Tan, K., Le Kernec, J., Imran, M. and Bremner, D. (2019) Clustering Algorithm in Vehicular Ad-hoc Networks: A Brief Summary. In: UK-China Emerging Technologies Conference, Glasgow, UK, 21-22 Aug 2019, ISBN 9781728127972 (doi:10.1109/UCET.2019.8881833)

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/193780/

Deposited on: 23 August 2019

Clustering Algorithm in Vehicular Ad-hoc Networks: A Brief Summary

Kang Tan
School of Engineering
University of Glasgow
Chengdu, China
2109534T@student.gla.ac.uk

Julien Le Kernec
James Watt School of Engineering
University of Glasgow
Glasgow, UK
Julien.LeKernec@glasgow.ac.uk

Muhammad Imran
James Watt School of Engineering
University of Glasgow
Glasgow, UK
Muhammad.Imran@glasgow.ac.uk

Duncan Bremner
James Watt School of Engineering
University of Glasgow
Glasgow, UK
Duncan.Bremner@glasgow.ac.uk

Abstract—**An Intelligent Transportation System (ITS) application requires vehicles to be connected to each other and to roadside units to share information, thus reducing fatalities and improving traffic congestion. Vehicular Ad hoc Networks (VANETs) is one of the main forms of network designed for ITS in which information is broadcasted amongst vehicular nodes. However, the broadcast reliability in VANETs face a number of challenges - dynamic routing being one of the major issues. Clustering, a technique used to group nodes based on certain criteria, has been suggested as a solution to this problem. This paper gives a summary of the core criteria of some of the clustering algorithms issues along with a performance comparison and a development evolution roadmap, in an attempt to understand and differentiate different aspects of the current research and suggest future research insights.**

Keywords— VANETs, Clustering Algorithm, Single-hop, Multi-hop, ITS

I. Introduction

The Intelligent Transportation System (ITS) [1] aims to increase on-road safety and improve traffic congestion, which is a crucial part of future smart cities [2]. To enable ITS, vehicles must connect to each other and also to roadside units. However, the increasing number of vehicles and the associated high mobility make it difficult to implement a traditional cellular network in this application. Therefore, Mobile Ad hoc Networks (MANETs) [3] have become the preferred option for its lower cost and flexibility. This approach also exploits the benefits of Vehicular Ad hoc Networks (VANETs) [4] because of the particular characteristics of on-road vehicles and the highway environment.

In VANETs, a vehicle in a formed ad hoc network is regarded as a node. Every node in the ITS plays the role of the sender, receiver, and router [5] to broadcast messages such as recently occurred accidents, storm, and traffic congestion, etc. Because of nodes' high mobility, VANETs have highly probable network partitions, leading to unguaranteed end-to-end communications [6]. Dynamic routing has been one of the most important issues in VANETs, and intermittent connection leads to potential severe pack loss, influencing traffic safety further. In addition to mobility, the traditional flat structured network also faces scalability problems as VANETs usually have large network size. A proactive routing protocol tends to have exponentially overhead with respect to node number; while a reactive one faces flooding of routing request and other messages causing serious resource wastage and high delay. For all these reasons, a flat structure cannot perform well in VANETs, and a hierarchical network structure is needed for the area [7].

On the other hand, a cluster structure is a hierarchical structure in VANETs which groups nearby vehicles and delivers information to the group as a whole and can reduce flooding, resource waste, and provide useful on-road information etc. Clustering algorithms have been improved and developed over the years to increase connection stability, clustering efficiency, and decrease clustering overhead, etc. This paper compares and contrasts a selection of clustering algorithm, and tries to draw insights for future research into VANET clustering. The paper is organized as follows: A selection of noteworthy clustering algorithms are introduced in Section 2; their performances are compared in Section 3, and a conclusion based on the results is given in Section 4.

