Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Babylon
College of Information Technology

A NEW CLUSTERING APPROACH FOR IMPROVING DATA DISSEMINATION IN VANET BASED ON MARKOV PROCESS

A Dissertation

Submitted to the Council of the College of Information Technology, University of Babylon in Partial Fulfillment of the Requirements for the Degree of Doctorate of Philosophy in Information Technology-Software

BY

SURA JASIM MOHAMMED ABD ALAMEER

SUPERVISED BY

PROF. DR. SAAD TALIB HASSON AL JEBORI

2022 A.D. **1443 A.H.**

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

﴿مَنْ عَمِلَ صَالِحًا مِّن ذَكَرٍ أَوْ أُنثَىٰ وَهُوَ مُؤْمِنٌ فَلَنُحْيِيَنَّهُ
حَيَاةً طَيِّبَةً ۖ وَلَنَجْزِيَنَّهُمْ أَجْرَهُم بِأَحْسَنِ مَا كَانُوا يَعْمَلُونَ﴾

النحل٩٧

صدق الله العلي العظيم

Supervisor Certification

I certify that the dissertation entitled (**A NEW CLUSTERING APPROACH FOR IMPROVING DATA DISSEMINATION IN VANET BASED ON MARKOV PROCESS**) was prepared under my supervision at the department of Software/ College of Information Technology/ University of Babylon as partial fulfillment of the requirements for the degree of Doctorate of Philosophy in Information Technology-Software.

Signature:

Supervisor Name: PROF. DR. SAAD TALIB HASSON

Date: / /2022

The Head of the Department Certification

In view of the available recommendations, I forward the dissertation entitled "**A NEW CLUSTERING APPROACH FOR IMPROVING DATA DISSEMINATION IN VANET BASED ON MARKOV PROCESS**" for debate by the examination committee.

Signature:

Assist. Prof. Dr. Ahmed Saleem Abbas

Head of Software Department

Date: / /2022

Certification of the Examination Committee

We hereby certify that we have studied the dissertation entitled (**A NEW CLUSTERING APPROACH FOR IMPROVING DATA DISSEMINATION IN VANET BASED ON MARKOV PROCESS**) presented by the student (**Sura Jasim Mohammed Abd Alameer**) and examined her in its content and what is related to it, and that, in our opinion, it is adequate with (**Excellent**) standing as a dissertation for the degree of Doctorate of Philosophy in Information Technology-Software.

Signature:
Name: Eng. Prof. Dr. Dhahir Abdulhade Abdullah
Title: Prof.
Date: / / 2022
(Chairman)

Signature:
Name: Dr. Naseer Ali Hussein
Title: Prof.
Date: / / 2022
(Member)

Signature:
Name: Dr. Soukaena Hassan Hashem
Title: Prof.
Date: / / 2022
(Member)

Signature:
Name: Dr. Mahdi S. Almhanna
Title: Assist Prof.
Date: / / 2022
(**Member**)

Signature:
Name: Dr. Raaid N. Alubady
Title: Assist Prof.
Date: / / 2022
(Member)

Signature:
Name: Prof. Dr. Saad Talib Hasoon
Title: Prof.
Date: / / 2022
(**Member and Supervisor**)

Approved by the Dean of the College of Information Technology, University of Babylon.

Signature:
Name: Dr. Hussein Attia Al-Khalidi
Title: Prof.
Date: / / 2022
(Dean of Collage of Information Technology)

Dedications

To MY MOTHER'S SOUL

To My Husband

,{ SALWAN KADHIM }

And To The Candles Glowing: My Sons

{ MUJTABA, MUAMAL And MURTAJA}

With All My Respect And Love…

Sura Jasim Mohammed

Acknowledgement

First and foremost, I would like to thank Allah, God, without divine care I could not even have contemplated all the work involved in this study. My dear mother, even though you left me and death took you from me, you will always be my support and encouragement, you are in my heart.. Thank you for your eternal support.

My thanks to my supervisor Dr. Saad Talib Hasson, for his help and support during the period of my study, he is the pillar of my work. I wish to express my love and gratitude to my beloved husband Salwan. My life is immeasurably richer in every way because of you. Words are inadequate to express my heartfelt thanks for your tremendous support and help. I feel motivated and encouraged whenever I speak and share my thoughts with you. Without your encouragement and support, this dissertation would not have materialized.

To my beloved children Mujtaba, Muamal and Murtaja. Thank you for your patience and endurance for the hardship of the road to success, my heartfelt thanks and love to all of you for your constant moral support and assistance. To my uncle, Kadhim your constant encouragement has made my journey towards my goal to obtain a Ph.D. degree a smooth sailing one.

Finally, but most importantly, I would like to thank all the lecturers in my Information Technology college who have imparted immense knowledge to me. My gratitude and thanks to the Dean Prof. Dr. Hussein Attiya Lafta you and all are highly appreciated.

