

4-2014

# A New Clustering Protocol for Hybrid Sensor Vehicular Networks (HSVNs)

Saleha Mubarak Mohammed Al Mheiri

Follow this and additional works at: [https://scholarworks.uaeu.ac.ae/all\\_theses](https://scholarworks.uaeu.ac.ae/all_theses)

Part of the [Engineering Commons](#)

---

## Recommended Citation

Mohammed Al Mheiri, Saleha Mubarak, "A New Clustering Protocol for Hybrid Sensor Vehicular Networks (HSVNs)" (2014). *Theses*. 230.

[https://scholarworks.uaeu.ac.ae/all\\_theses/230](https://scholarworks.uaeu.ac.ae/all_theses/230)

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarworks@UAEU. It has been accepted for inclusion in Theses by an authorized administrator of Scholarworks@UAEU. For more information, please contact [fadl.musa@uaeu.ac.ae](mailto:fadl.musa@uaeu.ac.ae).

United Arab Emirates University  
College of Engineering  
Department of Electrical Engineering

A New Clustering Protocol for Hybrid Sensor Vehicular Networks  
(HSVNs)

Saleha Mubarak Mohammed Al Mheiri

This thesis is submitted in partial fulfillment of the requirements for the  
Master of Science in Electrical Engineering degree

Under the direction of Dr. Hend Al Qamzi

April 2014

## DECLARATION OF ORIGINAL WORK

I, Saleha Al Mheiri, the undersigned, a graduate student at the United Arab Emirates University (UAEU) and the author of the thesis titled "A New Clustering Protocol for Hybrid Sensor Vehicular Networks (HSVNs)", hereby solemnly declare that this thesis is an original research work done and prepared by me under the guidance of Dr. Hend Al-Qamzi, in the College of Engineering at UAEU. This work has not been previously formed as the basis for the award of any degree, diploma or similar title at this or any other university. The materials borrowed from other sources and included in my thesis have been properly cited and acknowledged.

Signature:

*Al Mheiri*

Date: 17/4/2014

Copyright © 2014 by Saleha AlMheiri

All Rights Reserved

Approved by

Thesis Examination Committee:

1) Advisor: Dr. HEND AL-QAMZI

Title: Assistant Professor

Department of Electrical Engineering

College of Engineering

Signature: *Handwritten signature of Dr. HEND AL-QAMZI*

Date: 17-4-2014

2) Internal Examiner: Dr. Qurban Ali

Title: Assistant Professor

Department of Electrical Engineering

College of Engineering

Signature: *Handwritten signature of Dr. Qurban Ali*

Date: 17-4-2014

3) External Examiner: Prof. Anura Jayasumamana

Title: Professor

Department of Electrical & Computer Engineering

Institution: Colorado State University

Signature: *Handwritten signature of Prof. Anura Jayasumamana*

Date: 24/4/2014

Accepted by

Master's Program Coordinator: Dr. Abbas A.Fardoun

Signature: *Abbas Fardoun*

Date: *17-4-2014*

Dean of the College: Prof. Amr S. El-Dieb

Signature:

*4/5/2014*  
*Amr S. El-Dieb*

Date: *4/5/2014*

Copy \_\_\_ of \_\_\_

## ABSTRACT

As the UAE is classified among the highest countries in road accidents; it's an urgent necessity to derive proper and effective solutions. One of the well-known solutions is to move to the Intelligent Transportation Systems (ITS). ITS is achieved by the implementation of advanced technologies to help the roads to be smarter, safer and more coordinated. There are different studies and researches published in this area. This thesis is mainly conducted to achieve the same goal, which is increasing the safety on our roads.

The aim of this research is to build a new clustering protocol in the Hybrid Sensor Vehicular Networks (HSVN). HSVN is introduced as a new concept that integrates wireless sensor networks along with vehicular networks to increase the overall performance of both networks and the need of such new protocol in these networks is high. Although there are many published proposals, there is still a need to have a complete protocol that can reduce the shortcomings in the proposed solutions.

The main features in the proposed protocol are creating a balanced system by dividing the clustering process overheads between the vehicles in the network and the Road Side Unit (RSU), so instead of relying only on the vehicles' resources, the RSU shares the process's overhead by collecting vehicles data, calculating weighting factors, and electing suitable cluster heads. Moreover, the proposed protocol reduces the computational and the communication costs by electing two cluster heads for each cluster; one acts as the main cluster head and the other as a standby cluster head. By this feature, if a cluster head moves outside the cluster region, there is no need to run the clustering process again to elect a

new cluster head as a standby cluster head exists. Also, one of the main characteristics of the new protocol is minimizing the collision in the system, and consequently increasing the throughput by defining the upper bound of the number of members in each cluster.

The overall performance of the proposed protocol is very good and promises to solve many challenges in the existing protocols. In addition, the results show that this protocol outperforms one of the best existing mobility protocols in terms of the total number of clusters formed in the network, the number of single node clusters, the saturation throughput of the clusters, the communication overheads and energy consumption reductions.



## ACKNOWLEDGEMENT

First and always, all my appreciation is to Allah, who is always by my side, facilitates every difficulty facing me, and makes the impossible to be possible. Thanks to Allah.

I would like to express my sincere gratitude to my advisor Dr. Hend Al Qamzi for her continuous support and advice in my thesis research. Also, a special thanks to my friend Eng.Sarah Madi for reviewing the thesis report and providing me with valuable comments and encouragement.

Also, I would like to thank my examination committee for their valued inputs in improving the thesis, Dr. Qurban Ali and Prof. Anura Jayasumamana. In addition, my deepest thanks and appreciation to the chair and all members of the Department of Electrical Engineering at UAE University for assisting me all over my studies and research.

A special thanks to my family for all of the sacrifices that they have made for me to reach this level in education and life.

## **DEDICATION**

I dedicate this modest research to my country, and to the beloved people of the UAE. Also, I dedicate my research for all researchers who spent their time and efforts to find radical solutions to save the humans' lives.

## Table of Contents

Abstract .....	vii
Acknowledgment .....	ix
Dedication .....	x
Table of Contents.....	xi
List of Tables .....	xiii
List of Figures .....	xiv
List of Abbreviations .....	xv
<b>CHAPTER 1: Introduction.....</b>	<b>1</b>
1.1: Intelligent Transportation Systems .....	2
1.2: Mobile Ad-hoc Networks .....	2
1.3: Vehicular Ad-hoc Networks.....	3
1.4: Wireless Ad-hoc Networks.....	3
1.5: Hybrid Sensor Vehicular Networks.....	3
1.6: IEEE 802.11p.....	4
1.7: Clustering .....	7
1.7.1: Benefits of Clustering .....	8
1.8: Thesis Overview .....	9
<b>CHAPTER 2: Literature Review .....</b>	<b>10</b>
2.1: Classification and Comparison .....	10
2.1.1: ID-based Clustering .....	10
2.1.2: Nodal degree-based Clustering.....	11
2.1.3: Mobility-based Clustering .....	14
2.1.4: Direction-based Clustering .....	25
2.1.5: Leadership duration -based Clustering .....	27
2.1.6: Path loss-based Clustering .....	28
2.2: Drawbacks in existing algorithms .....	29
2.3: Classification summary .....	30
<b>CHAPTER 3: Protocol Design .....</b>	<b>34</b>
3.1: Problem Statement.....	34
3.2: Methodology.....	35
3.3: Design Challenges.....	36
3.4: Expected Impacts.....	36
3.5: Contributions and New Features .....	37
3.6: Assumptions .....	37
3.7: Protocol flowchart .....	38

3.8: Protocol Operation.....	40
3.8.1: Initial Stage.....	40
3.8.2: Cluster Formation.....	40
3.8.3: Cluster Head (CH) Election .....	41
3.8.4: Cluster Maintenance.....	44
3.9: Messages Exchanged .....	47
3.10: Throughput Calculation.....	50
<b>CHAPTER 4: Results and Discussion .....</b>	<b>54</b>
4.1: Simulation Scenario .....	54
4.2: Simulation Setup.....	56
4.3: Results .....	58
4.3.1: Total number of clusters .....	58
4.3.2: Single node clusters .....	60
4.3.3: Cluster size .....	61
4.3.4: Cluster lifetime .....	62
4.3.5: Standby cluster head .....	63
4.3.5.1: Communications Overhead .....	63
4.3.5.2: Energy Consumption .....	64
4.4: Discussion .....	67
<b>CHAPTER 5: Conclusion .....</b>	<b>69</b>
<b>Bibliography .....</b>	<b>71</b>

## List of Tables

<b>Name</b>	<b>Page</b>
Table 1: Summary of existing clustering algorithms.....	31
Table 2: Summary of existing clustering algorithms based on mobility.....	32
Table 3: The message fields' size.....	50
Table 4: Throughput analysis parameters.....	52
Table 5: Energy calculations parameters.....	65

## List of Figures

Name	Page
Figure 1.1: CSMA/CA access methods.....	7
Figure 3.1: Protocol Flowchart.....	39
Figure 3.2: Cluster formation.....	43
Figure 3.3: Flowchart for Scenario 1.....	44
Figure 3.4: Flowchart for Scenario 2.....	45
Figure 3.5: Flowchart for Scenario 3.....	46
Figure 3.6: Flowchart for Scenario 4.....	46
Figure 3.7: HELLO message fields.....	47
Figure 4.1: A sample of traffic data.....	55
Figure 4.2: Khaled Bin Sultan Street location.....	55
Figure 4.3: Khaled Bin Sultan Street Snapshot.....	56
Figure 4.4: Euclidean Distance Calculation Example.....	57
Figure 4.5: Total number of clusters in proposed and DCA algorithms for 10 times.....	58
Figure 4.6: Total number of clusters in the proposed and the DCA algorithms.....	59
Figure 4.7: Single node clusters in the proposed and the DCA algorithms.....	60
Figure 4.8: The average of single node clusters in the proposed and the DCA algorithms	61
Figure 4.9: Largest cluster size in the proposed and the DCA algorithms.....	61
Figure 4.10: Largest cluster size in the proposed and the DCA algorithms.....	62

## List of Abbreviations

- IITS:** Intelligent Transportation System
- VANET:** Vehicular Ad-hoc NETWORK
- WSN:** Wireless Sensor Networks
- MANET:** Mobile Ad-hoc NETWORK
- HSVN:** Hybrid Sensor Vehicular Network
- WAVE:** Wireless Access in Vehicular Environment
- IEEE:** Institute of Electrical and Electronics Engineers
- RSU:** Road Side Unit
- V2V:** Vehicle to Vehicle
- V2R:** Vehicle to Roadside unit
- FCC:** Federal Communications Commission
- DSRC:** Dedicated Short Range Communication
- AIFS:** Arbitration Inter Frame Space
- CW:** Contention Window
- OFDM:** Orthogonal Frequency Division Multiplexing
- DCF:** Distributed Coordination Function
- RTS/CTS:** Request to send/Clear to Send
- DIFS:** Distributed Inter-Frame Space
- SIFS:** Short Inter-Frame Space
- CSMA/CA:** Carries Sense Multiple Access /Collision Avoidance
- CH:** Cluster Head
- ACK:** Acknowledgment
- DOT:** Department Of Transport
- GPS:** Global Positioning System
- NS-2:** Network Simulator 2
- GloMoSim:** Global Mobile Information System Simulator

# CHAPTER 1

## Introduction

According to the latest road statistics, the accidents' rates are in continuous increase. Police records show that about 51,200 accidents happened in the period 2006 and 2011; these accidents led to 63,406 injuries. It is depressing to know that UAE is classified as one of the countries that have the highest rates in road accidents relative to its population. In addition, road accidents cost about 20 billion dirhams [1]. Moreover, different studies show that the main causes of road accidents are due to the irresponsible behavior of the driver and the high speeds [2]. As a result, several projects and researches are conducted in this area in order to reduce the traffic accidents rates and increase the roads' safety.

Through this chapter different topics that are strongly related to the thesis are reviewed and summarized. Firstly, Intelligent Transportation Systems (ITS) with different network configurations are presented. The networks' configurations include Mobile Ad hoc Networks (MANET), Vehicular Ad hoc Networks (VANET), Wireless Sensor Networks (WSN), and Hybrid Sensor Vehicular Networks (HSVN). Then, we shed the light on the IEEE 802.11p protocol in order to highlight its main features in the vehicular networks. Finally, we focus on the main goal of this thesis which is clustering. Clustering concept and benefits are presented in this chapter. Different clustering approaches and classifications are discussed in the following chapter.



## **1.1 Intelligent Transportation Systems (ITS)**

The number of vehicles on our roads is rapidly increasing year by year. This huge volume of vehicles leads to higher accidents' rates and more congested roads. Therefore, there is a severe need to exploit new technologies to minimize the fatalities in the streets and increase the traffic safety. The Intelligent Transportation Systems (ITS) aims at increasing the roads' safety and enhancing the roads' efficiency. ITS consists of communications systems, electronic technologies, processing and controlling units. Moreover, ITS have different applications, such as overtaking assistance, animal detection, road enforcement, Collision avoidance, etc. [3] . There are several forms of networks that are designed for intelligent transport applications. The main forms are the Mobile Ad hoc Networks (MANET), Vehicular Ad hoc Networks (VANET), Wireless Sensor Networks (WSN), and Hybrid Sensor Vehicular Networks (HSVN).

## **1.2 Mobile Ad hoc Networks (MANETs)**

Mobile Ad hoc Networks (MANETs) consists of mobile users or nodes that can communicate wirelessly. These networks work in a distributed manner with no fixed infrastructure. MANETs have different applications, such as battlefields, environmental monitoring, cars networks, etc. The main characteristics of MANETs are: the nodes' movement is not restricted to a certain pattern and can move freely in the field. Also, MANETs have more concentration on energy as the nodes depend on exhaustible batteries. In addition, these networks are prone to different security threats that influence their performance and their overall functions [4].

### **1.3 Vehicular Ad hoc Networks (VANETs)**

Vehicular ad hoc network has enormous researchers' attention due to its unique features. It can improve road safety, increase the traffic efficiency, and offer entertainments. VANETs consist of sensors that are embedded on vehicles and on roadside units. There are two forms of communications in this network: the vehicle to vehicle (V2V) and the vehicle to roadside (V2R) communications. These networks are different from other wireless networks in many aspects, for example, VANETs have predictable vehicles movements which are restricted to roads' structure and traffic regulations. Moreover, VANETs have powerful processing units and large storage capacities. Also, they do not have a problem with the energy constraints since they rely on the vehicles' batteries. However, VANETs suffer from link disconnectivities due to the high mobility of vehicles and low density in rural areas [5].

### **1.4 Wireless Sensor Networks (WSNs)**

Wireless Sensor Networks consist of tiny stationary sensors that have high efficiency detection capabilities with low costs. However, these networks have limited energy resources and low processing and storage capacities [6].

