

## Research Article

# Joint Resource Allocation and Power Control Based on Vehicle's Motion Characteristics in NOMA-Based V2V Systems

Xinyu Gu <sup>1,2</sup>, Yiyi Xu <sup>1</sup>, and Jiahan Zhang <sup>1</sup>

<sup>1</sup>School of Information and Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>2</sup>Purple Mountain Laboratories, Nanjing 211111, China

Correspondence should be addressed to Yiyi Xu; xuyiyi@bupt.edu.cn

Received 25 June 2020; Revised 15 September 2020; Accepted 9 November 2020; Published 29 November 2020

Academic Editor: Alessandro Bazzi

Copyright © 2020 Xinyu Gu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the high spectrum utilization of Nonorthogonal Multiple Access (NOMA), it becomes one of the potential candidate technologies for future wireless communication systems. Meanwhile, in New Radio, Vehicle to Everything (V2X) has been proposed as a promising issue in the 3<sup>rd</sup> Generation Partnership Project (3GPP). This paper studies the resource allocation mechanism with power control strategy which makes full use of vehicles' moving characteristics in the NOMA-based Vehicle to Vehicle (V2V) communication system. Firstly, vehicles are grouped according to their moving characteristics by spectral clustering. Then, vehicles which are in the same group are allocated the same wireless resource with NOMA strategy. Two grouping methods have been designed for freeway and urban scenarios separately. After that, the transmission power of vehicles is adjusted based on the result of power control strategy utilizing Q-learning. The simulation results show that the performance of the V2V system in terms of Packet Received Ratio (PRR) can be evidently improved by the proposed joint NOMA resource allocation and power control mechanism compared to typical energy sensing-based resource allocation method.

## 1. Introduction

With the development of Intelligent Transport System (ITS), Vehicle to Vehicle (V2V) communication has been the center of intensive research for several years. Due to the increasingly rare wireless spectrum resources along with more and more vehicles, resource allocation scheme design has become a focus of the research in both academic and industrial areas.

Among all academic researches, graph theory and optimization theory have been used the most. For example, resource allocation problem was transformed into a maximum weight matching problem in [1], while it was formulated as a three-dimensional matching problem in [2]. In industrial area, the specific group that belongs to the 3<sup>rd</sup> Generation Partnership Project (3GPP) has been doing the standardization work for V2V communications. In Radio Access Network (RAN) 80<sup>th</sup> meeting of 3GPP, New Radio-(NR-) Vehicle to Everything (V2X) has been put forward based on standards specified from Release 14 for Long-Term

Evolution- (LTE-) V2X. In past meetings, 3GPP RAN have discussed several resource allocation mechanisms. These mechanisms include mechanisms involving Base Station (BS) schedule resources among vehicles dynamically and the mechanisms by which vehicles select resources automatically without the aid from BS. These two types of resource allocation mechanisms are referred to as Mode 1 and Mode 2, respectively, in 3GPP [3]. Among all discussed resource selection methods, energy sensing-based mechanism [4] is the most typical one. Otherwise, making the use of vehicles' geographical position has also been another resource selection method put forward by 3GPP [5, 6].

In NR communication system, the whole frequency bandwidth is divided into subcarriers. In time domain, one transmission period includes dozens of time slots. Twelve subcarriers correspond to one Resource Block (RB) in frequency domain. Vehicles transmit signal on specific RBs and on specific time slots. Considering the limited bandwidth that V2V communication can use, for example, 10 MHz, if many vehicles exist or frequent interaction is required, the

wireless resources in RBs and time slots are not enough. Thus, resource collision where some vehicles have to transmit signals on the same RB greatly degrades the communication performance in the V2V system.

To alleviate performance degradation caused by resource collision, a promising technique that enables several vehicles to share the same resource to transmit messages and ensures these transmitting vehicles can still be distinguished by the receivers is put forward. This technique is Nonorthogonal Multiple Access (NOMA). With power domain NOMA, different transmitters can transmit their messages on the same RB simultaneously with different level of power, while receivers apply Successive Interference Cancellation (SIC) to decode received signals in the decreasing order of channel gains [7]. Therefore, it is suitable to apply NOMA in V2V communication system for its high spectrum utilization [8]. However, as far as authors know, there are few researches about applying NOMA in V2V communication [9–11].

When applying NOMA to V2V communication system, the key step is grouping vehicles. Vehicles in the same group will be allocated to the same resource. A straightforward idea is to make vehicles in the same group far from each other. As V2V communication system is a dynamic scenario, the movement of vehicles should be considered in grouping, which includes moving characteristics, for example, moving direction, speed, distance, and communication link type. Considering taking advantage of vehicles' different features in vehicle grouping process without resource allocation information, unsupervised clustering algorithm is desirable. After clustering, vehicles in the same cluster share the same wireless resource in NOMA manner.

Based on resource allocation results, power control is further conducted. The aim of power control is to maximize the Packet Reception Ratio (PRR) of receivers in the coverage range of transmitter. It is well known that Q-learning is a good method to learn in a new environment and give the nice strategy. Several papers have applied it into power control for Femtocell communication and D2D communication [12, 13] and we will investigate the feasibility of applying the Q-learning algorithm in vehicle power control in the NOMA-based V2V communication system.

The rest of the paper is organized as follows. In Section 2, system model and problem formulation of the V2V communication system are described. Resource allocation mechanism and power control strategy are proposed in Section 3. In Section 3.1, grouping method in freeway has been introduced. Section 3.2 shows resource allocation in urban scenario based on spectral clustering. Section 3.3 is the power control strategy based on Q-learning. Section 3.4 illustrates the signaling process for resource allocation mechanism. Then, simulation scenarios, simulation results, and computational complexity analysis are presented in Section 4. Finally, conclusion is given in Section 5.

## 2. System Model and Problem Formulation

**2.1. System Model.** In V2V broadcast system, each vehicle broadcasts its messages, while others receive messages and try to decode them as shown in Figure 1.  $T_1$  and  $T_2$  transmit

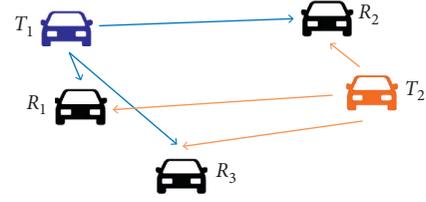


FIGURE 1: The communications between vehicles.

their messages in the form of data packets;  $R_1$ ,  $R_2$ , and  $R_3$  attempt to receive packets from  $T_1$  as well as packets from  $T_2$ .  $\text{SINR}_{ij}$  which denotes the Signal to Interference plus Noise Ratio (SINR) at receiver  $j$  from transmitter  $i$  can be calculated by

$$\text{SINR}_{ij} = \frac{P_i |h_{ij}|^2}{N_0 B_0 + \sum_{k=1, k \neq i}^N \alpha_{ik} s_{i,k}^j P_k |h_{kj}|^2}, \quad (1)$$

$$\alpha_{ik} = \begin{cases} 1, & i, k \text{ use same RB,} \\ 0, & \text{other,} \end{cases} \quad (1a)$$

$$s_{i,k}^j = \begin{cases} 1, & \frac{P_i |h_{ij}|^2}{n_{ij}} > \frac{P_k |h_{kj}|^2}{n_{kj}}, \\ 0, & \frac{P_i |h_{ij}|^2}{n_{ij}} < \frac{P_k |h_{kj}|^2}{n_{kj}}. \end{cases} \quad (1b)$$

In (1),  $P_i$  is the transmitting power of transmitter  $i$ ,  $|h_{ij}|^2$  denotes channel coefficient from transmitter  $i$  to receiver  $j$ ,  $N_0$  is the one-sided power spectral density of Additive White Gaussian Noise (AWGN),  $B_0$  is the bandwidth that transmitter  $i$  uses to transmit messages, and  $N$  represents the number of vehicles in V2V communication system. Binary variable  $\alpha_{ik}$  equals one when transmitters  $i$  and  $k$  transmit messages on the same wireless resource at the same time (1a). Another binary variable  $s_{i,k}^j$  shows the result of applying SIC. When  $(P_i |h_{ij}|^2 / n_{ij}) > (P_k |h_{kj}|^2 / n_{kj})$ , vehicle  $j$  decodes message from vehicle  $i$  firstly and decodes message from vehicle  $k$  afterwards. In this situation, binary variable  $s_{i,k}^j$  equals one, which means that interference from transmitter  $k$  exists. Otherwise,  $s_{i,k}^j$  equaling zero represents interference cancelled with SIC (1b).