II. Clustering Algorithms

Examining cluster topology in more detail, clustering algorithms are classified into two broad types: single-hop, and multi-hop algorithms, where the hop distance is defined as the number of times a data packet transmission must be rerouted after transmission until it reaches the receiver. In the VANETs application, a single-hop algorithm forms a cluster in which the cluster head (CH) directly connects to all cluster members (CMs) and transmits data. In a cluster formed by a multi-hop algorithm, a CH does not need to directly connect to all CMs but can transmit data to remote CMs through intermediate CMs within the predefined maximum hop distance (one, two or more hops). A simple data transmission model for two different hop distances is shown in Fig. 1.

Until now, several classic clustering algorithms [9, 10, 11, 12] are limited to single hops while recently, research has focused on the multi-hop approach.
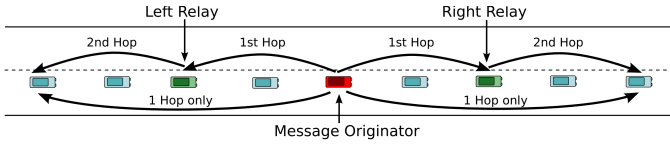
Fig. 1. Data transmission model for different hop distance in a simple highway scenario: Single-hop and Double-hop. Source: [8]

A. Single-hop Algorithms

The Lowest-ID algorithm [9] is a classic clustering algorithm originally developed for MANETs. It works within a defined area where every node is assigned a randomly generated ID, and each node periodically broadcasts a list of its immediate neighbors. A node which only receives broadcasts from node IDs higher than itself is then selected as the cluster head (CH) and broadcasts messages periodically to nearby cluster members (CM). A CM able to hear more than one CH which is then defined as a gateway node between the two clusters. Examining this topology, it is clear that there is a major drawback; specifically since the ID is randomly assigned, the algorithm cannot consider any nodal information that may be useful (such as vehicle mobility) to other cluster members. The topology factor is not taken into account during ID distribution leading to suboptimal clustering, thus causing lower energy efficiency for the CH and lower connection quality for distant CMs. This is due to the randomly distributed IDs ensure neither low average transmission distance between the CH and CMs after formation nor consider other physical factors such as fading effect caused by obstacles.

The Highest Connectivity Cluster Algorithm [10], is another clustering algorithm which assigns IDs and broadcasts the ID list and gateway node selection in the same way as the Lowest-ID method. However, it also considers the available neighbor numbers of a node into account and uses that information to select the node with the largest neighbor count as CH. Using this criterion, a CH will change less frequently, but the throughput of the cluster may be compromised due to the lack of cluster size specification.

As the above clustering algorithms both have particular flaws, an approach utilizing more parameters is preferred to improve clustering performance. MOBIC [11] is an early mobility-based [12] algorithm which is similar to the Lowest-ID but uses the received hello messages' received signal strength to calculate relative mobility metrics for CH selection. When two nodes have the same relative mobility, it reverts to the Lowest-ID method. The Weighted Clustering Algorithm (WCA) [13] is another early clustering algorithm that takes a number of mobility parameters into consideration, including the transmission power, node mobility, and node battery power. Each node then calculates its weighted index based on these parameters, and the node with the minimum weight is elected as the CH. The WCA also limits the cluster size to achieve load balancing by having a predefined threshold value for the CH to handle. However, these two algorithms also have limitations; MOBIC demands moderate to high transmission range to perform well whereas WCA relies on a global minima concept demanding that all nodes have the weighting indices of all other nodes before proceeding, leading to longer processing time and waste of computation resource [14].

