


Article

Machine Learning-Based Cooperative Spectrum Sensing in Dynamic Segmentation Enabled Cognitive Radio Vehicular Network

Mohammad Asif Hossain ¹, Rafidah Md Noor ^{1,2,*}, Kok-Lim Alvin Yau ^{3,*}, Saaidal Razalli Azzuhri ¹, Muhammad Reza Z'aba ¹, Ismail Ahmedy ¹ and Mohammad Reza Jabbarpour ⁴

- ¹ Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia; asif.ewu@gmail.com (M.A.H.); saaidal@um.edu.my (S.R.A.); reza.zaba@um.edu.my (M.R.Z.); ismailahmedy@um.edu.my (I.A.)
- ² Centre for Mobile Cloud Computing, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia
- ³ School of Science and Technology, Sunway University, Selangor 47500, Malaysia
- ⁴ Department of Information and Communications Technology, Niroo Research Institute, Tehran 1468613113, Iran; mrjabbarpour@nri.ac.ir
- * Correspondence: fidah@um.edu.my (R.M.N.); koklimy@sunway.edu.my (K.-L.A.Y.)

Abstract: A vehicle ad hoc network (VANET) is a solution for road safety, congestion management, and infotainment services. Integration of cognitive radio (CR), known as CR-VANET, is needed to solve the spectrum scarcity problems of VANET. Several research efforts have addressed the concerns of CR-VANET. However, more reliable, robust, and faster spectrum sensing is still a challenge. A novel segment-based CR-VANET (Seg-CR-VANET) architecture is therefore proposed in this paper. Roads are divided equally into segments, and they are sub-segmented based on the probability value. Individual vehicles or secondary users produce local sensing results by choosing an optimal spectrum sensing (SS) technique using a hybrid machine learning algorithm that includes fuzzy and naïve Bayes algorithms. We used dynamic threshold values for the sensing techniques. In this proposed cooperative SS, the segment spectrum agent (SSA) made the global decision using the tri-agent reinforcement learning (TA-RL) algorithm. Three environments (network, signal, and vehicle) are learned by this proposed algorithm to determine primary (licensed) users' activities. The simulation results indicate that, compared to current works, the proposed Seg-CR-VANET produces better results in spectrum sensing.

Keywords: spectrum sensing; cognitive radio; VANET; tri-agent reinforcement learning; machine learning



Citation: Hossain, M.A.; Md Noor, R.; Yau, K.-L.A.; Azzuhri, S.R.; Z'aba, M.R.; Ahmedy, I.; Jabbarpour, M.R. Machine Learning-Based Cooperative Spectrum Sensing in Dynamic Segmentation Enabled Cognitive Radio Vehicular Network. *Energies* **2021**, *14*, 1169. <https://doi.org/10.3390/en14041169>

Academic Editor: Kim Sunwoo

Received: 27 December 2020

Accepted: 17 February 2021

Published: 22 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A vehicular ad hoc network (VANET) enables transmission among smart vehicles for various purposes, including road safety, entertainment, congestion control, vehicle safety, etc. [1]. However, the vehicular network is in general composed of a large number of vehicles. Thus, spectrum availability for vehicles at all times becomes challenging [2]. A huge amount of data sharing is required for the implementation of VANET. According to Intel, a single smart car will share about 4 terabytes of data a day in the near future [3]. The VANET supports the IEEE 802.11p protocol, also known as dedicated short-range communications (DSRC), which has a 75 MHz bandwidth within the 5.85 GHz–5.925 GHz frequency spectrum. However, this spectrum range is not adequate for this massive volume of data exchange [4]. Therefore, a wider range is required to facilitate such high-volume data sharing.

Cognitive radio (CR) is a smart radio designed to utilize unlicensed bands known as spectrum holes [5]. As spectrum scarcity is addressed as the major problem for implementing VANET in recent times, CR becomes an emerging technology. In CR, unlicensed

users or secondary users (SUs) sense the vacant licensed spectrum used by the primary users (PUs). The spectrum's efficient utilization in the automotive environment integrates CR with VANET, which is termed CR-VANET. The vehicles are the SUs that include the CR technique by which they can sense the signals and utilize the licensed spectrum if the unlicensed spectrum is unavailable.

Spectrum sensing is the vital process of CR that identifies spectrum availability by sensing PU activities [6]. Through spectrum sensing, the available channels can be determined, which are further used for data transmission. Spectrum sensing can be made individually or cooperatively. Spectrum sensing is composed of conventional techniques, among which energy detection, cyclostationary feature detection, and matched filter detection are used most for sensing. The spectrum sensing conducted by a single vehicle is nevertheless vulnerable to multiple challenges, including multipath fading, shadowing, noise uncertainty, and hidden PU problems [7]. Moreover, vehicles are not familiar with PU activities in advance, and the PU system is not part of the sensing process. Thus, the sensing efficiency is not optimum while the individual vehicle performs, resulting in interference with the PU system. Furthermore, in a non-cooperating sensing system, it is difficult to model the PU activity pattern correctly. To mitigate these problems, cooperative spectrum sensing (CSS) has been introduced in the literature.

In CR-VANET, CSS allows all the vehicles to perform sensing at the same time, and then it reports to the fusion center or to themselves to make the global or the final decision. Once the sensing is performed, the channels are allocated by the RSU (roadside unit) and used for transmission. CSS is more accurate than individual sensing, since the global decision can be made based on multiple sensing reports [8]. In CSS, from the sensing reports of local information, the global decision is taken using fusion methods that can be the OR/AND rule and other methods. In CSS, the spatial diversity of the SUs is achieved. This spatial diversity, which is known as the cooperative gain, increases the spectrum sensing efficiency. This gain is also seen from a hardware perspective of the detector. The signal-to-noise ratio (SNR) of the primary signal can be very low, and identifying this becomes a challenging task due to multipath fading and shadowing. Since the sensitivity of recipients means the potential to detect the weak signals, it introduces a strict sensitivity criterion that dramatically raises the sophistication of the program and the related hardware costs. It is also difficult to increase detection efficiency by raising the sensitivity when the SNR of PU signals is below a certain threshold called the SNR wall. Fortunately, CSS will considerably alleviate the sensitivity criteria and hardware constraint problems. It can solve output deterioration by multipath fading and shadowing [9].

While cooperative gains can be obtained by cooperative sensing, as discussed previously, several variables can restrict the achievable cooperative gain. For example, when CR users with the same obstacle are in spatially associated shadows, their observations are correlated with each other. More collaborative spatially correlated CR users can be harmful to detection performance [6,10]. This raises the challenge of the user selection for CSS. The CR-VANET architecture is partitioned using segments or clusters to perform CSS [11]. The formation of clusters or segmentation of road lanes improves the sensing performance and results in a higher data delivery ratio. Moreover, cluster formation enables mobility management support by grouping vehicles with similar behavior, such as speed, direction, etc. The cluster head (CH) is selected to carry the sensing report to the fusion center in the clustering approach. There are additional time and resources involved in determining the CH. Moreover, CSS can also lead to overhead cooperation, in addition to gain-limiting factors. Overhead refers to any additional time, delay, energy, and operation dedicated to cooperative sensing compared to the individual (non-cooperative) spectrum sensing situation [9].

There are three types of CSS. They are centralized, distributed, and relay-assisted [12]. Individual vehicles sense the spectrum in the centralized CSS and then send local sensing results to a central node termed a fusion center (FC). An FC can be a roadside unit (RSU), a cluster head (CH), and others. The FC is responsible for deciding on the state of occupation

of the PU. The FC makes the global sensing reports with the individual vehicles' local sensing reports by using hard or soft fusion or other rules.

Although many research efforts have been carried over CR-VANET, the performance is still affected by many factors. Those critical factors are the following:

- Imperfect spectrum sensing without proper environmental knowledge
- High spectrum sensing errors due to lack of global decision-making
- CSS susceptibility to the overhead issues
- Considerable transmission delay due to improper network management and channel assignment

Due to the above factors, CR-VANET performance is degraded. Prior research works have either focused on VANET improvement or CR improvement, which are insufficient to handle all the above issues. A combined approach is necessary to enhance the overall performance of spectrum sensing of CR-VANET.

1.1. Motivation and Contributions

In recent times, road safety has become a significant concern due to the growth of smart vehicles. Hybrid CR-VANET architectures have been widely studied to provide better efficiency for data transmission. The primary motivation of this research work is the unsolved research problems that exists in those prior works. The main issue is that the combination of spectrum sensing and proper network management has not been focused. Spectrum availability detection alone cannot assure appropriate data transmission efficiency since it also depends on spectrum management and network management. This major issue is the primary motivation for this work.

Moreover, existing works have several limitations in spectrum sensing architecture for CR-VANET. For instance, several critical CR network parameters are not being considered, using a fixed threshold value for sensing, using a single sensing technique in all scenarios [6,13,14]. We then formulate our objectives to resolve this unsolved issue. This proposed architecture's primary research objective is to design spectrum sensing and make accurate global decisions using an advanced tri-agent reinforcement learning algorithm for a dynamic segmentation enabled CR-VANET.

Several advantages can be achieved by carrying out sub-segmentation of the road segment. They include the following:

- The involvement of the sub-segmentation process makes network management easy and improves data transmission efficacy.
- Cooperation overhead is reduced.
- Unlike the clustering system, there is no need for CH selection, which saves time, as selecting CH takes additional time and creates extra delays, which degrade the network's performance.
- By carrying out sub-segmentation, synchronization among the SUs is possible without added complexity.
- It helps to solve bandwidth requirement problems. A large amount of bandwidth is needed for sending the sensing reports by SUs to the FC.

Some of the significant contributions of this paper are as follows:

- A novel segment-based CR-VANET (Seg-CR-VANET) architecture is designed by segmenting the road lanes into equal distances. Segments are managed continuously by a probability-based sub-segment management approach. Each segment is further divided into sub-segments based on speed, segment size, and node degree. The proposed work improves VANET in two aspects, namely (i) accurate spectrum sensing and decision-making, and (ii) stable network management.
- Spectrum sensing accuracy is improved by a dynamically selecting sensing technique based on signal to noise ratio (SNR) and noise power. Each vehicle first selects an optimal sensing technique for the current situation by using the fuzzy-naïve Bayes algorithm.

- A dynamic threshold value is introduced for spectrum sensing. This novel solution assures that adaptive sensing results in accurate sensing reports.
- Segment spectrum agents (SSAs) then make a global decision on all vehicles' collected sensing reports. To avoid wrong decisions, SSA uses a novel tri-agent reinforcement learning (TA-RL) algorithm that learns three environments (signal, network, and vehicle behavior) by three agents. If channels are available, RSU allocates the available channels to the vehicles.

1.2. Paper Layout

The rest of this paper is organized as follows: Section 2 surveys significant research works carried out on CR-VANET. The section also includes the primary and sub-research problems that are solved in this paper. Section 3 details the proposed architecture with the proposed algorithms. Section 4 discusses the simulation set up and the theoretical comparison with prior works. Section 5 discusses the obtained results with comparative analysis. In Section 6, we conclude our contributions and highlight future research directions.

2. Related Works

This section reviews the current research work and summarizes the research gap, and it is focused on improving the efficiency of CR-VANET.

Spectrum sensing and management, which are the crucial processes of CR-VANET, are widely studied in the literature. A vehicular cognitive small cell network is presented to solve spectrum sharing problems using a game-theoretic approach [15]. A Bertrand competition model is proposed for optimal utilization of spectrum efficiency using a genetic simulated annealing algorithm. This algorithm determines the Nash equilibrium, and then the spectrum price is optimized using mutation and crossover operators. A genetic simulated annealing algorithm's high time complexity increases spectrum sensing and assignment time, which is not suitable for dynamic vehicular networks.

Cooperative spectrum sensing is performed by the RL algorithm with dynamic spectrum access (DSA) [14]. The problem in this work is that it is not suitable for dense vehicle environments, including urban scenarios, since the segments are maintained at a fixed size. Further, RL learns only the channel environment, and the decision is made based on a static threshold value, which is not suitable for a dynamic vehicular environment.