### **1.5 Hybrid Sensor Vehicular Networks (HSVNs)**

In order to strengthen the advantages of VANETs and WSNs and compensate for their weaknesses, the new concept of Hybrid Sensor Vehicular Networks (HSVNs) is introduced. In HSVNs, stationary nodes are deployed along the road sides to continuously collect information about the road conditions. Then, the gathered data is delivered to the passing vehicles where

they will transfer it to other regions using different dissemination protocols [7]. In HSVNs, there are two types of sensor nodes; sensors that are embedded on the vehicles which is known as vehicular nodes and sensors that are deployed on the lighting poles at road sides known as Road side sensor nodes [8]. It is to be noted that HSVNs do not suffer from energy constraints; this is due to the availability of road facilities which can provide an alternative power supply from the power lines or from solar panels [3].

### **1.6 IEEE 802.11 P**

The Federal Communications Commission (FCC) allocated 75 MHz of the frequency spectrum in the range 5.85-5.925 GHz for the vehicle to vehicle and vehicle to roadside communications. This spectrum is the Dedicated Short Range Communication (DSRC) which has one control channel and six service channels [5]. These channels are divided into a safety critical channel and non-safety channels with 10 MHz bandwidth. The safety critical channel has the priority over other channels and it has the highest transmission power. Each channel has four access categories which are AC0, AC1, AC2, and AC3. The AC3 has the highest priority over the other categories. For each different category, the Arbitration Inter Frame Space (AIFS) and the Contention Window (CW) are selected accordingly. The physical layer of the IEEE 802.11 p is based on OFDM which stands for Orthogonal Frequency Division Multiplexing.

It has been shown that the RSU distance should not exceed 1000 m in order to transmit steady data streams to the passing vehicles. Moreover, it is recommended that the maximum delay for the safety applications should not exceed 100 ms. Additionally, in order to minimize the delay in the control channel

for the safety applications, the total exchanged packets should be less than 1000 packets per second [9]. IEEE 802.11p is designed to provide V2V and V2R communications in the range up to 1000m and velocity up to 30 m/s. Also, it supports a wide transmission rate from 3 to 27 Mbps [10].

On the other hand, IEEE 802.11p depends on the Carrier Sense Multiple Access Collision Avoidance (CSMA/CA) as a medium access method. The IEEE 802.11p Distributed Coordination Function (DCF) defines two different access mechanisms which are the basic access mechanism and RTS/CTS (Request to send/Clear to Send) access mechanism. The former is a two way handshaking technique and it is summarized in the following steps:

1. A node that has a packet to transmit monitors the medium.
2. If the medium is idle for a period of time equal to the Distributed Inter-Frame Space (DIFS) then the packet will be transmitted.
3. However, if the medium is busy, the node will observe the channel until it becomes idle for a DIFS, then it will wait for a random back off interval before transmitting in order to minimize the probability of collisions.
4. Also, the node should wait a random back off time between two consecutive transmissions to avoid capturing the medium for a long time which affects the other transmissions.
5. When the packet is received successfully, an ACK is sent after a period of time called the Short Inter-Frame Space (SIFS).

DCF supports an exponential back off scheme, where the back off time is chosen from 0 to  $W-1$ .  $W$  stands for the contention window and its value depends on the number of failing transmissions. At the first attempt,  $W$  is equal to  $CW_{min}$

(minimum contention window) and each time is doubled for each unsuccessful attempt until it reaches  $CW_{max}$ . In addition, the back off time counter is decremented as the medium is sensed idle. On the other hand, the node transmits when the back off timer becomes zero. Also, this counter is frozen when there is a transmission in the medium, and is activated when the medium is sensed idle.

The second access technique which is the RTS/CTS is similar to the basic access method, but instead of sending the packet directly, a short RTS frame is sent as a request to reserve the medium. Then, the receiving node will wait for a SIFS before replying with CTS frame, giving permission to the requesting node to transmit. After receiving the CTS frame, the node then is allowed to send its packet. This method is beneficial in reducing the hidden terminal problem. Figure 1.1 illustrates both CSMA/CA access methods.

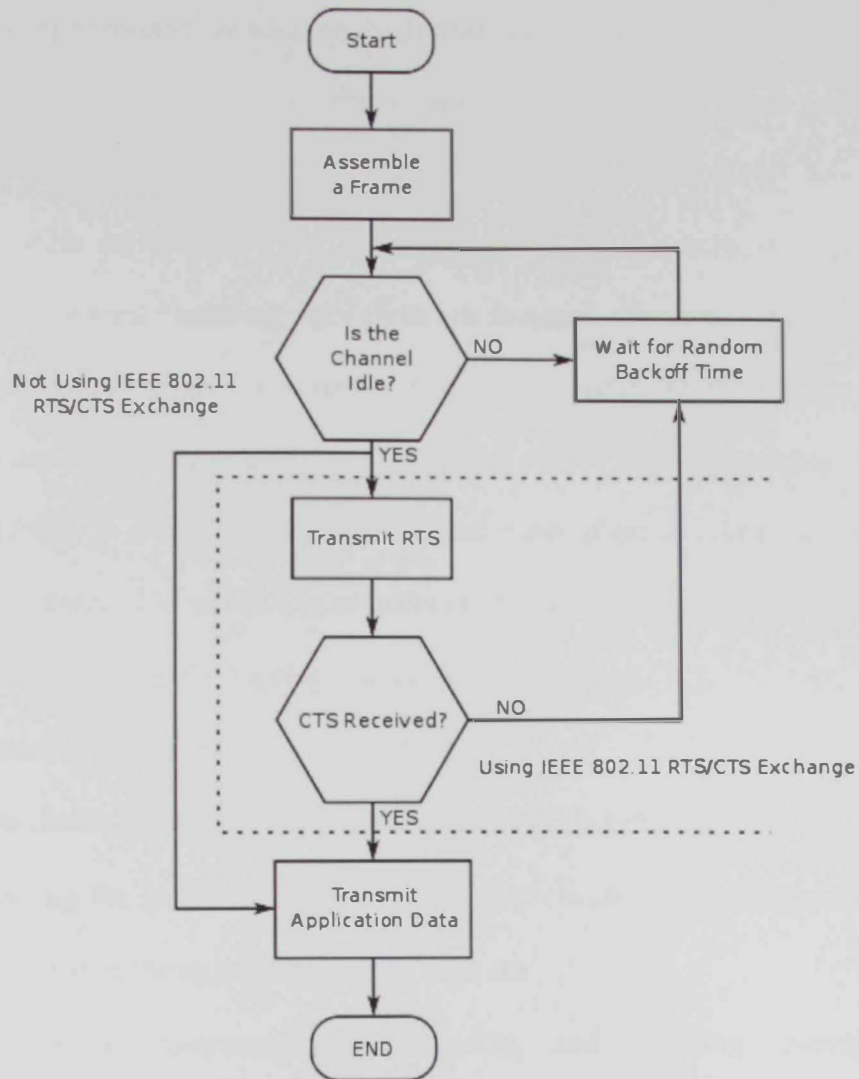


Figure 1.1: CSMA/CA access methods (source: Communication Networks, 2006)

### 1.7 Clustering

Messages in HSVNs are broadcasted among the nodes, however a number of challenges appear to affect reliable broadcasts. The straight forward method of broadcasting is flooding. In the flooding mechanism, each node in the network that receives a broadcast message will rebroadcast it to its nearby nodes. Although this method looks simple and easy to implement, it causes a broadcasting storm problem. This problem exists when there are redundant rebroadcasting messages as each node will receive multiple copies of the same message; wasting bandwidth

and processing resources. In addition, high contention occurs to access the wireless channel since all nodes want to rebroadcast at the same time. As a result, the number of collisions in the network increases as well as the delay [11].

In order to minimize the flooding caused by the network inefficient broadcasting, several clustering algorithms are designed. Clustering is a technique to group the nodes that have similarities according to some criteria together. Each algorithm has specific criteria that play a critical role in building a stable cluster. Cluster stability is a very important factor that most of the clustering algorithms attempt to achieve. The stable cluster is the cluster that does not change its cluster head very frequently as the topology of the network changes [12].

### **1.7.1 Benefits of Clustering**

A good clustering algorithm has many benefits [13] such as:

- Reusing the network resources: when two clusters are not adjacent, they can allocate the same frequency or code set.
- Conserving communication bandwidth and reducing transmission overhead: the clustering process eliminates the unnecessary exchange of messages among the nodes.
- Aggregation of topology information: as the number of nodes in the cluster is smaller than the number of nodes in the entire network, then each node is only required to store a small portion of the whole network routing information.
- Efficiency and Stability: the network with clusters seems to be more stable. For example, when a node leaves its cluster, only the corresponding clusters will update their data structures and the other clusters will be kept without any change.

## **1.8 Thesis overview**

The main objective of this thesis is to design a complete clustering protocol that is suitable for Hybrid Sensor Vehicular Networks by considering all the aspects that influence these networks.

The thesis report starts with the literature review in chapter 2 which discusses several existing clustering algorithms, their classifications, and a comparison between them. Chapter 3 includes the proposed protocol design, the problem statement, the methodology, the design challenges, the new features, the protocol operation, and finally the throughput calculations. Then, chapter 4 presents the simulation, the results found, and the discussion of the results. The last chapter is the conclusion of the thesis and the suggested future works.



## CHAPTER 2

### Literature Review

There are several clustering algorithms that were designed for both MANETs and VANETs. Each algorithm has different criteria that play a critical role in building a stable cluster. Some of these algorithms build their clusters according to a single metric while others design it considering multiple metrics.

In this chapter, we divide the clustering algorithms into different classifications. These classifications depend on how the nodes are grouped into clusters and how the cluster head is selected.

#### 2.1 Classification and Comparison

Clustering algorithms are classified based on the criteria for the selection of the cluster heads into different categories:

##### 2.1.1 ID-based clustering

In the Lowest ID (LID) algorithm each node is assigned a distinct ID that is distributed randomly among the nodes. A cluster head is selected based on the ID value, a node with the minimum ID is elected as a CH. LID shows a better throughput when compared with the highest degree algorithm which depends on the maximum number of neighbors (more details are in the next section). However, the highest degree algorithm is better than LID in terms of the number of clusters which are fewer in the highest degree scheme [14]. The main drawbacks are: first, considering the lowest ID as a parameter to distinguish between the ordinary nodes and the CHs is a weak decision; the reason is that the ID is distributed randomly among the nodes and does not represent any valuable information about the nodes.

Moreover, the node that is selected as a CH will run out of energy earlier than other nodes due to its role as cluster head that needs to send and receive long distance messages from cluster members, besides data aggregation of collected data.

### **2.1.2 Nodal degree-based clustering**

The first algorithm that is built according to the nodal degree is presented in [14]. The degree of a node is calculated based on its distance from other nodes. The node with the maximum number of neighbors is chosen as a cluster head. This algorithm is good due to the fact that the CH is not frequently being changed, but it has a low throughput compared to the LID algorithm. This algorithm does not specify certain number of nodes in each cluster. In other words, it has no limitations in the cluster size, which results in decreasing the throughput of the cluster.

Later on, in 2002 a connectivity based k-hop clustering (k-CONID) was proposed [15]. This algorithm combines two well-known clustering algorithms: LID and highest degree algorithms. It considers the connectivity degree as a primary criterion and the lowest ID as a secondary criterion in selecting the cluster head. A cluster consists of all the nodes that are at k-hops from the CH. The node that has the highest connectivity is chosen as a CH. However, if there are two nodes that have the same connectivity, then the node with the lowest ID is selected as a CH.

The cluster is updated whenever the network topology changes. There are four different cases that cause re-clustering. A node joining a network, a node leaving a network, an existing link is disconnected, and a new link is created. In the first case

when a new node intends to join the network, it checks if it is at a distance up to  $k$ -hops from any existing CHs, then it joins one of them. If not, it assigns itself as a new CH, creates its own cluster, and invites its  $k$ -hop neighbors to join its group. In the second case when a node leaves the cluster, there are two possibilities: if a node is an ordinary node then nothing will change in the cluster. However, if the node that leaves the cluster is a CH, then a new CH should be elected from the  $k$ -hop neighbors within the cluster. The third case happens when an existing link is disconnected. If the disconnected link exists between two nodes that belong to separate clusters, then no action is taken. On the other hand, if the link is between two nodes that are within one cluster, then the CH should check that all the nodes within the cluster are  $k$ -hop neighbors. If not, then nodes with more than  $k$ -hops will be disconnected and should form another cluster. For the last case when a new link is created between two nodes in two different clusters, there will be different cases that should be considered. For example, if the two nodes are not CHs, then nothing will change. But, if both nodes are CHs, then according to some criteria one node will be a CH and other nodes within  $k$ -hops distance will join the cluster.

When comparing this algorithm with the lowest ID algorithm, K-CONID algorithm shows better results than lowest ID in case  $k=1$ , but when  $k=2$  the difference between the two algorithms is small.

Maglaras and Katsaros proposed in [16] a clustering algorithm that is based on the force-directed methods. Every node applies a force to its neighbors based on their distances and velocities. If the total magnitude of the forces that are applied to a vehicle is negative, then this means that a vehicle is moving away from the surrounding vehicles. On the other hand, if the total magnitude of the forces that are applied to a vehicle is positive, then this means that a vehicle is moving in the

same direction and towards the other vehicles. The total force is used in order to select a suitable CH. The node that has the largest number of positive neighbors is selected as a CH. The size of the cluster is 2-hops at most.

In cluster maintenance phase, a new node that intends to join a cluster, checks the total force applied on it with respect to the total force applied to the CH. If the CH has a higher force, then it accepts the node as a cluster member. Otherwise, the new node tries to find nearby free nodes to form a new cluster. On the other hand, if a member node finds at any time that it has a higher force than any of the surrounding CHs; it should assign itself as a free node and try to form its own cluster. In addition, if two CHs become within the same range, then the one that has a lower force will give up its CH role to the other CH and become a member.

This algorithm was compared to the lowest ID algorithm and stationary local peer group architecture (LPG), and it was found that the proposed algorithm has smaller average number of clusters changed than the other two algorithms. Moreover, the total number of clusters formed is less compared to the lowest ID. Also, the average cluster lifetime is higher in the proposed algorithm than in the lowest ID.

Moreover, a Hierarchical Clustering Algorithm (HCA) is proposed in [17]. This algorithm handles channel access and schedules the transmissions in the cluster in order to have reliable communication and avoid collisions. The transmissions within the cluster are only allowed by the CH which assigns a specific slot for each member to transmit. The algorithm does not depend on the GPS to know the node's location; instead it gathers the connectivity from the messages sent. Three levels of hierarchy are defined: slave, cluster relay, and

cluster head. The slave node is the regular node, the cluster relay is the node that forwards messages from the CH to the slave nodes, and the CH node is the node that manages and coordinates the shared channel access. The control messages that are used are the SYNC messages which are created by the CHs and the ACK messages which are created by the slave nodes. The algorithm has four phases: the first three relates to the cluster formation and the fourth one relates to the cluster maintenance.

The main drawbacks of this algorithm are: The CH's election process is not perfect; this is because they select a slave node that receives two SYNC messages from two cluster relays as a CH and the other important factors such as: cars' velocities, locations, etc. are neglected. Moreover, the authors assumed that the first three phases of the algorithm take place while the nodes are static; obviously this assumption is not realistic, since it ignores the mobility of vehicles in VANETs.