Abstract

In the last few years, Vehicular Ad hoc Network (VANET) opens wide new research directions. It represents the core of the Intelligent Transportation System (ITS). VANET helps in providing users with safety and infotainment. However, there are many issues in VANET, such as efficient data dissemination, network scalability, and stable network topology. Therefore Firstly, to studying VANET's behaviour and evaluating its performance, the clustering technique must be the first step to controlling its changeable topology and building a meaningful group of vehicles.

A Highway scenario is assumed in this dissertation. In a proposed clustering approach, the moving vehicles are organized in non-overlapped clusters. The specific criteria are suggested to select the cluster head. Generally, there are two strategies to cluster formation: distributed (Ad hoc) and centralized. In this dissertation proposed system, the centre-based clustering strategy is selected to manage and control the clustering process with the assistance of the Road Side Units (RSUs).

Secondly , The most important issue is to predict the future movement to build a reliable system and enhance both the communication in VANET and the Intelligent Transportation System. The efficiency of the clustering algorithm depends on the mathematical approach called the stochastic process. The joining and leaving of vehicles to the clusters can be predicted mathematically using a Markov process. But in this study, practical implementation is applied to practically fit the mathematical approach of the Markov concept. The number of vehicles in each cluster was observed in different sequential times to observe the leaving and joining vehicles.

Finally, in VANET, Information can be spread among vehicles that contain communication equipment. Each vehicle plays the role of the sender, router, or receiver. Warning messages can propagate among vehicles. therefore, the process of modelling the data dissemination in VANET is also implemented in three types: broadcast, multicast, and unicast and evaluated based on the proposed clustering approach in different scenarios.

The experiments results show that the proposed clustering algorithm performance was better than the Centre-Based Secure clustering (CBSC) in term of clusters stability and number of created clusters, also the metrics refer to the improvement in results with increasing network size. In data dissemination models, the sparse scenario with unicast data dissemination achieves good performance metrics rather than broadcast and multicast. The proposed data dissemination model performance was better than Clustering and Probabilistic Broadcasting (CPB) algorithm in term of delay and packet delivery ratio.

Declaration Associated with this Thesis

I. **Published Paper 1**

S. J. Mohammed and S. T. Hasson, "**A developed Clustering Approach to Model the Data Dissemination Types in a Highway,**" in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, 2021, pp. 332–336, doi: 10.1109/BICITS51482.2021.9509917.

II. **Published Paper 2**

S. J. Mohammed and S. T. Hasson, "**Modeling and Simulation of Data Dissemination in VANET Based on a Clustering Approach,**" *2022 International Conference on Computer Science and Software Engineering (CSASE)*, 2022, pp. 54-59, doi: 10.1109/CSASE51777.2022.9759671.

III. **Published Paper 3**

S. J. Mohammed and S. T. Hasson (2022). " **A New Method For Cluster Stability Prediction In VANET Based on Stochastic process**". International Journal of Computer Information Systems and Industrial Management Applications. ISSN 2150-7988 Volume 14 (2022) pp. 010-018

Table of Contents

List of Figures

‘

List of Tables

List of Algorithms

List of Abbreviations

Abbreviation	Description
ABM	Agent Based Modeling
BFS	Breadth First Search
CBSC	Center-Based Secure and Stable Clustering algorithm
CG	Cluster Gateway
CH	Cluster Head
CM	Cluster Member
CMM	Continuous Markov Model
CPB	Clustering and Probabilistic Broadcasting
DBDC	Density-Based Dynamic Clustering
DSRC	Dedicated Short Range Communication
E2E	End to end Delay
GPS	Global Position System
HMI	Human-Machine Interface
I2I	Infrastructure to Infrastructure
IoV	Internet of Vehicles
ITS	Intelligent Transportation System
LET	Link Expiration Time
LTE	Long Term Evolution
MANET	Mobile Ad hoc Network
MCRE-DDP	Multi-Criteria-based Relay Election Protocol for Data Dissemination
NS2	Network Simulator 2
NUM_OF_CLU	Number of clusters
OBU	On-Board Unit
OMNET++	Objective Modular Network Testbed in C++

PDR	Packet Delivery Ratio
RSS	Received Signal Ratio
RSU	Road Side Unit
SNR	Signal To Noise Ratio
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
VANET	Vehicular Ad hoc Network

List of Symbols

Symbols	Description
${V_{asorting}}^{w}$	ascending sorted array
V_a^w	Array of vehicles speed
Δt_m	Time interval
A_k	Different in acceleration
chdur	Cluster head duration
cmdur	Cluster member duration
Cori-bri	Corresponding RSU
D_k	Difference in Distance
E_n	System state (the event at certain time)
$J_{backward}$	Winning vehicle due to its high speed
$J_{forward}$	Winning vehicle due to its low speed
$L_{backward}$	Losing vehicle due to its low speed
$L_{forward}$	Losing vehicle due to its high speed
M_k	Relatively Metric
N	Number of vehicles at time tm
$O(\Delta t_m)$	The error
$P_n (t_m)$	The probability of the cluster contains n vehicles at time t_m
$P_n (t_m, t_m+\Delta t_m)$	The probability of the cluster contains n vehicles at time $t_m +\Delta t_m$
A	Arrival
S	No change (static case)
${S_a}^w$	Average speed
stb	Stability counter
t_m	Current time
$t_m+\Delta t_m$	The next time
V_k	Difference in Velocity

X_t	Population size
λ_i	Birth event
μ_i	Death event

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Introduction

The smart city is a significant application due to the current and future development in Wireless communication technology (Hamdi *et al.*, 2020). The developments of the Internet of Vehicles (IoV) and Intelligent Transportation System (ITS) help in building the wanted smart cities (Hasson and Hasan, 2017). The ITS applications are divided into safety applications and non-safety. Safety applications improve and achieve safety for drivers and the roads environment, while the non-safety applications are to provide more services to passengers during their travel (Yeferny and Hamad, 2020).

