REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Sceitifique

Université Ferhat Abbas Sétif-1

Faculté des Sciences

Départment d'Informatique

THÈSE

Présenté par

Oussama SENOUCI

Pour l'obtention du diplôme de

Doctorat 3ème Cycle LMD en Informatique

Option : Systèmes Informatiques Intelligents et Communiquants

Thème

Application de l'Approche Clustering dans le Contexte de l'IoV

Soutenue le: 12 Juin 2019 Devant la commission d'examen:

Pr. Fouzi SEMCHEDDINE	Prof. Université Ferhat Abbas Sétif-1	Président
Dr. Zibouda ALIOUAT	MCA Université Ferhat Abbas Sétif-1	Rapporteur
Pr. Saad HAROUS	Prof. UAE University, UAE	Co-Rapporteur
Dr. Adel ALTI	MCA Université Ferhat Abbas Sétif-1	Examinateur
Dr. Abdelmalek BOUDRIES	MCA Université A. Mira Béjaia	Examinateur

PEOPLEs' DEMOCRATIC REPUBLIC of ALGERIA
Ministry of Higher Education and Scientific Research
Ferhat Abbas University Sétif-1
Faculty of Sciences
Department of Computer Science

THESIS

Presented by

Oussama SENOUCI

As a requirement to aim for the degree of

Doctorate 3rd Cycle LMD in Computer Science

Option : Intelligent and Communicating Computer Systems

Title

Applying the Clustering Approach in the Context of IoV

Defended on: June 12, 2019   Board of examiners:

Pr. Fouzi SEMCHEDDINE	Prof. Université Ferhat Abbas Sétif-1	Chairman
Dr. Zibouda ALIOUAT	MCA Université Ferhat Abbas Sétif-1	Supervisor
Pr. Saad HAROUS	Prof. UAE University, UAE	Co-Supervisor
Dr. Adel ALTI	MCA Université Ferhat Abbas Sétif-1	Examiner
Dr. Abdelmalek BOUDRIES	MCA Université A. Mira Béjaia	Examiner

**“I never dreamed about success,
I worked for it.”**

- Estee Lauder -

Declaration of Authorship

I, Oussama SENOUCI, declare that this thesis, “ Applying the clustering approach in the context of IoV ”, and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: ....

Date: ....

Acknowledgments

In the name of ALLAH, the most gracious and merciful. I thank ALLAH for the strength and blessing he gives me. I have never expected that I would arrive to complete this work without the aid of ALLAH.

My most sincere gratitude goes to my enthusiastic supervisor, **Dr. Zibouda ALIOUAT**, for her tremendous academic support.

Special mention goes to my co-supervisor **Prof. Saad HAROUS**, not only for his support and the advice he provided, but also for giving me so many wonderful remarks on different aspects of my research.

I wish to thank the members of my dissertation committee for generously offering their time, support, guidance and good will throughout the review of this document.

Last and not least, I would like to thank all my friends and colleagues for their encouragement.

Dedication

This dissertation is dedicated to:

My dear father and mother

My wife

My brother and sisters

All my family and friends

For all their love, patience, kindness and support.

Table of contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

MANET	Mobile Ad hoc NETwork
VANET	Vehicular Ad hoc NETwork
ITS	Intelligent Transportation System
IVC	Inter Vehicular Communication
V2V	Vehicle to Vehicle
V2I	Vehicle to Internet
V2R	Vehicle to Road infrastructures
V2D	Vehicle to Device
WATs	Wireless Access Technologies
IoT	Internet of Things
IoV	Internet of Vehicles
RSU	Road Side Unit
OBU	On Board Unit
TCC	Transportation Control Center
VCC	Vehicular Cloud Computing
GPS	Global Positioning System
CH	Cluster Head
CM	Cluster Member
WAVE	Wireless Access in Vehicular Environment
DSRC	Direct Short Range Communications
QoS	Quality of Service
GPS	Global Positioning System
IDM	Intelligent Driver Model

Introduction

This research work was done in the **LRSD Laboratory** of **Ferhat Abbas Sétif-1 University**. In our thesis, we focused on the design, modeling and simulation of new clustering approaches in the context of the Internet of Vehicles. In the introduction of this manuscript, we will first present the context and problematic of the research work, then we describe the goals and contributions of this thesis. Finally, we will describe the organization of this thesis.

Context and Problematic

In recent years, we have seen a significant evolution in Vehicular Ad hoc NETworks (VANETs), mainly due to the different needs currently expressed in terms of road safety and comfort. They can offer to both drivers and passengers. The emergence of the Internet of Things (IoT) has led to the evolution of conventional VANETs towards a new paradigm called Internet of Vehicles (IoV). This latter can be seen as a new Intelligent Transportation System (ITS) paradigm. According to recent statistics, 15 billion "things" are connected to the Internet in 2015 and 75 billion of "things" will be connected to the Internet by 2025 among which vehicles will constitute a significant portion[1].

The difference of the vehicle concept in VANET and IoV makes these two paradigms essentially different in the terms of devices, communications technologies, challenges, requirements, services and applications. In VANET, a vehicle is considered as a mobile node used to disseminate messages among vehicles and infrastructures. On the

[1]Statistics conducted by Statista (The portal for statistics) from 2015 to 2025.

other hand, a vehicle in IoV is considered as a smart mobile object with a powerful multi-sensor system, communications technologies and IP-based connectivity to the Internet, infrastructures and other vehicles [1]. Moreover, a vehicle in IoV is perceived as a dynamic mobile communication system, which allows the following communication modes: intra-vehicle components mode and Vehicle-to-X (V2X) mode, where X can be vehicle, road, infrastructure, human, cloud or Internet. As a result, the IoV enables the acquisition and processing of large amount of data (big data) from different geographical areas via intelligent vehicles computing systems to offer various categories of services for road safety and other non-safety services to drivers and passengers.

Compared to conventional VANETs, IoV has many specific advantages and characteristics, such as developing and extending the exploitation of the ITS in different fields of research and industry. As a main advantage, the ability to integrate multiple users, multiple vehicles and multiple networks using different communication technologies. This benefit allows sharing safety information between vehicles and providing useful information, such as availability of hotels, parkings, location, gas stations, and even drivers comfort application. Moreover, the IoV system allows to exploit the different wireless access technologies, which facilitates the network scaling with a significant number of connected vehicles. Furthermore, Vehicular Cloud Computing (VCC) technology can be integrated in the vehicular networks. This emergent technology allows applications, resources and data to be stored in the cloud, so that they can be used by clients with low capacity. Therefore, the VCC technology can manage the large amount of data generated by the high number of the connected vehicles [2].

Even though IoV has many advantages, it faces a number of challenges. As a main challenge, the integration of all components and object communications in the IoV ecosystem. Secondly, the rapid growing number of vehicles and other objects connected to the IoV system. Thirdly, the big data to be processed and stored in IoV network and that is because of the large number of connected vehicles. This challenge negatively affects the network overhead. Moroeover, it has been noticed that IoV faces many implementation problems, such as the security and reliability of inter-vehicles communications and multiple simultaneous requests for assistance that generate a large number of collisions [3,4]. Finally, the challenges that already exist in conventional VANET networks, such as high mobility of vehicles, dynamic topology, variable density, network fragmentation and communication models.

Considering all the above issues, the choice of an efficient network structuring technique for the IoV paradigm is a topic that is still relevant, especially since this type of network is expanding, because of its huge potential for beneficial contributions, especially for security in which the lives of individuals are at stake. In this context, researchers have turned to the use of new network structuring techniques, for managing their challenges and providing suitable conditions to run the different applications. As an important network structuring technique used in conventional VANETs, clustering has significantly improved their performances compared with classic flat structure in numerous applications, such as data dissemination and aggregation, Quality of Service (QoS), network overhead minimization, road safety, security, channel access control, topology discovery, drivers comfort and mainly routing prtocols [5, 6].

For this reason, we have focused our work on the clustering technique, which is widely deployed in vehicular networks to imporve the performances of the IoV system and meet their requirements. According to Yang and al. [7], clustering is the technique of dividing the network into groups of nodes called clusters. Each cluster has a cluster head and the rest of nodes in the cluster are called cluster members. This grouping is performed according to the application requirements to provide an easily manageable area. So, the vehicles are compared with each other, such that the most similar one, according to certain metrics, such as mobility metrics, weighted metrics, geographical location metrics and neighbors information metrics are selected to join the same cluster. Typically, the clustering method is composed of five main phases: neighborhood discovery, Cluster Head (CH) selection, announcement, affiliation and maintenance [8]. Clustering is therefore a key mechanism for the design and modeling of vehicular networks protocols and applications to ensure:

- Load balancing: Clustering aims to spread evenly the most expensive tasks in the network by observing node density and cluster size in order to avoid congestion points [9, 10]. This significantly minimizes the total network overhead.

- Scalability: Clustering allows to guarantee a good functioning even with a high number of nodes without affecting the network performances.

- Stable structure: Clustering makes it possible to form clusters with more stable structures in a very dynamic environment, because clusters are formed taking into account different metrics related to nodes, such as mobility, position, density, speed, etc.

- Optimal bandwidth exploitation: During the clustering process, the members interact only with the CH of their cluster, which avoids the unnecessary exchange of messages between nodes in the same cluster. As a result, the rate of collisions will decrease significantly.
- Quality of service (QoS): The high mobility of the nodes or other transitory faults can cause a frequent disconnection in the links of communications which causes a temporary stop of the service provided by the application of the network. It is therefore necessary to ensure and provide a high quality of service, even in the presence of breaks in communication links, to ensure the required service [11].

Goals and Contributions

In this PhD thesis, we are interested in the application of the clustering technique in the context of IoV to meet the requirements of such network. Therefore, the general goal of this thesis focuses on the design, modeling and simulation of new clustering algorithms for the IoV paradigm. These proposals are developed to improve the performances of these networks and confront their challenges. The contributions of our thesis is in the form of scientific papers, each dealing with a specific problem.

Firstly, in our papers: "A review of routing protocols in Internet of vehicles and their challenges" [3], published in *Sensor Review (Emerald)*, "Survey on VANET clustering algorithms: An overview, taxonomy, challenges and open research issues" [12], submitted paper and "Survey: Routing Protocols in Vehicular Ad Hoc Networks" [13], presented in *the Second International Conference on Advanced Wireless Information, Data, and Communication Technologies (AWICT 2017)*, we conduct a critical review of the literature of the clustering algorithms in VANETs. In this respect, we propose a new taxonomy to classify these algorithms with a detailed description of each existing algorithm, and discuss about their advantages and drawbacks. Moreover, a detailed comparison is provided for each classes of the proposed taxonomy considering relevant key parameters. This study allowed us to explored thoroughly the issue, in order to propose new clustering algorithms suitable for the IoV paradigm.

Secondly, we propose three different clustering approaches suitable for the IoV environment. The common goals between these proposals are to optimize the network overhead and improve the clusters stability in such environment. In our paper entitled "A New Heuristic Clustering Algorithm Based on RSU for Internet of Vehicles" [14], accepted for publication at *Arabian Journal of Science and Engineering (Springer)*, we propose a new one-hop heuristic clustering algorithm based on graph theory concepts and RSU, called HCAR. The latter entails the centralization of a clustering process at distributed RSUs. The basic motivation behind using the RSUs to perform the clustering process is that the RSUs are fixed infrastructure. It is much easier to send a message to a fixed target than to a moving one. Moreover, the corresponding algorithm uses a new recovery method, which is based on the election of the secondary CH using a weighted scheme.

In our paper entitled "MCA-V2I: A Multi-hop Clustering Approach over Vehicle-to-Internet communication for improving VANETs performances" [15], published in *Future Generation Computer Systems journal (Elsevier)*, we develop a new multi-hop clustering algorithm using Internet access to perform the clustering process. The corresponding algorithm propose a new multi-hop clustering model is proposed. Compared with one-hop clustering schemes, this model is designed to extend the coverage area of clusters, reduce the number of clusters, optimize the control overhead and improve cluster stability. Moreover, MCA-V2I provides an Internet access to vehicles to exchange the necessary information to perform the clustering process. Furthermore, a new parameter, called mobility rate is introduced to perform the clustering process. This parameter is calculated based on mobility metrics to satisfy the requirements of the new features of VANET, and to consider its mobility characteristics.

In our paper entitled "An Efficient Weight-Based Clustering Algorithm using Mobility Report for IoV" [16], presented in *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (IEEE UEMCON 2018)*, we present a new weight-based clustering algorithm for IoV, which aims to increase the stability of clusters and reduce communication overhead. The proposed algorithm uses both classic weighted parameters, such as degree, average distance and a new metric introduced by the corresponding algorithm, called mobility report. The latter is a parameter that combines mobility metrics: relative velocity and relative acceleration to satisfy the requirements of the new features of IoV, and to consider its mobility characteristics.

To sum up, the main contributions of this thesis can to be summarized as follows:

- analyze, through a critical review of the literature, a number of existing clustering algorithms in VANETs, in order to detect certain problems still open or whose proposed solutions are still to improve;
- propose new clustering approaches appropriate for the IoV network, which make it possible to improve its performances while guaranteeing a high level of service required by the function performed by the latter;
- evaluate the performance of the proposed algorithms, compared with known upper bounds and best known algorithms in recent literature.

Thesis organization

In this thesis, we opted for the "articles" format. Some chapters are therefore the transcription of articles published in, or submitted to, scientific journals or conferences. So, the thesis contains five chapters divided into two main parts: background and contributions, and it is organized as follows:

In the background part:

- Chapter 1 gives a background of knowledge about the Internet of Vehicle (IoV); their architecture, characteristics, challenges, communication technologies and applications; and all that belongs to this domain.
- Chapter 2 presents the state of the art on the clustering algorithms for VANETs in the literature. In this regard, we propose a new classification taxonomy based on different metrics, then we classify and survey a number of the recent clustering algorithms proposed in the literature according to this taxonomy. At this point, we attempt to detect certain problems still open or whose proposed solutions are still to improve, in order to propose new clustering approaches suitable for IoV.

The contributions part of this thesis introduces the proposed clustering approaches:

- In Chapter 3, we present a new one-hop heuristic clustering algorithm based on the graph theory concepts for IoV, known as HCAR, which is aimed at improving the total network overhead and clusters' stability. The proposed algorithm uses a central node that is the RSU to perform the clustering process. We carry out a theoretical and experimental analysis of our proposal.

- In Chapter 4, we propose a new multi-hop clustering approach over vehicle-to-Internet communication, known as MCA-V2I. This latter provides Internet access to vehicles to perform the clustering process. Moreover, the proposed approach aims to extend the coverage area of clusters, optimizes the control overhead and improves the clusters' stability. A study of the performances of the proposed approach is elaborated by simulation.

- In Chapter 5, we introduce an efficient weight-based clustering algorithm using a mobility report, namely WECA-MR for IoV. The proposed algorithm uses a new weighted metric called "mobility report" to perform the clustering process. The main goal of the proposed algorithm is to construct stable clusters with an acceptable load balancing. We evaluate our WECA-MR proposition through simulations.

- Finally, we conclude our thesis with a brief discussion and some future perspectives.

Background

Chapter 1

Internet of Vehicles: An overview

1.1 Introduction

In recent decade, the world has seen a great evolution in the automotive industry, which leads to increase the number of vehicles enormously. This revolution has certainly facilitated the daily lives of peoples and has even allowed the acceleration of economic growth around the world, but it has also generated some challenges, such as traffic congestion, traffic accident, energy consumption and environmental pollution [17]. These issues opened the door to global innovations, which contributed to the birth of the concept of the vehicular network. This concept is provided by a Vehicular Ad hoc Network (VANET), which allows vehicles to communicate with other vehicles (V2V) and road infrastructures (V2I) via messages. In VANET, each vehicle is either a source, a destination or a gateway in the network, which allows creating a network with a wide range [18]. However, when a vehicle goes outside the range of the network, it becomes an orphan node. As a result, it can be concluded that VANET covers only a very small mobile network due to mobility constraints and the number of connected vehicles [19].

Over the past several years, there has not been any classic or popular implementation of VANET. The desired commercial interests have not emerged either. Therefore, VANET's usage has begun to stagnate. The emergence of the Internet of Things (IoT) has led to the evolution of conventional VANETs toward a new paradigm called the Internet of Vehicles (IoV). IoV can be seen as a new Intelligent Transportation System paradigm. It extends VANETs scale, structure and applications. This evolution leads to the emergence of new interactions at the road level among vehicles, humans and infrastructure. IoV paradigm is a wide range of technologies and applications, including intelligent transportation, vehicle information service, modern information and communication technology and automotive electronics.

The present chapter introduces the fundamental concepts of IoV. Here, we point out on definitions, architecture, applications, communication technologies, characteristics, challenges and requirements that face such network.

1.2 IoVs' definitions

Alam et al. [20] defined IoV as an ITS integrated IoT from the perspective of intelligent transportation. Moreover, IoV is considered as the vehicular ITS, which regards the driving vehicle as the information sensing and intelligent mobile object.

Hartenstein et al. [21] consider that IoV is a global network that integrates three sub-networks: intra-vehicle network, inter-vehicles network, and vehicular mobile Internet.

Acording to Ang et al. [4], IoV is a convergence of the mobile Internet and the Internet of Things (IoT) where vehicles function as smart moving intelligent nodes or objects within the sensing network.

Jiacheng et al. [22] present the definition of IoV from the perspective of integration of on-board sensors and communication technology. They consider IoV as a set of intelligent vehicles equipped with advanced sensors, controllers, actuators, and other devices, which uses modern communications and network technology for providing the vehicles complex environmental sensing, intelligence decision-making and control functions.

It can be concluded that there are differences in the definitions of IoV when researchers provide their definition in view of their own research areas.

1.3 IoVs' network architecture

Because of the various definitions of IoV that proposed in the literature, there is no default or specific IoV network architecture. In similar trend, the current research proposes an architecture presented in Figure 1.1.

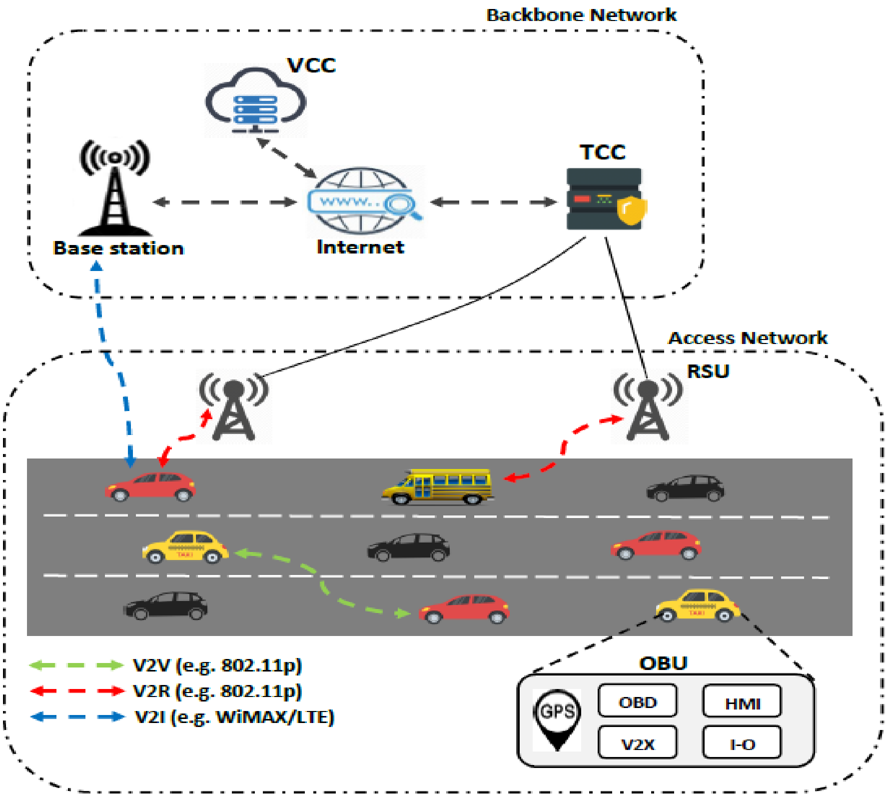


Figure 1.1: IoV network architecture.

The proposed network architecture of the IoV network is composed of two main parts: Access Network (AN) and Backbone Network (BN). The AN includes two main components: vehicle that equipped with On Board Unit (OBU) and RoadSide Unit (RSU). On the other hand, the BN includes three main components: Transportation Control Center (TCC), Vehicular Cloud Center (VCC) and Internet. The definition of the different components are as follows:

1. **Vehicle:** It is the mobile node and the main component of the architecture. It is seen a smart object equipped with a powerful multi-sensor platform, communications technologies, computation units, IP-based connectivity to the Internet and to other vehicles either directly or indirectly [3].

2. **On Board Unit (OBU):** It is a terminal equipment mounted on-board of a vehicle to provide a mutual wireless communication between vehicle and surrounding vehicles and infrastructures. It uses Wireless Access in Vehicular Environment (WAVE) standard which is based on the emerging IEEE 802.11p specification [23]. It operates on the 5.9 GHz Direct Short Range Communications (DSRC) band

compatible with vehicle systems to provide very low latency required for high speed events, such as crash avoidance. In addition, the OBU contains several modules: GPS device module, Human Machine Interface Device (HMI) module, On Board Diagnostics (OBD) module, Input-Output (I-O) devices and Vehicle-to-X (V2X) communication module. This latter allows the exchange of information from a vehicle to any entity that may affect the vehicle, and vice versa. It is a vehicular communication system that incorporates other more specific types of communication [24] as Vehicle-to-Vehicle (V2V), Vehicle-to-Road infrastructure (V2R), Vehicle-to-Internet (V2I), Vehicle-to-Sensors (V2S) and Vehicle-to-Device (V2D) (see Figure 1.2).

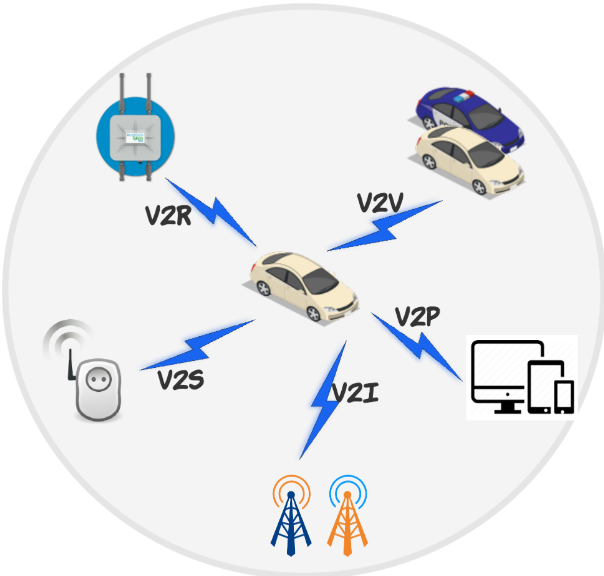


Figure 1.2: Communication types in IoV.

3. **Road Side Unit (RSU):** It is a computing device installed on the roadside that provides communication services to vehicles. Like OBU, the RSU uses WAVE standard which is based on the emerging IEEE 802.11p specification. It also operates on the 5.9 GHz DSRC standard. They are controlled and managed by the Transportation Control Center (TCC) through wired communication channels. It act as a gateway that allow vehicles to establish connection with Internet. The RSU provides different services to vehicles such as: communication services, security services, data dissemination and aggregation services, and Internet access. In addition, the RSU has a number of functionalities such as: extending the

network coverage, traffic directories, data dissemination, security management, location servers and service proxies [25].

4. **Transportation Control Center (TCC):** It is a trusted center that collects and maintains current information about vehicles. Based on the collected traffic and vehicle information, TCC can assist in controlling the traffic [26]. It has been built to support many safety applications such as: accident alert, traffic signal violation warning, traffic, Lane change warning and congestion detection; and no-safety applications such as: Mobile Internet access and music downloading [27]. It is responsible for network initialization, interconnecting RSUs and exchanging data between them. It represents the interface between the Access Network (AN) and the Internet network. that performs various functions such as: supervising and controlling the RSUs, road traffic management, interconnecting the conventional VANET network with other networks (e.g. cellular network) and facilitating the scaling of IoV network.

5. **Vehicular Cloud Computing (VCC):** It is a set of virtual servers that is based on cloud computing platform over the Internet. The VCC stores and shares data, resources and applications to serve vehicles on demand. Moreover, the VCC uses underutilized vehicle resources to form a cloud by aggregating vehicular network computing resources. Furthermore, the VCC refers to a group of largely autonomous vehicles, computing, sensing, communication and physical resources of which can be coordinated and dynamically allocated to end users [28]. The VCC provides several services include [29]:

   - Network as a service (NaaS): Objects having Internet access can provide this it to other network objects.
   - Storage as a service (STaaS): Objects having highest storage capacity can act as a storage platform for other network objects to run their applications remotely.
   - Cooperation as a service (CaaS): Objects having services subscribed provide necessary information to other objects according to their interest for the same services. Moreover, the CaaS allows users to obtain services with minimal capacity.

- Entertainment and Information as a Service (ENaaS / IaaS): Mobile objects can provide useful information for safe driving, such as traffic state, post-crash warning and difference emergency states. This service is knows as Information as a Service (IaaS). Furthermore, Entertainment as a Service (ENaaS) allows many commercials services, such as advertisements, photo, music and movies.

1.4 IoVs' applications

With the emergence of IoV, the already existing applications of VANETs have been improved and a number of new applications have emerged as well. This new infrastructure creates the basic network needed for many valuable applications that require seamless connectivity and addressability between their components. However, the overall of these applications can be divided into two major categories; safety-oriented applications and user-oriented applications [19] as shown in Figure 1.3.

1.4.1 Safety-oriented applications

Safety-oriented applications focus on reducing the damages and dangers, especially when the safety of the individuals is at stake. With the automatization and the intelligentization of vehicle, this latter can recognizing dangerous situation and can easily alerts alert others vehicles against these possible hazardous. There is a number research on safety applications, among them are the following applications:

- *Collision avoidance*: It is one of the important safety applications. At present, collision avoidance technologies are largely used in the literature [30–32].
- *Intelligent Intersection*: The intelligent intersection application, where such conventional traffic control devices as stop signs and traffic signals are removed, has been a hot area of research for recent years [19, 33].
- *Real-time traffic*: The real time traffic information can be stored at the road infrastructures or at vehicular cloud and can be available to the smart vehicles whenever and wherever needed. This benefit can play an important role in solving several issues such as traffic jams and emergency alerts [34].

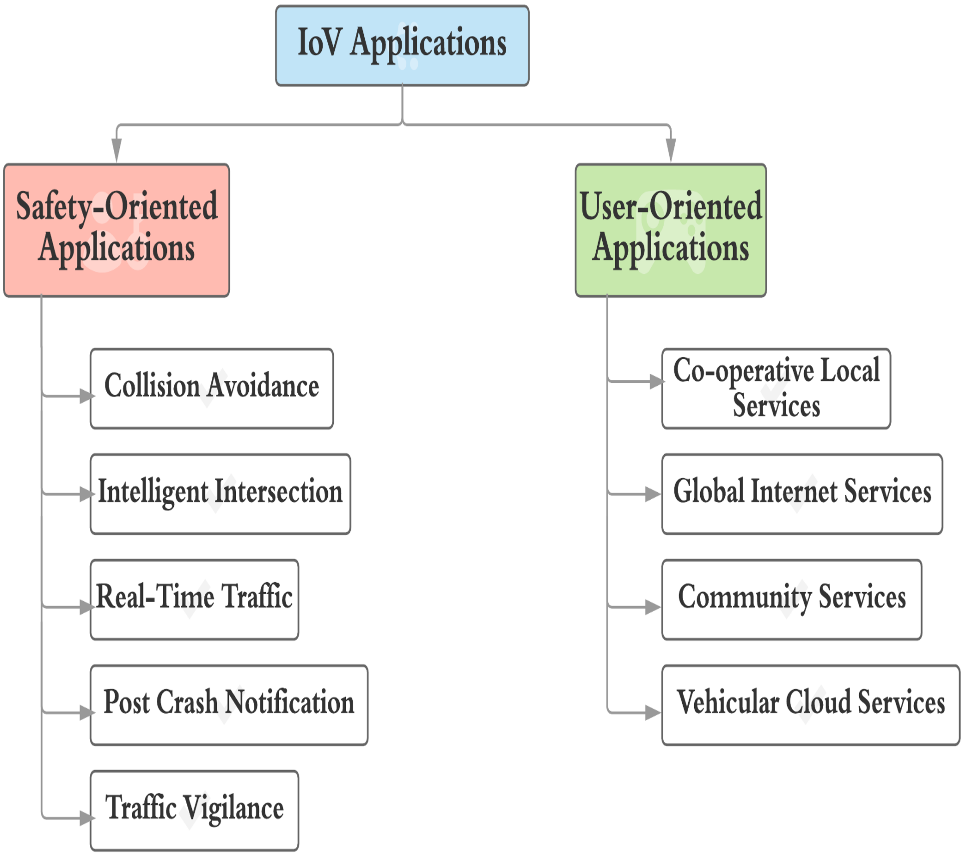


Figure 1.3: IoV applications.

- *Post crash notification*: When a vehicle be party to an accident, and it would transmit alert messages about its position to the neighboring vehicles, in order to these latter make decision.

- *Traffic Vigilance*: Example: cameras can be installed at the road infrastructures (RSU) in order to monitor the road. So, these cameras can work as input and act as the latest tool in low or zero tolerance campaign against driving offenses.

1.4.2 User-oriented applications

User-oriented applications refers to non-safety one, which provide value-added services to vehicles. Among these applications, let us site as examples:

- *Co-operative local services*: They are applications focusing on information that can be obtained from locally based services, such as point of interest notification, e-commerce, e-bank and media downloading [19].

- *Global Internet services*: Global Internet services focus on data that can be obtained from global Internet services [35].

- *Community services*: Example: insurance and financial services and parking zone management, which focus on software and data updates.

- *Vehicular cloud services*: It is the set of services provided by the Vehicular cloud technology, such as: web services and on-line storage.

1.5 IoV's features, challenges and requirements

The IoV paradigm is based on the integration of users, vehicles, things and networks, in order to provide the best connected communication capability that is manageable, controllable, operational, and credible [19]. However, the characteristics of the IoV environment in which the network works correctly and efficiently create new features and challenges of the former. For instance:

- *Conventional VANETs features:* Since IoV is evolution of conventional VANETs, many of its characteristics are similar to VANETs characteristics, such as high mobility, diversity of density, surrounding obstacles and dynamic topology.

- *Various Wireless Access Technologies (WATs):* IoV paradigm provides a number of wireless Access Technologies (WAT) such as WLANs, WiMAX, Cellular Wireless, and satellite communications [36]. The challenge is to consider the encapsulation of these multiple access technologies during the design of applications and protocols for IoV.

- *Enhancing communication ability:* There is always a dark spot associated with wireless access components, which is the bottleneck. This latter issue needs to be taken into consideration when designing a new applications and protocols for IoV paradigm.

- *Permanence and sustainability of the services provided:* A relevant example on this challenge is that vehicles in IoV, such as intelligent vehicle, normal vehicle and broken vehicle, cannot be expected to give the same type and level of services [19].

The outcome of IoV applications and services is affected by several circumstances. To succeed in their missions, the design of efficient an reliable applications and services for this paradigm should meet the following requirements:

- *Scalability:* One of the important characteristics of the IoV is that they can contain hundreds or even thousands of connected vehicles. Depending on the application needed, this number can further increase to millions of vehicles. The new proposed approaches must be able to guarantee a good functioning with this high number of vehicles without affecting network performances.

- *Fault tolerance:* Failure or crash of some vehicles in the network should not affect the rest of the network.

- *Latency:* It refers to the delay needed to transmit a packet in the vehicular network. This delay must be minimized as much as possible, in order to meet the requirements of IoV system, especially in critical applications.

- *Security and anonymity:* The importance of the information exchanged via vehicular communications makes the operation of securing these networks crucial task especially with access to the Internet, which is a prerequisite for the deployment of IoVs [37].

- *Quality of services (QoS):* The high mobility of the vehicles or other transitory faults can cause frequent disconnections in the communication links, which cause temporary stoppage of the service provided by the application in the IoV network. It is therefore necessary during the design of different applications to ensure a high quality of service, even in the presence of failure in communication links, to ensure the service sustainability.

- *Compatibility with Personal Devices (PD):* Compared to VANET, which is characterized by its limited communication technologies, IoV enables communication with the personnel devices, such as smart-phones and tablets.

- *Connectivity:* The vehicles must permanently be connected.

- *Network and environment awareness.*

- *Cloud Computing (CC) compatibility:* The main operations can be based on CC services.

1.6 Wireless Access Technologies (WATs) for IoVs

IoV system is based on multiple and heterogeneous wireless networks and technologies besides the traditional VANETs communication such as: Wireless Access in Vehicular Environment (WAVE) and Dedicated Short-Range Communications (DSRC) [38]. However, these WATs can be classified into four main technologies; WLANs, WiMAX, Cellular Wireless, and satellite technolgies [19, 36]. As shown in Figure 1.4, most of these technologies are used to connect vehicles to each other in IoV.

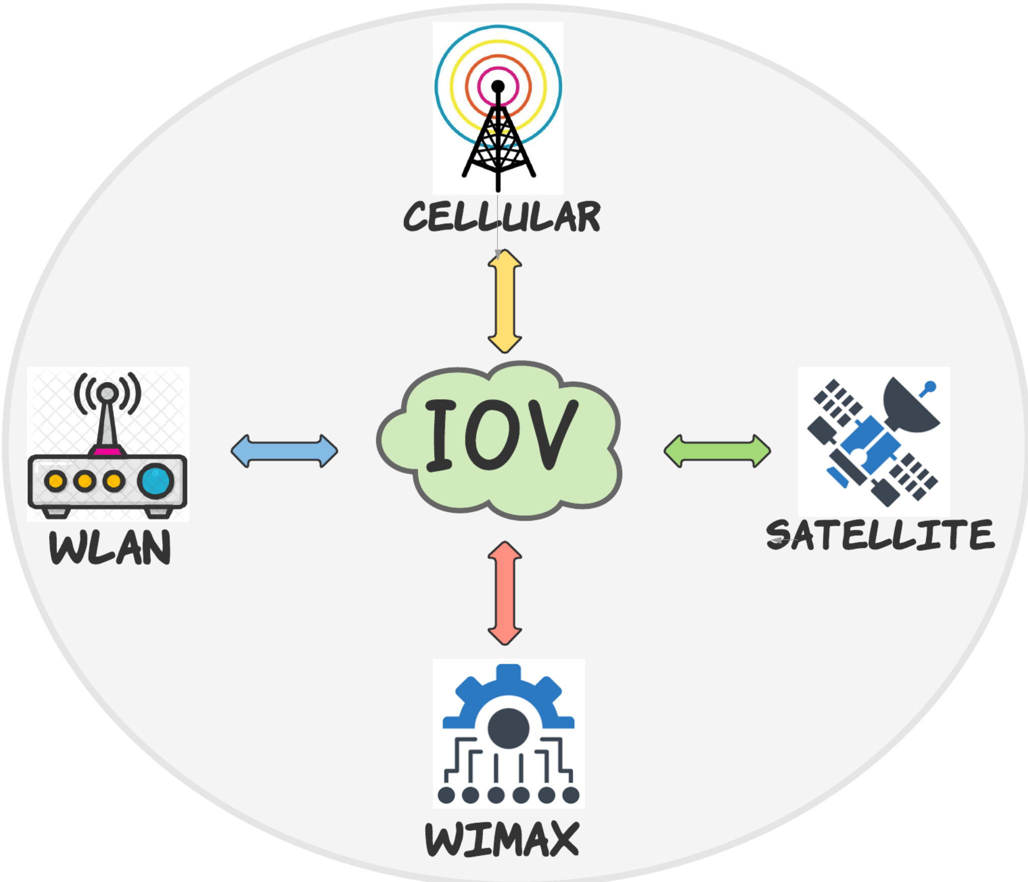


Figure 1.4: WATs in IoV.

1.6.1 Wireless LAN (WLAN) technology

Nowadays, WLANs are increasingly recognized as a versatile connectivity solution for different ad-hoc networks and environments. They are widely used to create ad-hoc networks due to their low cost, high data transfer rates provided and their easy deployment. WLAN technology uses IEEE 802.11a/b/g/n/p/ac standards. It has achieved great acceptance in both academic and industry areas, short-range supports, relatively high-speed data transmission. The IEEE 802.11n standard achieves a throughput of approximately 100 Mbps for a frequency of 5.9 GHz. IEEE 802.11p is a new communication standard in the IEEE 802.11 family which is based on the IEEE 802.11a standard. IEEE 802.11p is designed for wireless access in the vehicular environment to support intelligent transport system applications. So, it is the more suitable standard for IoV environment. The 802.11ac is an emerging standard of the WiFi family, which can provide a high throughput up to 1.3 Gbps for a frequency varies only from 5 to 6 GHz band.

1.6.2 WiMAX technology

WiMAX (Worldwide interoperability for Micro wave Access) supports IEEE 802.16 a/e/m standards. IEEE 802.16 standard-based WiMAX are able to provide an Internet access and to cover a large geographical area, up to 50 km, and can achieve a significant rate, up to 70 Mbps. While IEEE 802.16 standard only supports fixed broadband wireless technologies, the new IEEE 802.16e/m communication standards supports speeds up to 180 km/h and provides different classes of QoS [19]. In addition, WiMAX IEEE 802.16m uses advanced modulation techniques such as Adaptive Modulation and Coding (AMC), Hybrid Automatic Repeat Request (HARQ) and Fast Channel Feedback (CQICH) to offer broadband access to mobile objects [39]. Compared to WLAN, the key advantage of WiMAX is that the channel access method in WiMAX uses a scheduling process in which the subscribed user needs to compete only once for initial entry into the network.

