applied sciences

MDPI

Article

Information Relaying Methods in VANET: Algorithms, Standards and Tests

Tao Cui †, Chen Sun *,† and Lantao Li †

Sony R&D Center China, Beijing Laboratory, Beijing 100027, China
* Correspondence: chen.sun@sony.com; Tel.: +86-188-1028-6806
† These authors contributed equally to this work.

Abstract: This paper considers information transmission issue observed in a large scale test of cellular-based vehicular to everything (C-V2X) technology in China, which is the information blockage for the vehicle-to-vehicle (V2V) wireless communication. With such findings, we propose utilizing relaying-based information routing solutions by the support of K-means-type algorithm and an on-request relay selection method in a vehicular adhoc network (VANET). From the algorithm perspective, we refine the K-means-based VANET seeding principle with considerations on the global vehicle layout. On the other hand, base station and/or core network aided relay selection associated with vehicle maneuver is introduced for waking up relays when necessary in order to address the typical non-line-of-sight (NLOS) signal attenuation by taking the power consumption into consideration. To evaluate the system performance, we conduct Matlab and network simulator (NS)-3 simualtions to justify the algorithms, and adopt Veins-based simulations to show the data packets loss gain over non-relaying scenarios of realistic V2V scenario with NLOS signal dissemination.

Keywords: C-V2X; VANET; K-means; relay; NLOS; Veins

check for updates

Citation: Cui, T.; Sun, C.; Li, L. Information Relaying Methods in VANET: Algorithms, Standards and Tests. *Appl. Sci.* **2022**, *12*, 10748. https://doi.org/10.3390/app122110748

Academic Editor: Suhua Tang

Received: 26 July 2022
Accepted: 13 September 2022
Published: 24 October 2022

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

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Issuesfrom Real-World Tests

In recent years, in order to pave the way for the upcoming C-V2X commercialization, International Mobile Telecommunciations (5G), IMT-2020 (5G) for short, has launched a series of C-V2X tests within China, promoting the C-V2X technology to the real road test and discovering new potential issues. With our developed C-V2X hardware, i.e., onboard unit (OBU) and roadside unit (RSU) in Figure 1c,d, it is of great treasure for Sony for its very first attempt in a C-V2X real-world interoperability test in Wuhan, November 2021. During the 2-day activities, we observed the NLOS signal attenuation as a threat to the real deployment of C-V2X technology. A wide variety of reasons may be concluded including limited bandwidth, flooding data packets, and, the trickiest one, obstacles in the transmission path. Bushes, trees, buildings and even vehicles could become obstacles in V2V communications, leading to NLOS signal attenuation as a threat to the real deployment of C-V2X technology. Therefore, relaying-aided signal dissemination approaches (e.g., our location-based one) becomes fairly necessary to overcome the information blockage.

The test field is located in Wuhan economic development district, a blocked T-junction road area with 200 OBUs installed around to simulate C-V2X packet transmission in a crowded traffic area. The test result shows that the packet reception rate remains highly stable (e.g., 95% or above) within the transmission range of 50 ms, and gradually declines to 80% and even worse with transmission range growing up to 400 m. As observed in the bird's eye view Figure 1a, the road (i.e., path of B to D) is curved at a certain level, trees and bushes in the middle of the road happen to be a natural fence for long transmission between the sender and the receiver along the path of B to D, and the rough relief terrain between B and C also leads to typical NLOS transmission along the path of B to C. For such NLOS transmission, the worst test result shows that a data packet reception rate of 70% or even 60% may be achieved based on current 3GPP Rel-14 C-V2X chipset OBU.

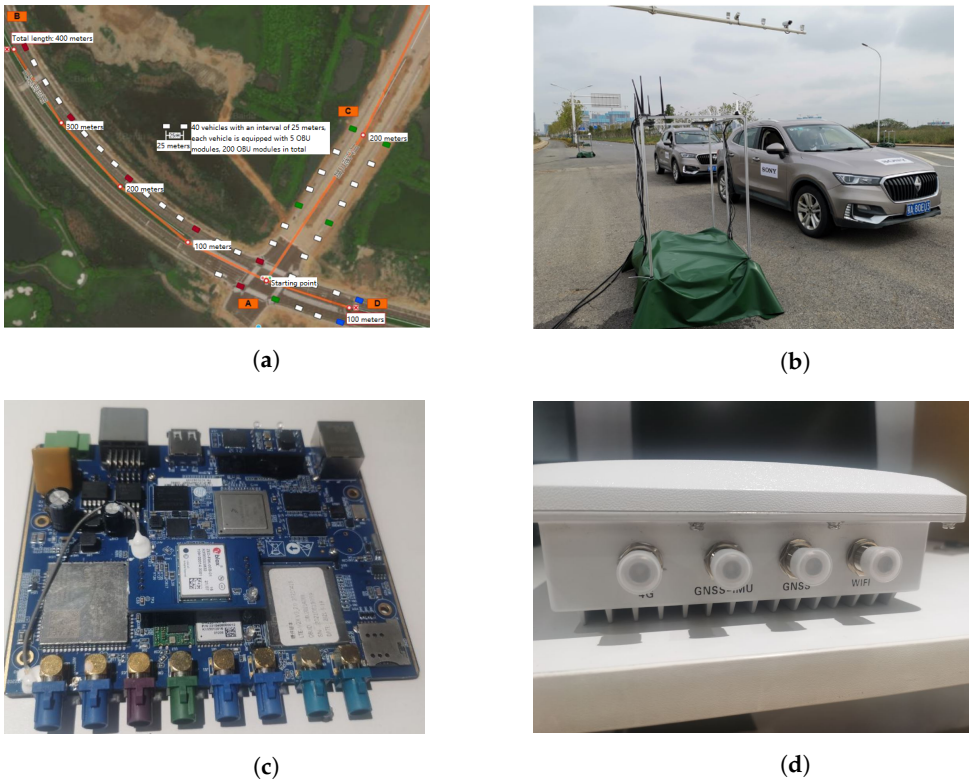


Figure 1. IMT-2020 Wuhan large-scale interoperability test and Sony test devices. (**a**) Test road bird view map. (**b**) Sony test vehicles. (**c**) Sony test OBU. (**d**) Sony test RSU.

