Vehicular Grouping Protocol: Towards Cyber Physical Network Intelligence

Duaa Zuhair Al-Hamid and Adnan Al-Anbuky
Electrical and Electronic Engineering
Auckland University of Technology
Auckland, New Zealand
Email: {Duaa.alhamid, Adnan.anbuky}@aut.ac.nz

***Abstract*— Vehicular network structures present a range of challenges and opportunities for efficiently managing awareness of road dynamics and network connectivity. An enhanced manageable organization can offer a better reaction to safety-related road events, facilitate dynamic topological flexibility, relate to road layout, and interact with unpredictable distribution of the vehicles. Vehicular grouping is one of the suggested structural techniques that offers a great benefit in grouping vehicles and modelling data routing, giving importance to road structure and the occurrence of a dynamic event within the associated group of vehicles. The approach discussed in this paper is based on a dynamic grouping through phases of self-formation, self-joining, self-leaving and self-healing as key components of the protocol operational cycle. Both vehicular physical connected resources and the remote computational cloud could be used for data processing and monitoring of road dynamics. This, in effect, encourages an Internet of Things (IoT) environment that enhances the dynamic performance through direct interaction between the virtualized network of vehicles and the physical network on the road leading to Internet of Vehicles (IoV). The objective of this paper is to develop a concept of network self-formation algorithm based on vehicle grouping strategy wherein the node can flexibly switch its function, be it an IoT gateway or a router node, based on the proposed fitness election model to be elected as group head. Testing using Contiki-Cooja simulator has been implemented on various road condition scenarios reflects the operational ability of the algorithm taking into consideration the network performance based on the ultimate capacity of the road.**

Keywords—Vehicular network, Self-formation, Wireless sensor network, Virtualization, IoT, IoV

I. INTRODUCTION

The vehicular network (VN) environment is an emerging field of research that offers various applications related to road safety, and involves many challenges related to communication, network structure and connectivity. Its dynamic nature and forceful mobility require stable communication to provide related information between vehicles and extend access to the Internet while on the road. This, in turn, allows the exchange of various types of data such as real-time traffic data extracted from sensors on board vehicles to facilitate passenger and vehicular services [1]. The main network design bottleneck is the existence of an appropriate topological organisation that can adapt to the dynamics of the system requirements. In a bid to overcome those challenges and manage the resource constraints of nodes like computational capabilities and node distribution, several recent works have focused on the routing algorithms and clustering/grouping approach for VN, virtualization, and utilizing cloud computing [2][3][4]. The clustering structure approach has been studied by researchers proposing various methods. From a topological design point of view, a topology clustering classification is demonstrated by Tan et al. [5]. The work examined the single-hop and multi-hop algorithms wherein the first forms a cluster based on a direct connection from the cluster head (CH) to all members, while the multi-hop cluster facilitates intermediate members to transmit data from the CH to remote members. The authors have reflected the core features of the clustering topologies, taking into consideration the performance aspect. For instance, the simplicity of single-hop algorithms could have an impact on the efficiency of the cluster, while more potential nodes can be isolated in large-scale clusters. On the other hand, the multi-hop approach promotes intra-cluster connectivity and covers more nodes to be connected than to be isolated. Several approaches [6-8] besides the above have been considered to enhance the aspect of clustering efficiency and increase connection stability. However, clustering algorithms lack the element of network flexibility under the dynamic changes and its scalability when complying with the road architecture and regulations. As vehicular clustering tends to be influenced by frequent changes in the network topology, the network lifecycle is a crucial part of the dynamics wherein vehicles may join or leave the cluster at any time. In this paper, this has been looked at from the self-formation of a cluster based on wireless sensor network (WSN) functions on a virtual platform allowing the performance analysis of a dynamic formed network associated with network downtime. Hence, forming a self-configured network that could be similarly formed for the physical network necessitates the support of the digital twin technology that generates a similar structure and behaviour to the physical environments [9]. This could then be utilised to facilitate planning for orchestrating the physical environment as demanded by road events. While the existing literature has highlighted the importance of this technology for various purposes, there is no provision for network digital twinning of vehicular grouping dynamics.

The remainder of this paper is structured as follows: Section II provides the proposed vehicular network grouping model including the system architecture and the fitness model. Section III describes the traffic case scenario of a highway. Section IV presents network modelling and analysis. Finally, the conclusion of the research work is presented in section V.

II. PROPOSED VEHICULAR NETWORK GROUPING MODEL

The structure of a vehicular network that could interact with the available resources in the cloud can be formed as discrete groups distributed along the stretch of the road. Each group containing vehicular nodes is distributed along the available lanes sharing the same direction of vehicular movements. Mainly, the group dynamics could be looked at as sub-processes based on aspects of self-formation, self-joining or departing, and self-healing to enhance the structured, organised data flow within the road systems. The nodes in a vehicular grouping approach are supported by the concept of WSN three core functions, namely the 'Leaf or Sensor Node', 'Router Node', and 'IoT Gateway Node'. This could be instrumental in forming physical or virtual organisations for a flexible vehicular sensor network [10]. As part of our vehicular grouping concept, we are looking at the self-formation approach considering the multi-hop managed tree topology structure based on a suggested highway traffic scenario in this paper.