1.6.3 Cellular wireless technology

Cellular system uses radio waves to transmit data over long distances. It has been utilized to provide mobile services since the 1970s. They use a technique of "frequency reuse" to increase coverage area and also for multiple transmissions simultaneously [40]. The smart cellular wireless comprises of 3G (UMTS) and 4G (LTE) technologies. Current 3G technology achieves a data rate of 384 kbps to moving vehicles, and can go up to 2 Mbps for fixed nodes or infrastructures. Compared to WLAN and WiMAX technologies, 3G systems deliver smoother hand-offs. On the other hand, the 4G technology were designed to provide high speed, broadband and cheaper mobile services over Internet. They support high mobility through soft hand-offs and seamless switching. Moreover, 4G system can provide a high transfer speed up to 129 Mbps for a frequency that varies between 1700 MHz and 2100 MHz band.

1.6.4 Satellite technology

Satellite technology is the use of the existing satellite system in the field of communications. It provides a number of mobile services such as: voice and video calling, Internet access, fax, television and radio channels. Moreover, satellite technology can provide ubiquitous coverage and communication capabilities spanning long distances and can operate under circumstances or conditions which are inoperable for other forms of communication such as: desert [41]. The short comings of this technology are high cost and large propagation delays, which affect negatively the network latency. In this respect, many protocols have been proposed for vehicular network using satellite technology. For example, Nasr et al. [42] propose a new VANET Clustering Based Routing Protocol Suitable for Deserts over the satellite technology.

Table 1.1 shows a comparison of the wireless access technologies in IoV.

Table 1.1: Comparison of wireless access technologies in IoV.

WATs	**Stand.**	**Freq. (GHz)**	**Rate (Gbps)**	**Cover. (Km)**	**Interf.**
WLANs	802.11a/b/g/n/p	5.9	1.3	1	Low.
WiMAX	802.11e/m	2.5/3.5	0.3	50	High
Cellular	UMTS/LTE	ND	0.3	50	Low
Satellite	MSIA/DVB	30	10	2000	Low

1.7 Simulations and mobility models for IoVs

The network simulation is the reference technique for predicting the performance of networks before they are physically built or deployed. Since the IoV network is composed of mobile objects, the choice of an appropriate mobility model becomes a very critical aspect, to meet the requirements of the proposed solution. The mobility models provide a precise description of vehicles mobility in the network.

According to the available literature, there exist many mobility models, which are widely used for IoV networks. Compared to conventional VANET networks, IoV vehicles are much more restrictive in terms of traffic on the network and require more traffic applications and services. These scenarios cannot be therefore handled by conventional networking solutions.

Therefore, IoV networks simulation should be realized via specially-designed VANETs simulators. The designer could either use a traffic simulator for generating realistic vehicular mobility traces that will be used as the input for a mobile ad hoc network simulator, such as NS3, NS2, OMNET, etc. Among the mobility generator in VANETs, we have: VanetMobiSim, SUMO and RoadSim. These mobility generators, called also vehicular traffic simulators, rely on a traffic flow theory.

For instance, VanetMobiSim utilizes the Intelligent Driver Model including Lane Change (IDM-LC) mobility model. This latter implements road intersection supervising strategy: making vehicle nodes slowing down, stopping or moving in accordance with traffic lights. Moreover, they are also capable to provide to vehicles for overtaking to change lane in multi-lane roads.

1.8 Conclusion

The IoV is one of the emergent technologies that attract growing attention from academia and industry areas in the last few years. This is because of the vital evolution provided to the conventional VANETs and ITS. The IoV paradigm inherit conventional VANETs communications (V2V and V2R) and offered a new communication types, such as Vehicle-to-Internet (V2I), Vehicle-to-Device (V2D), Vehicle-to-Human (V2H) and Vehicle-to-Sensors (V2S). This chapter provide an overview of the emergent IoV technology and an outlook for their future applications. An introduction presents the evolution of VANET towards the new paradigm IoV. Related definitions of IoVs and their applications are introduced. Since their peculiar characteristics create new challenges and requirements, we discuss the overall of IoV's issues. Moreover, the wireless access technologies related to IoV are depicted. Finally, simulations and mobility models for IoV networks are presented.

Next, we focus on the clustering technique that has been commonly used in VANET networks. In the literature, various clustering algorithms have been proposed. In this regard, we propose a new taxonomy to classify and review in depth a number of these clustering algorithms. It will be the subject of the next chapter.

Chapter 2

State of the art: Clustering in VANETs

2.1 Introduction

The emergence of the Internet of Things has led to the evolution of Vehicular Ad-hoc NETworks (VANETs) toward a new paradigm called the Internet of Vehicles (IoV). The difference of the vehicle concept in VANET and IoV makes these two paradigms different in the terms of devices, communications, services and applications. In VANET, the vehicle is considered as a mobile object used to disseminate messages among vehicles and infrastructures. However, each vehicle in IoV is considered as a smart object with a powerful multi-sensor system. Moreover, a vehicle in IoV is perceived as a dynamic mobile communication system, which allows the following communication modes: intra-vehicle components mode and Vehicle-to-X mode, where X can be vehicle, road, infrastructure, human or Internet. Therefore, IoV enables the acquisition and processing of large amount of data (big data) from different geographical areas via intelligent vehicles to offer various services for drivers and passengers. Most of the research works on VANETs and IoVs has focused on specific problems and applications, such as data dissemination and aggregation, channel access management, trafic safety, non-safety applications topology discovery and mainly routing schemes [43].

The design of effective application is one of important challenges that should not be neglected in VANET and even in IoV, especially with their special features and characteristics, such as scalability, high mobility of vehicles, not uniform density and rapid change of topology that make the design and the implementation of effective solutions for such networks a very difficult task [44]. Firstly, the high mobility which is the main factor distinguishing VANET and IoV from other classes of ad-hoc and wireless networks. The vehicles' speed vary according to the road conditions, it may be low or medium in urban areas and large on highways. This speed variation has a direct impact on the network stability and make the network topology very dynamic. Secondly, the nodes density in a vahicular network is not uniform with spatio-temporal variation. Typically, the density in urban areas is higher than in rural areas. On the other hand, the density is different depending on whether we consider night or day and rush hours [45]. Finally, the network fragmentation problem, which occurs generally when the vehicles' density is low and irregular. Thus, the vehicles move in disconnected isolated clusters, therfore it becomes difficult to end-to-end communications [3].

Considering all the above issues, researchers have turned to the use of new techniques and mechanisms that ensure a stable topology structure and an effective data routing and dissemination in the network. As an important network structure used in conventional VANETs, clustering presents an interesting solution for simplifying and optimizing network functions and services. It has significantly improved the performances in numerous applications compared to conventional flat structure [5]. The clustering process aims to structure the nodes of the network into small groups called clusters. Typically, the vehicles, geographically neighbors, are grouped in the same cluster according based on various parameters and metrics. The first contribution of this thesis is to provide a general overview of clustering process in VANETs and review the existent clustering algorithms in these networks, in order to propose new clustering approaches suitable for the IoV paradigm. The motivation behind this research is to the best of our knowledge the lack of clustering protocols suitable for the IoV network in the available literature.

As we know, clustering is a key mechanism for different networks. Thus, the design of efficient clustering algorithms must be guided by several factors and conceptual constraints:

- Load balancing: Clustering aims to spread evenly the most expensive tasks in the network by observing node density and cluster size in order to avoid congestion points [9, 10]. This significantly minimizes the total network overhead.

- Scalability: Clustering aims for a good functioning even with a high number of nodes without affecting network performances.

- Stable structure: Clustering makes it possible to form clusters with more stable structures in a very dynamic environment, because clusters are formed taking into account different metrics related to nodes, such as: mobility, position, density, speed, etc.

- Optimal bandwidth exploitation: During the clustering process, the members interact only with their CH, which avoids the unnecessary exchange of messages between nodes in the same cluster. As a result, the rate of collisions decreases significantly.

- Quality of service (QoS): The high mobility of the nodes or other transitory faults can cause frequent disconnections in the links of communications which causes a temporary stop of the service provided by the application of the network. It is therefore necessary to ensure and provide a high quality of service, even in the presence of breaks in communication links, to ensure the required service [11].

- Resource sharing: the main goal of distributed systems is to facilitate the access to shared resources, such as communication medium access.

- Concurrency: Multiple nodes working jointly on a specific application need an agreement about their common progress in real time. The timing sequence nature of these applications relies on an accurate clock synchronization mechanism.

- Fault-tolerance: fault-tolerance is an essential distributed system attribute that ensures reliability and maintainability. In distributed systems, nodes and communication links may fail independently at any time, so, fault recovering can be done by check-pointing or redundancy mechanisms. The lack of shared memory makes clock synchronization of a vital concern in any distributed applications needing to meet time requirements.

Clustering is largely used and well studied in VANETs. So far, many papers can be found in the current literature [7,8,46–49], which survey a number of proposed clustering algorithms. In relation to those surveys, we remark a lack of clustering algorithms suitable for the IoV paradigm. The available clustering algorithms for conventional VANETs do not meet the requirement of IoV system, especially its specific characteristics, such as heterogeneous networks, big data and connected vehicles. Therefore, the design of new clustering algorithms for IoV network becomes of critical importance.

The present chapter provides a deep review of clustering algorithms in VANET networks. Firstly, we introduce a general overview on clustering technique in VANETs. In this respect, we present a background of knowledge about the clustering process; their brief historical, definitions, clusters structure, cluster head selection metrics, general procedural flow and performance evaluation metrics. Secondly, we propose new taxonomy for clustering algorithms in VANETs with a detailed description of each algorithm. Thirdly, a detailed comparison is provided for each category of the proposed taxonomy considering relevant key parameters, such as vehicle density, cluster stability, latency, clusters overlapping and clustering overhead. Fourth, we highlight the main challenges encountered for each category and discuss some open research issues. Finally, we provide a general comparison of different clustering algorithms according to a number of selected parameters.

The rest of this chapter is organized as follows. Section 2.2 introduces an overview on clustering process in VANETs. Section 2.3 presents the related woks. Section 2.4 discusses the proposed taxonomy of clustering algorithms for VANETs in detail. Finally, a conclusion is presented in Section 2.5.

2.2 Clustering in VANETs: An overview

In recent years, the clustering is one of the most used control mechanism in VANETs, in order to overcome their different challenges and improve their performances. Bali et al. [46] define clustering as the mechanism of grouping vehicles into small groups called clusters based on a number of predefined metrics, such as nodes' density, average velocity, and geographical locations of the nodes. This grouping is done according to the requirements of the targeted application, in order to provide an easily manageable network.

2.2.1 Brief historical

The DARPA packet-radio network [50] is the earliest notable research work on clustering. The main goal being to autonomously form subnets the Mobile Ad hoc NETwork (MANET) in order to share the data and the network resources. Lin et al. [51] introduce the most popular Lowest ID and the Highest Degree (LID-HD) clustering protocols for MANETs. Mobility Clustering (MOBIC) [52] and Weighted Clustering Algorithm (WCA) [53] was later proposed, attempting to introduce a number of mobility and weighted metrics to the clustering process. Because of the significantly improvements of these proposed algorithms, several other algorithms were also designed, such as Distributed Group Mobility Adaptive (DGMA) clustering [54] and MobHiD [55]. Since the VANET is a subclass of MANET, the clustering algorithms designed for MANET is widely used in VANETs. This was the beginning of using clustering in vehicular network.

2.2.2 Clustering concepts

2.2.2.1 Cluster structure

The clustering is the process that aims to group the nodes of the network into small groups, called clusters, which gives the network a hierarchical structure. This structure can satisfy a number of network requirements, such as scalability, load balancing, network stability and Quality of Service (QoS). Typically, the nodes, geographically closer, have a high probability to join the same cluster upon based to certain rules and metrics. Typically, cluster structure have three main types of nodes (vehicles): Cluster Head (CH), Cluster Member (CM) and Cluster Gateway (CG) [56] as shown in Figure 2.1.

- **Cluster Heads (CH):** It is responsible for the cluster coordination and the communication with other clusters and network infrastructures. Moreover, the CH has other tasks, such as relaying information between nodes in the same cluster (intra-cluster communications) or between different clusters (inter-cluster communications). Compared to other nodes in the cluster, the CH has a number of additional functions, such as data aggregation and channel access management.

- **Cluster Members (CMs):** CMs are ordinary nodes that join a cluster according to their characteristics and similarities. They are responsible for sending their application-based information and data to the CH in specific time intervals. CMs in a cluster cannot communicate directly with other CMs or CHs from other clusters.

- **Cluster Gateway (CG):** It is a node that can provide inter-cluster communication and can communicate with neighboring clusters and transfer data between them. Typically, its position is at the end of clusters.

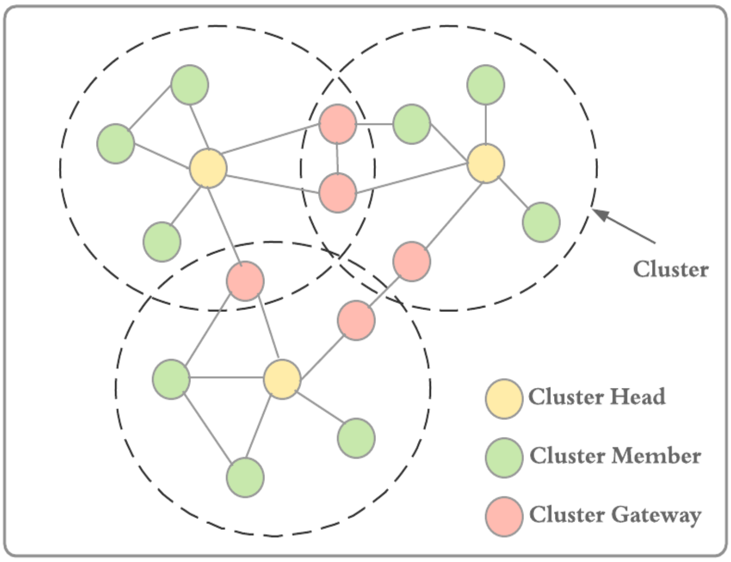


Figure 2.1: Example of clusters structure.

2.2.2.2 Cluster heads election criteria

According to the available literature, numerous metrics and criteria have been used to elect the Cluster Heads. There is a number of categories according to the type of metrics used when electing CHs [57]. For this reason, we propose to classify clustering algorithms into six categories. These categories are differentiated by the type of metrics used when selecting Cluster Heads. These six categories are:

1. **Neglected metrics:** This category groups algorithms that elect CHs without using any selection metrics. Typically, these schemes use heuristic algorithms. The latter have one or both of the following objectives: finding a solution with

reasonable run-time or finding the optimal solution. Moreover, heuristic algorithm leads to reasonable performances [58].

2. **Random metrics:** We can include in this category the early proposed clustering algorithms in the literature for Ad-hoc and MANET networks. These algorithms use metrics with random values, which are mostly not significant, such as nodes identifier [59].

3. **Position metrics:** This category includes all the algorithms, which are GPS-based on specific metrics related to the nodes position, such as geographical position, angle and euclidean distance.

4. **Mobility metrics:** This category includes all the solutions, which are based on mobility information related to mobile node, such as velocity, acceleration, sended/received signals power or link stability [60].

5. **Combined metrics:** This type refers to the weighted algorithms that combine multiple metrics of different types, such as degree, k-degree and density to elect the set of CHs.

6. **Destination metrics:** This class includes all the clustering algorithms, which are based on mobility information related to the destination object, such as relative speed, direction and current location.

2.2.2.3 Cluster radius

The cluster radius represents the maximum distance that separates the CH and its members. Typically, it is expressed in number of hops. Therefore, we then distinguish two main categories of algorithms: 1-hop algorithms and k-hop algorithms [61].

2.2.3 Clustering procedural flow

Typically, the clustering process consists of two main phases: cluster establishment and cluster maintenance phase. Therefore, a series of basic procedures is involved in the establishment and the maintenance of clusters, which may need to be repeated according to the conceptual rules of the algorithm and to meet the network requirements. According to Cooper et al. [8], the clustering process is composed of five main steps:

neighborhood discovery, CHs election, announcement, affiliation and maintenance as shown in Figure 2.2 as flowchart diagram.

2.2.3.1 Neighborhood discovery

Initially, when a vehicle enters the road and decides to join the network, its communication system is turned on. Then, to announce its existence, each vehicle broadcasts a periodic message to its neighbors, while simultaneously receiving similar messages from its neighbors. This messages includes a number of information, such as id, position, velocity and transmission range. Therefore, each node uses this gathering information during the clustering process. At the end of this step, the node will proceed to the cluster head election step.

2.2.3.2 Cluster heads election

After gathering the neighborhood information, a node will then check this information to elect the suitable node to act as its CH. The rules of choice of CH varies depending on the selected metrics, the clustering solution and the running application. If node itself is more suited to be CH, it will updates its state to CH and proceed to the announcement step; otherwise, if the chosen CH is found within the neighbor list, a node will proceed to the affiliation step.

2.2.3.3 Announcement

Each new CH must announce its election, therefore it broadcasts an announcement message to its neighbors to begin the formation and affiliation process. On the other hand, when the CH node accrued cluster members, it will proceed to the maintenance step.

2.2.3.4 Affiliation

When a node receives the announcement message from the CH, it compares this latter with its CH already elected in the previous step. If they are the same, the node sends a reply message to this CH, updates its state to CM, and therefore joins the cluster. Otherwise, the node ignores this event and waitfor another announcement message or moves directly to the maintenance step to join the appropriate cluster.

2.2.3.5 Maintenance

This step is different according to whether the node has become a CH or CM:

- As a CH: each CH monitors its CMs through the exchange of periodic messages to record the presence of members in the cluster. When a CM node moves out of the range of cluster, the CH detects this event and removes immediately this node from its members list. In addition, when two neighbor clusters have a big overlapping rate, clusters merging process will commence. Typically, the merging procedure produces two CHs at the same time for the final larger cluster obtained. On the other hand, when the CH receives a join request from a unclustered node, it must examine if the node is suited to join its cluster. If it is the case, the CH adds this node to its members list, then inform this node by sending a reply message.

- As a CM: the CM node will periodically examine the communication link with its CH through periodic messages by waiting for a reply from its CH. If the link is lost, the CM node must change its state to unclustered and try to join another cluster. On other hand, when an unclustered node approaches towards a cluster, it sends a join request to the CH of this cluster. Then, if the node receives a reply message from this CH, it joins the cluster. Otherwise, the node must look again for another cluster to join it.

2.2.4 Performance metrics for clustering algorithms

In order to validate the proposed clustering algorithm, it should be evaluated according to different performance metrics. Generally, these metrics are used in comparisons with other solutions that already exist in the literature. The main objective of these comparisons is to demonstrate the improvements performed by the corresponding proposed algorithms. According to the works of Fauziyyah et al. [62], Amudhavel et al. [63] and Guizani [64], we have deduce the following groups of metrics:

- **Messages complexity:** It is the number of control messages received by the nodes to accomplish a valid re-organization of the clusters when the network topology has been changed.

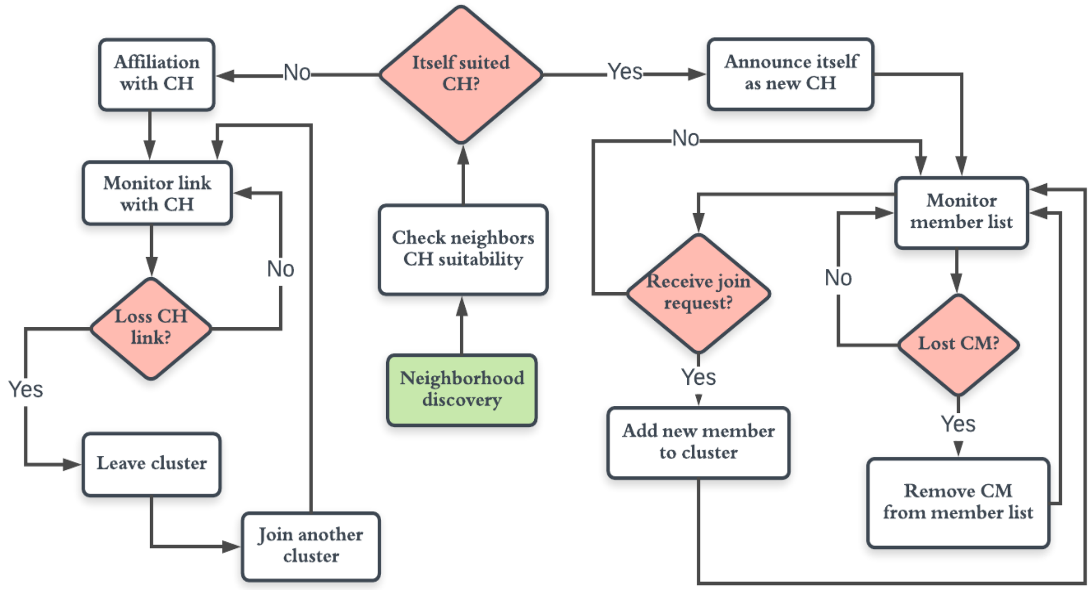


Figure 2.2: Clustering procedural flow.

- **Network overhead:** It is the amount of messages related the clustering process and resources used by every node during the network operation time such as bandwidth, energy, memory, time, etc.
- **Cluster structure stability:**
  - **Cluster Head Lifetime:** It is the time period from when a node changes its state to CH until the node leaves this state and changes to another state. Average cluster head lifetime is calculated by dividing the total cluster head lifetime by the total state change times from CH to another state.
  - **Cluster Member Lifetime** : It is the time period from when a node changes its state to CM to when this node switch off this state. Average cluster member lifetime is calculated by dividing the total cluster member lifetime by the total number of state change times from CM to another state.
  - **CH Change Number**: It represents the rate of state change times from CH to another state.
  - **Cluster Number**: It is the number of clusters constructed during the network operation time. A fewer clusters improves the efficiency of the clustering algorithm.

  - **Average size of cluster**: It is the average number of CM nodes managed by a CH during the total network operation time.

- **Time complexity:** Complexity in time for a topology change that represents the time spent performing a valid re-organization of clusters after a change in network topology.

2.3 Related work

Several taxonomy for clustering algorithms in VANETs have been proposed in the literature [7, 8, 46–48], where the authors categorize the algorithms based on different metrics parameters.

According to various key parameters, Bali et al. [46] classify them into six categories. They introduce predictive clustering including three sub-categories: position-based, destination-based and lane-based algorithms; backbone-based clustering including k-hop based algorithms, MAC-based clustering including IEEE 802.11 MAC based, TDMA-based and SDMA-based algorithms; traditional clustering including active and passive algorithms, hybrid clustering including intelligence-based, distributed and driver behavior-based algorithms; and secure clustering including authentication-based algorithms. The authors provide also a detailed description of each algorithm with a comparison of each categories with respect to various key parameters.

Accoridng to the clustering application, Cooper et al. [8] classified them into eight classes: general purpose applications, routing applications, channel access management applications, security applications, QoS applications, traffic safety applications, topology discovery applications and combination with cellular infrastructure application.

A survey on clustering protocols in VANETs was introduced by Yang et al. [7]. The objectives, challenges and issues for clustering protocols in VANETs were discussed. Moreover, the authors compared these protocols according to various parameters, such as relative velocity, nodes' density, size of clusters, hop distance and cluster establishment methodology.

Hosmani et al. [47] surveyed a number of clustering algorithms for VANETs and classified them based on cluster formation and CH election parameters. Moreover, the authors provide a comparison for these clustering algorithms based on various key metrics, such as nodes' density, scalability, velocity, cluster lifetime and feasibility.

Pal et al. [48] introduced an analytical model to evaluate the performance of cluster based algorithms in VANETs according to three important parameters: throughput, packet delivery ratio and end-to-end delay.

Finally, we summarize the previous survey papers related to ours in Table 2.1, which highlights the main contributions of those survey papers, including their year of publication.

2.4 Taxonomy of clustering algorithms in VANETs

Based on the cluster heads election metrics and cluster radius discussed above, a new taxonomy for VANET clustering algorithms is introduced, which is illustrated in Figure 2.3. Next, we discuss in depth the different VANET clustering algorithms in detail based on the proposed taxonomy.

2.4.1 Heuristic clustering algorithms

The heuristic is mathematical optimization aspect utilized for algorithms that attempt to find solution for a given problem among the possible ones, but they do not guarantee that the optimal one will be found, so they can be considered as approximate and not precise solutions. These type of algorithms, usually find a solution close to the best one and they find it fast [65]. Moreover, these algorithms are not based on particular metrics (metrics neglected). In this respect, many researches have addressed the design of heuristic clustering solutions to find sub-optimal solutions in a reasonable time period. However, most of the existing heuristic clustering solutions suffer from the problem of being sensitive to the initialization and do not guarantee high level of performance and services. According to the current literature, these algorithms tackle a number of graph theory problems, such as spanning tree construction, graph traversals, maximal independent sets and dominated sets in order to apply them in Ad-hoc networks and especially in VANETs networks. A number of proposals in this category are described next.

Graph domination is one of the main problems addressed largely by the heuristic algorithms. In this regard, Cha et al. [66] presented a new Connected Dominating Set

Table 2.1: Summary of previous surveys on clustering algorithms in VANETs.

Authors	**[Year]:Literature**	**Contributions**
Bali et al.	[2014]:Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions [46].	▶ Clustering in VANETs with respect to various challenges is explored. ▶ Detailed taxonomy for clustering algorithms in VANETs is provided. ▶ Comparison of different clustering approaches based on different key parameters.
Cooper et al.	[2017]:A Comparative Survey of VANET Clustering Techniques [8].	▶ Discussion of basic concepts related to the clustering process. ▶ A new taxonomy for clustering algorithms in VANETs is provided. ▶ Comparison of various clustering algorithms based upon some predefined metrics.
Yang et al.	[2015]:Clustering algorithm in VANETs: A survey [7].	▶ According to various parameters, a number of clustering algorithm is surveyed. ▶ An analysis of different VANETs clustering algorithms is provided.
Hosmani et al.	[2017]:Survey on cluster based routing protocol in VANET [47].	▶ Presentation of a taxonomy of typical clustering algorithms. ▶ Discussion of the challenging factors in VANET clustering.
Pal et al.	[2018]:Analytical model for clustered vehicular ad hoc network analysis [48]	▶ An analytical model to evaluate the performance clustering algorithm in VANETs. ▶ Comparison of clustered VANET performance with that of the non-clustered scenario. ▶ Analytical results are validated with simulation results.

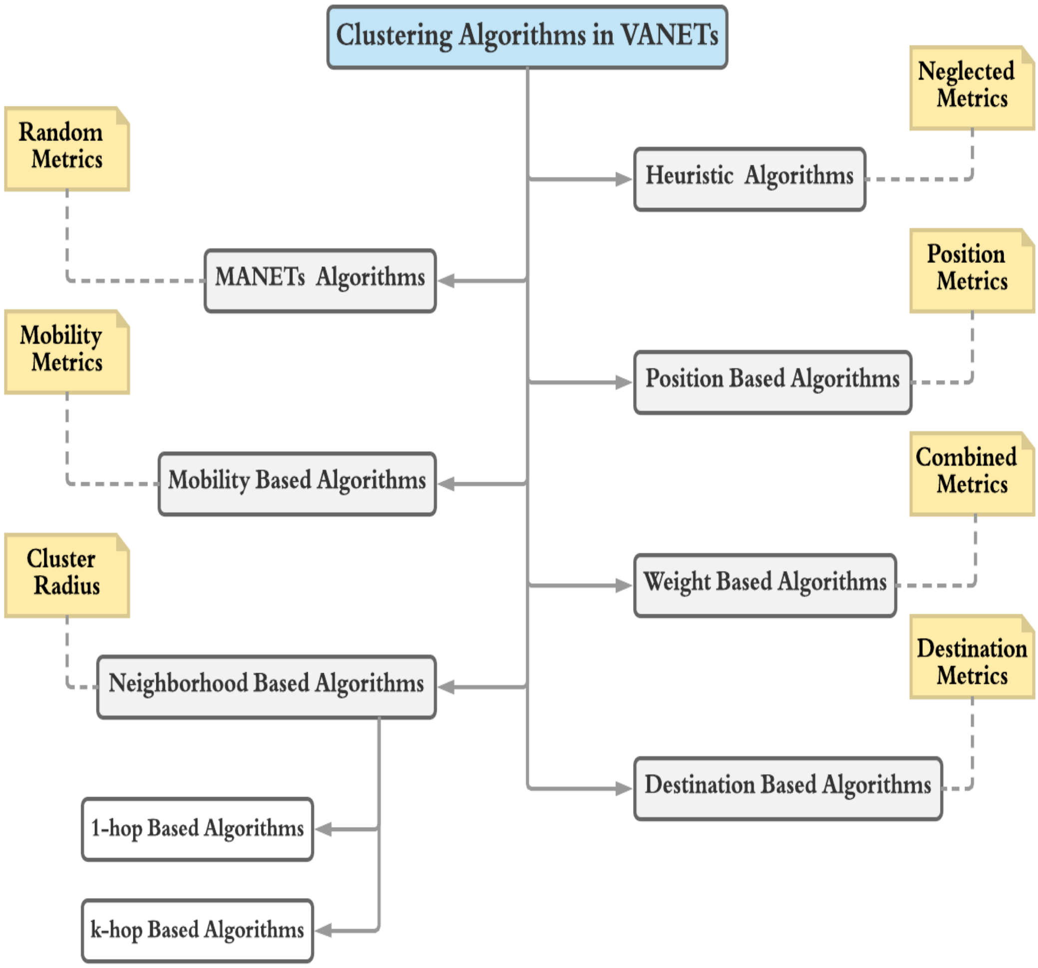


Figure 2.3: Taxonomy of clustering algorithms in VANETs.

(CDS) based clustering scheme to prevent and confront the broadcast storm problem in VANET networks. CDS scheme considers the high mobility and the connectivity of vehicles, and matches the dynamic topology of VANETs. CDS algorithm is a typical information dissemination techniques and suitable for different VANETs applications, such as emergency notifications and traffic safety. Moreover, a virtual core network can be constructed by searching for a CDS in the network graph. Therefore, the clustering using a CDS-based virtual core network can improve significantly the VANETs performance.

Yan et al. [67] introduced a new Distributed and Weighted Clustering algorithm based on Mobility Metrics (DWCM). Based on dominating set problem in graph theory, a distributed algorithm for cluster formation and cluster head election is proposed, where vehicles in the k-hop dominating set are selected as new CHs. Moreover, the cluster maintenance phase handles the topology changes caused by the high mobility of vehicles. Therefore, the main goal of DWCM solution is to construct and maintain stable k-hop clusters without incurring tremendous network overhead.

In Togou et al. [68], the authors developed a new Stable CDS-based Routing Protocol (SCRP). It is based on a distributed geographic location source forwarding technique that takes into account the vehicular network topology to select the data routing paths with low end-to-end delay. To achieve this objective, SCRP algorithm constructs stable backbones over road segments including two main metrics: vehicles' velocity and their spatial distribution. These backbones are connected at intersections via special gateway nodes that ensure an up-to-date vehicular network topology and monitor the expected delay for data routing on the road segments. According to this information, the SCRP assigns weights to the road segments. Therefore, the road segments with lowest weights are selected to form the routing paths. In this regard, the SCRP solution can avoid the local maximum problem and ensure load balancing on all possible routing paths.

Spanning tree construction problem, is also tackled largely by heuristic clustering algorithms. In this respect, Kponyo et al. [69] presented a VANET Cluster-on-Demand (CoD) Minimum Spanning Tree (MST) approach, which aims to construct clusters taking into account the intra-cluster quality of service (QoS). The main goal of this algorithm is that the clusters are constructed only to relay information on variable traffic density within the CH coverage area. Moreover, the proposed algorithm which implements the MST makes sure the size of clusters constructed does not affect the QoS in the network.

In Krishnakumar et al. [70], the authors proposed a new QoS enabled data dissemination using an improved Kruskal's algorithm to provide an efficient data dissemination and acceptable QoS in hierarchical VANET. The proposed algorithm extracts the minimum spanning trees based on Kruskal's algorithm in each road segment, where the nodes has been clustered according to an fuzzy c-means clustering technique taking into account the intra-cluster QoS. Therefore, every spanning tree will have a CH that is responsible to gather the data from the leaf nodes and forwards it to other coordinator nodes and vice versa.

Table 2.2: Comparison of heuristic based clustering algorithms.

Algorithm	Density	Stability	Latency	Overhead	Overlapping
CDS [66]	Medium	Medium	High	Medium	High
DWCM [67]	High	Medium	Medium	Low	High
SCRP [68]	Medium	High	High	Medium	Medium
CoD [69]	Medium	High	High	Low	High
QoS-VANET [70]	Medium	Medium	Medium	Low	High

2.4.1.1 Challenges and open research issues in heuristic clustering algorithms

The heuristic algorithm is a key solution for clustering in VANETs to improve their performances. However, most of the existing heuristic clustering schemes suffer from the problem of being sensitive to the initialization and do not guarantee high quality services and performance. Moreover, these existing algorithms suffer from the following two main issues: the data congestion and local maximum problem. Furthermore, these algorithms do not consider metrics related to mobile nodes, which makes the efficiency of such algorithms very weak. Table 2.2 compares the heuristic clustering algorithms according to a number of key parameters. From Table 2.2, it can be concluded that clusters overlapping and latency need to be further improved in order to the heuristic based clustering algorithms can be used effectively for vehicular networks.

2.4.2 MANETs clustering algorithms

Since the VANET is a subclass of MANET, convetional clustering algorithms designed for MANET might be applied in VANET networks with some modifications. The vehicles in VANET are characterized by their high mobility which implies a very dynamic topology. Moreover, the position of vehicles in the same geographical proximity does not guarantee that have the same mobility patterns. Thus, in order to apply MANET algorithms on VANET, researchers must take into consideration the specific characteristics of VANET networks, such as mobility, direction and location as well [56]. Typically, these algorithms use metrics with random values, which are mostly not significant, such as nodes identifier. A number of proposals in this category are described next.

One of the most popular algorithm is the Lowest-ID (LID) [51], where every node is assigned a unique ID. In order to announce its existence, each node in the network broadcast an "hello" message containing its ID, within a period called "Hello period" (HP). Then, each node constructs a table, that contains the neighbors ID information. Finally, the node that has the lowest ID will be elected as new CH; nodes which can hear two or more CHs at the same time become gateways, and the remaining nodes are cluster members.

In Gerla et al. [71], the authors presented Highest Degree (HD) algorithm. In order to perform the clustering process, each node broadcasts a beacon message periodically including its degree value to its direct neighbors, to compare its degree value with them. A node having the biggest connectivity rate in the neighborhood becomes the CH and the remaining neighboring nodes become members.

In Nguyen et. [59], the authors attempt to apply the lowest-ID algorithm in cluster-based TDMA system for VANETs. In this respect, a new mechanism for electing the CHs, which inherited the time division in TDMA- and MAC-frame format [72]. The algorithm uses the conventional lowest-ID algorithm to improve the network latency and to reduce significantly the number of re-clustering times. Therefore, the cluster is efficiently formed. Moreover, the algorithm uses cooperative MAC for inter-cluster communication to avoid collision and overlapping between clusters.

Gavalas et al. [73] introduced a new efficient distributed clustering scheme that uses a number of mobility metrics to construct stable clusters. According to the cost-efficient lowest-ID technique, CHs are initially elected. During the clustering maintenance stage, the node IDs are re-assigned based on different mobility parameters. As a result, the nodes with low mobility are assigned low IDs and, therefore are selected as CHs. Moreover, the proposed solution improves relatively the total network overhead because of use of the broadcast duration which is adjusted according to the network mobility model.

2.4.2.1 Challenges and open research issues in MANETs clustering algorithms

Although, VANET is a subclass of MANET and several VANET clustering algorithms are derived from MANET [16], but there are several differences between the two, such as the high mobility, dynamic topology and limited driving directions. Moreover, we

Table 2.3: Comparison of MANETs clustering algorithms.

Algorithm	Density	Stability	Latency	Overhead	Overlapping
LID [51]	Low	Low	High	High	High
HD [71]	Medium	Low	Medium	High	High
LID-TDMA [59]	Low	Medium	High	Medium	Medium
Adaptive-LID [73]	Medium	Low	High	Medium	High

notice that the MANET algorithms usually use arbitrary metrics and do not take into consideration the mobility patters of VANET nodes. As a result, the application of the MANETs clustering algorithm in VANET remains very limited, inefficient and does not meet the main requirements of this type of networks. Table 2.3 compares the MANETs clustering algorithms. From Table 2.3, it can be concluded that clusters stability, latency, total clustering overhead and clusters overlapping need to be further improved in order to the MANET clustering algorithms can be used effectively for VANET networks.

2.4.3 Position based clustering algorithms

Position based clustering algorithms use location and topology metrics, such as geographic position and spatial variation to perform the clustering process. They are among the key solutions for clustering in VANETs. In this regard, many clustering position algorithms have been proposed. These proposed solutions have greatly improved the VANET network performances due to their resistance to handling the vehicles position variation [46]. A number of position based algorithms are described next.

Wang et al. [74] proposed a new Position Based Clustering Technique for Ad Hoc Inter-vehicle Communication (PCTIC), which is a new clustering algorithm for large multi-hop VANET. The clusters are formed according to the geographic position of vehicles and the priorities assigned to the vehicles and traffic information. The algorithm elects one node to act as CH for each cluster based on the Minimal Dominating Set (MDS) problem. The size of cluster is defined according to the maximum distance between the CHs and their members. Thus, the clusters are independently monitored and dynamically re-configured as vehicles move and direction. As a result, vehicles are enables to move during cluster establishment and maintenance phases.