### **2.1.3 Mobility-based clustering**

A Weighted Clustering Algorithm (WCA) [18] is one of the first clustering algorithms that take into account the mobility parameters. It considers several parameters such as: the transmission power, nodes' mobility, ideal node degree, and battery power of the nodes. Each node is supposed to calculate its combined weight and the node that has the minimum weight is elected as a CH. This algorithm limits the cluster size to achieve load balancing by predefining a specific threshold that a CH can handle in order to facilitate the medium access control protocols operations. The cluster head election process is performed on-demand and not periodically to save the computation and communication costs. The major drawback of the WCA is that it uses the global minimal concept in order to form

the cluster. It means that all the nodes in the network has to know the weights of all other nodes to start the clustering process which consumes a lot of time and network resources.

MOBIC [19] which is mobility based metric for clustering in mobile ad hoc networks, is another algorithm that is based on the mobility. This algorithm is similar to the lowest ID clustering algorithm in execution and the difference is only that the MOBIC uses mobility metrics to form the clusters instead of the ID information. All nodes in the network send and receive hello messages from other nodes. Then, each node measures the received power levels and calculates the relative mobility metrics. After that, each node will have a neighbors' list that contains the relative mobility value for each neighboring node. Then, the node with the lowest mobility value is selected as a CH. Moreover, if two nodes have the same mobility values, then the node with the lowest ID is chosen as a CH.

The simulator used to check the performance of the proposed algorithm is NS-2. The results show that this algorithm has few improvements on the lowest ID algorithm. However, it has some limitations in implementation. For example, MOBIC algorithm is good only in the moderate and high transmission ranges and performs worse than the lowest ID in the short transmission ranges.

In [20], a mobility based d-hop clustering algorithm for mobile ad hoc networks is proposed. In this algorithm nodes with similar mobility are grouped together in one cluster. The diameter of the cluster is not restricted to two hops, but it is flexible and can be determined by the stability of the cluster. The distance between the nodes can be measured based on the received signal strength. From the variation of the estimated distances, the relative mobility pattern between the



nodes can be found using some statistical testing. Initially, the nodes are grouped into 2-hop clusters based on their mobility pattern, and then the cluster sizes are expanded by merging two or more clusters based on the same metric.

Each node in the network starts periodically by broadcasting a hello message that includes its local stability value which is the standard deviation of relative mobility values of all neighbors that is calculated from the signal strength of the received messages from other nodes. So, after the discovery period, each node will have a complete picture about all its neighbors. Then, after the assignment period, each node compares its stability value with its neighbors' values. Then, the node with the lowest stability value is elected as a CH. Then, in order to characterize the movement of its corresponding group members, each CH computes the estimated mean distance for the cluster. In case there are two nodes with the same stability values, the election of the CHs are postponed for certain back off period. After the back off period, the local stability value is recomputed to find the most suitable CH.

After the discovery stage, all the clusters are only 2-hop clusters. So, the merging stage comes after the discovery stage in order to expand the cluster size. So, when a non-clustered node initiates the merging process, it starts collecting samples for estimated distances between the nodes. Then, it computes the mean of the estimated distance and the relative mobility. Finally, it joins the cluster that has a neighbor with the lowest stability value.

In the maintenance stage, the network topology changes because either a node joins or leaves a cluster. If a new node joins a cluster, then the merging process takes place. On the other hand, if a node leaves a cluster and this node is a CH,

then the immediate neighbor of the CH initiates the discovery stage to elect a new CH.

The performance of this algorithm is tested using NS-2 simulator. It is compared with the lowest ID and MOBIC algorithms and found that it outperforms them in terms of the number of clusters formed.

A weight based adaptive clustering algorithm (WBACA) [21] is proposed in order to overcome the drawbacks of the WCA algorithm. WCA algorithm finds the global minima of the weights, whereas the WBACA finds the local minima of the weights. It depends on the transmission power, transmission rate, mobility, battery power and the nodal degree in forming the cluster. Each node is assigned a weight that represents its suitability to be either a CH or a member in the cluster. The node with the smallest weight is elected as a CH. This algorithm does not allow two CHs to be one hop neighbors of each other. Also, all member nodes are one hop from their CHs. Moreover, nodes that are connecting two clusters are called gateways.

Each node periodically broadcasts a hello message to its neighbors. The message contains its ID. Once the hello packets are received from all neighboring nodes within the node's transmission range, the neighbor's table is created and broadcasted to other nodes. Also, each node calculates its weight value and sends it to all its neighbors. A join-request message and a join-Ack message are used to form the cluster.

The algorithm is tested using a GloMoSim. IEEE 802.11 is used as a MAC layer and AODV is selected as a routing protocol. The mobility model used here is the random waypoint model. The experimental results show that the proposed



algorithm performs better than both the lowest ID and WCA algorithms in terms of number of the clusters formed. On the other hand, the WBACA outperforms the other two algorithms in terms of starting time delay since it depends on the local minima concept.

Hwang et al proposed a combined weight-based distribution clustering algorithm in [22] and a hierarchical structure that aims at maintaining a stable network topology. The algorithm has three different phases, namely: initialization, cluster setup, and cluster maintenance. The initialization phase is executed at the beginning of the system and at cluster reconfiguration due to the battery exhaustion of the nodes.

This algorithm starts by defining the nodes' ID and location coordinates. Then, the node degree within the transmission range is found. After that, the sum of the distances for each node with all its neighbors is calculated. Then, the average speeds for all nodes are computed. Finally, the combined weight for each node is calculated and the node with the smallest weight is selected as a CH.

Each node sends a hello message which includes its weight to its neighbors. Then, within certain time, each node will have a set consisting of the nodes' numbers and their weights. After setting the clusters, each member in the cluster sends a periodic hello message to the cluster head to manage the topology.

The topology change has been controlled through the maintenance phase. After the CH has been selected, the CH chooses a node from its neighbors with the smallest weight as a pre-defined CH. This node is selected in order to share the responsibility with the CH to have a load balancing in the network. In case a member node is leaving the cluster: if the node can connect to the CH through the

gateway, then there is no change in the cluster topology. But, if the node went outside the cluster and it cannot connect to the gateway, then the node should perform a re-clustering process. Such practice affects the configurations of other clusters and lowers the overall stability of the clusters.

This algorithm was tested using the GloMoSim (Global Mobile Information System Simulator) with uniform random distribution created by random way point model. The results show that the proposed algorithm maintains similar number of clusters although the speeds of the nodes are not fixed. Moreover, the algorithm has a small overhead when compared to WCA and WBACA algorithms. The reason behind this is that the proposed algorithm uses information of its neighbors only and neglects the information related to outside of its transmission range, i.e. local minima.

In 2008, a new clustering protocol for mobile ad hoc networks was designed [23]. In this protocol, clustering has been performed in two phases: cluster formation and cluster maintenance. Each node has four possible states: NORMAL, ISOLATED, CLUSTER-HEAD, and GATEWAY. In the network, each node is assigned with a unique ID, and each node maintains a neighbor table that consists of node's ID, cluster's ID, the status of the node, the combined weight, and the transmission power. There is another table which is named CH neighbor table. This table contains information about the other neighbor CHs. In this algorithm, each node sends a periodic message to notify its neighbors about its presence. Then, each node builds its neighbor table according to the received information. The election process starts when the topology is stable. The node with the largest weight is selected as a CH.

At the beginning, all the nodes are in the ISOLATED state. To start the clustering formation, each node sends a LIVE message that contains the node ID, the CH ID which is initially equal to -1, the weight group where each group has different priority, and the state of the node. When receiving this information, each node recalculates its weight.

In order to elect a CH, the nodes should wait for a predetermined period of time  $T_e$  to start electing a CH. During this period, different situations could happen:

- If an ISOLATED node receives a LIVE message from any CH node before the  $T_e$  period, then the node should change its CH ID field to the new CH ID and change its state to NORMAL.
- If an ISOLATED node does not receive a LIVE message from any CH node after the  $T_e$  period, it searches the neighbors' table for a node that has the largest weight value and elects it as a CH.
- If it does not find in its neighbors' table a node with a higher value than its value, then it elects itself as a CH and changes its state to a CLUSTER-HEAD.
- If a node receives a LIVE message from multiple CHs, then it should choose any of them as a CH.

The CH broadcasts a periodic LIVE message every  $T_i$  seconds within its cluster in order to maintain the cluster. Also, the cluster members are supposed to send a LIVE message every  $2T_i$ . If cluster members do not receive a LIVE message from the CH after  $2T_i$ , then they realize that the CH does not exist anymore and the cluster election should resume.

To test the performance of the proposed algorithm, NS2 simulator was used. The results showed that the proposed algorithm outperforms WCA in terms of the average number of created clusters, the connectivity of the nodes, and the overhead.

In [24], a Distributed Mobility-Adaptive Clustering (DMAC) is presented. At the initialization phase, a node decides its role by comparing its weight with its neighbors' weight, and if it finds a node with a higher weight than its weight, then it joins this cluster as a member. Moreover, there are two other phases that are running if a link fails or a new link is added. In both cases, the node always checks its neighbors and changes its role according to the surrounding neighbors.

There are some drawbacks in DMAC algorithm as it is a general clustering framework that neglects some important points; for example, the nodes do not update their weights which lead to consuming the cluster head's energy as it is not updated. As a result, a modified version of DMAC is proposed in [25]. This algorithm tries to avoid the re-clustering when the nodes move in different directions by using periodical Hello messages that estimate the connection time. So, if two nodes move within the transmission range and in opposite directions, it does not run the clustering process as the time is very short. Also, sending the Hello messages periodically helps in updating the information of the neighbors. Although this algorithm looks efficient, the simulation is not realistic. This is due to using the lowest ID or the maximum degree criteria to select the CHs and it does not use the mobility parameters for simulation.

A new algorithm for clustering formation in VANETs based on the lane detection is proposed in [26]. It assumes that each vehicle is embedded with a

digital street map and a lane detection system in order to know its exact lane. The algorithm is designed for urban scenarios with intersections. The cluster head is selected in the lane that most of the traffic flow in. Together with the lane information, each vehicle computes a cluster head level which depends on the network connectivity level, the average distance level, and the average velocity level to select a cluster head. The vehicle with the highest level is elected as a CH. The main drawbacks of this algorithm is that it considers only one traffic direction and the cluster formation is performed only for the vehicles that move with the majority of the traffic flow. Moreover, the cluster formation is processed periodically, which increases the overhead and wastes the network resources.

A Type-based Cluster-forming Algorithm (TCA) is proposed in [27] to minimize the number of cluster head updates in an emergency. The cluster formation is based on the type of the node; where all the nodes that belong to one type are grouped together. In this case, three different types of groups are considered: a rescue group, a fire-fighting group, and a paramedical group. Each node sends a periodically hello message to its neighbors that contains its ID, IP (for routing and distinguishing node's type), status (ordinary, gateway, and CH node), location, and stability factor. The stability factor relies on: the relative mobility of nodes with respect to each other, the average sum of distances, the connectivity degree, and the remaining battery power. The node that has the lowest stability factor is more likely to be a CH. Each CH within its cluster reassigns new IDs for each member according to their stability values; a node that has the highest stability value is assigned the highest ID and so on.

The simulation includes 50 nodes in 100 X 100 unit area; where in group 1 and 2 there are 20 nodes in each, and 10 nodes in group 3. The proposed algorithm is

compared with LID and WCA algorithms. The results show that the TCA outperforms the other two algorithms in terms of the average number of CHs; i.e. the proposed algorithm has a stable number of CHs when the velocity is varied. Moreover, the proposed algorithm has a lower cluster update rate than both LID and WCA algorithms.

A dynamic clustering algorithm (DCA) based on mobility metrics for VANETs is proposed in [28]. This algorithm depends on the spatial dependency. It captures the similarity of the mobility of two nodes within the transmission range. Two main mobility metrics are found for each node; the average velocity and the average acceleration. Each node broadcasts its mobility information to its neighbors by Hello message, and then a node finds its relative velocity and relative acceleration with its one hop neighbors. The authors define the spatial dependency as the multiplication of the node's relative velocity and relative acceleration. After that, each node finds its total dependency and the cluster relation which is defined as the average total dependency of all its neighbors. The higher value of the cluster relation indicates that the node has more similar mobility properties with its neighbors. Also, the node that has the highest cluster relation value among its neighbors is considered as a CH.

The performance of DCA is tested using the NS-2 simulator, and it is compared with the lowest ID and maximum degree algorithms. The DCA outperforms the other algorithms in terms of the CH lifetime and the number of cluster re-affiliations. Although this algorithm seems to be perfect, the performance comparison is weak as it is compared with two of the simplest algorithms.

A novel algorithm to form stable clusters in vehicular ad hoc networks on highways is presented in [29]. The cluster formation in this algorithm is limited to the vehicles that travel in the same direction. Additionally, the speed difference between the vehicles is the main key criterion to form stable clusters. Since the vehicles move in different speeds, the network is divided into different clusters in a way that the members of each cluster are stable with each other. In other words, the fast vehicles are grouped in one cluster and the slow ones in another cluster. Each vehicle in the network is supposed to classify its neighbors into stable and non-stable neighbors. This can be achieved by periodically broadcasting the velocity information to all neighbors. Then, by comparing the velocities of the vehicles, the slowest vehicle among all its neighbors within the transmission range should initiate the clustering process. All its neighbors that have a velocity less than the threshold velocity are grouped in one cluster, and the remaining vehicles reinitiate the process following the same procedure to create other clusters. The cluster head is selected by computing the suitability value. This value is calculated based on the mobility information. Each vehicle should find how close its position is to the mean position of its entire stable neighbors, and how close its velocity to the mean velocity of all its stable neighbors. The node with the highest suitability value is elected as a CH.

The cluster maintenance phase is initiated whenever the topology of the network changes. The topology could be changed by three different scenarios: a new vehicle joins a cluster, leaves a cluster, or two cluster heads come close to each other. In the first scenario, the new vehicle that intends to join a cluster should check its relative speed with the CH if it is within the threshold. In the second case, the member vehicle that leaves the cluster must be removed from the cluster



members' list, and it should find another nearby cluster to join it. In the last case, if two CHs come within the proximity of each other and the transmission ranges and their relative speeds are within the threshold, then a merging process is initiated. In this process, a CH that has less number of members waives its role to the other CH and become a member in the new cluster.

The performance of this algorithm was tested using C++ simulator. The performance metrics are: the cluster stability, average cluster lifetime, number of clusters that are created, and the clustering overhead. This algorithm was compared with two other algorithms which are: the weight based algorithm and the position based algorithm. It is noticed that the proposed algorithm has the smallest number of cluster changes, and has the highest cluster head lifetime. Moreover, the proposed algorithm has the smallest cluster formation rate.

#### **2.1.4 Direction-based clustering**

The direction based clustering algorithm C-DRIVE is proposed in [30] and [31]. It depends on the direction of the vehicles after passing the intersection. There are three different directions in each intersection: straight, right, or left. In each direction a cluster is formed. Before the intersection point, each vehicle sends a hello message to check if there is any existing cluster for a particular direction. If it receives a reply from a CH for its desired direction, then it joins that cluster. Otherwise, it announces itself as a CH for this direction. Once the cluster head is selected, it should compute the density of the vehicles in its cluster and then sends it to the existing infrastructure. Here, the CH is assumed to be at the front of the cluster. Moreover, each vehicle in the cluster should periodically send a live message to confirm its presence in the cluster.