In Figure 1, when  $T_1$  and  $T_2$  use the same resource in NOMA manner,  $R_1$  can decode the packets from both  $T_1$  and  $T_2$  successfully. Because  $R_1$  is close to  $T_1$  and far from  $T_2$ , it can decode message with the optimal order of decreasing channel gains normalized by the noise. More specifically,  $R_1$  decodes the signal  $s_1$  from  $T_1$  firstly, while it treats the signal  $s_2$  from  $T_2$  as interference when it receives signal  $s$  from both  $T_1$  and  $T_2$ . After signal  $s_1$  from  $T_1$  is decoded, it is cancelled in the signal  $s$  before  $R_1$  decodes signal  $s_2$  from  $T_2$  subsequently, which means it subtracts  $s_1$  from  $s$  and decodes signal  $s_2$  after. Therefore, the SINR of receiver  $R_1$  when decoding signal from  $T_2$  is  $\text{SINR}_{21} = ((P_2 |h_{21}|^2) / (N_0 B_0))$  rather than  $\text{SINR}_{21} = ((P_2 |h_{21}|^2) / (N_0 B_0 + P_1 |h_{11}|^2))$ , and so does SINR of  $R_2$ . However,  $R_3$

cannot decode any signal at strong possibility, neither from  $T_1$  nor from  $T_2$ , because it receives almost the same power of signals from  $T_1$  and  $T_2$ .

In V2V communication system, safety related messages are broadcasted by vehicles periodically. In 3GPP, PRR is defined as the typical performance metric, which is the statistical average of the probability of all packets received successfully. In essence, the probability of packets decoded correctly depends on the SINR at the receiver. The higher PRR is, the more reliable communications between vehicles are.

**2.2. Problem Formulation and Analysis.** In the regulation of 3GPP, one communication type is vehicle as transmitter broadcasts its messages, while other vehicles are as receivers receiving messages. As mentioned above, PRR is the statistical average of the probability of all packets to be decoded successfully. Hence, the objective for resource allocation and power control is maximizing PRR of all vehicles. Furthermore, since the probability of packets to be correctly decoded basically depends on the SINR at the receiver, the goal alerts to maximize the total SINR of all vehicles. The relationship between PRR and SINR is shown in Figure 2, which illustrates the feasibility to replace PRR with SINR as the objective of resource allocation and power control:

$$\max_{\{x_{m,t}^i\}} \sum_{t=1}^T \sum_{m=1}^M \sum_{j \in \text{Re}} \sum_{i \in \text{Tr}} \frac{P_i |h_{ij}|^2 x_{m,t}^i}{N_0 B_0 + \sum_{k=1, k \neq i}^N s_{i,k}^j P_k |h_{kj}|^2 x_{m,t}^k}, \quad (2)$$

$$\sum_{t=1}^T \sum_{m=1}^M x_{m,t}^i = 1, \quad \forall i = 1, 2, \dots, N, \quad (2a)$$

$$\sum_{i=1}^N x_{m,t}^i \leq c, \quad \forall m = 1, 2, \dots, M, \forall t = 1, 2, \dots, T, \quad (2b)$$

$$x_{m,t}^i = \begin{cases} 1, & i \text{ uses subchannel } m \text{ in slot } t, \\ 0, & \text{other,} \end{cases} \quad (2c)$$

$$s_{i,k}^j = \begin{cases} 1, & \frac{P_i |h_{ij}|^2}{n_{ij}} > \frac{P_k |h_{kj}|^2}{n_{kj}}, \\ 0, & \frac{P_i |h_{ij}|^2}{n_{ij}} < \frac{P_k |h_{kj}|^2}{n_{kj}}, \end{cases} \quad (2d)$$

$$P_i \leq P_{\max}, \quad \forall i = 1, 2, \dots, N. \quad (2e)$$

Assume that there are  $N$  vehicles in the V2V system. The total available bandwidth is divided into  $M$  subchannels (one subchannel consists of several RB) and one transmission period is divided into  $T$  slots. In this paper, one subchannel in frequency domain and one time slot in time domain are defined as a Resource Block Group (RBG) as the basis for resource allocation. Therefore, there are totally  $M * T$  RBGs in one period. Assume that BS has perfect channel state information of all vehicles in the system via dedicated

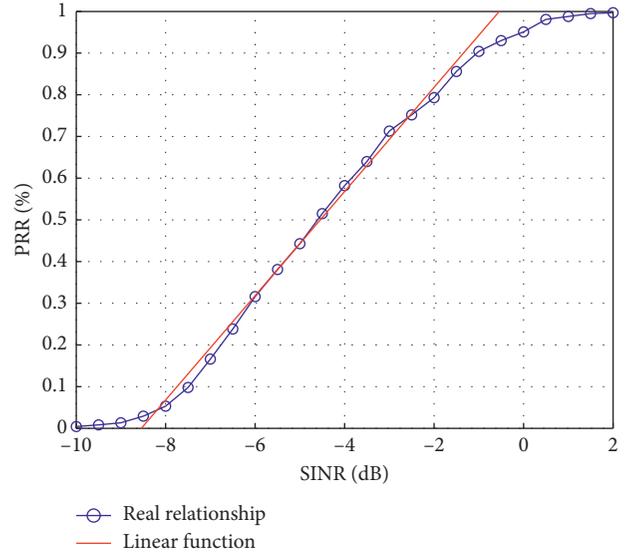


FIGURE 2: Relationship between PRR and SINR.

feedback channels. Each vehicle  $i$  transmits packets with power  $P_i$  on the RBG, which is allocated by BS. All of vehicles except transmitters decode messages with the optimal order of decreasing channel gains as discussed in Section 2.1. For each slot, specific subchannels are allocated to vehicles to transmit packets and receivers' SINR are calculated. The total SINR of all vehicles within one period can be calculated by (2).

In this formulation, set  $Tr$  represents vehicles using subchannel  $m$  in slot  $t$  to transmit messages, while set  $Re$  represents other vehicles receiving messages (2). Constraint (2a) corresponds to the transmitting limitation that one vehicle should transmit packets in one period once with only one subchannel, which means each vehicle uses one RBG in each period for transmission. Constraint (2b) means each RBG can only be assigned to at most  $c$  users in NOMA manner, which is also the maximum number of vehicles using the same RBG. The larger  $c$  is, the more complicated the process of decoding and receiver can be.  $c$  equaling one means no resource collision happens. Binary variable  $x_{m,t}^i$  equals one only when subchannel  $m$  in slot  $t$  is allocated to vehicle  $i$  (2c). Constraint (2d) shows the result of using SIC in NOMA manner: when  $(P_i |h_{ij}|^2 / n_{ij}) > (P_k |h_{kj}|^2 / n_{kj})$ , binary variable  $s_{i,k}^j$  equals one, which means interference from transmitter  $k$  exists;  $s_{i,k}^j$  equaling zero representing interference has been cancelled with SIC. Constraint (2e) shows the transmission power of each vehicle should not exceed the maximum power  $P_{\max}$ .

The optimization objective shown in (2) is a nonconvex problem and also a NP-hard problem because variables are binary and the existence of interference. However, the upper boundary of (2) can be given and that is the situation without interference between transmitting vehicles. Because of the existence of interference, this upper boundary can never be achieved. Therefore, a joint resource allocation and power control algorithm based on machine learning to solve (2) is proposed in the next section.

### 3. Resource Allocation Mechanism and Power Control Strategy

It is not easy to find the optimal solution with such constraints. Even with the greedy method, the optimal solution at this moment does not mean that it is the optimal solution at next moment as well because the distance between each two vehicles is changing from time to time. To simplify the problem in (2), it is decoupled into two stages.

