

# A Clustering Approach to Model the Data Dissemination in VANETs

Saad T. Hasson

*Information Networks Department  
College of Information Technology  
University of Babylon  
Babylon, Iraq  
saad.aljebori@uobabylon.edu.iq*

Abbas T. Abbas

*Information Networks Department  
College of Information Technology  
University of Babylon  
Baghdad, Iraq  
abbas.abbas@student.uobabylon.edu.iq*

**Abstract:** Recently, Vehicular Ad hoc Networks (VANETs) is focused on by many researchers. VANET's specific characteristics like high-speed, unstable communication link, network partitioning, and information transfer become a hard challenging problem. Access control depends on IEEE 802.11p and has been used for Vehicle to Everything (V2X) and Vehicle to Vehicle (V2V) communications. The high dynamics of VANET nodes which will cause frequent disconnection of links and even partitioning of networks causes the source vehicle is lost their message of data that disseminated to the destination region, and also may lead to the message redundancy, causing a broadcast storm. One of the suggested solutions for this in VANET is the clusters. This paper will focus on the dissemination of cluster-based data. The clustering algorithm is provided according to the highway road, the proposed algorithm seeks to increase the network stability by identifying the Cluster Head (CH) periodically. By using a suggested variable range of cluster head (R-of-CH) in this paper to be a measure of finding the CH for each new vehicle. The simulation results are performed by NetLogo, under different vehicle densities and different values of (R-of-CH). The results of the simulation show the proposed algorithm contributes to making the network more stable with a high delivery rate as the disseminated messages will reach all vehicles in the simulated area. It indicates that the coverage rate is also high and the lost message rate is very low.

**Index Terms**—VANETs, data dissemination, clustering, NetLogo, simulation.

## I. INTRODUCTION

There is a positive impact of VANETs on transportation systems. Whereas, vehicle intelligence helps develop both an intelligent environment for information sharing and secure transportation systems. The main components of this environment are the vehicles that communicate wirelessly with each other, as the vehicles have strong capabilities in networks, sensing, data processing, and communication, as well as information exchange [1]. VANET requires transferring data between them be continuous and wireless. The transmitting data task among vehicles becomes simple by wireless communication [2]. But vehicle speed affects wireless communication over a wide scale, and the network topology becomes unstable, then strong data dissemination protocol configuration is required [3]. Existing VANET data dissemination mechanisms will be categorized into three models: push, pull, and hybrid. In the model push, the data is published using the periodic broadcast proactively, but in the pull form, it is done upon request [4]. VANTs have been known as a "self-originating traffic information system" in

several applications [5]. Vehicles move on predetermined roads only and have not a problem with resource limitations in data storage and energy [6]. VANET has been Represent as a part of intelligent transportation systems (ITS). It is a very powerful and intelligent system for obtaining data transmission between vehicles and good communication while traveling on roads. Vehicles can communicate as "vehicles to everything" (V2X) where they use a phone tower nearby to send and receive long-range messages and communicate with the infrastructure and others. while systems that only allow communication between vehicles are called "vehicles to vehicles" (V2V). VANET can also combine V2V and V2X. with a roadside unit that makes it much easier to store local data or transmit messages over long distances[7].

Figure 1 shows the node communication types in VANETs. Vehicular systems can be considered complex systems due to their unpredictable behavior and high mobility [8]. Also, these systems are important nowadays which are used in avoiding accidents, transmitting messages between vehicles, implementing safety by detecting obstacles, transmitting an emergency message by using broadcast technologies, interactive multimedia service, and internet access at a high data rate especially video calls [9]. The information transmission technology in VANETs uses a peer-to-peer (P2P) communication method without any restrictions unless the vehicles are out of the transmission range [7]. But in last years, clustering has become in VANETs one of the more controls used and information transmission mechanisms to resolve their different challenges and enhance their performance. The cluster is characterized as the system for grouping vehicles into small groups named clusters by pre-defined metrics, including the density of nodes, the average speed, and the position of nodes. This clustering is performed to provide an efficiently controlled network according to the application requirements [10].