This section details the elements of the vehicular self-formation phase and its related virtualisation.

A. *Cyber Physical Vehicular Network Vision*

The organisation of a vehicular system supported by the cloud environment can reflect the overall structure of an IoT-based system, the virtualisation environment and perception with the physical environment. Cloud resources, mainly network virtualisation, could offer intelligent interaction with the vehicular physical network. The operational, functional role of a node could be tested at a virtual level before its actual physical implementation, preventing any major modification that could take place. Hence, the network operation and performance can be tested, modified and structured to meet the requirements of the physical network. The proposed architecture for the vehicular system, where the physical network is connected over the Internet via a gateway node to the cloud environment paving the way for establishing the digital twin, is illustrated in Figure 1.

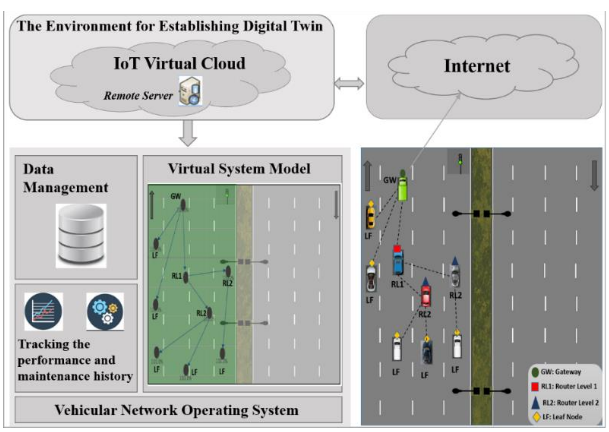

Fig. 1. Concept for a cloud-based vehicular network organization

It is worth mentioning here that dealing with a real-life application like VN requires real-time monitoring and analysis of data to deal with any changes or issues before they arise, and to recover the system if these changes happen. This could be enabled through the support of the structure and the related process of network virtualisation, which retain the structure and history of the dynamics of the physical network in the cloud. Considering the communication among the vehicular nodes in an ad hoc manner, the nodes attempt to establish a self-formation group wherein the role of a node is the key in organising this structure. Such structure is highly dynamic in nature, necessitating real-time and flexible mechanisms enhanced by the importance of the node's function. A node with the ability to access the cloud is a gateway function that can be elected as the head of the group, and be responsible for the connected nodes such as routers or leaf nodes. By enhancing such aspects related to the functional configuration process, the virtualisation platform offers a testing environment for running and analysing the formed group scenario. It is important to investigate the ability of the network to be functional under the various dynamic phases and allow any necessary adjustment needed to improve the network performance. Therefore, it is envisaged that better analysis of VN grouping organisation can be facilitated by the virtualisation concept for the physical network in capturing the various dynamics of traffic behaviour as an ongoing learning process.

B. *Vehicular Network Grouping Approach: Self-Formation*

The proposed dynamic grouping approach is assumed as a tree-based grouping that is formed along the stretch of the road covering the group. The lifecycle phases of the network dynamics could be reflected by networks running on a highway wherein each has its own lifecycle. The phases of being in the process of formation, already formed, or in the process of turning and diminishing, are to be assumed for the grouping approach. Each network sustains its integrity with members until most of its members leave the network. The focus here is on the network, that is, on the process of group formation when it initially groups vehicles within the vicinity of each other into a valid network of vehicular nodes.

The initial status of the group formation is being specified by the communication among the vehicles within the transmission range when more than one vehicle is available in the neighbourhood. This stage is crucial for gathering the data received by the nodes and analysing the ability of nodes to assume one of the three networking roles within the group. These are the gateway role, the router role and the leaf role. The node ability to assume more than one role could significantly benefit the election process when manoeuvring the orchestration of the organisation. Formulating the tree-based structure is initiated from the top to the bottom level, where the top level is represented by the head node that selects its connected nodes and forms the levels of connectivity. The vehicular network formation process flow chart and road example model is shown in Figure 2.

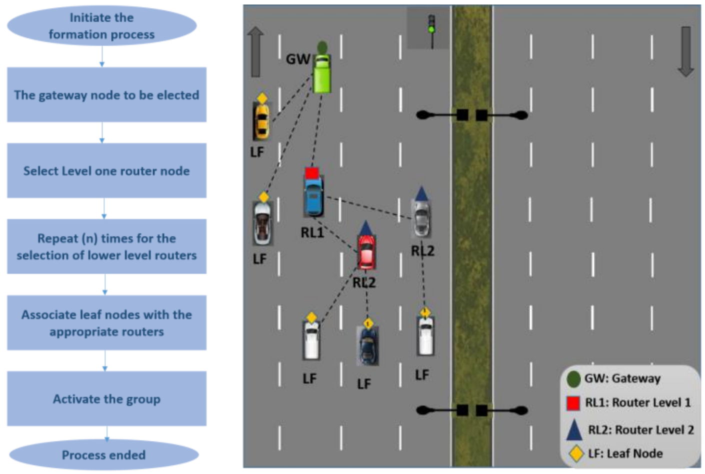

Fig. 2. The vehicular network formation process flow chart and road example model

The network structure process is designed through the cluster head fitness model and set of communication messages to be disseminated among the nodes as listed below.

