other methods in the literature based on different parameters such as the success rate, the data loss, the saved rebroadcasts, and the delay. We conclude by proving that the proposed technique combining deep learning along with reinforcement learning outperforms other recently proposed broadcasting schemes based on the results which show that the new solution increased the success rate by 16%, the saved rebroadcasts by 20%, and reduced the delay by 23%.

INDEX TERMS Broadcasting, VANET, relay selection, reinforcement learning, Viterbi algorithm.

I. INTRODUCTION

VANETs (vehicular ad hoc networks) enable a wide range of intelligent transportation system services. Vehicle-to-vehicle and vehicle-to-infrastructure communications are used to provide services varying from road safety to infotainment and traffic management [1]. VANET creates self-organizing networks composed of nodes with dedicated short-range communication (DSRC) implanted in automobiles. Due to their varied movement speeds and bounded covered area, automobiles constitute a mobile ad hoc network along the roads with very dynamic topology [2].

VANET provides many applications, which are either safe or non-safe [3]. Message dissemination presents the main

The associate editor coordinating the review of this manuscript and approving it for publication was Tariq Umer.

common issue of all these applications. Indeed, data delivery is one of the fundamental tasks in VANET applications deployment [4], [5]. In these kinds of networks with non-predefined infrastructure, every vehicle can send, forward or receive a disseminated packet of data. Thus, designing an efficient broadcasting model is considered as a challenge in order to guarantee a maximum accessibility with optimal performance.

Due to recurrent connection loss because of the interference, the unavailability of channel, and the possible topology variation, broadcasting protocols in VANETs are susceptible of malfunction. Consequently, the connections of vehicle-to-infrastructure and vehicle-to-vehicle can be intermittent. In addition, a malfunctioning of broadcasting protocol can lead to poor quality vehicular services and even their possible failure. Therefore, regarding such conditions,

it is greatly important to deploy efficient and stable data broadcasting.

Previously, we proposed eKMMPR (enhanced Kinetic Multi-Point Relay) [6], which is a dissemination technique combining the fuzzy system and a kinetic strategy. In eKMMPR scheme, we elected the relay set depending on some mobility parameters. It is an indisputable fact that the main challenge of VANET is to manage the dynamic behavior issue. In this research work, we aim to optimize the selection of relay set by combining deep learning techniques and reinforcement learning. In this regards, We introduce a novel hybrid relay selection technique for broadcasting in vehicular networks based on a reinforcement learning method using the Viterbi algorithm, and we apply artificial neural network for the classification of nodes to forward the message.

Deep learning, which is known for its ability to deal with complex problems that need automated decision making, could be a viable solution in a dynamic environment like in VANET. The huge data generated by transportation system traffic, on the other hand, can be relevant for useful data delivery. In our context, deep learning is used in order to use data properly and efficiently to select best relay nodes to accomplish the broadcasting task. In addition, reinforcement learning is the most suitable technique to deal with the dynamic behavior of VANET.

Deep learning has been used by researches, as an artificial intelligence tool, to solve many issues pertaining to broadcasting in VANET. To the best of our knowledge, combining it with reinforcement learning for the purpose of optimizing the relay set selection has not been proposed before.

The aim of this paper is to propose a novel hybrid relay selection technique for broadcasting in VANET through combining reinforcement learning together with the artificial neural network. In this regard, ANN is applied for the classification of nodes to forward the message. Subsequently, the Viterbi algorithm is used as a reinforcement phase to refine the first selection. To test the effectiveness of the proposed solution, we use a grid map scenario with varied traffic densities. Thereafter, we analyze and compare the simulation results with three other broadcasting techniques.

The remainder of the paper proceeds as follows. Section 2 discusses the related works. Section 3 deals with the proposed scheme of relay selection based on reinforcement learning. In section 4, we outline the simulation set up besides discussing and comparing the results of our technique with other similar schemes. Finally, section 5 concludes the paper.

II. RELATED WORKS

A. RELATED WORKS IN RELAY SELECTION FOR BROADCAST

Throughout most of the VANET-related research works, intelligent methods for information dissemination were prevalent. The authors of [7] introduced a multi-hop relay selection method based on intelligent greedy position. It improves the delivery of data while lowering the overall

percentage of hops. The authors of [8] presented an emergency message broadcasting protocol using an edge computing approach that employs the non-isomeric range of vehicular communication. This solution is implemented in a distributed scheme that takes advantage of the local topology data of VANET. It performs well in terms of latency average, packet delivery ratio and overhead, in both sparse and intensive traffic flow environments [9].

In [10], Jafer *et al.* presented an adapted genetic algorithm-based broadcast for freeway traffic environments. This scheme intends to minimize broadcast storms by applying an empirical fitness function to govern packet forwarding and to lower the overall propagation duration. [11] introduced another intelligent message transmission approach based on fuzzy logic for velocity prediction and relaying selection. It additionally targets connectivity time, number of hops, and travel guidance that are not dependent on GPS. Regarding the use of radio V2V communications, a machine learning model incorporated on the road side unit (RSU) performs the forwarding decision.

Heterogeneous communications-based information dissemination techniques have been developed in latest researches of automotive ad hoc networks. The authors in [12] propose a modified relay node selection algorithm for data broadcasting. The sender-oriented strategy increases accessibility in hybrid contexts by taking into account a variety of selection parameters for intermediary nodes, most notably that of the geographical configuration of neighbors.

In [13], the authors introduced a cluster based data dissemination approach using self-Assessment for Sparse and Dense traffic conditions for Internet of Vehicles. It uses the distance as a relay selection parameter and improves the packet delivery, the delay and the throughput. In [14], a geographical relay selection technique is presented based on road perception. The main selection metrics are: distance, direction and mid-range. The proposed technique is adapted to urban environment and it aims to improve the packet delivery, the average delay and the hop count. In table 1, we give a comparison of broadcasting techniques based on relay selection.

B. RELATED WORKS IN DEEP LEARNING BASED RELAY SELECTION FOR BROADCAST

In recent decades, machine learning and, in particular, deep learning have become prominent artificial intelligence tools for a myriad of issues, including those involving following hop decision. The three major disciplines of machine and deep learning are supervised learning, unsupervised learning, and reinforcement learning. The process of supervised learning entails first training the network with some classified material and then categorizing input characteristics using the training dataset. Unsupervised learning enables latent features to be extracted from unlabeled data. Eventually, by modifying variables and learning throughout multiple tests, reinforcement learning is utilized to achieve optimization goals.