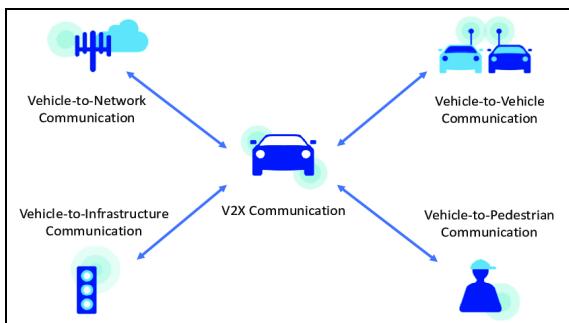


Fig. 1. VANET Architectures

## II. RELATED WORK

Given the importance of improving road safety and vehicular traffic quality, several data dissemination plans proposed. There are some related suggestions.

O. Urmonov et al. [11] suggested implementing relay node selection speed, regardless of a network scenario. In the relay assignment stage, all receivers that claim to have become a relay node have equal chances. Thus, a broadcast message can be posted in one or more directions. Also, the sending device can determine whether the message should be forwarded or not. According to the results, this method achieved a high relay coverage and a high message delivery rate. In [12] S. Shah et al. proposed a new way to spread emergency messages of time importance. The system uses clustering technology to address the problem of a broadcast storm and to reduce network congestion used the dynamic time barrier. the delay is reduced by 12% when compared with other known technologies. The transmission rate is increased. This mechanism works better at lower speeds and high density of the vehicle due to the network time increased. also Mr. Shahal and others. [13] Description of a future-based, reliable, and intelligent data dissemination protocol. During data transmission, the link duration will be used to choose the node that will be the next forwarding. Greed and recovery algorithms are based on edge weight that is calculated from the duration of the association. The results viewed that the proposed system is more secure with a slight improvement in latency in terms of message delivery rate. This method ensures that the most reliable data transfer paths are selected. Whereas proposition D. Wu et al. In [14]a new algorithm for distributing data based on neighbors' position in VANET trying to decrease the effect of vehicle movement, by creating a lot of effective set of relay nodes by predicting the true speed and location of neighboring nodes. view in the results of Simulation that the proposed algorithm file performs better than several conventional protocols. While O. Rehman et al. In [15] the correct selection of the relay nodes for the next stage of the messaging system was considered a major part of the VANETs, which control the reception of broadcast messages mainly in the wide coverage networks and the high density, This work introduces the hybrid relay node selection scheme as a new class, using the top features of existing message propagation protocols in the message accessibility, connection delay, etc. The assignment of the nodes of the next stage to the current propagation node is taken into account. An effort to increase VANET

performance if the density of nodes, traffic conditions, and speed scenarios are changed. Performance analyzes show that the new hybrid system enhances accessibility by up to 10%, based on current models.

## III. ASSUMPTIONS AND SYSTEM MODEL

### A. Assumptions

Based on the related literature [16][17], this paper assumes that the vehicles are equipped with a GPS device to be able to obtain their location and speed. Each vehicle contains Dedicated short-range communication (DSRC) and IEEE 802.11p wireless standard for routing within a range of 300 m. All vehicles are equipped with electronic devices that detect the presence of vehicles or roadside units (RSU) nearby inside a specific radius and determine the distance to them. Vehicles contain memory storage to store information about the received and sent messages.

### B. Simulation setup

To model the data dissemination in a highway environment, simulation experiments are implemented to observe the effect of the number of vehicles, density flow, the length of the radius over which the node itself sits in the cluster head, as well as the number of clusters formed along the road. Table I shows the most utilized parameters for simulation setup in this paper.

The extent of contact between vehicles set is set to 300 m, the extent of contact between the RSU and the vehicles is set to 1000 m, the rate of acceleration of the vehicle is 0.0075, the rate of deceleration of the vehicle is 0.025, and the number of RSUs are 2.

TABLE I.: SIMULATION PARAMETER

parameter	value
Flow Density	4 - 2 vehicle/second
Simulation time	1 min
Road Distance	12 km
Number of lanes	3
Number of vehicles	40 - 80
Model Simulator	Net Logo 6.2.0
Road Topology	High way
Maximum Speed	100 km/h
Channel type	Wireless
Coverage area	300 m
Radio standard	IEEE 802.11p

### C. System states

As mentioned, that VANET has a frequent changing topology, the states of this system can vary in each period. Varying is in the total number of vehicles and their speeds, traffic density, and clustering.

#### IV. THE FRAMEWORK OF PROPOSED DATA DISSEMINATION ALGORITHM

##### A. Cluster schema

Clustering technology is used to handle VANET features, and to ensure efficient data transfer. It is to create a hierarchical network structure. Grouped the vehicles into several collections. Driving directions, the vehicle's density on the road, and the location of the vehicles represent the important factors in each clustering process [13]. In this paper, a certain variable R-of-CH (range of CH) is adopted as a measure to determine the search range that the vehicle uses for the presence of cluster head (CH) or not. It can declare itself as a cluster member (CM) if it finds a CH and associates with the CH nearest to it or as a CH if it does not find it and associate with the CMs located in its range.