The Fitness Node Election Model: The suggested parameters for a vehicular node to be involved in the competition are speed variance in relation to the group, location among the group, ability to communicate with a roadside node, and ability to act as a data router node. While the first two parameters are associated with possible measurements and could be used as fuzzy or estimated values, the other two are binary values assuming availability or not of the related features. These are briefly as follows:

- *Distance*: the distance between two vehicles (*V1* and *V2*) based on their coordinates (*i*, *j*) represented by *D* could be calculated based on the Euclidean distance as shown in Equation (1):

$$D_{(V1,V2)} = \sqrt{(V1_i - V2_i)^2 + \left(V1_j - V2_j\right)^2} \quad (1)$$

The average distance *D (Avg)* of a vehicle (for example *V1*) with respect to other vehicles within the transmission range can be calculated using Equation (2). Then Equation (2) can be applied to the other vehicles that are participating in grouping formation. The distance offers an estimated indication of the location of the vehicles within the group. However, depending on a parameter like the radio signal strength (RSSI) could offer a reasonable indication for the possibility of communication.

$$D_{V1}(Avg) = \frac{\sum_{i=2}^{n} D_{V1-Vi}}{N} \quad (2)$$

Where $\sum_{i=2}^{n} D_{V1-Vi}$ is the sum of D_{V1} with respect to other nodes starting with node (V_2) and ending with (V_n), *N* is the total number of nodes.

- *Radio Signal Strength Intensity (RSSI)*: this factor ensures the communication quality within the group and the gateway node. It also has an impact on the network performance wherein the weak signal message could be dropped and considered as packet loss. The accuracy of measuring the *RSSI* could be enhanced by considering the *RSSI* tolerance level of the node that could reflect the difference among the nodes of being central or close to each other. This parameter could be part of the election process.

- *Speed*: this reflects the vehicle movement, whether it is going to maintain its presence within the group for some time or possibly accelerate through or lag behind the group. Hence, it is important to set the tolerance level to differentiate between stable nodes that could last more time as part of a group. The deviation of a vehicular speed in reference to the average speed of vehicles in the neighbourhood is significant to the sustainability of a given node with the group. The less the deviation, the more sustainable the node is. While the speed limit is assumed to be 100 km/h, vehicles may run at speed variances of $(80 - 110)$ km/h. The speed average S_{Avg} is calculated for the nodes based on Equation (3) wherein S_n is the speed of a node, *N* is the total number of nodes. Hence, the speed variance S_{var} can be obtained using Equation (4):

$$S(Avg) = \frac{\sum S_n}{N} \quad (3)$$

$$SVar1 = SAvg - Sn \quad (4)$$

The fitness model for the election is emphasised by those variables wherein each will be given weight (C) depending on its importance for the process as well as the level of connectivity within the group. This can be formed to select a cluster head *FCH* based on the selected parameters as shown in Equation (5):

$$FCH = C1 * D\,(Avg) + C2 * RSSI + C3 * S\,Var1 \quad (5)$$

Communication Messages: The communication among the vehicular nodes within the transmission range is based on a set of messages that needs to be disseminated to get the vehicular-tree network formed. Each node disseminates its message according to the suggested parameters which are used for running the fitness model to elect the cluster head and the other levels of connectivity. Description of each action undertaken by the nodes to disseminate the related messages are provided as follows:

- Cluster head process: this process is firstly initiated by any of the gateway-capable nodes on the road initiating a call for forming a group. The communication messages are numbered as follows:

1) In response, each node within the neighbourhood broadcasting message (1) $\mathbf{M_{Hello\_candidate}}$ that contains the information related to (node ID, speed, node coordinates and ability to route data and connect to the roadside or cellular). Depending on the available nodes within the line of sight (LoS), each node receives the “hello packet” of all other nodes (see Figure 3). When the packet is received by each node, the vehicle distance **D** based on the received coordinates of each node, average speed $\mathbf{S_{avg}}$, and **RSSI** are calculated at each node.

2) Upon receiving the information from the other nodes and calculating the **D** and $\mathbf{S_{avg}}$, message (2) $\mathbf{M_{Ack}}$ is acknowledgement transmitted back by each node along with its average **D(Avg)** values, speed variance $\mathbf{S_{var}}$, and **RSSI(Avg).** This message is to notify the node about receiving the data and being part of the fitness model computation for electing the head node of the group.

3) As per the information collected from each node within the transmission range, the election fitness model will be running on those nodes that have received the aforementioned data. Message (3) $\mathbf{M_{FCH}}$ is broadcast among those nodes along with their fitness values. Hence, each node will have the values of the fitness model of the other nodes stacked so that it can be compared with its own value which is based on what the election triggers.

4) The node that has the highest fitness value is the winner to be elected as the head of the first group of the tree structure, and to be switched to a gateway function and assigned as the cluster head (CH). Subsequently, the nodes will be notified about the outcome of the election by broadcasting the node ID of the elected node to them; message (4) $\mathbf{M_{CH}}$ performs this action.