A regional cluster-based approach is utilized by using a linear programming model [16]. This work fails to balance the density among various RSUs, and the involvement of linear programming introduces difficulties in defining the objective function. Moreover, channel estimation is performed a priori, which is unsuitable for VANET. In CSS, a binary decision-making approach is used on the aggregated local sensing reports [4]. All local reports are generated using an energy detection method with a static threshold value and are ineffective in a dynamic environment. Binary decision-making is also inaccurate since it only relies on the collected reports without knowledge of the channel and the current network environment.

A channel slotted contention protocol is designed with random single-channel sensing, slotted contention, and aggregation for high-density vehicle scenarios [17]. First, the vehicle selects one channel at random from all the channels, and then the sensing information is sent to RSU for decision-making. The OR rule is then applied for a decision and the PU signals are used and the data transmission is performed. A cooperative mechanism is presented with adjustable double thresholds that aim to reduce false alarm probability [18]. For sensing, the energy detection method is used, and it determines the decision-variant and independent threshold. Here the sensing decision is given using the OR fusion rule defined from the threshold value. The threshold value is determined concerning the probability of the detection of the signals.

CSS can also be performed using historical data information [19]. It focuses on the development of an attack model that has two different attacks, namely selfish and malicious. Based on the sensing results, it can differentiate the sensing of attackers and

normal SUs. The speed adjustment is also performed for the nodes in the network, which is suitable for highway scenarios. Hybrid cooperative spectrum sensing is performed by spatial–temporal correlation [20]. Here the historical sensing information is collected by the SUs that define temporal correlation, and it is combined with spatial correlation. The proposed scheme uses two steps: user selection and CSS. With the combination of spatial and temporal measurements, the optimal probability is predicted. With these methods, historical data collection and maintenance is a challenging issue, and it can be affected by many environmental factors such as noise.

In terms of spectrum allocation, quality of service (QoS) provisioning is the main focus [21]. Channel allocation is incorporated using the semi-Markov decision process (SMDP). The collision probability is determined and followed by the vehicle model. Here, the QoS is one of the significant constraints that is attained from proper channel assignment. Channel assignment follows the QoS factor but is not clearly described. Consideration of all QoS factors relatively increases delay.

From these recent works, two main problems were formulated. Those problems are the following:

Lack of Spectrum Management: The vehicles as SUs in CR-VANET use the conventional spectrum sensing technique, which is subject to limitations. Hence, each technique is suitable only for environmental features. Therefore, the use of a spectrum sensing technique for all conditions leads to a degraded spectrum sensing decision that increases the false alarm rate.

Lack of Network Management: In CR-VANET, the road lane is segmented concerning its length. The road lane segmentation is fixed and uneven, so vehicle traffic density differs from time to time, impacting abnormal detection of the spectrum and channel allocation.

3. Proposed CR-VANET Model

In this section, we explain the proposed work in detail with the proposed algorithms.

3.1. Network Model and Assumptions

The proposed CR-VANET model is comprised of N number of vehicles as V_1, V_2, \dots, V_N . From here, SUs and vehicles represent the same thing since the vehicles are the SUs in our work. The SUs sense the vacant spectrum of PUs. The network model also consists of RSU. The proposed network is disjointed into multiple segments as S_1, S_2, \dots, S_s . In each segment, we introduce a new segment sensing agent (SSA), i.e., the overall network has s number of SSAs for performing sensing and decision-making. Spectrum sensing and decision-making are managed by SSAs and SUs cooperatively.

The overall architecture is represented in Figure 1. As shown in the network model, we performed spectrum sensing, decision-making, and channel allocation.

At the same time, segment management and channel allocation management are carried by RSUs. Each entity in the designed network architecture has its work process. The entities and their responsibilities are illustrated in Figure 2 and described here:

- (1) The vehicles decide the sensing technique and sense the signal, and then report to the spectrum agent.
- (2) The SSA collects local sensing information and makes a decision using reinforcement learning and reports to CR-RSU.
- (3) The CR-RSU manages the segments, and it assigns channels to vehicles.

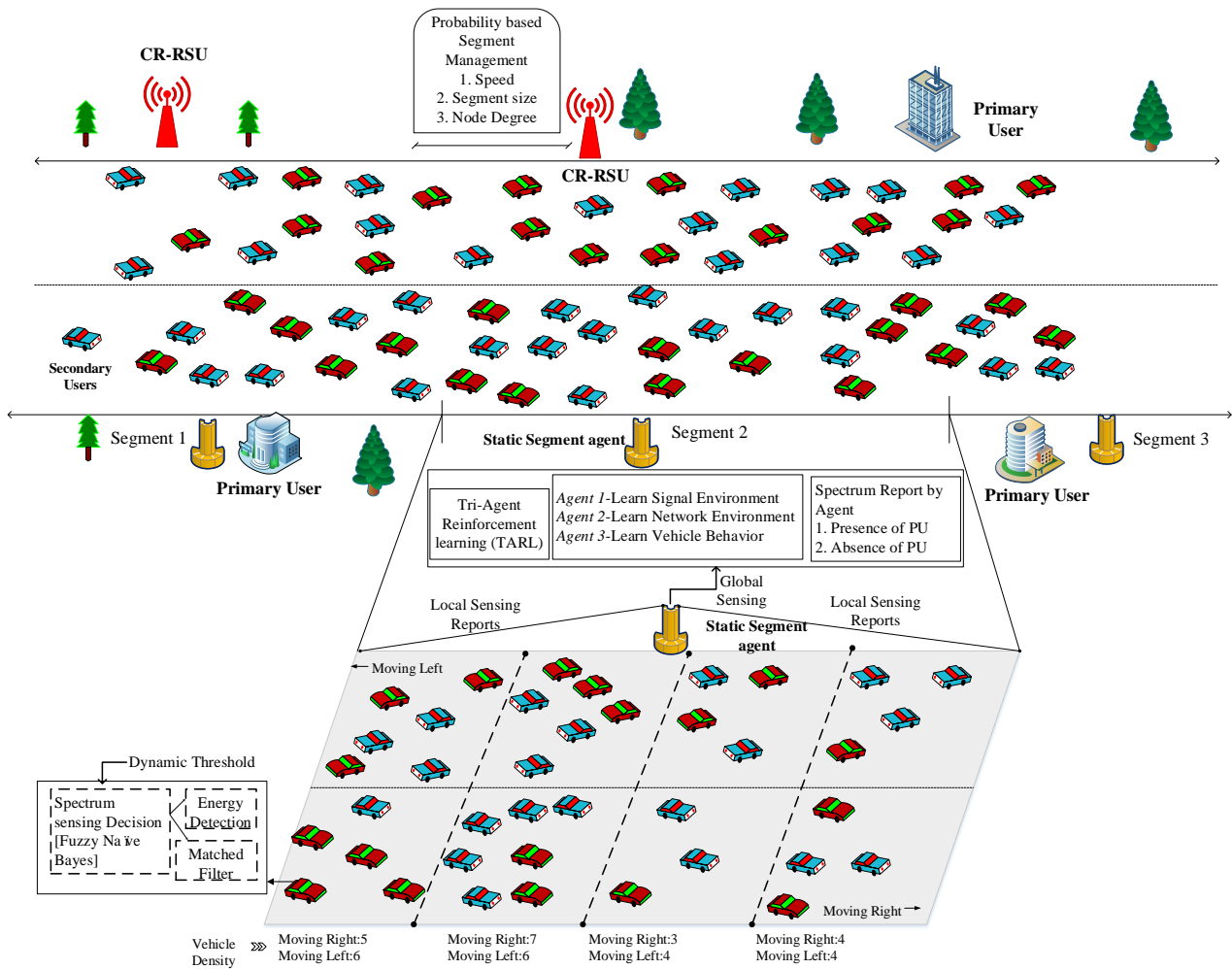


Figure 1. Proposed segmented cognitive radio vehicle ad hoc network (CR-VANET) model.

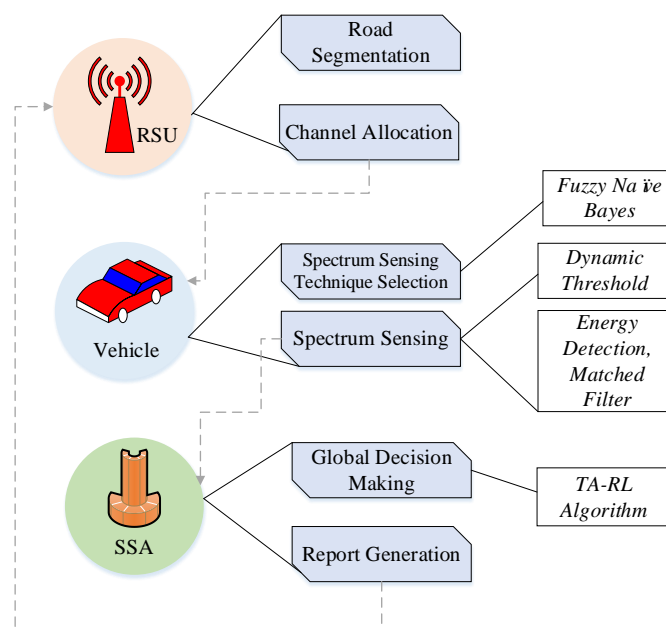


Figure 2. Entities of the proposed architecture and their responsibilities.

3.2. Dynamic Road Segmentation

In this proposed work, the network is considered with segments of equal length. However, it is not realistic that vehicle density will be the same all the time. Each time the segment's vehicle density varies, there may be a high density in a particular segment, leading to ineffectual network management. This paper proposes dynamic segment management to manage density variation through a probability value update. Here, the segment is further sub-segmented based on the probability value.

The probability value is formulated using multiple criteria including speed of vehicles (\mathbb{S}), segment size (ϕ), and node degree (Θ). The probability value for sub-segmentation can be obtained by normalizing the following value:

$$\Psi = \sum \mathbb{S}, \phi, \Theta \quad (1)$$

Let us assume that a segment can be divided into a maximum n_{sub} number of sub-segments. We assume that there is the same number of sensors available, i.e., every segment has a sensor. Each sensor is used to determine the speed of the vehicle, the number of the vehicle, and the direction.

Here, the speed value is computed as the average speed of all vehicles in the sub-segment, i.e.,

$$\mathbb{S} = \frac{\sum \mathbb{S}_1, \mathbb{S}_2, \dots, \mathbb{S}_m}{m} \quad (2)$$

where m is the number of vehicles in that sub-segment. Ψ is normalized (denoted as Ψ_{norm}) to compute the probability value. If the probability value is higher than the threshold value (Th_{seg}), the sub-segmentation process is initiated and enabled, i.e.,

$$\Psi_{norm} > Th_{seg} \quad (3)$$

Whenever any sub-segment's probability value Ψ is greater than the threshold value (Th_{seg}), that sub-segment will be treated as a separate sub-segment. Let us assume our operational segment is the i^{th} segment. In this segment, for example, if only the j^{th} sub-segment follows $\Psi_j > Th_{seg}$, then j^{th} sub-segment will be considered as a separate sub-segment, and the rest of the sub-segments ($n_{sub}-1$) will be treated as another sub-segment. This means that n_{sub} number of physical sub-segments will be considered as two logical sub-segments. If another k^{th} sub-segment's Ψ_k fulfills Equation (3), then the k^{th} sub-segment will be considered to be another group or coalition. In this case, there will be three sub-segments, namely the j^{th} , k^{th} , and the rest of the sub-segments ($n_{sub}-2$). Figure 3 depicts the scenario discussed above.

Let us consider that the i^{th} segment's j^{th} sub-segment has n number of vehicles. The participating SUs can be expressed as

$$SU_{ij} = \{SU_{ij1}, SU_{ij2}, \dots, SU_{ijn}\}. \quad (4)$$

Similarly, for the k^{th} sub-segment, if there are q number of vehicles, then it can be expressed as

$$SU_{ik} = \{SU_{ik1}, SU_{ik2}, \dots, SU_{ikq}\}. \quad (5)$$

SSA will consider these sub-segments as the clusters or the coalitions. For each sub-segment, different channels or the bands of interest will be different.