In 2017, a dynamic mobility-based clustering scheme (DMCS) for VANETs [15] was developed for improved clustering performance. The algorithm introduced a new metric called "safe transmission distance" to limit the size of a cluster. This distance threshold value D was defined as less or equal to the CH's transmission range TR, i.e. $D \leq TR$. The cluster size L was then defined with the $L \leq 2D$ restriction, resulting in more resilient message deliveries between the CH and CMs. At the initiation of cluster formation, a temporary CH (CH_t) is identified to assist in cluster formation. After the initial establishment of the cluster, the CH_t can be promoted to a CH or relegated to a CM after the process. The decision process to establish a CH demand that the relative distances of all members on the beacon list are within L. According to the design of DMCS, all CMs of a cluster should be in the same moving direction as the CH. Gateway nodes are selected based on the furthest relative distance to the CH and the Link Lifetime (LLT) [16] calculation is used when two or more nodes have the same relative distance. The node with a larger LLT value in this situation becomes the gateway node. DMCS also introduced cluster merging, which occurs when the two clusters' merged size $L_{merge} \leq 2D$ and hence a new CH is then selected thereby decreasing resource wastage and transmission collisions caused by two or more nodes transmitting messages simultaneously through the same channel. However, the DMCS assumes the use of GPS service and certain computing power of the vehicle, which can be a drawback depending on application scenarios as older-model vehicles may not have the computation ability or onboard GPS units while GPS service is not available in environments such as tunnels and mountain areas.

B. Multi-hop Algorithms

The N-Hop clustering algorithm [17] was the earliest multi-hop VANET clustering algorithm. Introduced in 2011, its design allowed vehicle nodes to broadcast beacon messages periodically. When a node received two consecutive beacon messages, it calculated its relative mobility with its N-hop neighbour nodes based on the message delay. This design also defines a cluster's diameter measured in hop number to be less than 2N. The algorithm then calculates the aggregate relative mobility metric using these relative mobility metrics and selects a CH node based on their lowest aggregate mobility. Other nodes will become a CM of this cluster when they receive the messages broadcasted from the CH. Because of its design, the N-Hop algorithm requires many control messages within the cluster resulting in a reduction in the overall cluster efficiency.

Introduced in 2013, the Vehicular Multi-hop algorithm for Stable Clustering (VMaSC) [18] is the first mobility-based multi-hop clustering algorithm simulated with realistic vehicle mobility data generated by the Simulation of Urban Mobility (SUMO) [19]. Its design goal was to minimize the number of CHs and maximize cluster duration while decreasing clustering resource cost and overhead. VMaSC introduced a periodically updated local knowledge base containing vehicles' parameters such as direction, current speed, etc. Relative average speed was then calculated for vehicles moving in the same direction after each update. Vehicles only checked this knowledge base for other vehicles moving in the same direction to increase cluster duration, and a vehicle node with the smallest average relative speed was selected as CH after comparing its speed with the

existing CHs' to ensure minimum CH number thus reduce overall resource wastage. However, similar to the DMCS, the VMaSC requires the support of GPS or similar location service to obtain mobility data, which may not always be available depending on application scenarios.

With the aim of improving the disadvantages of N-Hop and VMaSC, a distributed multi-hop clustering algorithm for VANETs based on neighborhood follow (DMCNF) [20] was developed. This algorithm used a cluster model based on the one-hop neighborhood follow. In this model, vehicles only need to follow targeted stable one-hop neighbor to transmit information directly while owning the same CH as this neighbor. A node with more followers and the smallest relative average mobility would be selected as a CH based on the neighborhood strategy. The DMCNF algorithm calculates relative mobility the same way as N-Hop and does not depend on GPS services. These features make DMCNF capable of forming more stable clusters with greater flexibility, reduced redundant message generation, and broadcast and be more adaptive to application and environmental scenarios.

The Passive Multi-hop Clustering algorithm (PMC) [21] is a new multi-hop clustering algorithm introduced in 2018, aiming to improve the poor cluster reliability because of the DMCNF inter-node link reliability ignorance. PMC keeps a neighbor information table in VANET including a vehicle's ID, location-related information, number of followers, etc. It also contains a newly proposed 'priority neighbor following strategy' that selects the optimal neighbor to join that neighbor's cluster by considering node following degree, expected transmission count [22] and LLT. A node's following degree is the sum of its connected node number and the number of neighbor nodes on the same lane. The expected transmission count is defined by the reciprocal of the product of transmission and reception rate. Respectively, larger node following degree and LLT value, smaller expected transmission count leads to higher cluster stability. The CH selection rule of the PMC algorithm shares a similar concept as DMCNF while including the advantages of cluster merging methodology of DMCS.