5) This process attempts to select number of nodes that connect to CH and to become router(s) level one (R1) based on the fitness values stored on CH to identify the eligible nodes; message (5) $\mathbf{M_{CH\_nominated\text{-}R1}}$ performs this action.

6) The selected nodes that have the best connectivity with respect to CH then transmit acknowledgement of connectivity with their node ID and router function is enabled on those nodes; message (6) $\mathbf{M_{R1\text{-}CH\_ACK}}$ performs this action (see Figure 3). However, the nodes that have poor connectivity will be unconnected and in a discovery status.

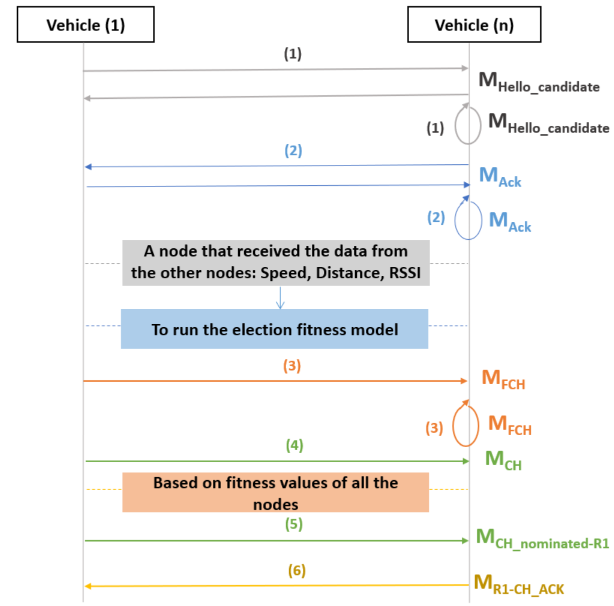

Fig. 3. The sequence diagram (UML) for the communication dialogue among the nodes for the election process

- Establishing the multiple levels of connectivity process: Based on CH election process, the stored fitness values are utilized to identify the nodes that are close to it for establishing the connectivity. The history of data based on “Hello packet” and the internal computation among the nodes including the fitness model can be utilized to form the tree automatically. The following communication messages are related to the tree levels of connectivity:

7) As R1 nodes are assigned, each has different fitness values stored in its buffer. Each will check the buffer and look into the nominated nodes that are eligible for establishing connectivity by sending message (7) $\mathbf{M_{R1\_nominated\text{-}R2}}$. As a result, each R1 is connected to multiple nodes that have the router function capability to switch to router level two (R2).

8) Those nominated nodes acknowledge the connectivity by transmitting message (8) $\mathbf{M_{R2\text{-}R1\_ACK}}$ to their R1 nodes. Hence, level two of the tree structure is established (see Figure 4).

9) Based on the number of nodes that are within the transmission range and the road structure, the levels of connectivity could be created. If few nodes are left for connectivity with certain criteria, then they could be leaf nodes seeking connection to a router. Hence, each R2 node tries to reach the unconnected nodes by sending message (9) $\mathbf{M_{R2\_nominated\_LF}}$ considering the fitness values stored in each R2 node.

10) Accordingly, the nominated leaf node(s) acknowledge the connectivity by transmitting message (10) $\mathbf{M_{LF\_R2\_ACK}}$ to each R2 node (see Figure 4).

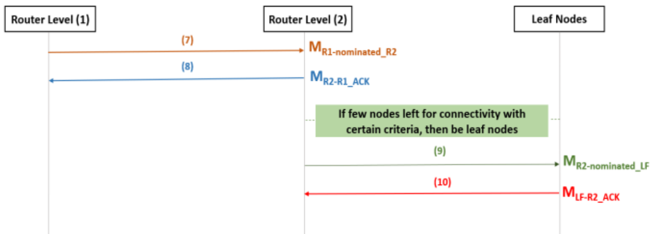

Fig. 4. The sequence diagram (UML) for the communication dialogue among the nodes to establish the multi-hop

III. TRAFFIC CASE AND SUGGESTED DYNAMIC GROUPING SCENARIO

The traffic on any road type, be it a highway or urban road, is characterised by various numbers of vehicles distributed in different lanes. The highway scenario reflects the continuous flow and sparse layout of vehicles to be the potential for grouping, taking into consideration the length of the stretch of road. The size of a vehicular group could be bounded by the road stretch and defined by various road parameters. This could, for example, be governed by the minimum two-second rule separation distance between any two following vehicles within the same lane. The rule is mainly to determine the safe and correct following distance that a driver should ideally keep from a vehicle in front of the driver's vehicle. The effective two-second rule only applies to normal weather conditions and should be extended for wet weather. Importantly, the type of communication, number of hops, and sensing measurements could also define the ideal capacity or modelling density of any group within the lane. The number of lanes on the road is another traffic parameter for indicating the density of groups accommodated within the highway sector. The speed on a $100\ \mathrm{km/h}$ highway could be assumed to vary from $80 - 110\ \mathrm{km/h}$, hence the calculation of the size of the group based on the two-second rule could vary accordingly. It is worth mentioning here that city roads use a much lower speed limit (possibly $50\ \mathrm{km/h}$) and could also be congested with a number of vehicles in a shorter portion of the road, hence the density and possible overlap among vehicular groups could take place. This could be considered subsequently as a more advanced level of vehicular grouping.