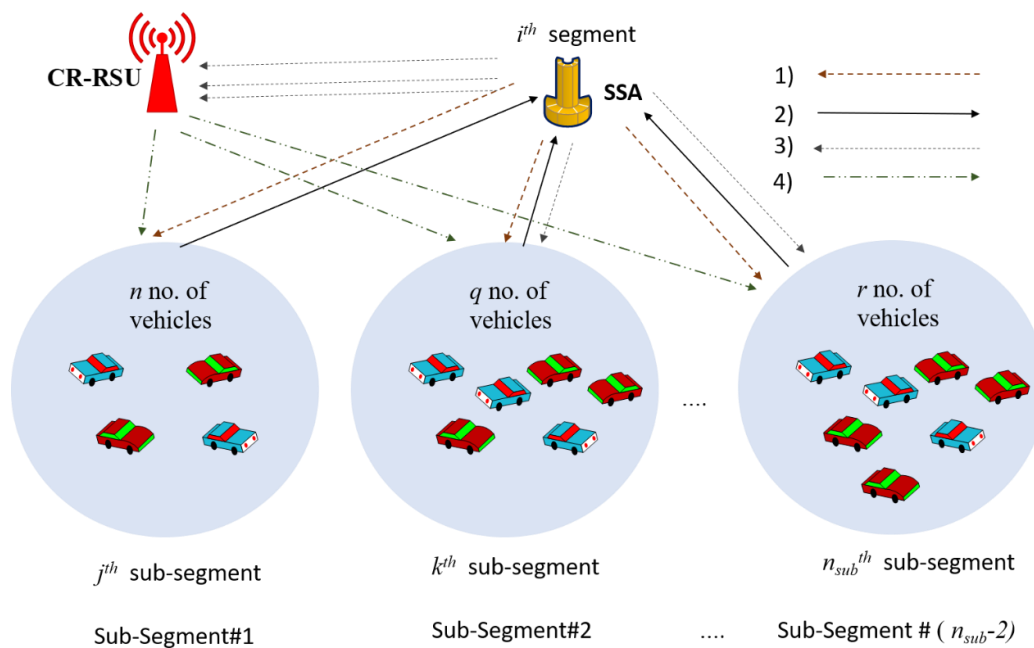


Figure 3. Sub-segmented cooperative spectrum sensing (CSS).

Figure 3 shows that SSA treats each sub-segment separately. The processes include the following:

- (1) SSA provides the channel or the band of interest to the SUs for spectrum sensing. Different channels or bands of interest will be provided to the different sub-segments.
- (2) All the vehicles of that sub-segment will send the local sensing results to the fusion center, i.e., SSA, in our case. SSA will combine its learning results and the local sensing reports of that sub-segment to make the global decision, i.e., the final decision of the PU's presence or absence.
- (3) After making the SSA's global decision, it sends the individual (each sub-segment) reports to the RSU.
- (4) RSU assigns the detected channel to the optimal vehicle.

3.3. Local Sensing and Dynamic Threshold Value

In CR-based networks, spectrum sensing plays a pivotal role in determining the available channels and in supporting data transmission through available channels. There are many conventional methods available for spectrum sensing [22]. However, the dynamic network and channel environment affect the spectrum sensing report. Thus, we presented a novel methodology to select the spectrum sensing method with the awareness of the current network situation.

In this work, SSA acts as an FC. SSA will send the spectrum or the band of interest to the SUs of a particular sub-segment for local sensing. We have concentrated on energy detection (ED) [23] and matched filter (MF) [24] as local sensing techniques, which can be denoted as ST_1 and ST_2 respectively. To select an optimal sensing technique for the current situation, we present the fuzzy-naïve-Bayes machine learning (ML) algorithm. The proposed algorithm computes SNR and noise power ranges for selecting the optimal sensing technique. The naïve Bayes algorithm is a classification technique that is used in many applications. In this paper, we improved the naïve Bayes algorithm by incorporating a fuzzy algorithm that fuzzifies the attributes before classification. In this work, the class denotes the optimal sensing technique. That is, the first class of signals belongs to ST_1 and the second class of signals belongs to ST_2 . Each class has multiple possible values, and the current technique is selected based on two major attributes: SNR and noise power. Further,

the ML technique is modelled as a combination of the Bayesian probabilistic model and the maximum a posteriori (MAP) rule. It can be given as

$$NB(\mathbf{a}) = \operatorname{argmax}_{c \in C} P(c) \prod_{i=1}^{na} P(x_i|c) \quad (6)$$

Here, \mathbf{a} is the complete set of attributes, x_i is the attribute that belongs to X_i , and c represents the corresponding class. In this work, na (i.e., the number of attributes) is 2 (SNR and noise power). The above equation models the conventional naïve Bayes algorithm. When it is combined with a fuzzy approach, the attributes are converted to crisp values to overcome the issue of information loss that occurs in naïve Bayes. In the hybrid ML, the degree of truth is considered as probabilities as $P(x_i|\mathbf{a}) = \mu_{x_i}$ and $P(c|\mathbf{a}) = \mu_c$. Thus, the fuzzy-naïve Bayes model is computed as follows:

$$NB(\mathbf{a}) = \operatorname{argmax}_{c \in C} P(c) \sum_{x_{1j} \in X_1} \frac{P(x_{1j}|c)}{P(x_{1j})} \mu_{x_{1j}} \cdots \sum_{x_{na_j} \in X_1} \frac{P(x_{na_j}|c)}{P(x_{na_j})} \mu_{x_{na_j}} \quad (7)$$

For each attribute, the probability value is computed as in the naïve Bayes algorithm. The probability computation can be performed as follows [25]:

$$P(C = c) = \frac{(\sum_{\sigma \in O} \mu_c^\sigma) + 1}{|O| + |D(C)|} \quad (8)$$

$$P(X_i = x_i) = \frac{(\sum_{\sigma \in O} \mu_{x_i}^\sigma) + 1}{|L| + |D(X_i)|} \quad (9)$$

$$P(X_i = x_i|C = c) = \frac{(\sum_{\sigma \in O} \mu_{x_i}^\sigma \mu_c^\sigma) + 1}{(\sum_{\sigma \in O} \mu_c^\sigma) + |D(X_i)|} \quad (10)$$

Here, O denotes the number of samples considered for classification, $D(X_i)$ is the finite domains $\operatorname{dom}(X_i)$ of the attributes x_i , and σ is the training sample set. In this way, the current network situation is classified based on a reference signal. The probability depends upon the attributes, including SNR and noise power. If SNR is high and noise power is low, then energy detection (i.e., ST_1) is performed.

We used the Neyman–Pearson (NP) binary hypothesis testing [26] in our scheme. The energy detection method senses the spectrum based on two hypotheses as follows:

$$R(n) = \begin{cases} w(n)_j, & H_0 \\ w(n) + \mathbb{p}(n), & H_1 \end{cases} \quad (11)$$

Here, $R(n)$ is the signal sample received by SU, $n = 1, 2, \dots, N$, N is the number of samples. H_0 denotes the absence of PU signal and presence of noise signal ($w(n)_j$), which is the additive white Gaussian noise (AWGN) with zero mean and variance of σ_w^2 , and H_1 denotes the presence of PU signal $\mathbb{p}(n)$ with the noise signal. The hypothesis is computed from the energy level computed from the sensed signal as follows:

$$E_{ST1} = \sum_{q=1}^N R(n)^2. \quad (12)$$

If the computed energy level is higher than the threshold value λ_1 in Equation (13) [14], then it is H_1 ; otherwise, it is denoted as H_0 .

$$\lambda_1 = \sigma_w^2 (\mathbb{Q}^{-1}(P_f)) \sqrt{2N} + N \quad (13)$$

Here, P_f is the probability of false alarm, and \mathbb{Q}^{-1} is the inverse Marcum \mathbb{Q} function. In this way, ST_1 determines the presence/absence of PU activity on the sensed channel. As the energy value is affected by SNR and noise power, the ST_1 is unsuitable for low SNR scenarios. This decision can be made from a hybrid ML algorithm a priori and based on the decision, and the SUs sense the spectrum. The test statistic for the ST_2 can be expressed as follows [27]:

$$E_{ST2} = \frac{1}{N} \sum_{n=1}^N X(n)x_p^*(n) \quad (14)$$

Here, $X(n)$ is the SU received signal and the $x_p^*(n)$ is the pilot samples.

According to the Neyman–Pearson criteria [26], probability of detection, P_d , and probability of false alarms, P_f of ST_2 , can be expressed as

$$P_d = \mathbb{Q}\left(\frac{\lambda_2 - E}{\sqrt{E\sigma_w^2}}\right) \quad (15)$$

$$P_f = \mathbb{Q}\left(\frac{\lambda_2}{\sqrt{E\sigma_w^2}}\right). \quad (16)$$

Here, E is the PU signal energy. For the fixed P_f , the threshold value for ST_2 is computed from Equation (11) as [27]

$$\lambda_2 = \mathbb{Q}^{-1}(P_f) \sqrt{E\sigma_w^2} \quad (17)$$

If E_{ST2} is higher than the value of λ_2 , then it is H_1 (presence of PU); otherwise, it is H_0 (absence of PU).

The spectrum availability decision is made based on these threshold values. Sensing accuracy depends on both threshold values; the fixed threshold value is unsuitable for a dynamic network environment. The threshold value depends on the value of false alarm (f_a), the PU signal energy, and noise variance. Because of the channel's uncertainty and noise aspect, for realistic situations, the traditional threshold value estimation is not optimal. Therefore, a dynamic threshold value is needed that considers both the noise factor and the channel's uncertainty. A dynamic threshold value can be estimated as

$$\lambda_{dyn} = \frac{\lambda}{\sum q_i P_e} \quad (18)$$

Here, λ is the predefined threshold value (λ_1 and λ_2 in our case), q_i is the noise uncertainty factor for the i^{th} SU, and P_e is the probability of sensing error.

P_e can be written as

$$P_e = \omega_1 P_f + \omega_2 P_m. \quad (19)$$

Here, ω_1 and ω_2 are the weighting factors, where, $\omega_1 + \omega_2 = 1$, P_f is the probability of false alarm, and P_m is the probability of missed detection.

Noise uncertainty factor q_i can be estimated using Tsallis entropy [28] as

$$q_i = \frac{1}{q-1} \left(1 - \sum_{i=1}^k p_i^q \right) \quad (20)$$

where p_i is the probability of the frequency of occurrences in the i^{th} bin, q is the Tsallis parameters or entropic index ($q > 1$ or $q < 1$), and k is the total number of possibilities of the system (total number of the bin).

$$p_i = \frac{m_i}{N} \quad (21)$$

where m_i is the total number of occurrences in the i^{th} bin, $i = 1, 2, \dots, k$, N is the total number of occurrences in all the bins.

As the noise uncertainty for each channel varies over time, this dynamic threshold is used in this paper for the optimal spectrum sensing.

Based on the Neyman–Pearson (NP) binary hypothesis testing,

$$E_{ST1}, E_{ST2} \begin{cases} > \lambda_{dyn} & H_1 \\ \leq \lambda_{dyn} & H_0 \end{cases} \quad (22)$$

3.4. Global Sensing and Final Sensing Result

In this work, SSA acts as an FC. SUs in particular sub-segments will send their local sensing reports to the corresponding SSA. Individual statistics E_{ST1} or E_{ST2} are quantized to one bit with $LS_{jn} \in \{0,1\}$. SUs send their individual local sensing as $LS_{j1} \in \{0,1\}$, $LS_{j2} \in \{0,1\}$, \dots , $LS_{jn} \in \{0,1\}$. Here “1” and a “0” represent PU’s presence (H_1) and absence (H_0), respectively. In summary we write, based on Equation (22),

$$LS_{jn} = \begin{cases} 1; & E_{ST1}, E_{ST2} > \lambda_{dyn} \quad i.e. \quad (H_1) \\ 0; & E_{ST1}, E_{ST2} \leq \lambda_{dyn} \quad i.e. \quad (H_0) \end{cases} \quad (23)$$

For the data fusion for the local sensing results of the SUs, we followed *Hard fusion with majority or voting rules*. In majority rule, the decision is taken from the k out of N rule if it follows $k \geq N/2$.