An important factor is that after cluster head selection, it can introduce large impact on CMs if the CH changes its speed and moving direction, which may cause CH switching and cluster reformation. Though later algorithms such as DMCS and VMaSC take relative mobility into account, the relative mobility between vehicles can change rapidly, leading to unreliable nodes and unstable cluster. This kind of issues remains in VANETs clustering methods because of on-road vehicles' high mobility and dynamic topology nature, which requires further research.

III. PERFORMANCE COMPARISON

Several factors can be used to measure cluster stability:

- CH duration time: The time over which a vehicle is selected as a CH to until it switches to other roles (such as CM, or in the initial state). Longer CH duration indicates more cluster stability.

- CH change number: The number of CH changing to other roles within a certain time. Smaller CH change number indicates higher cluster stability.

- CM duration time: The time over which a vehicle joins and leaves a cluster. A stable cluster has long CM duration

- Cluster overhead: The ratio of control message packets with the total message packet during cluster formation and maintenance phase. Smaller overhead indicates higher cluster efficiency.

- Cluster number: The number of clusters given a specific area. Smaller cluster number indicates higher cluster efficiency.

These metrics can broadly describe the stability and efficiency of a clustering algorithm. We use them and cluster overhead to compare the performances of the cluster algorithms mentioned above.

Qualitatively, single-hop clustering offers simpler cluster structure and fast cluster formation, while relatively smaller cluster size it introduces leads to a larger cluster number and more isolated nodes unable to connect to a CH directly. These factors reduce clusters' information throughput and overall efficiency. Multi-hop clustering, on the other hand, makes clusters more stable with fewer isolated nodes introduced by not being able to connect to the CH in the single-hop case. However, the topology of cluster formation becomes more complicated because of node connection extension, thus leads to longer formation time and the requirement for more complex computation.

A. Single-hop Algorithms

We use simulation results from [15] which evaluates the performances of Lowest-ID, VMaSC in one-hop mode and DMCS using average CH duration, CM duration, and average CH change rate with respect to the maximum lane speed (vehicle velocity). Instead of using the raw values, the authors chose normalized values for these factors defined by the percentage of the total simulation time respectively. The results are shown in Fig. 2 and Fig. 3. The "Our scheme" in the figures' legend represents the DMCS and "LID" represents the Lowest-ID algorithms.

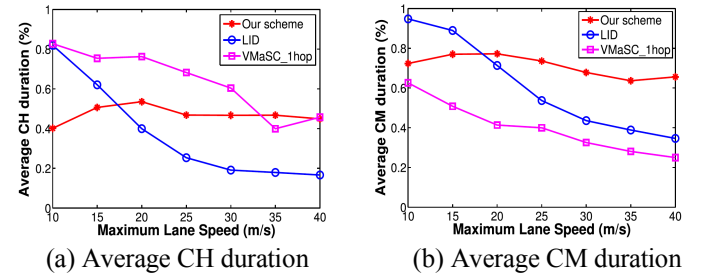(a) Average CH duration   (b) Average CM duration

Fig. 2. CH(CM) duration with respect to maximum lane speed. Source: [15]

According to Fig. 2, the average CH(CM) duration for Lowest-ID and VMaSc decreases with the maximum speed increment while the trend of DMCS seems to be overall stable with the increment happening in the lower speed part of the graph (smaller than 20m/s). VMaSC has the highest CH duration until the maximum speed reaches approximately 33m/s. DMCS becomes to have the highest CM duration after maximum speed rises above about 18m/s.