Liu et al. [75] introduced a Position Sensitive Clustering Algorithm (PSCA), which is based on the configuration and the section position information of the road, and the dynamic timestamps. In PSCA algorithm, the CH periodically broadcasts a control message at a certain time-stamp by direction-based broadcasting. As a result, each cluster member must send its position informations in real time to the CH before the next time-stamp, so that the CH stores and reuses this informations in the next iteration of the clustering process. The initial choice of CHs is performed according to the principle of "Randomly Competition, First Declare Win". Moreover, the PSCA also provides a special maintenance mechanisms that facilitates the re-configuration of cluster in the case where the CH leaves the cluster or new vehicles want to join the cluster.

In Alloulche et al. [76], the authors presented a new Cluster-Based Beacon Dissemination Process (CB-BDP). The main goal of this latter is to provide vehicles with more accurate information on the position of neighboring vehicles. The CB-BDP algorithm is designed to achieve two objectives. The first one is that the proximity map be wide and as accurate as possible. Secondly, this map is cooperated with vehicles located nearby, hereby to allow coordinated and synchronized reactions of nearby vehicles to the hazardous situations. In addition, CB-BDP algorithm proposed a new cluster-based gathering-based tagging process based on an optimized topology for sharing the informations collected by the proximity maps.

In Maslekar et al. [77], a new cluster-head selection policy for direction-based clustering approach, called Modified Clustering based on Direction in Vehicular Environment (MC-DRIVE) is introduced. This dissemination algorithm is based on another clustering protocol, called C-DRIVE [78]. MC-DRIVE approach constructs clusters that are composed of nodes close to the intersections and according to their future directions (left, right, straight, half-turn) after the next intersection. Moreover, the size of cluster is limited to a certain distance from the intersection. At this distance, every node broadcasts a hello message including its ID, driving lane and future direction provided that it finds another vehicle in front of itself.

Table 2.4: Comparison of position based clustering algorithms.

Algorithm	Density	Stability	Latency	Overhead	Overlapping
PCTIC [74]	Low	High	Low	Medium	High
PSCA [75]	High	High	Medium	High	Medium
CB-BDP [76]	Low	Medium	Low	High	Medium
MC-DRIVE [77]	Low	Medium	Low	High	High

2.4.3.1 Challenges and open research issues position based clustering algorithms

Position based algorithms present an effective solution for improving VANET networks performances. However, in the areas with high mobility and dynamic topology, position metrics become irrelevant, where the variation of position increases enormously with time. Moreover, position based algorithms assume that every vehicle is equipped with GPS system, which is not practically true. Furthermore, positioning services accuracy can significantly vary. Table 2.4 compares the position based clustering algorithms. According to Table 2.4, it can be concluded that clustering overhead and clusters overlapping need to be further improved in order to the position based clustering algorithms can be used effectively for vehicular networks.

2.4.4 Mobility based clustering algorithms

As VANETs are characterized by high mobility, dynamic topology, and limited driving directions, a number of mobility based clustering algorithms are recently proposed in the literature. These algorithms use metrics that relative to the mobile node (mobility metrics), such as relative velocity, acceleration and direction to perform the clustering process. A number of these proposals in this category are discussed next.

Hafeez et al. [79] proposed a new a fuzzy-logic-based clustering algorithm for VANETs where CHs are elected according to the relative speed and average distance from the neighbors nodes within the neighborhood. Moreover, the proposed protocol strengths the clusters stability through the election of secondary CH (to be used as CH when the primary one becomes unavailable). Furthermore, the maintenance step is adaptable to drivers behavior on the road based on a fuzzy logic inference system. Therefore, the algorithm is suitable to be applied in environments with high mobility.

In Ren et al. [80], the authors introduced an efficient Dynamic Mobility-based Clustering (DMC) algorithm for forming a stable core network for future data aggregation and dissemination. The proposed algorithm performs the clustering process according to the vehicles' mobility patterns, taking into account a number of metrics, such as moving direction, relative speed, relative average distance and link stability. Moreover, the algorithm proposed a new dynamic cluster formation phase, compared with existing algorithms where vehicles are static during this phase. Moreover, a "temporary cluster head" method is proposed to improve the cluster formation phase. Furthermore, the algorithm introduces a "safe distance threshold" in order to monitor the cluster size.

Morales et al. [81] presented an Adaptable Mobility-Aware Clustering Algorithm based on Destination positions (AMACAD) for accurately following the network mobility patterns to improve clusters' stability. AMACAD algorithm considers a number of destination metrics, such as the current position, relative velocity, final destination position of vehicles for performing the clustering process. In this respect, the algorithm provides a natural model of location references for improving the clusters' stability and the network performances. Therefore, the information is disseminated by clusters, this allows improving network latency, reliability and data delivery ratio.

Souza et al. [82] presented a new Aggregate Local Mobility (ALM) clustering algorithm that aimed at prolonging the network lifetime in VANETs. The ALM algorithm uses an aggregate local mobility mechanism for controlling and monitor the existing CHs. To elect the CHs, each vehicle calculates its mobility variance over all neighbors. A lower variance means less mobility and vehicle is more stable. As a result, a vehicle with less variance relative to its neighbors is suitable to be elected as new CH.

2.4.4.1 Challenges and open research issues in mobility based clustering algorithms

Table 2.5 compares mobility based clustering algorithms. From Table 2.5, it can be concluded that mobility-based clustering algorithms are more suitable for VANETs networks. Moreover, an improvement is observed on the clusters convergence thanks to the minimization of relative mobility as well as the distance of the CHs to theirs cluster members. As shown also in Table 2.5, the reviewed algorithms have a significant improvement compared to other types of algorithms. These improvements helps to meet the requirements of the new features of VANET, and to consider its mobility

Table 2.5: Comparison of mobility based clustering algorithms.

Algorithm	**Density**	**Stability**	**Latency**	**Overhead**	**Overlapping**
FUZZY [79]	Low	Medium	Medium	Medium	Low
DMC [80]	Medium	Medium	Low	High	Medium
AMACAD [81]	Low	Medium	Medium	Medium	Medium
ALM [82]	Medium	Low	Low	Medium	Low

characteristics. As a result, mobility based clustering algorithms are more suitable for VANETs, especially in the areas with very high mobility. However, the clustering overhead and clusters stability need to be further more improved.

2.4.5 Weight based clustering algorithms

Rather than using a single metric to elect CHs, the weight based clustering algorithms use a function that combines multiple metrics. This function represents the weight of a node. The general formula of the weight $W(i)$ for a given node i according to the metrics $A(i)$, $B(i)$ and $C(i)$ is as follows:

$$W(i) = \alpha \times A(i) + \beta \times B(i) + ... + \gamma \times C(i) \tag{2.1}$$

With α, β and γ are coefficients, where: $\alpha + \beta + \gamma = 1$.

This category of algorithms aims to elect a set of cluster-heads that are best suited to an environment to meet the requirements of the network environment. Therefore, a number of algorithms are proposed for VANETs that are described next.

A new Weight Clustering based TDMA-MAC Scheme (WCS) for VANETs is proposed in Xie et al. [83]. When selecting the set of CHs, WCS algorithm uses indexes and an entropy weight that combines a number of parameters, such as constrains of radio signal transmitting power, vehicles energy consumption and vehicles mobility. Based on the TDMA-MAC technique, a realistic clustering channel access mechanism is introduced for reducing the chance of collision and guarantee efficient end-to-end communication in VANET networks.

Hadded et al. [84] introduced a new Angle based Clustering Algorithm (ACA) for VANETs, which exploits the angular position and the direction of vehicles for electing the most stable vehicles to act as CHs as long as possible. In ACA algorithm, two or more vehicles can belong to the same cluster, if and only if the angle between their velocity vectors is acute.

A multi-objective genetic-based algorithm Adaptive Weighted Clustering Protocol (AWCP) [85] is designed for VANETs, which elects the CHs based on a weight that combines a number of key parameters, such as highway ID, vehicles direction, vehicle location, vehicle velocity and vehicles connectivity. Therefore, the main goal of AWCP algorithm is to improve the stability of the network topology. AWCP algorithm defines a genetic multi-objective method whose inputs are the metrics used to calculate the nodes weights. This method aims to provide a stable clusters structure as long as possible, improve the data delivery rate, and reduce the total clustering overhead.

In Khan et al. [86], the authors proposed a new mechanism based on Moth Flame Clustering optimization for Internet of Vehicles (MFCA-IoV). The authors introduced a Moth Flame Optimizer (MFO), which is a nature inspired process. The MFCA-IoV algorithm optimizes the generation of clusters to ensure a robust transmission.

Aadil et al. [87] proposed a new Clustering Algorithm for Internet of vehicles based on Dragonfly Optimizer, called CAVDO. The authors use a swarm-based multi-objective method to perform the clustering process for improving the topology stability in a dynamic environment. In addition, the CAVDO algorithm introduce a new Mobility Aware model (MA-DTR), which makes the transmission range more dynamic.

2.4.5.1 Challenges and open research issues in weight based clustering algorithms

Weight based algorithms present an effective solution to improve the performances of VANET networks. Typically, this type of algorithm use a multi-metric mechanism to perform the clustering process. However, the use of multiple parameters at once, makes the tuning of parameters a non-trivial problem in such conditions. Table 2.6 compares the weight based clustering algorithms. From Table 2.6, it can be concluded that clustering overhead needs to be further improved in order to the weight based clustering algorithms can be used effectively for vehicular networks.

2.4.6 Destination based clustering algorithms

Destination based clustering algorithms take into account metrics related to the final destination of vehicles, such as current geographical position, direction and relative

Table 2.6: Comparison of weight based clustering algorithms.

Algorithm	**Density**	**Stability**	**Latency**	**Overhead**	**Overlapping**
WCS [83]	Medium	Medium	Medium	High	Low
ACA [84]	Low	Medium	Low	Medium	Low
AWCP [85]	High	High	Medium	Medium	Medium
MFCA-IoV [86]	Medium	High	Low	High	Low
CAVDO [87]	High	High	Low	Medium	Medium

velocity to perform the clustering process. The destination is known in prior using GPS device and navigation systems. A number of algorithm in this category are described next.

Farhan et al. [88] proposed a Robust Localization using Cluster Analysis called LICA to improve the GPS devices accuracy. LICA algorithm utilizes a new modified tri-lateration mechanism for constructing a set of possible coordinates (x, y) based on a cluster analysis mechanism, allowing accurate data according to the given weights, resulting in each vehicle to improve their location estimate. Moreover, unreliable data from malicious vehicles are removed during the clustering process. As a result, vehicles are able for collecting real-time, reliable data and forward them to other vehicles, helping vehicles to reach their destinations safely.

Sethi et al. [89] presented a Destination Based Routing (DBR) algorithm for context based clusters in VANETs. DBR protocol provides two mechanisms for minimizing the data traffic and communication end-to-end in VANETs. Firstly, a context based clustering is designed to consider the various metrics about the destination vehicles: cluster formation-location, direction , relative velocity, interest list and final destination. On the other hand, a destination based routing protocol is proposed for these context based clusters to improve the inter-cluster communication. By using this context based clustering including interest list of vehicles, the overall end-to-end communication to relay the information from source vehicle to destination one is significantly improved.

Aravindhan et al. [90] designed a Destination-Aware Context-based Routing (DACR) scheme using a soft computing clustering algorithm for VANET. The authors introduced also two soft computing approaches. First, a hybrid clustering method is designed which combines the geographic and the context clusterings. As a result, this hybrid clustering aims to reduce the clustering overhead and avoid the network congestion. Second, the destination-aware routing protocol is introduced for inter-clusters communication for improving the overall packet delivery ratio and the end-to-end delay. DACR protocol improves significantly the inter-cluster routing by using the optimal next forwarding vehicle selection.

A new clustering routing algorithm for VANETs is based on the Euclidean distance is presented in Tian et al. [91], which uses the location and the direction information to group the vehicles into self-organized clusters. So, if the source vehicle and the destination one are located in different clusters, the direction information is utilized for restricting the data routing path along the same direction, which not only can help for finding a shorter and more stable route, but also can improve the flooding of the routing control message. To perform the clustering process, each vehicle announces its existence by broadcasting a beacon message including its geographical position, direction and current time. After receiving the beacon messages from the neighbors, every vehicle in the network updates its topology table. Therefore, the CHs are generated by electing the vehicles that having the minimum distance parameter.

2.4.6.1 Challenges and open research issues in destination based clustering algorithms

Destination algorithms use a various metrics related to the destination to form a stable clusters. In this repect, the ratio of cluster head changes is significantly reduced which make the clusters more stable. This reduction is justified by the fact that a vehicle can leave the cluster only when it encounters a CH whose destination is more similar to the destination of the current CH. Moreover, the destination based algorithms are based on the behavior of the vehicles taking into account a number of metrics related to the destination of vehicles, which allows to improve significantly the network performances [46]. This type of algorithms has a number of drawbacks. Among these latter, the large delay of transmission and global network overhead generated when the density of vehicles becomes high. This affects the network performances in a negative way. Moreover, this type of algorithms assumes that each vehicle is equipped with GPS system, which is not practically true. Table 2.7 compares the destination based clustering algorithms. From Table 2.7, it can be concluded that clustering overhead and clusters overlapping need to be further improved in order to the destination based clustering algorithms can be used effectively for vehicular networks.

Table 2.7: Comparison of destination based clustering algorithms.

Algorithm	Density	Stability	Latency	Overhead	Overlapping
LICA [88]	Medium	Medium	Medium	High	High
DBR [89]	Low	Medium	Medium	Medium	Medium
DACR [90]	Medium	Low	Medium	Medium	Medium
EUCLID [91]	High	High	Low	High	High

2.4.7 Neighborhood based clustering algorithms

According to the cluster radius, the proposed clustering algorithms can be categorized into two main categories: 1-hop clustering and k-hops clustering algorithms. Many of these algorithms construct clusters with 1-hop, where each node is one hop from its CH. On the other hand, a number of algorithms generate clusters with k-hop distance, where every node is at most k hops from its CH.

2.4.7.1 1-hop based clustering algorithms

Ahwazi et al. [92] presented a new MObility aware and SIngle-hop Clustering scheme (MOSIC) for VANETs suitable for highway environments. MOSIC algorithm presents a new clustering methodolgy that uses Gauss Markov mobility model (GMM) [93] for mobility prediction that enables the vehicles to prognosticate their relative mobility to their neighbors. Moreover, the GMM uses a memory-based mobility model, which is able for calculating the next hop vehicle position according to its current mobility metrics.

In Caballero-gil et al. [94], the authors proposed a new Self-Organized Clustering Algorithm (SOCA) for VANETs which based on 1-hop clustering for reducing the network overhead in dense road traffic areas and maintaining the communications security using a combining of the cryptography public- and secret-key. To elect the CHs, SOCA algorithm uses an improved algorithm of independent set problem with a secret-key agreement technique based on the generalized Diffie-Hellman protocol. Moreover, SOCA constructs a dynamic virtual backbone in the vehicular network. This backbone network is composed of all CHs and CGs already elected in the cluster formation phase.

Goonewardene et al. [95] presented a new clustering algorithm named Robust Mobility Adaptive Clustering (RMAC) to effectively enable highly dynamic VANETs for future ITS. RMAC algorithm uses a new vehicle priority method to adaptively and eas-

Table 2.8: Comparison of 1-hop based clustering algorithms.

Algorithm	**Density**	**Stability**	**Latency**	**Overhead**	**Overlapping**
MOSIC [92]	Low	Medium	Medium	Medium	High
SOCA [94]	Medium	Medium	Low	Low	High
RMAC [95]	Low	Low	Medium	Medium	High

ily identify 1-hop neighbors and select optimal CHs according to a number of metrics such as: of the relative speed of the vehicles, their position, and their direction. Moreover, RMAC introduces a new concept named zone of interest that helps the vehicles to maintains their neighbors table.

Challenges and open research issues in on 1-hop based clustering algorithms

According to the available literature, a number of 1-hop based clustering algorithm are proposed. These algorithms provide a very efficient coordination to CHs and a more reliable intra-cluster communications. However, the coverage area of clusters is small in such algorithms, which leads to increase the number of clusters formed and produce a high maintenance overhead. Therefore, the overlapping rate increases enormously. Table 2.8 compares the 1-hop based clustering algorithms. From Table 2.8, it can be concluded that clusters stability and clusters overlapping need to be further improved in order to the 1-hop based clustering algorithms can be used effectively for VANET networks.

2.4.7.2 k-hops based clustering algorithms

In Zhang et al. [96], the authors designed a new multi-hop clustering scheme for VANETs called (N-hop). This latter is considered as the first algorithm based on multi-hop clustering for VANETs. N-hop algorithm selects the vehicles, that have the low aggregate mobility as the CHs. The authors justified their choice because they believe that including the low aggregate mobility metric help to construct more stable clusters and topology.

Chen et al. [61] proposed a Distributed Multi-hop Clustering algorithm for VANETs based on Neighborhood Follow (DMCNF), which elects CHs according to the neighborhood follow relationship between vehicles. In DMCNF algorithm, a vehicle cannot directly identify the most suitable to be its CH in its multi-hop neighbors, although it can know the most stable and the most similar 1-hop neighbors, and therefore they probably belong to the same cluster. As a result, a vehicle can then elect its CH by following the most stable vehicle. This technique is identified as the neighborhood follow relationship.

Azizian et al. [97] introduced a new distributed D-hop clustering algorithm, called DHCV, which is based on the relative mobility information for constructing stable clusters. DHCV algorithm uses the velocity and the location differences of vehicles as metrics to perform the clustering process in D-hop communication range. To form multi-hop clusters, every vehicle in DHCV elects its CH according to the relative mobility within its D-hop neighbors.

A new Multi-hop Cluster-Based IEEE 802.11p and LTE Hybrid Architecture for VANET Safety Message Dissemination, namely, VMaSC-LTE, is proposed in Ucar et al. [98]. The authors claim that in networks based only on LTE technology, the delivery ratio and end-to-end delay of safety message relaying are degraded due to the broadcast storm and the disconnected network problems. Also the authors mention that a basic cellular based VANET communication is not feasible due to the high cost for installing and maintaining the infrastructures. Moreover, due to the high mobility of vehicles, a high number of hand-off occurrences at the base station is generated. In order to achieve a high data packet delivery ratio and improve the latency, the authors propose a new hybrid architecture, combining IEEE 802.11p based multi-hop clustering and LTE.

Challenges and open research issues in k-hops based clustering algorithms

Several k-hop clustering algorithms for VANETs have been proposed in the last few years. Compared to 1-hop clustering algorithm, these algorithms can extend the clusters coverage areas, decrease the number of clusters and improve the cluster stability. However, some issues remain in k-hop clustering for VANETs. For example, intra-cluster communication must be further improved and maintenance overhead must be improved. Table 2.9 compares the k-hop based clustering algorithms. From Table 2.9, it can be concluded that clusters stability and clustering overhead need to be further more improved in order to the k-hop based clustering algorithms can be used effectively for vehicular networks and environments.

Table 2.9: Comparison of k-hops based clustering algorithms.

Algorithm	Density	Stability	Latency	Overhead	Overlapping
N-hop [96]	Low	Medium	Low	High	Medium
DMCNF [61]	High	Medium	Medium	Medium	Low
DHCV [97]	Medium	Medium	Medium	High	Low
VMaSC-LTE [98]	High	High	Low	High	Low

2.4.8 General comparison of clustering algorithms in VANETs

Based on our review, we provide in Table 2.10 a general comparative analysis of various clustering algorithms in VANETs discussed above based upon various key parameters. The parameters selected for the comparison are: scalability, Cluster Radius, Load balancing, algorithm complexity and Quality of Service (QoS).

2.5 Conclusion

This chapter presents a deep review of clustering algorithms in VANET networks. Firstly, we introduced an overview on clustering technique in VANETs. In this respect, we present a background of knowledge about the clustering process; their brief historical, definitions, clusters structure, cluster head selection metrics, general procedural flow and performance evaluation metrics. Secondly, we proposed new taxonomy for clustering algorithms for VANETs with a detailed description of each algorithm. Thirdly, a detailed comparison is provided for each category of the proposed taxonomy considering relevant key parameters, such as vehicle density, cluster stability, latency, clusters overlapping and clustering overhead. Fourth, we highlight the main challenges encountered for each category and discuss some open research issues. Finally, we provided a general comparison of different clustering algorithms according to a number of selected parameters. This study allowed us to take advantage of the proposed algorithms and to take into account their deficiencies to propose new clustering algorithms suitable for the IoV environment. These algorithms aim to improve the conventional VANETs and generate a stable cluster structure. The design and evaluation of these algorithms makes the subject of the next part of our thesis.

Table 2.10: General comparison of clustering algorithms in VANETs.

Year	**Algorithm**	**Scalab.**	**Radius**	**Load balan.**	**Compl.**	**QoS**
Heuristic clustering algorithms						
2013	CDS [66]	Low	1-hop	Moderate	Moderate	No
2016	DWCM [67]	Moderate	K-hops	Good	High	No
2016	SCRP [68]	Low	1-hop	Good	Moderate	No
2013	CoD [69]	Moderate	1-hop	Good	Moderate	Yes
2017	QoS-VANET [70]	High	K-hops	Good	High	Yes
MANETs clustering algorithms						
1997	LID [51]	Low	1-hop	Bad	Low	No
1995	HD [71]	Low	1-hop	Bad	Low	No
2015	LID-TDMA [59]	Moderate	1-hop	Moderate	Moderate	No
2006	Adaptive-LID [73]	Moderate	1-hop	Bad	Moderate	No
Position based clustering algorithms						
2008	PCTIC [74]	High	K-hops	Good	High	No
2014	PSCA [75]	Moderate	1-hop	Moderate	Moderate	No
2015	CB-BDP [76]	High	1-hop	Good	Low	No
2011	MC-DRIVE [77]	Moderate	K-hops	Good	Moderate	Yes
Mobility based clustering algorithms						
2012	FUZZY [79]	Low	1-hop	Moderate	High	No
2016	DMC [80]	Moderate	1-hop	Bad	Moderate	No
2011	AMACAD [81]	Low	1-hop	Moderate	Low	No
2010	ALM [82]	Moderate	1-hop	Good	Moderate	Yes
Weight based clustering algorithms						
2016	WCS [83]	High	1-hop	Good	High	Yes
2017	ACA [84]	Moderate	K-hops	Moderate	Moderate	No
2015	AWCP [85]	High	K-hops	Good	High	Yes
2019	MFCA-IoV [86]	Moderate	1-hop	Good	High	Yes
2018	CAVDO [87]	High	1-hop	Good	High	Yes
Destination based clustering algorithms						
2011	LICA [88]	Moderate	K-hops	Good	High	No
2017	DBR [89]	High	K-hops	Moderate	High	Yes
2018	DACR [90]	High	K-hops	Good	Moderate	Yes
2010	EUCLID [91]	Moderate	1-hop	Bad	Low	No
1-hop based clustering algorithms						
2016	MOSIC [92]	Moderate	1-hop	Moderate	High	No
2015	SOCA [94]	Low	1-hop	Good	High	Yes
2009	RMAC [95]	Moderate	1-hop	Moderate	Moderate	No
k-hop based clustering algorithms						
2011	N-hop [96]	Low	K-hops	Bad	Low	No
2015	DMCNF [61]	High	K-hops	Good	Moderate	Yes
2016	DHCV [97]	Moderate	K-hops	Moderate	Moderate	No
2016	VMaSC-LTE [98]	High	K-hops	Good	High	Yes

Contributions

Chapter 3

A New Heuristic Clustering Algorithm Based on RSU for IoV

3.1 Introduction

Internet of Vehicle (IoV) is a very important promoter domain in the Intelligent Transportation System (ITS), covering a wide range of technologies and applications, including intelligent transport, cloud computing, vehicle information service, modern wireless technologies, Internet access and communication [17]. According to Wu et al. [99], the network structure of IoV is composed of four major components: On-Board Unit (OBU) in each vehicle, Road Side Unit (RSU), Control Center (CC) and Internet.

First, the OBU contains several modules: GPS device module, Vehicle-to-Vehicle and Vehicle-to-Infrastructure (V2V-V2I) communication module, Human Machine Interface Device (HMI) module, On-Board Diagnostics (OBD) module, GSM module and Input-Output (I/O) devices. Second, RSU is a roadside computing device that has a number of functionalities, such as traffic directories, data dissemination, security management, location servers, and service proxies [100]. RSU provides different services to vehicles such as: communication services, security services, data dissemination and aggregation services, and Internet access. In this respect, OBUs can exchange traffic information with RSU and all others infrastructure in the coverage area. Third, the CC is a main network component that performs various functions such as: supervising and controlling the RSUs, road traffic management, interconnecting the VANET network with other networks (e.g. cellular network) and facilitating the scaling of VANET network. Fourth, through the Internet, the messages could be shared between vehicles and different infrastructures and broad-casted more widely. Consequently, more information could be acquired.

IoV introduces the notion of encapsulation of communications, which enables the vehicles to communicate with different networks, adapt to them and use the services they provide. Due to different factors, such as the huge number of vehicles connected (scalability problem), the large volume of data (big data) to process and store, road conditions and traffic flows, and communication heterogeneity [3, 13], the design of effective applications becomes a challenge in IoV. As a reaction, several techniques are used among which is clustering. This latter, has recently, became a very effective approach to make the dynamic VANETs more manageable and stable with acceptable performance. Clustering is usable with different aims such as data dissemination and aggregation, network overhead minimization, road safety, drivers comfort and mainly routing schemes.

Aissaoui et al. [101] claim that clustering is the process of dividing the network into different groups of nodes called clusters. Each cluster has a member which plays the role of cluster head and allows communication among cluster members and among different clusters as well. Generally, the clustering algorithm is divided into two main steps: 1) cluster establishment, which in turn includes cluster heads selection and cluster formation, 2) cluster maintenance, which aims to maintain the stability of the clusters and recover failed links in internal or external communication in the cluster. Clustering has many advantages such as: the scaling up (scalability) advantage, which has to do with the division of the network into clusters to make it easier to manage and reduces the coordination messages exchanged between all nodes. The second advantage is the reduction of the execution load of the high-level protocols on all the individual nodes. Finally, in terms of radio transmission, the prioritization of access to channels makes it possible to limit access collisions and interferences.

Most of the proposed clustering protocols are based on totally distributed clustering approaches [61, 80, 92, 97, 102], which leads to a poor control and overloading the network caused by the lack of a global view of the network for each node. On the other hand, the centralized algorithm allows to improve network control and network overhead through the availability of a global view of the network by a central node or infrastructure (e.g., RSU). Moreover, these distributed clustering approaches rely on a huge number of broadcasting control messages to perform clustering algorithm, which has a negative effect on the network overhead. In addition, several proposed approaches do not take into account the unavailability of Cluster Head (CH) and the mobility characteristics of VANET, which leads to the use of re-clustering several times, affecting cluster stability and network performance negatively. Furthermore, according to the available literature, the proposed schemes neglect the RSU when performing the clustering process.

Nowadays, RSU is a must-have component especially with the evolution of conventional VANETs towards the IoV. It is used for different roles such as traffic directories, data dissemination, security management, location servers, network gateway and service proxies [100]. The basic motivation behind using RSUs to perform the clustering process is that RSUs are fixed infrastructure. It is much easier to send a message to a fixed target than to a moving one. In addition, there are many proposed RSU-based schemes [100, 103–105] hypothesizing that VANET network is totally covered by the RSUs. This hypothesis is justified by the importance of the RSU and the applications running on it. Mainly in critical applications such as safety applications, especially when the safety of the individuals is at stake. Moreover, these proposed RSU-based schemes clearly exhibit the efficiency of network performances. Furthermore, compared to the vehicles, the RSU has a wide transmission range leading to a significant reduction in network overhead since few control messages will be sent.

Considering all the above issues, this chapter aims to propose a new Heuristic one-hop Clustering Algorithm based on RSU called HCAR for IoV, which is suitable for highway environment. The proposed scheme focuses on Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. It uses centralized one-hop clustering method at the RSUs, using a simple heuristic algorithm based on graph theory concepts such as neighboring, adjacency matrix and node degree. To ensure good network coverage, the suggested approach assumes that several RSUs are distributed and installed in the highway area. Each RSU, taking charge of clustering in its coverage area, includes only cluster formation phase. The corresponding algorithm uses a new recovery method which is based on the election of the Secondary CH (SCH) using a weighted method that combines three metrics: node velocity, node degree and node transmission range. This method helps to detect the unavailability of primary CH and replace it quickly to increase cluster stability and avoid the re-clustering of the network. The proposed scheme uses distributed techniques to maintain the clusters in order to ensure clusters' stability and react to the frequent change in the network topology caused by the high mobility of the vehicles. In order to prove the workability of the proposed approach, simulations are conducted using the network simulator NS-2 and VanetMobiSim integrated environment. Simulation results show that the present scheme has better performances compared to the existing clustering schemes.

For a comprehensive view of the topic, the rest of this chapter is organized as follows: Section 3.2 describes the preliminaries. Section 3.3 gives details of the proposed approach. Section 3.4 provides the performance evaluation of the proposed scheme. Finally, a conclusion is presented in Section 3.5 with some highlighted future perspectives.

3.2 Preliminaries

This section describes the preliminaries of the proposed approach through presenting the network model and introducing the definitions and the notations used.

3.2.1 Network model

The proposed scheme is based on the following assumptions. First, each vehicle in the network has a unique *id*. Second, each vehicle can calculate its velocity and is able to know its current position and transmission range using GPS device. Third, the highway has two roads (one for each direction), and three lanes for each road. Fourth, several RSUs with a transmission range of 1.5 km are installed every 3 km on the side of the highway area, to cover the entire vehicular network. We assume that each RSU covers a sub-area of the highway, in order to perform cluster formation phase for each direction separately in a centralized way. Figure 3.1 shows the network structure for our proposed approach. If we assume that L is the length of the highway area, the approximate number of RSUs N necessary to cover the entire vehicular network is defined as follows:

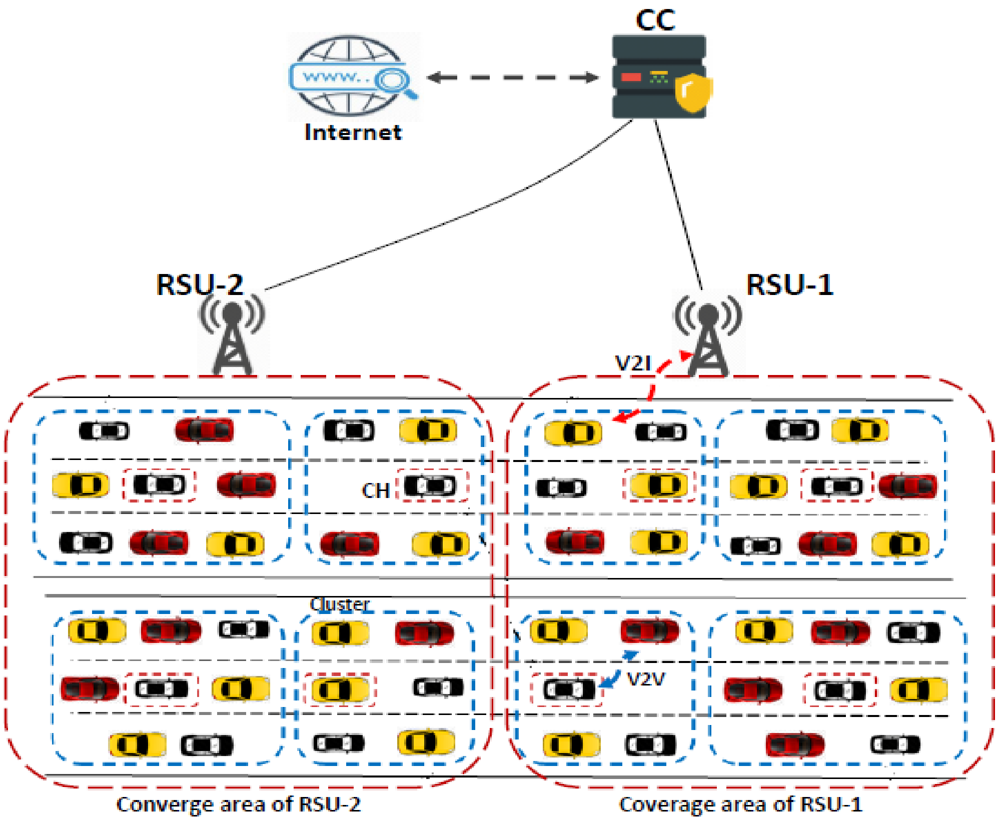


Figure 3.1: Network structure for our approach

$$N = \left\lceil \frac{L}{3} \right\rceil \tag{3.1}$$

The RSU models its coverage highway sub-area by an undirected graph $G = (V, E)$ for each sub-road (one for each direction), where V, is the set of vertex, representing the vehicles in the coverage area and E is the set of edges representing the set of the communication links between vehicles. There is an edge (link) $\{u, v\} \in E$ if and only if u and v are mutually in the coverage area of each other's (this implies that the link between these nodes are bidirectional). Then, we have the following basic concepts of graph theory that are used in our proposed approach:

- **Neighbor of node:** For each $v \in V$, the set of neighbors of v is defined by $N_v = \{u \in V \ / \ u \text{ is adjacent to } v\}$.

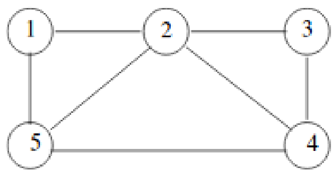
- **Degree of a node:** The degree of a node $v \in V$, denoted by d_v and is defined by the cardinality of its set of neighbors, which $d_v = |N_v|$.

- **Isolated node:** Node that has a zero degree.

- **Adjacency matrix:** Adjacency matrix is a mathematical tool used as a data structure in computer science to represent a graph [106].

  Either the graph $G = (V, E)$, all the vertex of graph are numbered from 1 to n with $n = |V|$. The adjacency matrix representation of G consists of a Boolean matrix A of size $n \times n$ such that $A[i][j] = 1$, if $edge(i, j) \in E$ and $A[i][j] = 0$ otherwise.

  Figure 3.2 shows a simple example of the adjacency matrix, where Figure 3.2a represents a non directed graph with five vertices and Figure 3.2b shows the corresponding adjacency matrix.



(a) Graph

	1	2	3	4	5
1	0	1	0	0	1
2	1	0	1	1	1
3	0	1	0	1	0
4	0	1	1	0	1
5	1	1	0	1	0

(b) Adjacency matrix

Figure 3.2: Example of adjacency matrix

3.2.2 Definitions and notations

This section presents definitions and notations of the important concepts used in the proposed clustering approach.

1. **Virtual backbone network:** It is a core network that interconnects all RSUs and allows them to share information about the vehicular network.

2. **On-Board Unit (OBU):** It is a terminal equipment mounted on board of a vehicle to provide a mutual wireless communication between vehicle and surrounding vehicles and infrastructures. It uses Wireless Access in Vehicular Environment (WAVE) standard, which is based on the emerging IEEE 802.11p communication technology [107].

3. **Road Side Units (RSUs):** These are fixed communication infrastructure units distributed on roadside for connecting vehicles to a larger infrastructure or to a

core network, such as a metropolitan traffic management system [108]. The RSUs have a wide transmission range of 1.5 km. It uses IEEE 802.11p communication technology for V2I communication and LTE for infrastructure to infrastructure (I2I) communication. These are the key elements in our approach and are responsible for performing the clustering process.

4. **Vehicle-to-Vehicle (V2V) communication:** It is the basic communication type in VANET. It allows a direct wireless transmission of data between vehicles and does not rely on fixed infrastructure. This type of communication is established if and only if the vehicles are in the coverage area of each other.

5. **Vehicle-to-Infrastructure (V2I) communication:** It takes place between vehicles and RSU fixed infrastructure through a wireless transmission. When a vehicle wants to send a message to RSU, it first determines whether the RSU is within its transmission range. If this is the case, the vehicle sends the message directly to the RSU through wireless communication. Otherwise, it checks whether it has a neighbor vehicle closer to the RSU. If it finds one, it sends the message to this neighbor vehicle so that the latter forwards the message to the RSU. Otherwise, this vehicle keeps the message (keeps carrying it) until it meets a neighbor vehicle closer to the RSU.

6. **Infrastructure-to-Vehicle (I2V) communication:** When an RSU wants to send message to vehicle, it first examines whether the vehicle is within its transmission range. If this is the case, the RSU sends the message directly to the vehicle. Otherwise, it looks for another RSU, which has the target vehicle in its coverage area via I2I communication. Then, the RSU sends the message to this RSU, which forwards the message to the target vehicle.

7. **Weight of node:** We introduce a metric: weight of a node, to be used later by our graph based clustering algorithm to elect the SCHs. For a node $v \in V$, we assign a weight ω_v that depends to three metrics: mobility indicator (node velocity), node degree and node transmission range. It is calculated as follows:

   If we suppose that a node v has been moved from position $P_1(x_1, y_1)$ at time T_1 to position $P_2(x_2, y_2)$ at time T_2, we denote Δd_v the distance traveled by the node v over time Δt ($\Delta t = T_2 - T_1$):

$$\Delta d_v = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{3.2}$$

Therefore, the velocity ψ_v of node v over time Δt, is calculated as follows:

$$\psi_v = \frac{\Delta d_v}{\Delta t} \tag{3.3}$$

Thus, the mobility indicator MI_v of node v, is calculated as:

$$MI_v = \left| \ln(\frac{\psi_v}{\psi_{max}}) \times 10^2 \right| \tag{3.4}$$

where ψ_{max} is the maximum velocity allowed on the road.