Recently, a considerable emergence in various technologies has been witnessed. These emerging are covering several areas such as wireless communication, simulation, mobility machine learning tools, and sensing technologies. Various applications have found their way to be used and applied. However, one specific application is combining the recently emerging fields in intelligent transportation systems (Patel and Kaushik, 2018). It implies the existence of communication infrastructure on the road, the vehicles sensing power, and the processing units that can carry the algorithms. ITS provides a base network for communicating messages and operating numerous services. Vehicle routing is implemented to avoid traffic jams, message dissemination in emergency cases (Yeferny and Hamad, 2020).

ITS supports wireless communication by utilizing Vehicular Ad hoc Networks (VANET) (Hasan and Sarker, 2020). VANET can interact with

Internet of Vehicles (IoV) to create suitable and useful wireless vehicular network communication. Over the last years, VANET has been considered an important field in wireless vehicular network communication (Yeferny and Hamad, 2021). There are three familiar scenarios in simulating VANET environments; urban, rural, and highway (Abdel-Halim and Fahmy, 2018).

VANET, in its initial establishment, represents a self-organizing wireless network to connect vehicles without any infrastructure. It connects to other devices wirelessly using an On-Board Unit (OBU) to achieve low cost and availability. There are different categories of VANET: Vehicle to Infrastructure (V2I), Infrastructure to Vehicle (I2V), Vehicle to Vehicle (V2V), and Infrastructure to Infrastructure (I2I) (Lee and Atkison, 2021).

The dynamic mobility in VANET causes frequent changes in topology and link disconnection. The prediction of future movement in VANET is a very important issue due to the constraints of environments, traffic conditions, and urban layout. Therefore, prediction plays a crucial role in building an efficient VANET communication system. Most prediction studies which are about 65% were focused on urban scenarios and only 10% focused on Highways (Abdel-Halim and Fahmy, 2018).

Clustering represents a sub-field in wireless networks to segment the network into collections known as clusters by achieving maximum similarity within the cluster and minimum similarity between distinct clusters (Ramalingam and Thangarajan, 2020). VANETs clustering produces a network topology that offers one-hop and multi-hop communication (Senouci, Harous and Aliouat, 2020).

There are numerous challenges in the clustering field that have not been completely overcome yet. Most of them are related to the resulting overhead while others are associated with the stability of the approach and its robustness to the frequent topology changes (Lee and Atkison, 2021).

The clustering process contains three stages: cluster formation, Cluster Head (CH) selection, and cluster maintenance. Each cluster is managed by one selected vehicle called CH. The others vehicles in a cluster are called Cluster Members (CMs) (Kannekanti and Nunna, 2017).

In clustering studies, some develop the clustering process especially a cluster formation phase (Regin and Menakadevi, 2021), while others develop or improve the CH selection phase (Cheng and Huang, 2019) and it is the most common. There are two algorithms in the cluster formation stage: distributed strategy and centralized strategy, in a distributed strategy, the clusters are formed without any centralized units (Road Side Units (RSUs)) assistance, in a centralized strategy, the cluster formation is performed and controlled by centralized units distributed along the roads called RSUs (Saleem *et al.*, 2021).

The center-based strategy is representing a recently developed clustering approach, according to (Senouci, Harous and Aliouat, 2020), there are only 6% of recent related works used a center-based strategy while 94% used distributed strategy. The CHs selection depends on many parameters, the popular ones are position, velocity, and acceleration. The cluster stability (cluster maintains stage) plays a crucial role in network reliability. It is an important measure of the VANETs clustering efficiency (Almheiri and Alqamzi, 2015)(Abboud, Omar and Zhuang, 2016).

Moreover, Most VANETs clustering approaches focused on CH selection phase while the few others were addressing the cluster creation and cluster maintenance phases (Cooper *et al.*, 2016). Most of the proposed clustering approaches are not settled to convey the maintenance steps in a generic vehicular network.

Data dissemination represents an essential challenge in VANET. Data can be disseminated among vehicles to inform drivers about traffic, predicted weather conditions, and expected risks (Yeferny and Hamad, 2021), (Yeferny, Arour and Bouzeghoub, 2013).

This chapter is organized as follows: section 2 refers to the problem statement, section 3 contains the aim of the dissertation. Section 4 contains related works. Section 5 contains contributions and section 6 contains dissertation organization.