At the first stage, assuming that transmit powers of all vehicles are the same, resource allocation mechanism based on vehicles' moving characteristics is conducted. At the second stage, power control is further done according to the resource allocation results obtained at stage one.

In this paper, two typical scenarios, freeway and urban, are considered, corresponding to comparatively simple and complicated vehicle traffic conditions. In different traffic scenarios, the importance of different moving characteristics of vehicles is different. For example, vehicles running on highway rarely change their moving statuses such as direction and speed. In most cases, link type between a transmitter and a receiver is Line of Sight (LOS). However, in urban scenario, vehicles may change their driving directions at any cross and their speed can change every now and then due to, for example, traffic jam and the control of traffic light. Even the link type between two vehicles may change because of the probability of building block. Therefore, different users grouping methods are proposed in this paper to cover different scenarios.

**3.1. Resource Allocation Mechanism for Freeway.** As mentioned above, the key step in the design of resource allocation mechanism is user grouping. In freeway scenario, there are some considerations in design. Firstly, vehicles which are as far away as possible should be chosen to use the same RBG in NOMA manner. The reason is that when two transmitters that are allocated to same resource are near to each other, their neighbors as receivers cannot receive and decode packets successfully from neither of them because of the large interference between them. To reduce the occurrence of the above situation, a parameter  $\gamma$  is defined to represent the minimum distance between vehicles sharing the same resource.

In user grouping, receivers which have similar distance to the transmitting vehicles should also be taken into account. The reason is that such kind of receivers usually cannot decode the messages from any of the transmitters; for example,  $R_3$  in Figure 1 cannot decode message from neither  $T_1$  nor  $T_2$  at great possibility.

A centralized scheduling mechanism that groups the vehicles according to their moving features is proposed based on the above considerations. The vehicles in the same group are allocated to the same resource in proposed resource scheduling algorithm. Vehicles in the same group are expected to have similar speed and similar moving direction. Furthermore, the distance between them should be larger than  $\gamma$ . By this means, distance between members in the same group is relatively stable and the interference caused by

close distance between transmitters sharing the same resource can be avoided.

The detailed resource allocation algorithm is described as follows: In the first step, all vehicles are divided into several categories according to their speed and direction. Vehicles having similar direction and speed are in the same category  $G_f$ . The roads in freeway are mainly designed for two or three kinds of vehicles' speed, like carriageway and passing lane for relatively slow and fast vehicles, respectively. Thus, there are two kinds of speeds adopted in simulation in Section 4.1 to simulate freeway in reality. The more dispersed the vehicle speed is, the greater the number of the category  $G_f$  will be, and the less vehicles each category has and vice versa. The following steps are done in each category  $G_f$  and the resource allocation algorithm will end until each vehicle belongs to a group. Hence, there is no influence on resource allocation mechanism proposed in the paper no matter how dispersed or intensive the vehicles' speed is.

In the second step, we will decide which vehicles can be chosen to be in the same group. Vehicle  $j$  is randomly chosen at the beginning of the algorithm, and suppose that vehicle  $j$  is in the  $r^{\text{th}}$  group,  $g_r$ . Vehicles that can be in the same group as vehicle  $j$  should be in the same category as vehicle  $j$ . Then we check each vehicle in this category according to (3) and the vehicle with the maximum argument is selected to be in the same group as vehicle  $j$ . Suppose that there are  $N$  vehicles in the system; (3) ensures that vehicle  $i$  is in the same category as  $j$  and is far away from  $j$ . At the same time, the number of receivers which have similar distance to transmitters using the same resource is minimized.

$$\arg \max_{\{i\}} \left( x_{ij} dm_{ij} - \beta \sum_{j \in g_r} \sum_{k=1, k \neq i, j}^N f_{kij} \right), \quad (3)$$

$$dm_{ij} = \min d_{ix}, \quad j \in G_f, \forall x \in G_f, \quad (3a)$$

$$f_{kij} = \begin{cases} 1, & |d_{kj} - d_{ki}| < \alpha, \\ 0, & |d_{kj} - d_{ki}| \geq \alpha, \end{cases} \quad (3b)$$

$$x_{ij} = \begin{cases} 1, & i, j \text{ in the same category,} \\ 0, & \text{other,} \end{cases} \quad (3c)$$

$$d_{ij} > \gamma. \quad (3d)$$

In (3),  $dm_{ij}$  denotes the minimum distance between vehicle  $i$  and vehicles in the same group with  $j$  (3a);  $\sum_{k=1, k \neq i, j}^N f_{kij}$  denotes the number of vehicles that have the almost equal distance to vehicle  $i$  and vehicle  $j$ ;  $d_{ij}$  denotes the distance between vehicle  $i$  and vehicle  $j$ ; binary variable  $f_{kij}$  equals one when the difference between distance from receiver  $k$  to transmitter  $i$  and distance from receiver  $k$  to transmitter  $j$  is within  $\alpha m$  and such kind of receiver  $k$  basically cannot decode messages from any transmitters correctly (3b); binary variable  $x_{ij}$  equals one when  $i, j$  are in the same category (3c); vehicle  $i$  should be far away from  $j$  and the distance between them should be larger than  $\gamma$  (3d). Considering the number of vehicles that have similar

distance between the two transmitters can be a quite small number compared to the distance between two transmitters,  $f_{kij}$  is suitably magnified through multiplying by  $\beta$  ( $\beta > 1$ ).

Repeat searching until no vehicle in the same category satisfies (3). Then the vehicle in the same category with  $j$  which has the minimum distance to all vehicles in the previous group is regarded as the first element of the next group. Repeat checking the satisfaction of (3) and finding the first member of next group until all vehicles in this category are chosen into groups. Then another category is chosen and the above steps are repeated until all the vehicles are the members in groups.

Vehicles in the same group use the same resource. If the number of groups is larger than the number of resources, the value of  $\gamma$  is repeatedly decreased and the vehicles are grouped until the number of groups is slightly less than the number of resources.

In order to get better performance with scarce spectrum, the minimum distance between vehicles sharing the same resource is calculated by the above steps, which means that distance depends on the amount of wireless resources. If wireless resources are adequate for vehicles information transmission, the minimum distance between vehicles sharing the same resource will be larger than that under the situation with little wireless resources.

The detailed description of Algorithm 1, vehicles grouping algorithm, is shown below.

The vehicle keeps transmitting on the allocated resource until it changes its motion status, like leaving the road or changing the speed, direction, and so forth. Once things like the above happen, it leaves the original group and needs to be regrouped. The group that satisfies (3) is the group for it to join in. If no group satisfies (3), it becomes the only member of a new group when idle resources exist. When no group satisfies (3) and no resource is idle, the group that satisfies (3) but relaxes (3d) can be the group for it to join in. The whole RBG reallocation algorithm for vehicles changing their moving status is shown in Algorithm 2.

**3.2. Resource Allocation Mechanism for Urban Scenario.** In urban scenario, the motion characteristics of vehicles are more complicated than those in freeway scenario. Vehicles change driving direction and speed frequently and the link type between two vehicles may change unpredictably due to, for example, blocking by a building or tree. In user grouping algorithm, these characteristics should be taken into account in addition to distance considered in freeway scenario.

Spectral clustering (SC) as an unsupervised method can partition the data into different groups according to multiple features. In most cases, the superiority of SC is attributed to the design of a metric function and the affinity graph [14]. As for the constant and unpredictable vehicle traffic changes in urban scenario, SC is therefore considered to solve the resource allocation problem while regarding vehicles as data points in this paper. Vehicles' geographical position, driving direction, speed, and communication link type are the main motion features taken into consideration.

Apart from selecting proper motion features, appropriate weight should also be built upon each feature with the consideration of their influence on communication between vehicles to design metric function. Proper grouping vehicles method in urban scenario needs quantizing features about vehicles and building appropriate weight among vehicles to assess the similarity of vehicles. According to similarity, the vehicles are clustered by, for example, BS during each transmitting period. After clustering, vehicles in the same cluster share the same resource in NOMA manner.