For the j^{th} sub-segment, for example,

$$GS(LS_j) = 1 \quad (H_1) \quad \text{when} \quad \sum_{l=1}^n LS_{jl} \geq \frac{n}{2} \quad (24)$$

$$GS(LS_j) = 0 \quad (H_0) \quad \text{when} \quad \sum_{l=1}^n LS_{jl} < \frac{n}{2} \quad (25)$$

where n is the total number vehicle in j^{th} sub-segment, and $GS(LS_j)$ is the global sensing result based on the local sensing LS_j .

However, this $GS(LS_j)$ is not the final decision. The next phase is making the decision of SSA by its own by using tri-agent reinforcement learning. These dual checking sensing results provide more reliability, are error free, and have enhanced performance.

Let $GS(SSA_j)$ be the global sensing result of SSA based on TA-RL for the j^{th} sub-segment. Here, $GS(SSA_j) \in \{0,1\}$. Now, $GS(LS_j)$ and $GS(SSA_j)$ together make the final decision regarding PU’s presence (H_1) and absence (H_0). For the final sensing result, we used OR hard rule. For the j^{th} sub-segment, the final result is

$$FS_j = GS(LS_j) \text{ OR } GS(SSA_j) \in \{0,1\} \quad (26)$$

Now, we focus on how $GS(SSA_j)$ can be achieved through SSA by using TA-RL.

In our proposed solution, we considered SSA instead of RSU as the RL agent; this is because of the proper management of the spectrum as well as for faster sensing. The RSU will communicate with the vehicles for the data transmission and for the final spectrum assignment (after the getting the confirmed sensing result from SSA). In other works, for instance in [14], RSU acts as an RL agent that deals all the spectrum sensing jobs and data transmission and other tasks. Therefore, there is a huge chance of network overhead that degrades the overall network performance.

The SSA is the intelligent agent [29] that continuously senses the spectrum and make the decision by considering all SUs sensing reports. Deployment of SSA improves the sensing accuracy. However, managing all SU reports in a single SSA becomes complex. Thus, we deployed SSA at each segment that collects the reports from all SUs presented in that segment. For optimal spectrum decision-making, the reinforcement learning (RL) approach is presented. We propose a novel TA-RL algorithm that learns the environment through three agents. The proposed novel decision-making methodology improves

decision-making by using three different agents. The detailed flow of the proposed work is shown in Figure 4.

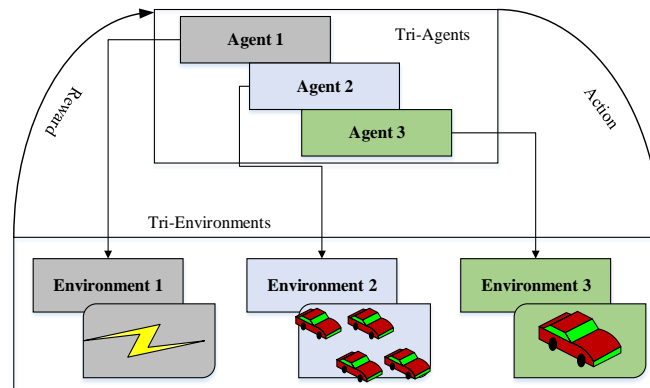


Figure 4. Tri-agent reinforcement learning (TA-RL)-based sensing decision-making.

Reinforcement learning is one of the branches of AI techniques. In RL, the agent is deployed to learn the environment and decide based on the current environment. However, single environment learning is slightly ineffective in our work. Thus, three environments are considered, and three agents are deployed. The considered environments are Signal Environment (SE) as environment 1, Network Environment (NE) as environment 2, and Vehicle Behavior (VB) as environment 3, which are learned by three agents: A_1 , A_2 , and A_3 , respectively. In our work, each agent has different responsibilities that are illustrated in Table 1.

Table 1. Tri-agents and responsibilities.

Agent	Responsibilities
A_1	Learns the SE Metrics Learned: {SNR, time (t), Channel Quality (CQ)}
A_2	Learns the NE Metrics Learned: {Global sensing result ($GS(LS_j)$), Sub-segment number (SN), number of vehicles in that sub-segment (N)}
A_3	Learns the VB Metrics Learned: {Mobility (M), Vehicle ID ()}

In the proposed TA-RL, spectrum availability is made based on these agents and sensing reports from SUs. The three agents are to achieve accurate sensing decisions since the sensing signal can be affected by all these three environments. Here, the state action pairs ($S - A$) define the decision on spectrum availability. The proposed TA-RL algorithm involves the following steps:

Q-value Initialization—Initially, the proposed algorithm defines the Q-table for the (s_t, a_t) pairs. Each pair in the table is denoted as $Q(s_t, a_t)$, and it is defined as per the target application.

At each step t , the SSA observes the states of its surrounding environment by using its three agents. Let us consider that S is a set of all possible states. Based on knowledge gained at s_t , the SSA selects an action $a_t \in A$, where A is a set of actions. Here, action refers to the declaration of the absence or presence of PUs. At the next step, $t + 1$, the environment transits to a new state s_{t+1} , and the agent gets a reward of r_t . Based on the reward table, the agent chooses the next action (it may be beneficial or may be harmful), and then they update a new value called Q-value mapping of state–action pairs $Q(s_t, a_t)$. Several Q-values are stored in the Q-table.

Perform Action—In this stage, the action is made by considering three environments that are learned by three agents. In contrast to fusing the sensing reports, this work considers the environmental parameters and the sensing reports. Each agent uses a ϵ -greedy exploration policy to update the Q-table. There are three states that are considered as S_{A1}, S_{A2}, S_{A3} , and each state is learned by each agent. For instance, the three states are as follows:

$$S_{A1} \rightarrow \{SNR, t, CQ\} \tag{27}$$

$$S_{A2} \rightarrow \{GS(LS_j), SN, N\} \tag{28}$$

$$S_{A3} \rightarrow \{M, U\} \tag{29}$$

For agent A_1 , the states are considered as the state of the channels' SNR, the time stamp, and the channel quality (congested or free). Similarly, for A_2 , the states are the global result of each sub-segment and how many participating vehicles there are. For A_3 , the states are the vehicle speed and the ID of the vehicle. In other words, overall TA-RL learns for a particular sub-segment its global sensing results, how many vehicles there are, what are their speeds and IDs, at what time the sensing result is created, and the what the channel condition is at that time.

The action is taken based on the above learning. Action either declares the band of interest as PU free or not. For instance,

$$GS(SSA_j) \in \{0, 1\} \tag{30}$$

Based on the action, the reward function (r_t) is updated for each action. The reward can be given as in Table 2.

Table 2. Reward table.

Value of $GS(LS_j)$	Value of $GS(SSA_j)$	Reward (r_t)
{0}	{0}	+ r_{t1}
{0}	{1}	- r_{t2}
{1}	{0}	- r_{t3}
{1}	{1}	+ r_{t4}

Here “1” and a “0” represent PU’s presence (H_1) and absence (H_0), respectively; r_{t1}, r_{t2}, r_{t3} , and r_{t4} are real integers values. When both global sensing and TA-RL’s own estimated result is the same, it would be given a “+” (positive) reward; otherwise, it would be a “-” negative reward (*punishment*).

The current state of three states can be written as

$$s_t = \sum S_{A1t}, S_{A2t}, S_{A3t} \tag{31}$$

After every action, the agent gets the reward and updates its Q-value based on the following equation:

$$Q_{new} (state, action) \leftarrow (1 - \alpha) Q_{old} (state, action) + \alpha (reward + \gamma \max Q_{old} (next\ state, all\ actions)) \tag{32}$$

$$Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha)Q_t(s_t, a_t) + \alpha \left[r_{t+1}(s_{t+1}, a_t) + \gamma \max_{a \in A} Q_t(s_{t+1}, a) \right] \tag{33}$$

Here, α is the *learning rate*, which determines how much the new Q-value overrides the previous Q-value. The α ranges from 0 to 1; γ is the *discount factor*, which implies how much importance is given to future rewards; and r is the *reward* received by the agent. The short-term reward is called the *delayed reward*, and the future reward is called the *discounted reward*.

Here, the action is to decide on spectrum availability as PU is presented, or PU is absent on the corresponding channel. The action is taken in a state where the reward is

the maximum for that action found in the past state–action pair. There are two policies for action. When an agent chooses to be exploited (using current knowledge to choose the best action), it uses an optimal policy, and it uses a random policy when deciding to be explored (needs more knowledge). The agent shall receive positive delayed rewards when choosing the required action for a specific state. The positive value increases and the respective Q-value increases and vice versa. Therefore, Q-learning aims to get an optimal policy (agent behavior) $\pi: S \rightarrow A$, which can maximize the reward at state S [30].

The optimal Q-value for a particular state can be written as

$$V^{\pi^*}(s_t) = \max_{a \in A} Q_t(s_t, a) \quad (34)$$

Therefore, the optimal policy can be written as

$$\pi^*(s_t) = \arg \max_{a \in A} Q_t(s_t, a) \quad (35)$$

It is evident from the discussions above that the convergence rate depends on the Q-table consistency and the values of α and γ . The more incentive the agent accumulates, the better the Q-table will be, and thus the convergence will be faster.

Based on the learning environment and the cooperative decision, the final decision is made by SSA, i.e., the spectrum is available or not. Then, this decision is exchanged with the corresponding segment RSU to support effective network management. The RSU assigns the available channels to the segment SUs to perform data transmission. On the allocated channels, the vehicles are allowed to transmit data.

The proposed solution discussed above is represented in the algorithms given below. Here, Algorithm 1 represents the complete solution while Algorithm 2 represents TA-RL algorithm.

Algorithm 1: Complete algorithm for the solution

```

1: Start
2: Initialize  $\mathbb{S}, \phi, \Theta, n_{sub}, Th_{seg}$ ;
3: //Sub-segmentation
4:  $i^{th}$  segment;
5: For ( $j \leq n_{sub}$ )
6:   Compute  $\Psi_{jnorm}$ ; //by Equation (2)
7:   If ( $\Psi_{jnorm} > Th_{seg}$ )
8:     Consider  $j^{th}$  as segment#1;
9:     Count ++;
10:    total_segment =  $n_{sub} - \text{count} + 1$ 
11: //Local sensing and hypothesis testing
12: For ( $j \leq \text{total\_segment}$ )
13:   number_vehicle =  $n$ ;
14:    $SU_{ij} = \{SU_{ij1}, SU_{ij2}, \dots, SU_{ijn}\}$ ; //Equation (4)
15:   For ( $r \leq n$ )
16:     Classify signal as low SNR or high SNR; //Equation (10)
17:     If (SNR == low)
18:       MFD is used;
19:       Compute  $E_{ST2}$ ; //Equation (14)
20:       Compute  $\lambda_2$  and  $\lambda_{dyn}$ ; //Equations (17) and (18)
21:       If ( $E_{ST2} > \lambda_{dyn}$ )
22:          $LS_{jn} = 1$ ;
23:       Else

```

```

24:       $LS_{jn} = 0;$ 
25:      If (SNR == high)
26:          ED is used;
27:          Compute  $E_{ST1}$ ; //Equation (12)
28:          Compute  $\lambda_1$  and  $\lambda_{dyn}$ ; //Equation (13), Equation (18)
29:          If ( $E_{ST1} > \lambda_{dyn}$ )
30:               $LS_{jn} = 1;$ 
31:          Else
32:               $LS_{jn} = 0;$ 
33:      Return  $LS_{jn}$ ;
34:      //Data fusion
35:      Compute  $GS(LS_j)$  based on Equation (24) and (25);
36:      //TA-RL
37:      Compute  $GS(SSA_j)$  by running Algorithm 2;
38:      //Final sensing result
39:      Return  $FS_j$  based on Equation (26);
40:      End

```

Algorithm 2: TA-RL Algorithm

```

1:      Start
2:      Initialize  $S_{A1}, S_{A2}, S_{A3}$  based on Equations (27)–(29);
3:      Initialize  $Q(s,a)$  arbitrarily;
4:      For t: =1 to T do
5:          Observe current state  $s_t$  based on Equation (31);
6:          Determine exploration or exploitation
7:          If (exploration)
8:              choose a random action  $a_t$ 
9:          Else if (exploitation)
10:             Choose the best-known action  $a_t$  using Equation (35); // i.e.,  $GS(SSA_j)$ 
11:             Receive reward:  $r_{t+1}(s_{t+1})$  based on Table 2;
12:             Update Q table  $Q_{t+1}(s_t, a_t)$  using Equation (33) for state–action pair  $(s_t, a_t)$ ;
13:      Return  $GS(SSA_j)$ ;
14:      End

```

3.5. Other Elements of CSS

There are other elements needed to perform CSS [12]. This sub-section discusses these elements aligned with our proposed solutions.