1.2 Problem statement

VANETs are characterized by their high mobility. Frequent changes in VANET topology cause continuous network communication failures. In such a dynamic environment, the creation and maintenance of a stable cluster are significant challenges (Kannekanti and Nunna, 2017).

One of the VANET environment is the highway, the towers in this type of environment is few, the distance among vehicles is far and the number of vehicles is little. For these reasons, the communication among vehicles and signals transmission from the beginning to the end faced many problems. To solve these problems we'll transfer data among vehicles in different data dissemination models based on the clustering approach.

Eventually, it can be concluded that building a stable cluster for the high mobility environment of VANETs is a big challenge, due to the frequent change in VANET topology. Hence, generated realistic traffic, as well as an appropriate clustering algorithm for high mobility of vehicles, is needed.

1.3 The Aim of the Dissertation

The aim of this research is to improve the data dissemination based on the clustering approach in VANET. Hence, in order to achieve the research aim, the following objectives are formulated:

1: To develop a cluster algorithm for selecting CH based on the relative speed, Mean and Median in Highway scenario with different parameters.

2: To design new method for data dissemination based on clustering approach to avoid the storm problem by using CH and RSU.

3: To develop a predictable mathematical model using Markov model and Birth-Death process to predicate the cluster’s stability.

4: To evaluate the performance of the proposed cluster approach in comparison with available solutions in a simulated network environment.

1.4 Related Works

In this section related works are divided into two categories: Clustering related works and Data dissemination-related works.

1.4.1 Clustering Related works

In (Zhang *et al.*, 2018), A Passive Multi-hop Clustering mechanism called PMC was proposed depending on the multi-hop. This method improved reliability and reduced the cost of clustering. They found increasing the cluster range increase the stability.

In (Alsuhli, Khattab and Fahmy, 2019), a new resilient algorithm was proposed for clustering. In this algorithm, they focused on clusters stability and decreased the number of clusters. Using a multi metrics to select primary and secondary CHs. These parameters were position, speed, and direction. Also the communication metrics like a Signal to Noise Ratio (SNR), Link Expiration Time (LET). The stability was enhanced by selecting two CHs, as well as it was improved the performance and stability.

In (Cheng and Huang, 2019), the researchers proposed a new clustering Center-Based Secure and Stable Clustering algorithm (CBSC) to decrease the vehicle motion differences. To select CHs firstly, they calculate multi metrics for each vehicle: summation of Euclidian distance among vehicles, summation of different velocities, and summation of difference acceleration. Then, put these metrics in one equation called relative mobility metric. The vehicle which has the smallest relative mobility metric was the more suitable to be CH.

In (Senouci, Harous and Aliouat, 2019), the researchers proposed a heuristic clustering method that depends on graph theory in the first phase of clustering (cluster formation) and a second phase (CH selection), there are primary and secondary CHs. Depending on nodes degree, neighbors, and adjacency matrix, they reduced the chance of CH unavailability. The proposed system achieves good performance metrics compared with the other methods.

In (Khayat *et al.*, 2020), the proposed clustering algorithm depended on a weighted formula that contains: trust, distance, and velocity in selecting CHs. The trust is a new factor which calculated for each vehicle and added to select suitable CHs.

In (Karthikeyan, 2021), the researcher proposed a clustering algorithm, called the adaptive clustering algorithm. As a first step, he selected RSUs as initial CHs due to the loss of data and communication disconnection. Then in the second step, They divided the network into static zones and selected CHs depending on positions, velocity, and buffer size. The final step contains using Nero Fuzzy Rules to predicate the CHs efficiency. The proposed system success in decreasing congestion, delay, and Packet Delivery Ratio (PDR).

In (Regin and Menakadevi, 2021), the researchers proposed a density-based dynamic clustering (DBDC). The algorithm clustered according to the dense regions in-network by finding the average vehicle density as a threshold using Received Signal Ratio (RSS). The proposed system improves stability and decreases congestion.

1.4.2 Data Dissemination Related Works

In (Shen *et al.*, 2014), a data dissemination strategy was proposed. The main challenge is to avoid collision and delay with maximum dissemination when the data was transmitted to the nodes. Their proposed system depends on a scheduling perspective to improve the data dissemination by sharing data with neighbor nodes. The authors used the average downloading delay as a performance metric to evaluate their study.

In (Li and Li, 2014), proposed a Multi-Criteria-based Relay Election Protocol for Data Dissemination in urban VANETs (MCRE-DDP). The relay node depends on several parameters in its selection, such as signal-to-noise ratio, speed, the distance between sender and receiver. The utilized performance metrics are redundancy ratio, link load, and speed.

In (Liu *et al.*, 2018), the data was disseminated based on clustering and probabilistic broadcasting. Firstly the clustering step is applied, then, the CM finds the forwarding probability depending on the value of a local-installed counter to guarantee the information coverage as well as packet delivery ratio.