##### B. Methodology

Figure 2 shows the main steps in implementing this paper.

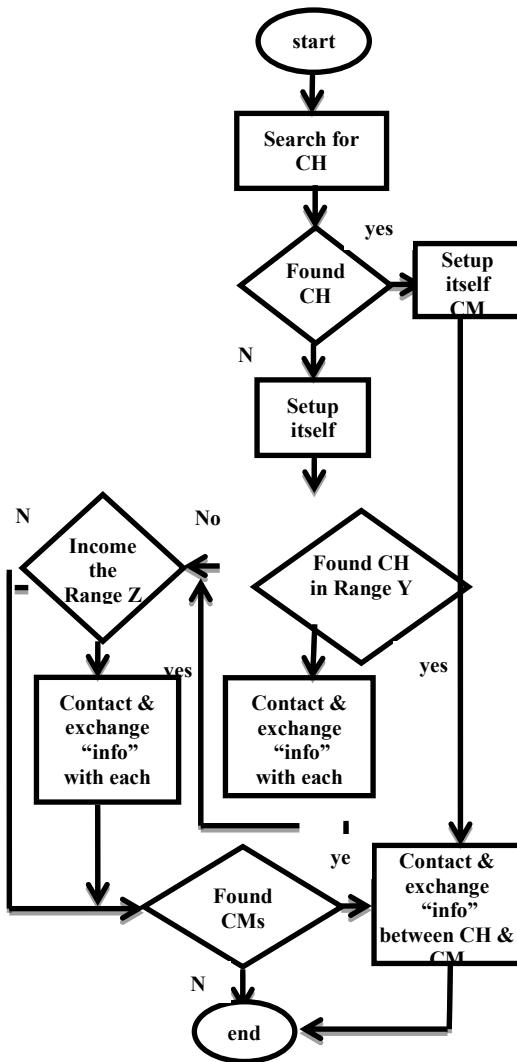


Fig. 2. The flow chart of the proposed approach.

##### C. Information transfer process

This paper suggests that the vehicles will send the message as soon as they get it. This process is also depending on the type of node (vehicle) and its location.

some case studies on a restricted highway are presented, to clarify what this work leads to. There are three types of application cases. Case 1, CM (with blue color) sends the message to its CH (with red color), case 2, the CH sends the message to all CMs that associated with it, and case 3, the CH sends the message to RSU and other associated CHs.

- **Case 1:** certain CM vehicle is suggested to detect a new event. It will send a message to its CH. Fig. 3 shows the process of data disseminating from one vehicle to another as a NetLogo snapshot. These data dissemination steps are as follows:

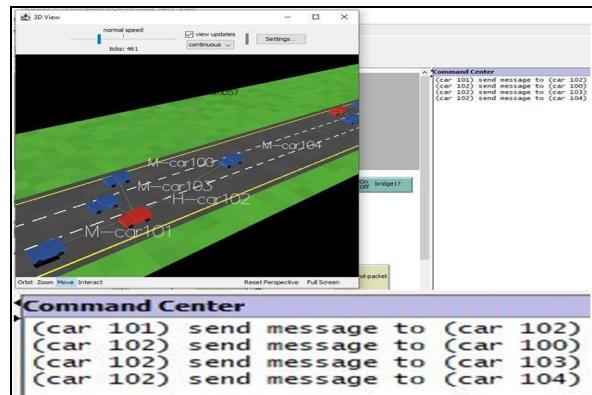


Fig. 3. Data dissemination for case 1.

- Step 1:** CM car (101) creates the message and adding its ID to his message list.
- Step 2:** CM sends the message to its associated CH car (102).
- Step 3:** CH receives the message by adding its ID to the message list.
- Step 4:** CH sends the message to (100,103,104) whose all CMs cars haven't received the message until that moment.

- **Case 2:** Another case is suggested in fig. 4, where the CH vehicle (number 19 as an example) detected the existence of a new event and send messages to its cluster members (vehicles 20, 21, and 22). The following steps represent this process sequence.