In keeping with the context of vehicular group size calculations, the lane occupancy on a highway with, for example, a stretch of 1 km can be calculated based on Equation (1) assuming various number of vehicles on the road:

$$O_L = \frac{V_T}{C_L} \qquad where\ O_L \leq 1\ if\ V_T \leq C_L \quad (1)$$

In which O_L is the lane occupancy per $1\ \mathrm{km}$, V_T is the total number of vehicles available on the road, C_L is the lane capacity (the maximum number of vehicles that the lane could accommodate based on speed and the two-second distance rule in normal ideal conditions). As Table I indicates, the lane capacity is the ultimate limit, and the vehicles may fill this limit or be part of it, based on random distribution and the number of lanes on the highway. The table offers the relationship between various numbers of vehicles and the speed limit of highway lanes within the road stretch, reflecting the estimated number of vehicles in a formed group as well as multiple cases of groups (various number of vehicles included) that could be formulated within the lane.

TABLE I. THE ESTIMATED NUMBER OF VEHICLES IN A GROUP

Speed Lane (km/h)	**Lane Capacity (CL)**	**2sec Distance (m)**	**(O_L) based on (V_T) Per 1Km**			
			V_T= 5	***V_T= 10***	***V_T= 15***	***V_T= 22***
80	22	44	O_L=23%	O_L=45%	O_L=68%	O_L=100%
90	20	50	O_L=25%	O_L=50%	O_L=75%	Saturation
100	17	56	O_L=29%	O_L=59%	O_L=88%	Saturation

Based on the calculated capacity of each condition, the distribution of vehicles for each would vary depending on the total number of vehicles entering/travelling along the lane. The lane density can be obtained accordingly. It is assumed in a multiple lane roading system that the high-speed lane could have the least number of vehicles due to its limited capacity while the left lane (the lower speed) could have a greater number of vehicles. Table II below depicts hypothesised distribution percentages of vehicles per lane represented by P_L, and the density of each lane as represented by K_L. The total number of vehicles entering the road is represented by V_T, and $V_{T(L)}$ is the total number of vehicles distributed per lane. A total number of 10-, 20-, and 30-vehicle scenarios have been considered for distribution on three-lane highways using the following two Equations (2) and (3). The random distribution various cases can be reflected on the vehicular grouping model to test the network structure.

$$K_L = \frac{V_{T(L)}}{C_L} \qquad (2)$$

$$P_L = \frac{V_{T(L)}}{V_T} \qquad (3)$$

TABLE II. THE ESTIMATED DENSITY AND DISTRIBUTION

V_T	**Three-Lane Highway**		
	Lane 1 (80km/h)	***Lane 2 (90km/h)***	***Lane 3 (100km/h)***
10	$V_{T(L)}$=5	$V_{T(L)}$=4	$V_{T(L)}$=1
	K_L =22.7%	K_L =20%	K_L =5.8%
	P_L =50%	P_L =40%	P_L =10%
20	$V_{T(L)}$=12	$V_{T(L)}$=6	$V_{T(L)}$=2
	KL =54.5%	KL =30%	KL =11.7%
	P_L =60%	P_L =30%	P_L =10%
30	$V_{T(L)}$=15	$V_{T(L)}$=10	$V_{T(L)}$=5
	K_L =68%	K_L =50%	K_L =29%
	P_L =50%	P_L =33%	P_L =16.6%

IV. NETWORK MODELLING AND ANALYSIS

The vehicular network self-formation grouping scenario has been modelled using Contiki-Cooja simulator as a WSN simulation and virtualisation tool. The vehicular network main scenario model is formulated based on pre-election and post-election processes, enhanced by vehicular-related parameters and the sequence of messages. The nodes are distributed exchanging sets of sense data to establish the communication wherein each node initially sends and receives the message. As the election takes place based on the proposed fitness model, one of the nodes is elected as a cluster head and consequently the level(s) of connectivity are structured. The functional role of the elected node can be switched to a gateway node via softwarization. The other node(s) functionality can be switched or remain as per the network requirements and the levels of connectivity with respect to other node(s). The following Table III depicts the simulation parameters that are utilised in Cooja scenarios taking into consideration the communication and road parameters.

TABLE III. SIMULATION PARAMETERS

Simulation Parameters	**Value**
Simulation time	5 minutes
Total number of vehicles	30
Number of lanes	3-highway lanes (same direction)
Number of hops	3-hops

The following Figure 5 depicts an example of 8 vehicles for the implementation and modelling of the proposed vehicular network structure. Three nodes (R2) are connected to router level one (R1) and three nodes (leaf) connected to (R2, node ID 4) based on the transmission range of the network. This offers an example of a multi-hop structure using Cooja simulator.