In the beginning of the next part, a brief introduction of normalized cut which is the major step in SC is given. After that, features selection and metric establishment are described and resource allocation algorithm in urban scenario is proposed.

**3.2.1. Brief Introduction of Normalized Cut.** When inputting a data set  $X$  with  $nd$ -dimensional samples, like  $X = \{x_1, x_2, \dots, x_n\} \in R^{d \times n}$ , the clustering algorithm can group  $X$  into  $k_2$  clusters  $c_i, i = 1, \dots, k_2$  with the aim of keeping data within the same cluster close to one another and data points from different clusters remain apart. That is to say, normalized cut not only minimizes weight of edges between different clusters but also maximizes weight of edges within cluster [15].

The main steps of normalized cut are shown as follows:

- (1) Construct similarity matrix  $S$ .
- (2) Construct adjacency matrix  $W$  and degree matrix  $D$ .
- (3) Calculate Laplace matrix  $L$ .
- (4) Do standardization of  $L$ .
- (5) Compute the smallest  $k_1$  eigenvalues of  $L$  and their respective eigenvectors  $f$ .
- (6) Normalize matrix  $L$  composed of corresponding eigenvector  $f$  by row and form the characteristic matrix  $F$  of  $n \times k_1$ . Each row in  $F$  is a  $k_1$ -dimensional sample.
- (7) Cluster the above  $n$  samples by, for example,  $k$ -means, and the clustering dimension is  $k_2$ .
- (8) Get cluster partition  $C = \{c_1, c_2, \dots, c_{k_2}\}$ .

### 3.2.2. Features Selection

**Feature 1: Distance between vehicles.** Similar to freeway scenario, the relative position among vehicles, namely, distance, is still important. In vehicles grouping, vehicles sharing the same resource are expected to be as far away as possible from each other. Weight calculating from distance feature is designed to be proportional to distance. The closer vehicles are, the less likely they are in the same cluster.

**Feature 2: Speed of vehicles.** Speed is another important feature because the speed of vehicles affects the distance between them dynamically. When there is a speed difference between vehicles, the distance between them changes drastically continuously. This brings about uncertainty of the distance. For instance, two vehicles

Input:

The sets of vehicles  $G_f$  based on the moving features;  $u = |G_f|$ ;  $d_{ij}$  denotes distance between  $i$  and  $j$ ;  $N$  is the number of vehicles;

Output:

The group  $g_r$  in which vehicles use the same RBG;

- (1)  $r = 1, g_r = \emptyset$
- (2) for each  $f \in [1, u]$  do
- (3) randomly choose  $j \in G_f, G_f = G_f \setminus \{j\}, g_r = g_r \cup \{j\}$
- (4) repeat
- (5) for each  $i \in G_f$  do
- (6) if  $i$  satisfies (3) then
- (7)  $g_r = g_r \cup \{i\}, G_f = G_f \setminus \{i\}$
- (8) end if
- (9) end for
- (10)  $r = r + 1, g_r = \emptyset$
- (11) for each  $i \in G_f$  do
- (12) if  $i = \arg_{\{i\}} \min d_{iy}, \forall y \in \{g_{r-1}\}$  then
- (13)  $G_f = G_f \setminus \{i\}, g_r = g_r \cup \{i\}$
- (14) end if
- (15) end for
- (16) until ( $G_f = \emptyset$ )
- (17) end for

ALGORITHM 1: Vehicles grouping algorithm.

Input:

Vehicle  $i$  which changes its moving status;  $m = |g_r|$  is the number of  $g_r$  depends on Algorithm 1

Output:

The new group  $i$  belongs to;

- (1) for each  $r \in [1, m]$  do
- (2) if  $i \in g_r$  then
- (3)  $g_r = g_r \setminus \{i\}$
- (4) end if
- (5) end for
- (6) if  $g_r$  satisfies (3) then
- (7)  $g_r = g_r \cup \{i\}$
- (8) else if  $m < \text{number of RBG}$  then
- (9)  $g_{r+1} = \{i\}, m = m + 1$
- (10) else if  $g_r$  satisfies (3) without (3d) then
- (11)  $g_r = g_r \cup \{i\}$
- (12) end if

ALGORITHM 2: RBG reallocation algorithm.

are far away from each other at the beginning and move in the same direction. If they are allocated to the same resource without considering their speed, their distance may become smaller after a short time if the back vehicle is faster than the front one. Thus, weight should be proportional to the difference of the speed inversely, which ensures that vehicles with large speed difference are in the same cluster with low probability. Considering that urban is the scenario with plenty of vehicles with different speeds, Simulation of Urban Mobility (SUMO) is used in Section 4.2. to model the real urban scenario. The speed of each vehicle is the result of normal running and complying with traffic rules.

Feature 3: Moving direction of vehicles. Similar to speed, moving direction also affects the change in

distance between vehicles, like approaching or separating.

Feature 4: Type of communication link between vehicles. The link types we refer to here are LOS and Non-Line-of-Sight (NLOS). If the link between two vehicles is blocked by buildings or other obstacles, the link is regarded as NLOS. Otherwise, it is LOS. When type of communication link is NLOS between transmitting vehicles, the communication links' types between receiving vehicle and each transmitting vehicle are more likely to be different. Thus, the receiving power values from different transmitting vehicles are more likely to have larger difference, which is helpful for receiving vehicle to decode the messages from both transmitters in NOMA manner. Therefore, the vehicles with NLOS

link are expected to share the same resource to reduce interference.

**3.2.3. Metric Function Establishment.** There are several strategies to construct adjacency matrix. Among all strategies, the most common way to compute the adjacency matrix, namely, weight matrix, is full connection (see the following equations):

$$w_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right), \quad (4)$$

$$W = \{w_{ij}\}.$$

Consider that different features have different effects on V2V communication. The method to compute weight matrix is designed upon different features.

$$w_{ij} = w_{ij}^{\text{dis}} \cdot \sqrt{(X_i^{\text{sd}} - X_j^{\text{sd}})^T \Sigma^{-1} (X_i^{\text{sd}} - X_j^{\text{sd}})} \cdot w_{ij}^l, \quad (5)$$

$$w_{ij}^{\text{dis}} = \begin{cases} \exp\left(\frac{d_{ij}}{\sigma_2} - 1\right), & d_{ij} > \sigma_1, \\ 0, & d_{ij} \leq \sigma_1, \end{cases} \quad (5a)$$

$$w_{ij}^l = \begin{cases} \sigma_3, & l_{ij} = \text{NLOS}, \\ \sigma_4, & l_{ij} = \text{LOS}. \end{cases} \quad (5b)$$

In (5),  $X_i^{\text{sd}} = [v_i, d_i]$ , where  $v_i$  is the numerical value of vehicle  $i$ 's speed and  $d_i$  is the numerical value of vehicle  $i$ 's driving direction; four moving directions, up, down, left, and right, are, respectively, represented by 1, 2, 3, and 4;  $\Sigma$  is two-dimensional covariance matrix in  $\sum_{ij} = E[(X_i^{\text{sd}} - \mu_i)(X_j^{\text{sd}} - \mu_j)]$ ,  $\mu_i = E(X_i^{\text{sd}})$ . The similarity of vehicle speed and moving direction between vehicle  $i$  and vehicle  $j$  is measured by Mahalanobis distance  $\sqrt{(X_i^{\text{sd}} - X_j^{\text{sd}})^T \Sigma^{-1} (X_i^{\text{sd}} - X_j^{\text{sd}})}$  which ensures that each variable is independent of the measurement scale.  $w_{ij}^{\text{dis}}$  represents the weight calculated from distance,  $d_{ij}$  is the distance between  $i$  and  $j$ ,  $\sigma_1$  and  $\sigma_2$  are fixed, and  $\sigma_2$  is bigger than  $\sigma_1$ . The closer the distance between vehicles sharing the same resource is, the more interference there will be for receivers receiving messages from any of them. As can be seen in [16], interference is relatively small when the distance between transmitting vehicles using the same resources is larger than 300 meters. Therefore, 300 meters is used as demarcation point  $\sigma_1$  of piecewise weight calculation formula derived from distance. Thus, here,  $\sigma_1 = 300$  and  $\sigma_2 = 600$ ; the value of  $w_{ij}^{\text{dis}}$  is large when the distance is large; when  $d_{ij}$  is bigger than  $\sigma_1$ ,  $w_{ij}^{\text{dis}}$  is equal to  $\exp((d_{ij}/\sigma_2) - 1)$ ; when  $d_{ij}$  is smaller than  $\sigma_2$  and bigger than  $\sigma_1$ ,  $w_{ij}^{\text{dis}}$  is smaller than one; when  $d_{ij}$  is bigger than  $\sigma_2$  and bigger than  $\sigma_1$ ,  $w_{ij}^{\text{dis}}$  is bigger than one; otherwise  $w_{ij}^{\text{dis}}$  is equal to zero. Because the multiplication of  $w_{ij}^{\text{dis}}$  and