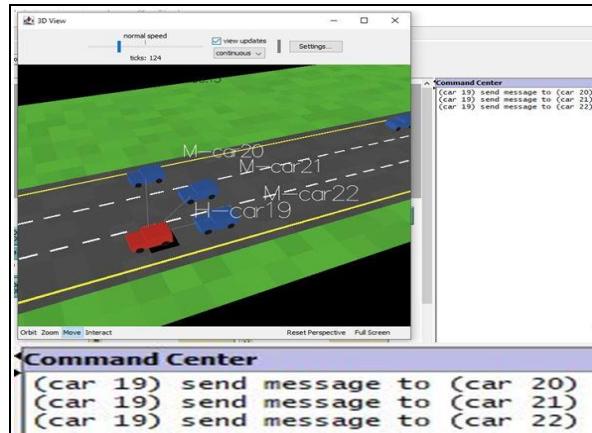


Fig. 4. Data dissemination for case 2.

- Step 1:** CH creates the message and announces it.
- Step 2:** CH sends the message to all associated CMs for which there is no message ID in their message list.
- Step 3:** All CMs add the message ID and announced that the message has been received.

- **Case 3:** Fig. 5, shows a CH vehicle linked to another CHs located within the range of RSU as follows:

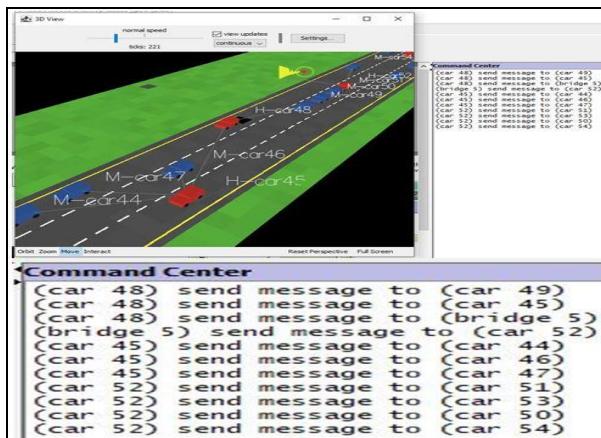


Fig. 5. Data dissemination for case 3.

- Step 1:** CH sends the message to all CMs that are associated with it and in range CHs. Each CH will send the message to all its CMs, other in range CHs.
- Step 2:** CH sends the message to the RSU within range, which in turn sends the message to all other CHs who associated with it and did not receive the message until now.

## V. RESULTS ANALYSIS AND DISCUSSION

NetLogo offers a good facility to observe the updating of each VANET variable with time in a visual manner.

To test this approach, several scenarios are created to see how the final results would be affected by the gradual changes of variables (number of vehicles, flow density, and distance between two CHs (R-of-CH)). These scenarios are as follows:

### A. Different results for different values

**1) Scenario 1:** in this scenario, 40 vehicles are generated, the density flow of vehicles on the road is suggested to be 4 vehicles/second, the R-of-CH is 150 m. The simulation results After 60 seconds are collected. Samples of these results are presented as continuous plots in fig. 6. In part (a) of fig. 6, the number of the generated links during the simulation time is found to be 45 links. In (b) the frequency of the resend messages between vehicles is shown, in this scenario the re-sent message is only one. Part (c) shows the maximum, minimum, and medium speed of vehicles, (d) shows the number each of Cluster Head vehicles is 10, out-layer vehicles are 2 and vehicles involved in the message transfer is 11, and (e) shows the number of delivery-rate is 40 that means the message was received by all vehicles and lost messages is 0. Fig. 7 shows Additional results and

details of vehicles whose become CH or out-layer vehicle and the number of CM associated with each CH as an example:

(car 6): " is cluster-head: 1 number of cluster-member: 1 >> [(car 7)]"

(car 9): " is out-layer : 1 number of cluster-member: 0 >> []"

(car 13): " is cluster-head: 2 number of cluster-member: 5 >> [(car 10) (car 11) (car 16) (car 12)]".

