Journal of Computers Vol. 34 No. 3, June 2023, pp. 31-47
doi: 10.53106/199115992023063403003

# Vehicle Clustering and Resource Allocation Algorithm Based on Cellular Network

Chen-Wei Feng, Xian-Guo Lu*, Yu Sun, Huang-Bin Zeng, Zhuo Li

School of Opto-electronic and Communication Engineering, Xiamen University of Technology, Xiamen, Fujian, China
chevyphone@163.com,{25905591,978995028,609615592,1358607215}@qq.com

*Received 6 June 2022; Revised 11 September 2022; Accepted 30 October 2022*

**Abstract.** As a special Mobile Ad-hoc Network (MANET), Vehicular Ad-hoc Network (VANET) plays a very important role in the future intelligent transportation system. In order to solve the problems of unstable communication connection, fast network topology change and low communication resource utilization caused by high vehicle mobility in VANET, a low-complexity resource allocation algorithm based on vehicle cluster is proposed. Firstly, considering the speed, position and moving direction of the vehicles, a vehicle clustering algorithm based on movement consistency is proposed to cluster the vehicles and keep the vehicle cluster stable. Secondly, a low-complexity resource allocation algorithm is proposed to improve the utilization rate of communication resources, which is constrained by the interference caused by the vehicle clusters to the cellular users. Simulation results show that the proposed algorithm has low complexity and can better maintain the stability of vehicle clusters and improve the system capacity in the common complex Internet of Vehicles (IoV) scenarios in cities.

**Keywords:** VANET, clustering algorithm, cellular communication, resource allocation

## 1 Introduction

As the future development direction of intelligent transportation system, Internet of Vehicles (IoV) is the core technology to realize intelligent connected vehicles and build intelligent transportation system. In recent years, vehicle Ad-hoc network (VANET) has been studied extensively. The most common communication modes in VANET include Vehicle to Vehicle (V2V), Vehicle to Pedestrian (V2P), Vehicle to Infrastructure (V2I) and Vehicle to Network (V2N) [1]. In VANET in urban environment, the network topology is complicated and changeable due to the high mobility of vehicles and the variable moving trajectory. This makes it impossible for vehicles in the network to establish long-term and stable communication with other vehicles, Road Side Units (RSU) or base stations [2]. In addition, the lack of spectrum resources gradually fails to meet the requirements of stable communication and high data transmission rate required by vehicle-mounted users [3]. Therefore, how to maintain the stable communication between vehicles and improve the system capacity is the focus of this paper.

The change of topology caused by the movement of vehicles will lead to additional communication overhead and unstable communication connections. The most common way to alleviate these problems is to cluster vehicles. The Cluster Head (CH) is used to communicate with cluster members, neighboring clusters, and infrastructure. However, in the complex urban environment, the high dynamics of vehicles makes them frequently leave clusters or join new clusters. This makes maintaining cluster stability and robustness extremely challenging. Literature [4] proposes the minimum ID algorithm to find the node with the minimum ID as the CH node, and then periodically obtain the IDs of surrounding nodes and update the CH. Therefore, the maintenance of the cluster is very simple, but the stability of the cluster is sacrificed to some extent. Literature [5] uses K-Medoid clustering model to cluster vehicle nodes. Firstly, energy-saving nodes are identified for forced communication, and the effective nodes are identified from each cluster by a meta-heuristic algorithm. The literature improves the energy efficiency of all vehicle nodes, but the cluster head is randomly selected and the cluster is very unstable. Literature [6] proposes a multi-hop clustering method MCA-V2I for IoV to improve network performance. When vehicles are connected to the Internet through a special infrastructure called RSU gateway, the necessary information about multi-hop neighbors can be obtained and shared, and the clustering is performed using Breadth-First Search (BFS) algorithm based on node movement rate. The clustering process uses few parameters, and the clustering results have certain local optimality. Literature [7] proposed a new clustering parameter - trust, and combined with the weighted formula composed of vehicle distance and speed to calculate the CH selection prob-

* Corresponding Author

ability, and further clustering. The algorithm improves the overall network performance, but it is only suitable for highway scenarios.

With the increasing number of vehicles in the IoV and the increasing rate requirements of users, how to use and allocate wireless spectrum resources more effectively has also been studied in depth. It can be found that the introduction of Device-to-Device (D2D) communication technology into cellular networks has the advantages of improving spectrum efficiency, reducing end-to-end delay and increasing channel gain. Therefore, the introduction of cellular network in V2V communication can also alleviate the shortage of spectrum resources in high-density vehicle environment, and ensure the requirements of low latency and high reliability of communication. But at the same time, the multiplexing of spectrum resources of cellular networks by V2V communication will cause additional interference to cellular users. It is of great significance to study how to reuse cellular resources as much as possible under the premise that cellular users are not seriously disturbed. Literature [8] proposed a greedy algorithm for reusing distance based on the location of vehicles. However, the greedy algorithm can only get the optimal solution under the current situation and cannot guarantee the global optimal solution. Literature [9] proposed a 3D matching algorithm for resource allocation. However, the algorithm does not consider the reliability of V2V, and only focuses on the throughput of the on-board system. In literature [10], in order to meet the constraints of maximum delay and dwell time of vehicles, the number of shared messages and restricted resources are balanced by fog-based vehicle network. The original problem is decoupled into two independent sub-problems and the solution is designed. However, the scheme does not consider the interference between vehicle pairs. Literature [11] proposed a joint solution of mode selection, base station selection, resource allocation and power allocation to solve the energy consumption problem. This scheme uses particle swarm optimization algorithm to maximize the system energy efficiency, but this scheme does not consider the interference between users, and the algorithm complexity is high. Literature [12] proposed a joint pairing and resource allocation algorithm based on continuous power allocation to solve the spectrum resource allocation problem and improve the network performance. However, this paper does not consider the instability caused by VANET network topology.

In view of the above problems, in order to solve the problems of unstable vehicle communication and low resource utilization in common urban intersection scenarios, this paper makes the following contributions:

(1) A vehicle clustering algorithm based on movement consistency is proposed. Firstly, three parameters of vehicle moving direction, vehicle speed and vehicle position are used to cluster the vehicles. Then the stable cluster head is selected by weighting the vehicle speed and vehicle spacing. Finally, the vehicle cluster is stabilized by the vehicle speed and the distance between the vehicle and the cluster head.

(2) In order to improve resource utilization, a resource allocation algorithm based on vehicle cluster is proposed for the optimization problem of vehicle cluster multiplexing cellular communication resources. Under the premise of satisfying the communication quality of cellular users first, the system is optimized with the constraint of the interference value of the vehicle cluster to the cellular users, and the total system capacity is effectively increased.

The rest of the paper is organized as follows: the second part introduces the communication model of the system, the third part discusses the vehicle clustering algorithm and resource allocation algorithm, the fourth part simulates and analyzes the algorithm, and finally concludes this paper.

## 2 System Model

### 2.1 Vehicle Clustering Communication Model

At present, most of the information flow of the IoV is based on the IEEE802.11p protocol. However, the transmission distance supported by this protocol is relatively short. On the one hand, with the acceleration of vehicle moving speed and the flexibility of driving routes, the communication time between vehicles is very short. On the other hand, the requirements of low delay, high reliability and high throughput for information transmission and allocation cannot be guaranteed in the case of dense vehicles. In order to effectively improve the stability of IoV information transmission, V2V communication is analyzed and studied. Aiming at the dynamic scene of vehicles at urban intersections, this paper proposes a clustering algorithm based on vehicle movement consistency in LTE cellular network architecture.