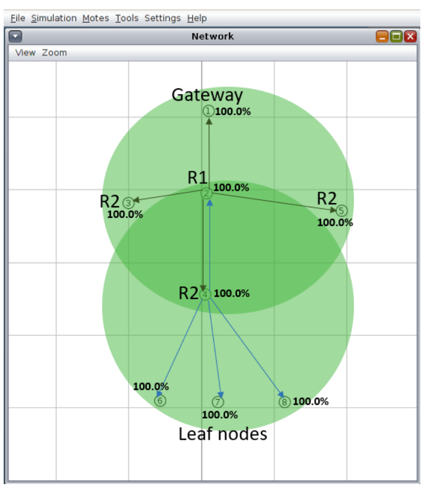

Fig. 5. Cooja network scenario

A. Pre-Election Process Network Behaviour

The initial stage as proposed in the communication dialogue is based on an un-clustered group (ad hoc network), where the communication among the nodes takes place with all the nodes within the transmission range. This could cause data traffic congestion, message loss, and latency in receiving the data due to the absence of a head node that could lead a group of nodes. For this phase, the message received and the delay in data dissemination are the performance measures to be evaluated, reflecting the network reliability and the need for the cluster formation phase.

Based on the literature, a variable sampling rate (communication message rate) seems more suitable for occasional disturbance in traffic flow; a lower sampling rate is adopted for stable traffic flow; and a higher sampling rate is required when a traffic disturbance occurs. The network scalability has been tested based on various communication messages rates that are (10, 20, and 30 messages/second), and number of vehicles (10, 20, and 30). It is shown in Figure 6 below that the percentage of messages received decreases with the increase in the message rate and number of vehicles. The lower message rates, ranging from 1-10 messages/second, could be a better choice for securing the nodes receiving more percentage of the data forwarded to that node so that it could be reflected on the formation process.

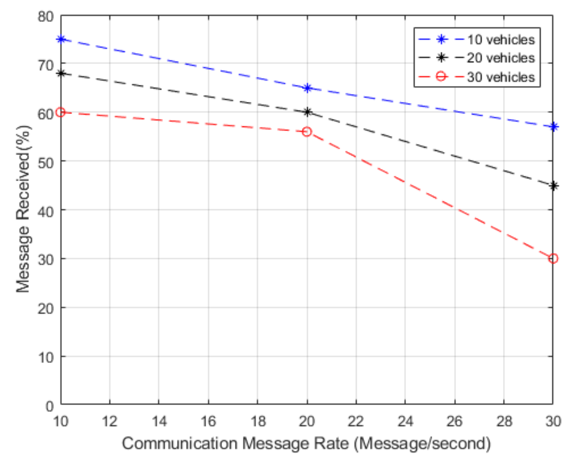

Fig. 6. Pre-election message received

Under this initial phase, the data dissemination has been analysed based on the time that the node takes to disseminate its data to the neighbour nodes. Figure 7 below shows that with the increase in number of vehicles (5, 10, 15, 20, 25, and 30), the latency in data dissemination occurs due to the ad-hoc flat communication nature among the nodes where the messages are being sent and could be received by every node within the transmission range.

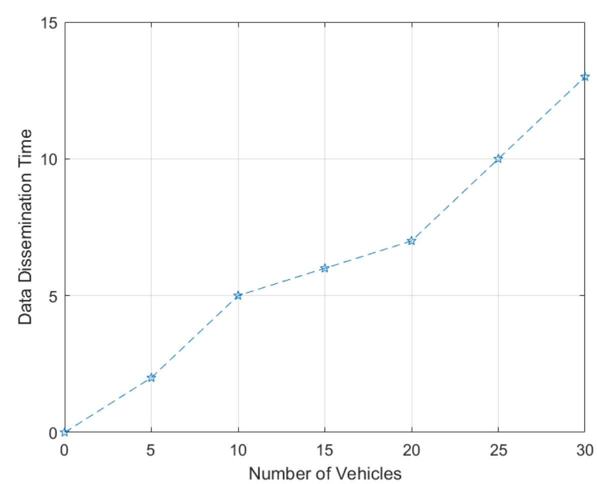

Fig. 7. Latency in data dissemination

B. *Router Capacity and Multi-Hop Network*

The suggested road stretch and its ultimate capacity based on the suggested highway traffic scenario indicate the possible distribution where the available vehicles within the lane may occupy full or partial capacity. Based on that, the number of hops and the time in formulating the vehicular group will vary accordingly. The other factor that can influence the structure of a given group is the transmission range of the communication protocol where the nodes should be within the LoS to formulate the level(s) of connectivity. The best size of a formulated group is tested through the analysis of a gateway or a router node capacity in covering the available nodes using Cooja. This, in turn, will be reflected on the number of hops that a group could have.

Based on the variance messages rates (message/second), Figure 8 below depicts the messages received by a router node with a different set of number of vehicles (5, 10, 15, and 20). As can be clearly seen from the figure, the messages received by a router connecting five vehicles range from 100% to 80% while the received messages range from 100% to 48%, from 100% to 4%, and from 97% to 1% for a router connecting 10 vehicles, a router connecting 15 vehicles, and a router connecting 20 vehicles respectively. The outcome here reflects the possible size of a vehicular group along with its boundary in terms of number of hops. In this case, a head node connecting to 10 vehicles formulating a single-hop group could be the ideal size with message rate ranging from (1-5 messages/second) as 100% of the messages are received. However, this is subject to the available vehicles within the road and the type of vehicular application for choosing the suitable message rate.