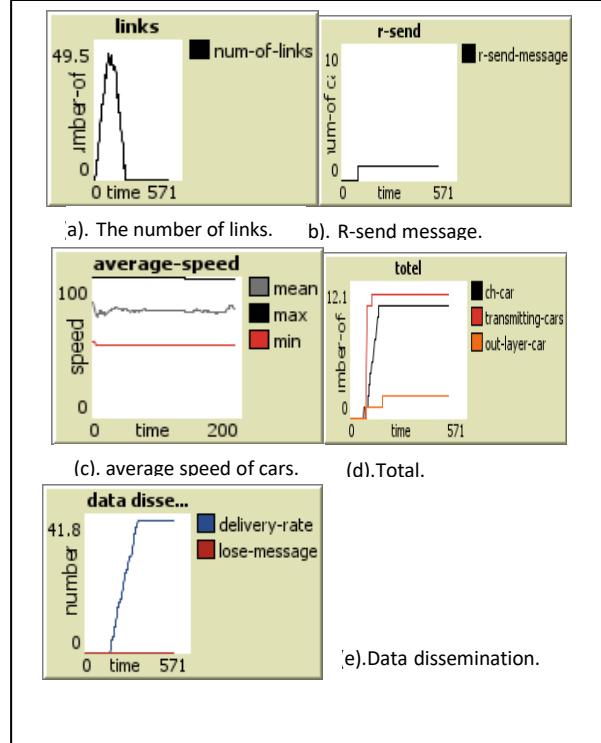


Fig. 6. a snapshot of the resulted performance metrics for scenario 1

```
(car 6): " is cluster-head : 1 number of cluster-member : 1 >> [(car 7)]"
(car 9): " is out-layer : 1 number of cluster-member : 0 >> []
(car 13): " is cluster-head : 2 number of cluster-member : 5 >> [(car 10) (car 11) (car 16) (car 12)]"
(car 19): " is cluster-head : 3 number of cluster-member : 2 >> [(car 20) (car 14)]"
(car 15): " is cluster-head : 4 number of cluster-member : 2 >> [(car 21) (car 23)]"
(car 24): " is cluster-head : 5 number of cluster-member : 4 >> [(car 22) (car 25) (car 18) (car 17)]"
(car 30): " is cluster-head : 6 number of cluster-member : 1 >> [(car 26)]"
(car 29): " is cluster-head : 7 number of cluster-member : 3 >> [(car 28) (car 31) (car 27)]"
(car 33): " is cluster-head : 8 number of cluster-member : 3 >> [(car 35) (car 32) (car 34)]"
(car 36): " is cluster-head : 9 number of cluster-member : 2 >> [(car 37) (car 38)]"
(car 40): " is cluster-head : 10 number of cluster-member : 5 >> [(car 39) (car 41) (car 42) (car 43) (car 44)]"
(car 45): " is out-layer : 2 number of cluster-member : 0 >> []"
```

That result for: 60 second

The number of cars entered into the model is : 40

at: 0:00:00 time has started

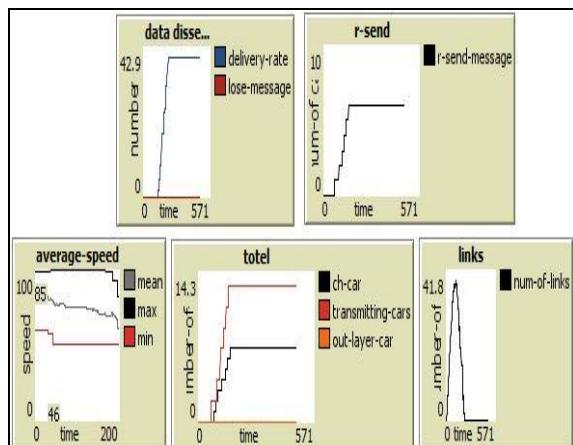
at: 0:01:00 time has ended

the number of cluster-head : 10      the number of out-layer : 2

Fig 7. detailed result for scenario 1

**2) Scenario 2:** in this scenario, 40 vehicles are generated, the density flow of vehicles on the road is 4 vehicle/second, the R-of-CH is 300 m. the collected results after 60 seconds of the simulation are presented in fig. 8. In part (a) seven

clusters are found, the number of out-layer is 0, and the number of links is 39, and the message has been re-sent 6 times, the message has reached all vehicles in the scenario. Fig. 8 shows in part (b) additional detailed results of the clusters formed and the number of CM in each group.