The LTE cellular IoV system model is shown in Fig. 1. Base stations, cellular users, cluster head vehicles and cluster member vehicles are included in the cellular IoV network. The base station clusters all the moving vehicles in the IoV. The relatively optimal vehicle is selected as the cluster head vehicle. The cluster head vehicle is the relay vehicle, and the cluster member vehicle is the ordinary vehicle. When the vehicle requests service information, it will be forwarded and reported to the base station through the relay vehicle. The base station will send the processed service information to the relay vehicle through the communication channel assigned to the vehicle cluster. The relay vehicle will forward the information to the ordinary vehicle again. Other vehicles only receive messages and do not forward the message.

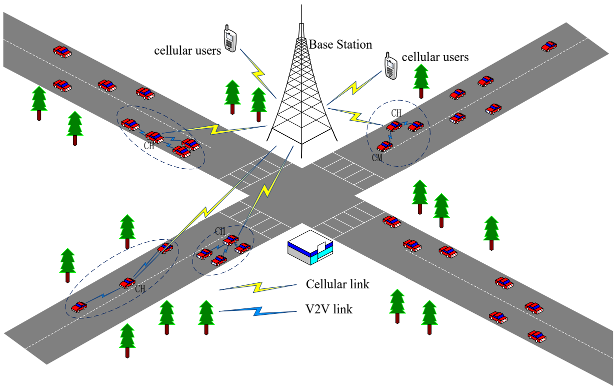


**Fig. 1.** LTE cellular car networking system model

It is supposed that there are $P$ cars in the current scene, which can be represented as $u_p$, $p = 1, 2, \ldots, P$. It is divided into $n$ clusters, which can be expressed as $C_n$, where $n \le P$. The vehicles in the cluster can be represented as $u_{n,l}$, $l = 1, 2, \ldots, L_n$. $L_n$ is the number of vehicles in each cluster. It is defined that the critical information *inf* related to the vehicle, such as vehicle speed $v$, acceleration $a$, vehicle position, and vehicle moving direction.

Considering the downlink of intra-cluster communication, the spectrum Resource Block (RB) is taken as the minimum resource unit to be allocated. Assuming that the number of vehicles in a cluster is l, and the resources of a cellular user have been reused in the cluster. As shown in Fig. 2, in the t1 period, the vehicle cluster multiplexes a RB of the cellular user. A RB can be divided into 7 eRBs on time slot t1. Each eRB can be used for data transmission between the cluster vehicles and the cluster head vehicles. The reuse of each RB by vehicles in the cluster is shown in Fig. 3.

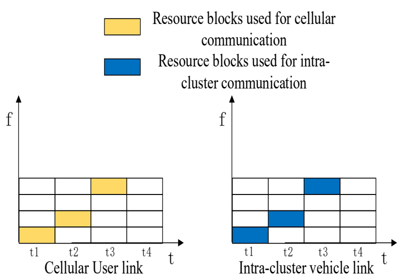


**Fig. 2.** Intra-cluster communication multiplexing cellular communication resources

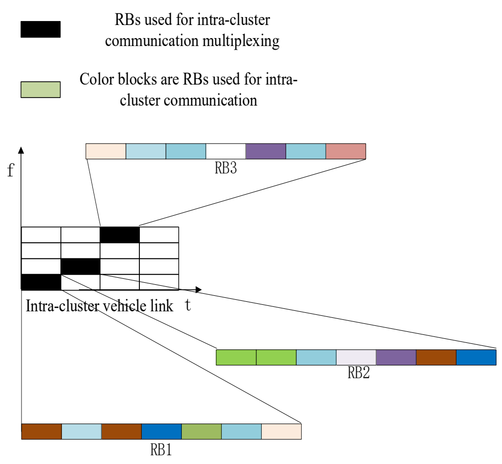


**Fig. 3.** Each RB reused by vehicles in the cluster

The vehicle communication model in the cluster is shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c) are respectively the system models when cluster vehicles CM1, CM2, CM3 communicate with cluster head vehicle CH. In a specific time slot, when only the CM1 in the cluster communicates with the cluster head vehicle, other vehicles CM2 and CM3 will temporarily retreat. It will not interfere with other users or base stations in the system. On the contrary, other users or base stations will not interfere with CM2 and CM3. At the same time, there will be a similar situation when other vehicles communicate with cluster head vehicles. However, at each eRB time slot, the vehicles will compete to choose which cluster vehicle to access to communicate with the cluster head vehicle.

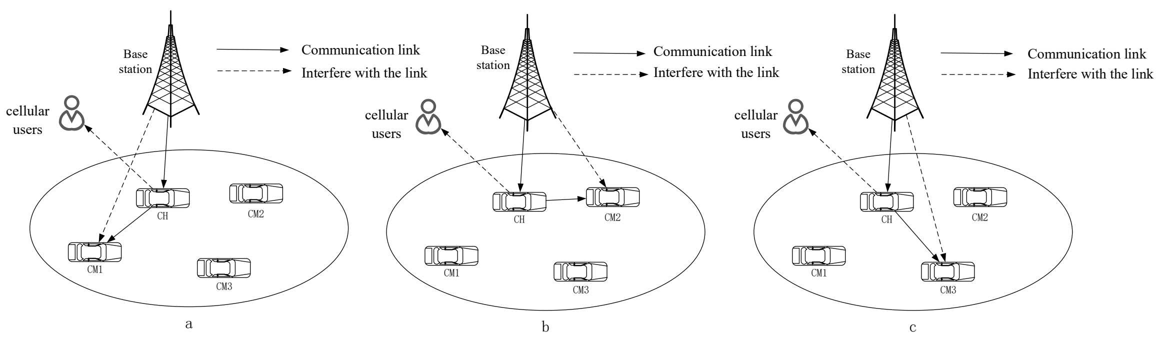


**Fig. 4.** Model of the vehicle communication system in the cluster

### 2.2 Cellular Communication Model for the IoV

According to the intra-cluster communication model in the previous section, this section takes the vehicle cluster (VCU) as the whole to form the V2V communication between CH and intra-cluster CM. This section mainly discusses the downlink scenario where VCU users multiplexing cellular user communication resources. In the system, vehicle CM in the cluster can communicate with CH, CH can communicate with the base station, and the base station can communicate with cellular users. However, because the spectrum resources of the cellular users are reused by the vehicle cluster, the base station will interfere with the CM in the cluster, and the CH will interfere with the cellular users. The system model is shown in Fig. 5.