2. Introduction

V2X (Vehicle to Everything) technology has been widely studied and implemented for road trial purposes in many places all over the world. Due to its natural advantages of the real-time traffic situation sharing, the road situation early warning can be easily realized for the driving safety and efficiency, which is very much superior to the environment detection results generated only by onboard sensors.

The VANET, as one of the well-known network architectures in the V2X use cases, has undergone significant investigations for its efficiency in routing messages to all other reachable vehicles and infrastructure within a given geographical area in a timely manner. One constraint of a VANET is the way that an appropriate relay vehicle is chosen and what the basic criteria is for determining the relay. Ref. [1] has studied a hybrid strategy for relay selection by considering the closest distance to the destination and best link status to bring in more robustness. Ref. [2] has considered the building shadow effect at intersections, utilizing the digital map and the Geohash coding system to relay the packets, and selecting the optimal relay node based on the relative position of the nodes and the intersection, to reduce the building shadow effect. Meanwhile, some other VANET relay papers focus on multi-hop relay, which is more practical in highway or suburban roadway scenarios. Ref. [3] has explored using the link quality as the key indicator for selecting multi-hop relay link paths to improve message reception performance over stringent communication conditions, and [4] has exploited the spatial diversity of the potential nodes from the sender for message reachability.

For improving vehicle grouping of a multihop vehicular network, clustering is introduced, which is a well-known unsupervised classification of data samples in a given space via the selection procedure to minimize an objective function. The K-means-type algorithm, as one of the most popular types of clustering, is formed by evaluating the summation of dissimilarities between each data point and its corresponding cluster center [5–9]. The dissimilarity refers to the mean squared distance from each data point to its nearest center. Note that these cluster centers are virtual centroids used for attracting data points within

their coverage and that no overlap is theoretically allowed between different data sections. In recent research, ref. [10] proposes an improved K-means algorithm by considering attributes with respect to vehicle movement, where the typical K-means tailored for actual VANET deployment adopts precise cluster head vehicle selection and threshold-based algorithm termination conditions. Ref. [11] also takes into account the initialization phase of cluster heads, which is to use a Maximum Stable Set Problem(MSSP) combined by the Continuous Hopfield Network (CHN) to secure a stable set of cluster head vehicles. Ref. [12] proposes the location based methods as well but with the assumption of the property of the destination vehicle as a cluster head.

Some other relevant works considering cluster-based approaches on message routing principles are widely studied as well. In [13], in order to increase the route stability and average throughput, the current and destination location of vehicle nodes are used to select the optimal route via several routing protocols in different scenarios. Another cluster stability solution in [14] proposes a combination of a bully algorithm and a Lamport timestamp for selecting the better cluster head and for switching the cluster head in an efficient manner. Ref. [15] concentrates on direction of movement similarity for the selected cluster head and the neighbor one, to use their position coordinates and message forwarding destination, combing the direction of transmission of the message by the node. All in all, five parameters—location, direction, speed, destination and list of interests—jointly form different routing protocols for a stable cluster network, but with no efforts to the investigation and analysis of power efficiency and algorithm convergence.

Our proposal, by contrast, holding the viewpoint of ensuring an efficient power and relaying network, considers a more realistic VANET topology with real centroids and a corresponding seeding criterion, implemented in a range-restricted manner. With the help of the enhanced K-means algorithm, the vehicle closest to the virtual mean center vehicle in the set is selected as a relay vehicle, and the seeding area in the initialization procedure is subject to the maximum transmission power of the relay vehicles. Speaking of the route optimization, our solution introduces connecting vehicles as a low power message dissemination node to bridge the selected cluster head and its neighbor cluster heads. To extend the scope after the trial test thinking, the base station and core network are incorporated with the potential of Third Generation Partnership Project (3GPP) standard enhancement of the architecture.

The rest of this paper is organized as follows: Section 3 gives a typical V2X radio propagation model with the help of a randomly distributed vehicle map on a four-lane highway. In Section 4, K-means-type algorithms with centroid and seeding enhancements are proposed. Through the experience of IMT-2020 organized C-V2X large-scale test, a high level location based solution with specification impact is summarized in Section 5. In Section 6, we carry out MATLAB, NS3 and Veins simulations to evaluate the performance of the proposed algorithms and the retro from the real-world test followed by corresponding analytics and speculations in Section 7. Finally, Section 8 concludes the paper.

3. System Model

A V2V communication scenario is considered in this section, as shown in Figure 2. Assume that the number of vehicles in a traffic area is m, which is denoted as $V = \{v_1, v_2, ..., v_m\}$. Among those that are randomly distributed exists a broadcasting vehicle we call the source vehicle, which broadcasts important messages (e.g., Emergency Situation Warning Message) with transmission power P_0. Define two sets of vehicles, i.e., relay vehicles and free vehicles, which are denoted as R and N, respectively. They satisfy $V = R \cup N$. Then, we have the following description of the two vehicle sets mentioned above:

$$R = \{v_n \mid \delta_n = 1, v_n \in V\}, \quad (1)$$

$$N = \{v_n \mid \delta_n = 0, v_n \in V\}, \quad (2)$$

where δ_n is a binary variable, with v_n being a relay vehicle when $\delta_n = 1$ and v_n being a free vehicle when $\delta_n = 0$. When $\sum \delta_n \geq 1$, at least one relay vehicle exists in the traffic area.