TABLE 1. Comparison of broadcasting techniques based on relay selection.

Ref	Type	Relay selection method	Performance metrics	Environment
[7]	Position based	Intelligent greedy position	Packet delivery Number of hops	Urban VANETs
[8]	Topology based	Non isomeric range of vehicles	Latency Overhead Packet delivery	Heterogeneous VANETs
[10]	Probability based	Empirical fitness function	Broadcast Storm Propagation duration	Freeway VANETs
[11]	Position based	Fuzzy logic machine learning model	Connectivity time Number of hops	Urban VANETs
[12]	Geographic based	Sender oriented strategy	Accessibility	Hybrid VANETs
[13]	Cluster based	Self assesment Distance	Packet delivery delay Throughput	Sparse and Dense urban VANETs
[14]	Geographic based	Distance; direction and midrange	Packet delivery Average delay Hop counts	Urban VANETs

TABLE 2. Comparison of broadcasting techniques based on deep learning for relay selection.

Ref	Type	Learning method	Performance metrics	Environment
[15]	Unsupervised learning	K-means	Acessibility Size of relay set	Urban
[16]	Reinforcement learning	Stochastic learning automata	Broadcsat time Success rate	LTE
[17]	Reinforcement learning	Sliding window Viterbi algorithm	Latency End-to-end SNR	LTE
[18]	Deep learning	CNN	Delay Throughput	Highway
[19]	Social learning	Store-Carry-Forward K-Clique algorithm	Average delay Packet delivery	Urban

These frameworks have been utilized to address broadcasting issues in a number of recent VANET studies. For example, the authors of [15] presented a multi-hop wireless broadcast relay selection technique based on unsupervised learning in urban vehicular networks. The authors devised a k-mean clustering technique using D2D communication and simulated it under various traffic densities. For mutual communications, a distributed learning-based relay placement is introduced in [16]. It is a completely decentralized system that employs a stochastic learning automaton for reinforcement learning that provides vehicle units with auto learning abilities in an LTE environment.

In [17], a Viterbi algorithm-based relay decision approach for subsequent hop-scheme optimization is described. Convolutional codes are used in this approach to choose paths and maximize the signal to noise ratio measure. A sliding window is also used to regulate processing latency and memory space.

Some data distribution broadcasting techniques place a greater emphasis on content. For example, [18] proposes a content-centric strategy for data broadcast that takes into account vehicle mobility and is based on deep learning algorithms. It aims to predict social ties between engaged cars in the network using a convolutional neural network to minimize delay and boost delivery. In [19], the authors presented a near-identical approach based on Store-Carry-Forward patterns. In table 2, we provide a comparison of broadcasting techniques using relay selection based on deep learning.

III. REINFORCEMENT LEARNING BASED RELAY SELECTION

A. MODEL OVERVIEW

RSUs are linked both wired and wirelessly in the VANET relay selection technique based on hybrid communications presented in this work. A joint unit is responsible for applying deep learning to classify nodes and for making relay set selection decisions based on the acquired results. This device, known as the Learning Processing Unit (LPU), is wired to each RSU and has its own energy, processing, and storage capabilities.

The vehicle nodal degree is a real-time number based on the neighboring association involving all of the cars in the broadcasting area. It is computed with the help of supervised learning and then it is fine-tuned through reinforcement learning. Each node continually finds its neighbors in real-time circumstances and creates a node vector for each of them including key categorization features.

Figure 1 depicts a broad overview of the suggested solution's surroundings.

At broadcast moment, roadside units acquire a set of node vectors from every car in relation to every other car its surroundings. As illustrated in figure 2, a node may have various node vectors corresponding to multiple neighbors and, as a result, may be allocated to differing degrees from different perspectives. The red node in this diagram is a neighbor of nodes A, B, and C at the same time. As a result, three equivalent node vectors out of these nodes are transmitted

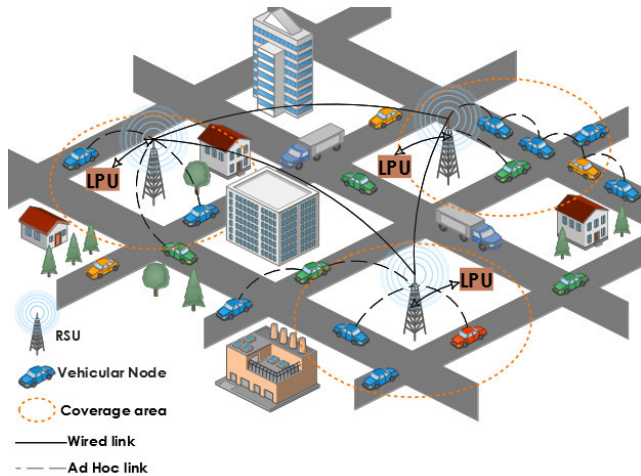


FIGURE 1. Hybrid communication environment for the proposed scheme.

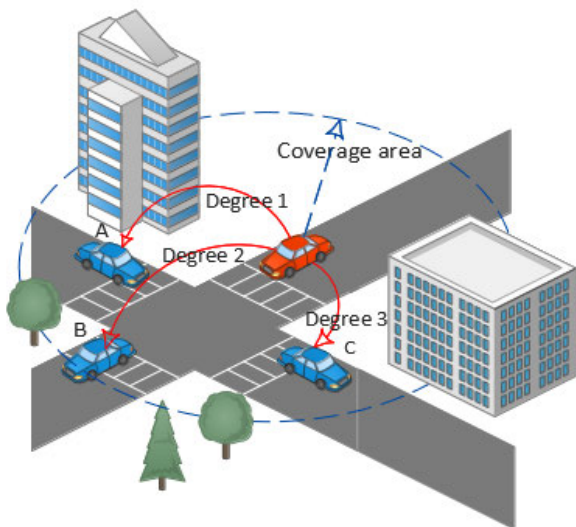


FIGURE 2. Node different degrees.

out of these nodes for the RSU, and three distinct degrees are allocated to them.