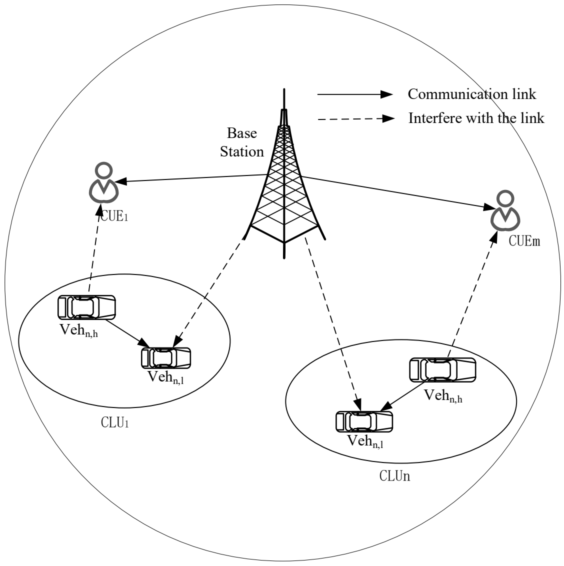


**Fig. 5.** System model

It is supposed that there are $M$ cellular users ($CUE_1$, $CUE_2$, …, $CUE_m$, …, $CUE_M$) are evenly distributed around the base station. $N$ clusters of vehicles ($CLU_1$, $CLU_2$, …, $CLU_n$, …, $CLU_N$) are evenly distributed on two roads covered by the base station. And each vehicle cluster contains a cluster head vehicle $Veh_{n,h}$ and $L$ other vehicles ($Veh_{n,1}$, $Veh_{n,2}$, …, $Veh_{n,l}$, …, $Veh_{n,L}$). The CH of each vehicle cluster is the transmitter, and the other CM is the receiver.

It is assumed that there are $K$ communicable channel resources for the base station to control and satisfy the relationship of $N < M < K < N + M$. That is, the number of vehicle clusters and cellular users to be multiplexed should be less than the number of channels provided by the system. At the same time, the number of channels provided by the system should be less than the sum of the two, and the number of vehicle clusters to be multiplexed should be less than the number of cellular users.

According to the system model, in the LTE cellular scenario, the base station assigns a subchannel to each cellular user. And the subchannels are orthogonal to each other, and the users do not interfere with each other. When the base station allocates resources to the vehicle cluster, due to the lack of idle channel, the base station will allocate two kinds of channel resources to the vehicle cluster: the idle channel resource and the shared channel resource. Idle channel resource refers to the channel resource that cellular users do not occupy. Shared channel resource refers to the channel resource that cellular users have occupied, that is, the vehicle cluster will be shared with the cellular users. When the base station transmits messages to cellular users, it will interfere with the CM user $Veh_{n,l}$ in the vehicle cluster that reuses the cellular user resources. At the same time, when the CH users in the cluster transmit messages to the CM users in the cluster, the communication interference will be caused to the cellular users in the multiplexed channel resources.

In the downlink, the signal-to-interference-noise ratio (SINR) of the CM (receiving end vehicle) $Veh_{n,l}$ in the vehicle cluster can be expressed as:

$$SINR_{n,l,k} = \frac{P_{n,l,k} G_{n,l}}{I_{BS,n,l} + N_0} . \quad \textbf{(1)}$$

In the formula, $P_{n,l,k}$ is the transmission power of the cluster head vehicle $Veh_{n,h}$ when the first vehicle user in the *nth* vehicle cluster occupies the *kth* channel resources. $G_{n,l}$ is the link gain between the cluster head and the vehicle $Veh_{n,l}$ in the *nth* vehicle cluster. $N_0$ is the received noise power of the $Veh_{n,l}$. $I_{BS,n,l}$ is the interference of the base station BS to the first vehicle user in the nth vehicle cluster. When the *nth* vehicle cluster multiplexes the link resources of the cellular user $CUE_m$:

$$I_{BS,n,l} = P_{BS,n,l} G_{BS,n,l} \quad . \tag{2}$$

Where, $P_{BS,n,l}$ represents the transmission power of the base station received by the CM user $Veh_{n,l}$ in the cluster. $G_{BS,n,l}$ is the link gain between the base station and the CM user $Veh_{n,l}$ in the cluster.

Because the relative distance between CM and CH in the cluster and the distance between CM and BS are different, the SINR generated by each CM communicating with CH is also different. From the discussion in the previous section, during intra-cluster communication, only one vehicle in each time slot communicates with the cluster head vehicle. If all CM with the lowest SINR in the cluster can satisfy the communication with the CH, then all the CM in the cluster can communicate. Therefore, the one with the lowest SINR of all the CM in the cluster can be regarded as a vehicle *Vehn*,*r* at the receiving end in the cluster. The link gain between the cluster head $Veh_{n,h}$ and $Veh_{n,r}$ in the vehicle cluster is $G_n$ . The interference of base station to the *nth* vehicle cluster $Veh_{n,r}$ is $I_{BS,n}$ . When the *nth* vehicle cluster is allocated to the idle channel resource, the base station transmission power received by the $Veh_{n,r}$ in the cluster is 0. The transmission power of base station received by $Veh_{n,r}$ in cluster is $P_{BS,n}$ . The link gain between the base station and the vehicle cluster $Veh_{n,r}$ is $G_{BS,n}$ .

From formula (1), the SINR of the vehicle $Veh_{n,r}$ at the receiving end in the cluster is:

$$SINR_{n,k} = \frac{P_{n,k} G_n}{I_{BS,n} + N_0} \quad . \tag{3}$$

Where:

$$I_{BS,n} = P_{BS,n} G_{BS,n} \quad . \tag{4}$$

To satisfy the condition that vehicles in the cluster can communicate, the SINR of the $Veh_{n,r}$ at the receiving end of the vehicle user in the cluster should meet the conditions as shown in formula (5):

$$SINR_{n,k} \geq Th_{clust} \quad . \tag{5}$$

Where, $Th_{clust}$ is the threshold of the SINR of the vehicle user $Veh_{n,r}$ in the cluster. Only when the SINR received by the CM users with the lowest SINR in the cluster is greater than the threshold, the communication quality of each CM in the cluster can be guaranteed.

Therefore, according to the formula (3) (5), the transmission power of the cluster head vehicle can be calculated as follows:

$$P_{n,k} \geq \frac{Th_{clust}(I_{BS,n} + N_0)}{G_n} \quad . \tag{6}$$

When the *nth* vehicle cluster multiplexes the channel resources of the cellular user $CUE_m$, the interference of the vehicle cluster to the cellular user is as follows:

$$I_{n,m} = P_{n,m} G_{n,m} \quad . \tag{7}$$

Where, $I_{n,m}$ is the interference caused by the transmitting end user $Veh_{n,h}$ of the *nth* vehicle cluster to the cellular user $CUE_m$. $P_{n,m}$ is the interference power of the transmitting end user $Veh_{n,h}$ of the vehicle cluster received by the cellular user $CUE_m$. $G_{n,m}$ is the link gain between the transmitting end user $Veh_{n,h}$ of the vehicle cluster and the cellular user $CUE_m$.

The total interference to all cellular users in the system is:

$$I_{sum} = \sum_{n=1}^{N}\sum_{m=1}^{M}\beta_{n,m} I_{n,m} \ . \quad \textbf{(8)}$$

Where $\beta_{n,m}$ is the reuse factor of shared cellular user resources in vehicle clusters, with a value of 0 or 1.

According to Shannon formula, the formula for calculating the total capacity of the system can be expressed as follows:

$$C_{sum} = B\sum_{m=1}^{M}\sum_{n=1}^{N}\log_2(1+SINR_m) + B\sum_{k=1}^{K}\sum_{n=1}^{N}\alpha_{n,k}\log_2(1+SINR_{n,k}) \ . \quad \textbf{(9)}$$

Formula (9) represents the capacity of channel resources shared by cellular users and vehicle clusters, and the total capacity of dedicated channel resources occupied by cellular users and vehicle clusters respectively. Where, $B$ is the channel bandwidth, and $\alpha_{n,k}$ is the dedicated channel resource occupation factor of the vehicle cluster, with a value of 0 or 1.