The authors in this paper highlighted an important point which is the overtaking in the cluster. In details, if a member vehicle moves with a speed higher than the CH's speed, then this vehicle has a chance to leave the cluster before the CH sends the information about the density of the cluster. As a result, this information is not accurate due to the overtaking of a vehicle. So, in order to solve such inaccuracy due to overtaking, a CH switching mechanism is proposed. Each vehicle in the cluster should constantly measure its relative distance with respect to the CH. Then, if a vehicle intends to overtake, it should send a message to the CH to inform it about its overtaking. Then, a CH checks if a vehicle can act as a CH till the intersection, it switches its role to the overtaking vehicle. So, the overtaking vehicle can handle transferring the density information. On the other hand, if the overtaking vehicle cannot take the role of the CH because of its high speed, then the original CH will reduce the density count due to the passing vehicle and no effect will exist in the cluster.

In this algorithm, the election of a CH is not perfect, since it only depends on the location parameter. Moreover, the goal of the clustering is limited to measuring the vehicles' densities in each direction of the intersection without considering the other clustering benefits such as reducing the messages exchanged. Also, the lifetime of the cluster is very short, since the cluster is destroyed after passing the intersection point.

In order to reduce the cluster head changing that is due to the vehicles' varying speeds, a modification of C-DRIVE algorithm is presented in [32]. The goal of the proposed modification is to implement a new cluster head election policy. The authors defined some imaginary points along the street before the intersections. These points are the starting point, the ending point, and the threshold point. The

start point is the point where the cluster formation is initiated whereas the end point is the point where all the clustering process and maintenance is terminated. On the other hand, the threshold point is the point that lies between the start and the end points and at this point the cluster head election process is stopped. Moreover, the threshold distance is defined as the optimal length of the cluster area. This distance is important for having a perfect cluster formation and cluster head election. This value allows the vehicles to be in the same cluster until reaching the intersection regardless of their speeds. In addition, this allows the cluster head to continue serving the cluster through the cluster lifetime.

The simulation showed that the proposed modification gives a better result than the original C-DRIVE solution in terms of the number of cluster heads changes and the overall overheads.

### **2.1.5 Leadership duration-based clustering**

The clustering algorithm in [33] takes into consideration a directional data, leadership duration and the ID value. The leadership duration means the latest CH relationship duration for a node. The longer leadership duration indicates more chance for the node to be more stable. The node is elected as a CH if it has the longest leadership duration, in other words, more experience. However, in case there are two nodes that have the same leadership duration, then the one with the lowest ID is selected as a CH.

If a member node loses its CH, then it looks for a nearby CH that satisfies the CH election conditions and joins its group. Otherwise, it elects itself as a CH.

The drawback of this algorithm is that it depends on the leadership duration, which is inefficient in case none of the nodes have ever been as a CH. So, in this case their leadership duration is equal zero, and the election will only be built according to the lowest ID.

#### **2.1.6 Path loss-based clustering**

A new clustering algorithm that is based on the physical constraints of the channel is proposed in [34]. It considers two of them, the path loss which is defined as the distance between the vehicle and the base station; the larger the distance between the vehicle and the roadside unit, the greater the path loss and accordingly the weaker the signal. The second physical constraint is the interference between the vehicles because of the overlapping.

When a new node wants to join a cluster, it looks for a CH within its transmission range, if it finds one; it compares its own distance to the base station and the CH distance to the base station. Then, the one that has the shortest distance is elected as a CH. Another possible case is that if the CH leaves the cluster, then the same procedure is followed to select another CH. If the member node leaves the cluster, it searches for another nearby cluster to join it. Furthermore, if two clusters come closer, then the CH with the shortest distance is chosen as a CH.

The author specifies three cases where a cluster only consists of one node. The first case is when the node does not have any vehicle within its transmission range. The second case, if the nearest cluster is overloaded and cannot handle more vehicles. The last case, if next to the nearest cluster has a weak signal quality to communicate and exchange messages with the node.

In order to measure the performance of this algorithm, a real input data is gathered from the Traffic Wales data of Swansea-London M4 motorway. This data is used to simulate a realistic vehicular traffic. The scenario of this simulation consists of three unidirectional lanes. A comparison between the analytical and the simulated model shows that there is a high agreement between them. The results show that the CH changes increase during the peak hours of the day due to the huge number of vehicles. Also, during these hours the cluster size and the average number of clusters increase. Moreover, the proposed protocol is compared with the Vehicle to Roadside (V2R) communication system. In the V2R system, each vehicle in the network directly communicates with the roadside unit, whereas in the proposed algorithm, all the communications to the roadside unit is done through the CH. The results show that the V2R has a greater end to end delay than the minimal path clustering algorithm. As a result, packet dropping probability is more in the V2R than in the proposed algorithm.

## **2.2 Drawbacks in existing algorithms**

The existing clustering algorithms have several drawbacks that influence their overall performances. The main drawback points are listed below:

- Some algorithms depend on parameters that do not represent any valuable information about the nodes in electing the CH, for example the lowest ID algorithm and the highest nodal degree.
- Some CHs are elected for long periods, so they are subjected to the resources drainage.
- Many algorithms do not have limitations in the cluster size which results in decreasing the throughput of the cluster and increasing the

number of lost safety messages that would affect the drivers' safety on the roads.

- Several algorithms have unrealistic assumptions. For example, the HCA algorithm assumes that the first three phases of the algorithm takes place while the nodes are static; obviously this assumption is not realistic, since it ignores the mobility of vehicles in VANETs.
- Different algorithms have some restrictions in their implementations, for instance the MOBIC algorithm is good only in the moderate and high transmission ranges and performs worse than the lowest ID in the short transmission ranges.
- Some algorithms consume a lot of time and network resources in the clustering formation process, such as WCA algorithm.
- Different authors validate their algorithm's performance by comparing them with the simplest and the inefficient algorithms such as the LID and highest nodal degree.

### **2.3 Classification Summary**

A summary of existing algorithms is listed in tables 1 and 2. They are divided into two tables because of spacing purposes. The first table shows a summary of all the mentioned algorithms except the mobility based algorithms which are listed in table 2. The tables highlight the types of networks using clustering techniques, the clustering metrics, the CH election, the performance, the simulator used and the year of publication. The publication year is mentioned here to show that the clustering concept is not a new concept and the researches in this area grow yearly.

Table 1: Summary of existing clustering algorithms

Classification	Algorithms	Ad-hoc networks	Metric	CH election	Performance	Simulator	Publishing Year
ID	LID	MANET	ID	Lowest ID	Better throughput than the highest degree algorithm	NS-2	1995
Nodal degree	Nodal degree	MANET	Nodal degree	maximum number of neighbors	More stable clusters than LID	NS-2	1995
	k-CONID	MANET	Nodal degree	the highest connectivity	Outperforms LID	-	2002
	Force directed	VANET	Distance and velocity	The largest number of positive neighbors	Outperforms LID and local peer group architecture (LPG)	Custom simulator	2012
	HCA	VANET	Nodal degree	Highest connectivity	Outperforms the K-CONID	OMNET+, SUMO	2011
Direction	C-DRIVE	VANET	Location	front node of the cluster	-	-	2011
	MC-DRIVE	VANET	Location	front node of the cluster	Outperforms C-DRIVE	NCTUns	2011
Leadership	Leadership duration	VANET	directional data, leadership duration and the ID value	Longest leadership duration	-	SWANS+	2007
Path loss	Path loss	VANET	path loss	Shortest distance to the BS	outperforms (V2R) communication system	Motorway Traffic Simulator	2013



Table 2: Summary of existing clustering algorithms based on mobility

Classification	Algorithm	Ad-hoc networks	Metric	CH Election	Performance	Simulator	Publishing Year
Mobility	WCA	MANET	transmission power, nodes' mobility, ideal node degree, and battery power	node with minimum weight	outperforms LID and highest degree algorithms	-	2000
	MOBIC	MANET	Mobility	lowest mobility	outperforms LID	NS-2	2001
	mobility based-d hop	MANET	Mobility	lowest stability value	outperforms lowest ID and MOBIC	NS-2	2004
	WBACA	MANET	Mobility, battery & transmission power, degree difference, transmission rate	Lowest weight	Outperforms both LID and WCA	GloMoSim	2005
	combined weight	MANET	nodes' ID, location, node degree, average speeds	Lowest weight	outperforms WCA and WBACA	CloMoSim	2007
	new clustering protocol	MANET	node's ID, cluster's ID, node's state, transmission power	largest weight	outperforms WCA	NS-2	2008
	DMAC	VANET	node connectivity, mobility, energy level	largest weight	-	-	1999
	Modified-DMAC	VANET	node connectivity, mobility, energy level	largest weight	outperforms DMAC	JiST/SWANS++, Vanet MobiSim	2008
	lane detection	VANET	lane information, connectivity level, average distance and velocity levels	highest level	outperforms LID and highest degree	NS-3	2010

TCA	VANET	Mobility, distance, nodal degree, energy	Lowest stability factor	Outperforms both LID and WCA	-	2010
DCA	VANET	average velocity, average acceleration	highest cluster relation value	outperforms LID and highest degree	NS-2, VanetMobi Sim	2011
novel algorithm	VANET	Position, velocity	highest suitability value	Outperforms weight and position algorithms	C++	2012



## CHAPTER 3

### Protocol Design

This chapter includes the problem statement, the methodology, the design challenges, the expected impacts, the new features, the protocol operation, and finally the throughput calculations.

#### 3.1 Problem statement

This thesis aims at designing a complete clustering protocol that takes into consideration all the aspects that affect the clustering structure, such as:

- The messages' format.
- The communication and the medium access technologies (between the RSU and CH, between the CH and the vehicles members).
  - IEEE 802.11p (CSMA/CA).
- The clustering algorithm that:
  - Considers all the significant parameters in the network, such as the nodal degree, the direction, and the velocity.
  - Creates a balanced system by dividing the clustering process overheads between the vehicles in the network and the RSU.
  - Reduces the computational and the communication costs of cluster formation process by electing two CHs, one as the main CH and the other one as a standby CH.
  - Minimizes the collision in the system.
  - Increases the throughput of the whole network by finding an optimal size of clusters.

### 3.2 Methodology

The starting point to achieve the above objectives of this project is to understand the main aspects of the HSVN, its main components and main features. Understand the clustering concept, design the clustering formation process, consider the clustering maintenance once the topology changes, choose the clustering election criteria that would maintain a balanced clustering group that has a minimum collision, maximum throughput, and a stable cluster. In addition, it is crucial to study most of the existing clustering techniques both in MANETs and VANETs, recognize their differences, and highlight their weaknesses.

In order to find the optimal size of a cluster, an analytical study is required. The system should be mathematically modeled to relate the probability of collision and the throughput to the cluster size.

Moreover, a real traffic data for a specific road in Al Ain region is collected from the Department Of Transport (DOT) in Abu Dhabi. Then, this data is used to conduct efficient analysis that is based on real scenario. Once the analysis result is obtained, it will be used to test the proposed clustering protocol using the Matlab simulator. The devised protocol will be compared with similar clustering algorithms according to some criteria that are crucial to clustering techniques such as, the stability of the cluster, the communication overheads, throughput, and minimum single node cluster.....etc.

### 3.3 Design challenges

Designing a new clustering protocol with certain features is a challenge in terms of:

- What are the most important parameters that the messages' fields should contain?
- What medium access method is appropriate for this algorithm that would guarantee high rate of message delivery to the cluster head and the road side unit?
- Several scenarios affecting the clustering process should be considered, such as traffic lights, intersections, roundabouts, etc.
- The analysis of this protocol to find the throughput and the optimal size of the cluster.
- Choosing the proper simulator.
- Testing other clustering algorithms to have a fair comparison with our algorithm.

### 3.4 Expected Impacts

To the best of our knowledge, the clustering methods are considered and studied only in MANETs and VANETs; we could not realize any clustering algorithm that is designed especially for the HSVNs. So, our proposal can be considered one of the first approaches in this area. In addition, the main concern for most of the proposed algorithms is only to establish stable clustering algorithms; however, in our algorithm we consider a very important parameter which is the throughput of the cluster to control the density of the vehicles in a

cluster. Also, our proposed protocol divides the complexity of the clustering algorithm between the roadside unit and the cluster head.

This project will open the door for researchers to explore more in this area. Additionally, further efficient clustering algorithms can be developed based on this protocol.

### **3.5 Contributions & New features**

The main contribution in this project is to design a complete clustering protocol that take into considerations all the aspects that influence the vehicular networks. Also, in designing the clustering algorithm, a solid analysis is performed in order to define the boundary of the cluster size, which depends on the probability of collision and the throughput. Furthermore, it is critical to preserve the time and the resources consumed in the clustering process, and this is achieved by selecting two CHs for each cluster.

### **3.6 Assumptions**

The proposed protocol assumes the following:

1. Each vehicle is equipped with a GPS device in order to know its exact coordinates.
2. Transmission range of the RSU is 800m.
3. Vehicle length is about 4m and considering the safety distance between vehicles as 1m, then, the total length is 5m.
4. The transmission range of a vehicle is 400m.
5. The design's environment is an urban area.
6. The speed limits ranges between 60 and 100km/h.
7. The street is straight with three unidirectional lanes.

### **3.7 Protocol flowchart**

The flowchart represents the algorithm steps and their orders, as it is shown in figure 3.1 below. The details of the protocol are illustrated in the next section.