3.5.1. Cooperation Models

The collaboration of CR users for spectrum sensing can be modelled on various approaches. Cooperative sensing modelling is mainly concerned with how CR users work together to perform spectrum sensing and achieve optimum detection efficiency. The most common and dominant approaches are the parallel fusion (PF) model for distributed detection and data fusion and the game theory approach. In this paper, PF model is used as the model of SU cooperation. In PF, SUs observe the physical phenomena H through the sensing observation and report to the central unit or FC. There are three steps in FC: local sensing, data reporting, and data fusion. All CR users are synchronized by the FC to sense the channel or frequency band of interest and to record the sensing data. The FC combines the local sensing data recorded and takes a global cooperative decision.

3.5.2. Control Channel and Reporting

In our CSS architecture, a common control channel (CCC) is used by the SUs to report local sensing results to the SSAs. There are three requirements to fulfil successful reporting: bandwidth, reliability, and security. Due to sub-segmentation, managing these

requirements is much easier. We assumed that SUs use dedicated CCC, which is not imperfect. However, focusing on improvement to these issues is beyond the scope of this work.

3.5.3. Knowledge Base

The efficiency of CSS schemes depends mostly on the knowledge of PU characteristics, including traffic flows, location and transmission of power, SNR, channel quality, etc. PU details, if available in a database, can facilitate the detection of PU. The database that holds all knowledge of the RF environment is called the knowledge base. It is an essential feature of CSS since it can support, supplement, or even substitute CSS to detect PU signals and classify the available spectrum. Our SSA acts like a knowledge database that maps the PU activities with the parameters shown in Table 2. After the convergence, the TA-RL agent can retrieve the PU information from its database (*Q-table with the best reward*). This retrieval of information saves time in spectrum sensing.

Table 3 shows the elements of CSS that we used in our proposed solution.

Table 3. Elements of proposed CSS.

Elements of CSS [12]	Used in our solution
Cooperation Models	Parallel fusion model
User Selection	Sub-segmentation of the segment
Sensing Techniques	Fuzzy-naïve Bayes algorithm chooses ED or MFD
Hypothesis Testing	Neyman-Pearson (NP) binary hypothesis testing
Control Channel and Reporting	Sub-segmented SUs via control channels
Data Fusion	Hard combining, majority rule, OR rules
Knowledge Base	SSA, TA-RL

4. Experimental Evaluation

This section discusses the simulation and parameter settings and the theoretical comparisons with prior works.

4.1. Simulation Setup

For evaluating the proposed concept, we modeled our proposed vehicular network using a network simulation tool, namely OMNeT++ with the SUMO framework. OMNeT is a C++-based simulation tool that supports the productive simulation of vehicular-based networks and many other network protocols. We used Veins, INET, and crSimulator frameworks in the OMNeT++ platform. Vehicle mobility type is considered based on Veins' submodule, TraCIMobility. In this work, a Rayleigh multi-path propagation model was considered. The channel vector was modeled as a zero-mean and complex Gaussian random vector. We considered the network area of 2750 m × 250 m with 100 vehicles as SUs, 10 static PUs, 2 RSUs, and 2 SSAs. We also considered a maximum of 4 sub-segments ($n_{\text{sub}} = 4$) per segment. In general, vehicles in non-congested network use DSRC channels (6 service channels or SCH) of 10 MHz bandwidth in the range of 5.9 GHz. For communication in the MAC/PHY layer, the WAVE/IEEE802.11p standard was used for the DSRC channel. TV channels of 6 MHz bandwidth in the range of 500 MHz–524 MHz were considered as CR bands. For the purpose of CR, we used 4 channels, which means that with DSRC and TV, we had a total of 10 channels.

Other parameter values used for the simulation are depicted in Table 4.

Table 4. Simulation and parameters settings.

Parameter	Value/Range	Parameter	Value/Range
Vehicle Speed	0–30 m/s	Vehicle acceleration	2.5 m/s ²
Number of Packets	10,000 (approx.)	WAVE Tx power	13 dBm
MAC Header Length	256 Bit	SNR	−20 to 5 dB
Packet Size	512 KB	Packet interval	0.1 s
Data Rate	20 Mbps	Simulation time	500 seconds
Sensing Duration	~6 ms	Transmission duration	~12 ms
Noise Power	−110 dBm	TV modulation	8 VSB
False Alarm	0.1	TV Tx power	21 dBm/Hz
Delay	0–18 ms	Learning rate and discount factor for RL	0.1 and 0.9
Jitter	10 ms	Number of signal samples	300

We first created a CR-VANET environment with the above configuration. We considered TV channels of 500 MHz–524 MHz for the CR usage. PUs were considered to be static, and they followed simple ON/OFF PU activity. SUs were equipped with two antennas, one for DSRC and another for CR usage. Then, we performed data transmission to test the proposed work performance. We then implemented segmentation, spectrum sensing, decision-making, and route selection processes on the created environment to measure the performance. The performance was measured in terms of performance metrics.

4.2. Comparative Analysis

This section evaluates the proposed work with existing works to prove our proposed approach's efficacy. We compared our work (Seg-CR-VANET) with existing works including RL-DSA [14], regional clustering [16], and binary decision-making [4]. A detailed comparison of the existing works is presented in Table 5. =

Table 5. Comparison of existing works.

Previous Work	Research Purpose	Spectrum Sensing	Limitations of the Work	Comparative Improvements Made in Our Work (Seg-CR-VANET)
RL-DSA [14]	To improve spectrum management by a dynamic spectrum access	Energy detection, cyclostationary	<ul style="list-style-type: none"> In the spectrum sensing, the threshold, λ is set fixed based on the probability of energy detection and noise power variance. However, the power variance-based fixed threshold for sensing decision is not optimal since the power differs based on the environment. Based on the threshold validation, two sensing methods are applied simultaneously that take time to report to RSU regarding a channel. The use of reinforcement learning in this work is used to learn only the channel environment and decide; however, the channel characteristics differ based on the network environment. 	<ul style="list-style-type: none"> We used dynamic threshold values. We used TA-RL that learns three environments (network, signal, and vehicle).

Table 5. Cont.

Previous Work	Research Purpose	Spectrum Sensing	Limitations of the Work	Comparative Improvements Made in Our Work (Seg-CR-VANET)
Regional clustering [16]	To improve data transmission through region-based clustering	Linear programming-based sensing	<ul style="list-style-type: none"> This work fails to balance the density among various RSUs. The involvement of linear programming introduces difficulty in defining the objective function. Channel estimation is performed priori, which is unsuitable for VANET. 	<ul style="list-style-type: none"> We used density aware segmentation and sub-segmentation-based network architecture. We used reinforcement learning-based channel estimation (learning) that does not need prior information.
Binary decision-making [4]	To support CSS-based sensing and decision-making	Hard and soft fusion	<ul style="list-style-type: none"> The energy detection method's local decision cannot detect PU's presence when the signal's SNR is low. The channel characteristics often change, so this technique is not suitable for accurate prediction of the PU signals. Static threshold value was used. 	<ul style="list-style-type: none"> We used dynamic machine learning approaches (fuzzy-naïve Bayes to choose the appropriate sensing technique). In low SNR, MFD is used and in high SNR, ED is used. We used dynamic threshold values.

The theoretical comparison shows that each existing work has some limitations and drawbacks. This can be tested through brief performance measures as shown in the following section.

5. Results, Discussion, and Highlights

This section discusses results obtained through the simulations. We compared our proposed solution with other works for evaluation purposes. We used several performance metrics.

5.1. Analysis of the Probability of Detection

The probability of detection metric measures a vehicle's probability of sensing the channel and accurately detecting the PU activity. This metric measures the effectiveness of the involved sensing technique.

The probability of a false alarm is the probability that a SU mistakenly detects the presence of a PU, where in reality, there is no PU present at that time. This means that a SU detects H_1 as true, but in reality, H_0 is true. On the other hand, the probability of missed detection is the opposite of a false alarm, and it is the probability that SU senses the channel as idle (absence of PU), but in actuality, the channel is not idle (occupied by the PU). In Figure 5, the proposed work is compared with the existing works regarding the probability of detection.

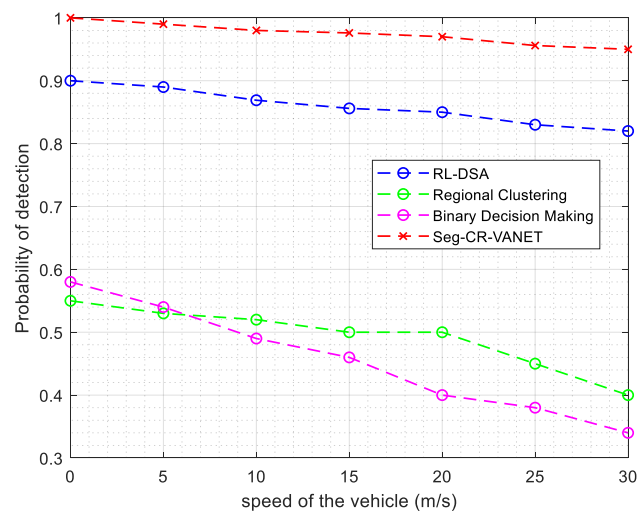


Figure 5. Comparison on probability of detection.

The analysis shows that the proposed work achieves a better probability of detection, i.e., the proposed work detects PU's presence on the sensing channel accurately. In general, sensing accuracy is significant in any CR-based network. Spectrum sensing in CR-VANET is much more challenging due to the dynamic movement of vehicles and the randomness of the network environment. Thus, the existing works have not yet achieved better results, since those work could not handle the dynamicity of the VANET environment effectively. As we focused on dynamic sensing technique selection by the hybrid ML algorithm, we achieved better detection accuracy. We attained a probability of detection in the range of 0.95 to 1, which is nearly 50% higher than in the previous works. This better result was achieved because sensing accuracy is greatly affected by channel and network errors, which were not considered in the existing works, but we considered them. A dynamic sensing technique is proposed with a dynamic threshold update in our work. Moreover, deployment of SSA in each segment assures high sensing accuracy.

As seen in Figure 6, we compared the proposed hybrid ML-based spectrum sensing method with the base spectrum sensing techniques such as energy detection and matched filter with static threshold values. The analysis shows that the base algorithms lack the probability of detection. When the vehicle speed is increased, then the probability of detection is decreased. We achieved better result than these base sensing techniques.

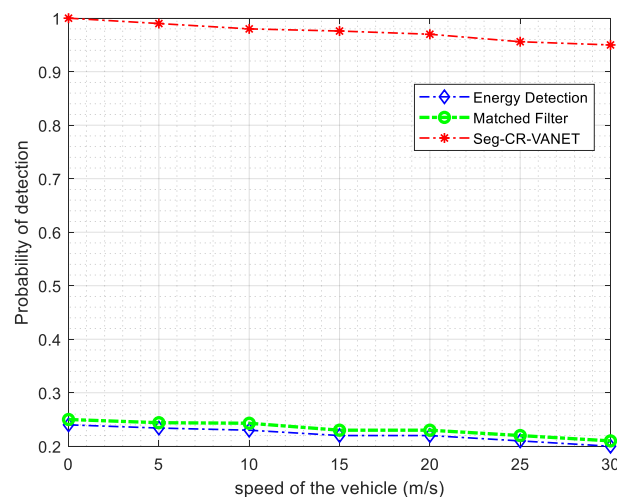


Figure 6. Probability of detection analysis.