## 3 Vehicle Clustering and Resource Allocation Algorithm

### 3.1 Vehicle Clustering Algorithm

This section proposes a Vehicle Movement Consistency Clustering Algorithm (VMCCA) based on the LTE Cellular IoV system scenarios. The algorithm can effectively reorganize fewer clusters and increase the stability of clusters. Through the analysis and discussion of the paper [13], in the process of cluster head election, the speed, direction, and distance between vehicles are prominent factors in optimizing clustering. Since the moving direction of the vehicle has been considered in the initial clustering process, the driving direction in each cluster has been consistent. Therefore, the comprehensive value *W* of vehicle speed and distance between vehicles is used to determine the cluster head of each cluster. The W value is represented by the following formula (10):

$$W_{n,l} = \omega_1 V_{n,l} + \omega_2 d'_{n,l} \ . \quad \textbf{(10)}$$

Where, $W_{n,l}$ is the *W* value of the *lth* vehicle in the *nth* cluster. $V_{n,l}$ is the difference between the speed of the vehicle $u_{n,l}$ and the average speed in the cluster, as shown in formula (11). $d'_{n,l}$ is the average distance between the vehicle $u_{n,l}$ and other vehicles in the cluster, as shown in formula (12). $\omega 1$ and $\omega 2$ are weighting coefficients and satisfy $\omega 1 + \omega 2 = 1$.

$$V_{n,l} = | v_{n,l} - \bar{v}_n | \ . \quad \textbf{(11)}$$

$$d'_{n,l} = \frac{1}{l-1}\sum_{i=1}^{l-1} | d_{n,l} - d_{n,i} | \ . \quad \textbf{(12)}$$

In formula (11), $v_{n,l}$ is the driving speed corresponding to vehicle $u_{n,l}$.

Then through the *W* value between vehicles in each cluster, the vehicle with the lowest W value is selected as CH. Because the smaller *W* value is, the closer the vehicle node is to the average speed in the cluster. The closer it is to other vehicle nodes in the cluster, the more stable it will be in the cluster. The vehicle members in the cluster periodically report their *inf* to the CH. As a relay vehicle, CH periodically broadcasts the information in the cluster to the base station and forwards the information of the base station to the vehicle members in the cluster. This achieves the function of maintaining the stability of the cluster.

The pseudo-code of the VMCCA algorithm is shown in Table 1, and the specific steps of the algorithm are as follows:

**Step 1.** Input the number of vehicles $P$ in the system, the number of initial clusters $n$, vehicle position information $D$, vehicle initial speed $V$, and initial vehicle acceleration $a$.

**Step 2.** With the support of LTE communication, all vehicles are clustered according to the vehicle location information $D$ and the number of initial clusters $n$. All vehicles are clustered according to the location distribution of vehicles to form initialization clusters. The initial clustering process is as follows:

(1). Input the vehicle data set $\boldsymbol{P} = \{u_1, u_2, \ldots, u_p\}$, the initial number of clusters. Calculate the proportion of vehicles in the north-south and east-west directions through the vehicle distribution. Allocate the initial number of clusters $n_{SN}$ and $n_{EW}$ according to the relevant proportion, where $n_{SN} + n_{EW} = n$.

(2). From the vehicle data set $P$, according to the number of vehicles in both directions, $n_{SN}$ and $n_{EW}$ vehicles are randomly selected as the initial $n$ pseudo cluster headsets $\boldsymbol{PCH} = \{pch_1, pch_2, \ldots, pch_n\}$.

(3). Initialize the cluster with pseudo-cluster head PCH. Establish the set $\boldsymbol{C} = \{C_1, C_2, \ldots, C_n\}$, where $pch_n$ is the pseudo cluster head of cluster $C_n$, respectively.

(4). For iTunes $i = 1, 2,\ldots, P$, the distance $d_{ij}$ between all vehicles $u_i$ in the north-south and east-west directions and the pseudo cluster head $pch_j$ ($j = 1, 2,\ldots, n$) in each direction is calculated by traversing, as shown in formula (13):

$$d_{ij} = \left| d_{u_i} - d_{pch_j} \right|^2 . \tag{13}$$

When the distance between $u_i$ and a pseudo-cluster head $pch_j$ is the minimum, the vehicle $u_i$ is classified into cluster $C_j$, $j = 1, 2, \ldots, n$.

(5). Output initial cluster $\boldsymbol{C} = \{C_1, C_2, \ldots, C_n\}$.

**Step 3.** For each cluster $C_n$, the average speed $\overline{v}_n$ and the number of vehicles per cluster $L_n$ are calculated. The difference factor $\Delta v$ between the vehicle speed in the cluster and the average speed in the cluster is calculated according to the formula (14). If $\Delta v \geq 0.5$, the vehicles in the cluster are deleted so that the speed and the moving direction of the vehicles in the remaining clusters are highly consistent.

$$\Delta v = \left| \frac{v_{n,l} - \overline{v}_n}{\overline{v}_n} \right|. \tag{14}$$

**Step 4.** According to the formula (10) (11) (12), the cluster head $u_{n,h}$ of each cluster is selected and marked as the set CH. The cluster head periodically broadcasts cluster information to nearby vehicles.

**Step 5.** The vehicles out of the cluster and the vehicles in the single cluster are formed into an unclustered vehicle set $Q$. According to the formula (15), the clustering factor $\eta_{n,q}$ between each unclustered vehicle $u_q$ and each cluster head $u_{n,h}$ is traversed.

$$\eta_{n,q} = \alpha \left| \frac{v_q - \overline{v}_n}{\overline{v}_n} \right| + \beta \frac{\left| d_q - d_n \right|}{R}. \tag{15}$$

Where, $\overline{v}_n$ is the average speed of the *nth* cluster. *vq* is the speed of the unclustered vehicle $u_q$. *dn* is the position of the *nth* cluster head. *dq* is the position of the unclustered vehicle. $R$ is the coverage radius of the cellular network. $\alpha$ and $\beta$ are the weighting coefficients $\omega 1$ and $\omega 2$ selected by CH in the cluster, respectively. The smaller $\eta_{n,q}$ is, the greater is the probability that the cluster head receives vehicle $u_q$ into the cluster. Because the smaller the relative difference between *vq* and $\overline{v}_n$, the smaller the difference between *dq* and *dn*, and the longer the $u_q$ stays in the cluster. The cluster will be more stable.

**Step 6.** If the entry factor $\eta_{n,q} \leq 0.5$, it indicates that the vehicle can be clustered. The vehicle is deleted from vehicle set $Q$, and each vehicle cluster set $C$ and vehicle set $Q$ is updated. Otherwise, the vehicle is clustered independently to form a single cluster. The vehicle is deleted from vehicle set $Q$ to update the vehicle cluster set $C$ and the vehicle set $Q$. Loop Step 6 until the vehicle set $Q$ is empty.

**Step 7.** According to the vehicle acceleration $a_p$, the vehicle speed $v_p$, and each iteration time, the speed and position of each vehicle are calculated. The data such as the number of vehicle clusters *n*, the number of clustered vehicles, and the cluster head *CH* are output.

**Step 8.** Determine whether the iteration termination condition is reached. If not, cycle Step 3, 4, 5, 6, 7, and continue iterative calculation for the IoV scenario.