Table 1.1 Related Works Summarization

Referen-ces	**Scena-rio**	**Algori-thm type**	**Techniques**	**Performance metrics**	**Simulator**
(Zhang *et al.*, 2018)	Rand-om	Distri-buted	Multi hop	Overhead, CH changing	NS2
(Alsuhli , Khattab and Fahmy, 2019)	Urban or High-way	Distri-buted	Select CHs according to vehicle speed, position, and direction,	Link expiration time (LET) The signal-to-noise ratio (SNR)	SUMO and NS3
Cheng and Huang, 2019)	High-way	Centra-lized	Combining velocity, acceleration and	CH lifetime, CM lifetime, Packet loss.	OMNET ++
(Senouc i,Harou s and Aliouat, 2019)	High-way	Centra-lized	Heuristic function in cluster formation	Overhead, CH lifetime, CM lifetime, CH change number	NS2
(Khayat *et al.*, 2020)	Rand-om	Distri-buted	The trust, the distance, and the velocity are used in CH selected	Delay	MATL-AB
(Karthik eyan, 2021)	Rand-om	Centra-lized	Divided roads to static regions then apply Neuro fuzzy prediction based CH selection.	End to end delay, Congestion	MATLA B
(Regin	Rand-	Distri-	density based	CH duration,	NS2

and Menaka devi, 2021)	om	buted	dynamic clustering (DBDC)	CM duration, Average delay, Overhead	
(Liu *et al.*, 2018)	High-way	Distri-buted	Data dissemination based on clustering with probability	Packet Delivery Ratio, Delay, Information coverage	NS2

In the proposed system, we differ from the related works in the CH selection method based on average speed and median concept which represent the key parameter in highways. Therefore the comparison with (Cheng and Huang, 2019) and (Liu et al., 2018) produces a good optimizer results. The most relate works are (Karthikeyan, 2021), (Alsuhli, Khattab and Fahmy, 2019), (Cheng and Huang, 2019), and (Liu et al., 2018).

1.5 The Contributions

In this dissertation, the simulation was applied to VANET Highway scenarios with different parameters such different number of vehicles, different cluster size , different acceleration and deceleration. The contributions in this dissertation abstract are as follows:

1: Proposed a clustering algorithm depending on the relative speed , Mean and Median in CHs selection as an essential parameter in a Highway scenarios.

2: Proposed a predictable mathematical model using the Continuous Markov Model (CMM) using the birth-death process to predicate the cluster's stability after a specific period of time. This process is achieved based on a derivative mathematical equation from the cases of vehicles leaving and joining the clusters.

3: Proposed a new method of Data dissemination in VANET based on the clustering approach. The proposed method avoids the storm problem, this can be done by disseminating messages among vehicles using CHs and RSUs. The proposed system excludes the rebroadcasting to send the message to the destination(s).

The developed clustering algorithm depends on RSUs in clusters formation and CH selection. The RSUs play the crucial role in management the clusters inside to controlling the data dissemination in each case.

1.6 Dissertation Organization

In addition to the first chapter, this dissertation contains the following chapters:

• **Chapter Two**: which is entitled Theoretical background. In this chapter, VANETs are clarified with challenges and applications. Clustering in the VANET approach is discussed with criteria and types. Prediction is explained especially prediction using Markov. As well, Data Dissemination in VANET is explained. The performance metrics are discussed in both clustering and data dissemination. Also, the benchmark is explained which is compared with the proposed clustering algorithm.

• **Chapter Three:** In this chapter, the proposed clustering algorithm is clarified as well as the method of applying the CBSC algorithm in the proposed environment. The prediction mathematical model is explained and its equations. The data dissemination in the three models is clarified in detail.

• **Chapter Four**: The results are discussions were presented in this chapter. The most effected performance metrics and evaluations and their relations were also presented and listed as graphs and tables. All the clustering and data dissemination calculations for each data dissemination model were listed and discussed. A comparison among many cases for different scenarios were also performed. Another comparison between proposed clustering algorithm and benchmark was also stated.

- **Chapter Five:** Conclusions And Future Works were presented in this chapter .

CHAPTER TWO

THEROTICAL BACKGROUND

2.1 Introduction

VANET helps in providing users with safety and infotainment. These goals require stable communication. However, the dynamic topology and frequent communication disconnection in VANET conflict with this requirement. The famous solution for this problem is clustering, there are many proposed algorithms in this field, The stability of the cluster represents the essential requirement in the clustering algorithm. On the other hand, Data dissemination represents the big challenge in VANET. In this chapter, all the previous concepts are clarified in detail. This chapter is organized as follows: Section 2 clarifies VANET and its applications. Section 3 refers to the clustering techniques. Section 4 explains prediction in VANET and Section 5 reviews data dissemination in VANET.