The MI depends on the velocity of vehicle. Typically, a vehicle with lower velocity has more stable communication links with its neighbors. Consequently, the vehicle is suitable to be CH. There should be created a negative correlational relationship between the MI and the velocity. Therefore, when the velocity decreases, the MI should be increased. Provided that $\psi_v > 0$ and $\psi_v <= \psi_{max}$, so $0 < \frac{\psi_v}{\psi_{max}} <= 1$. Therefore, we use the logarithm function, which is an increasing function on this interval with negative values. The absolute value is used to get a positive MI value. The MI is null when the velocity of vehicle reaches the maximum value allowed.

Therefore, the weight ω_v of node v is calculated based on the previous parameters as follows:

$$\omega_v = C_1 \times MI_v + C_2 \times d_v + C_3 \times Tr_v \tag{3.5}$$

where Tr_v is the transmission range of node v.
C_1, C_2 and C_3 are the weighting coefficients for the corresponding system parameters.

$$C_1 + C_2 + C_3 = 1 \tag{3.6}$$

The contribution of each weighting coefficient indicates the importance of the corresponding parameter relatively to the others. In this study, we consider that the most important factor is the one related to the mobility. For this reason, the MI is considered as the important parameter and therefore is assigned a higher weighting coefficient. The second important one is related to the connectivity, which represents the node degree. The last one is related to the transmission power, which represents the transmission range. Moreover, these weighting coefficients can be adjusted according to the system requirements. Therefore, we choose C1= 0.5, C2=0.3 and C3=0.2 as weighting coefficients values for the experiments in this work.

8. **Euclidean distance:** For each two vehicles v_i and v_j $(v_i, v_j \in V^2)$, Euclidean distance (Ed) between them is defined as follows:

$$Ed_{(v_i,v_j)} = \sqrt{(x_{v_i} - x_{v_j})^2 + (y_{v_i} - y_{v_j})^2} \tag{3.7}$$

   Where, $x_{v_i},\ y_{v_i},\ x_{v_j}\ and\ y_{v_j}$ represent the coordinates of the vehicles v_i and v_j, respectively.

9. **Communication link (edge):** There is a communication link between two vehicles v_i, v_j, noted as $l_{(v_i,v_j)}$, if and, only if the Euclidean distance between them is less than or equal to the shortest transmission range of them.

$$\exists\ l_{(v_i,v_j)} \Leftrightarrow Ed_{(v_i,v_j)} \leq min(Tr_{v_i}, Tr_{v_j}) \tag{3.8}$$

   Here, Tr_{v_i}, Tr_{v_j} are the transmission ranges of vehicle v_i, v_j, respectively.

10. **Notations:** Various notations used in the proposed scheme are given in Table 3.1.

11. **Vehicle state:** In the proposed clustering approach, vehicles can be in one of the following four status: Undefined Node (UN), Cluster Head (CH), Cluster Member (CM), and Secondary Cluster Head (SCH). Status are defined as follows:

    - **UN:** Initial state of a vehicle, it does not belong to any cluster and without role.
    - **CH:** Vehicle that has the task of coordination among cluster members and provides intra and inter cluster communication.

Table 3.1: Notations used in this study.

Symbols	Description
$\mathtt{ID}_v$	Identify of vehicle v
$\mathtt{RSU}_i$	Identify of RSU i
$\mathtt{Tr}_v$	Transmission range of vehicle v
$\mathtt{P}_v$	Position of vehicle v
$\mathtt{Ed}_{(v_i,v_j)}$	Euclidean distance between v_i, v_j
$l_{(v_i,v_j)}$	Communication link between v_i, v_j
$\mathtt{d}_v$	Degree of node v
ψ_v	Velocity of vehicle v
ω_v	Weight of node v
$\mathtt{Dir}_v$	Direction of vehicle v
$\mathtt{CH}_i$	Cluster Head of cluster i
$\mathtt{SCH}_i$	Secondary Cluster head of cluster i
$\mathtt{CM\_table}_i$	Member list of cluster i

- **SCH:** The vehicle that will replace the primary CH, in case it becomes unavailable or leaves the cluster.
- **CM:** A vehicle in a cluster but it is not a CH or a SCH.

12. **Message type:** Our clustering approach uses several types of messages. Table 3.2 presents the different types of messages with their descriptions.

Table 3.2: Message type in this study.

Message	Source	Destination	Description
BEACON	RSU	Vehicle	Notify vehicles
JOIN	Vehicle	RSU	Join the network
Reply_CH	RSU	Vehicle	Nominate a CH
Reply_SCH	RSU	Vehicle	Nominate a SCH
Reply_CM	RSU	Vehicle	Nominate a CM
update_C	Vehicle	Vehicle,RSU	Cluster update

3.3 Proposed approach

This section introduces a new Heuristic one-hop Clustering Algorithm based on RSU (HCAR) for IoV, which provides improved and reliable performance for classic VANET with better cluster stability and lower communication overhead. The proposed algorithm focuses only on Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. As it was mentioned previously, the proposed approach uses a hybrid method that combines the centralization of the heuristic one-hop clustering algorithm at distributed RSUs. This latter are responsible for performing the cluster formation phase for all vehicles within its coverage area according to a fully centralized way based on a heuristic algorithm using graph theory concepts. In addition, the current approach uses a new algorithm to recover the problem of the unavailability of primary CH, through the selection of Secondary CH (SCH).This latter is selected from the cluster members based on a weighted mechanism that combines three metrics: node velocity, node degree and node transmission range. The proposed approach also takes care of the maintenance phase using distributed methods to maintain the stability and structure of clusters. The rest of this section shows the main steps of the suggested approach including cluster formation and maintenance phases.

3.3.1 Cluster formation

The main steps of this phase are defined as follows:

- **Step 1:** Each RSU broadcasts periodic *BEACON* message containing its identity (RSU_i) and position.

- **Step 2:** Once a vehicle v enters the road, it is in UN state, and when entering into the coverage area of an RSU_i and receiving its broadcast, and after calculating its velocity ψ_v based on Eqs. (3.2) and (3.3), and getting its position (P_v), its direction (Dir_v) and its transmission range (Tr_v), it determines whether the RSU_i is within its communication range. If so, the vehicle v sends the JOIN request to this RSU_i directly. Otherwise, it checks out whether it has a neighbor vehicle closer to the concerned RSU_i. If it finds one, the vehicle v sends a JOIN request to this neighbor vehicle so that it forwards the JOIN request to the RSU_i. Otherwise, vehicle keeps the JOIN request (keeps carrying it) until it meets a neighbor vehicle closer to the RSU_i.

- **Step 3:** Each RSU, after receiving the *JOIN* requests from vehicles within its coverage area (inside of its transmission range), creates an adjacency matrix for each direction based on information received from vehicles. Algorithm 1 presents the procedure executed by the RSU for the creation of the adjacency matrix.

Algorithm 1: Adjacency Matrix Creation

Input: $(v_1 ... v_n)$;$(P_1 ... P_n)$;$(Tr_1 ... Tr_n)$
Output: A[n][n]: adjacency matrix
for $i \leftarrow 1$ ***to*** n **do**
  for $j \leftarrow 1$ ***to*** n **do**
    $Ed_{(v_i,v_j)} = \sqrt{(x_{v_i} - x_{v_j})^2 + (y_{v_i} - y_{v_j})^2}$
    if $(Ed_{(v_i,v_j)} \leq min(Tr_{v_i}, Tr_{v_j}))$ **then**
      $A[i][j] \leftarrow 1$
    else
      $A[i][j] \leftarrow 0$
return *A[n][n]*

- **Step 4:** Each RSU executes a heuristic algorithm (Algorithm 2) to perform clusters' formation phase.

  In our algorithm, we have two types of isolated nodes. First, the permanent isolated nodes are the nodes having a degree zero according to the initial adjacency matrix (initial graph). Second, the temporary isolated nodes are the nodes having a degree one, which becomes zero at the end of the current iteration when the adjacency matrix is updated.

  The algorithm starts by extracting the permanents isolated nodes from the adjacency matrix (nodes that have degree zero), and puts them in $isoNode$. Next, the algorithm selects the node that has the maximum degree according to the adjacency matrix and assigns as a new CH for the cluster that is under construction. If two or more nodes have the same maximum degree, the node with the lowest ID is chosen. Then, the algorithm adds all its neighbors to its members list (CM_table), and sets all the rows and columns in the adjacency matrix corresponding to its neighbors to zero. After these updates occur on the adjacency matrix and at the end of each iteration, if any node becomes isolated (temporary isolated nodes), it is added immediately to the cluster members table (CM_table) of CH. This process repeats iteratively until all the elements of the adjacency matrix become zero. At the end, the permanent isolated nodes are handled. For each permanent isolated node, the algorithm chooses the nearest CH for this isolated node to join its cluster. Algorithm 3, executed by RSU, describes the handling of the permanent isolated nodes in our approach. We denote the set of

Algorithm 2: Cluster Formation at RSU

Input: A[n][n]: Adjacency matrix of Graph G(V,E)
Output: $C_1, C_2, ..., C_m$

```
isoNode: Set of permanent isolated nodes
iC: index of cluster under creation
Extract permanents isolated nodes (isoNode)
iC ← 1
while sum(A[n][n]) != 0 do
    Select node that has max degree as CH_iC
    if many nodes has the same max degree then
        Select node that has lowest ID as CH_iC
    Add all neighbors of CH_iC to members table CM_table_iC
    foreach node v in CM_table_iC do
        A[id_CH][id_v] ← 0
        A[id_v][id_CH] ← 0
    foreach temporary isolated node v do
        Add v to members table CM_table_iC
    iC ← iC + 1
Permanent_isolated_node_handling(isoNode)
for i ← 1 to iC do
    SCH_i ← SCH_selection(C_i,CM_table_i)
    Remove SCH_i from CM_table_i
    Send reply_CH(C_i, CM_table_i,SCH_i) to CH_i
    Send reply_SCH(C_i, CM_table_i,CH_i) to SCH_i
    foreach node v in CM_table_i do
        Send reply_CM(C_i, CH_i) to v
```

isolated nodes by *isoNode*.

Figure 3.3 shows a simple example of clusters formation phase of our proposed approach. At first, the algorithm starts with the extraction of the permanent isolated nodes. In this example and according to the adjacency matrix, node 15 is a permanent isolated node. This latter is put in the set *isoNode* to handle it next. Then, node 16 (colored in green) is chosen as CH (because it has the maximum degree) and all its neighbors become members (colored in yellow). In parallel, the elements of the rows and columns corresponding to this CH and its members become zero in the adjacency matrix. At the end of this iteration, the first cluster C1 is formed: C1{CH:16; CM: 1,7,9,10,14}. In the second iteration, the same procedure is repeated, node 3 is chosen as CH2 and its neighbors as cluster members. At the end of this iteration, there is a new temporary isolated

Algorithm 3: Permanent Isolated Node Handling

Input: $C_1, C_2, ..., C_{iC}$; isoNode

Output: Final clusters state

idC: id of cluster to join

foreach *node v* ***in*** *isoNode* **do**

  Select the nearest CH to v

  $idC \leftarrow$ id of cluster of nearest CH

  Add v to members table CM_table_{idC}

node (node 17) that appears (colored in blue), this node is added immediately to the cluster table (CM_table) of this CH2. The state of the second cluster becomes as follows: C2 {CH:3; CM: 11,8,6,14,17}. Afterward, the node 5 is chosen as CH3 and its neighbors as cluster members C3 {CH:5; CM: 2,12,13}. The process ends as soon as all the elements of the adjacency matrix became zero. At the end, the permanent isolated node 15 (colored in red) is handled, the algorithm chooses CH3 because of its closeness to this isolated node. Therefore, the node 15 joins the cluster C3.

- **Step 5:** After clusters formation, the RSU must choose for each cluster a node that has the maximum weight (ω) among the cluster members of cluster as a SCH. This latter replaces the primary CH in case of its unavailability. The stage of selecting the SCHs is described in Algorithm 4.

Algorithm 4: SCH Selection

Input: C_i

Output: SCH_i

maxW: the maximum weight

$maxW \leftarrow 0$

foreach *node v* ***in*** CM_table_i **do**

  Calculate weight ω_v of v based on Eqs. (3.4) and (3.5)

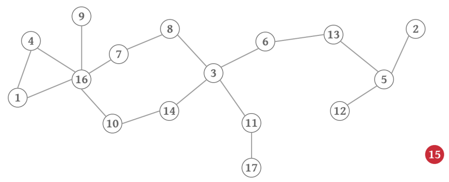
  if $\omega_v > maxW$ **then**

    $maxW \leftarrow \omega_v$

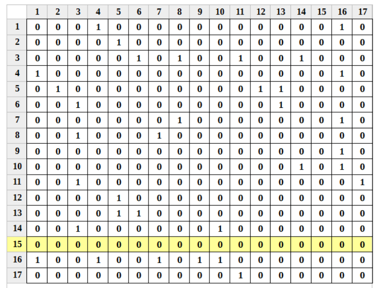
    $SCH_i \leftarrow ID_v$

return SCH_i

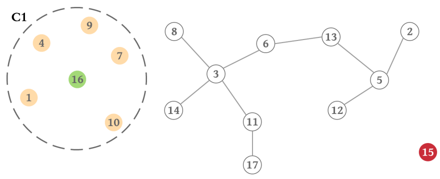
- **Step 6:** Then, the RSU informs all the vehicles of their role in the network, by sending a reply: reply_CH, reply_SCH and reply_CM to CH, SCH and CMs respectively for each cluster. The process of the cluster formation phase at RSU is illustrated in Figure 3.4.



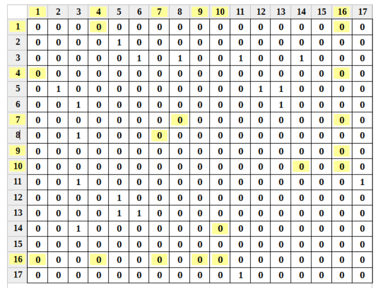
(a) Initial network



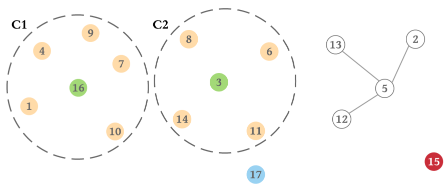
(b) Initial associated adjacency matrix



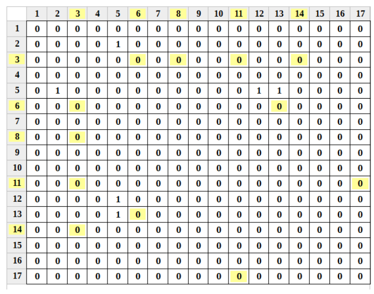
(c) First cluster formation



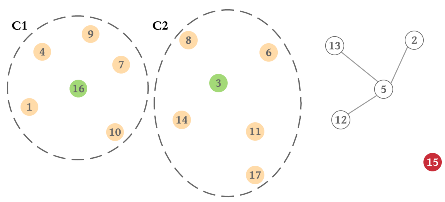
(d) Associated adjacency matrix



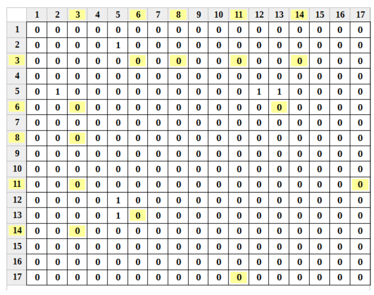
(e) Second cluster formation



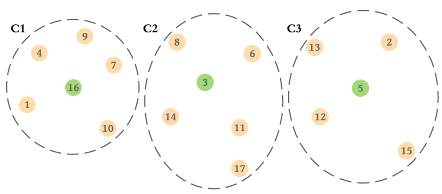
(f) Associated adjacency matrix



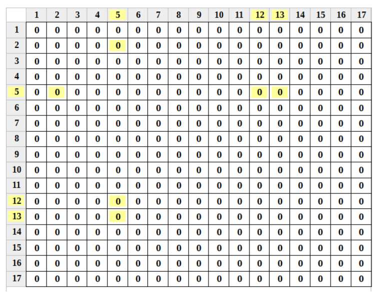
(g) Temporary isolated node handling



(h) Associated adjacency matrix



(i) Permanent isolated node handling and final clusters formation



(j) Final associated adjacency matrix

Figure 3.3: Example showing clusters formation phase

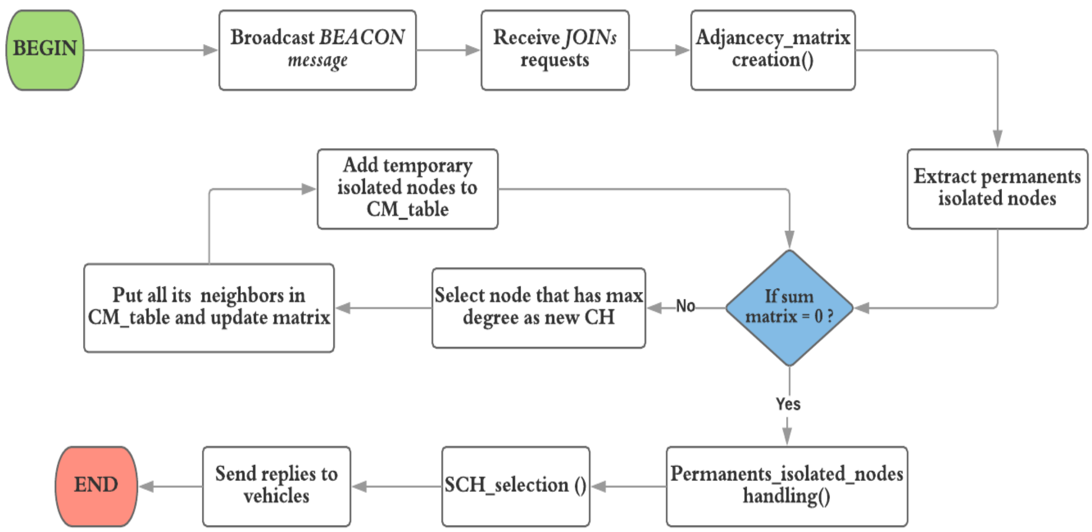


Figure 3.4: Clusters formation phase function at the RSU

- **Step 7:** After receiving the reply from the RSU, each vehicle updates its state to CH, SCH and CM respectively with reply_CH, reply_SCH and reply_CM and starts the transmission phase. Algorithm 5 describes the execution of this step by the vehicles. The process of the cluster formation phase at vehicle can be given by the flowchart as illustrated in Figure 3.5.

Algorithm 5: Cluster Formation at Vehicle

Receive(*reply*) from RSU i
switch *reply* **do**
    case *reply_CH* **do**
        state $\leftarrow$ CH
        Update (C,SCH,CM_table)
    case *reply_SCH* **do**
        state $\leftarrow$ SCH
        Update (C,CH,CM_table)
    case *reply_CM* **do**
        state $\leftarrow$ CM
        Update (C,CH)

3.3.2 Cluster maintenance

Because of the special characteristics of VANETs, such as high mobility and frequent change of topology, vehicles keep joining and leaving clusters frequently. Therefore, an

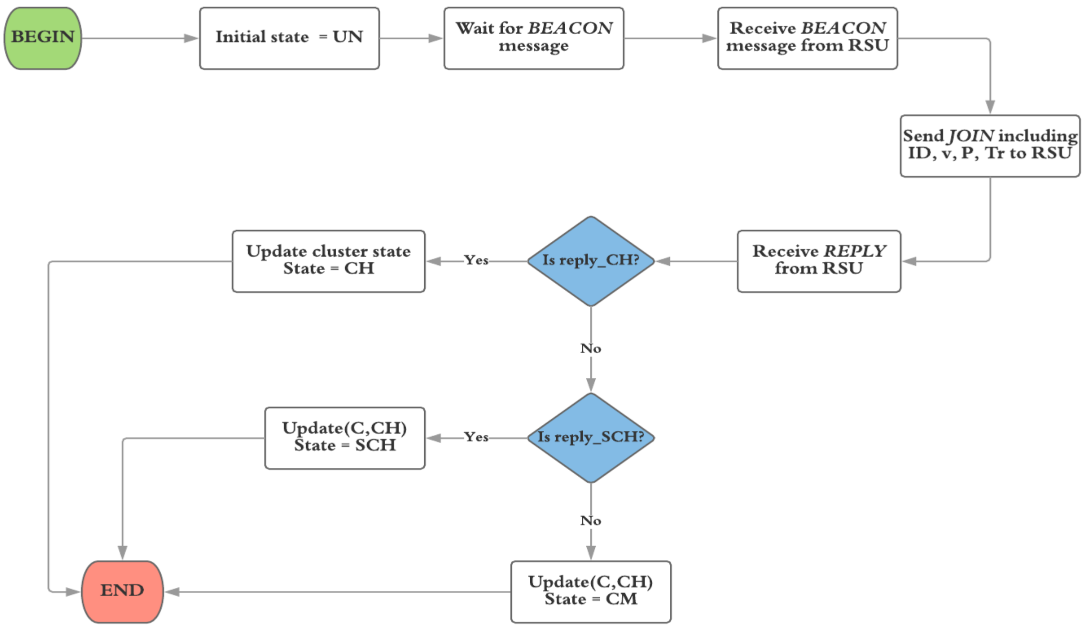


Figure 3.5: Clusters formation phase function at the vehicle

extra maintenance overhead is present. In our proposed approach, cluster maintenance phase is described next.

3.3.2.1 CH or SCH leaving discovery

When the clusters' formation phase is finished, a private communication link is established between the CH and SCH in the cluster. Both of them exchange periodic messages stating their state in the cluster. The CH vehicle periodically detects the state of the SCH vehicle and vice versa. If a SCH does not receive a periodic message from its CH over a time period. It means that the CH has left the cluster because it is outside the transmission range of SCH. This latter must replace it and serve as a new CH of cluster, and change its state to CH. Then, it selects a new SCH for the cluster from cluster members based on the weights of this cluster member that has already been calculated and sent by the RSU during the formation phase. The new CH (old SCH) must inform all vehicles within its cluster to make the necessary updates by sending an *update_C* message. The CM that becomes the new SCH changes its status to SCH. The remaining CMs update their CH and the state of the old CH changes to UN. At the end, the new CH informs the nearest RSU of the update of the state of the cluster in order to share it via backbone network with the other RSUs. The pseudo code of CH leaving discovery at SCH is shown in Algorithm 6. On the other hand, if CH does not receive a periodic message from its SCH over a time period, it means that the SCH has left the cluster. Therefore, the CH selects a new SCH for the cluster among cluster members based on the weights of members. Then, the CH sends an *update_C* message to the node which becomes a new SCH. This latter changes its status to SCH. The state of the old SCH changes to UN. At the end of this phase, the CH informs the nearest RSU by this update in order to share it via backbone network with the other RSUs.

Algorithm 6: CH Leaving Discovery at SCH

```
Detect leaving of CH_i
state ← CH
Select a new SCH_i among members
Remove SCH_i from CM_table_i
Send update_C(C_i, CM_table_i,CH_i) to SCH_i
foreach node v in CM_table_i do
    Send update_C(C_i, CH_i) to v
```

3.3.2.2 Leaving a Cluster

On the highway, vehicles may join and leave the clusters several times. When the cluster formation phase terminates, each CH triggers a monitoring process of its members *(CM_table)* to keeps track of their presence in the cluster. In this respect, each cluster member periodically sends a beacon message to its CH according to its proper slot time. On the other hand, the CHs use an intra-cluster gathering process to collect the beacon messages from their cluster members to monitor their presence in the cluster. Therefore, when a CM moves out of the cluster range, the CH detects this event and deletes this CM from its members table, then it sends an update_C message to its SCH. Furthermore, if a CM does not receive the periodic message from its CH over a time period, the state of that node changes to UN and this node can then join another cluster.

The process of leaving a cluster is illustrated in Figure 3.6.

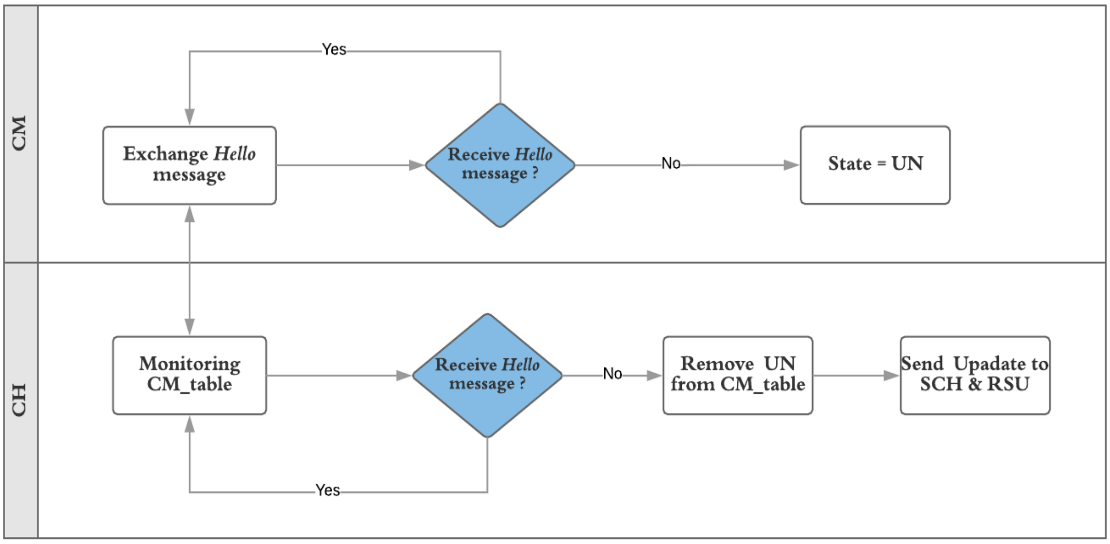


Figure 3.6: Cluster leaving function

3.3.2.3 Joining Cluster

When the clusters formation phase is finished, several UN nodes seek to join the network. These nodes are either new nodes or nodes leaving other clusters. When a UN vehicle approaches a CH (comes inside its transmission range), sends a beacon message containing its velocity to this CH. Then, this latter calculates the relative velocity to this UN vehicle, and if the velocity difference is within +/- ΔV_{th}, the UN vehicle will join the cluster and subsequently, the CH adds it to its members list, otherwise it ignores it. On the other hand, if the UN node receives the reply from the CH, it must change its state to CM and join the cluster. The different steps to join a cluster are shown in Figure 3.7. Algorithm 7 describes the process followed by a CH during the joining a cluster phase. .

Algorithm 7: Joining cluster discovery at CH

```
Receive periodic message from UN
Calculate relative velocity (rv) to UN
if rv +/- ΔV_th then
    Add UN to CM_table of CH
    Send update_C(C, CM_table,CH) to SCH
    Send reply to UN
else
    Ignore UN
```

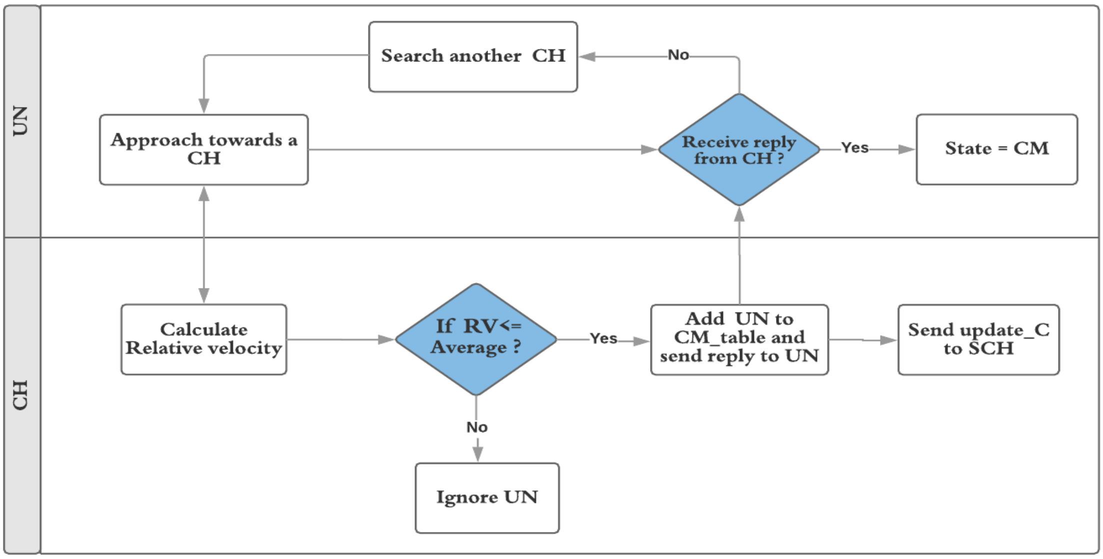


Figure 3.7: Cluster joining function

3.3.3 Theoretical analysis

In this section, we present the theoretical analysis in order to prove the rationality and performances of the proposed clustering algorithm.

3.3.3.1 Clustering Overhead Analysis

The major disadvantage introduced by the majority of clustering protocols is the additional overhead messages including control messages in order to maintain the cluster structure. The clustering overhead, OV_c can be classified as follows:

(1) Overhead due to broadcast of BEACON message, OV_b.

(2) Overhead due to cluster formation phase, OV_{cf}.

(3) Overhead due to cluster maintenance phase, OV_{cm}.

Therefore, the total aggregation overhead is the sum of the contributing factors above.

$$OV_c = OV_b + OV_{cf} + OV_{cm} \tag{3.9}$$

To simplify the analysis, the following parameters are used.

- ***n:*** It is the number of vehicles in the network.
- ***N:*** It is the number of RSUs installed on the roadside (see Eq. 3.1).

- ***b_i*:** It is the number of vehicles which received the BEACON message broad-casted by the RSU i.
- ***c*:** It is the number of CHs elected.
- ***m_i*:** It is the number of members in the cluster i.

The BEACON messages are broad-casted by the RSUs to attract vehicles to join the network, the BEACON overhead in the worst case is:

$$OV_b = O(\sum_{i=1}^{N} b_i) \tag{3.10a}$$

Knowing that $(\sum_{i=1}^{N} b_i) \leq n$, so:

$$OV_b = O(n) \tag{3.10b}$$

The cluster formation overhead depends on the RSU and vehicles at the same time. After receiving the BEACON message, every vehicle in the RSU area sends a JOIN request to the RSU and waits for a reply. After receiving all JOINs requests, the RSU in turn, sends replies to vehicles in its area. Thus, the message complexity in the worst case of the formation phase is:

$$OV_{cf} = O(2.\sum_{i=1}^{N} b_i) \tag{3.11a}$$

Knowing that $(\sum_{i=1}^{N} b_i) \leq n$, so:

$$OV_{cf} = O(n) \tag{3.11b}$$

The cluster maintenance overhead depends on the events that change and affect the cluster structure, that are: CH (or SCH) leaves a cluster, UN joins cluster, or CM leaves a cluster. The CH periodically detects the state of the SCH and vice versa. If the SCH (or CH) detects the unavailability of CH (or SCH), the SCH must take the role of CH and inform the members of the cluster by broadcasting an update messages. Therefore, the complexity in the worst case for this event is $O(\sum_{i=1}^{c} m_i)$. In our proposal, the CH has a function of monitoring its members (CMs) to keep track of their presence. When a CM moves out of the cluster range, the CH detects this event and deletes this CM from the members table, then sends an update message to its SCH. The complexity corresponding to this event is $O(\sum_{i=1}^{c} m_i)$. When a UN vehicle approaches a CH (enters its transmission range), the UN sends its relative velocity to CH which checks whether the relative velocity of this node is less than or equal to the average relative velocity of the cluster. If it is the case, the CH adds the UN to its members table and sends a reply message to this UN. The complexity corresponding to this event is $O(c)$. The maintenance phase overhead is as follows:

$$OV_{cm} = O(2.\sum_{i=1}^{c} m_i) + O(c) \tag{3.12a}$$

Knowing that $(\sum_{i=1}^{c} m_i) \leq n$ and $c \leq n$, so:

$$OV_{cm} = O(n) \tag{3.12b}$$

The total overhead OV_c for our proposed scheme is therefore:

$$OV_c = O(3n) \tag{3.13a}$$

$$OV_c = O(n) \tag{3.13b}$$

3.3.3.2 Clustering Properties:

In order for our proposed solution to work properly, some clustering properties must be satisfied.

Definition 1 ***Safety property****: Each cluster has one and only one CH. Every node can belong to only one cluster.*

The safety property has to ensure that the vehicles in the network are grouped into clusters and each cluster has only one CH. In addition, this propriety ensures that each node belongs to a single cluster to avoid overlap between clusters.

Lemma 1 *The safety property is satisfied.*

Proof 1 *Based on Algorithm 2, node v_i is a new CH, if it satisfies following conditions:*

1. *It has the maximum degree.*

2. *It has the lowest ID, if two or more nodes have the same maximum degree :*

$$d_i = d_j = .. = d_k \implies id_i = min(id_i, id_j, .., id_k) \tag{3.14}$$

This implies a unique choice of CH for each cluster and the first part of the property is verified.

In addition, each node v_i can belong to a single cluster with two possible cases:
Case 1: *if it has a direct link (one-hop) to a CH :*

$$\exists\ CH_j:\ A[j][i] = 1.$$

Consequently, it joins the cluster of CH_j, and the elements of its row and its column are set to zero in the adjacency matrix. Therefore, this node cannot join another cluster in the following iterations.
Case 2: *if it has indirect link (two-hops) to a CH (temporary isolated node) :*

$$\exists\ CH_j,\ \exists v_k : (A[j][k] = 1) \wedge (A[i][k] = 1).$$

And the elements of its row and its column become zero in the current iteration.

This implies that each node can belong to only one cluster and the second part of the property is verified.

Lemma 2 *Cluster formation algorithm terminates.*

Proof 2 *Since every node can determine its cluster (lemma 6), the sum of elements of the adjacency matrix corresponding to the virtual graph will eventually become zero.*

$$\sum_{i,j=1}^{n} A[i][j] = 0 \tag{3.15}$$

Thus, the algorithm will terminate.

Definition 2 ***Dominance property****: In a cluster, every ordinary node has a cluster head as its neighbor with two hops at most.*

The dominance property ensures that each ordinary node has direct access to its cluster head, thus allows fast inter and intra-cluster communication.

Lemma 3 *The dominance property is satisfied.*

Proof 3 *Each ordinary node can reach its CH. v_i is an ordinary node. Two cases are possible.*

Case 1: *if it has a direct link (one-hop) to a CH:*

$$\exists\ CH_j\text{: } A[j][i] = 1.$$

Thus, the node v_i can reach CH_j in one hop.

Case 2: *if it has indirect link (two-hops) to a CH (temporary isolated node) :*

$$\exists\ CH_j,\ \exists v_k : (A[j][k] = 1) \wedge (A[i][k] = 1).$$

Thus, the node v_i can reach CH_j in two hops via node v_k.

This implies that each node can reach its CH in two hops at most, the property is verified.

Definition 3 ***Independence property****: No two cluster-heads can be neighbors.*

The independence property ensures an efficient clustering method with a few clusters.

Lemma 4 *The independence property is satisfied.*

Proof 4 *N_i is the set of neighbors of node i. We assume that we have two neighboring cluster heads, CH_i and CH_j of clusters C_i and C_j respectively. The following condition is satisfied:*

$$(CH_i \in N_j) \wedge (CH_j \in N_i)$$

That implies

$$(CH_i \in\ CM\_table_j) \wedge (CH_j \in\ CM\_table_i)$$
$$\text{and}$$
$$(CH_i \in (C_i, C_j)) \wedge (CH_j \in (C_j, C_i))$$

Since every node can belong to only one cluster (lemma 6), the conditions above present a contradiction.

Definition 4 ***Liveness property***: *At time t, each node is in one of these states UN, CH, SCH or CM.*

The liveness property ensures that each node is in a stable state at a given time.

Lemma 5 *The liveness property is satisfied.*

Proof 5 *According to the transition state (Figure 3.8), every transition from one state to another is due to an event. The possible transitions are as follows:*

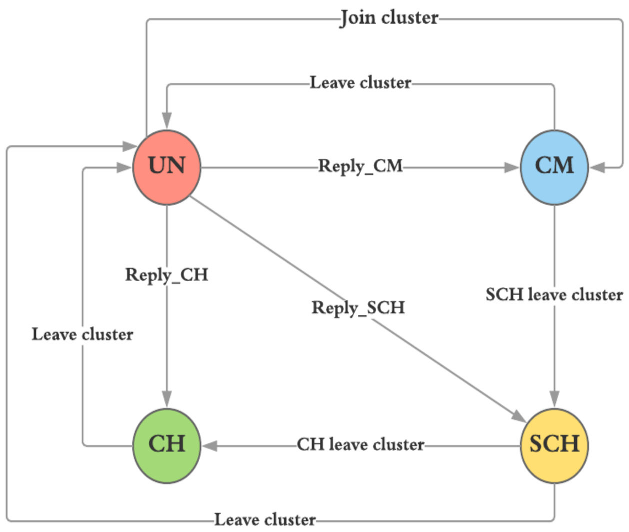


Figure 3.8: Transition state

1. ***From UN state to CM state***: *this transition is triggered by two events: the node receives a reply_CM from RSU or receives a reply from CH to join a cluster during the maintenance phase.*

2. ***UN state to CH state***: *this transition is triggered when a node receives a reply_CH from RSU.*

3. ***UN state to SCH state***: *this transition is triggered by the reception of a reply_SCH from RSU.*

4. ***Transition from CH state to UN state***: *is triggered when a node in CH state leaves a cluster.*

5. ***CM state to UN state***: *this transition is initiated when a node in CM state leaves a cluster.*

6. ***The transition from CM state to SCH state****: is activated by the following event: the CH assigns a new SCH for a cluster because the current SCH left the cluster.*

7. ***SCH state to CH state****: this transition is triggered when the CH leaves the cluster.*

8. ***SCH state to UN state****: this transition is activated when a node in SCH state leaves the cluster.*

3.4 Simulation

3.4.1 Experimental analysis

This section is devoted to simulation in order to prove the effectiveness and performances of the proposed approach using the network simulator NS-2 (version 2.34) [109] and VanetMobiSim [110] integrated environment on a machine with intel $i5$ (4th generation) processor and 8 GB of RAM. The simulation is based on one directional highway with three lanes. There are three RSU installed at the road segment. Physical and MAC layers are configured according to 802.11p standard with channel throughput of 11 Mbps. The velocity varies uniformly between 10 m/s and 35 m/s. The transmission range is also varied from 100 m to 300 m. Many scenarios were used in which important parameters have been modified, such as the velocity and the transmission range. The vehicles were assigned to random positions and they moved according to the mobility model, named the Intelligent Driver Model including Lane Change (IDM-LC) [111, 112], which is integrated in VanetMobiSim. The prorogation model used is Two-ray Ground. The different simulation parameters are listed in Table 3.3.