**Table 1.** Pseudo code of VMCCA algorithm

| |
|---|
| Input: vehicle set *P* related information *inf*, initial vehicle clustering number *n* |
| Output: number of vehicle clusters *n*, number of vehicles out of the cluster, and cluster head ***CH*** |
| 1)The initial clustering numbers $n_{SN}$ and $n_{EW}$ on the two roads are calculated according to the vehicle position information. |
| 2)$n_{SN}$ and $n_{EW}$ vehicles are randomly selected as pseudo-cluster head $pch_j$ on two roads. |
| 3)Using the formula (13) to calculate the distance $d_{ij}$ between each vehicle and $pch_j$ |
| 4)for $i$ =1 to $P$ do |
| 5) Search for the minimum distance from the *ith* car min_dist |
| 6) for j to n do |
| 7) if min_dist==$d_{ij}$ |
| 8) The vehicle $u_i$ is grouped into cluster $C_j$ |
| 9) end if |
| 10) end for |
| 11)end for |
| 12)The average speed factor $\Delta v$ of vehicles in each cluster is calculated by the formula (14). |
| 13)if $\Delta v \geq 0.5$ |
| 14) Vehicles leave the cluster, update the clustered vehicle set ***Q*** |
| 15)end if |
| 16)According to the formula (10) (11) (12), the cluster head $u_{n,h}$ of each cluster is selected and marked as set ***CH***. |
| 17)while t less than the number of iterations |
| 18) t = t + 1 |
| 19) The clustering factor $\eta_{n,q}$ of the vehicle and each cluster ***CH*** node in the vehicle set ***Q*** is calculated by traversing formula (15). |
| 20) for $i$ = 1 to $q$ do |
| 21) if $\eta_{n,q} < 0.5$ |
| 22) Vehicle $u_q$ clustering |
| 23) else |
| 24) Vehicle $u_q$ are clustered independently to form single clusters. |
| 25) end if |
| 26) Delete the vehicle from ***Q*** and update ***Q*** and ***C*** |
| 27) end for |
| 28) Output the number of vehicle clusters *n*, number of vehicles out of the cluster, and cluster head ***CH*** |
| 29) All vehicle nodes move according to acceleration $a_p$, vehicle speed $v_p$ and intersection direction. |
| 30) Put forward the steering and U-turn vehicle node information, and update ***Q*** and ***C*** |
| 31) end while |

The algorithm considers a dynamic scene of a vehicle node at an urban intersection. Each vehicle node will move flexibly in the iteration time. There is no complex traversal in the first clustering to reselect the CH. In the maintenance process of the cluster, the speed factor and the clustering factor are calculated only once for each in and out of the cluster. Therefore, it can be known that the algorithm is relatively simple and the complexity is $O(n*q)$.

### 3.2 Vehicle Cluster Resource Allocation Algorithm

Aiming at the cluster multiplexing LTE cellular resource system model, a spectrum resource allocation algorithm is proposed in this section. On the premise of ensuring the communication of cellular users, the reasonable allocation of cellular channel resources can reduce the interference of vehicle clusters to cellular users and further improve the overall capacity of the system. In the process of allocation, to ensure the communication quality between the cellular users and the vehicle users in the cluster, the maximum interference threshold of the communication in the vehicle cluster is $I_{th}^{CLU}$, the maximum interference threshold of the cellular user is $I_{th}^{CUE}$, and the communication RB of all channels are the same. The following are the specific steps of the algorithm:

**Step 1.** Input the number of cellular users *M*, the number of vehicle clusters *N*, the number of communication RBs *K*, the maximum interference threshold of vehicle clusters $I_{th}^{CLU}$, the maximum interference threshold of cellular users $I_{th}^{CUE}$, and the number of other vehicles in each vehicle cluster *L*.

**Step 2.** By traversing the number of other vehicles in each cluster *L*, if L = 0, the vehicle cluster can reuse any cellular user resources without any interference, and the vehicle cluster is set to *NaN*. The users of the vehicle cluster are deleted, and the number of vehicle clusters is recorded as . $N^*$ is the number of vehicle clusters that can be allocated, namely $N = \bar{N} + N^*$.

**Step 3.** By selecting the base station interference threshold $I_{BS}$ as a variable. The minimum transmission power of CH users in vehicle clusters satisfying intra-cluster communication conditions is calculated by the formula (6). Then the interference caused by all $N^*$ vehicle clusters multiplexing all M cellular user channel resources can be calculated by the formula (7), denoted as $N^* \times M$ two-dimensional interference matrix $I_{N^* \times M}$.

**Step 4.** Cycle through the matrix $I_{N^* \times M}$. If the element $I_{n \times m} > I_{th}^{CUE}$ in the matrix, means that the communication of the cellular users cannot be guaranteed when the vehicle cluster multiplies the cellular users. Therefore, the element is set to *NaN*. When a row of elements in $I_{N^* \times M}$ is all *NaN*, it means that the vehicle cluster cannot meet the communication requirements of multiplexing any one cellular user and deleting the element of that line. When there is no matrix with a *NaN* row, the matrix is marked as $I_{N' \times M}$, where *N*' represents the number of vehicle clusters that can participate in multiplexing cellular user resources. The interference matrix between the vehicle cluster with reusable cellular user resources and all cellular CUE users is $I_{N' \times M}$, and the candidate vehicle cluster user set *CLU* is initialized.

**Step 5.** The *K-M* vehicle cluster is selected from the non-reusable vehicle clusters of $N^*$-*N*' to occupy the idle communication channel. The number of occupancies is $P = \begin{cases} N^* - N' & (N^* - N') < (K - M) \\ K - M & (N^* - N') \geq (K - M) \end{cases}$.

**Step 6.** Cycle through the interference matrix $I_{N' \times M}$ to find the minimum interference value of the matrix $I_{n,m}^{\min}$.

**Step 7.** By calculating the interference value of the corresponding base station to the vehicle cluster $I_{BS,n}$. If the interference value is satisfied $I_{BS,n} < I_{th}^{CLU}$, the cellular user $CUE_m$ and the vehicle cluster $CLU_n$ can share the RB m. At the same time, all the elements of the *nth* row and *mth* column in the matrix $I_{N' \times M}$ are set as *NaN*. In addition, the vehicle cluster $CLU_n$ is deleted from the vehicle cluster user set *CLU*, indicating that the *nth* vehicle cluster reuses the communication channel resources of the *mth* cellular user and no longer participates in the allocation. Otherwise, only $I_m$, is set to *NaN*, which means that the *nth* vehicle cluster user cannot reuse the communication resources of the *mth* cellular subscriber. Finally, the matrix $I_{N' \times M}$ and the vehicle cluster user set *CLU* are updated.

**Step 8.** If there is an element $I_{m,n}$ that in the matrix $I_{N' \times M}$ is not NaN, and the number of vehicle cluster users to be allocated in the vehicle cluster user set *CLU* is greater than the number of dedicated idle channels (K-M-P), the cycle executes Step 6, 7. Otherwise, idle dedicated channel resources are allocated to the vehicle cluster.

**Step 9.** Calculates the number of vehicle cluster users connected to the communication system and the system's total capacity.

After implementing the algorithm, priority will be given to extracting the single clusters that will not cause interference to the cellular users. Priority will be given to allocating dedicated channel resources or not allowing access to the system to the vehicle clusters that cause more significant interference to the cellular users. In comparison, the vehicle clusters with less interference to the cellular users are allocated the cellular user channel resources or dedicated channel resources. All the resources will be allocated to satisfy the cellular user communication. The proposed algorithm does not have too much computation, the algorithm's complexity is *O*(*N*'**M*), and the algorithm complexity is low. The algorithm pseudo-code is shown in Table 2.