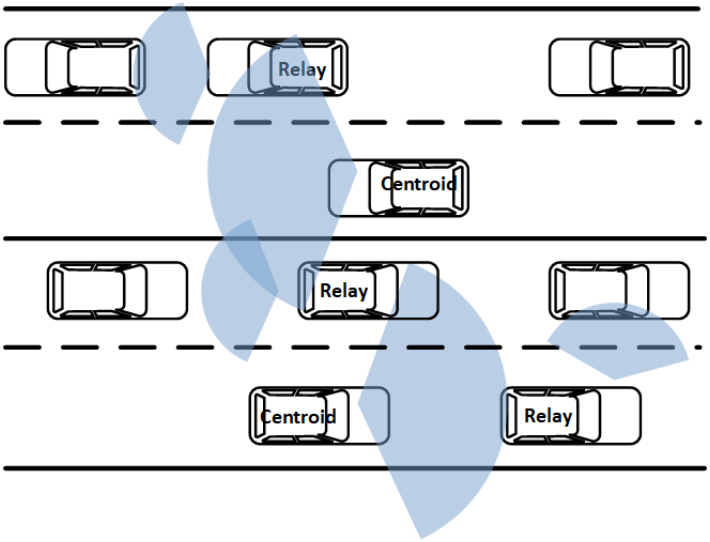


Figure 2. Information broadcasting relaying in a four-lane and two-direction highway.

4. Enhanced K-Means-Type Algorithm

Recall that the conventional K-means-type algorithm was designed for describing the process of partitioning an N-dimensional observation into k sets on the basis of a sample [5]. When clustering points with the squared Euclidean distance metric, the procedure described in [5] gives the following objective function that updates the mean center of a cluster to find the closest center for each point:

$$\phi = \sum_{v \in V} \min_{c \in C} \|v - c\|^2, \tag{3}$$

where $v_1, v_2, \cdots, v_m \in V$ are the m vehicle nodes defined in the system settings of Section 3 and C is the centroid set, with each element c representing the center of a corresponding cluster.

4.1. Real Center Mean Exploration

The center mean $\mu(C_i) = |C_i|^{-1} \sum_{x \in C_i} x$ is well defined in [9], and the intermediate consequence keeps updating until the mean squared error (MSE) of two consecutive mean centers is smaller than a given threshold. By an analysis of the iterative procedure, it is not difficult to find that the cluster center selected is not a real data point in the set but a statistical mean value, which we call a virtual center. To make the algorithm fit the system model proposed in the VANET, the closest vehicle to the virtual mean center vehicle in the set should be chosen as the cluster center for the information rebroadcasting. In this regard, we provide two candidate enhanced K-means-type algorithms with consideration of the update timing of different real centers.

The K-means-type1 algorithm given in Algorithm 1 shows a virtual-to-real centroid transition in each iteration via the centroid recalculation once the subspace (cluster) has changed. Specifically, step 4 to step 6, which we designed for the VANET, differentiate this algorithm from the typical K-means algorithm. It is well worth noting that θ_{th} is a numerical threshold to assure two nodes selected adjacent in time are close enough.

Algorithm 1 K-means-type1 Algorithm for the VANET

Require:
Number of clusters required for the network;
The set of initialized cluster centers, c_i;
Ensure:
Real cluster center closest to the virtual center mean, v_i;
Connecting vehicle, v_j;
1: Arbitrarily select k clusters with the center initialization $C = \{c_1, c_2, ..., c_k\}$;
2: Partition the vehicle set via ϕ by calculating the distance between each individual vehicle and all centers to find the nearest cluster;
3: Update the centroid c_i via the mean center calculation function $\mu(C_i) = |C_i|^{-1} \sum_{x \in C_i} x$ in each cluster C_i;
4: Map the mean center $\mu(C_i)$ to the closest vehicle $v_i \in V$ within each cluster C_i;
5: Repeat 3 and 4 until $\|v_i(t) - v_i(t+1)\|^2 \leq \theta_{\text{th}}$ is satisfied;
6: Pick $v_j \in V(j \neq i)$ that is closest to the center of two sequential centroid vehicles;
7: **return** v_i and v_j;

The other enhanced K-means type algorithm, called K-means-type2, which is shown in Algorithm 2 in parallel, calculates the real vehicle center after the algorithm converges in the form of virtual center means. Similar to the Type1 algorithm, a numerical threshold λ_{th} is set for the convergence of a virtual center mean. The two algorithms theoretically share many similarities, with the only difference being the order of step 4 and step 5 and the parameters therein. It is also well noted that the algorithm selects the connecting vehicle v_j by finding the closest approximation to the midpoint of the distance between two selected consecutive real center vehicles, and this procedure fits for both algorithms. The outcome achieved via the enhanced K-means algorithms matches the actual network layout better, even though a double approximation on the cluster centers and connected relay vehicles is implemented.