(a): Resulted performance metrics

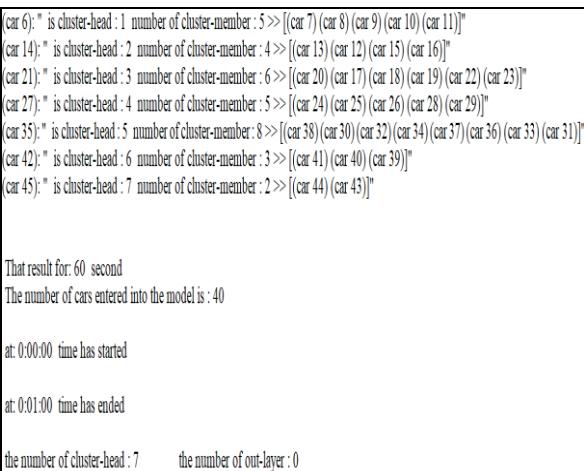
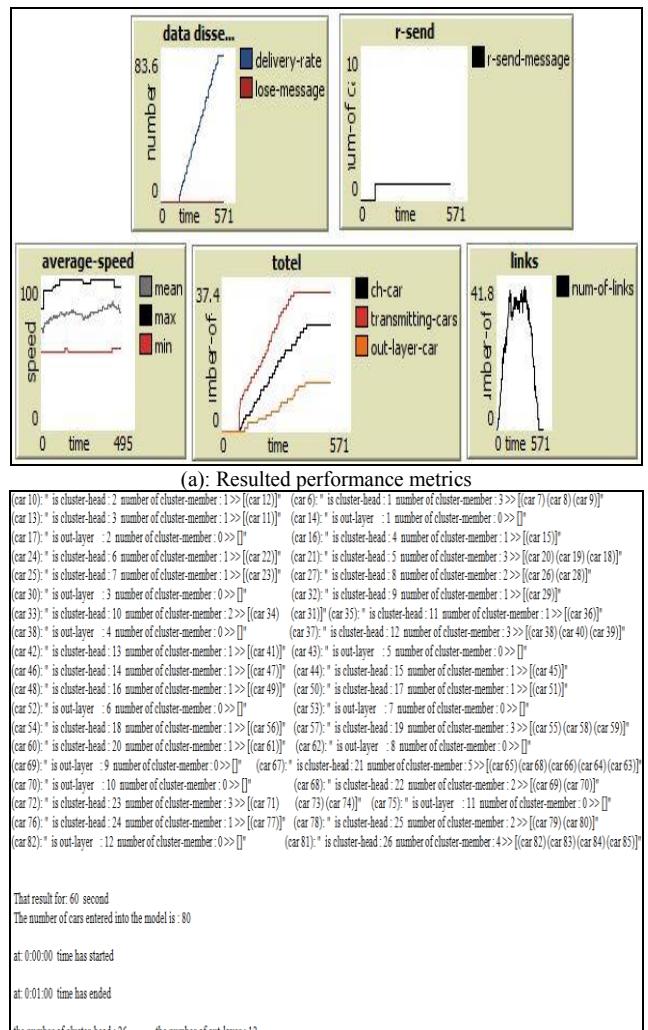


Fig 8. Snapshot of the results for scenario 2.

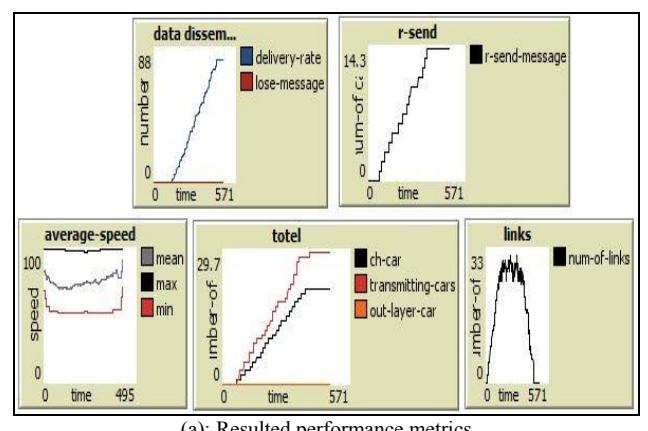
**3) Scenario 3:** in this scenario, 80 vehicles are generated, the density flow of vehicles on the road is 2 vehicles/second, the R-of-CH is 150 m. Simulation results show that the number of clusters is 26 and the number of out-layer is 12 and the number of generated links is 39, and the message was not re-sent only once, through which the message reached all vehicles in the scenario. This is indicated in fig. 9 (a) and (b).



(b): detailed results

Fig 9. a snapshot of the results for scenario 3.

**4) Scenario 4:** in this scenario, 80 vehicles are generated, the density flow of vehicles on the road is 2 vehicles/second, the R-of-CH is 300 m. The number of clusters is found to be 21 and the number of out-layer is 0 and generated 32 links, the message has been re-sent 14 times and in the end, the message has reached all vehicles in the scenario. This is indicated in Fig. 10 (a) and (b).