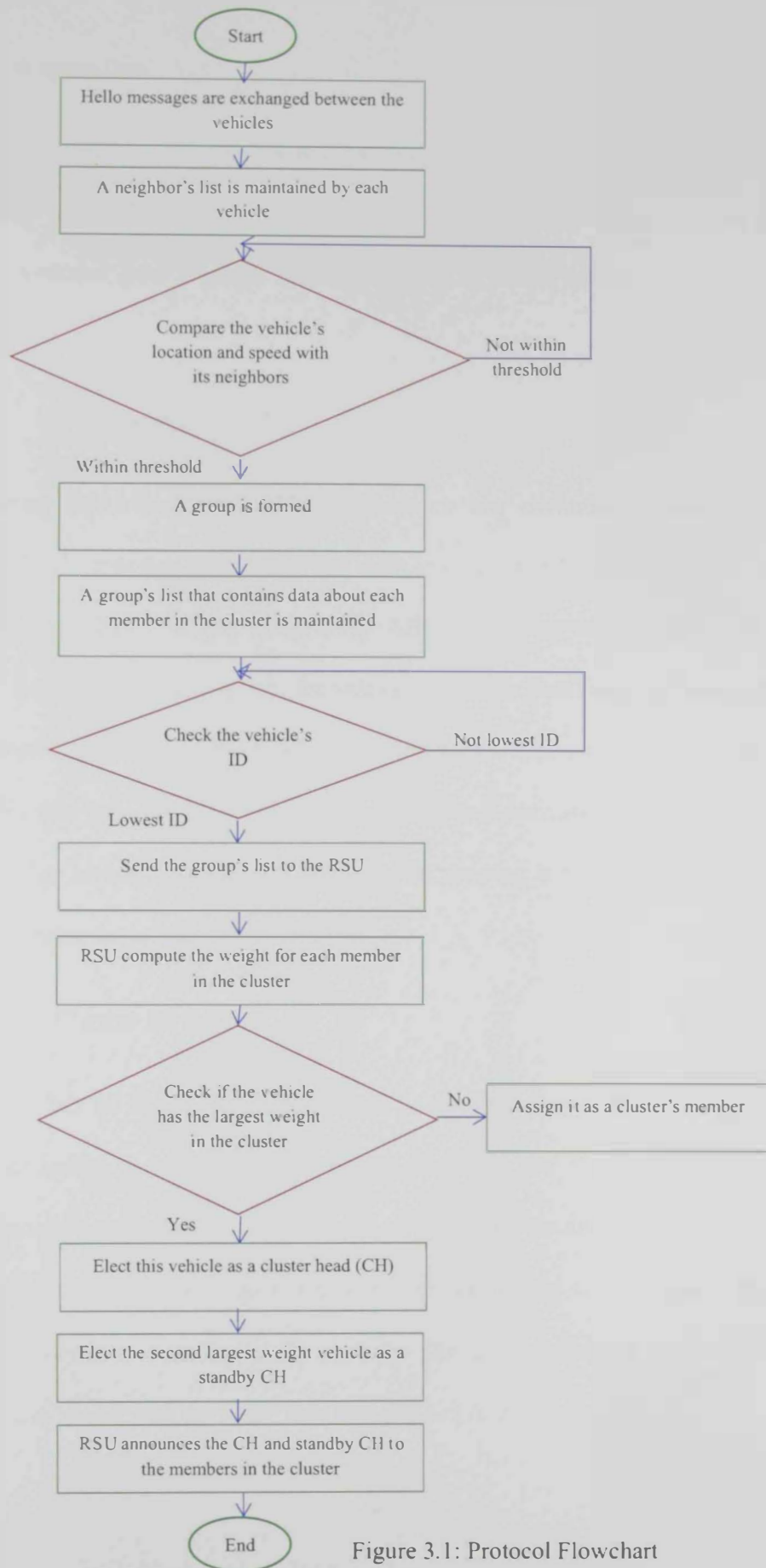


Figure 3.1: Protocol Flowchart

### **3.8 Protocol operation**

In this protocol, the road is divided into several zones and each zone is controlled by one RSU. So, the zone is defined by the transmission range of the RSU. The protocol goes through different stages in order to form clusters and maintain them.

#### **3.8.1 Initial stage**

During the initial stage, Hello messages are exchanged between the vehicles. A Hello message consists of the vehicle's ID, coordinates, speed and the nodal degree (total of two hop neighbors). After that, each vehicle maintains a neighbors' list. Then, in each group, the vehicle that finds itself with the lowest ID in the collected table should back-off for a certain time and transmits its table to the RSU. By this way, the RSU will have a complete picture about all the vehicles in its zone. The lowest ID vehicle is in charge of sending the table to the RSU in order to minimize the overheads in the RSU.

#### **3.8.2 Cluster formation**

The RSU has all the information related to the vehicles moving in its zone, which is defined by the transmission range of the RSU. So, the RSU divides the vehicles into different clusters according to their velocities and locations. As a result, all the existing vehicles at that moment are associated with a cluster. Now, each cluster needs a coordinator to manage the communication between the vehicles in the cluster and the RSU; this is explained in the next section.

### 3.8.3 Cluster Head (CH) election

A cluster head is the main coordinator in the cluster. Each cluster should have a cluster head to manage the communication between the vehicles' members and the RSU. But, the question is which vehicle in the cluster is eligible to be a cluster head among other vehicles?

Once the network is divided into clusters, the RSU computes a weight for each vehicle in the zone. The weight calculation depends on the total number of two hop neighbors, and the distance between the vehicle and the RSU. The vehicle that has the largest number of two hop neighbors and the shortest distance to the RSU is the most desirable vehicle to be elected as a cluster head for its cluster.

In case there are more than one vehicle that has the same weight, then there will be two possible scenarios: the first scenario is: if more than one vehicle has the same weighting factor and they are all candidates to be as a CH, then the RSU selects the vehicle with the lowest ID as a CH, the next lowest ID vehicle as a Standby, and the remaining vehicles as members. The second possible scenario is: if more than one vehicle has the same weight and want to join a cluster, but the cluster has only one free space, then the vehicle with the lowest ID will be accepted as a member. To summaries this situation, the proposed algorithm only relies on the vehicle's ID in case the vehicles' weights are equal.

Furthermore, the RSU, once it distributes the clusters informations, it deletes all the records of the clustering process except the CH and Standby CH data. This is necessary to conserve the RSU memory.



The weighting formula (1) is obtained according to some facts that are explained in details in the following paragraph.

$$W = \frac{Nd}{Nd_{Max}} \alpha_1 + \frac{1}{d/d_{Max}} \alpha_2 \quad (1)$$

Where:

$W$  is the vehicle weight,

$Nd$  is 2-hop nodal degree,

$Nd_{Max}$  is Maximum number of 2-hop nodal degree,

$d$  is distance between the vehicle and the RSU,

$d_{Max}$  is Maximum distance between the vehicle and the RSU,

$\alpha_1, \alpha_2$  are weighting factors,

In the above formula, the relation between the nodal degree and the weight is a direct relation, as the number of 2-hop neighbors increases that means the vehicle can communicate or can reach large number of vehicles which is a desired feature in the cluster head vehicle. Moreover, the distance between the vehicle and the RSU is inversely proportional to the weight. The reason is, when the vehicle has the smallest distance to the RSU, it will gain more weight, since the path losses and interferences will be less and this is preferred in selecting the cluster head node. Also, each parameter in the equation is divided by the maximum possible value in order to normalize it. Normalization is applied since we are dealing with parameters that have different units and scales. So, we need to normalize all the parameters under the same scale to have a fair comparison. In addition,  $\alpha_1$  and  $\alpha_2$

which are the weighting factors are assigned the same value which is 0.5 since both parameters (nodal degree, distance) have the same significance in selecting the cluster head.

The figure below illustrates the cluster formation and the main nodes that are contained in any cluster.

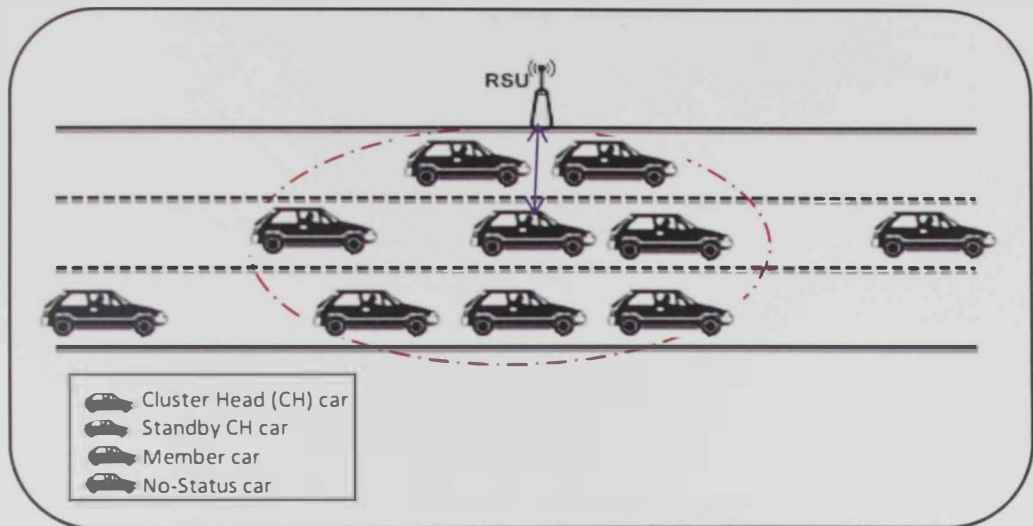


Figure 3.2: Cluster formation

### 3.8.4 Cluster maintenance

There are different scenarios affecting the clustering process and its stability.

These scenarios are summarized in the following:

- Scenario 1: if the CH leaves the cluster early due to speeding

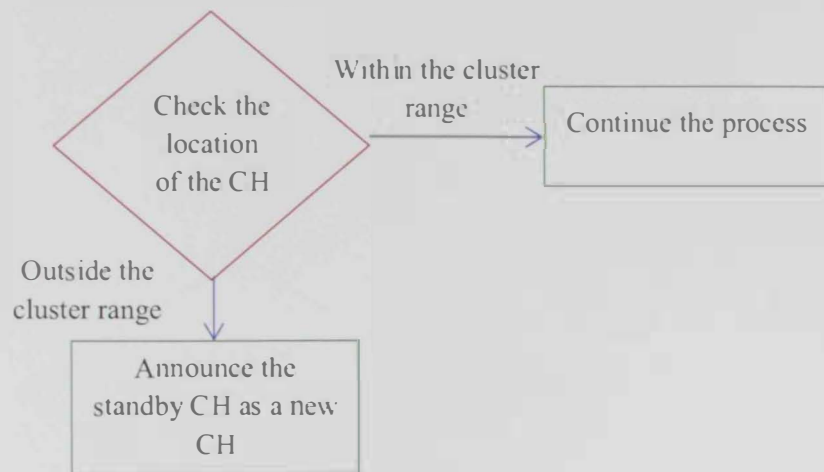


Figure 3.3: Flowchart for Scenario 1

In this scenario, if the CH moved outside its cluster range, then the standby CH should take over as a CH to the members and as a coordinator to the RSU.

- Scenario 2: standby CH leaves the cluster while the CH is operating in the cluster, or both CHs leave the cluster at the same time.

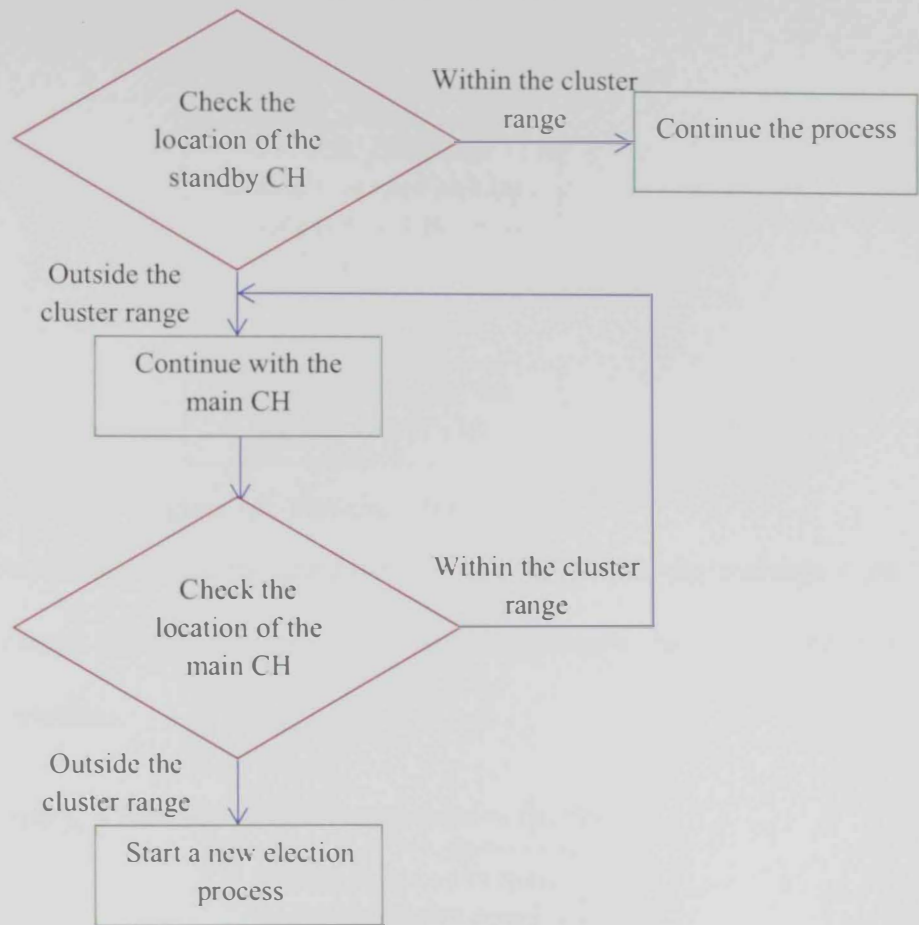


Figure 3.4: Flowchart for Scenario 2

In case the standby CH leaves the cluster, then nothing should be done as the main CH still exists in the cluster. However, if the main CH also leaves the cluster, then there will be no coordinator for the cluster. So, a new election process should commence.

Scenario 3: a new vehicle wants to join the cluster

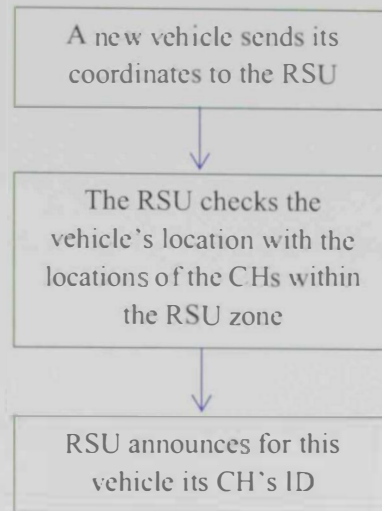


Figure 3.5: Flowchart for Scenario 3

A new vehicle sends its coordinates to the RSU by Hello message with higher transmitted power than the regular Hello messages that are exchanged between the vehicles.

- Scenario 4: a member vehicle wants to leave the cluster

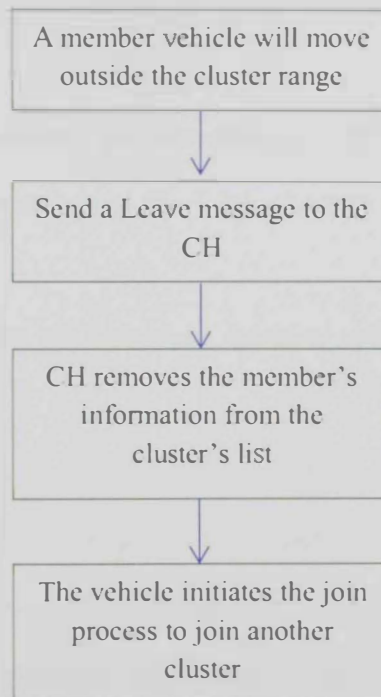


Figure 3.6: Flowchart for Scenario 4

### 3.9 Messages Exchanged

There are different types of messages exchanged through the clustering process. These messages are: HELLO message, JOIN-REQUEST message, JOIN-ACCEPT message, LEAVE message, and the Periodic CH-Data message.

#### a. HELLO message

Hello messages are exchanged between the vehicles at the start point of each zone. These messages are very critical in the clustering process since they consist of the most important parameters that identify the mobility of vehicles. The main fields of these types of messages are described in figure 3.7.

Vehicle-ID	X-coordinate	Y-coordinate	Direction	Speed	Weight	Cluster-ID	Status
------------	--------------	--------------	-----------	-------	--------	------------	--------

Figure 3.7: HELLO message fields

- a) Vehicle-ID: each vehicle has a unique ID that will distinguish it from other vehicles.
- b) X-coordinate and Y-coordinate: these coordinates specify the location of the vehicle and are obtained from the GPS which is embedded in the vehicle.
- c) Direction: this field shows the movement direction of the vehicle, either forward or backward.
- d) Speed: the speed field describes the speed of the vehicle.
- e) Weight: this field is initially equal to zero; but after exchanging the messages, the RSU calculates the weight for each vehicle. This field helps in defining the status of the vehicle, either CH or member.