We compare the results of our proposed scheme to two well-known protocols for VANET belonging to the same family of clustering named MOSIC [92] and N-hop [96].

3.4.2 Comparison metrics

The analysis of the stability performance of the proposed algorithm is based on five metrics: CH lifetime, CM lifetime, CH change number, number of clusters and control overhead.

Table 3.3: Simulation parameters.

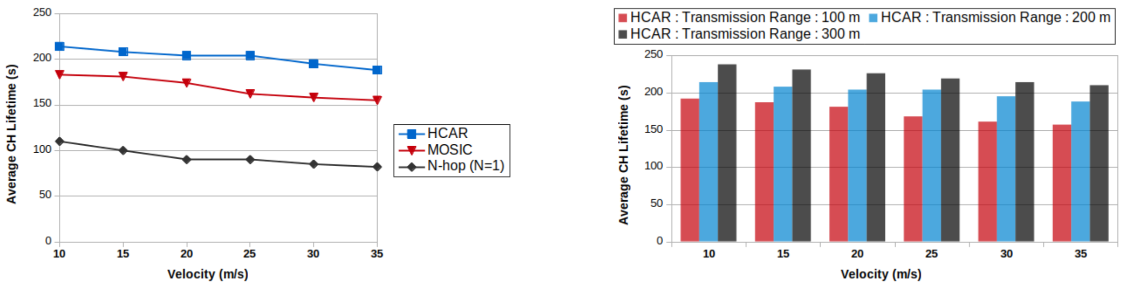
Parameter	Value
Number of vehicles	80 – 100
Simulation time	360 s
Simulation area	5000 × 100 m
Number of lanes	3
Road Side Unit	2
Transmission range	100 – 300 m
Mobility model	IDM-LC
Propagation model	Two-ray Ground
MAC/PHY standard	IEEE 802.11p
MAC data transfer rate	11 Mbps
Velocity of vehicles	10 – 35 m/s
Maximum allowed velocity (ψ_{max})	40 m/s
C1, C2 and C3 values	0.5, 0.3 and 0.2

3.4.2.1 CH Lifetime

CH lifetime is an important parameter to show the efficiency of the technique used for CHs' election and clusters' stability. Average CH lifetime represents the time during which the vehicle is in a CH state. Figure 3.9a shows the average CH lifetime of the proposed approach versus MOSIC and N-hop protocols under different velocities value. Figure 3.9b presents the average CH lifetime of the proposed scheme under different transmission ranges.

According to Figure 3.9a, for the lowest simulated speed (10 m/s), the average CH lifetime of HCAR is equal to 214 s, while the average CH lifetime of both MOSIC and N-hop is, respectively, equal to 183 s and 110 s. For the highest simulated speed (35 m/s), the average CH lifetime of HCAR is equal to 188 s, while the average CH lifetime of both MOSIC and N-hop is, respectively, equal to 155 s and 82 s. We can make three observations. Firstly, as the velocity gets higher, the decrease in the average CH lifetime of the three protocols is very moderate because the network topology is very dynamic due to the high mobility of the vehicles. Secondly, the average CH lifetime of HCAR is longer than those of MOSIC and N-hop. Thirdly, the average CH lifetime achieved by our proposed scheme in the case of the highest speed (188 s) is greater than those achieved by MOSIC (183 s) and N-hop (110 s) for the lowest speed. This shows that our approach is efficient even for high speed.

As shown in Figure 3.9b, for the lowest simulated speed (10 m/s), the average CH lifetime, when HCAR is used for 100 m, 200 m and 300 m transmission ranges, is 192 s, 214 s and 238 s, respectively. For the highest simulated speed (35 m/s), the average CH lifetime, when HCAR is used for 100 m, 200 m and 300 m transmission ranges is 157 s, 188 s and 210 s, respectively. We can make the following two observations: Firstly, the average CH lifetime, when HCAR is used, increases as the transmission range increases. This is because increasing the transmission range leads to increasing the radius of the cluster's coverage area. Secondly, as the speed increases, the average CH lifetime decreases moderately for the three transmission ranges (100 m, 200 m and 300 m). This is due to the efficient clustering process used in our proposed scheme.



(a) Average CH lifetime versus MOSIC & N-hop

(b) CH lifetime under different transmission ranges

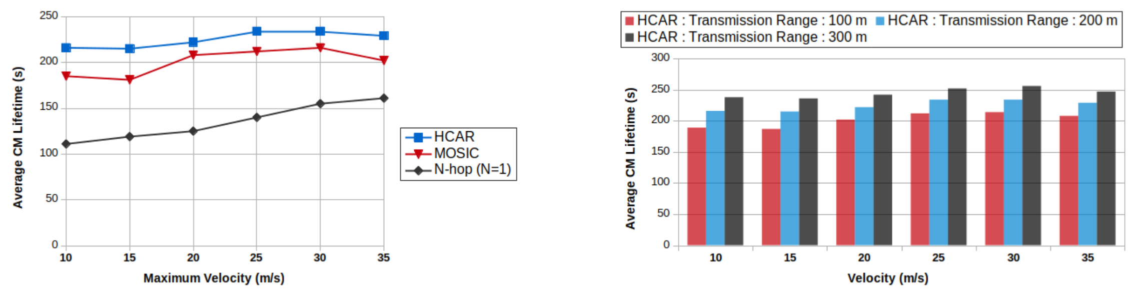
Figure 3.9: Average CH lifetime

3.4.2.2 CM Lifetime

CM lifetime is a metric used to show the performance of the clustering algorithm and demonstrate the stability of the constructed clusters. Average CM lifetime represents the time during which the vehicle is in a CM state. Figure 3.10a shows the average CM lifetime of the proposed scheme versus MOSIC and N-hop protocols under different velocity values. Figure 3.10b depicts the average CM lifetime of the proposed algorithm under different transmission ranges.

As illustrated in Figure 3.10a, for the lowest simulated speed (10 m/s), the average CM lifetime of HCAR is equal to 219 s, while the average CM lifetime of both MOSIC and N-hop is, respectively, equal to 186 s and 113 s. For the highest simulated speed (35 m/s), the average CM lifetime of HCAR is equal to 225 s, while the average CM lifetime of both MOSIC and N-hop is, respectively, equal to 202 s and 161 s. We can make three observations. Firstly, as the velocity gets higher, the irregular change in the average CM lifetime of the three mechanisms, that is due to the high mobility of vehicles, which makes the network topology very dynamic. Secondly, the average CM lifetime of HCAR is longer than those of MOSIC and N-hop. Thirdly, the average CM lifetime achieved by our proposed scheme in the case of the highest speed (225 s) is greater than those achieved by MOSIC (186 s) and N-hop (113 s) for the lowest speed. This shows that our approach is efficient even for high speed.

As shown in Figure 3.10b, for the lowest simulated speed (10 m/s), the average CM lifetime, when HCAR is used for 100 m, 200 m and 300 m transmission ranges, is 189 s, 219 s and 238 s, respectively. For the highest simulated speed (35 m/s), the average CM lifetime, when HCAR is used for 100 m, 200 m and 300 m transmission ranges is 208 s, 225 s and 247 s, respectively. We can make the following two observations: Firstly, as the speed increases, the average CM lifetime decreases moderately for the three transmission ranges (100 m, 200 m and 300 m). This is due to the efficient maintenance phase of our proposed scheme. Secondly, the average CM lifetime, when HCAR is used, increases as the transmission range increases. This is because increasing the transmission range leads to increasing the radius of the cluster's coverage area.



(a) Average CM lifetime versus MOSIC & N-hop

(b) CM lifetime under different transmission ranges

Figure 3.10: Average CM lifetime

3.4.2.3 CH Change Number

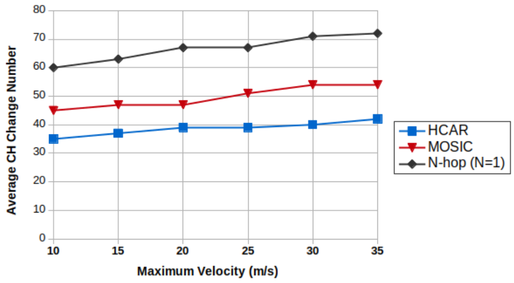
CH change number is a metric used to demonstrate the performance of the clustering algorithm and the stability of the constructed clusters. Average CH change number represents the number of vehicles whose state changes from CH state to UN state during a simulation period. Typically, a low CH change number leads to a stable cluster structure. Figure 3.11a shows the average CH change number of the proposed scheme versus MOSIC and N-hop protocols under different velocity values. Figure 3.11b depicts the average CH change number of the proposed algorithm under different transmission ranges.

According to Figure 3.11a, for the lowest simulated speed (10 m/s), the average CH change number of HCAR is equal to 35, while the average CH change number of both MOSIC and N-hop is, respectively, equal to 45 and 60. For the highest simulated speed (35 m/s), the average CH change number of HCAR is equal to 42, while the average CH change number of both MOSIC and N-hop is, respectively, equal to 54 and 72. We can make three observations. Firstly, as the velocity gets higher, the increase in the average CH change number of all three protocols is very moderate because the network topology is very dynamic due to the high mobility of the vehicles. Secondly, the average CH change number of HCAR is less than those of MOSIC and N-hop. Thirdly, the average CH change number achieved by our proposed scheme in the case of the highest speed (42) is less than those achieve by MOSIC (45) and N-hop (60) for the lowest speed. This shows that our approach is efficient even for high speed.

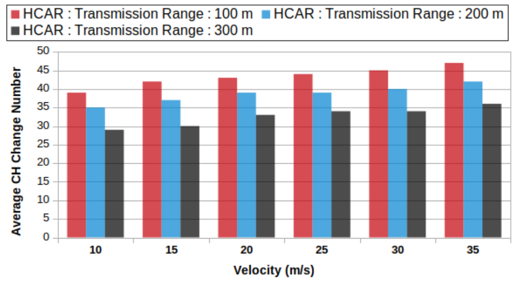
As shown in Figure 3.11b, for the lowest simulated speed (10 m/s), the average CH change number, when HCAR is used for 100 m, 200 m and 300 m transmission ranges, is 39, 35 and 29, respectively. For the highest simulated speed (35 m/s), the average CH change number, when HCAR is used for 100 m, 200 m and 300 m transmission ranges is 47, 42 and 36, respectively. We can make the following two observations: Firstly, the average CH change number, when HCAR is used, decreases as the transmission range increases. This is because increasing the transmission range leads to increasing the radius of the cluster's coverage area. Secondly, as the speed increases, the average CH change number increases moderately for the three transmission ranges (100 m, 200 m and 300 m). This is due to the efficient clustering process of our proposed scheme.

3.4.2.4 Number of Clusters

The average number of clusters is a metric used for the evaluation of the quality of clustering scheme. Figure 3.12a depicts the average number of clusters of the proposed approach versus MOSIC and N-hop protocols under different velocity values. Figure 3.12b presents the average number of clusters of the proposed scheme under different transmission ranges.



(a) Average CH change number versus MOSIC & N-hop

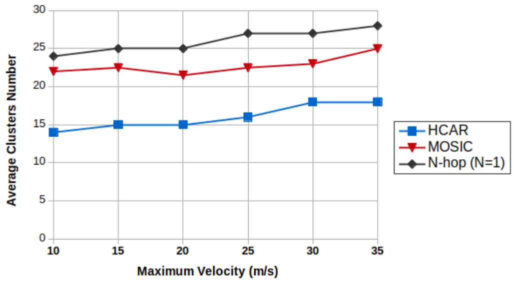


(b) CH change number under different transmission ranges

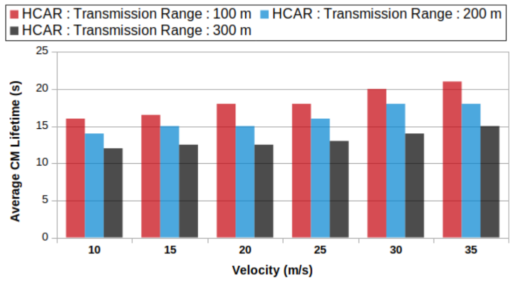
Figure 3.11: Average CH change number

According to Figure 3.12a, for the lowest simulated speed (10 m/s), the average number of clusters of HCAR is equal to 14, while the average number of clusters of both MOSIC and N-hop is, respectively, equal to 22 and 24. For the highest simulated speed (35 m/s), the average number of clusters of HCAR is equal to 18, while the average number of clusters of both MOSIC and N-hop is, respectively, equal to 25 and 28. We can make two observations. Firstly, as the velocity gets higher, the increase in the average number of clusters of the three protocols is very moderate because the network topology is very dynamic due to the high mobility of the vehicles. Secondly, the average number of clusters of HCAR is less than those of MOSIC and N-hop.

As shown in Figure 3.12b, for the lowest simulated speed (10 m/s), the average number of clusters, when HCAR is used for 100 m, 200 m and 300 m transmission ranges, is 16, 14 and 12, respectively. For the highest simulated speed (35 m/s), the average number of clusters, when HCAR is used for 100 m, 200 m and 300 m transmission ranges is 21, 18 and 15, respectively. We can conclude the following two observations. Firstly, as the speed increases, the average number of clusters increases moderately for the three transmission ranges (100 m, 200 m and 300 m). This is due to the efficient clustering process of our proposed scheme. Secondly, the average number of clusters, when HCAR is used, decreases as the transmission range increases. This is because increasing the transmission range leads to increasing the radius of the cluster's coverage area.



(a) Clusters number versus MOSIC & N-hop



(b) Clusters number under different transmission ranges

Figure 3.12: Average clusters number

3.4.2.5 Control Overhead

The major issue introduced by the majority of proposed clustering schemes is the additional overhead messages including control messages in order to maintain the cluster structure. Therefore, the control overhead is the number of control messages received by each node during the clustering process. Figure 3.13 shows the average control message overhead of our proposed algorithm, MOSIC, and N-hop for different velocity values.

According to Figure 3.13, for the lowest simulated speed (10 m/s), the average control message overhead of HCAR is equal to 11.5 kbps, while the average control message overhead of both MOSIC and N-hop is, respectively, equal to 13.4 kbps and 17 kbps. For the highest simulated speed (35 m/s), the average control message overhead of HCAR is equal to 14,7 kbps, while the average control message overhead of both MOSIC and N-hop is, respectively, equal to 15 kbps and 20 kbps. We can make two observations. Firstly, as the velocity gets higher, the increase in the average control message overhead of the three protocols is very moderate because the network topology is very dynamic due to the high mobility of the vehicles. Secondly, the average control message overhead of HCAR is less than those of MOSIC and N-hop.

3.5 Conclusion

In this chapter, a new Heuristic one-hop Clustering Algorithm based on RSU (HCAR) for IoV is presented. Clustering in HCAR is performed in a centralized way at the RSUs using a simple heuristic algorithm based on graph theory concepts, such as neighboring,

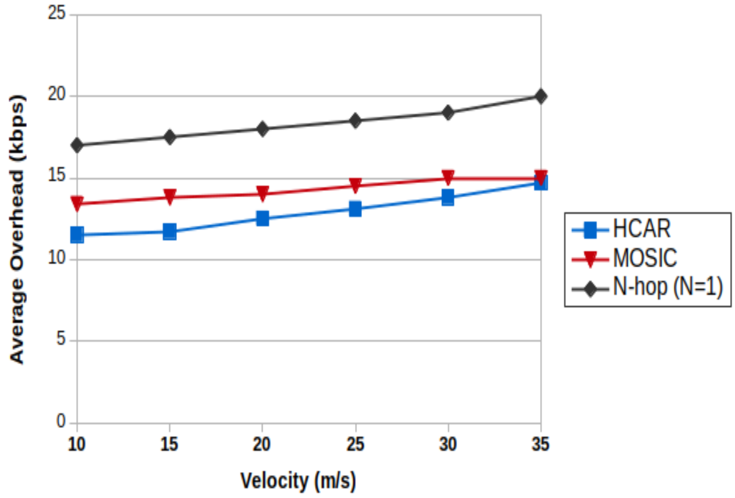


Figure 3.13: Average overhead

adjacency matrix and node degree. Furthermore, HCAR uses a new recovery method which is based on the election of the Secondary CH (SCH), using a weighted method that combines three metrics: vehicle velocity, vehicle degree and vehicle transmission range. This method helps to detect the unavailability of primary CH and replace it quickly to increase cluster stability and avoid the re-clusterings of network. Moreover, HCAR maintains the clusters using a distributed mechanism to ensure clusters' stability and react to the frequent changes in network topology. HCAR is evaluated using simulation by comparing it to other protocols under different scenarios. Simulations are conducted using the network simulator NS2 and the VanetMobiSim integrated environment. Results clearly show the efficiency of HCAR as compared to MOSIC and N-hop protocols. HCAR increases the average CH lifetime and the average CM lifetime, improves the overview of the network using a central node (RSU) to perform the cluster formation phase, decreases the average CH change number and clusters number, and reduces significantly the number of control messages. As future work, we aim to investigate the use of HCAR in urban areas by designing an effective protocol and comparing its performance to existing clustering protocols.

The next chapter presents the second contribution of our thesis. In this respect, we propose a new multi-hop clustering approach over Vehicle-to-Internet communication for improving conventional VANETs performances.

Chapter 4

MCA-V2I: A Multi-hop Clustering Approach over Vehicle-to-Internet communication for improving VANETs performances

4.1 Introduction

4.1.1 VANET toward IoV: An overview

The Internet of Vehicles (IoV) is an evolution of conventional VANET. It extends VANET's scale, structure and applications. This evolution leads to the emergence of new interactions at the road level among vehicles, humans and infrastructure [43]. It is an important field of research to improve conventional VANETs and their performances. Researchers have proposed several protocols for different aims and applications, such as data dissemination and aggregation, network overhead minimization, road safety, traffic management and mainly routing schemes.

Compared with VANET, IoV has many specific advantages and characteristics, such as developing and extending the exploitation of the Intelligent Transportation System (ITS) in different fields of research and industry [3]. The first main advantage is the quick and easy access to the Internet. This allows sharing safety information between vehicles and providing useful information, such as the availability of hotels, parking's location, gas stations and even drivers' comfort applications. The second advantage is the ability to support a significant number of connected vehicles (scalability). As a third advantage, Cloud Computing (CC) technology can be integrated into the vehicular networks. This emergent technology allows applications, resources and data to be stored in remote stations and servers that represent the cloud, so that they can be used by clients with low capacity. The CC technology manages the large amount of data generated by the connected vehicles. Finally, the IoV expands basic types of VANET communications such as Vehicle-to-Vehicle (V2V) and Vehicle-to-Road Side Unit (V2R) to new types of communications, such as Vehicle-to-Internet (V2I), Vehicle-to-Person and Vehicle-to-Device.

4.1.2 Motivation

According to the available literature, most of the proposed clustering algorithms [92, 94, 113–115] are focused only on one-hop clustering, which only allows communication between a Cluster Member (CM) and its CH with one-hop distance at most. Consequently, the coverage area is very small, and many clusters are formed, which affects the network performance and increases the rate of overlapping between clusters. Moreover, because the VANET is a subclass of MANETs, several proposed protocols are derived from the MANET clustering schemes [53, 71]. However, these schemes do not consider the mobility characteristics, the dynamic topology and the limited driving directions of VANET; moreover, they do not consider energy problems [116]. Furthermore, most of the proposed clustering protocols do not use mechanisms that exploit the Internet and to take advantage of their large services to improve the performances of VANET. Several proposed mobility-based clustering approaches [52, 79, 80, 117] are based on the broadcast of control messages, which causes overloading of the networks and leads to many collisions, especially because the number of messages is high due to the multi-metric mechanism used.

4.1.3 Contribution

In this chapter, we propose a new Multi-hop Clustering Approach over Vehicle-to-Internet communication called MCA-V2I to improve VANETs' performance. The main idea of this work is to perform a clustering algorithm using Internet access. MCA-V2I is based on the reasonable assumption that a vehicle can connect to the Internet via a special infrastructure called a Road Side Unit Gateway (RSU-G) to obtain and share the necessary information about its multi-hop neighbors to perform the clustering algorithm. It is performed using a Breadth-first search (BFS) algorithm for traversing the graph and based on a Mobility Rate (MR) that is calculated according to some mobility metrics such as node connectivity, average relative velocity, average distance and link stability. In MCA-V2I, a vehicle with low MR is suitable to be elected as the Master CH (MCH). Therefore, all the multi-hop neighbors of the new elected MCH become Cluster Members (CMs). The MCA-V2I scheme strengthens clusters' stability through the election of a Slave Cluster Head (SCH). We evaluate the performances of MCA-V2I using network simulator NS-2 and the VanetMobiSim integrated environment.

The main contributions of this work are as follows:

1. A new multi-hop clustering model is proposed. Compared with one-hop clustering schemes, this model is designed to extend the coverage area of clusters, reduce the number of clusters, optimize the control overhead and improve cluster stability.

2. A Mobility Rate is introduced for the clustering algorithm. This parameter is calculated based on mobility metrics to satisfy the requirements of the new features of VANET, and to consider its mobility characteristics.

3. MCA-V2I provides Internet access to vehicles to obtain and share the necessary information to perform the clustering algorithm. This benefit significantly reduces the rate of control messages used in traditional clustering algorithms. Therefore, MCA-V2I can significantly improve the network overhead.

4. MCA-V2I strengthens clusters' stability through the election of an SCH in addition to the MCH.

The rest of this chapter is organized as follows. Section 4.2 describes the preliminaries of the proposed approach. Section 4.3 introduces the proposed approach in details. Section 4.4 presents the experimental results. Finally, conclusion is presented in Section 4.5.

4.2 Preliminaries

The following section describes the preliminaries of the proposed approach presenting the network model and system description.

4.2.1 Network model

The proposed approach is based on the following assumptions. First, each vehicle in the network has a unique *id*, which is the MAC address of the OBU interface. Second, every vehicle is equipped with an OBU device using WAVE technology. Third, we have a highway with two roads (one for each direction), and three lanes for each road. Finally, several RSUs with a transmission range of 1.5 km are installed every 3 km on the sides of the highway area, to cover the entire vehicular network. If we assume that L is the length of the highway area, the approximate number N_{RSU} of RSUs necessary to cover the entire vehicular network is defined as follows:

$$N_{RSU} = \left\lceil \frac{L}{3} \right\rceil \tag{4.1}$$

The vehicular network topology is modeled as an undirected graph $G(V, E)$, where V is the set of vertices representing the vehicles in the network, and E is the set of edges representing the communication links between vehicles. There is a link $(i, j) \in E$, if and only if vehicles i and j are mutually in the coverage area of each other:

$$\exists\ (i, j) \in E \implies distance(i, j) \leq min(Tr_i, Tr_j) \tag{4.2}$$

where Tr_i and Tr_j are the transmission ranges of vehicles i and j, respectively.

Then, we have the following basic concepts of graph theory that will be used in our proposed scheme.

- **Node's neighbors:** It is the set of one-hop neighbors of node i, N_i, where

$$N_i = \{j \in V \mid \exists\ (i, j) \in E\} \tag{4.3}$$

- **Node degree:** It is the cardinality of one-hop neighbors set N_i of node i, where

$$Deg_i = |N_i| \tag{4.4}$$

- **Multi-hop neighbors of node:** It is the set of all nodes within multi-hops from node i, denoted by MN_i.
- **Multi-hop degree of node:** It is the cardinality of the multi-hop neighbors set of node i, denoted by MD_i, where:

$$MD_i = |MN_i| \tag{4.5}$$

- **Graph traversal:** Graph traversal means visiting every node (vertex) exactly once in a well-defined order from a given node v ($v \in V$) [118]. According to the order in which the nodes are visited, there are two main algorithms of traversals: Depth-First Search (DFS) and BFS [119]. In the proposed approach, we are interested in the BFS algorithm to perform the clustering process.

  The implementation of a simple BFS algorithm starting from a given source node s is shown in Algorithm 8. The purpose of the implementation is to visit every node exactly once. For this reason, the implementation uses a queue to mark nodes already visited. The algorithm works as follows:

  1. Start by adding the (given) node s to the queue and mark it as visited.
  2. Remove the head of the queue.
  3. Add single-hop neighbors of the removed node, that are not already visited to the queue and mark them as visited.
  4. Keep repeating steps 2 and 3 until the queue becomes empty.

4.2.2 System description

4.2.2.1 Network architecture

Our proposed scheme's architecture is illustrated in Figure 4.1. It is mainly composed of vehicles, OBUs, RSUs-G, TTC, CC and Internet. The definition of the different components and communication types between them are as follows.

Algorithm 8: BFS algorithm

Input: graph $G(V, E)$, start node s

$Q\{\}$: BFS queue

$Q \leftarrow \{s\}$ ▷ Initially, Q contains s

Mark s as visited

while ! *empty* Q **do**

  Remove the head u of Q

  foreach *neighbor* v *of* u **do**

    if v *is unvisited* **then**

      Add v to the back of Q

      Mark v as visited

    else

      ignore v

1. **Vehicle:** It is the mobile node and the main component for our network architecture. Each vehicle is equipped with a GPS devise.

2. **On-Board Unit (OBU):** It is a terminal equipment mounted on board a vehicle to provide a mutual wireless communication between the vehicle and surrounding vehicles and infrastructures. It uses the Wireless Access in Vehicular Environment (WAVE) standard, which is based on the emerging IEEE 802.11p specification [107].

3. **Road Side Units Gateway (RSUs-G):** These are fixed communication infrastructure units distributed on the roadside. They are controlled and managed by the Transportation Control Center (TCC) through wired communication channels. They use use IEEE 802.11p communication technology for V2I communication. Compared with conventional RSUs, the RSUs-G have an extension in terms of features. First, they provide a registration feature for vehicles to join the network. Second, they contain an integrated DHCP server to ensure automatic IP address configuration for the vehicles. They act as gateways for the vehicles to allow them to access the backbone network and to exploit the services provided by the Internet and Cloud Center (CC). Finally, they are responsible for aggregation, updating and distributing real-time traffic information to the vehicles.

4. **Transportation Control Center (TCC):** It is responsible for network initial-

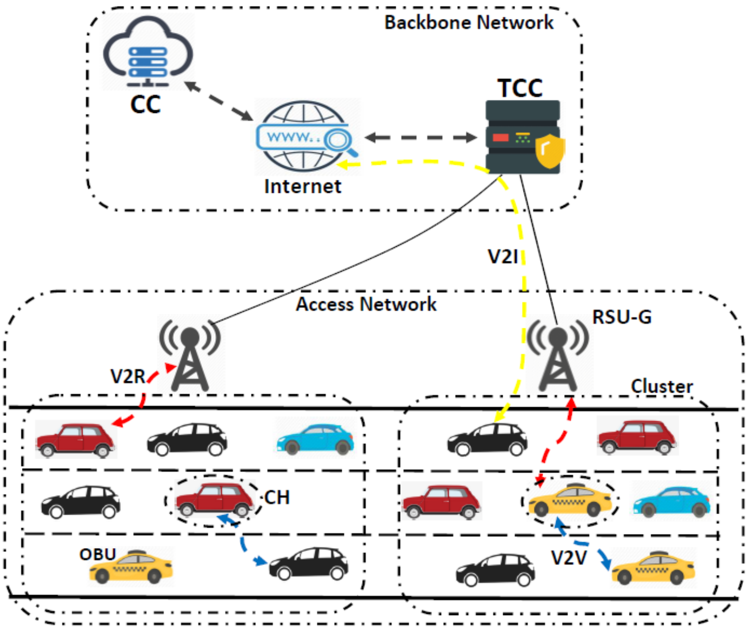


Figure 4.1: Network archetecture for the proposed approach

ization, interconnecting RSUs-G and exchanging data between them. It represents the interface between the Access Network (AN) and the Internet network.

5. **Cloud Center (CC):** It is a virtual server that is based on a cloud computing platform over the Internet. It has features similar to a standard server. CC contains cloud servers to store and share data, resources and applications to serve vehicles on demand to perform the clustering algorithm.

6. **Vehicle-to-Vehicle (V2V) communication:** It is the basic communication type in VANET. It allows the direct wireless transmission of data between vehicles and does not rely on fixed infrastructure. This type of communication is established if and only if the vehicles are mutually in the coverage area of each other. The choice of efficient relay selection process in V2V communications is considered as one of the significant challenges in vehicular network. In this regard, numerous relay selection mechanisms have been proposed in the literature [120, 121]. Consequently, proper selection of relay selection process can provide a high delivery ratio, acceptable overall end-to-end communication delays and efficient bandwidth usage.

7. **Vehicle-to-RSU-G (V2R) communication:** It takes place between vehicles and RSU-G fixed infrastructure through wireless transmission. It is the first step

for vehicles to access the Internet. In the proposed approach, when a vehicle wants to send a message to an RSU-G, it first determines whether the RSU-G is within its transmission range. If this is the case, the vehicle sends the message directly to the RSU-G through the wireless communication. Otherwise, it uses a greedy forwarding mechanism and checks whether it has a neighbor vehicle closer to the RSU-G. If it finds one, the vehicle sends the message to this neighbor vehicle so that the latter forwards the message to the RSU-G. Otherwise, the vehicle keeps the message (keeps carrying it) until it meets a neighbor vehicle closer to the RSU-G.

8. **RSU-G-to-Vehicle (R2V) communication:** It takes place between RSU-G and vehicles. When an RSU-G wants to send a message to a vehicle, it first examines whether the vehicle is within its transmission range. If this is the case, the RSU-G sends the message directly to the vehicle. Otherwise, it looks for another destination RSU-G which contains the target vehicle in its coverage area via the backbone network. Then, the RSU-G sends the message to this destination RSU-G so that the latter forwards the message to the target vehicle.

9. **Vehicle-to-Internet (V2I) communication:** It is a virtual communication type that allows the vehicle to access the Internet via RSU-G and TCC.

4.2.2.2 Definition and notation

In this section, definitions and notations used in the proposed clustering approach are introduced.

1. **Multi-hop Clustering Record (MCR):** In our proposed approach, each node has a record called MCR that contains a set of information needed for the clustering process. It is composed of three fields: node identifier (*id*), node mobility rate (MR) and set of single-hop neighbors of this node (SN). Figure 4.2 shows the structure of MCR with a simple example.

id: 7	MR: 1.51	SN: {4, 9, 11, 18, 21}

Figure 4.2: MCR structure

2. **Mobility Rate (MR):** It is a parameter introduced by our approach to be used during the clustering process. It is based on a combination of the mobility metrics described below.

   - **Node Connectivity (NC):** It depends on the degree Deg_i of node i, where

   $$NC_i = Deg_i \tag{4.6}$$

   - **Average Relative Velocity (ARV):** A lower ARV of the node relative to its neighbors indicates that the node has a more stable state. Let us assume that $P_1(x_1, y_1)$ is the position of node i at time T_1 and $P_2(x_2, y_2)$ is the position of node i at time T_2. Δd_i is the distance traveled by node i over time Δt ($\Delta t = T_2 - T_1$).

   $$\Delta d_i = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{4.7}$$

   Thus, the velocity v_i of node i over time Δt is computed as:

   $$v_i = \frac{\Delta d_i}{\Delta t} \tag{4.8}$$

   Finally, the average relative velocity ARV_i of node i is computed as:

   $$ARV_i = \frac{1}{NC_i} \sum_{j=1, j \neq i}^{NC_i} |v_i - v_j| \tag{4.9}$$

   - **Average Distance (AD):** It is the average distance between a node and its neighbors. A node that has a minimum AD is closer to the center of its neighborhood. The AD_i of node i is computed as the cumulative mean square distance to its neighbors divided by its NC_i as follows:

   $$AD_i = \frac{1}{NC_i} \sum_{j=1, j \neq i}^{NC_i} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4.10}$$

   - **Link Stability (LS):** It represents the link stability of the node relative to its neighbors. It depends on the AD variation rate. Let us assume that $AD_i(t1)$ is the average distance of node i at time $t1$ and $AD_i(t2)$ is the average distance of node i at time $t2$. The link stability $LS_i(T)$ of node i over a time $T(T = t2 - t1)$ is calculated as follows:

$$LS_i(T) = |AD_i(t1) - AD_i(t2)| \tag{4.11}$$

Therefore, the mobility rate MR_i of node i is calculated based on the previous parameter as follows:

$$MR_i = \frac{LS_i(T)}{NC_i} + \left|(\ln(1 - \frac{ARV_i}{v_{max}}))\right| + \frac{AD_i}{maxD_i} \tag{4.12}$$

where $maxD_i$ is the maximum distance between node i and its neighbors. v_{max} is the maximum velocity allowed on the road.

3. **Notations:** Various notations used in the proposed approach are given in Table 4.1.

Table 4.1: Notations used in this study.

Symbols	Description
$\mathtt{id}_i$	Identity of vehicle i
$\mathtt{RSU\text{-}G}_i$	Identity of RSU-G i
$\mathtt{Tr}_i$	Transmission range of vehicle i
$\mathtt{P}_i$	Position of vehicle i
$\mathtt{Ed}_{(i,j)}$	Euclidean distance between vehicles i and j
$\mathtt{Deg}_i$	Degree of node i
v_i	Velocity of vehicle i
$\mathtt{C}_i$	Cluster i
$\mathtt{MCH}_i$	Master Cluster Head of cluster i
$\mathtt{SCH}_i$	Slave Cluster Head of cluster i
$\mathtt{CM\_list}_i$	Member list of cluster i

4. **Message types:** Our clustering scheme uses several types of messages. Table 4.2 describes the different types of messages and their descriptions.

5. **Vehicle States:** In the proposed clustering scheme, vehicles can be in one of the following states: Undefined Node (UN), Master Cluster Head (MCH), Slave Cluster Head (SCH) and Cluster Member (CM). Statuses are defined as follows:

   - **UN:** Initial state of a vehicle, it does not belong to any cluster.
   - **MCH:** Vehicle that has the task of coordination among cluster members.

Table 4.2: Message type in this study.

Message	Source	Destination	Description
HELLO	RSU-G	Vehicles	Notify vehicles
REGISTER	Vehicle	RSU-G	Vehicle registration
BEACON	Vehicle	Neighbors	Exchange information
SHARE	Vehicle	RSU-G	Share the MCR
ANNOUNCE	MCH	RSU-G	Announce new MCH
REPLY	Vehicle	MCH	Confirm membership
NOMINATION	MCH	Vehicle	SCH nomination

- **SCH:** The vehicle that will replace the MCH, in case it becomes unavailable or leaves the cluster.
- **CM:** A vehicle in a cluster but it is not an MCH or an SCH.

4.3 Proposed approach

In this section, we introduce a new Multi-hop Clustering Approach over Vehicle-to-Internet communication called MCA-V2I for improving VANETs' performance. The main idea of this work is to execute a clustering algorithm using Internet access. MCA-V2I is based on the reasonable assumption that a vehicle can connect to the Internet via a special infrastructure called a Road Side Unit Gateway (RSU-G) to obtain and share the necessary information about its multi-hop neighbors to perform the clustering process. This latter is performed using a BFS algorithm for traversing a graph and based on a Mobility Rate (MR), which is calculated using mobility metrics such as node connectivity, average relative velocity, average distance and link stability. In MCA-V2I, a vehicle with low MR is suitable to be elected as Master CH (MCH). The MCA-V2I scheme also strengthens clusters' stability through the election of a Slave Cluster Head (SCH). The MCA-V2I approach is composed of six phases: registration, neighborhood discovery, MCH selection, announcement, affiliation and maintenance. The rest of this section describes these phases in detail.

4.3.1 Registration

Initially, when a vehicle enters the road and decides to join the network, its OBU system is turned on. On the other hand, each RSU-G is required to broadcast a HELLO message which includes its location and identity information to the vehicles. When a vehicle comes into the coverage area of an RSU-G and receives the broadcast message, it sends a REGISTER request to register with the backbone network (Internet) and the RSU-G. When an RSU-G receives the REGISTER request, it forwards the registration request to the TCC to confirm the registration of this vehicle and provide it an IP address by sending a CONFIRM message. When a vehicle receives the confirmation, it changes its state to UN and starts the clustering algorithm. Figure 4.3 summarizes the steps of the registration phase in a sequence diagram format.