**Table 2.** Pseudo-code of the resource allocation algorithm

| Input: $K$, $M$, $N$, $I_{th}^{CLU}$ , $I_{th}^{CUE}$ , $L$<br>Output: num_accc, apa_prop |
|---|
| 1) $\overline{N}$ =length(find(L==0)) |
| 2) Delete the vehicle cluster |
| 3) $N^* = N - \overline{N}$ |
| 4) Calculate the minimum transmit power $P_{n,k}$ of vehicle cluster head users by formula (6) |
| 5) The interference matrix $I_{N^*\times M}$ of all cellular user resources by all clusters multiplexing is calculated by the formula (7). |
| 6) for $n$=1 to $N$* do |
| 7) for $m$=1 to $M$ do |
| 8) if $I_{n\times m} > I_{th}^{CUE}$ |
| 9) $I_{n\times m} = NaN$ |
| 10) end if |
| 11) end for |
| 12) end for |
| 13) The matrix $I_{N'\times M}$ is obtained by deleting the nth row elements of all $NaN$. |
| 14) Assign idle communication channels to N*-N' vehicle cluster that cannot be multiplexed, and find out the $P$ |
| 15) Initialize the interference matrix as $I_{N'\times M}$, and the vehicle cluster user set ***CLU*** |
| 16) if $N' > K\text{-}M\text{-}P$ |
| 17) for $I$ = 1 to $N'\text{-}K\text{+}M$ do |
| 18) $I_{n,m}^{\min}$ = min(min($I_{N'\times M}$)) |
| 19) [n,m]=find($I_{N'\times M} == I_{n,m}^{\min}$ ) |
| 20) Calculate the corresponding base station interference value $I_{BS,n}$ |
| 21) if $I_{BS,n} < I_{th}^{CLU}$ |
| 22) Delete row n and column m in $I_{N'\times M}$ |
| 23) Update matrix $I_{N'\times M}$ and vehicle cluster user set ***CLU*** |
| 24) Calculate the system capacity capa_prop through formula (9) |
| 25) end if |
| 26) Calculate the number of reused vehicles num_acc by ***CLU*** |
| 27) end for |
| 28) else |
| 29) num_acc= $N'$ |
| 30) end if |
| 31) num_acc= num_acc+ $\overline{N}$ |

## 4 Simulation and Analysis

### 4.1 Selection of Simulation Parameters

Based on the LTE cellular communication model of the cell, it is assumed that the radius of the cell is 500m and the base station is in the center of the cell. There are two roads intersecting near the base station in the community, and the CLUs vehicle cluster is evenly distributed on the two roads. Both roads are 1000m long, and the cellular user CUEs are evenly distributed in the cell. The vehicle's maximum speed is 60km/h, and the acceleration is generated by the normal distribution with the mathematical expectation of 0 and a standard deviation of 1. By setting the number of vehicles in the scene as 50, different weighting coefficients $\omega 1$ and $\omega 2$ are simulated and compared. The number of simulations is 200 times, as shown in Fig. 6.

As can be seen from the figure above that when $\omega1$ is 0.9 and $\omega2$ is 0.1, the algorithm's stability is higher. Therefore, the selection of vehicle clustering simulation parameters is shown in Table 3. The uniform clustering algorithm based on mobility, the initial clustering algorithm based on K-means, and the clustering algorithm based on location distribution proposed in this paper are simulated.

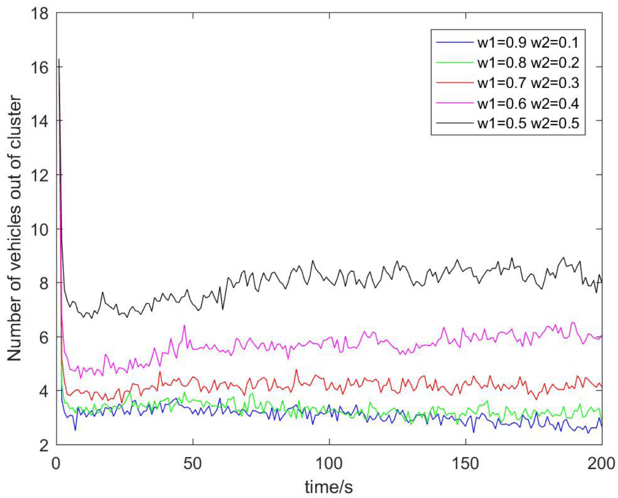


**Fig. 6.** Variation of the number of vehicles leaving the cluster over time

**Table 3.** System simulation parameters

| Parameters | Numerical value |
|---|---|
| Cell radius/m | 500 |
| Road length/m | 1000 |
| Maximum vehicle speed/km/h | 60 |
| Total number of vehicles | 300 |
| Initial clustering number | 20 |
| Broadcast interval /s | 1 |
| $\omega_1$ | 0.9 |
| $\omega_2$ | 0.1 |
| Number of simulations | 50 |

**Table 4.** System simulation parameters

| Parameters | Numerical value |
|---|---|
| Cell radius/m | 500 |
| Number of cellular users | 30 |
| Number of vehicle clusters | 10 |
| Number of link channels | 33 |
| Total system bandwidth/MHz | 10 |
| Path lose model/dB | 128.1+37.6lg(d) |
| Noise power spectral density/(dBm/Hz) | -174 |
| SINR threshold of vehicle cluster users /dB | 1.8 |
| SINR threshold for cellular users /dB | 2.6 |
| Maximum distance between vehicle cluster users /m | 50 |
| Number of simulations | 5000 |

It is assumed that every RB in the system is the same, and each RB can only be occupied or shared by at most one cellular user or at most one vehicle cluster. The selection of stimulation parameters of the resource allocation algorithm is shown in Table 4. The algorithm proposed in this paper is simulated and compared with the heuristic algorithm, Hungarian algorithm, and random algorithm.

### 4.2 Analysis of Simulation Results

**Vehicle Clustering Algorithm.** Cluster stability refers to the ability of vehicles and cluster heads to establish long-term and stable communication. In this paper, the number of vehicles leaving clusters, the number of clusters and the number of continuous communication vehicles are used as performance indicators to analyze. Fig. 7 shows the relationship diagram of the number of vehicles leaving the cluster with time. As can be seen from the figure, the number of vehicles leaving clusters of the proposed algorithm is significantly reduced compared with the other two algorithms. This is because the proposed algorithm maintains the vehicle cluster through the vehicle speed and position, and has better stability than the other two algorithms.