- f) Cluster-ID: this field is initially equal to zero; then after defining the neighbors for each vehicle, it will have a value.
- g) Status: this field is initially equal to zero; then after defining the groups and finding the weight, it will contain the status of the node. The status can be a CH, a Standby CH, a member, or no-status.

- b. Periodic message

A periodic message will be sent by the RSU that consists of all the CHs data in its zone. This message will contain the CH-IDs, their coordinates, and the total number of their members. The benefit of this type of message is to help the new vehicles that enter the zone to find the nearest CH to join its group. Also, the members in each cluster will send periodically a message to notify the CH about their presence in the cluster.

- c. JOIN-REQUEST message

This type of messages is only exchanged if the vehicle loses its connectivity with its CH. The vehicle sends join request message to the corresponding CH after receiving the CH ID from the RSU, as it is explained in scenario 3. The main fields of JOIN-REQUEST messages are the same as the Hello message fields, except that the Cluster-ID and Status fields are equal to zeroes.

In the join request message, the cluster ID and status fields are zero until the vehicle receives the accept message from the CH, and then these two fields will have values. However, the weight field can have a value or it can be zero. This depends on the exchange messages. If the vehicle enters the zone at the initial stage where the Hello messages are exchanged between the vehicles, then the vehicle will have a weighting value. On the other hand, if the vehicle enters the region

after the clusters are formed, then the vehicle will exchange the messages with the RSU to select for it the corresponding cluster and the weight field is zero since it does not have any role in this stage. The main purpose of the weight value is to define the status for each vehicle in the cluster. However, the new vehicle is only accepted as a member in the cluster and not as a CH or Standby CH. The reason is to keep the system simple and not complicated with high number of exchanged messages.

d. JOIN-ACCEPT message

The JOIN-ACCEPT message is sent by the CH after receiving the JOIN-REQUEST message to the vehicle. But, it is mandatory that the cluster does not reach the maximum size in order to accept the new vehicle. The main fields of the JOIN-ACCEPT messages are the same as the Hello message fields.

e. LEAVE message

The LEAVE message is sent by the member that will leave the cluster to the CH in order to be removed from the CH records. Also, the CH sends a Leave message to the standby CH in order to take its role.

The average payload of the messages exchanged during the clustering formation is the total summation of the message fields' lengths:



Table 3: The message fields' size

Field	Size (byte)
ID	8
X coordinate	20
Y coordinate	20
Direction	14
Speed	15
Weight	20
Cluster ID	8
Status	20
Total	125

The total message length is 125 bytes which equals 1000 bits, and this is the average message payload. This message's size is regular in VANETs.

### 3.10 Throughput calculation

Throughput is a very important parameter that should be considered in network communications. It describes the number of successful messages that are delivered through the communication link [35].

In order to find the throughput of the system, Bianchi model was studied and implemented [36]. Bianchi model helps in finding the saturation throughput of the channel by using the two dimensional Markov chain process. It assumes that the channel is ideal by ignoring the hidden terminal problem and the capture effect. Briefly, two dimensional Markov chain process depends on the backoff stages, considering the minimum and maximum contention windows. The result of this model is a closed form expression of the saturation throughput which is derived from the probability of transmission, probability of collision, and other significant parameters. The resulting equations are as follow:

$$\text{Probability of collision: } p = 1 - (1 - t)^{N-1} \quad (2)$$

$$\text{Probability of accessing the channel: } t = \frac{2(1-2p)}{(1-2p)(W_{\min} + 1) + pW_{\min}(1-(2p)^m)} \quad (3)$$

$$\text{Saturation throughput: } \tau = \frac{p_s p_{tr} L}{p_s p_{tr} T_s + p_{tr}(1-p_s)T_c + (1-p_{tr})T_{slot}} \quad (4)$$

Saturation throughput can be defined as the average payload transmitted in a slot time over the average duration of the slot time.

The variables in the above equations are:

$p$  is the probability of collision

$t$  is the channel access probability of a node

$N$  is the total number of nodes in the system

$$W_{\min} = CW_{\min} + 1$$

$$CW_{\max} = 2^m W_{\min}$$

$p_s$ : is the probability of a successful transmission and it is equal to:

$$p_s = \frac{Nt(1-t)^{N-1}}{1-(1-t)^N}$$

$p_{tr}$ : is the probability that at least one transmission occurs in a slot and it is equal

$$\text{to: } p_{tr} = 1 - (1-t)^N$$

$L$ : is the average packet payload size

$T_{slot}$ : is the duration of a single slot time

$T_s$ : is the average time needed to transmit a packet of size  $L$ ,

$$T_s = H + L + SIFS + \sigma + ACK + DIFS + \sigma$$

$T_c$ : is the average collision time,  $T_c = H + L + DIFS + \sigma$

Where  $H$  represents the physical and MAC headers,  $ACK$  is the acknowledgment time,  $\sigma$  is the propagation delay,  $DIFS$  is the Distributed Inter Frame Space time, and  $SIFS$  is the Short Inter Frame Space time.

In order to find the values of  $p$  and  $t$ , the two nonlinear equations (2) and (3) were solved numerically using Matlab for different number of nodes. Then, the saturation throughput was found by evaluating equation (4). The result showed that the optimal number of vehicles in each cluster should not exceed 10 vehicles. The parameters set are listed in table 4 according to the IEEE standards as stated in [37].

Table 4: Throughput analysis parameters

Parameter	Value	
Data rate	3 Mb/s	
Slot time	13 $\mu$ s	
DIFS	58 $\mu$ s	
SIFS	32 $\mu$ s	
Propagation delay	2 $\mu$ s	
$CW_{min}$	15	
$CW_{max}$	1023	
Payload	1000 bits	333.33 $\mu$ s
PHY header	192 bits	64 $\mu$ s
MAC header	256 bits	85.33 $\mu$ s
ACK	304 bits	101.33 $\mu$ s

CSMA/CA has two access methods, either the basic access method or the RTS/CTS method. The basic access method is the method that sends the data packets directly without exchanging Request To Send (RTS) and Clear To Send (CTS) packets, as mentioned in chapter 1. In our analysis, a basic access method is implemented since the RTS/CTS method requires several handshakes which increases the overhead in the network and sequentially reduces the performance of the system specifically in high mobility networks such as vehicle networks [38].

All the previous algorithms that are discussed in chapter 2 do not consider the throughput of the clustering; whereas in the proposed algorithm, the throughput

is taken into account to specify the optimal number of vehicles in a cluster. The optimal size means finding the upper bound of the cluster size that has an acceptable throughput which does not affect the network performance. In addition, some of the existing algorithms set the upper bound of a cluster without highlighting the method of getting that value.

## CHAPTER 4

### Results & Discussion

In this chapter, the main results of the algorithm are presented and discussed.

#### 4.1 Simulation scenario

In order to check the validity of our algorithm, a real traffic scenario is implemented. The data is collected from the Department Of Transport (DOT) in Abu Dhabi. The area that is chosen is the school zone in Al Ain, specifically Khaled Bin Sultan Street (147<sup>th</sup>) as shown in figure 4.1, 4.2, and 4.3. The total length of the street is approximately 800m, and it has three lanes, each one of width 3.5m. Moreover, in our case, only the light cars were considered and the large vehicles were neglected since they only represent 5% from the total vehicles in the street as obtained from the DOT data. Also, the total number of cars that can be at the same time within the street length is 480 cars during the peak hours. In addition, the algorithm was simulated using Matlab 7.6.0 (R2008a),







Figure 4.3: Khaled Bin Sultan Street Snapshot (Source: DOT)

## 4.2 Simulation Setup

In order to start the simulation, a traffic data file is generated. First of all, each vehicle in the network; i.e. 480 vehicles is assigned with a unique ID. Also, the vehicles are assigned random speeds that range from 60 to 100 Km/h (16.66 to 27.77 m/s). In addition, each vehicle has a random (X, Y) coordinates. Initially, the weight values, the group IDs and the vehicles' status are all set to zero.

After the traffic data file is generated, the clusters' borders are defined by calculating the Euclidean distances and the speed differences between each point and all its neighbor nodes in the zone. An example of Euclidean distance calculation is illustrated in figure 4.4 which shows the maximum possible distance between the CH and a 2-hop neighbor. Then, if the Euclidean distance is less than 10.051 and the speed difference is less than 5.5556m/s, then the nodes are grouped in one cluster. Notice that we assume that the vehicle within 5.5556m/s (20km/h) speed difference can be in the same cluster.

Furthermore, the RSU is assigned a location in the middle distance of the zone at the roadside. As described previously, the weighting factor depends on the

number of neighbors for each vehicle and the distance between the vehicles and the RSU. So, the weighting factors of each vehicle are found in order to define the status for each vehicle. The status values could be zero which means that the vehicle has no-status yet, or one as a CH, or two as a standby CH, or three as a member vehicle.

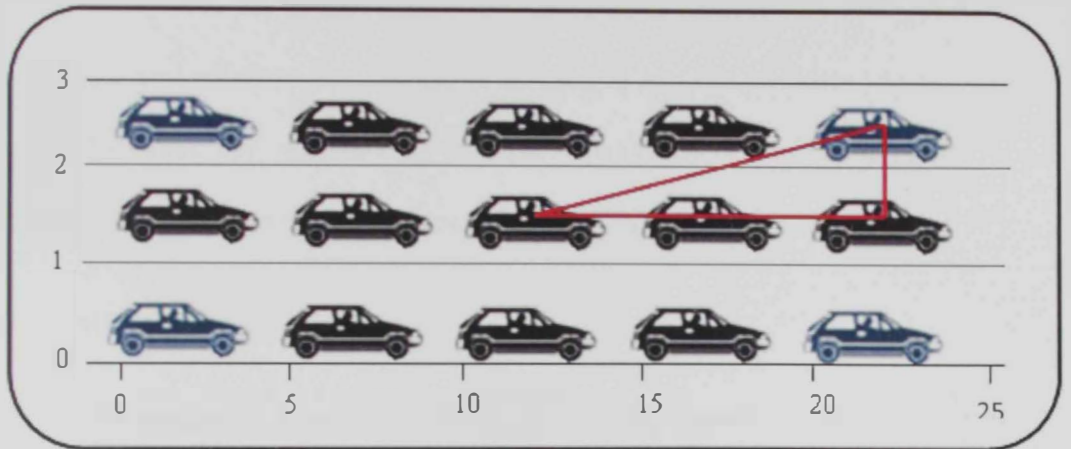


Figure 4.4: Euclidean Distance Calculation Example

$$\text{Euclidean Distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \sqrt{(12.5 - 22.5)^2 + (1.5 - 2.5)^2} = 10.05 \quad (5)$$

Moreover, according to the previous calculations in chapter 3, the optimal cluster size is 10 vehicles. So, in case more than 10 vehicles are eligible to join the cluster, then their weight factors should be compared. The vehicles that have the highest weighting factors are kept together in one cluster, and the remaining vehicles should join the cluster where they have the shortest distance to their CH. Moreover, if a new vehicle wants to join a cluster and it has a higher weight than the existing members, the method of accepting this vehicle is different. The method depends on the total number of members in the cluster, if they are less than the threshold value (10 vehicles) then the new vehicle is accepted as a new



member. On the other hand, if the cluster reaches its maximum size; then the vehicle should try to join another cluster.

### 4.3 Results

In order to evaluate the performance of the proposed algorithm, one of best existing mobility based algorithms (DCA) is implemented [28]. This way, a fair comparison between the proposed and the DCA algorithm is presented. Also, to have reliable and efficient results, the simulation is run for 10 times varying the vehicles' locations and speeds, and then the average of the collected results is taken. The simulation results are presented and discussed below:

#### 4.3.1 Total number of clusters

The total number of clusters that are created in DCA and proposed algorithms is illustrated as shown in figure 4.5.

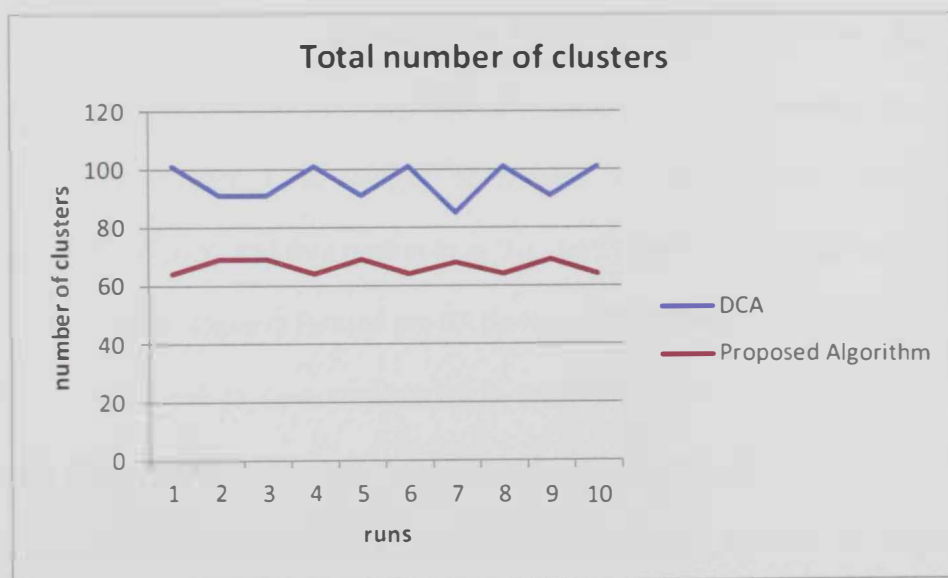


Figure 4.5: Total number of clusters in the proposed and the DCA algorithms

The average number of clusters formed within the zone is 95 clusters in the DCA algorithm and 66 clusters in the proposed algorithm, as shown in figure 4.6.

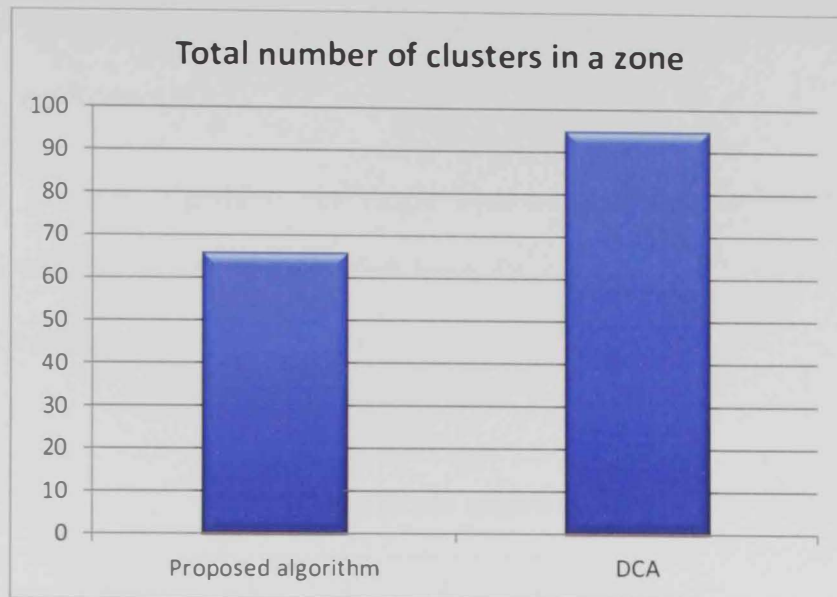


Figure 4.6: Total number of clusters in the proposed and the DCA algorithms

It is clear that DCA algorithm has a larger number of clusters than the proposed algorithm. This is because there is no upper or lower bounds that limit the size of the cluster. Less number of clusters means less contention in the medium, less collision and less packet loss. To clarify more, in the DCA algorithm the total number of clusters formed are 95 clusters with 480 cars in the zone, which means 95 CHs need to communicate with the RSU. On the other hand, in the proposed algorithm only 66 clusters are formed with the same total number of cars, which means only 66 CHs communicate with the RSU. So by this, the proposed clustering algorithm looks better than the DCA.