Based on Fig. 3, the CH change rate of Lowest-ID appears to have a close-linear relationship with maximum lane speed, and the slope of the curve is larger. However, the CH change rate of DMCS and VMaSC are in a similar trend, and they rise much slower. The CH change rate of DMCS is a little higher than that of VMaSC throughout the graph.
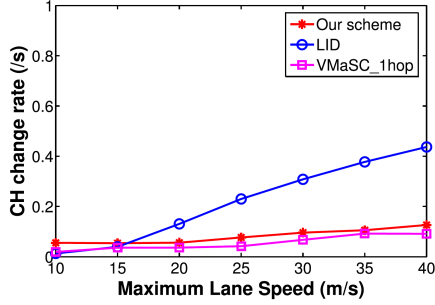

Fig. 3.   CH change rate with respect to maximum lane speed. Source: [15]

To sum up, DMCS has a more stable and overall good performance among the chosen group. It has an average CH(CM) duration and CH change rate around 50%, 70%, and 8% respectively. The Lowest-ID performs better at lower maximum speed values while VMaSC gains better performance in CH duration and change rate overall.

B. Multi-hop Algorithms

We use simulation results from [21] which presents and compares the average CH duration, average CM duration, average CH changes and clustering overhead ratio with respect to maximum vehicle velocity. The results are presented directly without normalization. The simulation contains three sets of trials with different transmission range (100m to 300m) and other settings fixed. We take the 300m data-set and present the results for example. The results are shown in Fig. 4 and Fig. 5.

Fig. 4(a) and Fig. 4(b) present the same decreasing trend of average CH(CM) duration with the maximum velocity increment for all four clustering algorithms. Similarly, Fig. 4(c) and Fig. 5 present a similar increasing trend of average CH changes and clustering overhead ratio with the maximum velocity increment. The four graphs show that all four clustering algorithms become less stable and efficient in higher maximum speed situations.

On the other hand, the relative positioning of curves gives the fact that DMCNF and PMC have approximately 20 and 35 sec longer CH(CM) duration and about 100-unit smaller average CH changes, which are significant differences. When inferring from the figures, the four curves can be classified into two groups with PMC and DMCNF forming a group with better data while N-Hop and VMaSC forming the other group with worse data. Within two groups, PMC has slightly better results compared with DMCNF, and VMaSC has better results compared to N-Hop.

Therefore, the presented results imply that DMCNF and PMC have made significant performance improvement compared with N-Hop and VMaSC while VMaSC performs slightly better than N-Hop in for every factor. The fact of using GPS and local knowledge base for a more optimal cluster formation may be the underlying reason for it. PMC, on the other hand, performs better than DMCNF in a similar way that VMaSC does to N-Hop. It may benefit from the PMC's updated priority neighborhood selection strategy. In addition, the two other sets of data presented in [21] show similar trends and relative positions, which further strengthen our conclusion.
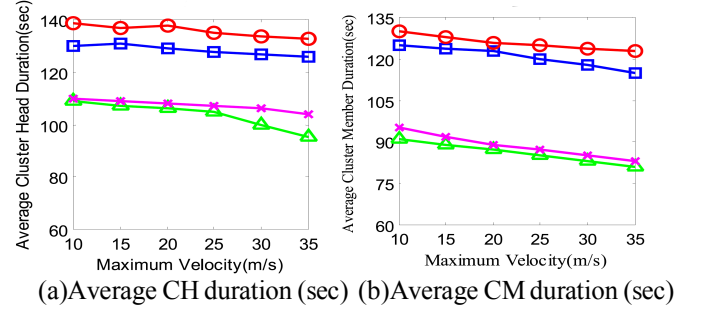
(a)Average CH duration (sec) (b)Average CM duration (sec)
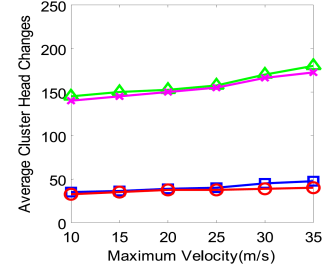
(c) Average CH changes