The RSUs proceed to gather node vectors in a variety of scenarios in order to build up training data for the learning algorithm. RSU sends all node vectors to the LPU, as well as all potential temporary relay sets, based on the furthest node in its transmission range. A potential temporary relay set is a relaying path that guarantees transmission of data from a source node to the rest of the nodes in a given circumstance without taking into account any optimization factors. The data acquired is used as input by the LPU for deep learning assessment. It designates a degree to each node in relation to every one of its neighbors as an output.

B. TRAINING AND LEARNING PHASE

The LPU requires a learning process in order to conduct nodal degree classification. The node vectors are collected

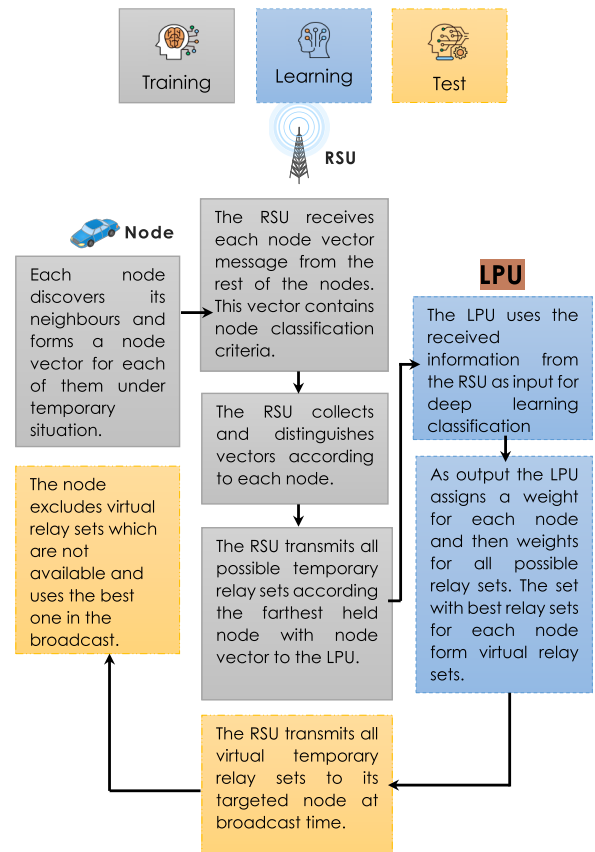


FIGURE 3. Learning process of the proposed scheme.

by base stations. The LPU then uses a large number of collected vector instances as input data for the construction of neural networks. The deep learning corpus extracted from the dataset [21] includes each of the vector components as a feature. The relevance of each feature value is determined depending on the current vehicular node situation as seen through the eyes of the relative base station at the moment of node vector collection.

Road side units deliver all potential temporary relay sets to the LPU, together with node vectors, that are instantly accessible for aggregation relative to the farthest node in their ranges. All previous classification data in the LPU are linked to the relay information held in the RSUs. The LPU creates virtual relays based on node degrees and temporary relay sets.

The LPU saves so much information on the temporary sets that were acquired before the broadcast. At the moment of broadcast, virtual relay sets are ranked according to their optimality and relayed back to the source node, which selects the best one based on its accessibility.

The LPU runs a supervised classifier and allocates degrees from 0.2 to 1.0 to each of the candidate nodes, using previously stored training knowledge from past scenarios. It sorts the relay sets based on these levels and gives back the outcomes. The source node processes the broadcast after checking the accessibility of the best selection. The different

steps of the training and learning phase are described in figure 3.

C. REINFORCEMENT LEARNING PHASE

1) THE VITERBI ALGORITHM

Maximum likelihood decoding of convolutional codes is performed by the Viterbi algorithm (VA). The most likely sequence is found by searching the trellis using this approach. A (n, k) convolutional code's trellis diagram contains 2^k branches emanating from and converging into each state [20].

If L is the length of the information sequence vector \mathbf{b} . The convolutional code encodes this sequence into a codeword \mathbf{c} with a length $(L + M)$. The received sequence is then denoted as \mathbf{y} with the same length as \mathbf{c} . Based on \mathbf{y} , a likelihood decoder selects a path denoted as $\hat{\mathbf{c}}$ with a maximization criterion given by the following function:

$$\log p(\mathbf{y}/\hat{\mathbf{c}}) = \sum_{i=0}^n \log p(\mathbf{y}_i/\hat{\mathbf{c}}_i) \tag{1}$$

where $\log p(\mathbf{y}/\hat{\mathbf{c}})$ is the metric of the path and $\log p(\mathbf{y}_i/\hat{\mathbf{c}}_i)$ represents the metric of the branch i in the path. The Viterbi algorithm matches the metrics of all paths ongoing each state and selects the path with the largest metric and saves its metric. The selected path is entitled the survivor. At the last phase, the survivor is decoded. The different steps of the VA are detailed in Algorithm 1.

Algorithm 1 The Viterbi Algorithm (VA)

Input: Set of all paths $(e_1 i_1(1), \dots, e_{M-1}, i_M(M))$
Output: Last survivor and convolutional codes(metrics)

1. **Begin**
2. At time $t = 0$ Calculate the branch metric for each state's single branch. For each state, save the branch (the survivor) and its metric.
3. **For** t from 1 to $L + m$ **do**
 - a. For each path inputting to each state calculate the path metric by summing the metric of the branch entering to that state with the metric of the connecting survivor at the previous trellis layer.
 - b. Select the path with the largest metric as survivor and save its metric for each state.
 - c. Eliminate all other paths.
4. **End For**
5. Save the last survivor and decode the sequence of codes.
6. **End**

2) APPLICATION OF THE ALGORITHM

We consider a general M -hops broadcasting area, where $M - 1$ relay sets in each hop are equally spaced as described in figure 4. This figure presents the network model with a source node S broadcasting a message. At each hop, a subset of relaying nodes are selected according to the method described in the subsection B.