### 4.3.2 Single node clusters

The number of single node clusters is shown in figure 4.7, and the average number of single node clusters is 29 in DCA and 3 clusters in the proposed algorithm, see figure 4.8.

In the proposed algorithm, the single node clusters are eliminated to the maximum possible extent. On the other hand, DCA has a very large number of single node clusters.

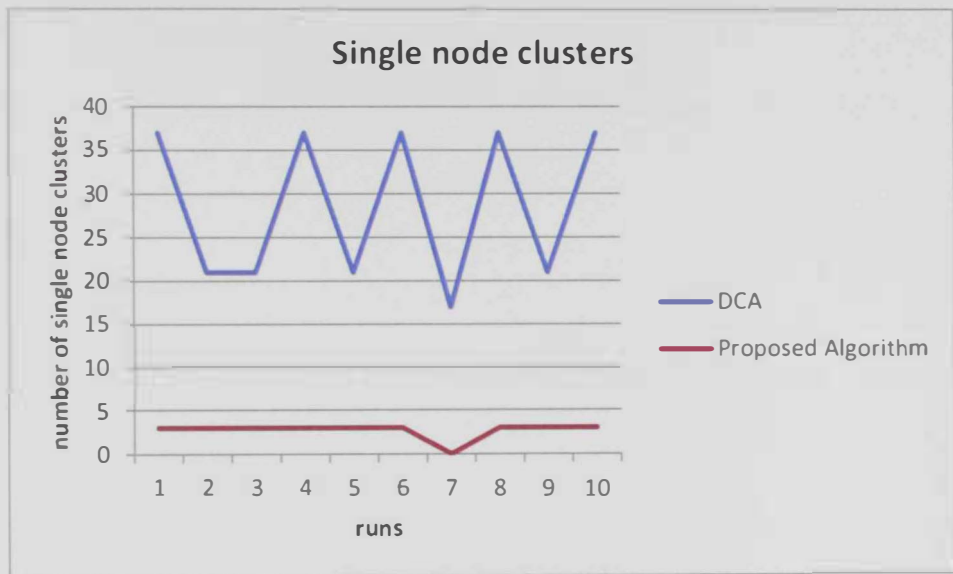


Figure 4.7: Single node clusters in the proposed and the DCA algorithms

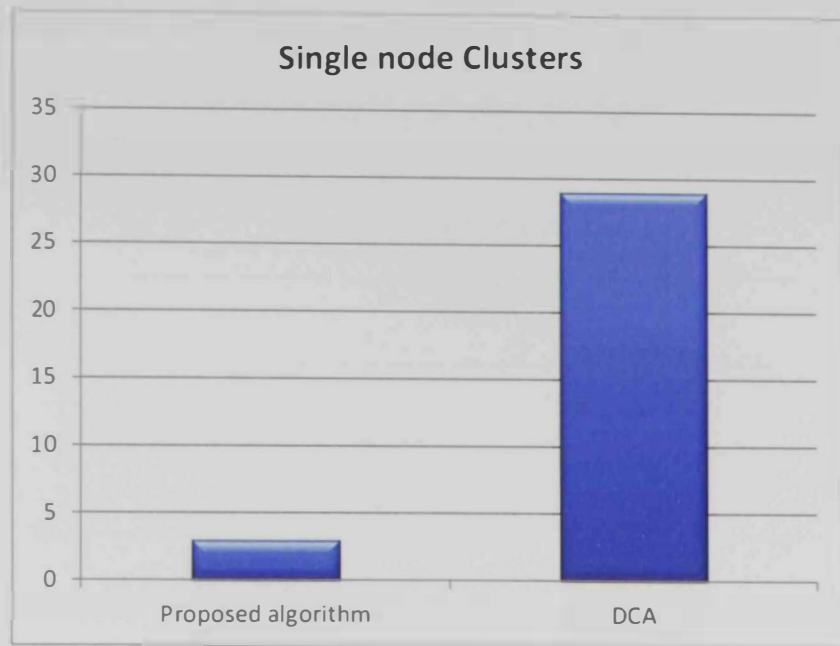


Figure 4.8: The average of single node clusters in the proposed and the DCA algorithms

### 4.3.3 Cluster size

The largest cluster size for the DCA and the proposed algorithms is shown in figure 4.9.

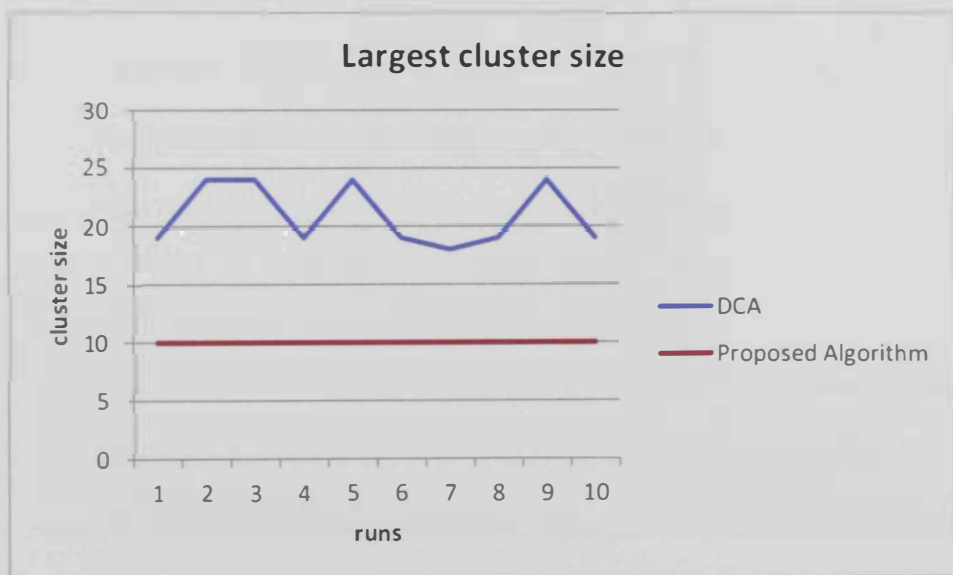


Figure 4.9: Largest cluster size in the proposed and the DCA algorithms

The average value of the largest cluster size in the DCA is 21 vehicles, where it is reduced to 10 vehicles in the proposed algorithm, see figure 4.10.

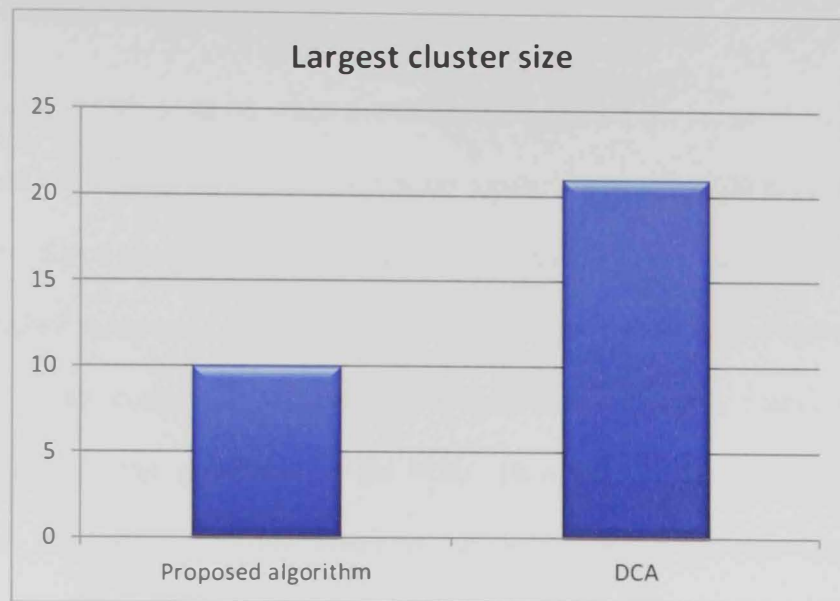


Figure 4.10: Largest cluster size in the proposed and the DCA algorithms

#### 4.3.4 Cluster lifetime

The cluster lifetime is directly linked to the cluster head lifetime. If the cluster head keeps moving in the same direction with the same velocity, then the cluster lasts for a long period. In the vehicular networks, since the vehicles' movements are restricted to the roads' structures and the speeds' limits; especially in the urban areas, then the cluster lifetime is long enough to exchange important data between vehicles.

In the proposed protocol, the cluster lifetime is related to the RSU zone. This is because the clustering formation process is only initiated at the beginning of each RSU's zone. It means that the cluster vanishes when the CH reaches the end of the RSU zone. So, if we assume that a CH moves with the maximum allowed speed which is 27.77 m/s, then the cluster lifetime is equal to 28.8 s.

$$Cluster\_Lifetime = \frac{RSU\_zone}{max\_speed} = \frac{800m}{27.77m/s} = 28.8s \quad (6)$$

#### 4.3.5 Standby Cluster Head

Standby CH is useful when the CH moves outside the range of its cluster. So, instead of running the clustering process again, the standby CH takes the role of the CH. Since the cluster formation process cause a delay in exchanging the normal safety messages, the existence of a standby cluster allow more safety messages to be exchanged without any discontinuation of the communication process between the vehicles and the RSU. However, in the other clustering algorithms, when the CH moves outside its region, the cluster formation process is initiated to elect a new CH. As a result, the communication overheads and the energy consumptions increase in the network. In order to highlight energy saving and the importance of having a standby CH in our clustering protocol, the communication overheads and energy consumption to form and maintain the cluster are calculated for the following cases: a cluster without Standby CH, and a cluster with Standby CH.

##### 4.3.5.1 Communications Overhead

The communication overheads depend on the total number of messages that are exchanged in the cluster formation and when the CH leaves its cluster.

Case1: Without Standby CH

Initially, Hello packets are sent between the vehicles. Then, join-request and join-accept messages are exchanged between the CH and the new vehicles to join a cluster. In case the CH moves outside its cluster's range, it sends a leave

message to its members and the cluster vanishes. Again, a new clustering formation process is resumed, and new hello messages are exchanged to form new clusters and elect new CHs.

#### Case2: With Standby CH

Similar to the first case, hello, join-request and join-accept messages are sent between vehicles to form the clusters. However, when the CH leaves its cluster, it sends a leave message to its standby CH, and the standby CH sends its ID to the members as a new CH. So, no new clustering process is needed in this case. By this way, the overheads of the communication messages that are sent in the absence of the CH are reduced approximately to half when compared with the first case.

#### 4.3.5.2 Energy Consumption

The energy consumption is calculated according to the energy model that is described in [39]. There are energy consumptions in transmitting and receiving messages. The equations below represent the transmitting energy consumption and the receiving energy consumption. Equation (7) describes the transmitting energy for the CH and RSU communications, whereas Equation (8) describes the transmitting energy for the CH and members' communications. Since the distance between the RSU and the CH is long, the two ray propagation model is assumed. On the other hand, the free space model is used for the communications between the CH and its members as they are very close to each other, so it is more practical to use the free space model. In addition, equation (9) represents the receiving energy consumption.



$$E_{Tx} = lE_{ELEC} + l\varepsilon_{2r}d_1^4 \quad (7)$$

$$E_{Tx} = lE_{ELEC} + l\varepsilon_{fs}d_2^2 \quad (8)$$

$$E_{Rx} = lE_{ELEC} \quad (9)$$

The parameters that are used in the calculations are listed in table 5.

Table 5: Energy calculations parameters

Parameter	Value
Radio electronics energy $E_{elec}$	50nJ/bit
Transmit amplifier energy (free space) $\varepsilon_{fs}$	10pJ/bit/m <sup>2</sup>
Transmit amplifier energy (two ray) $\varepsilon_{2r}$	0.0013pJ/bit/m <sup>4</sup>
Number of bits in the packets $l_{data}$	1000bits
Average distance between the cluster head and the RSU $d_1$	200
Distance from the member node to the cluster head $d_2$	10.05

- a. Energy consumption in the proposed protocol:

Energy consumption for one cluster formation is explained in the following steps:

1. Neighboring nodes send their information to the lowest ID vehicle, the maximum number of neighboring vehicles in the proposed protocol is 10, including the lowest ID vehicle. Using equation (8), the transmitting energy is equal to 459.09μJ.
2. The lowest ID vehicle receiving the transmitted messages from the neighboring vehicles, which costs receiving energy equal to 450μJ as equation (9).



3. The lowest ID sends all the information about its group to the RSU, consuming  $2130\mu\text{J}$  from its energy, using equation (7).
4. RSU broadcasts for all the group members their new status after performing necessary calculations, which needs about  $2130\mu\text{J}$ .
5. The cluster members receive the broadcast message, which is calculated as  $500\mu\text{J}$ .
6. The total energy consumed is equal  $5669.09\mu\text{J}$ .

Energy consumption when the CH leaves the cluster and the standby CH takes over in the cluster head duties:

1. CH sends a Leave message to its standby CH, consuming  $51.01\mu\text{J}$ .
2. Standby CH receives the CH's message which consumes a receiving energy of  $50\mu\text{J}$ .
3. Standby CH sends a message to its member to announce its new role consuming about  $51.01\mu\text{J}$ .
4. The members receive the message that totally consume  $400\mu\text{J}$ .
5. The total energy consumed in this case is equal to  $552.02\mu\text{J}$ .

By adding the energy consumed in the cluster formation and the energy consumed when the CH leaves its cluster, the total energy is equal to  $6221.11\mu\text{J}$ .

b. Energy consumption in the DCA:

In the DCA algorithm, the RSU is not eligible to take a place in the cluster formation process. The energy consumption for one cluster formation is explained in the following steps:

1. Neighboring nodes exchange their information together; the maximum number of neighboring vehicles in the DCA is 21 vehicles. So, the total energy consumed in transmitting and receiving Hello messages is equal to  $2121.21 \mu\text{J}$ .
2. After the calculations, the vehicle with the highest weight announces its role as a CH to its members. The energy consumed is  $1020.2 \mu\text{J}$ .
3. The members receiving the announced message consume  $1000 \mu\text{J}$ .
4. The total energy consumed is equal  $4141.41 \mu\text{J}$ .