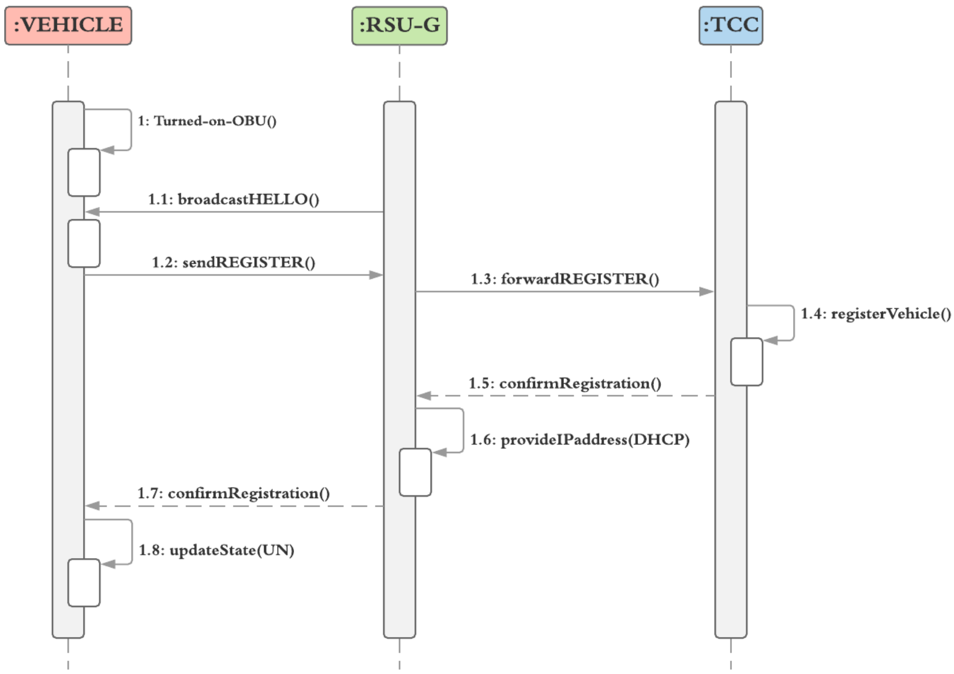


Figure 4.3: Sequence diagram of registration phase

4.3.2 Neighborhood discovery

To announce its existence, each vehicle broadcasts a periodic BEACON message to its single-hop neighbors, including its MAC address (*id*), its velocity, its transmission range and its position (two-dimensional coordinates). After receiving the BEACON message from all its single-hop neighbors, each vehicle calculates the following parameters: Node Connectivity (NC), Average Relative Velocity (ARV), Average Distance (AD) and Link Stability (LS). The values of these parameters are used to compute its Mobility Rate (MR). Then, each vehicle sends a SHARE message containing its MCR to the RSU-G to share its MCR in the backbone network (CC) with all vehicles in the network. Figure 4.4 shows the steps of the neighborhood discovery phase in a sequence diagram format. The neighborhood discovery process including the MCR construction is described in Algorithm 9.

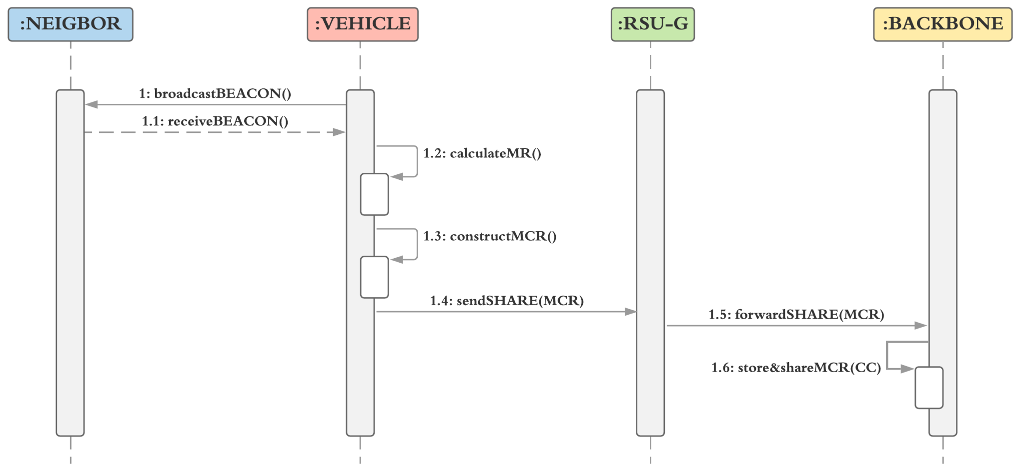


Figure 4.4: Sequence diagram of neighborhood discovery phase

4.3.3 Master CH selection

To elect the MCHs, each vehicle tries to establish a connection to the Internet via an RSU-G, to traverse its multi-hop neighbors using the BFS algorithm based on its MCR and the MCRs (of other vehicles) shared in the backbone network. During the traversal, each vehicle saves all visited vehicles (Multi-hop neighbors (MN)) and compares its MR with their MRs. If its MR has the lowest value, the vehicle must update its state to MCH and add all the vehicles crossed before in its CM_*list*. Otherwise, the vehicle elects the vehicle that has the lowest MR value as its new MCH and updates its variable myMCH (the variable that indicates the *id* of the MCH). Then, it moves to the affiliation phase. If there are two or more vehicles that have the lowest MR, the vehicle that has the lowest *id* will be elected as MCH. The MCH selection process using BFS traversal is described in Algorithm 10.

Figure 4.5 illustrates a simple example of the MCH election phase using the BFS algorithm with source vehicle 5 colored in red and its multi-hop neighbors colored in black. Table 4.3 presents the different parameters of the vehicles. First, the queue (Q) contains the source vehicle {5}. Then, the algorithm removes the head of Q {5} and adds its neighbors {1,2,8} one by one to the back of Q. For each added node, the

Algorithm 9: Neighborhood Discovery

Struct multiHopRecord
{
id : identifier
MR : Mobility Rate
SN{} : Single-hop Neighbors set
}
MCR_i: multiHopRecord
$\mathrm{N}_i\{\}$: Neighbors set
Broadcast BEACON(i, v_i, Tr_i, P_i) message
Receive BEACON(id, v, Tr, P) messages from neighbors
foreach *received BEACON(j) message* **do**
    if *distance(i,j)* $<$ *min(*Tr_i, Tr_j*)* **then**
        $\mathrm{N}_i \leftarrow \mathrm{N}_i \cup \mathrm{j}$
$Deg_i \leftarrow |N_i|$
Calculate MR_i based on equations 4.6 to 4.12
$\mathrm{MCR}_i[\mathrm{id}] \leftarrow i$
$\mathrm{MCR}_i[\mathrm{MR}] \leftarrow \mathrm{MR}_i$
$\mathrm{MCR}_i[\mathrm{SN}] \leftarrow \mathrm{N}_i$
Send SHARE(MCR) to RSU-G

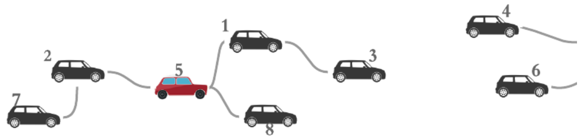
algorithm marks it as visited (colored in red) and checks if this node has an MR less than the MR of the source node. This process repeats iteratively until all the multi-hop neighbors are visited. At the end of this example, the source vehicle 5 has the lowest MR compared with its multi-hop neighbors. Consequently, vehicle 5 becomes the new MCH and all its multi-hop neighbors become CMs. The state of this cluster becomes as follows: Cluster {MCH: 5; CMs: 1,2,8,3,7}.

4.3.4 Announcement

Each vehicle, having determined itself as the new MCH, must announce its election. For this reason, each MCH must try to establish a connection to the Internet and send an ANNOUNCE message to the RSU-G including its *id*, its CM_list and its cluster *id* to share its cluster state in the backbone network (CC).

4.3.5 Affiliation

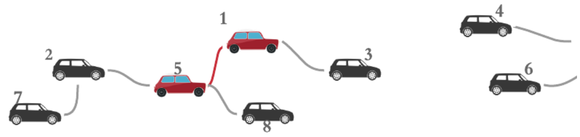
Each ordinary vehicle (not MCH) accesses the backbone network via the RSU-G, to find its affiliation and the cluster to which it belongs. When an ordinary vehicle finds



(a) Initial network

Queue (Q)	minMR	id_minMR
{5}	1.46	5

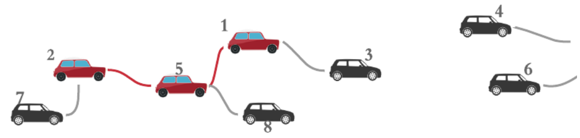
(b) Initial values



(c) First iteration

Queue (Q)	minMR	id_minMR
{1}	1.46	5

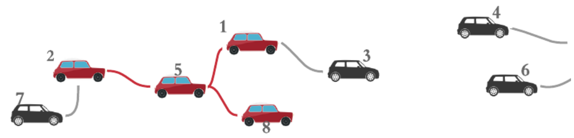
(d) Corresponding values



(e) Second iteration

Queue (Q)	minMR	id_minMR
{1,2}	1.46	5

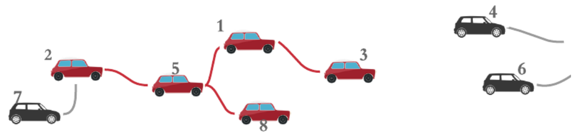
(f) Corresponding values



(g) Third iteration

Queue (Q)	minMR	id_minMR
{1,2,8}	1.46	5

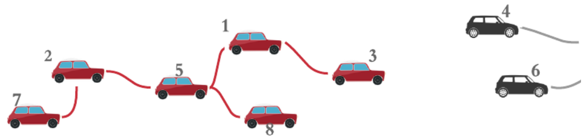
(h) Corresponding values



(i) Fourth iteration

Queue (Q)	minMR	id_minMR
{2,8,3}	1.46	5

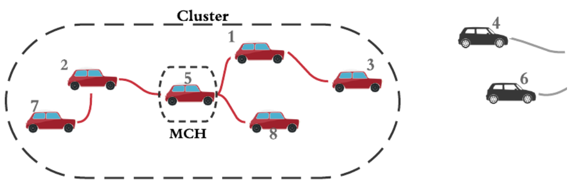
(j) Corresponding values



(k) Fifth iteration

Queue (Q)	minMR	id_minMR
{7}	1.46	5

(l) Corresponding values



Queue (Q)	minMR	id_minMR
{}	1.46	5

(m) Final cluster state

(n) Final corresponding values

Figure 4.5: Example showing MCH selection phase using BFS algorithm

Algorithm 10: MCH Election

```
Input: MCR_i
Q{}: BFS queue
Q ← {i}                                          ▷ Initially, Q contains i
MN_i{}: Multi-hop Neighbors set
minMR: minimum MR
id_minMR: id of node that has minMR
minMR ← MCR_i[MR]
id_minMR ← i
while ! empty Q do
    Remove the head j of Q
    foreach node k in MCR_j[SN] do
        if k is unvisited then
            Add k to the back of Q
            MN_i ← MN_i ∪ k
            if minMR > MCR_k[MR] then
                minMR ← MCR_k[MR]
                id_minMR ← k
            Mark k as visited
        else
            ignore k
if id_minMR = i then
    state_i ← MCH                                ▷ Update state to MCH
    CM_list_i ← MN_i
else
    myMCH ← id_minMR                             ▷ Update its MCH
```

the cluster to which it belongs, it compares the MCH *id* of this cluster with its myMCH variable. If they are the same, the vehicle sends a REPLY packet to this MCH, updates its state to CM and its cluster *id*. Otherwise, the vehicle ignores this event and moves to the maintenance phase to join the appropriate cluster. On the other hand, each MCH, after receiving all the REPLY packets, updates its CM_list. At the end, each MCH must select a vehicle with the lowest MR value among the cluster members (except itself) as Slave CH (SCH). Then, the MCH sends a NOMINATION message to the designated vehicle. The vehicle that receives the NOMINATION message updates its state to SCH. Figure 4.6 illustrates the steps of the MCH election, announcement and affiliation phases in a sequence diagram format.

Table 4.3: Vehicles parameters.

MCR[id]	MCR[MR]	MCR[SN]
5	1.46	{1,2,8}
1	1.69	{5,3}
2	1.74	{5,7}
8	1.93	{5}
3	1.81	{1}
7	2.07	{2}

4.3.6 Maintenance

The aim of this phase is to maintain the cluster structure and stability as long as possible. Because of the high mobility of vehicles, the cluster structure and network topology change frequently. For this reason, several events are triggered at the cluster level. The different events with their maintenance are described as follows.

4.3.6.1 MCH leaving discovery

In each cluster, the SCH vehicle periodically monitors the state of the MCH vehicle using a private communication link. If an SCH does not receive a periodic message from its MCH over a time period T, it means that the MCH has left the cluster. The SCH must replace the MCH immediately and takes over as the the new MCH of cluster. Therefore, it must change its state to MCH. Then, it elects a new SCH among the cluster's members based on their MRs. Furthermore, it broadcasts an update message to its CMs to inform them to update their MCH (myMCH variable). Finally, the new MCH must inform the CM that it has been elected as the new SCH to change its state to SCH. The old MCH must change its state to UN and join another cluster.

4.3.6.2 Clusters merging

The proposed approach can react with clusters overlapping. However, when two neighbor clusters have a big overlapping rate over a period Tm (Time merging), a cluster merging process is invoked. Typically, the merging procedure results in two MCHs at the same time for the final cluster obtained. Therefore, only one MCH is selected to manage all of the CMs of the merged clusters. Thus, the MCH that has the largest number of cluster members (cardinality of the CM_list) is elected as the new MCH for the cluster obtained and its SCH becomes also the SCH for the cluster obtained. The

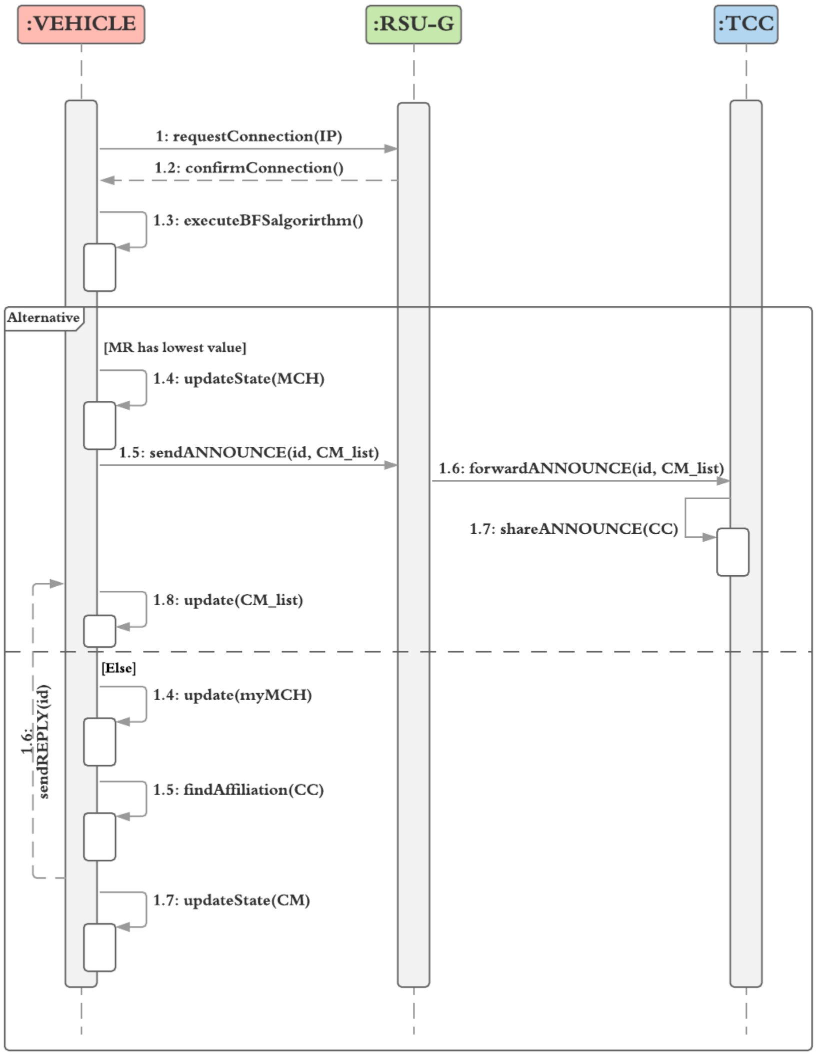


Figure 4.6: Sequence diagram of MCH election, announcement and affiliation phases

other MCH and SCH must change their state to CM.

4.3.6.3 Leave a cluster

Each MCH monitors its CMs through the exchange of periodic messages to keep track of members in the cluster. When a member moves out of the cluster range over a time period T, the MCH detects this event and immediately removes this node from its members' list (CM_list). Then, the MCH sends a message to its SCH indicating this change to perform the necessary updates. On the other hand, if CM does not receive the periodic message from its MCH over a time period T, it must change its state to UN and join another cluster.

4.3.6.4 Join a cluster

When a UN vehicle approaches toward a cluster (comes inside its communication range), it sends a join request including its position and velocity to the nearest CM of the cluster. The CM forwards this join request to its MCH, which calculates its relative velocity with this UN vehicle. If this relative velocity is less than or equal to the average relative velocity of the cluster, the MCH adds this UN vehicle to its CM_list and sends a reply to this UN to confirm its cluster membership. Consequently, this UN changes its state to CM and joins the cluster. Furthermore, the MCH must send an update message to its SCH indicating this change.

4.3.7 Theoretical analysis

In this section, we discuss the rational and performance of the proposed clustering approach.

4.3.7.1 MCH selection algorithm complexity

Based on Algorithm 10, we assume that every vehicle and its multi-hop neighbors are modeled by an undirected graph $G(V, E)$, where V is the set of vertices, representing the vehicles and E is the set of edges representing the set of communication links between vehicles. Assume n ($n = |V|$) is the number of vehicles and m ($m = |E|$) is the number of communication links between them. According to Algorithm 10, for a given vehicle i to browse its multi-hop neighbors (MN_i), it must execute a BFS algorithm. Each vehicle visited by i is inserted into the queue and marked as visited. Because the insertion to the queue is done in $O(1)$, the time complexity in the worst case to traverse all the multi-hop neighbors is $O(n)$. Moreover, the edges between the traversed vehicles are visited at most m times. Therefore, the complexity of Algorithm 10 is $O(n + m)$.

4.3.7.2 Message overhead analysis

The message overhead counts all the control messages received by each vehicle in the network during the clustering process. To simplify the analysis, the following definitions are used.

- ***N:*** The number of vehicles in the network.

- ***R:*** The number of RSUs-G installed on the roadside (see Eq 4.1).
- ***h_i:*** The number of vehicles which have received the HELLO message broadcasted by the RSU-G i.
- ***b:*** The number of the BEACON messages broadcasted by a vehicle. $b = \Theta(1)$ because b is proportional to node velocity v and inversely proportional to the transmission range Tr, and both v and Tr are less than or equal to some constants [122].
- ***c:*** The number of the elected MCHs.
- ***r_i:*** The number of REPLY messages received by an MCH i.
- Φ_{REG}**:** The overhead of the registration phase.
- Φ_{NEIGH}**:** The overhead of the neighborhood discovery phase.
- Φ_{ANN}**:** The overhead of the announcement phase.
- Φ_{AFF}**:** The overhead of the affiliation phase.
- Φ_{TOTAL}**:** The total overhead.

During the registration phase, each RSU-G broadcasts a HELLO message to invite the vehicles to join the network. Then, each vehicle sends a REGISTER request to the appropriate RSU-G. Thus, the registration phase message overhead Φ_{REG} may be expressed as follows:

$$\Phi_{REG} = \Theta(N) + \Theta(\sum_{i=1}^{R} h_i) \tag{4.13a}$$

Knowing that $(\sum_{i=1}^{R} h_i) \leq N$, so:

$$\Phi_{REG} = \Theta(N) \tag{4.13b}$$

During the neighborhood discovery phase and to announce its existence, each vehicle sends a BEACON message to its single-hop neighbors. Then, each vehicle sends a

SHARE message (to share its MCR) to the appropriate RSU-G. Therefore, the neighborhood discovery phase message overhead Φ_{NEIGH} can be expressed as follows:

$$\Phi_{NEIGH} = \Theta(b.N) + \Theta(N) \tag{4.14a}$$

$$\Phi_{NEIGH} = \Theta(\Theta(1).N) + \Theta(N) \tag{4.14b}$$

$$\Phi_{NEIGH} = \Theta(N) \tag{4.14c}$$

During the announcement phase, each elected MCH must send an ANNOUNCE message to the RSU-G. Thus, the announcement phase message overhead Φ_{ANN} may be expressed as follows:

$$\Phi_{ANN} = \Theta(c) \tag{4.15}$$

During the affiliation phase, every MCH node receives a number of REPLY messages from its multi-hop neighbors. Then, each MCH sends a NOMINATION message to the member elected as SCH. Therefore, the affiliation phase message overhead Φ_{AFF} can be expressed as follows:

$$\Phi_{AFF} = \Theta(\sum_{i=1}^{c} r_i) + \Theta(c) \tag{4.16a}$$

Knowing that $(\sum_{i=1}^{c} r_i) \leq N$ and $c \leq N$, so:

$$\Phi_{AFF} = \Theta(N) \tag{4.16b}$$

Finally, the total message overhead Φ_{Total} is as follows:

$$\Phi_{TOTAL} = \Phi_{REG} + \Phi_{NEIGH} + \Phi_{ANN} + \Phi_{AFF} \tag{4.17a}$$

$$\Phi_{TOTAL} = \Theta(3N) + \Theta(c) \tag{4.17b}$$

Knowing that $c \leq N$, so:

$$\Phi_{TOTAL} = \Theta(N) \tag{4.17c}$$

4.3.7.3 Clustering properties

To meet the requirements imposed by the VANET characteristics and to demonstrate the effectiveness of the proposed approach, the following clustering properties must be verified.

Definition 5 *Safety property: Each cluster has one and only one MCH, and each ordinary vehicle can belong to only one cluster.*

The safety property ensures that every cluster has a unique MCH. It also ensures that each ordinary vehicle belongs to only one cluster at a time. A safety property asserts that nothing bad happens during the clustering algorithm.

Lemma 6 *The safety property is satisfied.*

Proof 6 *According to Algorithm 10, vehicle i is an MCH if it satisfies the following conditions:*

Condition 1: *It has the lowest Mobility Rate (MR) compared with its multi-hop neighbors.*

Condition 2: *It has the lowest id, if two or more nodes have equal MR (the smallest one):*

$$MR_i = MR_j = .. = MR_k \implies id_i = min(id_i, id_j, .., id_k) \tag{4.18}$$

This implies that each cluster has a single MCH. On the other hand, each ordinary vehicle (not MCH) elects the node that has the lowest MR value among its multi-hop neighbors as its MCH (myMCH). Then, it must send a REPLY message to the appropriate MCH to confirm its membership. So, each node can belong to only one cluster. As a result, the safety property is verified.

Definition 6 *Liveness property: Cluster formation phase terminates and each vehicle is either a UN, MCH, CM or an SCH at a given time.*

The liveness property ensures that the clustering algorithm progresses normally and ends after a finite time and each vehicle is in a stable state at a given time. Typically, a liveness property asserts that something good eventually happens.

Lemma 7 *The liveness property is verified.*

Proof 7 *First, since every vehicle can determine its cluster according to Lemma 6, the cluster formation phase will terminate. Second, based on transition state (see Figure 4.7), each transition from one state to another is due to an event. The possible transitions are described in Table 4.4.*

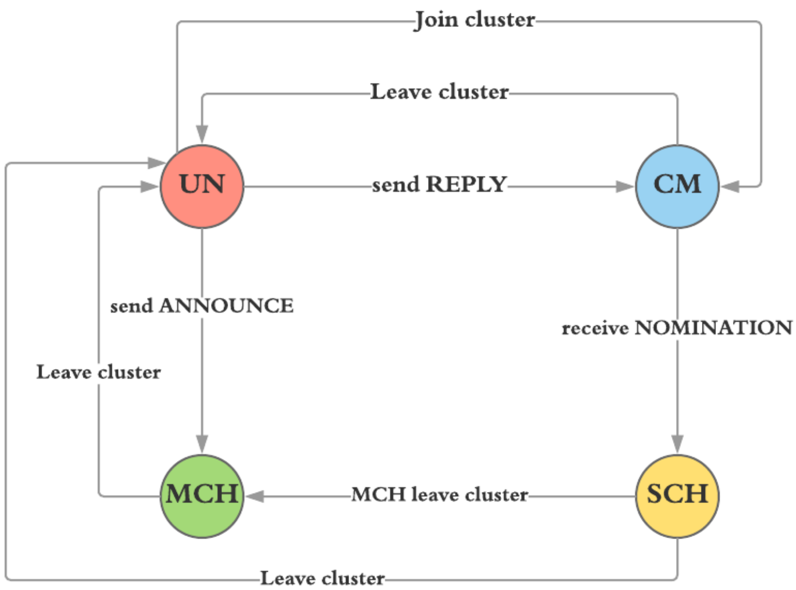


Figure 4.7: Vehicles' state transition

Table 4.4: Transitions and their corresponding events.

Transition	Event	Phase
UN to CM	UN sends REPLY message to MCH	Cluster formation
UN to CM	UN joins a new cluster	Maintenance
UN to MCH	UN sends ANNOUNCE message to RSU-G	Cluster formation
CM to SCH	CM receives NOMINATION message from MCH	Cluster formation
SCH to MCH	MCH leaves the cluster	Maintenance
SCH to UN	SCH leaves the cluster	Maintenance
CM to UN	CM leaves the cluster	Maintenance
MCH to UN	MCH leaves the cluster	Maintenance

4.4 Performance evaluation

In this section, we study the performances of the proposed MCA-V2I approach using the network simulator NS-2 [109] and VanetMobiSim [110] integrated environment. The simulation is performed on a machine with Intel $i5$ (4th generation) processor and 8 GB of RAM. Mobility is simulated on a one-directional highway of 6 km length with three lanes. There are 2 RSUs-G installed on the roadside. Physical and MAC layers are configured according to the 802.11p standard. The speed of vehicles varies uniformly between 10 m/s and 35 m/s ($\simeq$ 40 km/h – 125 km/h). Moreover, the transmission range of vehicles is varied from 100 m to 300 m. The simulation period in this work is 360 s. The vehicles were assigned to random positions and they move according to the mobility model, named the Intelligent Driver Model including Lane Change (IDM-LC) [111, 112], which is integrated into VanetMobiSim. The propagation model used is Two-ray Ground. The different simulation parameters are listed in Table 4.5.

Table 4.5: Simulation parameters.

Parameter	Value
Simulation time	360 s
Simulation area	6000 × 50 m
Transmission range	100 – 300 m
Number of RSU-G	2
Number of vehicles	60 – 180
Propagation model	Two-ray Ground
Mobility model	IDM-LC
MAC/PHY protocol	802.11p
Velocity of vehicles	10 – 35 m/s
Maximum allowed velocity (v_{max})	40 m/s

We compare the results of our proposed approach MCA-V2I to two well-known protocols for VANET belonging to the same family of multi-hop clustering, named N-hop [96] and DMCNF [61]. The comparison is based on the following metrics:

- **Cluster Head Lifetime (CHL):** The interval of time from when a vehicle changes its state to CH until this vehicle leaves this state and changes to another state (e.g., UN). The average CHL is calculated by dividing the total CHL by the total number of state changes from CH to another state. A longer CHL leads to more reliable communication with minimized overhead.

- **Cluster Member Lifetime (CML):** The interval of time from when a vehicle changes its state to CM (join a cluster) to when this vehicle changes from this state to another state. The average CML is computed by dividing the total CML by the total number of state changes from CM to another state. A longer CML can show the stability of the constructed clusters and the effectiveness of the maintenance techniques used.

- **Cluster Head Change Number (CHCN):** The number of state changes from CH to another state (e.g., UN). Low CHCN can demonstrate the cluster's stability.

- **Cluster Number (CN):** The number of clusters formed during the simulation period. Fewer clusters can indicate the efficiency of the clustering algorithm.

- **Clustering Overhead (CO):** The total number of control messages received by each vehicle in the network during the phase of cluster's formation.

- **Message Delivery Latency (MDL):** Refers to the average delay or time taken for a message to be transmitted from a source to a destination.

- **Message Delivery Ratio (MDR):** The average number of messages that have been successfully received by the destination divided by the average number of messages sent by the source.

4.4.1 Cluster Head Lifetime (CHL)

Figure 4.8 shows the average CHL of the proposed MCA-V2I approach versus DMCNF and N-hop protocols under different transmission ranges. According to Figure 4.8, we observe that when a vehicle's velocity increases, the average CHL of MCA-V2I, DMCNF and N-hop decreases relatively. This is because the network topology becomes very dynamic due to the high mobility of vehicles, which makes it difficult for the cluster's heads to maintain stable connections with their CMs. On the other hand, when the transmission range increases, the average CHL also increases. This can be justified by the fact that in a wide range of transmission, the coverage area of the cluster increases and that the CH can find at least one CM to serve it, so that a vehicle continues to reside in the state CH for a longer period of time. In both DMCNF and

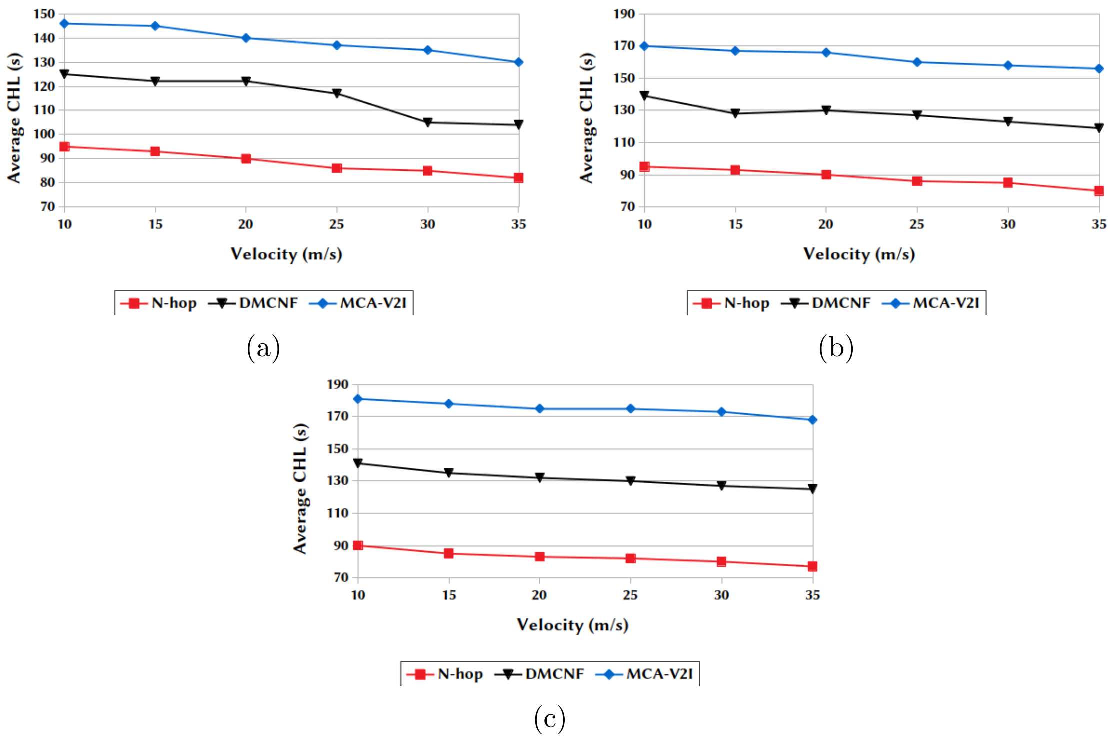
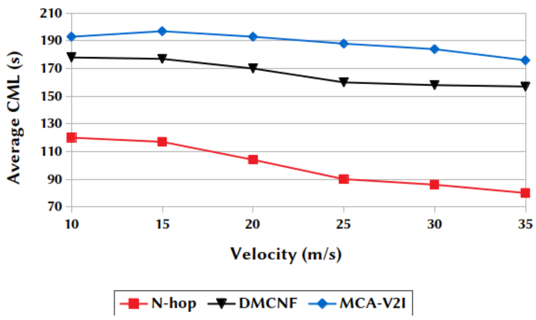


Figure 4.8: Average CHL under different transmission ranges. (a) 100 m. (b) 200 m. (c) 300 m.

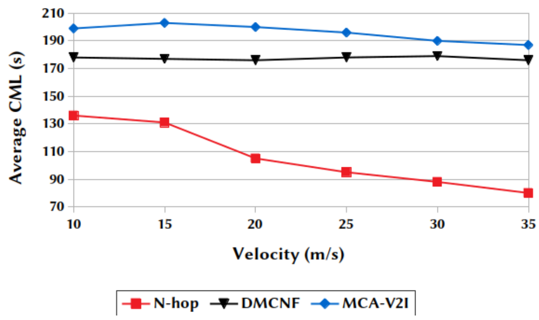
N-hop, the vehicles that have the smaller average relative velocity with their single-hop neighbors are suitable to be elected as CH. Consequently, this metric alone may lead both protocol DMCNF and N-hop to elect CHs which have very low connectivity with their CMs. However, in MCA-V2I, the election of MCHs is performed using mobility rate, which combines more than one metric, such as node connectivity, average relative velocity, average distance and link stability with their single-hop neighbors, in which the vehicles that have the lowest MR are elected as MCH. Thus, as shown in Figure 4.8, MCA-V2I outperforms both DMCNF and N-hop in term of CHL.

4.4.2 Cluster Member Lifetime (CML)

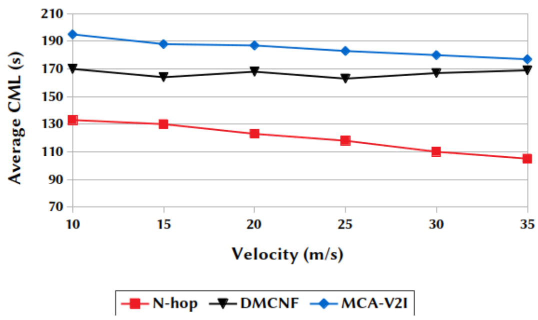
Figure 4.9 shows the average CML of MCA-V2I versus DMCNF and N-hop scheme under different transmission ranges. As shown in Figure 4.9, the vehicle velocity moderately affects the CML for MCA-V2I compared with DMCNF and N-hop, owing to the effective clustering algorithm used. This latter allows the CMs to maintain stable connections with their MCHs. Furthermore, the election of SCHs in addition to MCHs makes it possible to increase the cluster's stability and avoid the reclustering. On the other hand, when the transmission range increases, the average CML also increases. This can be justified by the fact that in a wide range of transmission, the coverage area of the cluster increases, which gives the CMs a large area of movement without the loss of communication links with their MCHs. Thus, the MCA-V2I scheme outperforms both N-hop and DMCNF in terms of CML.



(a)



(b)



(c)

Figure 4.9: Average CML under different transmission ranges. (a) 100 m. (b) 200 m. (c) 300 m.

4.4.3 Cluster Head Change Number (CHCN)

Figure 4.10 shows the average CHCN of MCA-V2I versus N-hop and DMCNF approaches under different transmission ranges. Figure 4.10 demonstrates that the average CHCN when using N-hop and DMCNF is higher than when MCA-V2I is used. The reason for this improvement is the effective initial MCHs selection using mobility metrics and Internet access, which allows the MCHs to keep stable connections with their CMs as long as possible.

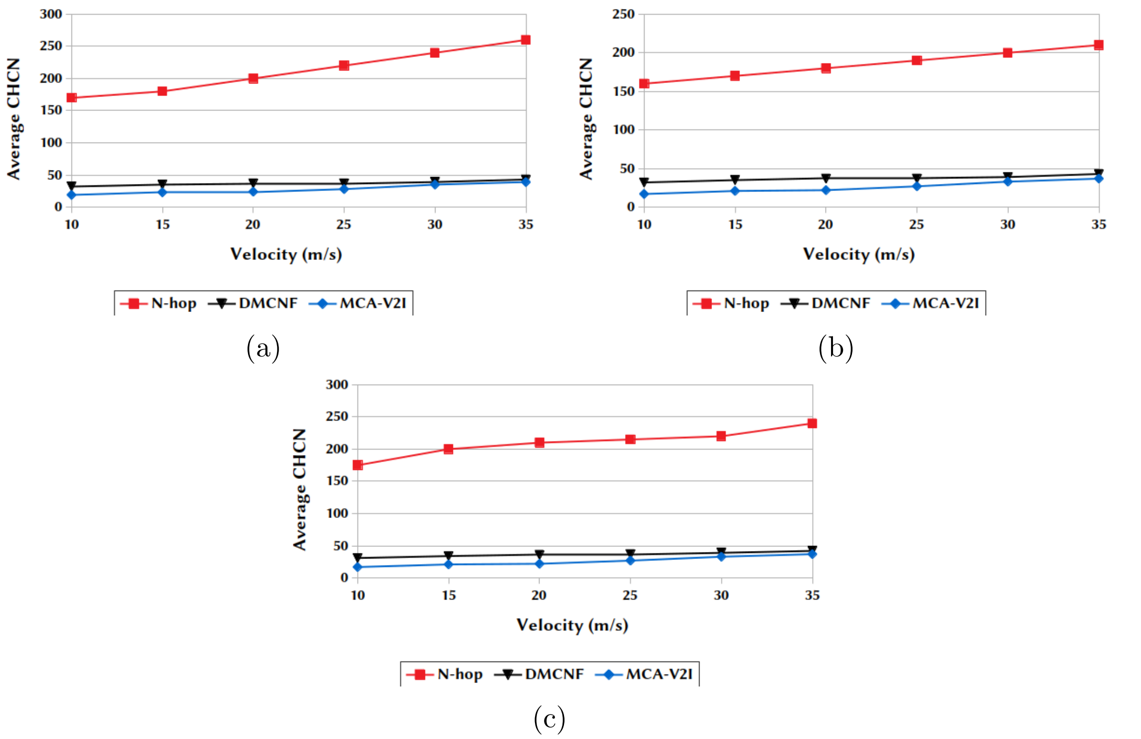


Figure 4.10: Average CHCN under different transmission ranges. (a) 100 m. (b) 200 m. (c) 300 m.