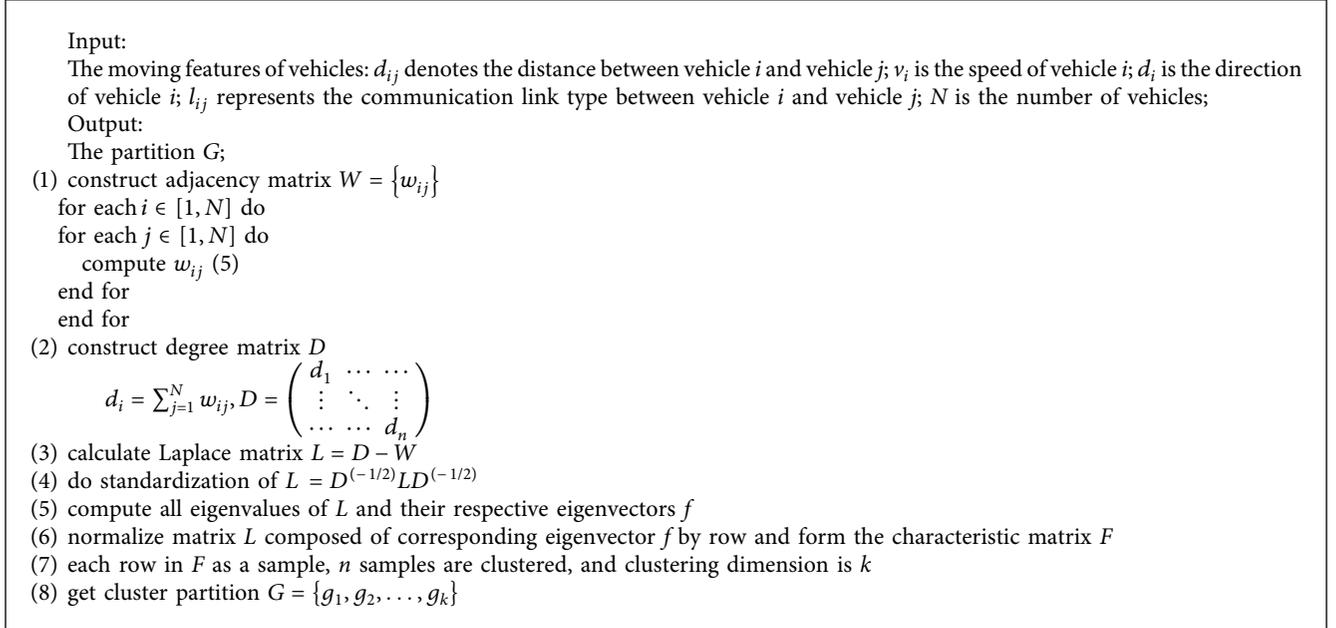
other terms is equal to  $w_{ij}$ ,  $w_{ij}^{\text{dis}} > 1$  means that the possibility of vehicle  $i$  and vehicle  $j$  to be divided into the same group is increased and the possibility of them using the same wireless resources for transmission is increased accordingly. Otherwise, it means that they are less likely to be divided into the same group. Because vehicles are in half duplex mode, the vehicles within a certain range cannot send information at the same time, or they definitely cannot receive information from the others. Thus, when  $d_{ij} \leq \sigma_1$ ,  $w_{ij}^{\text{dis}} = 0$  means that the possibility of vehicle  $i$  and vehicle  $j$  being divided into a cluster and using the same wireless resources for transmission is basically zero (5a). In (5b),  $w_{ij}^l$  refers to the weight calculated according to the type of communication link between vehicle  $i$  and vehicle  $j$ ,  $l_{ij}$  represents the type of communication link between vehicle  $i$  and vehicle  $j$ , and  $\sigma_3$  ( $\sigma_3 > 1$ ) and  $\sigma_4$  ( $\sigma_4 < 1$ ) are fixed, such as  $\sigma_3 = 1.2$  and  $\sigma_4 = 0.9$ ; when the communication link from vehicle  $i$  to vehicle  $j$  is NLOS, the possibility of them using the same wireless resources for information transmission is increased; otherwise, it is decreased.

**3.2.4. Resource Allocation Mechanism.** Algorithm 3 (vehicle clustering algorithm) based on normalized cut shown below is conducted in the beginning of each transmission period. Vehicles in the same partition share the same resources; therefore, the number of clusters  $k$  equals the number of RBGs.

**3.3. Signaling Process for Resource Allocation Mechanism.** Resource allocation mechanism proposed in this paper is done by BS, which means that BS needs to collect information about vehicles. Consensus has been formed that NR supports that users (UE) report assistance information to the generation Node B (gNB) after the discussion in 3GPP 94<sup>th</sup> meeting [17]. The whole signaling process is shown in Figure 3. In the beginning, each vehicle reports geographic and speed related information to BS. Next, BS executes resource allocation mechanism. Then, BS transmits message about resource allocation results to vehicles and vehicles receive that message. After decoding corresponding message, vehicles obtain which RBG it can use to transmit message to other vehicles.

**3.4. Power Control Based on Q-Learning.** Q-learning is a model-free reinforcement learning algorithm, and it can learn to find optimal policy through maximizing expected reward. It shows very good performance in the complex system. In this paper, Q-learning is introduced to solve the power control problem.

In Q-learning, state, action, and reward are three main elements. Detailed contents about Q-learning can be found in [18]. In this NOMA-based V2V power control problem, state is formulated as the set of transmission power values of all transmitting vehicles using the same wireless resources. In order to limit the number of elements in the state set, it is assumed that the vehicle can use one possible discrete power value in  $P$  to send information (see the following equation):



ALGORITHM 3: Vehicle clustering algorithm.

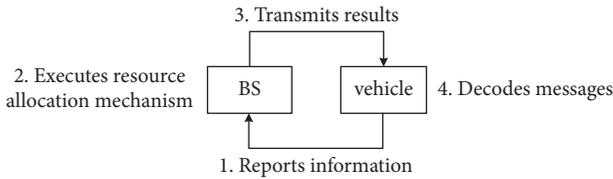


FIGURE 3: Signaling process for resource allocation mechanism.

$$P = \{p_1, p_2, \dots, p_L\}, \quad p_i \leq P_{\max}, 1 \leq i \leq L. \quad (6)$$

Because the upper limit of the transmission power value of each vehicle is  $P_{\max}$ , each possible power value  $p_i$  is smaller than that. When  $K$  vehicles use the same resource, the state set  $S$  is expressed as follows:

$$S = \{(P_1, P_2, \dots, P_K)\}, \quad P_i \in P, 1 \leq i \leq K. \quad (7)$$

Action set is the change of transmission power of all transmitting vehicles using the same resource. In order to limit the number of elements in action set, the change of the transmission power value of each transmitting vehicle can only be reduced, increased, and kept unchanged, which are represented by  $-1$ ,  $0$ , and  $1$ , respectively. When  $K$  vehicles use the same resource, there are  $3^K$  possible actions. At this time,  $a_i$  is the power changing action of vehicle  $i$ , and action set  $A$  is represented by the following equation:

$$A = \{(a_1, a_2, \dots, a_K)\}, \quad a_i \in \{-1, 0, 1\}, 1 \leq i \leq K. \quad (8)$$

The upper limit of transmitting power value of each vehicle is  $P_{\max}$ . In the circumstance that vehicle uses the power value  $p_i$  to transmit messages, if continuously increasing  $p_i$  causes transmission power to be greater than  $P_{\max}$ , the action of increasing the transmission power will not be taken; similarly, if continuously decreasing  $p_i$  causes

transmission power to be equal to or less than zero, the action of reducing the transmission power will not be taken.