The j^{th} hop contains a number of N_j candidate nodes. We assume that the forwarding radio within a hop can only

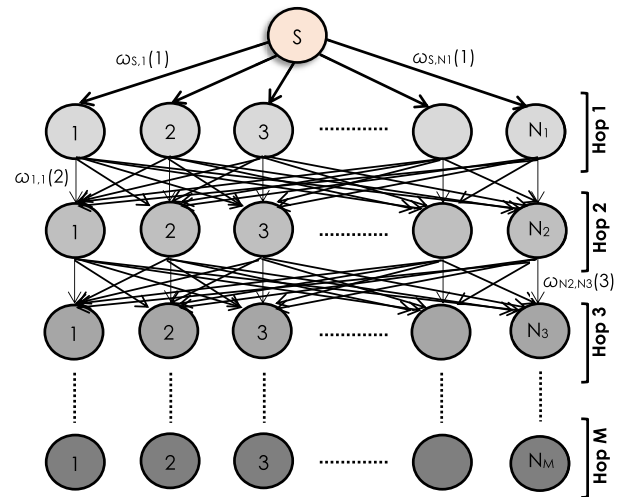


FIGURE 4. Network model: M -hops broadcasting area starting from the source node S with $M-1$ relay sets.

reach the adjacent hops. At each receiving node in a hop, a unique relay from the precedent hop is selected to relay the message. Each branch connecting the i^{th} to the j^{th} relay in the m^{th} hop is described with the corresponding SNR value $\omega_{i,j}(m)$ which is defined later in this section by equation (2).

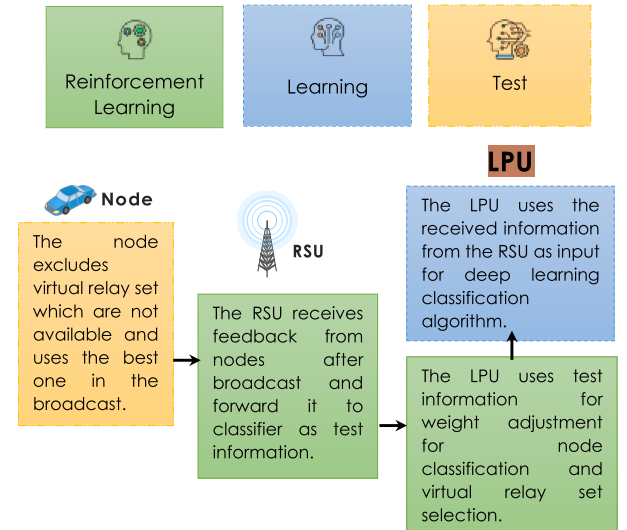


FIGURE 5. Reinforcement learning process of the proposed scheme.

As described in figure 5, the feedback required for the reinforcement procedure is dependent on the accessibility of the selected nodes. Therefore, we use the Signal to Noise Ratio (SNR) for the application of the VA. The metric of each branch is computed as the inverse of the SNR of the channel.

We also consider, $e_{i,j}(m)$ the branch connecting the relaying node i at the hop $m - 1$ to the relay node j at the hop m . Thence, the general representation of a path; e.g. a selected relay set, is $(e_{i_1,i_1}(1), \dots, e_{i_1,i_2}(m), \dots, e_{M-1,M}(M))$. Let the

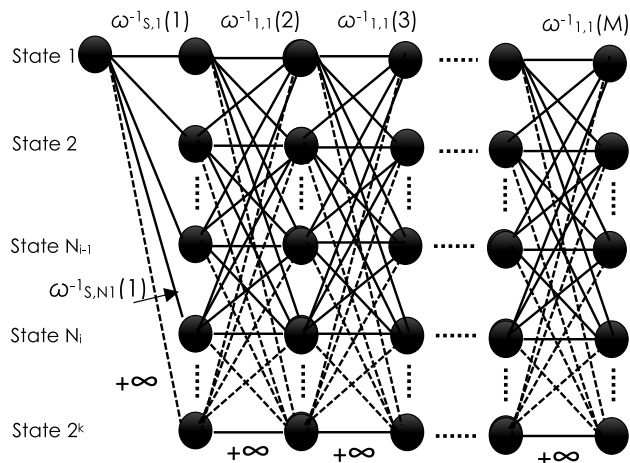


FIGURE 6. Abstraction trellis of virtual relay sets generation.

channel value corresponding to the branch $e_{i,j}(m)$ be denoted as $s_{i,j}(m)$. Then according to [22], the SNR is defined by:

$$\omega_{i,j}(m) = \bar{\omega} |s_{i,j}(m)|^2 \quad (2)$$

where $\bar{\omega}$ is the average instantaneous power at each node. Then the path's equivalent SNR, computed as the inverse of the SNR of the path is given by:

$$\omega_i \leq \min_{m-1,m} \omega_{i_{m-1},i_m}(m) \quad (3)$$

There are many other works that consider the SNR in the relay selection and which use other analytical models such as in [23] and [24].

A path is considered to be in outage and the involved nodes would have the node degree adapted in future selection if the path metric goes below a certain threshold value $\omega_i < \omega_{th}$. In this respect, the outage probability of a path is the probability of the path metric going under the threshold.

$$P_i = Pr \{ \omega_i < \omega_{th} \}$$

$$P_i \geq Pr \left\{ \min_{m-1,m} \omega_{i_{m-1},i_m}(m) < \omega_{th} \right\} \quad (4)$$

(4) demonstrates that the instantaneous SNR of the weakest branch among all branches on path i defines the outage probability of that path.

3) OUTAGE PROBABILITY ANALYSIS

The path metric computation is based on the SNR value. Figure 7 describes the average variation of the SNR in each node according to the distance. A path is considered to be in outage if the cumulated SNR is under a threshold value that we assume in our case $\omega_{th} = 1dB$ in the simulation of the proposed solution.