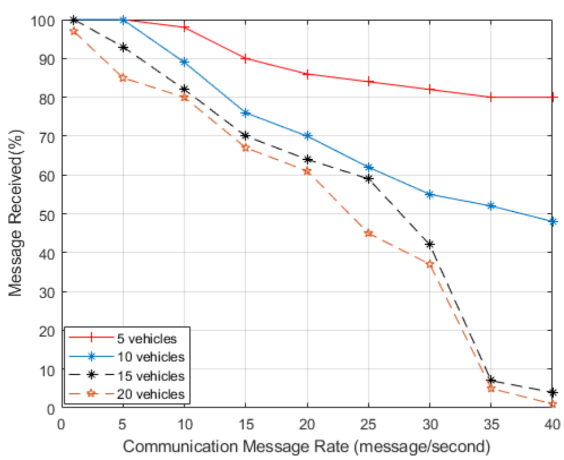

Fig. 8. The router node capacity

Upon the analysis of a node connection capacity, the communication parameter of a single and multi-hop structure can be correlated to the road parameters of lane capacity and vehicles distribution. A Cooja scenario based on the formation of network levels is designed with the emphasis on the ideal number of hops for a highway lane. The distance factor among the vehicles that is mainly related to RSSI plays an important role in formulating the group as well as in structuring the levels of network. This also is related to the use of the communication protocol that is enabled by the nodes in setting the transmission range among the nodes. The ZigBee protocol has been used as an option for developing the model while the protocols performance testing is the next stage of the work. Herein, the ZigBee protocol used in the Cooja simulation provides $200-250m$ LoS for managing the hop(s) that can be correlated to the capacity of a given road stretch. As based on the traffic scenario analysis, the $1km$ stretch ultimate capacity for the $80Km/h$ lane speed is 22 vehicles.

When a given transmission range in Cooja is applied, the possibility of establishing the communication among the nodes depends on the signal strength of the nodes and the number of nodes available within the space. Figure 9 below depicts the result of a Cooja scenario designed to test a number of vehicles under various transmission ranges in terms of messages received. Even though the router node can connect 10 vehicles with good performance as was shown in Figure 8, the nodes could be available within the stretch of the road, but with different distances enforcing the need for other router(s) to connect some of the nodes based on multi-hop. This is assumed when the RSSI of a given node is below $-90dB$. As can be seen from Figure 10, a group of 15 vehicles with a transmission range of 50, and 100m has message received percentages as 60, and 78 respectively. Even though this indicates the better performance under a larger transmission range, this is dependent on the RSSI, distance among the vehicles within the stretch of the road, and the number of vehicles. Hence, considering a single-hop group could be challenging for a vehicular network system. Our approach implies the need for having a managed group based on a tree structure where the number of hops should meet the requirements of the protocol range, the distance among the vehicles, and the road stretch. Furthermore, the flow of data within the tree-based group can be enhanced through a load balanced levels of router nodes.

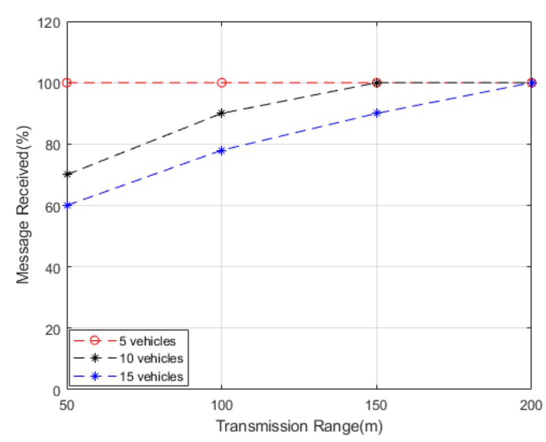

Fig. 9. Message received based on a given transmission range in Cooja

C. *Network Formation- Election Process*

As per the proposed UML and the analysis of the first stage of the algorithm, the election process takes place when the data is exchanged among the nodes for electing a node to be the head of the group based on the proposed fitness model. The number of transactions reflects the communication messages exchanged among the nodes proposed in Figure 3 to facilitate the election process. The different set of numbers of vehicles, and the 5 messages/sec have been considered for the evaluation, as shown in Figure 10 below. Each set of vehicles has various number of transactions depending on the communication process.

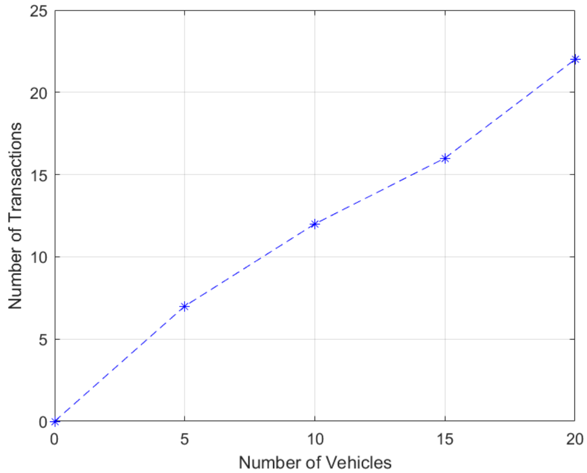

Fig. 10. Network formation- election process based on number of transactions

D. Single- and Multi-Hop Network Formation Time

A Cooja scenario is designed to investigate the time of a group formation post-election, taking into consideration a single- and multi-hop structure, and the lane capacity. Based on the above analysis of the ideal number of hops, along with the lane capacity, a test is conducted to analyse single-hop, two-hop, and three-hop networks with 5, 10, and 15 nodes, as shown in Figure 11 below. The multi-hop-based group time is based on the number of messages involved within the process.