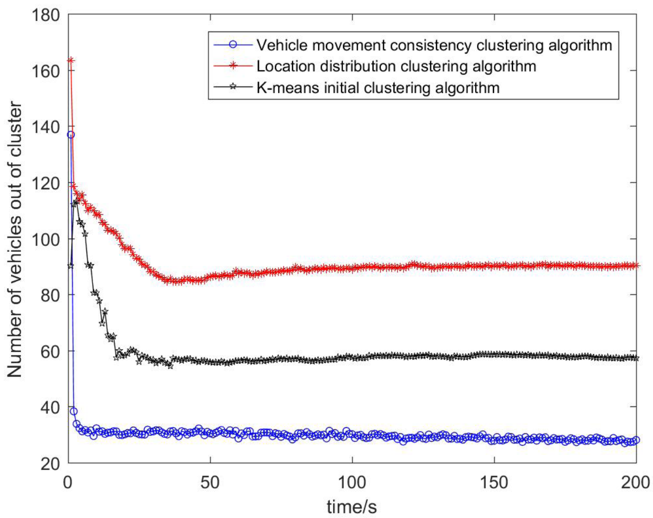


**Fig. 7.** The relationship between the number of vehicles leaving the cluster and the time

Fig. 8 shows the relationship between the number of vehicle clusters and time. As can be seen from the figure, the number of vehicle clusters proposed in this paper is also much smaller than that of the other two algorithms. That is to say, the algorithm proposed in this paper has more vehicles clustered on average than the other two algorithms, which is more convenient for vehicle management in the vehicle network. On the other hand, the number of vehicle clusters in the proposed algorithm becomes more and more stable with the change of time, which can further reflect the stability and robustness of the proposed algorithm.

Continuous communication means that the vehicle communicates with the same target in both the previous time slot and the next time slot. That is, there is no need to switch network connections between vehicles. Fig. 9 shows the relationship between the number of continuous communication vehicles and time. As shown in the figure, the proposed algorithm is also superior to the other two algorithms in the number of continuous communications between vehicles. Because the cluster of vehicles is stable, the continuous communication connection of vehicles will also be stable. Through calculation, it can be concluded that the stability rate of the proposed algorithm is more than 80%.

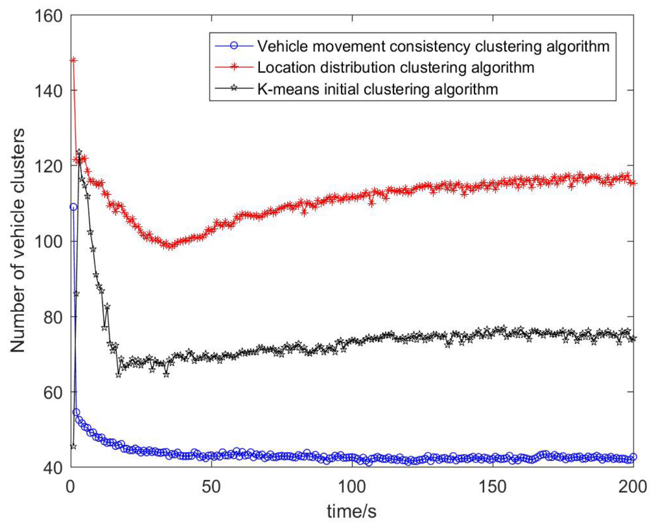


**Fig. 8.** The relationship between the number of vehicle clusters and time

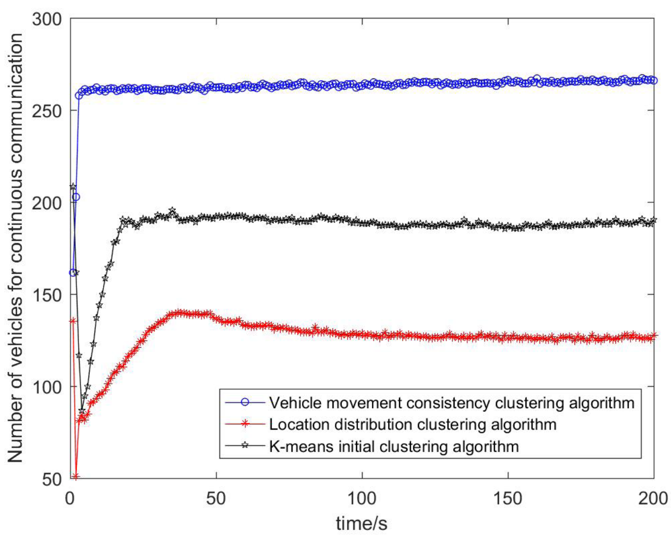


**Fig. 9.** Relationship between the number of continuous communication vehicles and time

From the above simulation analysis and algorithm implementation, the proposed algorithm is simple and stable. A large number of incoming and outgoing clusters and cluster recombination are reduced, the stability of clusters is better maintained, and the performance is further improved.

**Resource Allocation Algorithm.** In order to verify the effectiveness of the proposed algorithm, the sender and receiver of cellular users and vehicle clusters are randomly distributed in the cell. The distribution map of users in a single cell scene is simulated. The base station is located in the center of the cell. The random distribution scene of cellular users, vehicle clusters, and base stations in the cell is shown in Fig. 10.

Fig. 11 shows the relationship between the base station interference threshold and the number of vehicle clusters allowed to communicate by the system under different algorithms. As can be seen from the figure, as the interference threshold of the base station increases gradually, the number of vehicle clusters allowed to communicate also increases gradually. Moreover, the algorithm proposed in this paper increases the allowed access number of vehicle clusters faster than the other three algorithms. When the base station interference threshold is -110dbm, the proposed algorithm first reaches the upper limit of the number of vehicle cluster communication, which is lower than the other three algorithms. On the basis of satisfying the communication performance of cellular users, the proposed algorithm preferentially assigns the vehicle clusters with large interference to the dedicated channel. Therefore, when the cellular user reuse resources are allocated, the vehicle clusters with less interference can meet the system requirements faster and connect to the system faster.