In our proposed reinforcement learning algorithm, the topology of the broadcasting area is first transformed into a coding trellis. The smallest value matching the condition term K such as $2^k \geq \max_{i=1,2,\dots,M} \{N_i\}$ in the resulting trellis code is chosen. The convolutional encoder's number of entries, k ,

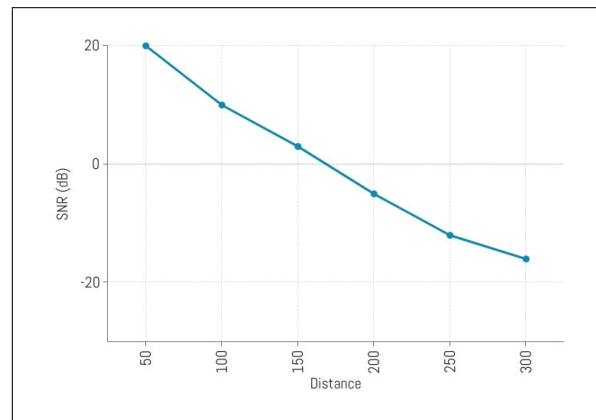


FIGURE 7. SNR variation.

is considered to be the minimum number such that 2^k equals or exceeds the highest number of branches reaching every node. Then, depending on these two components, a convolutional code is determined, and the matching trellis as created in figure 6. In this figure the topology of the broadcasting area modeled in 4 is transformed using the convolutional encoder to an abstraction trellis with 2^k states describing all possible combinations of virtual relay sets with corresponding path metric represented in the figure by the elementary branch metrics $\omega_{i,j}^{-1}(m)$. According to (3), the summation of the inverse instantaneous SNRs of all the branches along the path at large SNRs firmly outer bounds the inverse of the equivalent SNR of a path. As a result, the metric of a branch is chosen as the inverse of the instantaneous SNR of the relating channel. The cumulated metrics of all branches of a path are defined as path metric. The path metric at the relay r of the hop n is defined as:

$$\lambda(r, n) = \sum_{m=1}^n \omega_{i,j}^{-1}(m) \quad (5)$$

A relay of a hop r outside the window w (chosen according to the duration of the relay in previous broadcasting scenario, and whose feedback has been returned by the classification unit) will not be selected and will have a degree refining in the future. The relevance of this refining will be tested in the next relay selection procedure.

Figure 8 shows a measured outage probability variation with the SNR. The value of the probability gets higher in accordance with the decrease of the SNR. Specially, the lower is the SNR, the higher the value of the probability gets.

In figure 9, we analyze the outage probability fluctuation as a function of the vehicular traffic density and the SNR threshold value. Indeed, the number of vehicles in the broadcasting area affects the quality of radio links between the nodes, when the density increases, the probability of loss is likely to decrease. However, when the fixed threshold value is higher, the outage probability may increase because it would be hard to get SNR that is higher than the threshold.

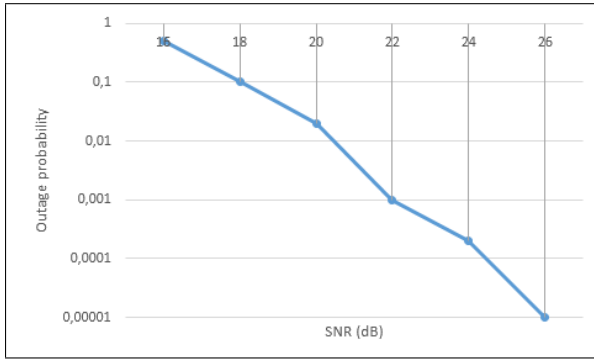


FIGURE 8. Outage probability variation.

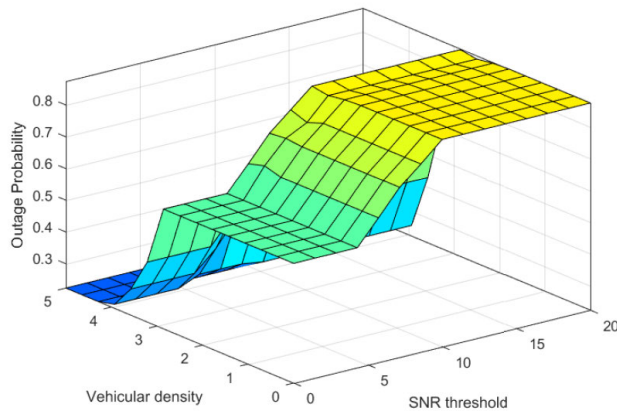


FIGURE 9. Analysis of the outage probability variation with the number of vehicles and the SNR threshold value.

IV. SIMULATION, RESULTS AND COMPARATIVE STUDY

A. DESCRIPTION OF THE SIMULATION SCENARIO

To evaluate the different broadcasting methodologies, we use the simulation model of [25] which is based on a grid-map scenario as modeled in figure 10, in which we examine a 1 kilometer perimeter with a vehicle count ranging from 100 to 600.

We used a shared MATLAB environment for simulation and learning to facilitate connections across different interfaces. From a mobility trace generated by the SUMO traffic simulator, the vehicle movement on the road is retrieved and merged into MATLAB. SUMO is an open source simulator that is used to create a simulated traffic scenario for a vehicular network. The SUMO interface application is employed to automatically construct itineraries for cars having mobility in a particular area of the map model. We collect mobility data traces that are then used to dynamically produce exchanged messages for our algorithm. The implementation of broadcasting and RL based optimization is carried out using MATLAB. Table 3 summarizes the simulation environment parameters.

An example of a broadcasting scenario is shown in figure 11. As soon as an accident occurs at a road junction, the red car transmits an alert notification message to all the other



FIGURE 10. Grid map like road model.

TABLE 3. MAC, PHY, environment and simulation parameters of the SUMO traffic simulator and MATLAB network simulator.

Parameter	Value	SUMO	MATLAB
Modulation scheme	802.11 PHY		✓
Frequency	2.4 GHz		✓
Data rate	6 Mbp		✓
Number of nodes	100-600	✓	✓
Radius of emission	250 m		✓
Maximum speed	16 m/s	✓	✓
Mac protocols	802.11 MAC		✓
Medium Access	CSMA/CA		✓
Broadcast message size	512 bytes		✓
Feedback message size	14 bytes		✓
Environment type	Urban	✓	
Roads area	1 km ²	✓	✓
Simulation duration	100 s		✓

vehicles in the vicinity. We tested the four different broadcasting techniques on the identical situation and compared their results in terms of efficiency in packet loss rate (PLR), success rate, saved rebroadcasts and average delay.