Energy consumption when the CH leaves the cluster:

When the CH leaves the cluster, the cluster formation process is performed again since there is no standby CH. So, the total energy consumed in the cluster formation and the energy consumed when the CH leaves its cluster is equal to  $8282.82 \mu\text{J}$ .

By comparing the two algorithms, the proposed algorithm reduces the energy consumed when the CH leaves its cluster by 25% from the DCA algorithm. This is the result of electing two cluster heads in each cluster.

#### **4.4 Discussion**

An efficient clustering algorithm should consider the rate of clustering process in the network to be minimized to an acceptable level. It is clear that the DCA algorithm had a larger number of clusters than the proposed algorithm. This is because there is no upper or lower bounds that limit the size of the cluster. So, the cluster could be as large as it can be, or a very small size as in single node case.

Furthermore, as the total number of clusters increases, the contention increases too. So, the probability of collision is increased, and the throughput of the system is decreased.

In the proposed algorithm, the single node clusters are eliminated since they only consume the bandwidth and the resources of the network. The single node clusters are treated as any cluster exists in the network. They send and receive messages from the RSU as other clusters. As a result, these single node clusters increase the RSU energy consumption, the medium contentions, and the computations processes. The results showed that there is 3 single node clusters in the proposed algorithm. But to clarify this point, the algorithm does not accept these remaining nodes as single node clusters, actually the RSU will not save their history as an existing node in the zone and it will not reserve any resources for them. Moreover, these single nodes are remaining without any cluster because either they are far away from any CH or they are near clusters that are already filled with the possible maximum number of members (10 in our case).

On the other hand, the proposed algorithm has a defined maximum cluster size which is 10 vehicles. This size is derived considering the collision and the throughput calculations explained in chapter 3. As a result, by substituting the parameters that are defined in table 3 in the equations that are described in section 3.10, it was found that the probability of collision equal 0.384 and the saturation throughput equal 0.9315 (93%). However, in the DCA algorithm since the cluster size was undefined; the largest cluster size reached was 21 members. As a result, the probability of collision increased to 0.48 and the throughput decreased by 9% to 0.854 (85%).

## CHAPTER 5

### Conclusion

This chapter represents an overview of the thesis, the main contributions and features, and the future works that can be done for further improvements of the proposed protocol.

Building a complete clustering protocol in order to coordinate the communication between the vehicles themselves and between the vehicles and the RSU is a great challenge. It requires a careful design and overall considerations of all the aspects that affect the performance of the clustering. As our system is a vehicular network system; it is important to consider the mobility factors in the design. The mobility factors are the vehicles' locations, their nodal degrees, their speeds, and finally their movement directions. The weighting factor in the clustering technique depends on these parameters. According to the values of the weighting factors, the cluster heads can be elected. Once the cluster heads are selected, the communications between the cluster heads and the RSU will be coordinated, and the communications between the cluster heads and their members as well.

The main contribution of this thesis is designing a complete protocol that takes into considerations all the aspects that influence the vehicular networks. Also, this protocol has defined the boundary of the cluster size which depends on the probability of collision and the saturation throughput using analysis and calculation. Furthermore, it is critical to preserve the time and resources consumed in the clustering process, and this was done by selecting two CHs for each cluster.

Moreover, a realistic weighting factor has been created by considering all the important factors in the vehicular networks.

The overall performance of this protocol is very good. It reduces the total number of clusters that are formed within the RSU zone. Also, it limits the infinite size of clusters by defining an upper bound for the cluster size. In addition, it eliminates the single node clusters which consume the resources of the network.

For further improvements related to the proposed clustering protocol, there are some suggestions:

- This protocol is implemented for urban areas, so it needs few modifications to be implemented for highway scenarios.
- Propose a new communication technique between the RSU and the Cluster heads (e.g. TDMA) to have free collision communications.
- Consider all the urban area features in the design, such as traffic lights, intersections, roundabouts, opposite direction of the road, etc.

## Bibliography

- [1] "Emirates 24/7," 28 2 2012. [Online]. Available: <http://www.emirates247.com/news/emirates/road-accidents-kill-20-emirati-every-month-2012-02-28-1.445610>. [Accessed 31 12 2013].
- [2] "Health Statistics 2012." Health Authority Abu Dhabi, 10 10 2013. [Online]. Available: <http://www.haad.ae/statistics>. [Accessed 31 12 2013].
- [3] F. Losilla, A.-J. Garcia-Sanchez, F. Garcia-Sanchez, J. Garcia-Haro and Z. Haas, "A comprehensive Approach to WSN-based ITS Applications: A Survey," *Sensors*, vol. 11, pp. 10220-10265, 2011.
- [4] J. Hoebeke, I. Moerman, B. Dhoedt and P. Demeester, "An Overview of Mobile Ad Hoc Networks: Applications and Challenges," *Journal of the Communications Network*, vol. 3, pp. 60-66, 2004.
- [5] S. Zeadally, R. Hunt, Y. S. Chen, A. Irwin and A. Hassan, "Vehicular ad hoc networks (VANETs): Status, Results, and Challenges," *Telecommunication Systems*, 2010.
- [6] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, p. 393-422, 2002.
- [7] S. Djahel and Y. G. Doudane, "A Framework for Efficient Communication in Hybrid Sensor and Vehicular Networks," in *The 9th Annual IEEE Consumer Communications and Networking Conference-Smart Spaces and Personal Area Networks*, 2012.
- [8] M. J. Piran, G. R. Murthy and G. P. Babu, "Vehicular Ad Hoc and Sensor networks: principles and challenges," *International journal of ad hoc, sensor & ubiquitous computing (IJASUC)*, vol. 2, no. 2, pp. 38-49, 2011.
- [9] S. Gräfling, P. Mahonen and J. Riihijärvi, "Performance evaluation of IEEE 1609 WAVE and IEEE 802.11p for vehicular communications," in *Second International Conference on Ubiquitous and Future Networks (ICUFN)*, Jeju Island, Korea (South), 2010.
- [10] C. Han, M. Dianati, R. Tafazolli and R. Kernchen, "Throughput Analysis of the IEEE 802.11p Enhanced Distributed Channel Access Function in Vehicular Environment," in *IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, Ottawa, 2010.
- [11] S. Y. Ni, Y. C. Tseng, Y. S. Chen and J. P. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," in *Fifth Annual International*

*Conference on Mobile Computing and Networking (MobiCom'99)*, Seattle, Washington, 1999.

- [12] S. Vodopivec, J. Bester and A. Kos, "A survey on clustering algorithms for vehicular ad-hoc networks," in *35th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, 2012.
- [13] "Benefits of clustering for mobile ad hoc networks," REGREE Essays, [Online]. Available: <http://essaybank.degree-essays.com/computer-science/benefits-of-clustering-for-mobile-ad-hoc-networks.php>. [Accessed 28/5/2013].
- [14] M. Gerla and J. T. Tsai, "Multiuser, Mobile, Multimedia radio network," *Wireless Networks*, vol. 1, pp. 255-265, 1995.
- [15] G. Chen, F. G. Nocetti, J. Gonzalez and I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," in *35th Annual Hawaii International Conference on System Sciences, HICSS*, Hawaii , 2002.
- [16] L. Maglaras and D. Katsaros, "Distributed clustering in vehicular networks," in *8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Barcelona, 2012.
- [17] E. Dror, C. Avin and Z. Lotker, "Fast randomized algorithm for hierarchical clustering in Vehicular Ad-Hoc Networks," in *The 10th IFIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, Favignana Island, Sicily, 2011.
- [18] M. Chatterjee, S. Sas and D. Turgut, "An on-demand weighted clustering algorithm (WCA) for ad hoc networks," *GLOBECOM*, vol. 3, pp. 1697-1701, 2000.
- [19] P. Basu, N. Khan and T. Little, "A Mobility Based Metric for Clustering in Mobile Ad Hoc Networks," *IEEE ICDCS*, Massachusetts, 2001.
- [20] I. Er and W. Seah, "Mobility-based d-hop clustering algorithm for mobile ad hoc networks," in *Wireless Communications and Networking Conference*, 2004.
- [21] . S. Dhurandher and G. Singh, "Weight based adaptive clustering in wireless ad hoc networks," in *International Conference on Personal Wireless Communications, ICPWC* , 2005.
- [22] Y. Hwang, Y. Jeong, S. Lee and J. Song , "Advanced Efficiency and Stability Combined Weight Based Distributed Clustering Algorithm in MANET," in



*Future Generation Communication and Networking (FGCN 2007)*, Jeju, 2007.

- [23] M. Rezaee and M. Yaghmaee, "A new clustering protocol for Mobile Ad-Hoc networks," in *International Symposium on Telecommunications, IST 2008.*, Tehran, 2008.
- [24] S. Basagni, "Distributed Clustering for Ad Hoc Networks," in *Fourth International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN '99)*, Perth/Fremantle, WA, 1999.
- [25] G. Wolny, "Modified DMAC Clustering Algorithm for VANETs," in *3rd International Conference on Systems and Networks Communications ICSNC '08*, Sliema, 2008.
- [26] M. Almalag and M. Weigle, "Using traffic flow for cluster formation in vehicular ad-hoc networks," in *IEEE 35th Conference on Local Computer Networks (LCN)*, Denver, CO, 2010.
- [27] W. Kaimin, Z. Zhangdui and L. Hanzo, "A Cluster-Head Selection and Update Algorithm for Ad Hoc Networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, Miami, FL, 2010.
- [28] F. Wei , S. Yan, C. Shanzhi and Z. Longhao , "A mobility metrics based dynamic clustering algorithm for VANETs," in *IET International Conference on Communication Technology and Application (ICCTA 2011)*, Beijing, 2011.
- [29] Z. Rawashdeh and S. Mahmud, "A novel algorithm to form stable clusters in vehicular ad hoc networks on highways," *EURASIP Journal on Wireless Communications and Networking*, vol. 15, pp. 1-13, 2012.
- [30] N. Maslekar, M. Boussedjra, J. Mouzna and H. Labiod, "A stable clustering algorithm for efficiency applications in VANETs," in *7th International Wireless Communications and Mobile Computing Conference (IWCMC)* , Istanbul, 2011.
- [31] N. Maslekar, M. Boussedjra, J. Mouzna and H. Labiod, "C-DRIVE: Clustering Based on Direction in Vehicular Environment," in *4th IFIP International Conference on New Technologies, Mobility and Security (NTMS)* , Paris, 2011.
- [32] N. Maslekar, J. Mouzna, H. Labiod and M. Devisetty, "Modified C-DRIVE: Clustering based on direction in vehicular environment," in *IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, 2011.



- [33] F. Peng, "Improving Broadcasting Performance by Clustering with Stability for Inter-Vehicle Communication." in *IEEE 65th Vehicular Technology Conference, VTC2007-Spring*, Dublin, 2007.
- [34] Y. Harikrishnan and J. He, "Clustering algorithm based on minimal path loss ratio for vehicular communication," in *International Conference on Computing, Networking and Communications (ICNC)*, San Diego, 2013.
- [35] "Throughput," Wikipedia, [Online]. Available: <http://en.wikipedia.org/wiki/Throughput>. [Accessed 5 12 2013].
- [36] M. H. Manshaei and J. P. Hubaux, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function: Bianchi Model," in *Mobile Networks*, 2010.
- [37] C. Han, M. Dianati, R. Tafazolli and R. Kernchen, "Throughput Analysis of the IEEE 802.11p Enhanced Distributed Channel Access Function in Vehicular Environment," in *IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, Ottawa, 2010.
- [38] S. D. Jung, S. W. Park and S. S. Lee, "Effect of MAC Throughputs According to Relative Velocity in Vehicle Ad Hoc Network," in *Proc. International Conference on Convergence Information Technology*, 2007.
- [39] H. Al Qamzi and T. J. Li, "A Distributed Topology Management Protocol for Wireless Sensor Networks," in *IEEE International Conference on Communications (ICC)*, Cape Town, 2010.



## ملخص الرسالة

صُنفت دولة الإمارات العربية المتحدة ضمن أعلى الدول في عدد حوادث الطرق؛ لذلك هناك ضرورة ملحة لإيجاد الحلول المناسبة والفعالة لتقليل أعداد هذه الحوادث. أحد هذه الحلول هو الانتقال إلى أنظمة المواصلات الذكية (ITS). و المقصود بها هو تنفيذ التقنيات المتطورة في الطرقات لتكون أكثر ذكاءً، أكثر أمناً وأكثر تنسيقاً. هناك عدد من الدراسات والبحوث المختلفة التي نُشرت في هذا المجال. و هذه الأطروحة هي أحد هذه البحوث التي تسعى بشكل رئيسي إلى تحقيق السلامة في طرقاتنا.

إن الهدف الأساسي من هذا البحث هو بناء بروتوكول جديد لتجميع المركبات في الشارع في شبكات السيارات و الاستشعار اللاسلكي (HSVN). حيث أن هذه الشبكات عبارة عن مفهوم جديد من الشبكات التي تدمج شبكات الاستشعار اللاسلكية مع شبكات المركبات و ذلك لزيادة كفاءة الشبكتين معاً. على الرغم من أن هناك العديد من المقترحات و الحلول التي نشرت، لكن لا تزال هناك حاجة إلى وجود بروتوكول متكامل له القدرة في حل أو تقليل أوجه القصور الموجودة في الحلول السابقة.

من أهم صفات البروتوكول المقترح في هذه الأطروحة، أولاً أن هذا البروتوكول له القدرة على خلق نظام متوازن بين المركبات الموجودة في الشبكة و وحدة الاتصال المثبتة على جانب الطريق (RSU) و ذلك من خلال تقسيم عملية التجميع بينهما. بالإضافة إلى ذلك ، فإن هذا البروتوكول يقلل من الحسابات في عملية التقسيم و التصنيف و كذلك الاتصالات من خلال انتخاب اثنتين من المركبات في كل مجموعة إحداهما تُمثل السيارة القائدة للمجموعة و الأخرى كقائدة احتياطية في حال فُقدت القائدة الرئيسية.

وعلاوة على ذلك، يهدف هذا البروتوكول الجديد للحد من عملية فقدان الرسائل المتبادلة بين المركبات و الوحدة الجانبية (RSU) مما يؤدي إلى زيادة إنتاجية الشبكة، و هذا لا يمكن أن يحدث إلا بتحديد الحد الأعلى لعدد المركبات الممكن تواجدتها في كل مجموعة.

أظهرت نتائج هذا البحث أن البروتوكول المطروح في هذه الرسالة تَفُوق على عدد من البروتوكولات الموجودة حالياً التي تعتمد على نظام التنقل (mobility-based) من عدة أوجه، أولاً: العدد الإجمالي للمجموعات التي تشكلت في الشبكة. ثانياً: عدد المجموعات التي تتكون فقط من مركبة واحدة. و أخيراً: إنتاجية المجموعات في الشبكة.



جامعة الإمارات العربية المتحدة

كلية الهندسة

قسم الهندسة الكهربائية

تصميم بروتوكول لتقسيم و تجميع السيارات في شبكة المواصلات الذكية

صالحه مبارك محمد المهيري

رسالة مقدمة لاستكمال متطلبات الحصول على درجة ماجستير العلوم في الهندسة الكهربائية

قسم الهندسة الكهربائية

بإشراف د. هند القمزي

إبريل 2014