Algorithm 2 K-means-type2 Algorithm for the VANET

Require:
Number of clusters required for the network;
The set of initialized cluster centers, c_i;
Ensure:
Real cluster center closest to the virtual center mean, v_i;
Connecting vehicle, v_j;
1: Arbitrarily select k clusters with the center initialization $C = \{c_1, c_2, ..., c_k\}$;
2: Partition the vehicle set via ϕ by calculating the distance between each individual vehicle and all centers to find the nearest cluster;
3: Update the centroid c_i via the mean center calculation function $\mu(C_i) = |C_i|^{-1} \sum_{x \in C_i} x$ in each cluster C_i;
4: Repeat 3 until $\|c_i(t) - c_i(t+1)\|^2 \leq \lambda_{\text{th}}$ is satisfied;
5: Map the mean center $\mu(C_i)$ to the closest vehicle $v_i \in V$ within each cluster C_i;
6: Pick $v_j \in V(j \neq i)$ that is closest to the center of two sequential centroid vehicles;
7: **return** v_i and v_j;

4.2. VANET Seeding Principle

Because the clustering problem is NP-hard with only local optimality, the cluster centers selected have very low correlation with each other. For a system of the VANET, this seeding principle may not work well since a proper network topology has to ensure that the relay vehicles stay in the effective coverage area of one or more of them. In other words, to cause the information broadcasting to occur in a continuous way, the maximum

transmission power of the relay vehicles should be considered. As such, we modify the existing objective function (3) by making it subject to the following condition:

$$\phi(v_i, v_{i+1}) \leq D^* \quad i = 1, \ldots, k-1, \tag{4}$$

where D^* is the maximum distance allowed between two adjacent centroid vehicles. Note that nonoverlaping centroid vehicles correspond to the k clusters initialized in the algorithms mentioned above. Directly incorporating (4) into the algorithms proposed in Section 4 would more or less slow down the algorithm convergence due to the mismatch between globally searching and locally restricting cluster centers. To avoid this drawback, we propose applying the random seeding principle with restriction of the range at the very beginning of the algorithm. Specifically, the centroid vehicles are initialized separately and locally fulfill (4) until the algorithm becomes stable instead of arbitrarily selecting the cluster centers globally. To this end, Section 6 takes these limits into account when implementing the two proposed algorithms.

4.3. Mathematical Formulation of the VANET Seeding Principle

Recalling the vehicle nodes and cluster centers defined in (3), we recalculate the formulation of the objective function in [6] by imposing the constraints mentioned in Section 4.2. The following mathematical problem is obtained with a restriction on the distance between two consecutive cluster centers, which is formulated in (4) and corresponds to the proposed seeding principle for the VANET.

$$\text{minimize} f(W, C) = \sum_{i=1}^{k} \sum_{j=1}^{m} w_{ij} D\left(v_j, c_i\right) \tag{5}$$

subject to

$$\begin{aligned} &\sum_{i=1}^{k} w_{ij} = 1,\ j = 1, 2, \cdots, m, \\ &w_{ij} = 0 \text{ or } 1,\ i = 1, 2, \cdots, k, \\ &\sum_{i=2}^{k} \left| c_i^{(l)} - c_{i-1}^{(l)} \right| \leq (k-1) D^*, \end{aligned}$$

where W is a $k \times m$ real matrix, $C = [c_1, c_2, \cdots, c_k]$ is the set of cluster centers, $D\left(v_j, c_i\right)$ measures the dissimilarities between v_j and c_i, and l represents the number of iterations.

Proposition 1. *c_k^* is the local minimum cluster center of the objective function (5).*

Proof. To obtain the local minimum cluster center, we adopt the Lagrange multipliers method to form a Lagrangian function with the KKT condition as follows.

$$\begin{aligned} L(c, \mu) = &\sum_{i=1}^{k} \sum_{j=1}^{m} w_{ij} D\left(v_j, c_i\right) \\ &+ \sum_{j=1}^{m} \mu_j \left[\sum_{i=2}^{k} \left| c_i^{(l)} - c_{i-1}^{(l)} \right| - (k-1) D^* \right] \\ &+ \sum_{j=1}^{m} \lambda_j \left(\sum_{j=1}^{k} w_{ij} - 1 \right) \end{aligned} \tag{6}$$

By setting $\bigtriangledown_{c_k} L = 0$, we obtain a potential local minimum through

$$\frac{\partial L}{\partial c_k} = \frac{\partial A}{\partial c_k} + \frac{\partial B}{\partial c_k} = 0, \tag{7}$$

where

$$A = \sum_{i=1}^{k}\sum_{j=1}^{m} w_{ij} D(v_j, c_i)$$

$$B = \sum_{j=1}^{m} \mu_j \left[\sum_{i=2}^{k} \left| c_i^{(l)} - c_{i-1}^{(l)} \right| - (k-1)D^* \right]$$

$$\frac{\partial}{\partial c_k} \sum_{j=1}^{m} \lambda_j \left(\sum_{j=1}^{k} w_{ij} - 1 \right) = 0.$$

Then, we obtain

$$c_k^* = \frac{\sum_{j=1}^{m} w_{ij} v_j + \mu_j c_{k-1}}{\sum_{j=1}^{m} w_{ij} \mu_j}, \tag{8}$$

$$\sum_{j=1}^{m} \mu_j = \frac{\sum_{j=1}^{m} w_{ij} (c_k - v_j)}{c_{k-1} - c_k}. \tag{9}$$

Suppose that c_k^* fulfills the local minimum requirement; then, a unique μ^* exists, subject to the following:

$$\nabla_{c_k} L(c_k^*, \mu^*) = 0, \tag{10}$$

$$\mu^* \geq 0, \tag{11}$$

$$\mu^* g(c_k^*) = 0, \tag{12}$$

where $g(c_k^*) = \sum_{i=2}^{k} \left| c_i^{(l)} - c_{i-1}^{(l)} \right| - (k-1)D^* \leq 0$. For (12), consider two cases: $\sum_{j=1}^{m} \mu_j = 0$ and $\sum_{j=1}^{m} \mu_j > 0$. Thus,

$$c_k^* = \max \left\{ v_j \& w_{ij} = 1, \frac{\sum_{j=1}^{m} w_{ij} v_j + \mu_j c_{k-1}}{\sum_{j=1}^{m} w_{ij} + \mu_j} \right\}. \tag{13}$$