B. RESULTS AND COMPARATIVE STUDY

The novel relay selection introduced in this paper brings an improvement to the initial eKMPPR method. It combines reinforcement and supervised learning tools to increase the relevance of relaying set selection for more effective message broadcasting amongst vehicles. In the novel strategy, the same three selection factors as in eKMPPR were combined with four additional elements (vehicle altitude and velocity, location, and number of hops from the source node) to generate a seven parameters selection vector. This vector serves as the entry for a neural network, which will classify the nodes based on their ability to complete the transmission. Following the choice of the appropriate set, the shortlisted variables are refined using a reinforcement algorithm based on prior outcomes. Afterwards, we conduct a comparative study of the proposed

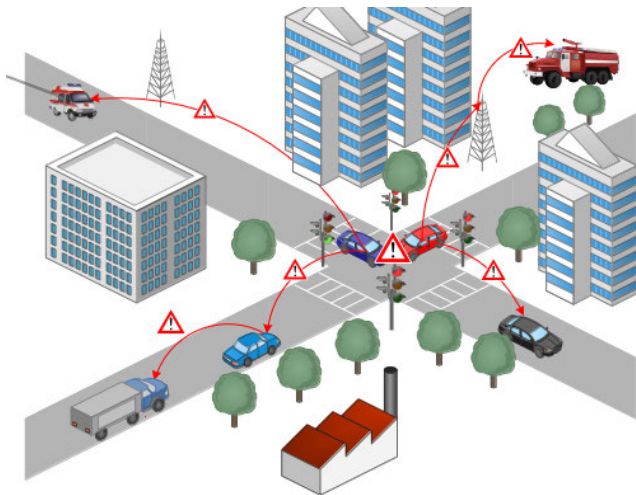


FIGURE 11. Example of broadcasting scenario: red vehicle is broadcasting an alert message to all vehicles.

techniques with other methods from literature to demonstrate its performance.

1) TAXONOMY OF THE COMPARED TECHNIQUES

For the comparative study, we choose three other methods: eKMPr [1], FBBPA [2] and ASPBT [26]. The eKMPr technique is a statistical approach for estimating vehicle mobility. It selects the optimum relay set for broadcasting the information in a specific road region using a fuzzy controller. The optimal relay node, according to this method, is the candidate with the best linear combination of coverage density, distance to the geometric median of its neighbors, and speed. FBBPA is a Fuzzy-based Beaconless Probabilistic Broadcasting Algorithms (FBBPA) dedicated to the notification of vehicles about an incident with less broadcasting messages. It is a receiver oriented broadcast suppression technique, where the relaying probability of a packet by a node is decided on the basis of their distance, angular orientation, mobility direction and buffer load delay.

ASPBT is an adaptive scheduled partitioning and broadcasting technique. This protocol dynamically adapts the number of partitions and the beacons periodicity to decrease the number of rebroadcasts. It uses the network density and the transmission schedule for each partition estimated by lion optimization algorithm (LOA). LOA is an optimization bio-inspired algorithm mitigating the behavior of lions. Its corporate and solitary behaviors such as prey capturing, roaming, mating and defense help to identify the optimal partition to broadcast the emergency messages first. Table 4 describes the taxonomy of the aforementioned three approaches compared to the technique proposed in this paper.

2) PERFORMANCE EVALUATION

Figure 12 describes the accessibility (the success rate) analysis of the different techniques by varying the traffic density. According to the graph, we can notice that the broadcasting

TABLE 4. Taxonomy of the studied broadcasting schemes.

Broadcasting approach	Type	Techniques
eKMPr [1]	Statistical	Fuzzy logic and mobility prediction
FBBPA [2]	Probabilistic	Fuzzy controller
ASPBT [26]	Bio-inspired	LOA
Proposed	Hybrid	Deep and reinforcement learning

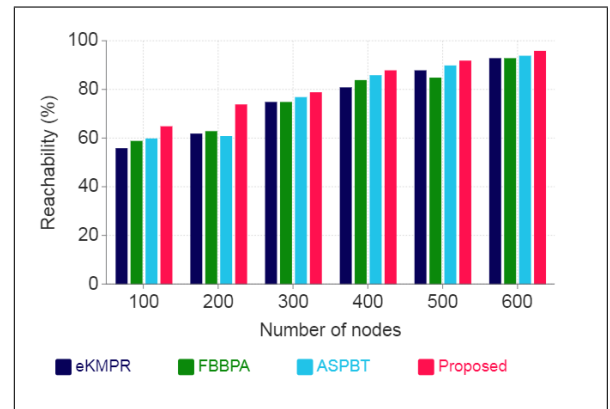


FIGURE 12. Success rate.

based on the proposed relay selection solution outperforms the existing protocols because our method uses the vehicular neighboring density as a selection criterion. This ensures that our proposed solution is applicable in both sparse and dense traffic scenarios, in particular, for the applications that require a high success rate such as emergency message delivery systems.

In figure 13, we analyze the Packet Loss Rate (PLR). For low traffic density, we note that all the compared techniques have a high rate of packet loss, which is due to the frequent link breakage. When the number of nodes increases, the rate of data loss becomes lower and the proposed technique has the lowest values owing to the consideration of the received transmission power in the classification of nodes and the SNR rate in refining the selection degree of each node. The PLR is a crucial performance parameter especially for multimedia applications requiring a high delivery rate.

By maintaining a balance between the success rate and the latency, an effective broadcast technique should be able to make the warning message reach all of the drivers. In most cases, there is a trade-off between saving rebroadcast and accessibility. The urge to conserve network resources increases the likelihood that messages will not reach all of the targeted automobiles. As a result, the primary goal is to minimize the replayed packet and increase its chances of being received. As the number of vehicles on the road grows, so does the number of retransmissions; thereby, vehicles within transmission range of each other are more likely to be recognized. Unnecessary retransmission of the same packet is not required in this instance. This supports the claim that the number of preserved rebroadcasts grows with traffic density.

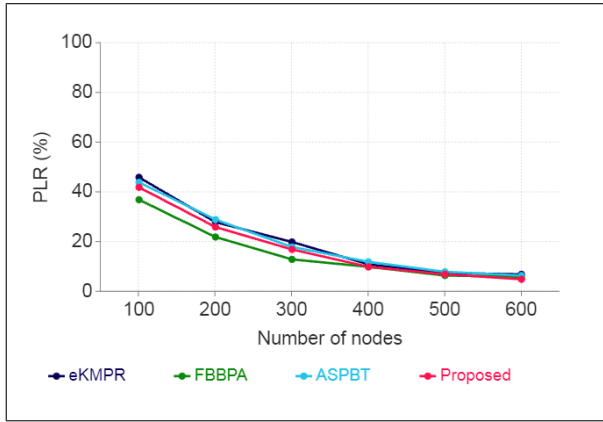


FIGURE 13. Packet loss rate.