4.4.4 Cluster Number (CN)

Figure 4.11 illustrates the average number of clusters of the MCA-V2I scheme versus N-hop and DMCNF schemes under different transmission ranges. According to Figure 4.11, the proposed approach has fewer clusters compared with both N-hop and DMCNF due to the effective multi-hop clustering process that is based on combined mobility metrics and the BFS algorithm. On the other hand, when the transmission range increases, the average CN decreases. This can be justified by the fact that in a wide range of transmission, the coverage area of the cluster increases, which gives the MCHs the ability to handle more vehicles.

4.4.5 Clustering Overhead (CO)

Figure 4.12 depicts the average clustering overhead of MCA-V2I, N-hop and DMCNF for different velocity values. MCA-V2I significantly decreases the number of overhead messages compared with N-hop and DMCNF. In MCA-V2I, every vehicle can access the Internet and exploit the shared MCRs to perform the clustering process. This results in a significant reduction in the number of control overhead messages. On the

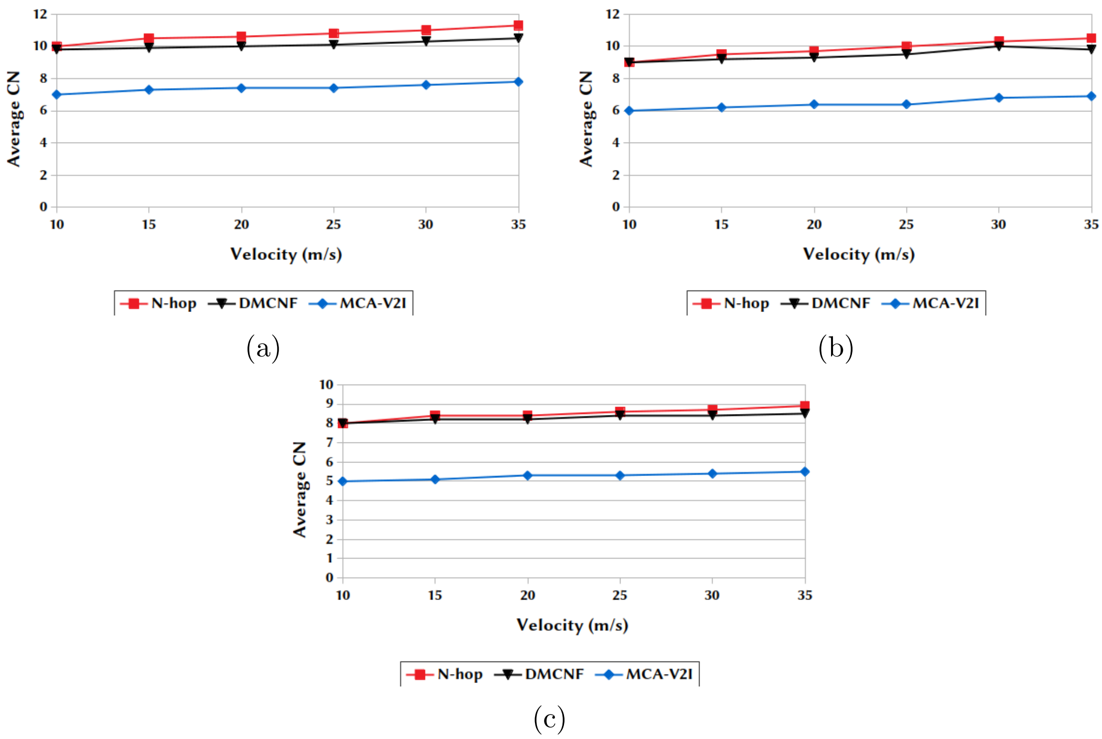


Figure 4.11: Average CN under different transmission ranges. (a) 100 m. (b) 200 m. (c) 300 m.

other hand, each vehicle in N-hop and DMCNF exchanges a control message with all its single-hop neighbors to calculate the relative mobility between them, to elect the CHs. This leads to an increase in the number of control overhead messages in both N-hop and DMCNF.

4.4.6 Message Delivery Latency (MDL)

Figure 4.13 illustrates the message delivery latency (in ms) of MCA-V2I, N-hop and DMCNF as a function of the number of simulated vehicles. The message delivery latency inversely proportional to the number of vehicles. High density of the vehicles improves the connectivity of the networks and therefore there are more chances to deliver the message with the shorter expected delivery delay to the destination. Compared with DMCNF and N-hop, our proposed MCA-V2I scheme exhibits the lowest message delivery latency for all numbers of simulated vehicles.

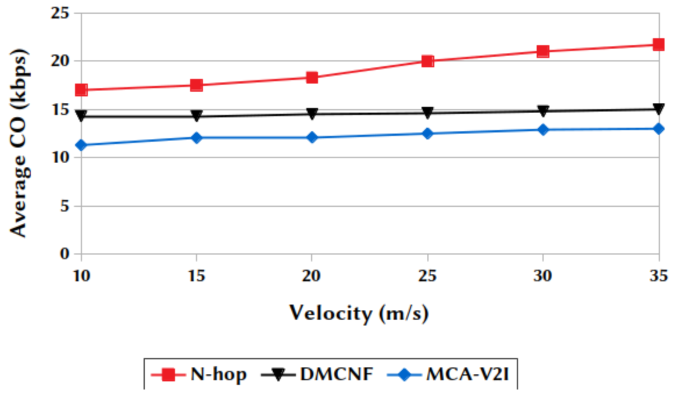


Figure 4.12: Average CO.

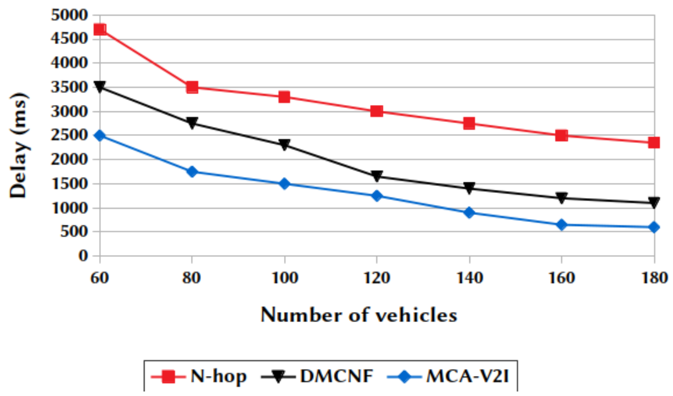


Figure 4.13: Message Delivery Latency (MDL).

4.4.7 Message Delivery Ratio (MDR)

Figure 4.14 illustrates the message delivery ratio of MCA-V2I, N-hop and DMCNF as a function of the number of simulated vehicles. The message delivery ratio increases quickly with the increase in the number of vehicles. This is because the growth of the density of vehicles improves the connectivity of the network and therefore more chances to deliver the message successfully. Compared with DMCNF and N-hop, our proposed MCA-V2I scheme exhibits the highest message delivery ratio for all numbers of simulated vehicles.

4.5 Conclusion

In this chapter, we propose a new Multi-hop Clustering Approach over Vehicle-to-Internet communication called MCA-V2I for improving VANETs' performances. MCA-V2I allows vehicles to connect to the Internet via a special infrastructure called a Road Side Unit Gateway (RSU-G) so that each vehicle can obtain and share the necessary

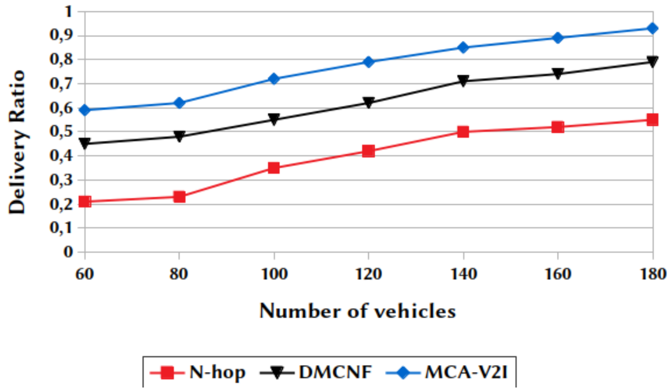


Figure 4.14: Message Delivery Ratio (MDR).

information about its Multi-hop neighbors to perform the clustering process. This latter is performed using a BFS algorithm for traversing the graph and based on a Mobility Rate (MR), which is calculated according to mobility metrics. The MCA-V2I approach strengthens the cluster's stability through the election of a Slave Cluster Head (SCH) in addition to the Master Cluster Head (MCH). Our simulation uses network simulation NS-2 and the VanetMobiSim integrated environment. The simulations' results show that the proposed scheme MCA-V2I outperforms N-hop and DMCNF schemes in terms of CH lifetime, CM lifetime, CH change number, number of clusters, clustering overhead, message delivery latency and message delivery ratio.

The next chapter presents the last contribution of our thesis. In this regard, we propose an efficient weight-based clustering algorithm using mobility report for IoV.

Chapter 5

An Efficient Weight-Based Clustering Algorithm using Mobility Report for IoV

5.1 Introduction

With the emergence of the Internet of Things (IoT), the IoV has recently been the focus of many researchers as a new paradigm of ITS and the industry to provide an effective solution to ensure high safety on the road and comfort for drivers [123]. The mobility is a major challenge for Mobile Ad-hoc NETworks (MANET) generally and for Vehicular Ad-hoc NETwork (VANET) in particular [107]. VANET has a high mobility due to the high speed of the vehicles, which leads to a very dynamic topology. This causes a large number of control messages to react to the frequent change of the topology. For this reason, providing effective and reliable routing solutions for IoV becomes a big challenge.

Clustering is one of the methods that is being used to properly control the network by optimizing and reducing the flow of control [3]. It is the process of dividing the network into smaller sized groups called clusters in a hierarchical manner. Each cluster has a Cluster Head (CH), one or more nodes that act as gateways between clusters and many other ordinary nodes that act as members [101]. Recently many clustering protocols have been proposed for IoV. The main challenge for these proposed algorithms is to ensure the stability and the maintenance of the cluster structure. Most of the existing VANET clustering protocols are derived from the classic MANET clustering schemes [124]. These proposed algorithms do not take into consideration the mobility characteristics of VANET.

In this chapter, we introduce an Efficient Weight-based Clustering Algorithm using a Mobility Report (WECA-MR) for IoV suitable for highway environment. It uses a simple distributed algorithm using both classic weighted metrics such as degree, average distance and new metric introduced by our proposed approach called Mobility Report (MR). This new metric combines other mobility metrics: velocity and acceleration. Our proposed approach takes into account the possibility that the CH leaves the cluster by introducing the election of Backup CH (BCH), in order to increase the stability of clusters and reduce the communication overhead. We compare the performance of the proposed algorithm using network simulator NS-2 and VanetMobiSim integrated environment to show the performances of the proposed clustering approach.

The rest of this chapter is organized as follows: Section 5.2 describes the proposed approach in details. Section 5.3 presents the experimental results. Finally, conclusion is presented in section 5.4.

5.2 Proposed approach

In this section, we introduce an Efficient Weight-Based Clustering Algorithm using Mobility Report for IoV called WECA-MR. The proposed algorithm focuses only on Vehicle-to-Vehicle (V2V) communication, and its main objective is to maintain the clusters and increase their stability to face the different changes, that affect the topology of the network due the high mobility of vehicles. To elect the CHs, the proposed scheme uses both classic weighted metrics: node degree, average distance and a new metric introduced by our scheme called Mobility Report (MR). This new metric combines other mobility parameters such as: velocity and acceleration. WECA-MR takes into account the possibility that the CH may leave the cluster, by the election of Backup CH (BCH) based also on weight of nodes. Next, we will present the network model for our proposed scheme and describe the main steps of the suggested approach including cluster initialization and maintenance phases.

5.2.1 Network model

The proposed scheme is based on the following assumptions. First, each vehicle has a unique *id* in the network. Second, every vehicle is equipped with digital road map and a GPS device, that allows it to obtain its real-time geographic location, instant velocity and location of its direct neighbors. Third, each vehicle is equipped with wireless interface, in order to communicate with other vehicles. Fourthly, each node broadcasts its mobility informations to its direct neighbors using *Hello* messages. Finally, we assume that we have a highway with two roads (one for each direction), and three lanes for each road.

In addition, the vehicular network is modeled theoretically as an undirected graph $G(V, E)$ where V is the set of vertices representing the vehicles in the network, and E is the set of edge representing the set of the communications link between vehicles.

5.2.2 Cluster initialization

5.2.2.1 Combined weight

WECA-MR algorithm performs the clustering process based on a combined weight of the metrics described below. The weight W_i of node i is calculated using the following parameters:

- **Node degree:** It is the size of the set of single-hop neighbors of node i, N_i, where

$$N_i = \{j \in V \mid \exists\ (i, j) \in E\} \tag{5.1}$$

  Consequently, the degree of node i is defined as follows:

$$deg_i = |N_i| \tag{5.2}$$

  A node , with maximum degree (it has high connectivity), is suitable to be elected as CH.

- **Distance indicator:** The second parameter indicates the average distance between a node and its single-hop neighbors. A node, that has a minimum average distance, will have the highest distance indicator. This kind of node is suitable to be elected as CH. For a node i, its average distance AD_i is computed as the cumulative mean square distance to its single-hop neighbors (N_i) divided by its degree (deg_i):

$$AD_i = \frac{1}{deg_i} \sum_{j=1}^{deg_i} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{5.3}$$

Therefore, the distance indicator DI_i of node i, is calculated as:

$$DI_i = \left| \ln(\frac{AD_i}{maxD_i}) \times 10^2 \right| \tag{5.4}$$

Where $maxD_i$ is maximum distance between node i and its neighbors.

- **Mobility report:** In our proposed approach, we introduce a new parameter, called Mobility Report (MR), which depends on two mobility metrics: relative velocity and relative acceleration. A high MR rate indicates that the node is suitable to be selected as CH. Let us assume that $P_1(x_1, y_1)$ is the position of node i at time T_1 and $P_2(x_2, y_2)$ is the position of node i at time T_2. d_i is the distance traveled by the node i over time $\Delta t(\Delta t = T_2 - T_1)$:

$$d_i = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{5.5}$$

Thus, the velocity of node i over time Δt, is computed as:

$$v_i = \frac{d_i}{\Delta t} \tag{5.6}$$

Let us assume that v_1 is the velocity of node i at time T_1 and v_2 is the velocity of node i at time T_2. a_i is the acceleration of the node i over time $\Delta t(\Delta t = T_2 - T_1)$. Consequently, the acceleration of node i is computed as follows:

$$a_i = \frac{\Delta v_i}{\Delta t} \tag{5.7}$$

Where $\Delta v_i = |v_1 - v_2|$.

Based on the exchanges of velocity and acceleration metrics between neighboring nodes using periodic messages, each node calculates its relative velocity (Rv) and relative acceleration (Ra) with respect to its single-hop neighbors.

$Rv_{(i,j)}$ is the relative velocity of node i relative to node j, calculated as follows:

$$Rv_{(i,j)} = \ln(1 - \frac{|v_i - v_j|}{v_{max}}) \tag{5.8}$$

Where v_{max} is maximum velocity allowed on the road.

$ARv_{(i,N_i)}$ is the average relative velocity of node i relative to its single-hop neighbors (N_i). It is defined as follows:

$$ARv_{(i,N_i)} = \frac{\sum_{j=1}^{deg_i} Rv_{(i,j)}}{deg_i} \tag{5.9}$$

$Ra_{(i,j)}$ is the relative acceleration of node i relative to node j, calculated as follows:

$$Ra_{(i,j)} = \ln(1 - \frac{|a_i - a_j|}{a_{max}}) \tag{5.10}$$

Where: a_{max} is maximum acceleration allowed according to v_{max}.

$ARa_{(i,N_i)}$ is the average relative acceleration of node i relative to its single-hop neighbors (N_i). It is defined as follows:

$$ARa_{(i,N_i)} = \frac{\sum_{j=1}^{deg_i} Ra_{(i,j)}}{deg_i} \tag{5.11}$$

Finally, the mobility report of node i is calculated as follows:

$$MR_i = ARv_{(i,N_i)} \times ARa_{(i,N_i)} \tag{5.12}$$

The weight W_i of node i is calculated based on the previous parameter as follows:

$$W_i = C_1 \times deg_i + C_2 \times DI_i + C_3 \times MR_i \tag{5.13}$$

Where C_1, C_2 and C_3 are constant coefficients.

$$C_1 + C_2 + C_3 = 1 \tag{5.14}$$

5.2.2.2 Cluster head election & cluster formation

- **Step 1:** Initially, all vehicles are in the Undefined State (US). Each vehicle broadcasts a *HELLO* message to its single-hop neighbors, including its *id*, its velocity, its acceleration, its transmission range and its position (two-dimensional coordinates). The initialization steps executed by each vehicle in the network are described by Algorithm 11.

Algorithm 11: Vehicle Initialization

state $\leftarrow$ *US*
Calculate velocity v_i *based on Eqs. 5.5 and 5.6*
Calculate acceleration a_i *based on Eq. 5.7*
Get transmission range Tr_i
Get position $P_i(x_i,y_i)$
Broadcasts HELLO(id_i, v_i, a_i, Tr_i, P_i)

- **Step 2:** After receiving the *HELLO* message from all its single-hop neighbors (N_i), each node calculates the different parameters such as degree, distance indicator and mobility report, in order to computes its weight based on Eqs. 5.2 to 5.13. Then, the node broadcasts a *REQUEST* packet to its single-hop neighbors, including its *id* and its weight.

- **Step 3:**

  When a node receives the *REQUEST* packets from all its single-hop neighbors, it compares its weight with all weights received. If its weight has the largest value, the node must announce its election as new CH and updates its state to CH. Thus, the newly elected CH will broadcast a *ANNOUNCE* message to its single-hop neighbors, and wait for their replies. Otherwise, the node elects the node that has the largest weight value as its new CH, updates the variable *currentCH* (variable that indicates the *id* of its current CH) at *id* of new CH and waits for a *ANNOUNCE* message from this CH. If there are two or more nodes that have the largest weight, the node that has the lowest *id* will be elected as CH. The CH selection process is described in Algorithm 12.

Algorithm 12: CH Selection

```
state_i: state of vehicle i
N_i: set of single-hop neighbors of node i
currentCH_i: variable contains id of the CH of node i

Receive HELLO(id, v, a, Tr, P) messages from neighbors
foreach HELLO(id_j, v_j, a_j, Tr_j, P_j) message do
    if distance(i, j) ≤ min(Tr_i, Tr_j) then
        N_i ← N_i ⋃ j
deg_i ← |N_i|
Calculate the distance indicator DI_i based on Eqs. 5.3 to 5.4
Calculate the wight W_i based on Eqs. 5.8 to 5.13
Broadcasts REQUEST(id_i, W_i)
Receive REQUEST(id, W) messages from neighbors
max_W ← max(W_i, W_j, ..., W_m)
id_max ← id of node that has max_W
if W_i = max_W then
    currentCH ← id_i
    state ← CH
    Broadcasts ANNOUNCE(id_i)
    Cluster_formation_&_BCH_election(i)
else
    currentCH ← id_max
    Receive ANNOUNCE(id_j) messages from CH
    if id_j = currentCH then
        Send REPLY(id_i) message to CH
        state_i ← CM
```

- **Step 4:** Each node, after it receives *ANNOUNCE* message from a new CH, compares the *id* of sender of the *ANNOUNCE* message with *currentCH*, if they are the same, the node sends a *REPLY* to this CH and updates its state to CM. Otherwise, the node ignores the message and waits for a new *ANNOUNCE* message from another CH that has the same *id* recorded in *currentCH*.

- **Step 5:** Each CH, for each *REPLY* received, adds the sender node to its cluster members list.

- **Step 6:** After cluster formation, the CH must selects a node with the highest weight value among cluster members (except itself) as Backup CH (BCH). Then, the CH sends a *SET_BCH* to the designated node.

- **Step 7:** The node that receives the *SET_BCH* updates its state to BCH. The cluster formation and BCH election are described in Algorithm 13. The process of the cluster initialization phase executed by each vehicle is illustrated in Figure 5.1 as flowchart diagram.

Algorithm 13: Cluster Formation

CM_list_i: cluster member list of CH i;
CM_list $\leftarrow \emptyset$
Receive REPLY(id) messages from neighbors
foreach *REPLY(id_j) message* **do**
    $\text{CM\_list}_i \leftarrow \text{CM\_list}_i \bigcup id_j$
$max\_W \leftarrow maxWeight(\text{CM\_}list)$
$id\_max \leftarrow id$ of node that has max_W
Send SET_BCH(CM_list_i) to node that has id_max

Figure 5.2 shows a simple example of cluster initialization phase of our proposed scheme with 11 nodes. Table 5.1 illustrates the different parameters of the nodes.

Table 5.1: Nodes parameters.

Node id	Neighbors	MR	Weight	CH
1	11, 5	16.25	8.53	5
2	3, 6, 7, 9, 10	18.15	10.23	-
3	2, 6	16.04	8.42	2
4	5	15.95	8.11	5
5	1, 4, 7, 8, 11	18.25	10.57	-
6	2, 3	16.42	8.74	2
7	2, 5	17.19	9.28	5
8	5	16.72	8.77	5
9	2	15.72	7.97	2
10	2	16.37	8.08	2
11	5	16.24	8.16	5

At first, each vehicle broadcasts a HELLO message to its single-hop neighbors (see Figure 5.2b). Then, each node broadcasts a REQUEST packet to its neighbors, including its weight (see Figure 5.2c). Thus, each node compares its weight with the received weights. If its weight has the largest value, the node must announce its election as new CH (in this case nodes 2 and 5 are going to announce that they are the CH). On the other hand, the remaining nodes elect the node that has the largest weight value as its new CH among its neighbors. Therefore, after receiving the ANNOUNCE

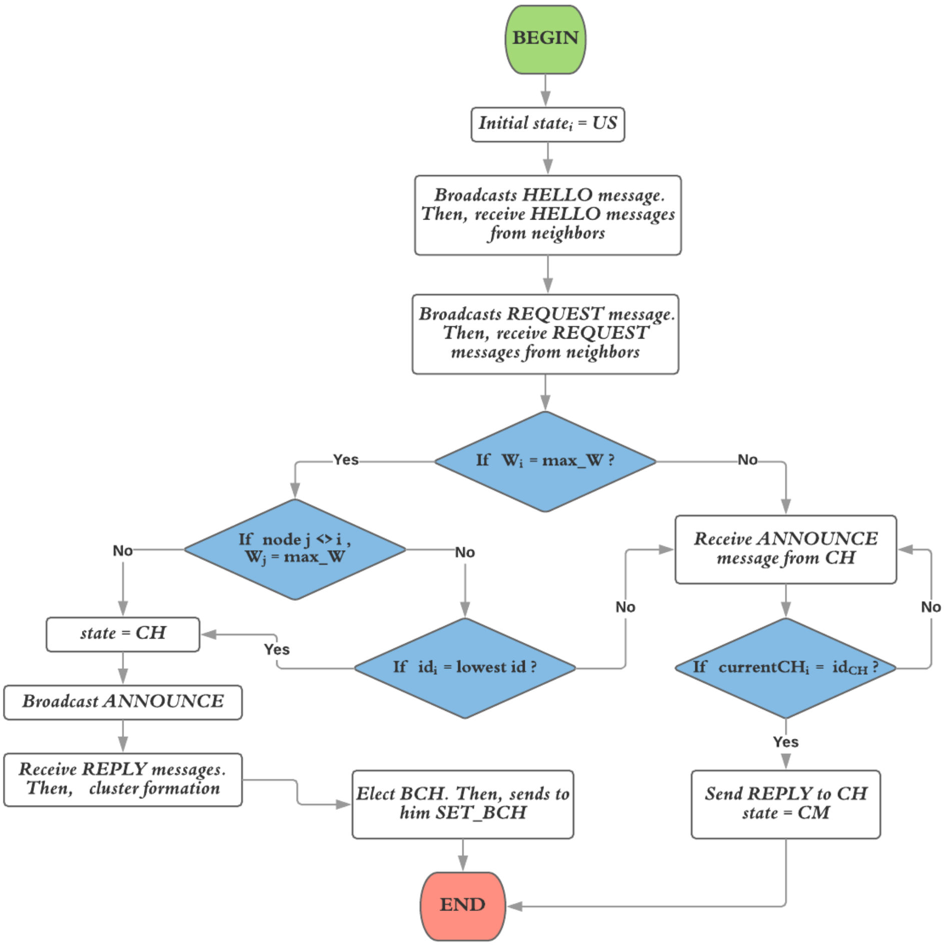
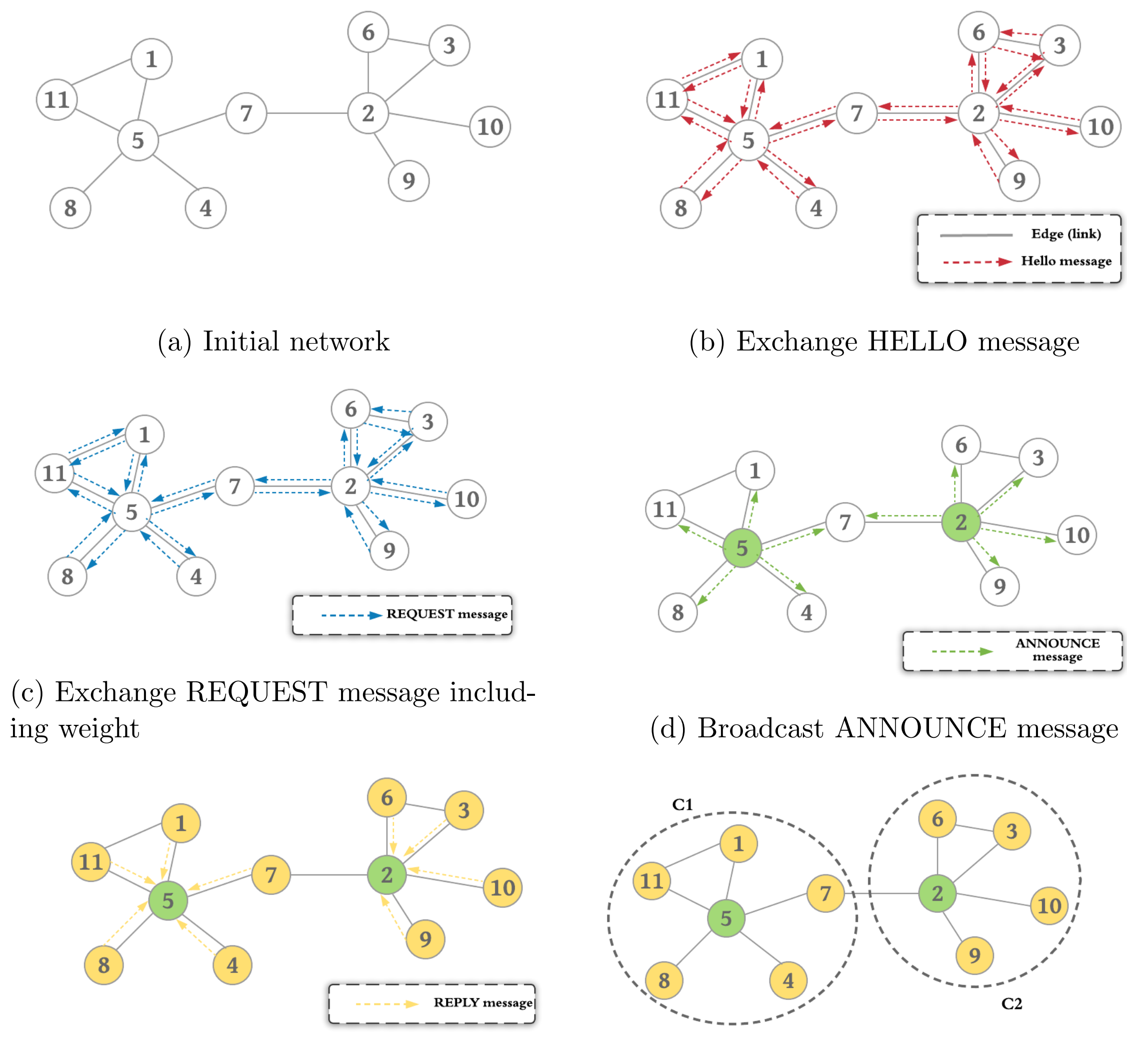


Figure 5.1: Cluster initialization phase

message from the CH, each node responds by sending a REPLY message. If a node receives two or more ANNOUNCE messages, it will respond to the CH that has the largest weight, in this example the node 7 received two ANNOUNCE messages from nodes 2 and 5 (see Figure 5.2e), but it sends a reply to node 5 only. At the end, the state of clusters becomes as follows: C1{CH:5; CM: 1,4,7,8,11} and C2{CH:2; CM: 3,6,9,10} (see Figure 5.2f).

5.2.3 Cluster maintenance

The aim of this phase is to maintain the cluster structure and stability as long as possible. Due to the nature of VANET (vehicles with high mobility), vehicles keep joining and leaving the clusters frequently. Therefore, the cluster maintenance phase is described as follows:



(a) Initial network

(b) Exchange HELLO message

(c) Exchange REQUEST message including weight

(d) Broadcast ANNOUNCE message

(e) Send REPLY message to CH

(f) Final clusters state

Figure 5.2: Example showing cluster initialization phase

5.2.3.1 CH leaving discovery

After cluster formation phase, a private communication link is established between the CH and BCH in the cluster. Thus, the BCH vehicle periodically detects the state of the CH vehicle. If a BCH does not receive a periodic message from its CH over a time period Δt, it means that CH has left the cluster. The BCH takes over as a new CH of cluster. Therefore, it must change its state to CH state, then selects a new BCH from cluster's members based on their weights. Then, it broadcasts an update message to its members to inform them to update their *currentCH* variable. In addition, the new CH must inform the CM that has been elected as new BCH to change its state to BCH. The old CH must change its state to US state and join another cluster.

5.2.3.2 Leave a cluster

In our proposed scheme, we assume that each CH monitors its cluster members through the exchange of periodic messages to record the presence of members in the cluster. When a vehicle member moves out of the cluster range, the CH detects this event and removes immediately this node from its members list. Then, the CH sends a message to its BCH indicating this update (BCH will update its member list to reflect this change). On the other hand, if CM does not receive the periodic message from CH over a time period Δt, it must change its state to US.

5.2.3.3 Join a cluster

After cluster formation phase, US nodes join the different clusters. When a US node approaches towards a CH (comes inside its transmission range), this CH calculates its relative velocity with this US vehicle over a time period Δt. If the relative velocity is less than or equal to the average relative velocity of its members, the CH adds the US node to its members list and informs this US node. On receiving the message, the US node changes its state to CM and joins the cluster. At the end, the CH must send a message to its BCH indicating this update.

5.2.4 Theoretical analysis

In this section, we discuss the rational and performance of the proposed clustering scheme.

5.2.4.1 Message complexity

Assuming n is the number of vehicles in network, m ($m < n$) is the maximum number of neighbors for a node in the network, l ($1 \leq l \leq m$) is the number of neighbors for a node and c ($c < n$) is the number of CHs. Initially, each vehicle sends l HELLO message to its single-hop neighbors, this implies a cost in the worst case of $O(nm)$. Second, every node after calculating its weight, sends l REQUEST messages to its single-hop neighbors, with a complexity in the worst case $O(nm)$. Third, each elected CH sends l ANNOUNCE message to its neighbors, with a message complexity in the worst case of $O(cm)$. Fourthly, every CH node receives r ($r \leq l$) REPLY messages in order to join its cluster, with cost in the worst case of $O(cm)$. Consequently, total overhead in the algorithm is $O(2nm + 2cm)$, that leads to the total cost in the worst case of $O(n)$.

5.3 Simulation

The proposed algorithm is implemented in NS-2 with VanetMobiSim integrated environment. The simulations are performed with 100 vehicles on a highway with single direction. The simulation is based on one directional highway with three lanes. The velocity varies between 10 m/s and 35 m/s. The parameters and the settings of the proposed scheme are listed in Table 5.2.

Table 5.2: Simulation parameters.

Parameter	**Value**
Number of vehicles	100
Simulation time	300 s
Area	1000 x 50 m
Speed of vehicles	10 - 35 m/s
Maximum allowed velocity (v_{max})	40 m/s
Coefficients C1, C2 and C3	0.5, 0.3 and 0.2

We compare the results of our proposed scheme to MOSIC protocol [92] using the following metrics: CH lifetime, CM lifetime, number of clusters and clustering overhead.

5.3.1 CH lifetime

Figure 5.3 shows the average CH lifetime of our proposed scheme WECA-MR versus MOSIC protocol under different velocities. According to Figure 5.3, our proposed scheme WECA-MR outperforms MOSIC protocol in term of CH lifetime.

5.3.2 CM lifetime

Figure 5.4 depicts the average CM lifetime of our scheme WECA-MR versus MOSIC algorithm under different velocities. WECA-MR outperforms MOSIC protocol in term of CM lifetime.

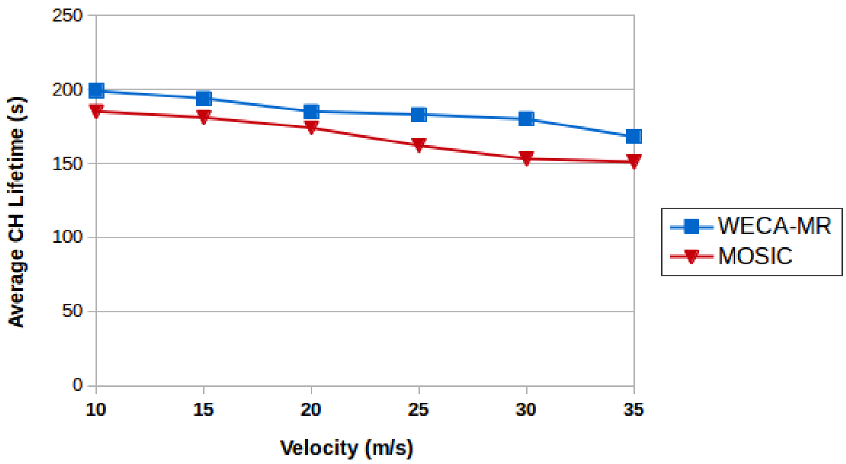


Figure 5.3: CH lifetime

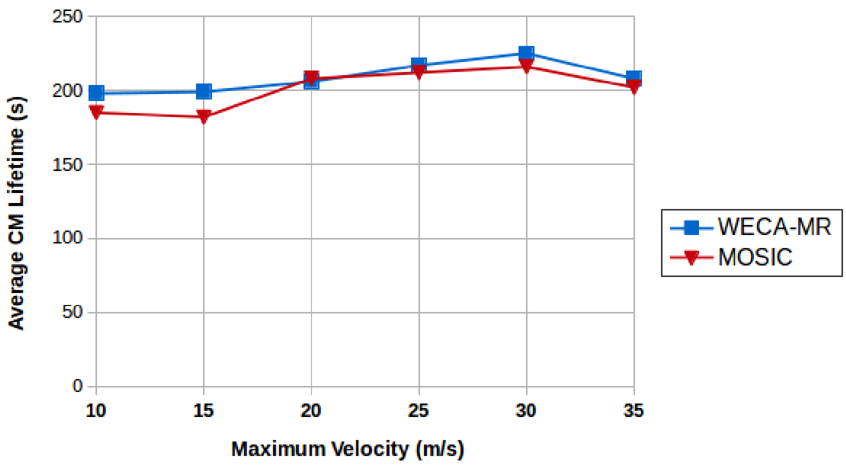


Figure 5.4: CM lifetime

5.3.3 Clusters number

Figure 5.5 illustrates the average numbers of cluster for our scheme WECA-MR versus MOSIC algorithm under different velocities. WECA-MR outperforms MOSIC protocol in term of clusters numbers.

5.3.4 Clustering overhead

Figure 5.6 illustrates the clustering overhead for our approach WECA-MR versus MOSIC protocol under different velocities. WECA-MR outperforms MOSIC protocol in term of clustering overhead.

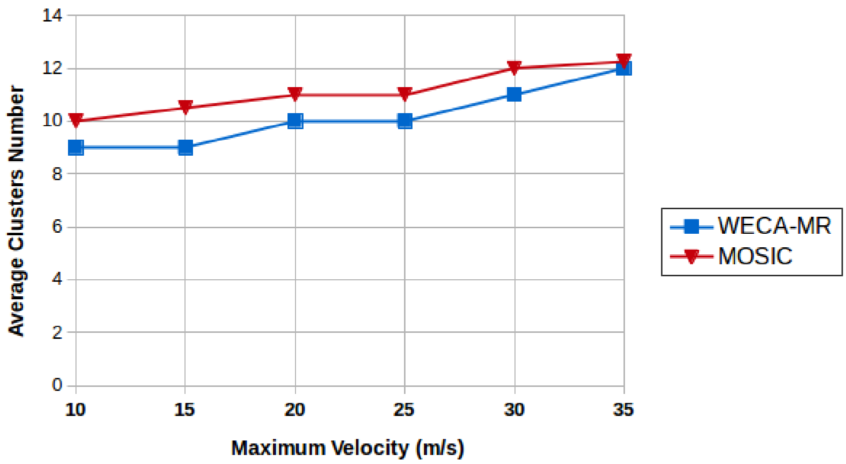


Figure 5.5: Clusters number

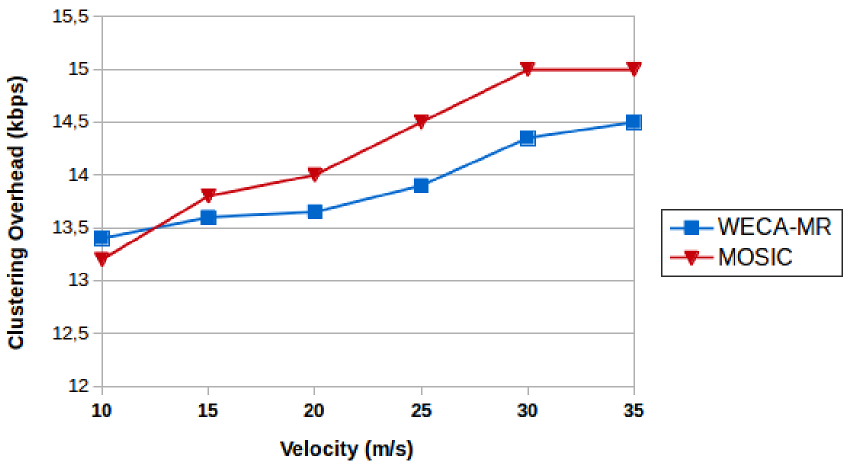


Figure 5.6: Clustering overhead

5.4 Conclusion

In this chapter, we propose an Efficient Weight-Based Clustering Algorithm using a Mobility Report (WECA-MR) for IoV suitable for highway environment. It uses a simple distributed algorithm using both classic weighted metrics such as degree, average distance to neighbors and a new metric called Mobility Report (MR). The new metric combines mobility metrics: velocity and acceleration. Our proposed approach takes into account the leaving of the CH, by the election of Backup CH (BCH) for each cluster by the CH, in order to increase the stability of clusters as long as possible. The main contributions of the proposed scheme are: fast completion of clustering procedure, where both cluster initialization and cluster maintenance phases are taken into account. Our simulation uses network simulation NS-2 and VanetMobiSim integrated environment. Simulation's results show that the proposed scheme WECA-MR outperforms MOSIC protocol in term of CH lifetime, CM lifetime, clusters number and clustering overhead.