It is obvious that $\mu^* g(c_k^*) = 0$ if $\mu^* = 0$. When $\mu^* \geq 0$, the following approximation is obtained:

$$\lim_{l} g(c_k^*) = \sum_{i=2}^{k} \left| c_i^{(l)} - c_{i-1}^{(l)} \right| - (k-1)D^* \approx 0. \tag{14}$$

Hence, c_k^* is the local minimum cluster center position of the objective function (5). This ends the proof. □

As an NP-hard problem, the result could be easily falling into the local optimal circle instead of a global one due to its nature. Any deterministic algorithms targeting the optimization of the system performance in terms of minimizing or maximizing the objective function may have negative impacts on the overall performance. As such, we proposed to partition the clustering layout for the sake of local optimal relaying network, then connect different clusters for a nearly global optimization outcome.

4.4. Standards Impact

How do the real-time local topology and the calculated optimal centroids work in the actual V2X implementation? Based on the design of the application layer of a typical communication protocol [16,17], we propose extending C-V2X messages, e.g., *Emergency Situation Warning Message* and *Special Vehicle Warning Message* by adding new data elements of the broadcast relay node and lane direction as needed [18]. Note that the specific information included in the broadcast relay node related to the target relay vehicles is the identification, position and relay

type, with which several applications such as intervehicle security are well supported. In the case of vehicle-controlled broadcasting, the local topology is signalled either by the base station through a Uu link or by the RSU through the PC5 interface. In addition to a V2V information relaying, the V2I scheme is also considered to be applicable though we did not make it into our enhanced K-means algorithm. In the following section, we elaborate a scenario with the consideration of participated RSU/base station/core network.

5. Real-World Test Summary

It is a popular design in the C-V2X demonstration area for the RSU to be deployed at either the intersection or highway road side for information collection, information refinement and relay function. However, requiring substantial infrastructure costs, the vehicle-based relay may provide another route. Though such relay candidate selection area seems easy to find, the various real road conditions introduce a wide variety of scenario-specific difficulties in accurately identifying such area in detail. However, with the development of the network data analytics function (NWDAF) and other AI topics such as Federated Learning introduced into 3GPP, the exposure of VANET communication characteristics from UE to gNodeB and core network may provide a platform for data analysis, relay on-request area and relay candidate selection area determination. Similar to how the base station manages the area of the cell, the high-definition map can also be divided into different VANET communication blocks with the intersection area, the bend area and semi open location close to obstacles as relay node selection areas. The mechanism to define transmission block and intersection should be based on the quality of service (QoS) of packet transmission; the QoS of the packet transmission within one VANET communication block shall remain aligned and stable, while Tx UE and Rx UE belonging to different communication blocks differ largely with location differences. The relay node selection area shall bridge multiple communication blocks, while the UEs inside the relay node selection area can establish a stable short range transmission link with the UE in those blocks within the surroundings. The base station may take the role of data collector to identify short range data transmission quality characteristics with the locations of both sender and receiver, reporting the statistics to the core network, enabling NWDAF to analyze the VANET transmission quality in relation to different areas with the help of the location management function (LMF), to decide VANET communication blocks and relay node selection areas. Vehicle type UE entering a relay node selection area can be validated by the base station or Core Network functions as relay nodes dynamically minimize the NLOS transmission from obstacle (e.g., buildings, vehicles) shadow effects.

6. Results

6.1. Relaying Algorithm Evaluation

To evaluate the system performance when the proposed algorithms were adopted for the VANET relay selection, a MATLAB-NS3-based method is utilized in this simulation. A brief files coordination flow is shown in Figure 3 to give a general understanding of different platform connections. The simulation platform is developed based on the scenario and parameters defined in *3GPP TR36.885* [19]. Our simulation focuses on the highway scenario, where the relay node selection targets those mobile vehicles heading in the same direction. When the vehicles reach the margin of the designed area, they are reassigned to the entrance of the highway, and this procedure repeats as a closed-loop process. The specific parameter configuration of the simulation is given in Table 1. It is well worth noting that two parameters involved in our proposed algorithms, which are the number of clusters required for the network and the initialized cluster center for each cluster. For simplicity, 2 clusters are selected to cover the overall traffic and the initialized cluster centers are randomly nominated within each cluster correspondingly. It is certain that if the number of clusters increases, the message routing delay expands as multiple hops are introduced. Additionally, the range of the blind zone shrinks as the location of clusters adjustment becomes more flexible.

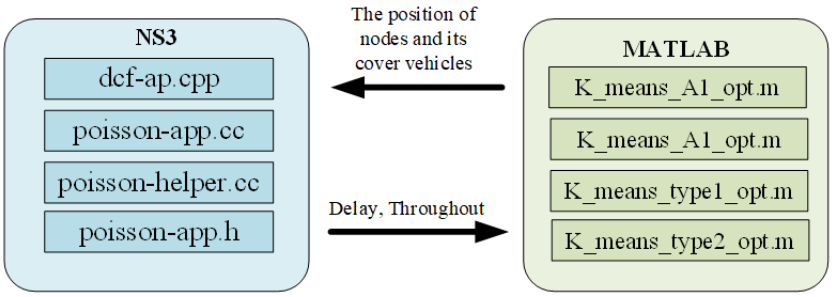


Figure 3. Matlab-NS3 files coordination.

Table 1. Simulation-1 parameter settings.