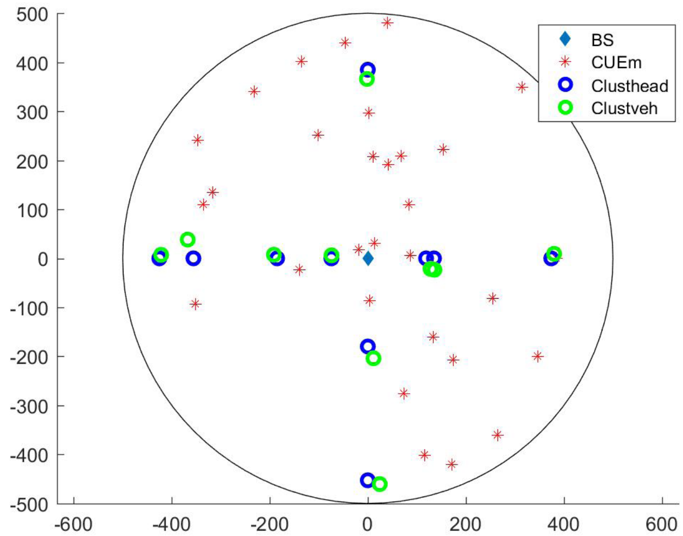


**Fig. 10.** Scene distribution map of the cell

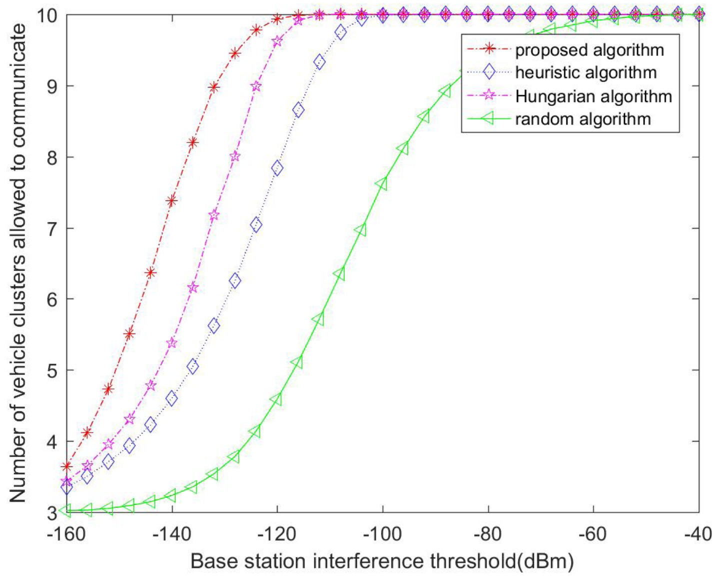


**Fig. 11.** The number of vehicle clusters allowed to communicate under different algorithms

Fig. 12 shows the relationship between the base station interference threshold and the total system capacity under different algorithms. On the one hand, as the interference threshold of the base station increases gradually, the total capacity of the system decreases gradually. As the interference value of the base station increases, the number of vehicle cluster users connected to the system increases, and the total capacity of the system becomes smaller and stable. On the other hand, the total capacity of the system in the proposed algorithm decreases slowly compared with the other three algorithms. This is because the algorithm in this section gives priority to the vehicle cluster users with less interference to reuse the cellular user resources. For the vehicle cluster users with large interference, the dedicated channel resources should be allocated as much as possible. Therefore, the use of system capacity will be better than the other three algorithms.

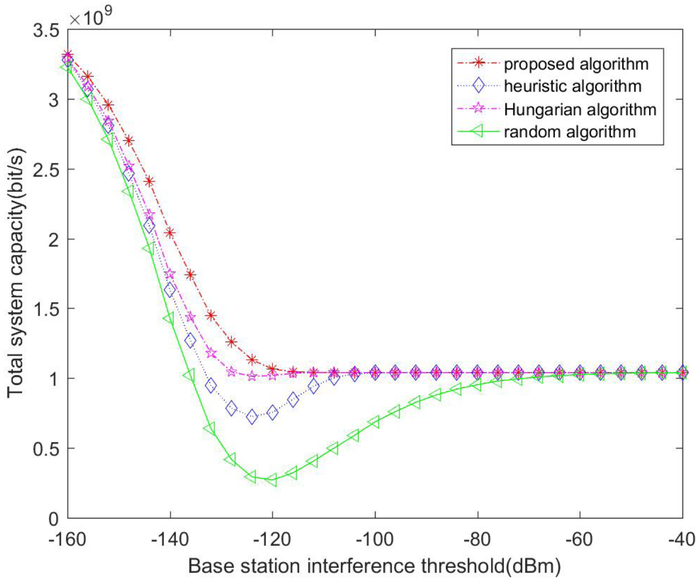


**Fig. 12.** Total system capacity under different algorithms

From the above simulation analysis and algorithm implementation, the proposed algorithm can reasonably analyze and consider the vehicle clusters with different interference. The communication resources of cellular users are given priority, and then the interference of vehicle clusters is allocated, so the utilization of communication resources is improved.

## 5 Conclusion

In order to solve the problems of unstable communication connections and low utilization of communication resources caused by high mobility of vehicles in common urban intersection scenarios, this paper proposes a low-complexity resource allocation algorithm based on vehicle clusters. The vehicles are clustered by three parameters: moving direction, speed and position. Then the stable cluster head is selected by weighting the vehicle speed and distance between vehicles. Finally, the vehicle cluster is stabilized by the vehicle speed and the distance between the vehicle and the cluster head. After the vehicle cluster is stabilized, the wireless spectrum resources are optimally allocated with the constraint of the interference value generated by the vehicle cluster users to the cellular users on the premise of satisfying the communication quality of the cellular users first. Finally, the simulation results show that the proposed algorithm has low complexity and can form more stable vehicle clusters. At the same time, the resource utilization rate is further improved to improve the overall system capacity. Although the proposed algorithm effectively reduces the communication instability in the IoV and improves the system capacity, the resource allocation algorithm has not been discussed in combination with power control. Future work will further study the control of vehicle transmit power to improve the overall performance.

## 6 Acknowledgement

This work was supported by Scientific Research Climbing Project of Xiamen University of Technology (Grant No. XPDKT19006), High-level Talent Project of Xiamen University of Technology (Grant No. YKJ20013R, and No. YKJ22030R), Natural Science Foundation of Fujian Province (Grant No. 2022J011276), and Education and Scientific Research of Young Teacher of Fujian province (Grant No. JAT190677, No. JAT200471, and No. JAT200479).

## References

[1] M. Chen, Y. Wang, On the Key Technologies of Internet of Vehicles and Its Innovative Application of Integration with 5G, in: Proc. 2020 International Conference on Artificial Intelligence and Electromechanical Automation, 2020.
[2] A. Gupta, R. Singh, D. Ather, R.-S. Shukla, Comparison of various routing algorithms for VANETS, in: Proc. 2016 International Conference System Modeling & Advancement in Research Trends, 2016.
[3] K. Abboud, H.-A. Omar, W. Zhuang, Interworking of DSRC and Cellular Network Technologies for V2X Communications: A Survey, IEEE Transactions on Vehicular Technology 65(12)(2016) 9457-9470.
[4] V.-D. Nguyen, O. Kim, D. Dang, S.S. Kim, C.S. Hong, Application of the lowest-ID algorithm in cluster-based TDMA system for VANETs, in: Proc. International Conference on Information Networking, 2015.
[5] M. Elhoseny, A. Hassanien, Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks. Studies in Systems, Decision and Control, vol. 242, Springer, Cham, 2020.
[6] O. Senouci, Z. Aliouat, S. Harous, MCA-V2I: A Multi-hop Clustering Approach over Vehicle-to-Internet communication for improving VANETs performances, Future Generation Computer Systems 96(2019) 309-323.
[7] G. Khayat, C.-X. Mavromoustakis, G. Mastorakis, J.M. Batalla, H. Maalouf, E. Pallis, VANET Clustering Based on Weighted Trusted Cluster Head Selection, in: Proc. 2020 International Wireless Communications and Mobile Computing (IWCMC), 2020.
[8] J. Yang, B. Pelletier, B. Champagne, Enhanced autonomous resource selection for LTE-based V2V communication, in: Proc. Vehicular Networking Conference, 2016.
[9] B. Wang, R. Zhang, C. Chen, X. Cheng, L. Yang, Y. Jin, Interference Hypergraph-Based 3D Matching Resource Allocation Protocol for NOMA-V2X Networks, IEEE Access 7(2019) 90789-90800.
[10] X. Zhang, Z. He, Y. Sun, S. Yuan, M. Peng, Joint Sensing, Communication, and Computation Resource Allocation for Cooperative Perception in Fog-Based Vehicular Networks, in: Proc. 2021 13th International Conference on Wireless Communications and Signal Processing, 2021.
[11] C. Fan, B. Li, Y. Wu, J. Zhang, Z. Yang, C. Zhao, Fuzzy Matching Learning for Dynamic Resource Allocation in Cellular V2X Network, IEEE Transactions on Vehicular Technology 70(4)(2021) 3479-3492.
[12] M.-I. Parizi, S. Rajabi, M. Ardebilipour, Joint Pairing and Resource Allocation for V2X Communications, in: Proc. 2020 10th International Symposium on Telecommunications (IST), 2020.
[13] Y.A. Shah, H.A. Habib, F. Aadil, M.F. Khan, M. Maqsood, T. Nawaz, CAMONET: Moth-Flame Optimization (MFO) Based Clustering Algorithm for VANETs, IEEE Access 6(2018) 48611-48624.