Conclusion and future work

As a new ITS paradigm, IoV becomes one of the most active research fields and plays an important role in solving various driving and traffic problems by advanced information and communication technologies. Despite these benefits, IoVs have many design challenges and limitations that must be considered by researchers to suggest new solutions for these networks and improve the existing ones. In this regard, the choice of an efficient network structuring technique for the IoV paradigm is a topic that is still relevant, especially since this type of network is expanding, because of its huge potential for beneficial contributions, especially for security in which the lives of individuals are at stake. Clustering is one of the promising techniques for structuring the network, which aims at optimizing network performance and scalability. In this thesis, we have thoroughly explored the issue to propose new clustering algorithms suitable for the IoV system.

Firstly, we presented a deep literature review on clustering algorithms in VANETs. In this context, we introduced a new taxonomy to review and classify these algorithms, and discuss about their advantages and drawbacks. Moreover, a detailed comparison is provided for each classes of the proposed taxonomy considering relevant key parameters. The in-depth study of previous solutions has guided our work towards the concept of clustering to achieve of our goal. Secondly, we proposed three different clustering approaches suitable for the IoV network. The common objectives of these proposals are to optimize the total network overhead and improve the clusters stability, while guaranteeing a high level of service required by the function performed by these environments.

In our first contribution, we designed a new one-hop new Heuristic Clustering Algorithm based on RSU called HCAR for IoV that aims to improve the network performances. HCAR entails the centralization of a clustering algorithm at distributed RSUs. These latter are responsible for performing the cluster formation phase based on a simple heuristic algorithm, using graph theory concepts such as node degree and adjacency matrix. The suggested algorithm used a new method to recover the problem of the unavailability of CH, through the election of Secondary CH (SCH) using weighted mechanism. Moreover, this scheme takes care of the maintenance phase to keep clusters stability and structure.

In our second contribution, we proposed a new Multi-hop Clustering Approach over Vehicle-to-Internet (V2I) called MCA-V2I to extend the clusters coverage area and reduce the number of clusters. This proposal is based on a reasonable assumption that a vehicle can connect to the Internet via a special infrastructure called Road Side Unit Gateway (RSU-G), in order that each vehicle can obtain and share the necessary information about its Multi-hop neighbors to perform the clustering process. The clustering is performed using BFS algorithm for traversing graph based on a Mobility Rate (MR) that is calculated according to mobility metrics. MCA-V2I strengthens clusters stability through the election of Slave Cluster Head (SCH) in addition to the Master Cluster Head (MCH).

Finally, a new Efficient Weight based Clustering Algorithm using Mobility Report (WECA-MR) for IoV is presented, which aims to increase the stability of clusters and reduces the communication overhead. To elect the CHs, the designed algorithm used both classic weighted metrics, such as degree, average distance and a new metric introduced by the proposed approach called mobility report. It is a metric that combines mobility metrics: relative velocity and relative acceleration.

The experimental results prove demonstrate that the proposed approaches show better performances compared to other referenced algorithms in terms of different performance parameters, such as CH lifetime, CM lifetime, CH change rate, cluster number and clustering overhead, message delivery delay and message delivery ratio.

Future work and perspectives

Although the proposed clustering approaches for IoV are very promising and can provide very acceptable performance, there are other significant topics that we plan to carry out and study further in the research and applications of clustering approach in IoV. We summarize them as follows:

- Investigate the use of the proposed algorithms in urban areas by designing effective corresponding algorithms and comparing its performance to existing clustering protocols.
- Improve the proposed solutions by integrating/proposing a mechanism to improve the consensus rate between the vehicles during the CHs selection phase, especially in the distributed approaches.
- Consider the impact of collisions and hidden terminals, that have an impact on the delivery ratio when the number of connected vehicles increase.

List of publications

This thesis gave rise to several publications. Here is the list:

Articles published in indexed journals

1. [3] Oussama Senouci, Zibouda Aliouat, and Saad Harous. A review of routing protocols in Internet of vehicles and their challenges. Published in *Sensor Review (Emerald publisher)*, 39(1):58-70, 2019. **Impact Factor: 1.07**, SJR H-Index: 29 (Q2). DOI: `https://doi.org/10.1108/SR-08-2017-0168`.

2. [15] Oussama Senouci, Zibouda Aliouat, and Saad Harous. MCA-V2I: A Multi-hop Clustering Approach over Vehicle-to-Internet communication for improving VANETs performances. Published in *Future Generation Computer Systems (Elsevier publisher)*, 96:309-323, 2019. **Impact Factor: 4.639**, SJR H-Index: 85 (Q1). DOI: `https://doi.org/10.1016/j.future.2019.02.024`.

3. [14] Oussama Senouci, Saad Harous, and Zibouda Aliouat. A New Heuristic Clustering Algorithm Based on RSU for Internet of Vehicles. Accepted for publication at *Arabian Journal of Science and Engineering (Springer publisher)*. **Impact Factor: 1.09**, SJR H-Index: 27 (Q2).

Articles submitted for publication at indexed journals

1. [125] Oussama Senouci, Zibouda Aliouat, and Saad Harous. DCA-DS: A Distributed Clustering Algorithm Based on Dominated Set for Internet of Vehicles. Submitted paper.

2. [12] Oussama Senouci, Saad Harous, and Zibouda Aliouat. Survey on VANET

clustering algorithms: An overview, taxonomy, challenges and open research issues. Submitted paper.

Conference articles

1. [13] Oussama Senouci, Aliouat Zibouda, and Saad Harous. Survey: Routing proto- cols in vehicular ad hoc networks. In *Proceedings of the Second International Conference on Advanced Wireless Information, Data, and Communication Technologies*, AWICT '17, pages 8:1–8:6, Paris, France, November 2017, ACM. DOI: http://doi.acm.org/10.1145/3231830.3231838.

2. [16] Oussama Senouci, Saad Harous, and Zibouda Aliouat. An efficient Weight-Based clustering algorithm using mobility report for IoV. In *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE UEMCON '18, New York, USA, November 2018, IEEE. Being indexed in IEEE Xplore Digital Library.

Bibliography

[1] C. A. Kerrache, N. Lagraa, R. Hussain, S. H. Ahmed, A. Benslimane, C. T. Calafate, J. Cano, and A. M. Vegni. Tacashi: Trust-aware communication architecture for social internet of vehicles. *IEEE Internet of Things Journal*, 2018.

[2] A. Boukerche and R. E. De-Grande. Vehicular cloud computing: Architectures, applications, and mobility. *Computer Networks*, 135:171 – 189, 2018.

[3] O. Senouci, Z. Aliouat, and S. Harous. A review of routing protocols in internet of vehicles and their challenges. *Sensor Review*, 39(1):58–70, 2019.

[4] L. M. Ang, K. P. Seng, G. Ijemaru, and A. M. Zungeru. Deployment of iov for smart cities: Applications, architecture and challenges. *IEEE Access*, 2018.

[5] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan. A comprehensive survey on vehicular ad hoc network. *Journal of Network and Computer Applications*, 37:380 – 392, 2014.

[6] I. Salhi, M. O. Cherif, and S. M. Senouci. A new architecture for data collection in vehicular networks. In *2009 IEEE International Conference on Communications*, pages 1–6, June 2009.

[7] P. Yang, J. Wang, Y. Zhang, Z. Tang, and S. Song. Clustering algorithm in vanets: A survey. In *2015 IEEE 9th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, pages 166–170, Sep. 2015.

[8] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan. A comparative survey of vanet clustering techniques. *IEEE Communications Surveys Tutorials*, 19(1):657–681, Firstquarter 2017.

[9] Y. Huo, Y. Liu, L. Ma, X. Cheng, and T. Jing. An enhanced low overhead and stable clustering scheme for crossroads in vanets. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):74, Mar 2016.

[10] Y. Zheng, Y. Wu, Z. Xu, and X. Lin. A cluster–on–demand algorithm with load balancing for vanet. In Ching-Hsien Hsu, Shangguang Wang, Ao Zhou, and Ali Shawkat, editors, *Internet of Vehicles – Technologies and Services*, pages 120–127, Cham, 2016. Springer International Publishing.

[11] O. Abdel-Wahab, H. Otrok, and A. Mourad. Vanet qos-olsr: Qos-based clustering protocol for vehicular ad hoc networks. *Computer Communications*, 36(13):1422 – 1435, 2013.

[12] O. Senouci, S. Harous, and Z. Aliouat. Survey on vanet clustering algorithms: An overview, taxonomy, challenges and open research issues. *Submitted paper.*

[13] O. Senouci, Aliouat Zibouda, and Saad Harous. Survey: Routing protocols in vehicular ad hoc networks. In *Proceedings of the Second International Conference on Advanced Wireless Information, Data, and Communication Technologies*, AWICT 2017, pages 8:1–8:6, Paris, France, 2017. ACM.

[14] O. Senouci, S. Harous, and Z. Aliouat. A new heuristic clustering algorithm based on rsu for internet of vehicles. *Accepted for publication at Arabian Journal of Science and Engineering (Springer).*

[15] O. Senouci, Z. Aliouat, and S. Harous. Mca-v2i: A multi-hop clustering approach over vehicle-to-internet communication for improving vanets performances. *Future Generation Computer Systems*, 96:309 – 323, 2019.

[16] O. Senouci, S. Harous, and Z. Aliouat. An efficient Weight-Based clustering algorithm using mobility report for IoV. In *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (IEEE UEMCON 2018)*, New York, USA, November 2018.

[17] F. Yang, J. Li, T. Lei, and S. Wang. Architecture and key technologies for internet of vehicles: a survey. *Journal of Communications and Information Networks*, 2(2):1–17, Jun 2017.

[18] A. Dua, N. Kumar, and S. Bawa. A systematic review on routing protocols for vehicular ad hoc networks. *Vehicular Communications*, 1(1):33 – 52, 2014.

[19] F. Yang, S. Wang, J. Li, Z. Liu, and Q. Sun. An overview of internet of vehicles. *China Communications*, 11(10):1–15, Oct 2014.

[20] K. M. Alam, M. Saini, and A. E. Saddik. Toward social internet of vehicles: Concept, architecture, and applications. *IEEE Access*, 3:343–357, 2015.

[21] H. Hartenstein and K. Laberteaux. *VANET: Vehicular Applications and Inter-Networking Technologies*. Intelligent Transport Systems. Wiley, 2009.

[22] C. Jiacheng, Z. Haibo, Z. Ning, Y. Peng, G. Lin, and S. Xuemin. Software defined internet of vehicles: architecture, challenges and solutions. *Journal of Communications and Information Networks*, 1(1):14–26, Jun 2016.

[23] X. Li, Y. Feng, F. Wang, and Q. Qian. When smart phone meets vehicle: A new on-board unit scheme for vanets. In *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pages 1095–1100, Oct 2015.

[24] S. U. Bhover, A. Tugashetti, and P. Rashinkar. V2x communication protocol in vanet for co-operative intelligent transportation system. In *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pages 602–607, Feb 2017.

[25] Y. Park, C. Sur, S. Noh, and K. Rhee. Self-controllable secure location sharing for trajectory-based message delivery on cloud-assisted vanets. *Sensors*, 18(7), 2018.

[26] A. Soua and H. Afifi. Adaptive data collection protocol using reinforcement learning for vanets. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1040–1045, July 2013.

[27] I. A. Soomro, H. Hasbullah, and J. bin Ab Manan. User requirements model for vehicular ad hoc network applications. In *2010 International Symposium on Information Technology*, volume 2, pages 800–804, June 2010.

[28] S. Bitam, A. Mellouk, and S. Zeadally. Vanet-cloud: a generic cloud computing model for vehicular ad hoc networks. *IEEE Wireless Communications*, 22(1):96–102, February 2015.

[29] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya. A survey on vehicular cloud computing. *Journal of Network and Computer Applications*, 40:325 – 344, 2014.

[30] H. Kowshik, D. Caveney, and P. R. Kumar. Provable systemwide safety in intelligent intersections. *IEEE Transactions on Vehicular Technology*, 60(3):804–818, March 2011.

[31] A. Colombo and D. Del Vecchio. Supervisory control of differentially flat systems based on abstraction. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 6134–6139, Dec 2011.

[32] S. Sampath Kumar, D. Rajaguru, T. Vengattaraman, P. Dhavachelvan, A. Juanita Jesline, and J. Amudhavel. Intelligent collision avoidance approach in vanet using artificial bee colony algorithm. In L. Padma Suresh and Bijaya Ketan Panigrahi, editors, *Proceedings of the International Conference on Soft Computing Systems*, pages 545–551, New Delhi, 2016. Springer India.

[33] B. Zheng, M. O. Sayin, C. Lin, S. Shiraishi, and Q. Zhu. Timing and security analysis of vanet-based intelligent transportation systems. In *Proceedings of the 36th International Conference on Computer-Aided Design*, ICCAD '17, pages 984–991, Piscataway, NJ, USA, 2017. IEEE Press.

[34] A. H. Khosroshahi, P. Keshavarzi, Z. D. KoozehKanani, and J. Sobhi. Acquiring real time traffic information using vanet and dynamic route guidance. In *2011 IEEE 2nd International Conference on Computing, Control and Industrial Engineering*, volume 1, pages 9–13, Aug 2011.

[35] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil. Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions. *IEEE Communications Surveys Tutorials*, 13(4):584–616, Fourth 2011.

[36] E. Hossain, G. Chow, V. Leung, R. McLeod, J. Mišić, V. Wong, and O. Yang. Vehicular telematics over heterogeneous wireless networks: A survey. *Computer Communications*, 33(7):775 – 793, 2010.

[37] F. Sakiz and S. Sen. A survey of attacks and detection mechanisms on intelligent transportation systems: Vanets and iov. *Ad Hoc Networks*, 61:33 – 50, 2017.

[38] Y. Li. An overview of the dsrc/wave technology. In Xi Zhang and Daji Qiao, editors, *Quality, Reliability, Security and Robustness in Heterogeneous Networks*, pages 544–558, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[39] P. Tiwari and R. S. Kushwah. Traffic analysis for vanet using wave and wimax. In *2015 International Conference on Communication Networks (ICCN)*, pages 343–346, Nov 2015.

[40] H. Moustafa and Y. Zhang. *Vehicular Networks: Techniques, Standards, and Applications*. Auerbach Publications, Boston, MA, USA, 1st edition, 2009.

[41] D. G. Oh, P. Kim, J. J. Song, S. Jeon, and H. Lee. Design considerations of satellite-based vehicular broadband networks. *IEEE Wireless Communications*, 12(5):91–97, Oct 2005.

[42] M. M. Mohammed Nasr, A. M Abdelgader, Z. Wang, and L. Shen. Vanet clustering based routing protocol suitable for deserts. *Sensors*, 16(4), 2016.

[43] J. Cheng, J. Cheng, M. Zhou, F. Liu, S. Gao, and C. Liu. Routing in internet of vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2339–2352, Oct 2015.

[44] M. S. Rayeni and A. Hafid. Routing in heterogeneous vehicular networks using an adapted software defined networking approach. In *2018 Fifth International Conference on Software Defined Systems (SDS)*, pages 25–31, April 2018.

[45] Z. Cao, K. Shi, Q. Song, and J. Wang. Analysis of correlation between vehicle density and network congestion in vanets. In *2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC)*, pages 409–412, July 2017.

[46] R. S. Bali, N. Kumar, and J. Rodrigues. Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions. *Vehicular Communications*, 1(3):134 – 152, 2014.

[47] S. Hosmani and B. Mathpati. Survey on cluster based routing protocol in vanet. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pages 1–6, Dec 2017.

[48] R. Pal, A. Prakash, R. Tripathi, and D. Singh. Analytical model for clustered vehicular ad hoc network analysis. *ICT Express*, 4(3):160 – 164, 2018.

[49] W. Zhu, Q. Zhang, and A. C. M. Fong. Performance analysis of a hierarchical structured vanet. In *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pages 1352–1356, Aug 2013.

[50] A. Ephremides, J. E. Wieselthier, and D. J. Baker. A design concept for reliable mobile radio networks with frequency hopping signaling. *Proceedings of the IEEE*, 75(1):56–73, Jan 1987.

[51] C. R. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 15(7):1265–1275, Sep. 1997.

[52] P. Basu, N. Khan, and T. D. C. Little. A mobility based metric for clustering in mobile ad hoc networks. In *Proceedings 21st International Conference on Distributed Computing Systems Workshops*, pages 413–418, April 2001.

[53] M. Chatterjee, S. K. Das, and D. Turgut. Wca: A weighted clustering algorithm for mobile ad hoc networks. *Cluster Computing*, 5(2):193–204, Apr 2002.

[54] Y. Zhang and J. M. Ng. A distributed group mobility adaptive clustering algorithm for mobile ad hoc networks. In *2008 IEEE International Conference on Communications*, pages 3161–3165, May 2008.

[55] C. Konstantopoulos, D. Gavalas, and G. Pantziou. Clustering in mobile ad hoc networks through neighborhood stability-based mobility prediction. *Computer Networks*, 52(9):1797 – 1824, 2008.

[56] M. Sood and S. Kanwar. Clustering in manet and vanet: A survey. In *2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA)*, pages 375–380, April 2014.

[57] S. Mehta, P. Sharma, and K. Kotecha. A survey on various cluster head election algorithms for manet. In *2011 Nirma University International Conference on Engineering*, pages 1–6, Dec 2011.

[58] A. Maizate and N. El Kamoun. A new metric based cluster head selection technique for prolonged lifetime in wireless sensor networks. *International Journal on Communications Antenna and Propagation (IRECAP)*, 3(4), 2013.

[59] V. Nguyen, O. Kim, D. Dang, S. Kim, and C. Hong. Application of the lowest-id algorithm in cluster-based tdma system for vanets. In *2015 International Conference on Information Networking (ICOIN)*, pages 25–30, Jan 2015.

[60] W. Fan, Y. Shi, S. Chen, and L. Zou. A mobility metrics based dynamic clustering algorithm for vanets. volume 2011, pages 752–756, 01 2011.

[61] Y. Chen, M. Fang, S. Shi, W. Guo, and X. Zheng. Distributed multi-hop clustering algorithm for vanets based on neighborhood follow. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):1–12, 2015.

[62] A. K. Fauziyyah, S. Sulistyo, and I. W. Mustika. Performance analysis of cluster formation method in vehicular ad-hoc networks. In *2017 7th International Annual Engineering Seminar (InAES)*, pages 1–6, Aug 2017.

[63] J. Amudhavel, K. Prem Kumar, R. Sai Smrithi, S. Banumathi, D. Rajaguru, and T. Vengattaraman. Performance evaluation of dynamic clustering of vehicles in vanet: Challenges and solutions. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)*, ICARCSET '15, pages 58:1–58:4, New York, NY, USA, 2015. ACM.

[64] B. Guizani. *Algorithme de clusterisation et protocoles de routage dans les réseaux Ad hoc.* PhD thesis, Université de Technologie de Belfort-Montbéliard, 2012.

[65] Y. Zong, G. Xu, P. Jin, Y. Zhang, and E. Chen. Hc-ab: A new heuristic clustering algorithm based on approximate backbone. *Information Processing Letters*, 111(17):857 – 863, 2011.

[66] S. Cha, M. Ryu, K. Kim, and B. Jeon. Applying connected dominating set to broadcasting in vehicular ad hoc networks. In *2013 International Conference on Information Science and Applications (ICISA)*, pages 1–2, June 2013.

[67] Yan Shi, X Xu, C Lu, and Shanzhi Chen. Distributed and weighted clustering based on d-hop dominating set for vehicular networks. *KSII Transactions on Internet and Information Systems*, 10:1661–1678, 04 2016.

[68] M. A. Togou, A. Hafid, and L. Khoukhi. Scrp: Stable cds-based routing protocol for urban vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 17(5):1298–1307, May 2016.

[69] J. J. Kponyo, Y. Kuang, E. Zhang, and K. Domenic. Vanet cluster-on-demand minimum spanning tree (mst) prim clustering algorithm. In *2013 International Conference on Computational Problem-Solving (ICCP)*, pages 101–104, Oct 2013.

[70] K.G. Krishnakumar and E.J. Thomson Fredrik. Qos enabled data dissemination in hierarchical vanet using machine learning approach. *International Journal of Signal and Imaging Systems Engineering*, 10:231, 01 2017.

[71] M. Gerla and J. Tzu-Chieh Tsai. Multicluster, mobile, multimedia radio network. *Wireless Networks*, 1(3):255–265, Sep 1995.

[72] T. Sheu and Y. Lin. A cluster-based tdma system for inter-vehicle communications. *Journal of Information Science and Engineering*, 30:213–231, 01 2014.

[73] D. Gavalas, G. Pantziou, C. Konstantopoulos, and B. Mamalis. Lowest-id with adaptive id reassignment: A novel mobile ad-hoc networks clustering algorithm. volume 2006, page 5 pp., 02 2006.

[74] Z. Wang, L. Liu, M. Zhou, and N. Ansari. A position-based clustering technique for ad hoc intervehicle communication. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2):201–208, March 2008.

[75] H. Liu, L. Yang, Y. Zhang, and L. Wu. A position sensitive clustering algorithm for vanet. *International Journal of Online Engineering (iJOE)*, 10:14, 02 2014.

[76] Y. Allouche and M. Segal. Cluster-based beaconing process for vanet. *Vehicular Communications*, 2(2):80 – 94, 2015.

[77] N. Maslekar, J. Mouzna, H. Labiod, M. Devisetty, and M. Pai. Modified c-drive: Clustering based on direction in vehicular environment. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 845–850, June 2011.

[78] N. Maslekar, M. Boussedjra, J. Mouzna, and H. Labiod. A stable clustering algorithm for efficiency applications in vanets. In *2011 7th International Wireless Communications and Mobile Computing Conference*, pages 1188–1193, July 2011.

[79] K. A. Hafeez, L. Zhao, Z. Liao, and B. N. Ma. A fuzzy-logic-based cluster head selection algorithm in vanets. In *2012 IEEE International Conference on Communications (ICC)*, pages 203–207, June 2012.

[80] M. Ren, L. Khoukhi, H. Labiodt, J. Zhangt, and V. Veque. A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (vanets). *Vehicular Communications*, 9:233–241, 2017.

[81] M. M. C. Morales, C. S. Hong, and Y. Bang. An adaptable mobility-aware clustering algorithm in vehicular networks. In *2011 13th Asia-Pacific Network Operations and Management Symposium*, pages 1–6, Sep. 2011.

[82] E. Souza, I. Nikolaidis, and P. Gburzynski. A new aggregate local mobility (alm) clustering algorithm for vanets. In *2010 IEEE International Conference on Communications*, pages 1–5, May 2010.

[83] J. Xie and C. Li. Weight clustering based tdma-mac scheme in vanet. *Automatika*, 57, 07 2016.

[84] M. Hadded, P. Muhlethaler, A. Laouiti, and L. Azzouz Saidane. A novel angle-based clustering algorithm for vehicular ad hoc networks. In Anis Laouiti, Amir Qayyum, and Mohamad Naufal Mohamad Saad, editors, *Vehicular Ad-Hoc Networks for Smart Cities*, pages 27–38, Singapore, 2017. Springer Singapore.

[85] M. Hadded, R. Zagrouba, A. Laouiti, P. Muhlethaler, and L. A. Saidane. A multi-objective genetic algorithm-based adaptive weighted clustering protocol in vanet. In *2015 IEEE Congress on Evolutionary Computation (CEC)*, pages 994–1002, May 2015.

[86] M. F. Khan, F. Aadil, M. Maqsood, S. H. R. Bukhari, M. Hussain, and Y. Nam. Moth flame clustering algorithm for internet of vehicle (mfca-iov). *IEEE Access*, 7:11613–11629, 2019.

[87] F. Aadil, W. Ahsan, Z. Rehman, P. A. Shah, S. Rho, and I. Mehmood. Clustering algorithm for internet of vehicles (iov) based on dragonfly optimizer (cavdo). *The Journal of Supercomputing*, 74(9):4542–4567, Sep 2018.

[88] A. Farhan, T. Javid, and Z. Albert. Lica: Robust localization using cluster analysis to improve gps coordinates. In *Proceedings of the First ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, DIVANet '11, pages 39–46, New York, NY, USA, 2011. ACM.

[89] V. Sethi and N. Chand. A destination based routing protocol for context based clusters in vanet. *Communications and Network*, 9:179 – 191, 2017.

[90] K Aravindhan and C Suresh Gnana Dhas. Destination-aware context-based routing protocol with hybrid soft computing cluster algorithm for vanet. *Soft Computing*, 12 2018.

[91] D. Tian, Y. Wang, G. Lu, and G. Yu. A vanets routing algorithm based on euclidean distance clustering. In *2010 2nd International Conference on Future Computer and Communication*, volume 1, pages V1–183–V1–187, May 2010.

[92] A. Ziagham Ahwazi and M. NooriMehr. Mosic: Mobility-aware single-hop clustering scheme for vehicular ad hoc networks on highways. *International Journal of Advanced Computer Science and Applications*, 7(9), 2016.

[93] B. Liang and Z. J. Haas. Predictive distance-based mobility management for pcs networks. In *IEEE INFOCOM '99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and*

Communications Societies. The Future is Now (Cat. No.99CH36320), volume 3, pages 1377–1384 vol.3, March 1999.

[94] C. Caballero-Gil, P. Caballero-Gil, and J. Molina-Gil. Self-organized clustering architecture for vehicular ad hoc networks. *International Journal of Distributed Sensor Networks*, 11(8):1–12, 2015.

[95] R. T. Goonewardene, F. H. Ali, and E. Stipidis. Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks. *IET Intelligent Transport Systems*, 3(2):148–158, June 2009.

[96] Z. Zhang, A. Boukerche, and R. Pazzi. A novel multi-hop clustering scheme for vehicular ad-hoc networks. In *Proceedings of the 9th ACM International Symposium on Mobility Management and Wireless Access*, MobiWac '11, pages 19–26, New York, NY, USA, 2011. ACM.

[97] M. Azizian, S. Cherkaoui, and A. S. Hafid. A distributed d-hop cluster formation for vanet. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2016.

[98] S. Ucar, S. C. Ergen, and O. Ozkasap. Multihop-cluster-based ieee 802.11p and lte hybrid architecture for vanet safety message dissemination. *IEEE Transactions on Vehicular Technology*, 65(4):2621–2636, 2016.

[99] H. Wu, H. Tang, and L. Dong. A novel routing protocol based on mobile social networks and internet of vehicles. In Robert C.-H. Hsu and Shangguang Wang, editors, *Internet of Vehicles – Technologies and Services*, pages 1–10, Cham, 2014. Springer International Publishing.

[100] K. Mershad, H. Artail, and M. Gerla. Roamer: Roadside units as message routers in vanets. *Ad Hoc Networks*, 10(3):479 – 496, 2012.

[101] R. Aissaoui, A. Dhraief, A. Belghith, H. Menouar, H. Mathkour, and A. Abu-Dayya. Hcbls: A hierarchical cluster-based location service in urban environment. *Mobile Information Systems*, 2015:1–16, 2015.

[102] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, and B. Zhou. A data dissemination scheme based on clustering and probabilistic broadcasting in vanets. *Vehicular Communications*, 13:78 – 88, 2018.

[103] P. Li, T. Zhang, C. Huang, X. Chen, and B. Fu. Rsu-assisted geocast in vehicular ad hoc networks. *IEEE Wireless Communications*, 24(1):53–59, February 2017.

[104] R. Tomar, S. Verma, and G. Tomar. *Cluster Based RSU Centric Channel Access for VANETs*, pages 150–171. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[105] C. Wang, X. Li, F. Li, and H. Lu. A mobility clustering-based roadside units deployment for vanet. In *The 16th Asia-Pacific Network Operations and Management Symposium*, pages 1–6, Sept 2014.

[106] G. Staples. A new adjacency matrix for finite graphs. *Advances in Applied Clifford Algebras*, 38(3):979–991, 2008.

[107] W. Liang, Z. Li, H. Zhang, S. Wang, and R. Bie. Vehicular ad hoc networks: Architectures, research issues, methodologies, challenges, and trends. *International Journal of Distributed Sensor Networks*, 11(8):745303, 2015.

[108] C. Maglaras, A. Al-Bayatti, Y. He, and I. Wagner. Social internet of vehicles for smart cities. *Journal of Sensor and Actuator Networks*, 5(1):1–22, 2016.

[109] The network simulator NS-2. `http://nsnam.isi.edu/nsnam/index.php/Main_Page`, [Online]. Accessed: 2018-09-13.

[110] M. Fiore, J. Harri, F. Filali, and C. Bonnet. Vehicular mobility simulation for vanets. In *40th Annual Simulation Symposium (ANSS'07)*, pages 301–309, March 2007.

[111] A. M. Mezher, J. J. Oltra, L. U. Aguiar, C. I. Paredes, C. T. Barba, and M. A. Igartua. Realistic environment for vanet simulations to detect the presence of obstacles in vehicular ad hoc networks. In *Proceedings of the 11th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, 38; Ubiquitous Networks*, PE-WASUN '14, pages 77–84, New York, NY, USA, 2014. ACM.

[112] K. Medani, M. Aliouat, and Z. Aliouat. Fault tolerant time synchronization using offsets table robust broadcasting protocol for vehicular ad hoc. *AEU-International Journal of Electronics and Communications*, 81:192–204, 2017.

[113] L. Rivoirard, M. Wahl, P. Sondi, M. Berbineau, and D. Gruyer. CBL: a Clustering Scheme for VANETs. In *VEHICULAR 2017 - The Sixth International Conference on Advances in Vehicular Systems, Technologies and Applications*, pages pp.19–25, Nice, France, July 2017.

[114] P. Singh, R. Pal, and N. Gupta. Clustering based single-hop and multi-hop message dissemination evaluation under varying data rate in vehicular ad-hoc network. In Ramesh K. Choudhary, Jyotsna Kumar Mandal, Nitin Auluck, and H A Nagarajaram, editors, *Advanced Computing and Communication Technologies*, pages 359–367, Singapore, 2016. Springer Singapore.

[115] K. Abboud and W. Zhuang. Stochastic modeling of single-hop cluster stability in vehicular ad hoc networks. *IEEE Transactions on Vehicular Technology*, 65(1):226–240, 2016.

[116] S. Tayal and M. R. Triphathi. Vanet-challenges in selection of vehicular mobility model. In *2012 Second International Conference on Advanced Computing Communication Technologies*, pages 231–235, Jan 2012.

[117] B. Hassanabadi, C. Shea, L. Zhang, and S. Valaee. Clustering in vehicular ad hoc networks using affinity propagation. *Ad Hoc Networks*, 13:535 – 548, 2014.

[118] J. Chhugani, N. Satish, C. Kim, J. Sewall, and P. Dubey. Fast and efficient graph traversal algorithm for cpus: Maximizing single-node efficiency. In *2012 IEEE 26th International Parallel and Distributed Processing Symposium*, pages 378–389, May 2012.

[119] J. Lu, N. Xu, J. Yu, and T. Weng. Research of timing graph traversal algorithm in static timing analysis based on fpga. In *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, pages 334–338, Oct 2017.

[120] C. La Palombara, V. Tralli, B. M. Masini, and A. Conti. Relay-assisted diversity communications. *IEEE Transactions on Vehicular Technology*, 62(1):415–421, Jan 2013.

[121] G. Wu, P. Xu, W. Zhu, and T. Bui. Relay-assisted based af in two-hop vehicular networks over rayleigh fading channels. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–4, May 2016.

[122] J. Sucec and I. Marsic. Hierarchical routing overhead in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 3(1):46–56, Jan 2004.

[123] P. Papadimitratos, A. D. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza. Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation. *IEEE Communications Magazine*, 47(11):84–95, November 2009.

[124] A. Malathi and N. Sreenath. An efficient clustering algorithm for vanet. *International Journal of Applied Engineering Research*, 12(9):2000–2005, 2017.

[125] O. Senouci, Z. Aliouat, and S. Harous. Dca-ds: A distributed clustering algorithm based on dominated set for internet of vehicles. *Submitted paper.*

ملخص:

تتناول أطروحة الدكتوراه هذه بشكل أساسي مشكلة تحسين النفقات العامة للشبكات وتحسين استقرار الشبكة في بيئة إنترنت السيارات. في هذا السياق، نحن مهتمون بنهج التجميع لتلبية متطلبات هذه البيئة. وبالتالي، فإن الهدف العام من هذا العمل البحثي يركز على تصميم ونمذجة ومحاكاة خوارزميات تجميع جديدة لشبكات انترنت السيارات. في هذا الصدد، نبدأ بتحليل، من خلال مراجعة نقدية للأدبيات، لعدد من خوارزميات التجميع المقترحة لشبكات السيارات التقليدية، لاكتشاف بعض المشكلات التي لا تزال مفتوحة أو التي لا تزال حلولها المقترحة تحتاج إلى تحسين أكثر. بعد ذلك، نقترح ثلاث خوارزميات تجميع جديدة لشبكات انترنت السيارات، مما يجعل من الممكن تحسين أدائها، مع ضمان مستوى عالٍ من الخدمة التي تتطلبها الوظيفة التي تؤديها هذه الشبكات. تتمثل أهداف الخوارزميات المقترحة في التقليل إلى حد كبير من النفقات العامة للشبكة وتقليل زمن الوصول إلى الشبكة وزيادة نسبة التسليم بين عدد كبير من المركبات المتصلة التي تسير بسرعة عالية. لقد أسفرت الدراسة التحليلية والمحاكاة لتقييم الخوارزميات المقترحة، والتي تم إجراؤها بواسطة مجموعة من محاكي الشبكة ومولد الحركة، عن نتائج مقنعة، تفوقت على تلك التي عرضتها الخوارزميات الأساسية المقترحة في الأدبيات.
مفاتيح: شبكات المركبات، انترنت المركبات، التجميع، وحدة جانب الطريق، حركية.

Abstract :

This PhD thesis deals mainly with the problem of optimizing networking overhead and improving network stability in the Internet of Vehicles (IoV) environment. In this context, we are interested in the clustering approach to meet the requirements of such environment. Therefore, the general goal of this research work focuses on the design, modeling and simulation of new clustering algorithms for the IoV network. In this respect, we start with an analysis, through a critical review of the literature, a number of the clustering algorithms in VANETs, to detect certain problems still open or whose proposed solutions are still to improve. Then, we propose three new clustering algorithms for IoV networks, which make it possible to improve their performances, while guaranteeing a high level of service required by the function performed by these networks. The objectives of the proposed algorithms is to minimize significantly the network overhead, decrease the network latency and increase the delivery ratio between a large number of connected vehicles traveling at high speed. The analytical study and the simulation for evaluating the proposed algorithms, carried out by a combination of the network simulator NS2 and mobility generator VanetMobiSim, have yielded convincing results, outperforming those exhibited by the basic referred algorithms.
Keywords: VANET, Internet of Vehicles, Clustering, RSU, Mobility.

Résumé:

Cette thèse de doctorat porte principalement sur le problème de l'optimisation de la charge et de l'amélioration de la stabilité du réseau dans un environnement Internet des Véhicules (IoV). Dans ce contexte, nous sommes intéressés par l’approche de regroupement pour répondre aux exigences d’un tel environnement. Par conséquent, l'objectif général de ces travaux de recherche est de concevoir, modéliser et simuler de nouveaux algorithmes de regroupement pour le réseau IoV. À cet égard, nous commençons par analyser, à travers une revue critique de la littérature, un certain nombre d’algorithmes de regroupement dans des VANET, afin de détecter certains problèmes encore en suspens ou dont les solutions proposées doivent encore être améliorées. Ensuite, nous proposons trois nouveaux algorithmes de regroupement pour les réseaux IoV, qui permettent d’améliorer leurs performances, tout en garantissant un haut niveau de service requis par la fonction réalisée par ces réseaux. Les algorithmes proposés ont pour objectif de réduire de manière significative les frais généraux du réseau, de réduire la latence du réseau et d’augmenter le rapport de transmission entre un grand nombre de véhicules connectés se déplaçant à grande vitesse. L'étude analytique et la simulation d'évaluation des algorithmes proposés, réalisées à l'aide d'une combinaison du simulateur de réseau NS2 et du générateur de mobilité VanetMobiSim, ont donné des résultats convaincants, supérieurs à ceux des algorithmes de base cités.
Mots clés: VANET, Internet des Véhicules, Regroupement, RSU, Mobilité.