Parameter	Configuration
Num. of Lane	6, Single Direction
Wid. of Lane	3 m
Len. of Lane	3000 m
Num. of Vehicles	200
Velocity of Vehicles	≤120 km/h
Bandwidth	10 MHz
PSD of Noise	−204 dB/Hz
Carrier Freq.	5.9 GHz
Tx Power of Source Vehicle	24 dBm
Tx Power of Centroid Vehicle	27 dBm
Tx Power of Connecting Vehicle	24 dBm

6.2. Relaying Mechanism Evaluation

Except for the above algorithm itself helping to find the optimal information relaying vehicles, a high level approach is considered from the perspective of the relaying mechanism. We as such conduct Veins-based simulation including relays to virtually recover and relieve information blockage exposed in the real-world test. The reason behind choosing Veins is that, besides the signal path loss model provided, the Obstacle Shadowing model [20] and Vehicle Obstacle Shadowing model [21] added in the latest version of Veins make it easy to focus on how to define a better optimized relay selection algorithm in the upper layers to enhance VANET communication coverage.

In the simulation, a very typical urban intersection area near Berlin (see in Figure 4) is selected with around 30 vehicles near or approaching the intersection center, the vertical road has 6 lanes and the horizontal road has only 2 lanes; thus, different densities of vehicles are applied to different roads. The transmission parameters setting is as follows (in Table 2).

Table 2. Simulation-2 parameter settings.

Parameter	Configuration
Layout size	4800 × 1600 × 50 m
Velocity of Vehicles	≤108 km/h
Bandwidth	20 MHz
PSD of Noise	−204 dB/Hz
Carrier Freq.	5.9 GHz
Tx Power of Vehicle	40 mW
Bitrate	6 Mbps
Minimum power level	−110 dBm
noise power level	−98 dBm

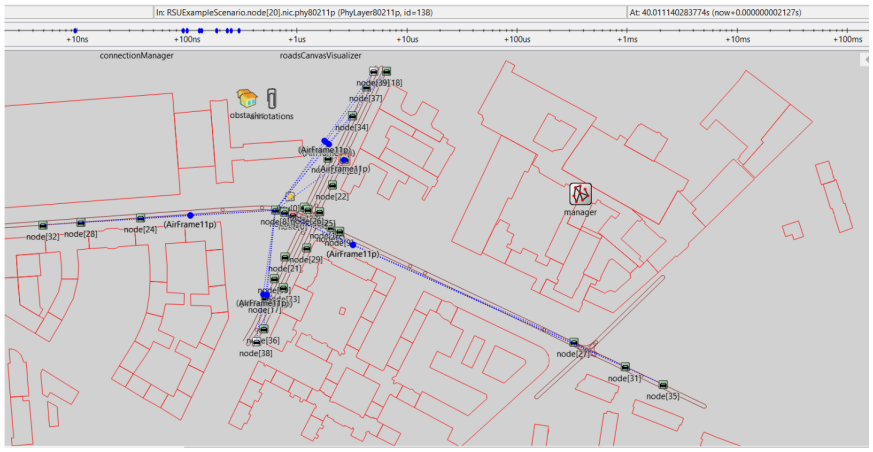
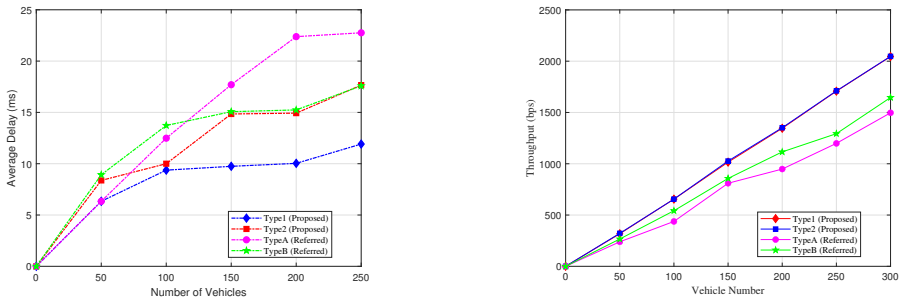


Figure 4. Real-world simulation map.

7. Discussion

One of the important features to evaluate our enhanced K-means-type algorithm is the information relaying delay. In Figure 5a, a local maximum information transmission delay is calculated, i.e., a clustering-led maximum delay is nominated to evaluate the performance. It is well worth noting that our proposed K-means-type1 algorithm keeps a fairly low profile with the ascending number of vehicles on the road, and the end-to-end delay from either source to cluster head or cluster head to the ego vehicle maintains around 10 ms which is sufficiently smaller than two of the algorithms referenced [10,11].



(**a**) (**b**)

Figure 5. Proposed algorithms delay and throuput performance evaluation. (**a**) Information broadcasting delay. (**b**) System throughput vs. the number of vehicles.

To better evaluate the performance of the proposed clustering algorithms, it is necessary to observe the throughput of the system in comparison with the legacy approaches in [10,11]. In Figure 5b, we evaluate how the algorithms work from a sparse scattering scenario to a dense one. It can be seen that the throughput calculated by implementing our proposed algorithms shows an ascending superiority over two referred algorithms when dropping more vehicles in the layout. The advantage could be greater than 30% in terms of performance gain @300 vehicles scenario. In particular, two proposed algorithms show almost no difference due to the reason that the order of centroids selection contributes little to system performance instead of the algorithm itself.