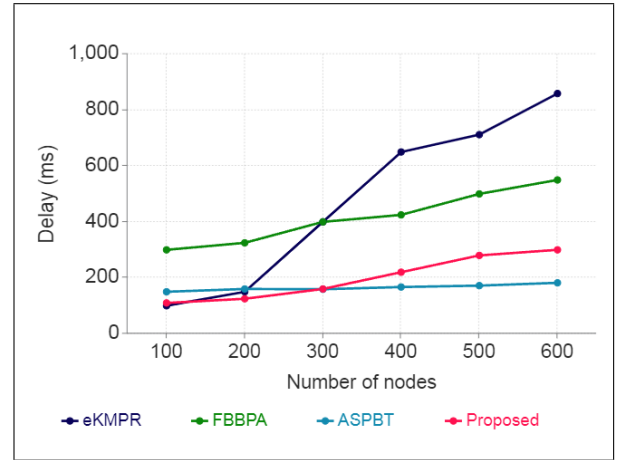


FIGURE 16. Analysis of the average delay.

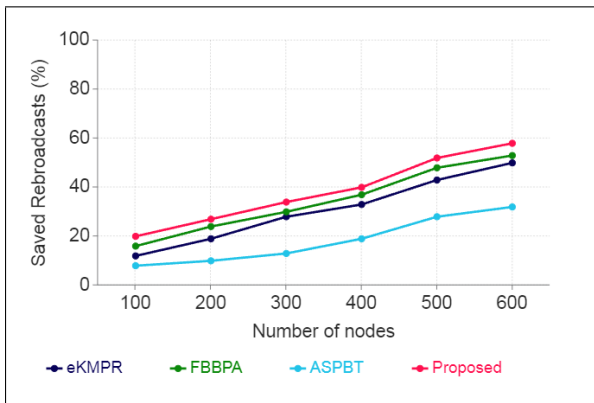


FIGURE 14. Saved rebroadcasts.

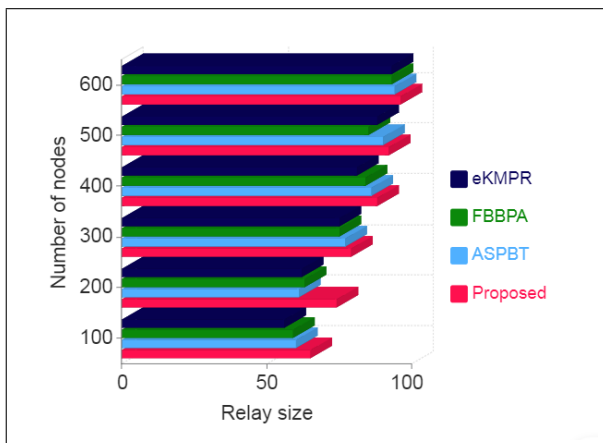


FIGURE 15. Forwarding set size.

Compared to the other protocols, the proposed solution has the best performance in terms of saved rebroadcasts analyzed in figure 14. The rate of its overperformance increases with the increment of the vehicle density and that is due to the reinforcement technique. Indeed, reinforcement algorithm allows to refine the selection of the vehicles to forward the message

which helps to reduce the relay set size (shown in figure 15) and thus increase the number of saved rebroadcasts.

Figure 16 presents the results obtained from measuring the average delay achieved by each one of the compared techniques in delivering the broadcasting message. The measured values are highly dissimilar especially for high density networks. The proposed solution achieves the lowest values with a small number of nodes in the network. However, these values increase when the network density augments. These results prove that our method fulfills maximum QoS requirement in terms of delay, such as indicated in [27] and [28]. As shown in figure 16, this requirement is fulfilled even for 150 vehicles. Then it slowly increases when the node density exceeds 200 vehicles. Nevertheless, the maximum value is stills better than the highest levels reached by the other methods.

V. CONCLUSION

In this paper, we introduced a novel hybrid relay selection technique for broadcasting in vehicular networks based on a reinforcement learning method. In addition to the artificial neural network applied for the classification of nodes to forward the message, we adopted the Viterbi algorithm as a reinforcement phase to refine the first classification. To apply this algorithm, we used the SNR for determining the outage probability of the radio link at each node. The value of this probability, which depends on the SNR threshold value, allows to rectify the classification degree of nodes in future scenarios.

To test the performance of the proposed solution, we adopt a grid map scenario with varied traffic density. Subsequently, we analyze and compare the simulation results with three other broadcasting techniques. The comparison was based on different parameters such as the success rate, the data loss, the saved rebroadcasts and the delay. The obtained results and the comparative study prove the overperformance of the proposed technique owing to its hybrid nature which ombines deep learning and reinforcement learning.