Reward is the sum of SINR when other vehicles receive messages from transmitting vehicles since the possibility of information being decoded successfully is proportional to the SINR. When  $K$  vehicles use the same resource, the sum of SINR when receiving vehicles receive information is as follows:

$$r = \sum_{j=1, j \neq i}^N \sum_{i=1}^K \frac{P_i |h_{ij}|^2}{N_0 B_0 + \sum_{k=1, k \neq i, k \neq j}^K s_{i,k}^j P_k |h_{kj}|^2}. \quad (9)$$

Greedy search which achieves the balance between exploration and utilization is used. Explore rate  $\epsilon$  equals one and decreases gradually. The action is selected based on the information that has already known when the random number generated in each step is bigger than  $\epsilon$ ; otherwise, the action is chosen randomly. The power control strategy learning process based on Q-learning is shown in Algorithm 4. When estimated value functions of the  $t$ -th  $Q_t(s_t, a_t)$  sampling and  $t + 1$ -th sampling  $r_{t+1}$  are known,  $Q_{t+1}(s_t, a_t)$  can be obtained by incremental summation (10), where  $\alpha_t$  is the learning rate, which indicates how fast to give up the old value.

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t [r_{t+1} - Q_t(s_t, a_t)]. \quad (10)$$

After learning, the action that gets the maximum Q value is the power control strategy.

## 4. Simulation Results

In this section, the proposed joint resource allocation and power control mechanism is evaluated through system-level simulation.

**4.1. Freeway Scenario.** In this section, the proposed mechanism under the freeway scenario defined by 3GPP [19]

```

(1) initialize Q-table to zeros
(2) for time  $t$  do
(3)   if  $\text{rand}(\cdot) < \epsilon$  then
(4)     select action randomly
(5)   else
(6)     choose action  $a_{t+1} = \arg_{a_{t+1}} \max Q(s_{t+1}, a_{t+1})$ 
(7)     end if
(8)   calculate reward value as (9)
(9)   update Q-table as (10)
(10)   $t = t + 1$ 
(11)  end for
(12) choose  $a = \arg_a \max Q(s, a)$ 

```

ALGORITHM 4: Q-learning based power control algorithm.

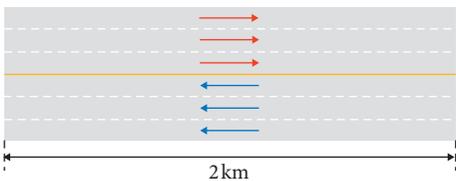


FIGURE 4: Road configuration for simulation.

TABLE 1: Simulation parameters of freeway scenario.

Parameters	Value
Carrier frequency	6 GHz
Bandwidth	10 MHz
Thermal noise density	-174 dBm/Hz
Transmission power	23 dBm
Number of vehicles	230
Absolute speed	140 km/h, 70 km/h
Number of slots in one period	100
Number of subchannels	2
Length of lanes	2000 m

shown in Figure 4 is evaluated. The major simulation parameters are shown in Table 1.

Here, sensing mechanism is selected as the comparison algorithm, which is one kind of resource allocation scheme regulated by 3GPP. Each vehicle senses the energy on every RBG and ranks them from low to high. When reallocation happens, it converts to a RBG which is one of RBGs corresponding to the lowest 20% energy ones as long as the 20<sup>th</sup> minimum value of energy of RBGs is 3 dB less than the energy value of RBG it is using now [1]. This mechanism needs vehicles to sense energy constantly.

Figure 5 shows the performance of PRR for sensing mechanism and the proposed mechanisms. It is clear that the proposed mechanism with or without power control strategy's PRR is higher than that of the sensing mechanism. Among three mechanisms, the proposed mechanism with power control has the best performance. When the distance between vehicles is larger than 250 m, the performance of sensing mechanism decreases greatly compared to that of the proposed mechanisms. The reason is that the vehicles' moving characteristics are taken fully into consideration to determine the resource allocation proposed in this paper,

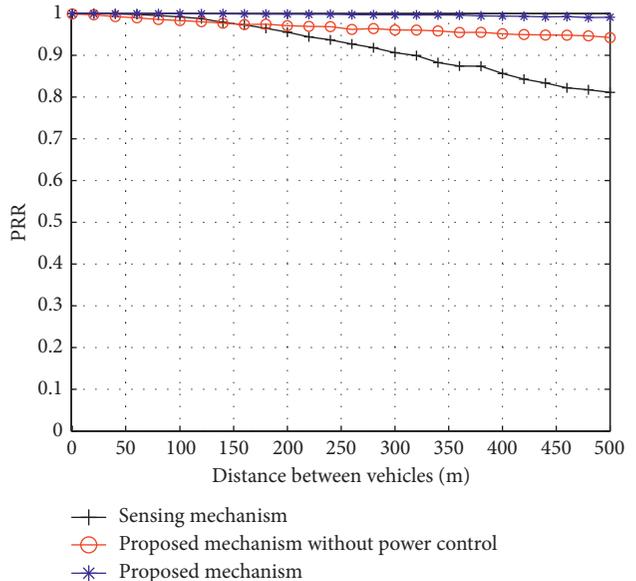


FIGURE 5: PRR among distance between vehicles.

which makes vehicles as far as possible use the same resource.

To further evaluate the resource utilization efficiency, the utilization ratio of resource is defined as the number of RBGs that have been allocated to vehicles divided by the total available number of RBGs. As shown in Figure 6, the RBG utilization ratio is constant in the proposed mechanism, while the utilization ratio changes frequently for the sensing mechanism. This means that resource reallocation in sensing mechanism happens more frequently than that in proposed mechanism. In freeway scenario, where vehicles do not easily change moving characteristics, the proposed mechanism capturing these steady characteristics has very stable performance.

To get better performance in V2V communication system, the amount of vehicles that are allocated to the same resource should be less and the distance between them should be far to eliminate interference. Figures 7 and 8 show the relationship between the number of vehicles sharing the same resource and their distance at different time for the sensing mechanism and proposed mechanism between vehicles sharing resource in proposed mechanism is relatively farther, while that in sensing mechanism is equally distributed along the whole length of the road. It is clear that the proposed mechanism taking distance characteristic into account in the resource allocation has better performance.

**4.2. Urban Scenario.** The joint resource allocation and power control mechanism for urban scenario is evaluated under the real map of Manhattan in Figure 9. Simulation scenario is further abstracted by SUMO [20] as shown in Figure 10. Moving features of vehicles are also obtained by SUMO according to the map, for example, the change of moving speed and direction at the cross. Major simulation parameters are shown in Table 2, including parameters of clustering as described in Section 3.

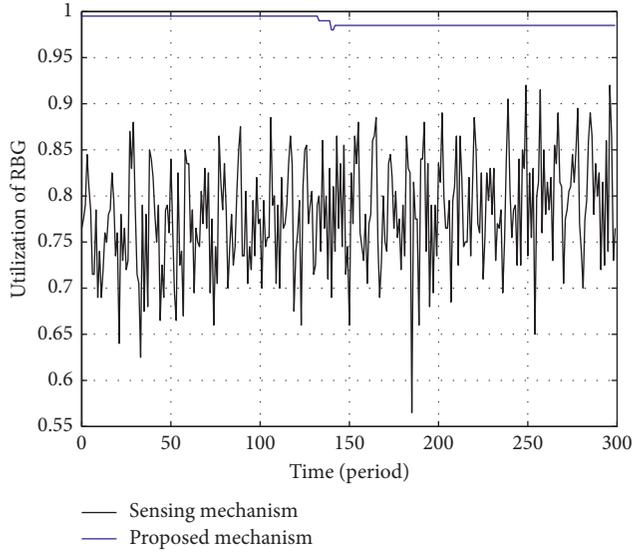


FIGURE 6: Utilization of RBG.