2.2 Vehicular Ad hoc Network

VANET is created spontaneously according to the Mobile Ad hoc Network (MANETs) principle. It represents the most technology that assists vehicles communication in many applications such as pedestrian, traffic, entertaining passengers, and supporting drivers. The purpose of VANET is to provide information to enhance safety. This information is about traffic, accident, and weather conditions. Vehicles can communicate together as Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), and Infrastructure to Infrastructure (I2I). The combination of V2V and V2I also called V2X. It is a main component of ITS

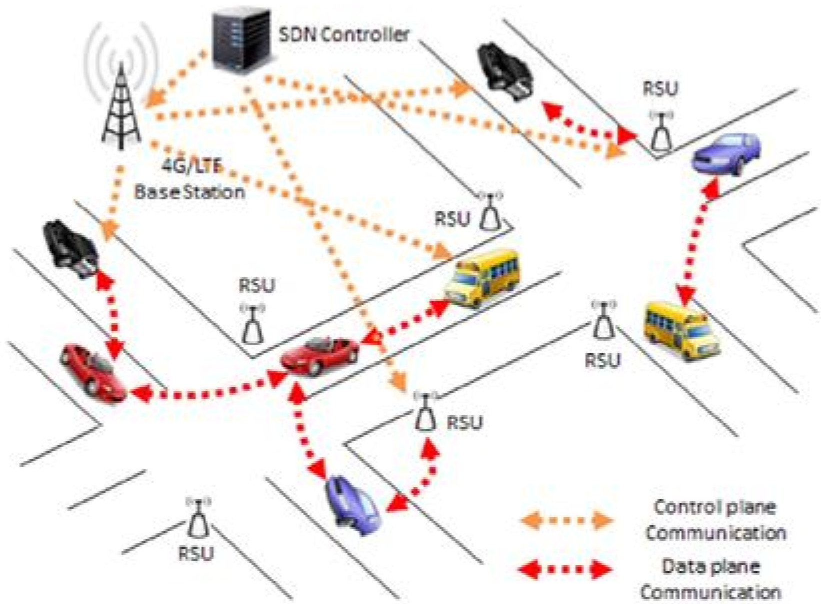
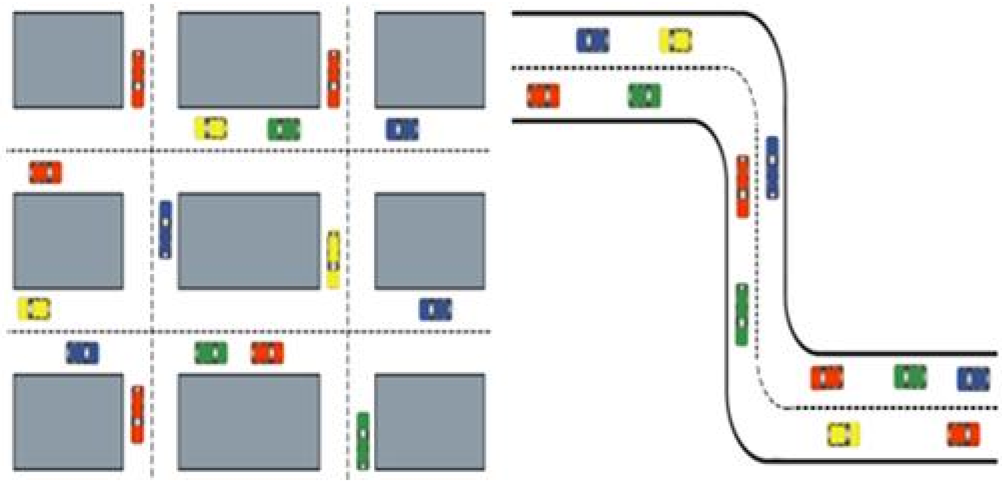


Figure 2.1 VANET overview (Shahid, 2019)

as shown in Figure 2.1. Previously, VANET was known as a car-to-car ad hoc in 2001, where the information can be transmitted among cars. From this time, the principle of the vehicle to vehicle and vehicle to RSU communication was proved that it will be important in the future in navigation, road safety, and other road services (Talib *et al.*, 2019). Vehicles are used by people daily intensively. There are lots of roads accidents caused by vehicles, which caused humans death, therefore, sharing information among vehicles is very important and supportive to ITS. This information is founded on two types: safety information and non-safety information (Saleh and Hasson Aljebori, 2019). The first one is important to inform the drivers about predictable accidents and avoid them. The second contains transmit general information about the trip or weather.

There are three environments in VANET (Najafzadeh, Ithnin and Abd Razak, 2014):



(a) Urban (b) Highway

Figure 2.2 VANET environments sample (Allani, 2018)

- **Rural Scenario**: this type is characterized by low traffic density and the disconnection of the network. The network in this scenario is divided into many distinct groups which cannot connect directly because they are far from each other. The communication of this type is done based on any near infrastructure (Najafzadeh, Ithnin and Abd Razak, 2014).
- **Urban Scenario**: in this type, there are lots of vehicles in the road segment which helps in the process of finding the path from source to destination. The speed is frequently changed between high and low according to road conditions as shown in Figure 2.2a (Najafzadeh, Ithnin and Abd Razak, 2014).
- **Highway Scenario**: in this type, the traffic style is different from the previous types, the speed is very high without traffic lights or crossovers. The density of vehicles is different according to time, for example, at the night, vehicles are sparse, unlike the rush hours.