(a): Resulted performance metrics

```

(car 6): " is cluster-head : 1 number of cluster-member : 1 >> [[car 7]]"
(car 9): " is cluster-head : 2 number of cluster-member : 2 >> [[car 8) (car 10]]"
(car 12): " is cluster-head : 3 number of cluster-member : 4 >> [[car 11) (car 13) (car 14) (car 15]]"
(car 20): " is cluster-head : 4 number of cluster-member : 4 >> [[car 19) (car 17) (car 16) (car 18]]"
(car 23): " is cluster-head : 5 number of cluster-member : 1 >> [[car 21]]"
(car 24): " is cluster-head : 6 number of cluster-member : 4 >> [[car 22) (car 25) (car 26) (car 27]]"
(car 30): " is cluster-head : 7 number of cluster-member : 3 >> [[car 28) (car 29) (car 31]]"
(car 34): " is cluster-head : 8 number of cluster-member : 3 >> [[car 33) (car 32) (car 36]]"
(car 37): " is cluster-head : 9 number of cluster-member : 2 >> [[car 38) (car 35]]"
(car 40): " is cluster-head : 10 number of cluster-member : 2 >> [[car 39) (car 41]]"
(car 43): " is cluster-head : 11 number of cluster-member : 6 >> [[car 46) (car 45) (car 44) (car 42) (car 48) (car 47]]"
(car 49): " is cluster-head : 12 number of cluster-member : 1 >> [[car 50]]"
(car 51): " is cluster-head : 13 number of cluster-member : 2 >> [[car 52) (car 53]]"
(car 55): " is cluster-head : 14 number of cluster-member : 2 >> [[car 54) (car 56]]"
(car 58): " is cluster-head : 15 number of cluster-member : 4 >> [[car 57) (car 59) (car 60) (car 61]]"
(car 63): " is cluster-head : 16 number of cluster-member : 2 >> [[car 62) (car 64]]"
(car 65): " is cluster-head : 17 number of cluster-member : 1 >> [[car 66]]"
(car 67): " is cluster-head : 18 number of cluster-member : 4 >> [[car 68) (car 70) (car 72) (car 69]]"
(car 75): " is cluster-head : 19 number of cluster-member : 4 >> [[car 74) (car 76) (car 73) (car 71]]"
(car 78): " is cluster-head : 20 number of cluster-member : 1 >> [[car 77]]"
(car 81): " is cluster-head : 21 number of cluster-member : 6 >> [[car 79) (car 80) (car 82) (car 83) (car 85) (car 84]]"

```

The result for: 60 second  
The number of cars entered into the model is : 80  
at 0:00:00 time has started  
at 0:01:00 time has ended  
the number of cluster-head : 21 the number of out-layer : 0

(b): detailed results

Fig 10. Snapshot of the results for scenario 4.

### B. Scenarios analysis

The system behaves as planned as the process of controlling the value of the R-of-CH variable directly contributes to network stability and improved data dissemination. Since the number of cluster head vehicles is inversely proportional to the value of the R-of-CH variable, while the number of times the message is retransmitted is directly proportional to it. This is what is observed in the second and fourth scenarios. The third and fourth scenarios can also be compared. Whereby doubling the value of the R-of-CH variable in the system, it appears that the number of out-layer vehicles decreased from (12) to (0) with a decrease in the count of the links generation during the simulation period.

## VI. CONCLUSIONS

In this paper, a dynamic approach is proposed to improve data dissemination using the clustering method. The system is implemented in Net Logo 6.0.2. Where a new approach is suggested by adopting a variable (R-of-CH) to be used as a measure to search for the availability of a CH in the surrounding area of each vehicle. The aim is to properly distribute the CHs along the road. A unique message identifier and a message list in each vehicle will prevent receiving the same message from more than one source to alleviate the problems of message redundancy. The performance of different variables is observed and compared. Increasing the value of R-of-CH will reduce the number of created clusters. Decreasing the number of generated clusters is a good indication of cluster stability. Some out-layer vehicles (not clustered) appear but they do not have a significant influence on the process of data dissemination. When the value of the R-of-CH is decreased, the message is sent only once. With the presence of several

out-layers vehicles, however, the messages sent reach all vehicles in all scenarios.