This better result is because energy detection fails to sense the spectrum in low SNR, and the matched filter fails to sense the spectrum in high SNR scenarios. Thus, both methods achieve less than 0.3 as the probability of detection. In our work, the involvement of hybrid ML-based dynamic spectrum sensing improved detection probability up to 0.98.

In Figure 7, the average probability of detection is compared by varying mean detection time. This analysis was carried out to assure that the proposed work attained better accuracy, even with lower detection time. Here, detection denotes the sensing time allotted to the SUs for spectrum sensing. Although high sensing time improves detection probability, it degrades the data transmission ability. Thus, an optimal sensing technique must use minimum sensing time to achieve higher detection accuracy. The increase was encountered in the proposed curve, which varied from around 0.7 to 1, thus increasing sensing time. Simultaneously, the previous works have a sensing accuracy of 0.2 when the sensing time is low.

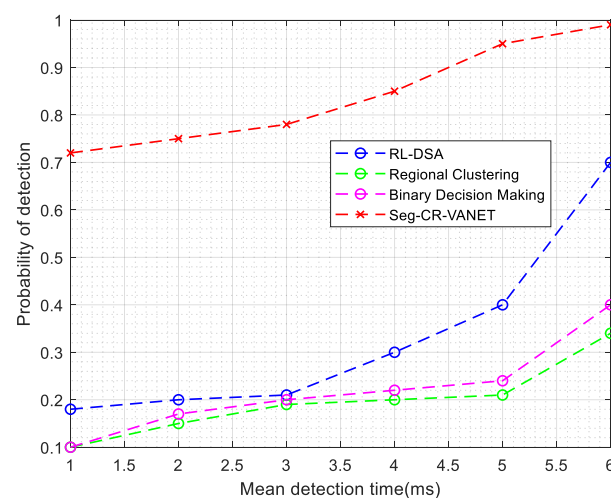


Figure 7. Probability of detection vs. mean detection time.

This analysis shows that the proposed work can assure better sensing and transmission efficiency in a dynamic CR-VANET environment. Due to the use of TA-RL, SU adapts the environment very quickly, and as a result, it takes much less time to detect the spectrum hole.

Figure 8 shows receiver operating characteristics (ROC) curves, where the average probability of detection is compared by varying values of the average probability of a false alarm.

We considered the SNR value as -10 dB. The figure shows that the value of the probability of detection is increased as the value of the probability of false alarm is increased. The proposed Seg-CR-VANET showed a better result than the previous works. Our sensing scheme could maintain a probability of detection of 0.9 (i.e., 90%), compared to RL-DSA with 0.8 (i.e., 80%); and regional-clustering, and binary decision-making with 0.7–0.75 (i.e., 70%–75%) based on a probability of false alarm of 0.2. However, the higher value of the probability of a false alarm makes the SUs limit the reuse of the radio spectrum.

Figure 9 shows the probability of missed detection with the varying values of the false alarm.

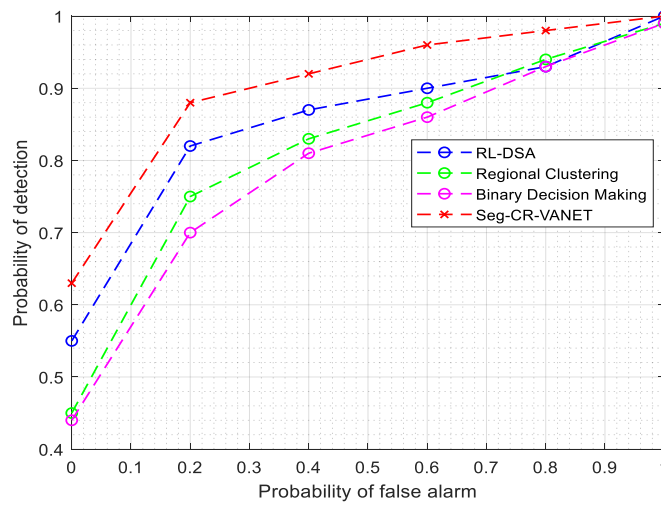


Figure 8. ROC curves of proposed solution and prior works.

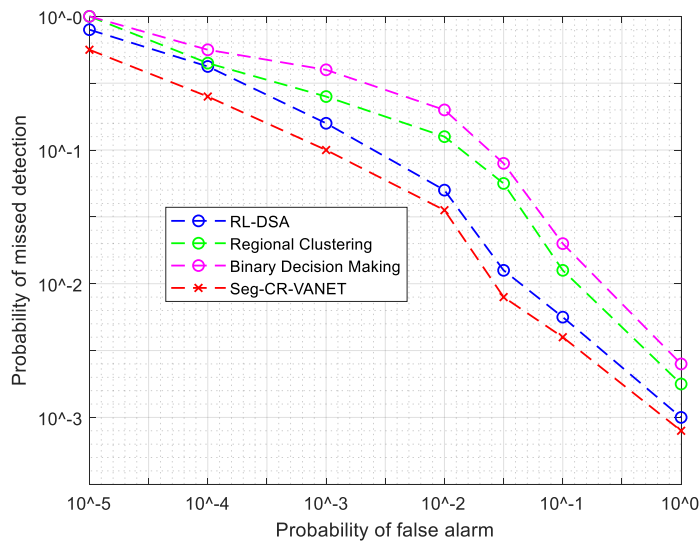


Figure 9. Probability of missed detection vs. probability of false alarm.

The probability of missed detection value should be kept low for better sensing performance, which causes interference, while the probability of false alarm causes losses of spectral opportunities. For best performance, both values should be at a minimum level, while the probability of detection should be at the maximum level. Figure 9 shows that our proposed scheme provides lower missed detection compared with the previous works. Achieving a better result is due to the spectrum’s proper management by using the segment and sub-segment concept and the TA-RL algorithm as it deals with three environments.

5.2. Analysis of Throughput

Throughput is defined as the amount of data transmitted over the network over the given time slot. In the case of CR-VANET, it depends greatly on the channel availability. Thus, we compared throughput with varying sensing time.

In Figure 10, a comparison of throughput and sensing time is shown.

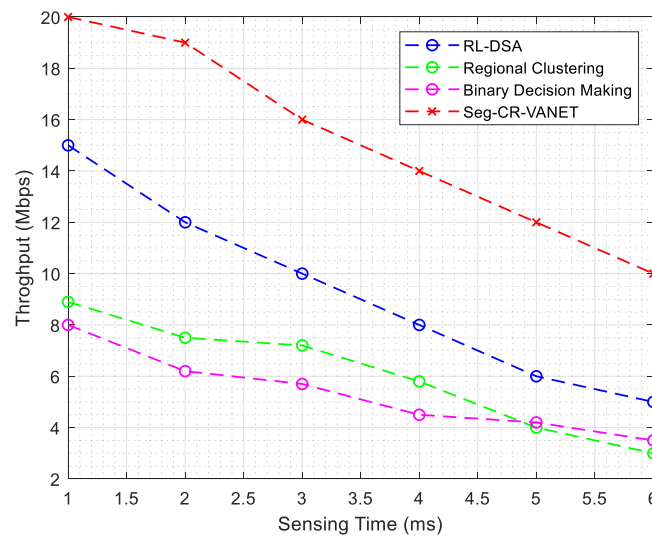


Figure 10. Comparison of throughput and sensing time.

The analysis shows that all works decrease the throughput with an increase in sensing time. As the vehicles use more time for sensing, they have less time for data transmission, which is why throughput is decreased as sensing time is increased. Our work maintains throughput within a better range and achieves up to 20 Mbps, which is relatively more than prior works. The primary reason for this achievement is that the proposed work considered several significant parameters including noise power, vehicle density, vehicle behavior and speed, network quality, etc. As a result, we achieved a stable spectrum sensing and stable channel allocation scheme. For these reasons, we obtained good throughput, while other works' throughput is minimized to 2.5 Mbps. Moreover, using dynamic threshold values provide more accurate and stable spectrum sensing results. If the spectrum is available, then the spectrum must be utilized efficiently to achieve better throughput. In RL-DSA, the spectrum sensing is performed by the energy detection method, and the road is segmented into an equal length of segments. Here, maintaining a fixed segment increases data loss.

Similarly, regional cluster-based CSS is presented with binary decision-making. In this method, the sensing decision is made inaccurate, and it lacks the throughput range. Due to inaccurate sensing and improper network management, throughput is very minimal in these prior works.

5.3. Analysis of Packet Delivery Ratio

Packet delivery ratio (PDR) is defined as the ratio between the total number of packets generated to the number of packets successfully transmitted to the destination.

In Figure 11, PDR is compared with a varying number of vehicles. The PDR decreases with a varying number of vehicles. We achieved PDR up to 96%–99%, which is around 10% better than RL-DSA, which provides the closest results to ours. PDR is decreased with the increase in the number of vehicles. This is due to contention in the wireless channel, as the number of nodes in connection grows. As a consequence, several packets are lost due to a collision. However, our proposed algorithm maintains a good PDR due to the adaptive spectrum sensing technique, dynamic threshold values, and proper learning of the network using the TA-RL algorithm. Thus, the PDR is achieved between 96% to 99%, since we have performed optimal spectrum sensing based on the current network environment, and decision-making is also performed based on three environments. Unlike the proposed work, the existing results have achieved lower PDR. For achieving data transmission successfully, accurate spectrum availability is mandated. This analysis shows that the proposed approach, which focuses on both spectrum and road segmentation, improves PDR effectively.

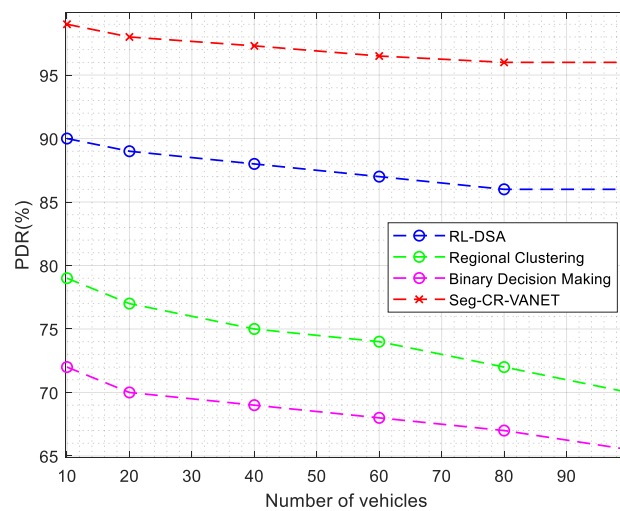


Figure 11. Comparison of packet delivery ratio (PDR).

5.4. Analysis of Average Delay

Delay is defined as the time taken by a data packet to reach the destination from the source. The delay is measured as the function of propagation time, waiting time, and transmission time. In Figure 12, the delay is compared concerning the number of vehicles.

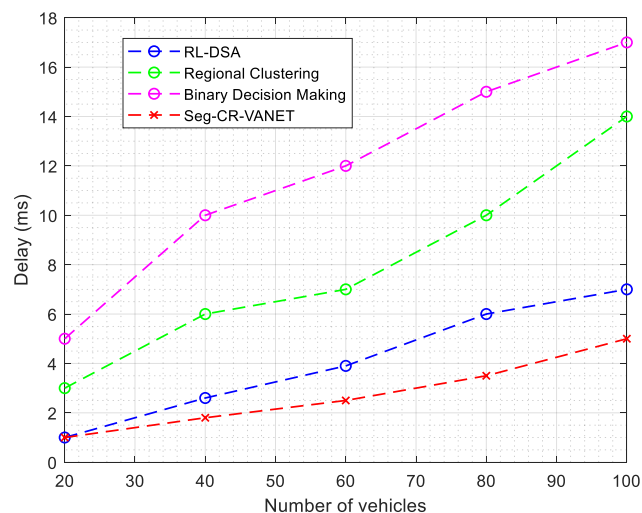


Figure 12. Comparison of delay.