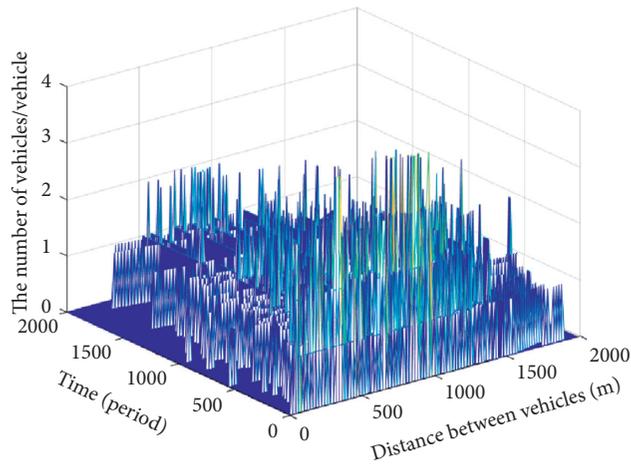


FIGURE 7: The number of vehicles allocated to the same resource in the sensing mechanism.

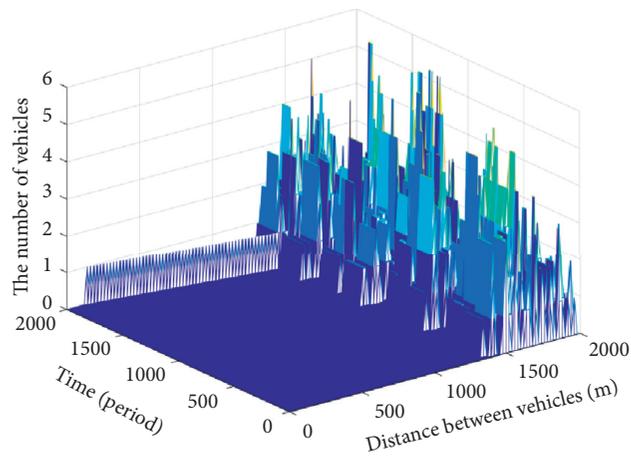


FIGURE 8: The number of vehicles allocated to the same resource in the proposed mechanism.



FIGURE 9: Real map of Manhattan for simulation.

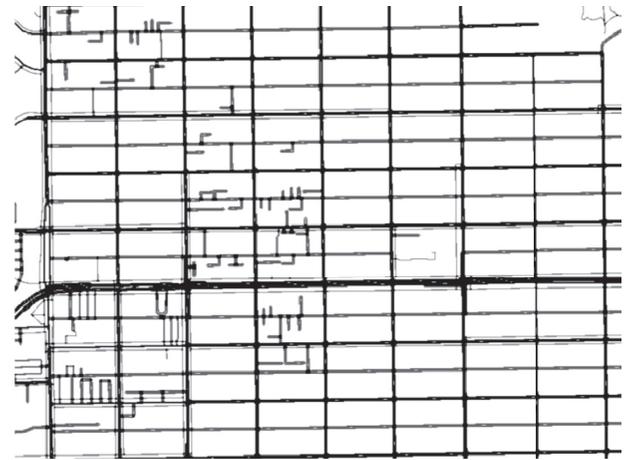


FIGURE 10: Simulation scenario abstracted by SUMO.

TABLE 2: Simulation parameters of urban scenario.

Parameters	Value
Carrier frequency	6 GHz
Bandwidth	10 MHz
Path loss model	Log-distance
Shadowing distribution	Log-normal
Decorrelation distance	25 m
Thermal noise density	-174 dBm/Hz
Transmission power	23 dBm
Number of vehicles	750
Number of slots in one period	100
Number of subchannels	3
Length of areas	1185 m
Width of areas	820 m
$\sigma_1, \sigma_2$	300, 600
$\sigma_3, \sigma_4$	1.2, 0.9

Figure 11 shows the PRR performance of the proposed mechanism without power control and with power control and sensing mechanism. Compared with the PRR performance in freeway scenario, PRR in urban scenario decreases more sharply when the distance between vehicles increases. Similar to the results under freeway scenario, the PRR of the proposed mechanism is higher than that of the sensing

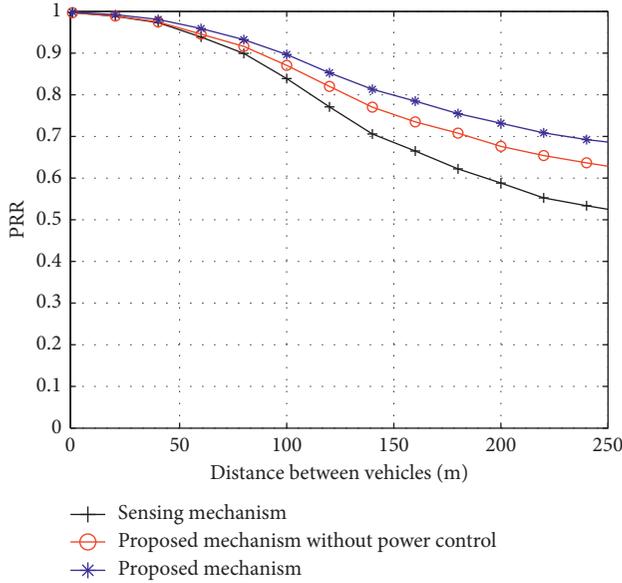


FIGURE 11: PRR among distance between vehicles.

mechanism. Among three mechanisms, the proposed mechanism with power control has the best performance.

Figures 12 and 13 show the relationship between the number of vehicles sharing the same resource and their distance at different time for the sensing mechanism and proposed mechanism, respectively, in urban scenario. It is clear that vehicles with larger distance have greater possibility to share the same resource. However, in urban scenario, there are more characteristics to consider in vehicles clustering, including distance. Therefore, the proposed mechanism has more possibility to group vehicles with similar moving characteristics, not just vehicles that are far away from each other, to share the same resource.

Figure 14 demonstrates the average number of vehicles sharing the same resource at different distance ranges. It can be observed that less vehicles share the same resource in the proposed mechanism compared to the sensing scheme at most distance values, which indicates that the proposed scheme can utilize resource more efficiently and has better PRR performance.

**4.3. Analysis of Computational Complexity.** In the sensing mechanism, vehicles need to collect Schedule Assignment (SA) messages, detect received energy on each RBG, and exclude RBGs based on SA messages. Each vehicle ranks RBGs according to their own average received energy and selects RBGs for itself. All of the above steps are done by vehicles themselves, which means that vehicles' computing capabilities influence delay greatly. With the existence of the ranking procedure, the computational complexity is between  $O(n \log n)$  and  $O(n^2)$  according to which kind of sorting algorithm is adopted, while  $n$  is the number of RBGs.

The complexity of the vehicle grouping resource allocation algorithm for freeway scenario proposed in this paper

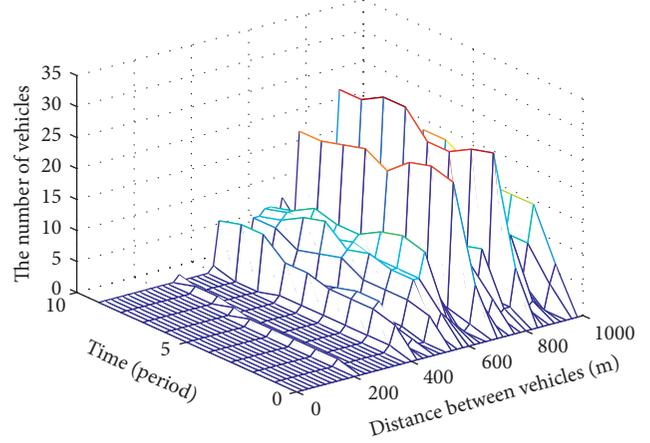


FIGURE 12: The number of vehicles allocated to the same resource in the sensing mechanism.