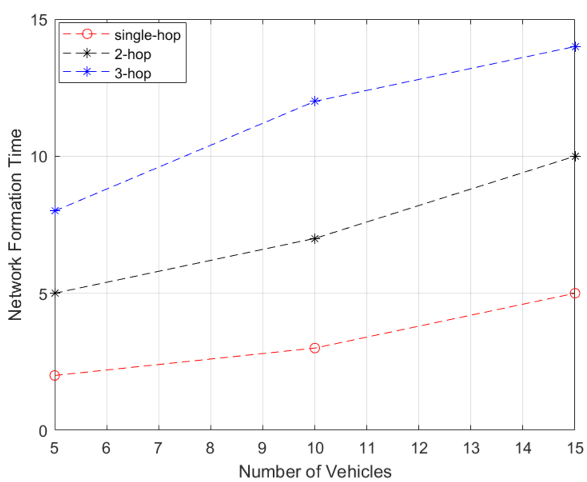

Fig. 11. The single- and multi-hop network formation time

V. Conclusion

The paper has explored a novel approach for the vehicular self-formation based-grouping considering the flexibility of the nodes to have different functions to be switched as needed in the model. Herein, the approach emphasis the vehicle connectivity node as either gateway, router or leaf function within the network structure promoting a flexible network structure. The proposed protocol has been modelled using Contiki-Cooja simulation tool. Both the network communication parameters and the physical traffic parameters have been considered in the fitness model for the tree network formation. The group formation process has been tested based on a router node capacity and the number of hops considering the road stretch and its ultimate capacity. This analysis reflects the need of the vehicular network structure to be complied with the road layout. Herein, the data received by a router node with a different set of number of vehicles reflects the possible size of a vehicular group along with its boundary in terms of number of hops. The other crucial performance measure used in network formation analysis is the time of the process based on number of transactions. The formation time of various number of hops, along with the different sets of numbers of vehicles available within each hop, have been considered.

For future work, the dynamic grouping structure model will have the key components of self-formation, self-joining, self-leaving and self-healing integrated as a full solution. In addition, the model will be validated using the physical network, paving the way for the digital twin paradigm.

References

[1] Sharif, A., Li, J.P., & Saleem, M.A., "Internet of things enabled vehicular and ad hoc networks for smart city traffic monitoring and controlling: a review". *International Journal of Advanced Networking and Applications*, 2018. **10**(3): pp. 3833-3842.

[2] Senouci, O., Harous, S., & Aliouat, Z., "Survey on vehicular ad hoc networks clustering algorithms: Overview, taxonomy, challenges, and open research issues". *International Journal of Communication Systems*, 2020. pp. e4402.

[3] Acharyya, I.S., A. Al-Anbuky, and S. Sivaramakrishnan. “Software-defined sensor networks: towards flexible architecture supported by virtualization”. in *2019 Global IoT Summit (GIoTS)*. 2019. IEEE.

[4] S. Ezdiani, I. S. Acharyya, S. Sivakumar, and A. Al-Anbuky, "An IoT Environment for WSN Adaptive QoS". in *2015 IEEE International Conference on Data Science and Data Intensive Systems*, 2015, pp. 586-593.

[5] Tan, K., Le Kernec, J., Imran, M., & Bremner, D. "Clustering Algorithm in Vehicular Ad-hoc Networks: A Brief Summary". In *2019 UK/China Emerging Technologies (UCET)*. pp. 1-5. 2019. IEEE.

[6] Khan, Z., Fan, P., Fang, S., & Abbas, F., "An unsupervised cluster-based VANET-oriented evolving graph (CVoEG) model and associated reliable routing scheme". *IEEE Transactions on Intelligent Transportation Systems*, 2019. **20**(10): pp. 3844-3859.

[7] Rashid, S.A., Audah, L., Hamdi, M.M., & Alani, S., "Prediction Based Efficient Multi-hop Clustering Approach with Adaptive Relay Node Selection for VANET". *Journal of Communications*, 2020. **15**(4).

[8] Qureshi, K.N., Abdulla, A.H., Bashir, F., Iqbal, S., & Awan, K.M., "Cluster-based data dissemination, cluster head formation under sparse, and dense traffic conditions for vehicular ad hoc networks". *International Journal of Communication Systems*, 2018. **31**(8): pp. e3533.

[9] Madni, A.M., Madni, C.C., & Lucero, S.D., "Leveraging Digital Twin Technology in Model-Based Systems Engineering". *Systems*, 2019. **7**(1): pp. 7.

[10] Al-Hamid, D.Z. and A. Al-Anbuky. “Vehicular grouping and network formation: virtualization of network self-healing”. in *International Conference on Internet of Vehicles*. pp. 106-121. 2018. Springer.