Delay is an important performance measure that shows the efficacy of the proposed spectrum sensing and network management. In the proposed work, the delay is minimized to 5 ms since the available spectrum is utilized by the proposed algorithm effectually. In the proposed work, the available spectrum is determined by the hybrid ML technique. The road is segmented and sub-segmented using a probabilistic approach considering vehicle density, mobility, and node degree. In the prior research, the delay is increased up to 17 ms due to a lack of optimal spectrum sensing and network management, since inaccurate spectrum sensing decision-making decreases the availability of the spectrum for vehicular nodes.

5.5. Analysis of Packet Loss Ratio

Packet loss rate (PLR) is defined as the ratio of the number of packets lost and the total number of packets transmitted over the network. In Figure 13, PLR is compared based on the number of vehicles.

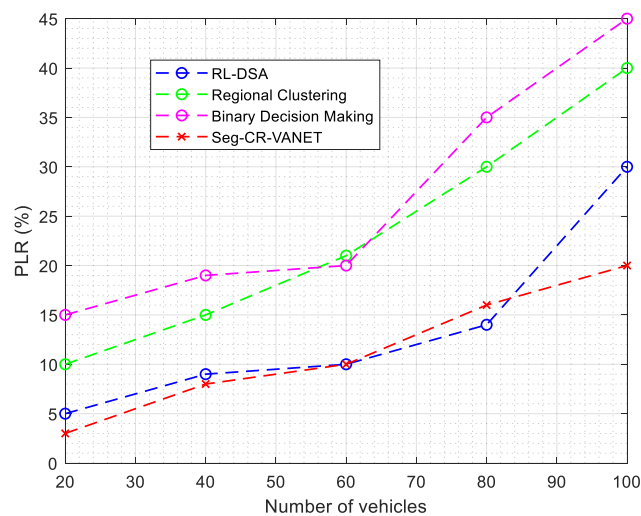


Figure 13. Comparison of PLR.

In this work, PLR is nearly 20%, i.e., 0.2, which is relatively lower than that of previous research works. In the proposed work, the sensing technique is chosen based on the environment, a dynamic threshold is used, and clusters are made based on the road segmentation and sub-segmentation. Thus, the PLR is reduced even with an increase in the number of vehicles. On the other hand, spectrum-based RL-DSA works focus on spectrum allocation, the regional clustering method concentrates on CSS, and binary decision-making uses OR rule-based decision-making. The spectrum is underutilized in all these works, which leads to a PLR of up to 30% to 50%. From this analysis, it is clear that the proposed work, which includes multiple sensing techniques and adaptive threshold values, improves the PLR by transmitting most of the packets successfully. The more accurate the sensing results, the less the PLR.

In Table 6, the obtained results are summarized with mean and standard deviation (SD) values. It can be noted that the proposed Seg-CR-VANET achieves better results in all metrics due to the involvement of optimum spectrum management and road management. Thus, the results confirmed our problems, including lack of spectrum and road segment management. In particular, we also achieved better probability detection, which assures that the sensing technique selection must rely on the current network environment. Optimal spectrum decision-making with dynamic threshold and proper network management by using a probabilistic approach improves data transmission performance effectually.

Table 6. Numerical comparison of obtained results.

Work	PDR (%)		Delay (ms)		PLR (%)	
	Mean	SD	Mean	SD	Mean	SD
RL-DSA	88	±1.5	4.1	±2.4	13.6	±0.1
Regional Clustering	75.4	±2.7	8	±4.1	23.2	±0.11
Binary Decision-making	70.2	±1.1	11.8	±4.6	26	±0.13
Seg-CR-VANET	97.4	±1.0	2.76	±1.1	11	±0.07

The performance of throughput, PDR, PLR, and delay can be further improved by optimizing the route properly. For simplicity, we used the AODV (ad hoc on demand distance vector) routing protocol in our simulations. Although we used this simple routing protocol, we achieved very good results in all aspects. However, there is scope to improve these performances by incorporating the proper routing method, which is beyond the scope of this paper, but we will address this issue in future work.

5.6. Detection Performance Measures

To compare the performance of proposed Seg-CR-VANET sensing with the mentioned prior works, we used the performance metrics shown in Table 7) [31,32].

Table 7. Performance metrics and definition and formula.

No.	Performance Metrics	Definition	Formula
1	Accuracy	It represents the proportion of correctly identified results, both positives and negatives.	$accuracy = \frac{T_p + T_N}{T_p + F_p + F_N + T_N}$
2	Recall (also known as sensitivity or true positive rate (TPR))	It represents the fraction of correctly identified positives.	$recall = \frac{T_p}{T_p + F_N}$
3	Precision (also known as positive predictive value (PPV))	It is fraction of positive results that are true positives.	$precision = \frac{T_p}{T_p + F_p}$
4	Specificity (also known as true negative rate (TNR))	It measures the proportion of negatives that are correctly identified	$specificity = \frac{T_N}{F_p + T_N}$
5	Negative predictive value (NPV)	It is the fraction of negative results that are true negatives.	$NPV = \frac{T_N}{F_N + T_N}$
6	False positive rate (FPR) (also known as fall-out)	It is the proportion of negatives that are incorrectly identified.	$FPR = 1 - specificity = \frac{F_p}{F_p + T_N}$
7	False negative rate (FNR) or miss rate	It is the proportion of positives that are incorrectly identified.	$FNR = 1 - recall = \frac{F_N}{F_N + T_p}$
8	F1 Score	It is needed when we want to make a balance between precision and recall.	$F1 = 2 \times \frac{precision * recall}{precision + recall}$

Where:

- T_p (true positives): These are cases in which the sensing technique detected the presence of a PU signal, and an actual PU signal is there in the environment.
- T_N (true negative): These are cases in which the sensing technique detected the absence of a PU signal, and there is in actuality no PU signal in the environment.
- F_p (false positive): In this case, the sensing technique detected the presence of a PU signal; however, there is actually no PU signal in the environment.
- F_N (false negative): In this case, the sensing technique detected the absence of a PU signal; however, there is actually a PU signal in the environment.

The confusion matrix is a matrix in which the number of correct and incorrect detections are summarized. Table 2 shows the confusion matrix for our proposed solution (Seg-CR-VANET) along with the three other works compared in the previous subsection.

We took 600 samples of the signals, out of which 300 samples contained PU signals along with noise signals, and the other 300 samples contained only noise signals. We considered the SNR value of -10 for all cases. Our proposed Seg-CR-VANET sensing correctly detected 268 signals as PU signals out of 300 PU signal samples, whereas out of 300 noise samples, it detected 274 correctly. In the above matrix, we also included the corresponding values for the benchmark works.

After using the formulas mentioned in Table 1 and the values provided in Table 3, we achieved the results shown in Table 8.

Table 8. Confusion matrix.

		Actual	
		Presence of PU Signal (1)	Absence of PU Signal (0) (i.e. Only Noise Signal)
Detected/Predicted	[Presence of PU signal (1)]	T_P 281 (Seg-CR-VANET) 242 (RL-DSA) 224 (Regional clustering) 213 (Binary decision)	F_P 19 (Seg-CR-VANET) 58 (RL-DSA) 76 (Regional clustering) 87 (Binary decision)
	Absence of PU signal (0) (i.e. only noise signal)	F_N 17 (Seg-CR-VANET) 41 (RL-DSA) 72 (Regional clustering) 83 (Binary decision)	T_N 283 (Seg-CR-VANET) 259 (RL-DSA) 228 (Regional clustering) 217 (Binary decision)

Based on these performance measures, our proposed solutions performed significantly better than other prior works (Table 9). We achieved an accuracy of 0.940: RL-DSA had 0.835, regional clustering had 0.7533, and binary decision had 0.7167. We also achieved very low FPR and FNR compared to the other works. Higher values of accuracy, precision, recall, and F1 scores confirmed the better performance of our proposed solutions.

Table 9. Performance comparison.

No.	Measure	Seg-CR-VANET	RL-DSA	Regional Clustering	Binary Decision
1	Accuracy	0.9400	0.8350	0.7533	0.7167
2	Recall (Sensitivity)	0.9430	0.8551	0.7568	0.7196
3	Precision	0.9367	0.8067	0.7467	0.7100
4	Specificity (True Negative Rate (TNR))	0.9371	0.8170	0.7500	0.7138
5	Negative Predictive Value (NPV)	0.9433	0.8633	0.7600	0.7233
6	False Positive Rate (FPR)	0.0629	0.1830	0.2500	0.2862
7	False Negative Rate (FNR)	0.0570	0.1449	0.2432	0.2804
8	F1 Score	0.9398	0.8302	0.7517	0.7148

5.7. Performance of TA-RL

In this subsection, we evaluated our TA-RL algorithm's detection performance as well as its convergency. Figure 14 shows the improvement of our proposed TA-RL. We ran the simulation for 3000 episodes. Here, episode denotes all the stages that fall between an initial state to the terminal state of a sensing cycle. At the end of each episode, the agents integrate local decisions and take a cooperative sensing decision. We achieved good detection performance even before our optimum solution was achieved. Figure 14 shows the enhancement of detection performance during the TA-RL process. In the figure, we showed two cases: one was with the use of TA-RL (i.e., *GS(SSA)*), and the other was without the use of TA-RL (i.e., *GS(LS)*). We calculated P_d based on the PU activity and with the initial 500 sensing decisions made by TA-RL. We found that P_d was improved steadily and reached above 0.92 after 2200 episodes. Thus, the efficiency of detection increased with TA-RL-based CSS as soon as learning from the environment took place.

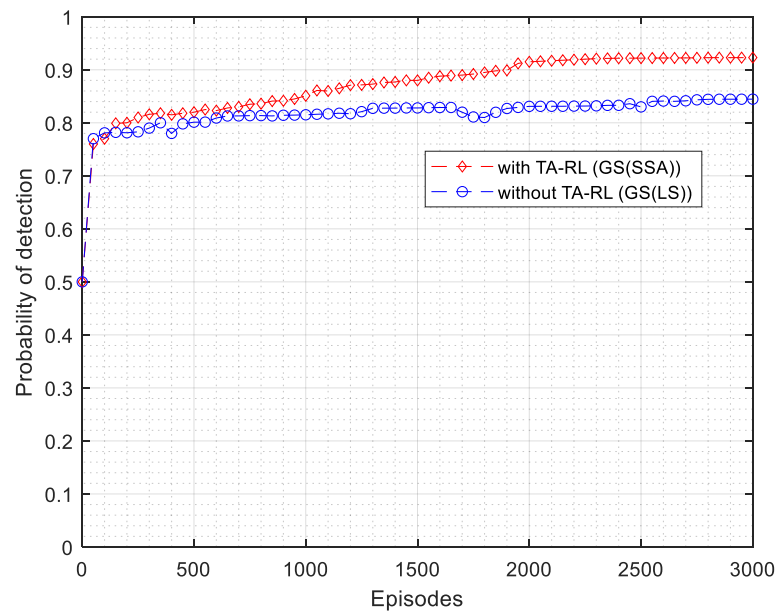


Figure 14. Improvement of P_d during TA-RL-based CSS.

Figure 15 shows the average rewards of all the three agents over the most recent 100 episodes for a total of 3000 episodes and averaged the results to validate the performance of TA-RL based on Q-learning with ϵ -greedy. We considered discount factor $\lambda = 0.9$ with ϵ -greedy, $\epsilon = 0.1$. Since the reward observed at each state was constrained and the number of states was finite for each episode, the expected reward asymptotically approached its upper bound when the algorithm converged. We obtained the convergence of the algorithm after 2200 episodes with the maximum average rewards of 3.84.