Figure 6 investigates a comparison of the convergence speed by implementing the proposed algorithms and typical K-means algorithm under different metrics. The metrics in regards to the distance calculation are derived based on the metric criterion in [6]. The cluster centers calculated via the proposed algorithms become stable after only three to four iterations, in contrast to the eight iterations for the typical K-means algorithm in the case of evaluating the member quantity in each cluster. With more stringent restrictions on the algorithm seeding range, it is well understood that the shrinking search space of the

proposed algorithms effectively contributes to finding the optimal relay vehicles with a priori information of the real-time topology.

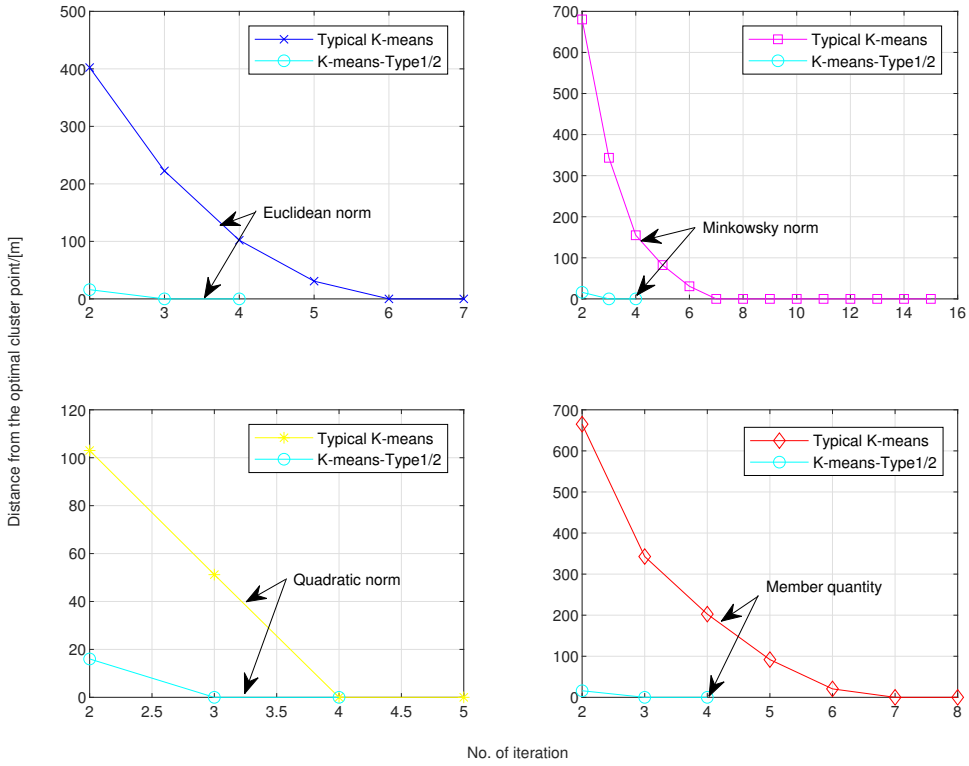


Figure 6. Proposed algorithm convergence speed under different evaluation metrics.

In Figure 7, without any relays, vehicles alongside the horizontal road achieved the data packet reception rate at 29.5% while vehicles in the vertical road at 75.1%, and vehicles within 50 meters range to the intersection center could achieve 95% above data packet reception if dividing simulation periods and focusing on packet reception rate variance corresponding to location variance. This is a natural consequence since the line-of-sight (LOS) transmission path dominates the scenario when either the sender or the receiver is at the intersection, while NLOS is typical when both sides are far from the center and not in the same section of road. The square area with 100 m side length is specifically selected as the relay selection area, all vehicles within this area could switch their relay function on, and the maximum relay nodes number is applied to minimize redundancy of relayed transmission. Various methods could be used for relay node candidate confirmation, in the Veins simulation we use our proposed K-means type1 algorithm to initially pick the relay vehicles up, after that the announcement message is used to identify the existence of relay nodes to eliminate potential relay redundancy. It is well noted that the candidate node must listen for a period before switching on relay function. With relay switched on in the simulation, the performance increased to 91.3% for horizontal direction vehicles and 98.9% for vertical direction vehicles, at the cost of 1.99 times more transmission budget for relaying data packet.

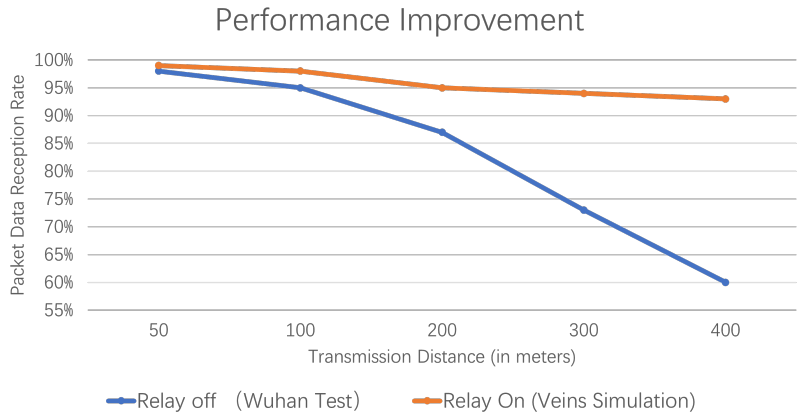


Figure 7. Data reception rate performance improvement with relation to transmission distance.

8. Conclusions

To address the issues found in the real-world test, we propose a cluster-based relaying vehicle selection scheme with a restricted seeding procedure. Taking celluar-based network architecture into account, Veins-based simulation is facilitated to recover the situation in the test, and the relaying based approach is nominated and compared. From the simulation results, the performance gain in terms of system throughput, latency, and packet received rate could be achieved by the comparison with the state-of-the-art and the no relaying case.