Fig. 4.   Average CH(CM) duration and CH changes with respect to maximum vehicle velocity. Source: [21]
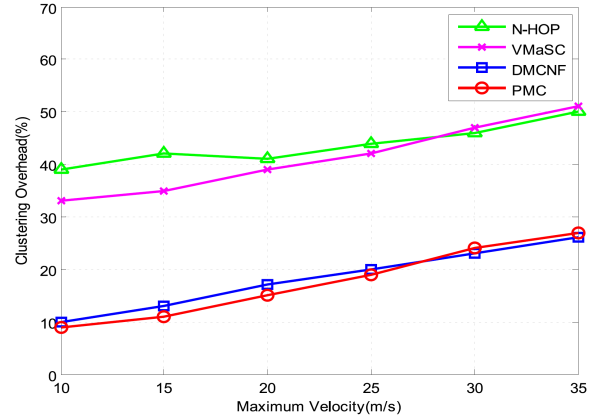

Fig. 5.   Cluster overhead ratio with respect to maximum vehicle velocity. Source: [21]

IV. Conclusion

We introduced a topology classification of clustering algorithms in VANETs, single-hop, and multi-hop clustering. Based on this classification, we also summarized some classic and recent clustering algorithms' core design features in both types with some performance comparison.

Qualitative inferences based on the summarized information show that single-hop clustering gives simple-structured clusters and fast cluster formation while cluster efficiency can be lower with more potential isolated nodes and a larger cluster number. While multi-hop clustering produces larger and fewer clusters

leading to fewer potential isolated nodes and higher intra-cluster connectivity and cluster stability, the more complicated-structured clusters slow down the cluster formation speed and require more computation power.

According to the simulation analysis available, both recent single-hop and multi-hop clustering algorithms (DMCS and PMC) show significant performance improvement compared to their predecessors. These results also prove that higher vehicle speeds can reduce cluster stability and efficiency.

Based on additional supplementary materials on VANETs, there appears to be significant research opportunity for further work on this subject including the potential fusion of clustering algorithms. Swarm intelligence shows particular promise for future VANETs application with its ability to resolve the spanning tree algorithm [23] for more optimal routing. Multi-cast routing algorithms have been designed [24] to operate within the upcoming 5G protocol using existing VANET communication protocols. This area also appears to offer a promising future research direction through the combination of LTE and the current protocol's success in raising clustering stability and efficiency [25].

REFERENCES

[1] Figueiredo, L., Jesus, I., Machado, J. T., Ferreira, J. R., & De Carvalho, J. M. (2001). Towards the development of intelligent transportation systems. In ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585) (pp. 1206-1211). IEEE.

[2] Nam, T., & Pardo, T. A. (2011, June). Conceptualizing smart city with dimensions of technology, people, and institutions. In Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times (pp. 282-291). ACM.

[3] Macker, J. (1999). Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations.

[4] Raya, M., & Hubaux, J. P. (2007). Securing vehicular ad hoc networks. *Journal of computer security*, *15*(1), 39-68.

[5] Zeadally, S., Hunt, R., Chen, Y. S., Irwin, A., & Hassan, A. (2012). Vehicular ad hoc networks (VANETS): status, results, and challenges. *Telecommunication Systems*, *50*(4), 217-241.

[6] Toor, Y., Muhlethaler, P., Laouiti, A., & De La Fortelle, A. (2008). Vehicle ad hoc networks: Applications and related technical issues. *IEEE communications surveys & tutorials*, *10*(3), 74-88.

[7] Yu, J. Y., & Chong, P. H. J. (2005). A survey of clustering schemes for mobile ad hoc networks. *IEEE Communications Surveys & Tutorials*, *7*(1), 32-48.