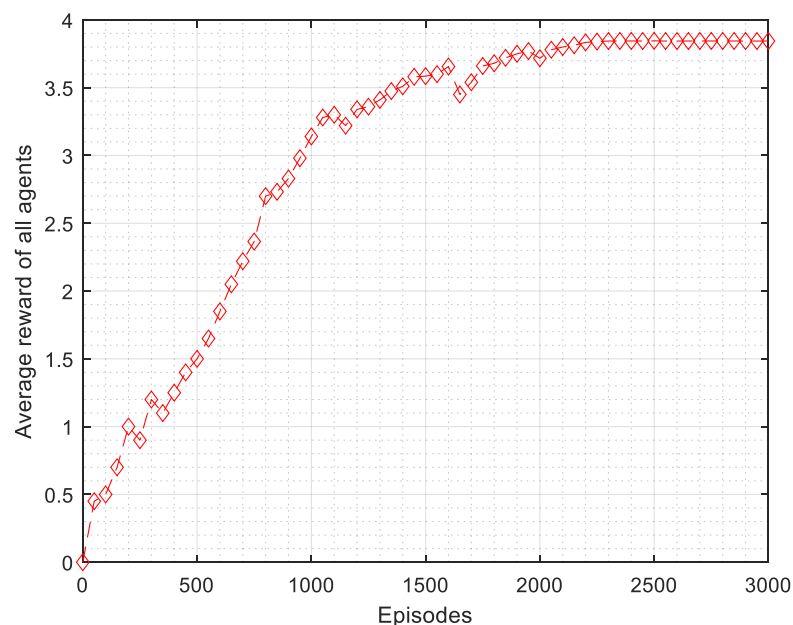


Figure 15. The average reward performance of TA-RL-based CSS.

We listed our research highlights below:

- Two conventional spectrum sensing techniques can sense even at different noise levels to ensure higher accuracy. Thus, between two spectrum sensing techniques, one was chosen using the fuzzy-naïve Bayes algorithm.

- Usage of dynamic threshold values is more accurate, feasible, and adaptive, especially in the CR-VANET environment due to its rapid change and noise uncertainty.
- The management of vehicle density was obtained by merging and splitting segments into sub-segments; a probability value-based division of sub-segments was performed for cooperative spectrum sensing.
- For efficient global decision-making of the spectrums, the tri-agent reinforcement learning algorithm was proposed to learn three different environments and decide the spectrum concerning the collected local sensing reports from the secondary users, i.e., vehicles.

6. Conclusions

This paper introduced a novel Seg-CR-VANET (segment-based cognitive radio vehicular ad hoc network) architecture to achieve better data transmission efficacies for the vehicular environment. The proposed Seg-CR-VANET relies on spectrum sensing and road segmentation management efficiency, which improved the overall network performance. We introduced a novel spectrum sensing technique using a hybrid ML algorithm that combines the fuzzy and naïve Bayes algorithms. The spectrum sensing technique is dynamically chosen based on the current network condition between energy detection and the matched filter. Due to the uncertainty of noise, static threshold value usage is not feasible, which is why we used dynamic threshold values calculated using Tsallis entropy. Based on the sensed reports, a cooperative sensing decision is made with the TA-RL (tri-agent reinforcement learning) algorithm. It is executed by SSA (segment spectrum agent), which is responsible for managing spectrum availability in each segment. The roads are managed by equal segmentation and further sub-segmented dynamically if the vehicle density increases at a certain threshold level. The proposed architecture provides much better results than previous works. We achieved better spectrum detection, throughput, and packet delivery ratio; lower delay and lower packet loss; higher accuracy; and good convergence rate. In the future, we will focus on route optimization by using the 2HMO-HHO (2-Hop Multi-Objective Harris Hawks Optimization) algorithm. We will also focus on the resource allocation scheme for secondary users by considering multiple parameters.

Author Contributions: The conceptualization and the main design schema were made by M.A.H. Methodology of research was discussed by M.A.H., R.M.N., K.-L.A.Y. and S.R.A. Implementation was made by M.A.H., M.R.Z. and I.A. Validation and formal analysis was by M.A.H., R.M.N. and M.R.J. Original draft was prepared by M.A.H., R.M.N. and M.R.J. Visualization was by M.A.H., I.A. and M.R.J. Supervision was by R.M.N., M.R.Z., S.R.A. and K.-L.A.Y. Funding acquisition was by R.M.N. and K.-L.A.Y. All the authors participated in the review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Partnership Grant under Grant CR-UM-SST-DCIS-2018-01 and Grant RK004-2017 between Sunway University and University of Malaya.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon request.

Acknowledgments: The authors would like to thank the anonymous reviewers for their comments and constructive suggestions which helped them to improve this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hossain, M.A.; Noor, R.M.; Yau, K.-L.A.; Azzuhri, S.R.; Z'Abu, M.R.; Ahmedy, I. Comprehensive Survey of Machine Learning Approaches in Cognitive Radio-Based Vehicular Ad Hoc Networks. *IEEE Access* **2020**, *8*, 78054–78108. [[CrossRef](#)]
2. Zhang, D.; Zhang, T.; Liu, X. Novel self-adaptive routing service algorithm for application in VANET. *Appl. Intell.* **2018**, *49*, 1866–1879. [[CrossRef](#)]

3. Krzanich, B. Data is the New Oil in the Future of Automated Driving. 2016. Available online: <https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving/#gs.0j9aua> (accessed on 10 February 2019).
4. Qian, X.; Hao, L. Performance Analysis of Cooperative Sensing over Time-Correlated Rayleigh Channels in Vehicular Environments. *Electronics* **2020**, *9*, 1004. [[CrossRef](#)]
5. Li, J.; Peng, Y.; Yan, Y.; Jiang, X.-Q.; Hai, H.; Zukerman, M. Cognitive Radio Network Assisted by OFDM With Index Modulation. *IEEE Trans. Veh. Technol.* **2019**, *69*, 1106–1110. [[CrossRef](#)]
6. Arjoune, Y.; Kaabouch, N. A Comprehensive Survey on Spectrum Sensing in Cognitive Radio Networks: Recent Advances, New Challenges, and Future Research Directions. *Sensors* **2019**, *19*, 126. [[CrossRef](#)]
7. Hossain, M.A.; Rafidah, N.; Saaidal, R.A.; Muhammad, R.Z.; Ismail, A.; Kok-Lim, A.Y.; Christopher, C. Spectrum sensing challenges & their solutions in cognitive radio based vehicular networks. *Int. J. Commun. Syst.* **2021**. [[CrossRef](#)]
8. Lee, W.; Kim, M.; Cho, D.-H. Deep Cooperative Sensing: Cooperative Spectrum Sensing Based on Convolutional Neural Networks. *IEEE Trans. Veh. Technol.* **2019**, *68*, 3005–3009. [[CrossRef](#)]
9. Liu, X.; Zhang, X.; Ding, H.; Peng, B. Intelligent clustering cooperative spectrum sensing based on Bayesian learning for cognitive radio network. *Ad Hoc Netw.* **2019**, *94*, 101968. [[CrossRef](#)]
10. Zhang, T.; Zhang, D.; Qiu, J.; Zhang, X.; Zhao, P.; Gong, C. A Kind of Novel Method of Power Allocation with Limited Cross-Tier Interference for CRN. *IEEE Access* **2019**, *7*, 82571–82583. [[CrossRef](#)]
11. Benkerdagh, S.; Duvallat, C. Cluster-based emergency message dissemination strategy for VANET using V2V communication. *Int. J. Commun. Syst.* **2019**, *32*, e3897. [[CrossRef](#)]
12. Akyildiz, I.F.; Lo, B.F.; Balakrishnan, R. Cooperative spectrum sensing in cognitive radio networks: A survey. *Phys. Commun.* **2011**, *4*, 40–62. [[CrossRef](#)]
13. Arjoune, Y.; El Mrabet, Z.; El Ghazi, H.; Tamtaoui, A. Spectrum sensing: Enhanced energy detection technique based on noise measurement. In Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 828–834.
14. Chembe, C.; Kunda, D.; Ahmedy, I.; Noor, R.M.; Sabri, A.Q.M.; Ngadi, A. Infrastructure based spectrum sensing scheme in VANET using reinforcement learning. *Veh. Commun.* **2019**, *18*, 100161. [[CrossRef](#)]
15. Wu, G.; Chu, H. Spectrum Sharing with Vehicular Communication in Cognitive Small-Cell Networks. *Int. J. Antennas Propag.* **2020**, *2020*, 6897646. [[CrossRef](#)]
16. Pal, R.; Prakash, A.; Tripathi, R.; Naik, K. Regional Super Cluster Based Optimum Channel Selection for CR-VANET. *IEEE Trans. Cogn. Commun. Netw.* **2019**, *6*, 607–617. [[CrossRef](#)]
17. Khattab, A.; Elgaml, N.; Mourad, H. Single-channel slotted contention in cognitive radio vehicular networks. *IET Commun.* **2019**, *13*, 1078–1089. [[CrossRef](#)]
18. Hill, E.M.; Sun, H. Double Threshold Spectrum Sensing Methods in Spectrum-Scarce Vehicular Communications. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4072–4080. [[CrossRef](#)]
19. Li, H.; Gu, Y.; Chen, J.; Pei, Q. Speed Adjustment Attack on Cooperative Sensing in Cognitive Vehicular Networks. *IEEE Access* **2019**, *7*, 75925–75934. [[CrossRef](#)]
20. Li, X.; Song, T.; Zhang, Y.; Chen, G.; Hu, J. A Hybrid Cooperative Spectrum Sensing Scheme Based on Spatial-Temporal Correlation for CR-VANET. In Proceedings of the 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, Portugal, 3–6 June 2018; pp. 1–6.
21. Li, R.; Zhu, P. Spectrum Allocation Strategies Based on QoS in Cognitive Vehicle Networks. *IEEE Access* **2020**, *8*, 99922–99933. [[CrossRef](#)]
22. Salah, I.; Saad, W.; Shokair, M.; Elkordy, M. Cooperative spectrum sensing and clustering schemes in CRN: A survey. In Proceedings of the 2017 13th International Computer Engineering Conference (ICENCO), Cairo, Egypt, 27–28 December 2017; pp. 310–316.
23. Arshid, K.; Jianbiao, Z.; Hanif, I.; Munir, R.; Yaqub, M.; Tariq, U. Energy Detection Based Spectrum Sensing Strategy for CRN. In Proceedings of the 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS), Dalian, China, 20–22 March 2020; pp. 107–112.
24. Kabeel, A.A.; Hussein, A.H.; Khalaf, A.A.; Hamed, H.F. A utilization of multiple antenna elements for matched filter based spectrum sensing performance enhancement in cognitive radio system. *AEU Int. J. Electron. Commun.* **2019**, *107*, 98–109. [[CrossRef](#)]
25. Störr, H.-P.; Xu, Y.; Choi, J. A compact fuzzy extension of the Naive Bayesian classification algorithm. In Proceedings of the Third International Conference on Intelligent Technologies and Vietnam-Japan Symposium on Fuzzy Systems and Applications, Hanoi, Vietnam, 3–5 December 2002; pp. 172–177.
26. Zhao, Y.; Paul, P.; Xin, C.; Song, M. Performance analysis of spectrum sensing with mobile SUs in cognitive radio networks. In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 2761–2766.
27. Salahdine, F.; El Ghazi, H.; Kaabouch, N.; Fihri, W.F. Matched filter detection with dynamic threshold for cognitive radio networks. In Proceedings of the 2015 International Conference on Wireless Networks and Mobile Communications (WINCOM), Marrakech, Morocco, 20–23 October 2015; pp. 1–6.
28. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [[CrossRef](#)]

-
29. Raman, G.P.; Perumal, V. Neuro-fuzzy based two-stage spectrum allocation scheme to ensure spectrum efficiency in CRN–CSS assisted by spectrum agent. *IET Circuits Devices Syst.* **2019**, *13*, 637–646. [[CrossRef](#)]
 30. Sutton, R.; Barto, A. Reinforcement Learning: An Introduction. *IEEE Trans. Neural Netw.* **1998**, *9*, 1054. [[CrossRef](#)]
 31. Bin Ahmad, H. Ensemble Classifier Based Spectrum Sensing in Cognitive Radio Networks. *Wirel. Commun. Mob. Comput.* **2019**, *2019*, 1–16. [[CrossRef](#)]
 32. Haider, D.; Ren, A.; Fan, D.; Zhao, N.; Yang, X.; Tanoli, S.A.K.; Zhang, Z.; Hu, F.; Shah, S.A.; Abbasi, Q.H. Utilizing a 5G spectrum for health care to detect the tremors and breathing activity for multiple sclerosis. *Trans. Emerg. Telecommun. Technol.* **2018**, *29*, e3454. [[CrossRef](#)]