Author Contributions: Conceptualization, T.C. and C.S.; methodology, T.C.; software, T.C. and L.L.; validation, T.C., C.S. and L.L.; formal analysis, T.C.; investigation, T.C. and L.L.; resources, C.S.; data curation, T.C. and L.L.; writing—original draft preparation, T.C. and L.L.; writing—review and editing, C.S.; visualization, T.C.; supervision, C.S.; project administration, C.S.; funding acquisition, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: Special thanks to Sony RDC Beijing Lab head Takenaka -san for his consistent support of our V2X for autonomous driving project. Meanwhile, we are also gradeful for the real-world opportunity of China C-V2X large scale test organized by IMT-2020 (5G) promotion group and CAICT.

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

Sample Availability: Samples of the compounds are available from the authors.

Abbreviations

The following abbreviations are used in this manuscript:

C-V2X	Cellular-based Vehicle to Everything
VANET	Vehicular Ad-Hoc Network
NLOS	Non Line-of-Sight
MSSP	Maximum Stable Set Problem
CHN	Continuous Hopefield Network
MSE	Mean Square Error
RSU	Road Side Unit
NWDAF	Network Data Analytics Function
QoS	Quality of Service
LMF	Location Management Function
CAICT	China Academy of Information and Communications

References

1. Liu, K.; Niu, K. A hybrid relay node selection strategy for VANET routing. In Proceedings of the IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, China, 22–24 October 2017; pp. 1–6.
2. Wu, J.; Lu, H.; Xiang, Y.; Wu, R.; Wang, F. MBR: A Map-Based Relaying Algorithm for Reliable Data Transmission through Intersection in VANETs. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 3661–3674. [CrossRef]
3. Rehman, O.; Ould-Khaoua, M. Link Quality Driven Hybrid Scheme for Multi-hop Message Broadcast in VANETs. In Proceedings of the IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 1–3 November 2018; pp. 479–485.
4. M. Hussain Rehman, O.; Bourdoucen, H.; Ould-Khaoua, M. Hybrid Relaying for Multi-Hop Broadcast in VANETs. In Proceedings of the IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB), Montreal, QC, Canada, 4–7 October 2015; pp. 1–5.
5. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In *Berkeley Symposium on Mathematical Statistics and Probablity*; University of California Press: Berkeley, CA, USA, 1976; Volume 1, pp. 281–296.
6. Selim, S.Z.; Ismail, M.A. K-means-type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 81–87. [CrossRef] [PubMed]
7. Arthur, D.; Vassilvitskii, S. K-Means++: The advantages of careful seeding. In Proceedings of the SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, LA, USA, 7–9 January 2007; pp. 1027–1035.
8. Morissette, L.; Chartier, S. The k-means clustering technique: General considerations and implementation in Mathematica. *Tutor. Quant. Methods Psychol.* **2013**, *9*, 15–24. [CrossRef]
9. Telgarsky, M.; Vattani, A. Hartigan's Method: k-means Clustering without Voronoi. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 820–827.
10. Ye, L.; Chen, Y.; Han, Q.; Zeng, L.; Cheng, S.; Xiao, L.; Ding, X.; Vattani, A. Vehicle Message Distribution Mechanism Based on Improved K-means Adaptive Clustering Algorithm. In Proceedings of the IEEE 23rd International Conference on Intelligent Transportation System, Rhodes, Greece, 20–23 September 2020.
11. Kandali, K.; Bennis, L.; Bennis, H. A New Hybrid Routing Protocol Using a Modified K-Means Clustering Algorithm and Continuous Hopfield Network for VANET. *IEEE Access* **2021**, *9*, 47169–47183. [CrossRef]
12. Yu, F.; Zhang, C.; Tang, S. Unified Performance Analysis of Stochastic Clustered Cooperative Systems with Distance-based Relay Selection. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 6180–6194. [CrossRef]
13. Abuashour, A.; Michel, K. Performance Improvement of Cluster-Based Routing Protocol in VANET. *IEEE Access* **2017**, *5*, 15354–15371. [CrossRef]
14. Dharmawardena, P.K.; Wang, Z. Cluster Head Selection Based Routing Protocol for VANET Using Bully Algorithm and Lamport Timestamp. *Int. J. Comput. Theory Eng.* **2017**, *9*, 218–222. [CrossRef]
15. Anas, A.T. VANET Routing Protocols and Architectures an Overview. *J. Comput. Sci.* **2018**, *14*, 423–434.
16. *DSRC Implementation Guide—A Guide to Users of SAE J2735 Message Sets over DSRC*; SAE International: Warrendale, PA, USA, 2016.
17. *Cooperative Intelligent Transportation System-Vehicular Communication Application Layer Specification and Data Exchange Standard (Phase II)*; China SAE: Beijing, China, 2020.
18. Li, J.; Cui, T.; Sun, C. Electronic Equipment, Wireless Communication Method and Computer Readable Storage Medium. European Patent Application WO2021185157 (A1), 23 September 2021.
19. *3GPP TR 35.885 v2.0.0*; Study on LTE-Based V2X Services (Release 14). 3GPP: Valbonne, France, 2016.
20. Sommer, C.; Eckhoff, D.; Dressler, F. IVC in Cities: Signal Attenuation by Buildings and How Parked Cars Can Improve the Situation. *IEEE Trans. Mob. Comput.* **2014**, *13*, 1733–1745. [CrossRef]
21. Sommer, C.; Joerer, S.; Segata, M.; Tonguz, O.K.; Cigno, R.L.; Dressler, F. How Shadowing Hurts Vehicular Communications and How Dynamic Beaconing Can Help. *IEEE Trans. Mob. Comput.* **2015**, *14*, 1411–1421. [CrossRef]