## REFERENCES

- [1] A. Botta, W. De Donato, V. Persico, and A. Pescapé, “Integration of Cloud computing and Internet of Things: A survey,” *Futur. Gener. Comput. Syst.*, vol. 56, pp. 684–700, 2016, doi: 10.1016/j.future.2015.09.021.
- [2] M. A. Mahdi and S. T. Hasson, “Grouping vehicles in Vehicular Social Networks,” *Kurdistan J. Appl. Res.*, vol. 2, no. 3, pp. 218–225, 2017, doi: 10.24017/science.2017.3.42.
- [3] A. R. Deshmukh, S. A. Dhawale, and S. S. Dorle, “Analysis of Cluster Based Routing Protocol (CBRP) for Vehicular Adhoc Network (VANet) in Real Geographic Scenario,” pp. 1–5, 2020, doi: 10.1109/conecct50063.2020.9198669.
- [4] M. Chaqfeh, A. Lakas, and I. Jawhar, “A survey on data dissemination in vehicular ad hoc networks,” *Veh. Commun.*, vol. 1, no. 4, pp. 214–225, 2014, doi: 10.1016/j.vehcom.2014.09.001.
- [5] S. T. Hasson and Z. Y. Hasan, “Simulating road modeling approach’s in Vanet environment using net logo,” *Res. J. Appl. Sci.*, vol. 11, no. 10, pp. 1130–1136, 2016, doi: 10.3923/rjasci.2016.1130.1136.
- [6] I. Achour, T. Bejaoui, A. Busson, and S. Tabbane, “SEAD: A simple and efficient adaptive data dissemination protocol in vehicular ad-hoc networks,” *Wirel. Networks*, vol. 22, no. 5, pp. 1673–1683, 2016, doi: 10.1007/s11276-015-1050-9.
- [7] M. Lee and T. Atkison, “VANET applications: Past, present, and future,” *Veh. Commun.*, p. 100310, 2020, doi: 10.1016/j.vehcom.2020.100310.
- [8] M. A. Niazi, *Introduction to the modeling and analysis of complex systems: a review*, vol. 4, no. 1. 2016.
- [9] O. T. Abdulhae, M. Ismail, M. A. Altahrawi, and M. A. Saad, “Data dissemination of vehicular Ad-Hoc network in highway scenario,” *2019 14th IEEE Malaysia Int. Conf. Commun. Emerg. Technol. Icet 5G, MICC 2019*, no. December, pp. 25–30, 2019, doi: 10.1109/MICC48337.2019.9037571.
- [10] O. Senouci, S. Harous, and Z. Aliout, “Survey on vehicular ad hoc networks clustering algorithms: Overview, taxonomy, challenges, and open research issues,” *Int. J. Commun. Syst.*, vol. 33, no. 11, pp. 1–21, 2020, doi: 10.1002/dac.4402.
- [11] O. Urmonov and H. Kim, “A multi-hop data dissemination algorithm for vehicular communication,” *Computers*, vol. 9, no. 2, 2020, doi: 10.3390/computers9020025.
- [12] S. S. Shah, A. W. Malik, A. U. Rahman, S. Iqbal, and S. U. Khan, “Time Barrier-Based Emergency Message Dissemination in Vehicular Ad-hoc Networks,” *IEEE Access*, vol. 7, pp. 16494–16503, 2019, doi: 10.1109/ACCESS.2019.2895114.
- [13] M. Chahal and S. Harit, “A stable and reliable data dissemination scheme based on intelligent forwarding in VANETs,” *Int. J. Commun. Syst.*, vol. 32, no. 3, 2019, doi: 10.1002/dac.3869.
- [14] D. Wu, H. Li, H. Zhao, and J. Zhang, “A Data Distribution Algorithm Based on Neighbors’ Locations in VANET,” *Procedia Comput. Sci.*, vol. 174, no. 2019, pp. 530–534, 2020, doi: 10.1016/j.procs.2020.06.121.
- [15] O. Rehman and M. Ould-Khaoua, “A hybrid relay node selection scheme for message dissemination in VANETs,” *Futur. Gener. Comput. Syst.*, vol. 93, pp. 1–17, 2019, doi: 10.1016/j.future.2018.10.042.
- [16] A. Reyes, C. Barrado, M. López, and C. Excelente, “Vehicle density in VANET applications,” *J. Ambient Intell. Smart Environ.*, vol. 6, no. 4, pp. 469–481, 2014, doi: 10.3233/AIS-140271.
- [17] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, and B. Zhou, “A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs,” *Veh. Commun.*, vol. 13, no. May, pp. 78–88, 2018, doi: 10.1016/j.vehcom.2018.05.002.