REFERENCES

- [1] A. Mchergui, T. Moulahi, M. T. B. Othman, and S. Nasri, "QoS-aware broadcasting in VANETs based on fuzzy logic and enhanced kinetic multipoint relay," *Int. J. Commun. Syst.*, vol. 33, no. 5, p. e4281, 2020, doi: 10.1002/dac.4281.
- [2] A. Srivastava, A. Prakash, and R. Tripathi, "Fuzzy-based beaconless probabilistic broadcasting for information dissemination in urban VANET," *Ad Hoc Netw.*, vol. 108, Nov. 2020, Art. no. 102285.
- [3] A. Mchergui, T. Moulahi, B. Alaya, and S. Nasri, "A survey and comparative study of QoS aware broadcasting techniques in VANET," *Telecommun. Syst.*, vol. 66, pp. 253–281, Feb. 2017, doi: 10.1007/s11235-017-0280-9.
- [4] K. Maha, H. Otrok, H. Barada, M. Al-Qutayri, and Y. Al-Hammadi, "A cluster-based QoS-OLSR protocol for urban vehicular ad hoc networks," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, 2018, pp. 554–559.
- [5] T. Moulahi, H. Guyennet, M. Lehsaini, and S. Nasri, "An energy aware MPR-based broadcasting algorithms for wireless sensor networks," in *Proc. Int. Conf. Sel. Topics Mobile Wireless Netw. (iCOST)*, Oct. 2011, pp. 13–18.
- [6] A. Mchergui, T. Moulahi, and S. Nasri, "Relay selection based on deep learning for broadcasting in VANET," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2019, pp. 865–870.
- [7] S. Agrawal, N. Tyagi, A. Iqbal, and R. S. Rao, "An intelligent greedy position-based multi-hop routing algorithm for next-hop node selection in VANETs," *Proc. Nat. Acad. Sci., India A, Phys. Sci.*, vol. 90, pp. 39–47, Sep. 2018.
- [8] W. Liu, X. Tang, S. Jia, and J. Pu, "Safety message dissemination using edge computing in heterogeneous VANETs," in *Proc. IEEE 27th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2018, pp. 1276–1281.
- [9] R. Hajlaoui, H. Guyennet, and T. Moulahi, "A survey on heuristic-based routing methods in vehicular ad-hoc network: Technical challenges and future trends," *IEEE Sensors J.*, vol. 16, no. 17, pp. 6782–6792, Sep. 2016.
- [10] M. Jafer, M. A. Khan, S. U. Rehman, and T. A. Zia, "Broadcasting under highway environment in VANETs using genetic algorithm," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–5.
- [11] Y. Zhou, "A fuzzy-rule based data delivery scheme in VANETs with intelligent speed prediction and relay selection," *Wireless Commun. Mobile Comput.*, vol. 2018, Aug. 2018, Art. no. 7637059.
- [12] O. Rehman and M. Ould-Khaoua, "A hybrid relay node selection scheme for message dissemination in VANETs," *Future Gener. Comput. Syst.*, vol. 93, pp. 1–17, Apr. 2019.
- [13] K. N. Qureshi, M. M. Idrees, J. Lloret, and I. Bosch, "Self-assessment based clustering data dissemination for sparse and dense traffic conditions for internet of vehicles," *IEEE Access*, vol. 8, pp. 10363–10372, 2020.
- [14] K. N. Qureshi, A. H. Abdullah, and J. Lloret, "Road perception based geographical routing protocol for vehicular ad hoc networks," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 2, Feb. 2016, Art. no. 2617480.
- [15] W. Song, F. Zeng, J. Hu, Z. Wang, and X. Mao, "An unsupervised-learning-based method for multi-hop wireless broadcast relay selection in urban vehicular networks," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–5.
- [16] Z. Chen, T. Lin, and C. Wu, "Decentralized learning-based relay assignment for cooperative communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 813–826, Feb. 2016.
- [17] Q. You, Y. Li, Z. Chen, and M. S. Rahman, "A simple near-optimal path selection scheme for multi-hop wireless relay networks based on Viterbi algorithm," *Trans. Emerg. Telecommun. Technol.*, vol. 27, no. 10, pp. 1294–1307, Oct. 2016, doi: 10.1002/ett.2849.
- [18] A. Gulati, G. S. Aujla, R. Chaudhary, N. Kumar, and M. S. Obaidat, "Deep learning-based content centric data dissemination scheme for internet of vehicles," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
- [19] R. Tian, Z. Jiao, G. Bian, Z. Huang, and Y. Hou, "A social-based data forwarding mechanism for V2 V communication in VANETs," in *Proc. 10th Int. Conf. Commun. Netw. China (ChinaCom)*, Aug. 2015, pp. 595–599.
- [20] Q. You, Y. Li, M. S. Rahman, and Z. Chen, "A near optimal Viterbi algorithm based routing scheme for multi-hop relay networks," in *Proc. Austral. Commun. Theory Workshop (AusCTW)*, Jan. 2012, pp. 150–153.
- [21] University of South Carolina. (2017). *Dept of computer Science and Engineering, USA*. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/DSRC_Vehicle_Communications
- [22] M. Obaidat, I. Shahwan, A. Hassebo, S. Obeidat, M. Ali, and M. Khodjaeva, "SNR-based early warning message scheme for VANETs," *J. Mobile Multimedia*, vol. 15, pp. 163–190, Feb. 2020.
- [23] M. A. Karabulut, A. F. M. S. Shah, and H. Ilhan, "OEC-MAC: A novel OFDMA based efficient cooperative MAC protocol for VANETs," *IEEE Access*, vol. 8, pp. 94665–94677, 2020.
- [24] A. F. M. S. Shah, H. Ilhan, and U. Tureli, "RECV-MAC: A novel reliable and efficient cooperative MAC protocol for VANETs," *IET Commun.*, vol. 13, no. 16, pp. 2541–2549, 2019.
- [25] A. Mchergui, T. Moulahi, and S. E. Khediri, "On the comparison of broadcasting techniques in vehicular ad hoc networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2021, pp. 1599–1603.
- [26] M. Selvi and B. Ramakrishnan, "Lion optimization algorithm (LOA)-based reliable emergency message broadcasting system in VANET," *Soft Comput.*, vol. 24, no. 14, pp. 10415–10432, Jul. 2020.
- [27] A. F. M. S. Shah, M. A. Karabulut, H. Ilhan, and U. Tureli, "Performance optimization of cluster-based MAC protocol for VANETs," *IEEE Access*, vol. 8, pp. 167731–167738, 2020.
- [28] A. F. M. S. Shah, H. Ilhan, and U. Tureli, "CB-MAC: A novel cluster-based MAC protocol for VANETs," *IET Intell. Transp. Syst.*, vol. 13, no. 4, pp. 587–595, 2019.



ABIR MCHERGUI received the Ph.D. degree from the National School of Engineer of Sfax (ENIS), Sfax University, Tunisia, in 2020. She is currently an Assistant Professor with the Computer Sciences Department, Higher Institute of Management, Gabes University, Tunisia. Her research interests include vehicular *ad hoc* networks, cybersecurity, and pervasive computing.



TAREK MOULAH received the Ph.D. degree from the University of Franche-Comté, Besançon, France, in March 2015, and the Sfax National School of Engineering, Tunisia. He is currently an Assistant Professor with the Mathematics and Computer Science Department, Faculty of Science and Technology of Sidi Bouzid (FSTSB), University of Kairouan, Tunisia, and with the Department of Information Technology, College of Computer, Qassim University, Saudi Arabia. His research interests include wireless sensor networks, vehicular *ad hoc* networks (VANET), the Internet of Things (IoT), and blockchain. He received the 2019 IEEE Sensors Council Sensors Journal Best Paper Runner-Up Award.

• • •