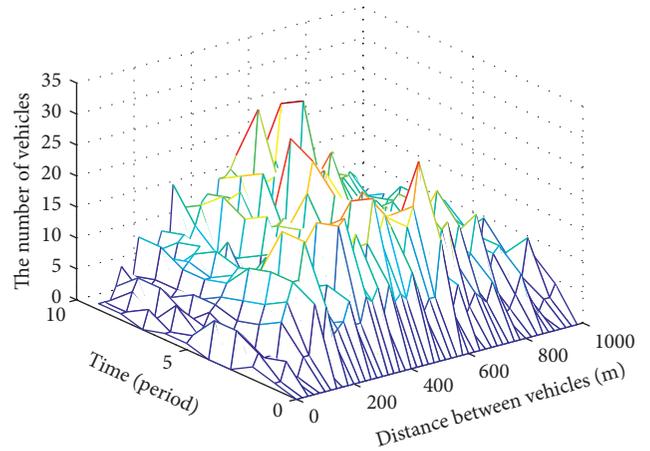


FIGURE 13: The number of vehicles allocated to the same resource in the proposed mechanism.

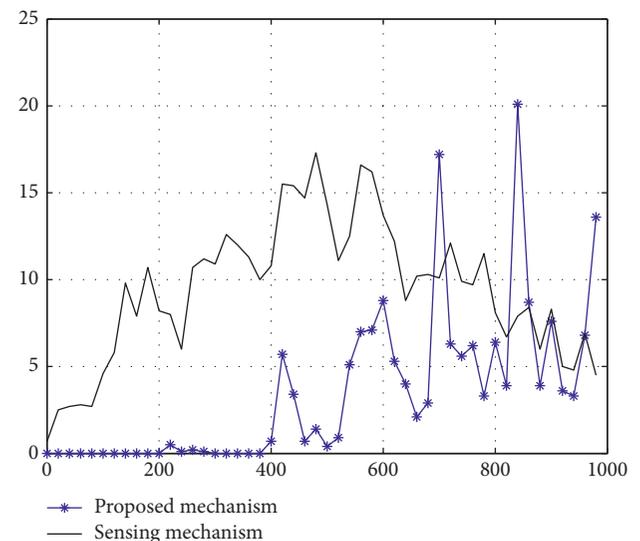


FIGURE 14: The average number of vehicles that are allocated to the same resource.

is  $O(N^*m)$ , while  $N$  is the number of vehicles and  $m$  is the number of vehicles in the group divided by their velocity. Because in a typical freeway scenario [19] the number of vehicles  $N$  is a little bigger than the number of RBGs, while the number of vehicles in the group divided by their velocity  $m$  is a bit smaller than the number of RBGs, the computational complexity of vehicle grouping algorithm is between the lower limit of the sensing mechanism and the upper limit of the sensing mechanism. Computational complexity of the vehicle clustering resource allocation algorithm for urban scenario utilizing SC is  $O(N^3)$  which is larger than the sensing mechanism. However, considering that computation process in the resource allocation mechanisms for freeway and urban scenarios proposed in the paper is done by the BS which collects vehicles' geographic related information and so forth from vehicles and that BS's computing capacity is far beyond that of the vehicle equipment, delay is less or at least comparable to the sensing mechanism with higher PRR.

After receiving resource allocation results from BS, vehicles adjust their power according to power control strategy utilizing Q-learning. Owing to learning process in reinforcement learning, the computational complexity is relatively high, which depends on the number of the steps in each episode. The advantage of this method is that the transmitting power can be adjusted along with the change of the environment. In the future, some effective methods can be adopted to reduce the number of iterations and reduce the computational complexity.

## 5. Conclusion

In this paper, NOMA is introduced into V2V communication system to enhance the utilization of limited frequency resource and a joint resource allocation and power control mechanism based on vehicles' moving characteristics is proposed. According to different moving conditions in freeway and urban scenarios, two resource assignment algorithms are designed, which divide vehicles into several groups according to their moving features. After that, power control strategy is obtained through Q-learning. System-level simulation results show that PRR of the proposed mechanism can be improved compared to that of the energy sensing mechanism.

## Data Availability

The simulation codes' data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the State Major Science and Technology Special Projects (Grant no. 2018ZX03001024).

## References

- [1] W. Sun, E. G. Ström, F. Brännström, K. C. Sou, and Y. Sui, "Radio resource management for D2D-based V2V communication," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6636–6650, 2016.
- [2] Q. Wei, W. Sun, B. Bai, L. Wang, E. G. Ström, and M. Song, "Resource allocation for V2X communications: a local search based 3D matching approach," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, May 2017.
- [3] 3GPP, "Final report of 3GPP TSG RAN WG1 94 v1.0.0," Technical Report R1-1810051, ETSI, Gothenburg, Sweden, 2018.
- [4] 3GPP, "3rd generation partnership project; technical specification group radio access network; study on LTE-based V2X services, Rel-14 V0.5.0," Technical Report TR 36.885, ETSI, Sophia Antipolis, France, 2016.
- [5] 3GPP, "Final report of 3GPP TSG RAN WG1 #83 v1.0.0," Technical Report R1-157773, ETSI, Anaheim, CA, USA, 2015.
- [6] X. Y. Gu and Y. Ding, "A dynamic geo-based resource selection algorithm for LTE-V2V," *Journal on Wireless Communications and Networking*, vol. 2018, p. 186, 2018.
- [7] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Proceedings of the 2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 770–774, Naha, Japan, November 2013.
- [8] B. Di, L. Song, Y. Li, and G. Y. Li, "Non-orthogonal multiple access for high-reliable and low-latency V2X communications in 5G systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2383–2397, 2017.
- [9] S. Pfletschinger and M. Navarro, "Non-orthogonal multiple access for vehicular communication," in *Proceedings of the 2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–5, Lisbon, Portugal, August 2018.
- [10] B. Wang, R. Shi, C. Ji, and J. Hu, "Joint precoding and user scheduling for full-duplex cooperative MIMO-NOMA V2X networks," in *Proceedings of the 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pp. 1–6, Honolulu, HI, USA, September 2019.
- [11] B. W. Khoueiry and M. R. Soleymani, "An efficient Noma V2X communication scheme in the Internet of vehicles," in *Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–7, Sydney, Australia, June 2017.
- [12] Z. Gao, B. Wen, L. Huang, C. Chen, and Z. Su, "Q-learning-based power control for LTE enterprise femtocell networks," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2699–2707, 2017.
- [13] S. Nie, Z. Fan, M. Zhao, X. Gu, and L. Zhang, "Q-learning based power control algorithm for D2D communication," in *Proceedings of the 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, Valencia, Spain, September 2016.
- [14] Z. Li, F. Nie, X. Chang, Y. Yang, C. Zhang, and N. Sebe, "Dynamic affinity graph construction for spectral clustering using multiple features," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 12, pp. 6323–6332, 2018.
- [15] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

- [16] 3GPP, “System-level evaluations on sidelink for NR V2X, 3GPP TSG RAN WG1 meeting #95,” Technical Report R1-1812210, HiSilicon, Spokane, WA, USA, 2018.
- [17] 3GPP, “Final report of 3GPP TSG RAN WG1 #94bis v1.0.0,” ETSI, Chengdu, China, R1-1812101, 2018.
- [18] B. J. A. Kröse, “Learning from delayed rewards,” *Robotics and Autonomous Systems*, vol. 15, no. 4, pp. 233–235, 1995.
- [19] 3GPP, “3rd generation partnership project; study on evaluation methodology of new vehicle-to-everything (V2X) use cases for LTE and NR,” Technical Report 37.885, 2017, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3209>, Anaheim, CA, USA.
- [20] SUMO, *Simulation of Urban Mobility*, SUMO, Berlin, Germany, 2019, <http://www.sumo.dlr.de/userdoc/SUMO.html,2019-08-27>.