Similarly, at weekends, during holidays, at matches, and on occasions, the traffic conditions are different as shown in Figure 2.2b (Najafzadeh, Ithnin and Abd Razak, 2014).

2.2.1 Communication in VANET

In VANET, there are many types of communications: Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), and Infrastructure to Infrastructure (I2I) which is called hybrid communication. The next subsections will give an overview of each type of them:

- **Vehicle to Vehicle**

This type of communication permits the vehicles to talk with others. It uses a 5.9 GHz frequency band using protocol IEEE 802.11p. This communication type is used for safety and data exchange. Data may be the vehicle's speed, position, number of persons in the vehicle, brake status, etc. V2V communications help a vehicle to sense the danger depending on the position of the other vehicles. This communication type is usually uses Dedicated Short Range Communication (DSRC) (Tanuja *et al.*, 2015). The advantage of V2V: short and medium communication range is allowed, it works without the need for RSUs, its cost is low, it is fast and reliable in providing passenger's and vehicle safety (Al-Sultan *et al.*, 2014).

- **Vehicle to Infrastructure**

In this type, vehicles can be connected with RSUs to exchange data and information. Vehicles connect to RSUs to extend their communication range. When the source and destination aren't in the same range, the source sends packets to RSU. V2I provides a wide range of communication. In V2I, the vehicle connects to RSU through On-Board Unit (OBU). It contains a radio transceiver (DSRC), GPS, application processor, memory unit to save information, sensors, and a Human-Machine Interface (HMI). OBU provides the connection from a vehicle to RSU and from a vehicle with its nearby vehicles. It also transfers messages regularly to other OBUs for supporting safety and applications among vehicles (Tanuja *et al.*, 2015).

- **Infrastructure to Infrastructure**

RSUs can communicate together to exchange the data and important information for the users. Each vehicle contains an application with GPS which helps in determining its position and the nearby RSU. Each RSU sends a HELLO message for getting confirmation. All the information is provided to the users (Shanmuga Priya.S, 2014).

2.2.2 VANET Applications

Over the last years, many researchers trying to improve and propose VANET applications. In VANET, the connection is very important to receive all the information about the weather, surrounding vehicle's status, etc. The latest VANET development represents online file sharing, real-time video updates, and entertainment via connection to the Internet through RSUs.

2.2.2.1 Safety Applications

These applications aim to prevent accidents, damage and save human lives on the road, also deliver safety-related information to the vehicle's divers. In this type, an alarm message was associated with road events to notify the driver. The information contains incident management and nearby vehicles (Rasheed *et al.*, 2017). Most of these applications are collision avoidance and incident management (Hamdi *et al.*, 2020).

2.2.2.2 Unsafety Applications (Infotainment)

This application provides passengers with suitable information and entertainment to achieve people's comfort and improve traffic efficiency. It provides the drivers with Internet connectivity, the location of shopping malls, restaurants, etc. (Ahmad *et al.*, 2018). This application contains two types: entertainment and background information (Hamdi *et al.*, 2020).

2.2.3 VANET Characteristics

VANET was characterized by the following features (Domingo and Reyes, 2012), (Jarupan and Ekici, 2011)

1) VANET is composed of a huge number of vehicles, particularly when the traffic intensity is high.
2) VANET's applications are aimed to save and improve ITS and future safe driving vehicles.

3) The multi-safety applications connected with VANET are trying to transfer the corresponding information about the traffic to all vehicles in the region of interest (Li, Zeng and Lou, 2011).
4) In VANET, there is a predefined network structure that allows vehicles locations to be determined. The connection among vehicles can be disconnected due to obstacles and other reasons on roads (Arshad *et al.*, 2018).
5) The dynamic topology of VANET is a distinctive trait, due to the high movement of vehicles. Therefore, the topology connections in fast movements are not highly secure (Hamdi *et al.*, 2020).
6) Vehicles are already having a power source. So, they have no problem with power (Hamdi *et al.*, 2020).

2.2.4 VANET Challenges

There are many considerations in VANET's design such as dynamic topology, high speed, driver behavior, and mobility constraints which represent the famous features of VANET. These features have an effective impact on networks efficiency, therefore, some of the challenges can be addressed (Misener, Biswas and Larson, 2011)(Fernandes and Nunes, 2007):

2.2.4.1 Node Velocity

In VANET, nodes either RSU's or vehicles, the first has zero velocity, the second can be reached 200 km/hour on Highway. The velocity will represent a crucial component of VANET because the connection in this case will have an exceptional challenge (Hamdi *et al.*, 2020). The coverage area of RSU's is about 1 kilometer, while the vehicles are forced to leave the coverage area in a few seconds due to their high velocities . The clustering approach is the appropriate solution for inflated velocity especially in highway scenarios (Bao *et al.*, 2020).