[8] Mittag, J., Thomas, F., Härri, J., & Hartenstein, H. (2009, September). A comparison of single-and multi-hop beaconing in VANETs. In *Proceedings of the sixth ACM international workshop on VehiculAr InterNETworking* (pp. 69-78). ACM.

[9] Ephremides, Anthony, Jeffrey E. Wieselthier, and Dennis J. Baker. "A design concept for reliable mobile radio networks with frequency hopping signaling." Proceedings of the IEEE75.1 (1987): 56-73.

[10] Gerla, Mario, and Jack Tzu-Chieh Tsai. "Multicluster, mobile, multimedia radio network." *Wireless networks* 1.3 (1995): 255-265.

[11] Basu, P., Khan, N., & Little, T. D. (2001, April). A mobility based metric for clustering in mobile ad hoc networks. In *Proceedings 21st International Conference on Distributed Computing Systems Workshops* (pp. 413-418). IEEE.

[12] Hong, X., Gerla, M., Pei, G., & Chiang, C. C. (1999, August). A group mobility model for ad hoc wireless networks. In *Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems* (Vol. 53, p. 60).

[13] Chatterjee, M., Das, S. K., & Turgut, D. (2002). WCA: A weighted clustering algorithm for mobile ad hoc networks. *Cluster computing*, *5*(2), 193-204.

[14] AlMheiri, S. M., & AlQamzi, H. S. (2015, February). MANETs and VANETs clustering algorithms: A survey. In *2015 IEEE 8th GCC Conference & Exhibition* (pp. 1-6). IEEE.

[15] Ren, M., Khoukhi, L., Labiod, H., Zhang, J., & Veque, V. (2017). A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (VANETs). *Vehicular Communications*, *9*, 233-241.

[16] S.-S. Wang, Y.-S. Lin, Passcar: a passive clustering aided routing protocol for vehicular ad hoc networks, Comput. Commun. 36 (2) (2013) 170–179.

[17] Zhang, Z., Boukerche, A., & Pazzi, R. (2011, October). A novel multi-hop clustering scheme for vehicular ad-hoc networks. In *Proceedings of the 9th ACM international symposium on Mobility management and wireless access* (pp. 19-26). ACM.

[18] Ucar, S., Ergen, S. C., & Ozkasap, O. (2013, April). VMaSC: Vehicular multi-hop algorithm for stable clustering in vehicular ad hoc networks. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 2381-2386). IEEE.

[19] Simulation of urban mobility, sumo @ONLINE. [Online]. Available: http://http://sumo.sourceforge.net/

[20] Chen, Y., Fang, M., Shi, S., Guo, W., & Zheng, X. (2015). Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow. *Eurasip journal on Wireless communications and networking*, *2015*(1), 98.

[21] Zhang, D., Ge, H., Zhang, T., Cui, Y. Y., Liu, X., & Mao, G. (2018). New multi-hop clustering algorithm for vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, (99), 1-14.

[22] N. Javaid, A. Bibi, and K. Djouani, "Interference and bandwidth adjusted ETX in wireless multi-hop networks," in *Proc. GLOBECOM Workshops*, Dec. 2010, pp. 1638–1643.

[23] Zhang, X., & Zhang, X. (2017). A binary artificial bee colony algorithm for constructing spanning trees in vehicular ad hoc networks. *Ad Hoc Networks*, *58*, 198-204.

[24] Zhang, X., Zhang, X., & Gu, C. (2017). A micro-artificial bee colony based multicast routing in vehicular ad hoc networks. *Ad Hoc Networks*, *58*, 213-221.

[25] Ucar, S., Ergen, S. C., & Ozkasap, O. (2015). Multihop-cluster-based IEEE 802.11 p and LTE hybrid architecture for VANET safety message dissemination. *IEEE Transactions on Vehicular Technology*, *65*(4), 2621-2636.