2.2.4.2 Movement Patterns

In VANET, there are three scenarios: Urban, Highway, and Rural. High mobility is essential for the nature of Highway roads. The vehicles move in two directions using predefined roads. The irregular changes always happen at road crossing points. In this challenge, there are many proposed protocols. In (Contreras and Gamess, 2020) the authors proposed a protocol used to select the best way for communication using a weighted graph. In (Soleymani *et al.*, 2020) the authors proposed a protocol used to count the vehicles in traffic lights and propagate BEACON messages from RSU to the vehicles at the end of the waiting line. The BEACON messages are periodically sent among vehicles and neighbors in (1-hop). Therefore, in the movement pattern, the speed, distance, reliability, and path must be described in the proposed protocol. Also, the spectrum sensitivity measure depends on the number of specific channels and consistency.

2.2.4.3 Node Density

Density is one of the important challenges in VANET. The number of vehicles may be increased from zero to hundreds or thousands in the mutual radio range. Many studies dealt with nodes density. In (Hamdi *et* *al.*, 2020) the authors depend on trust to measure credibility and collect information like the vehicle's position.

2.2.4.4 Scalability

Scalability means there is no performance degradation of a network despite the expansion in network size. Since hundreds or thousands of vehicles are deployed on roads, the designer should be careful about dealing with possibility of extending the network. However, this high density of vehicles has to be exploited to cover as large area as possible. (Tanuja *et al.*, 2015).

2.2.4.5 Frequent Disconnection:

The high speed in VANET causes frequently changes in its topology. In this case, RSUs as infrastructure are very important to assist the V2I and V2V communications (Tanuja *et al.*, 2015).

2.3 Clustering in VANET

However, there are many issues in VANET, such as efficient data dissemination, network scalability, and stable network topology. To study VANET's behavior and evaluate its performance, the clustering technique must be the first step to controlling its changeable topology and building a meaningful group of vehicles (Yeferny and Arour, 2012) (Talib *et al.*, 2020)**.**

In VANET, the process of grouping vehicles in a virtual separated group is called the clustering approach. It represents a controlling mechanism in VANET. Clustering represents an Unsupervised Machine Learning tool. At least there is one vehicle called CH in each cluster. Also, there are several vehicles in the cluster called CMs. The CH represents the leader of a cluster which joins CMs with the outside clusters (Ren *et al.*, 2021).

Also, in overlapped clusters, there is a Cluster Gateway (CG) which shares two or more clusters at the same time (Katiyar, Singh and Yadav, 2020).

The clustering process gives the network a hierarchical system. This system meets specific requirements, such as Quality of Service (QoS), scalability, load balancing, and network stability (Senouci, Harous and Aliouat, 2020). There are two cluster topologies, single-hop and multi-hop (Wang *et al.*, 2008)(Maslekar *et al.*, 2011).

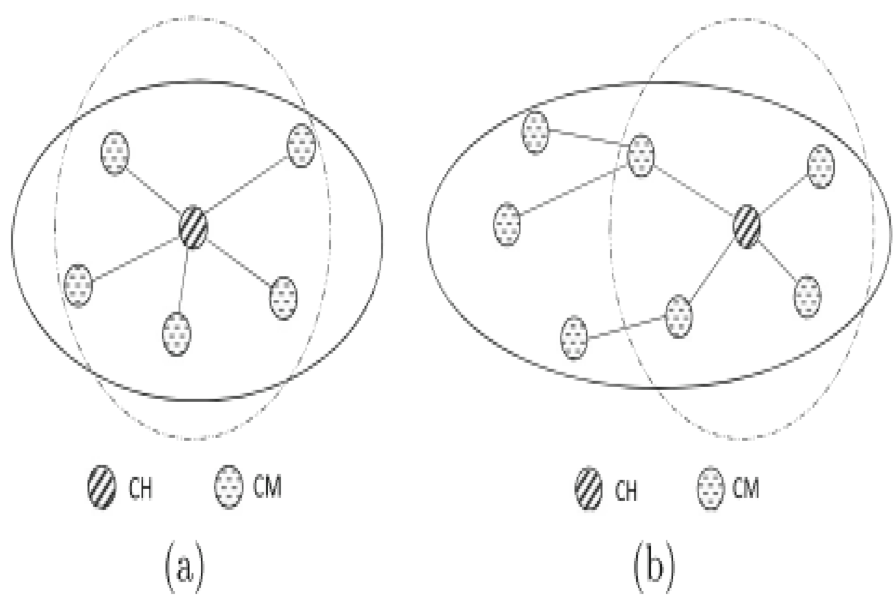
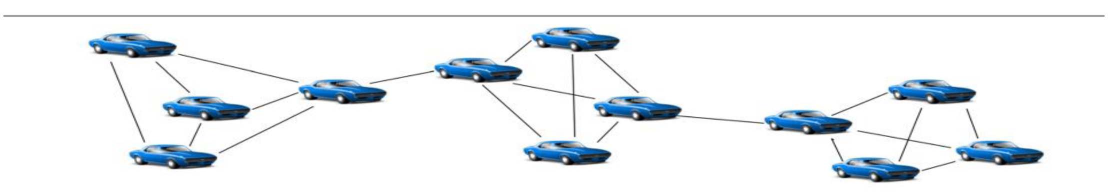


Figure 2.2: Clustering Topology
(a) single-hop clustering, (b) multi-hop clustering (Ren *et al.*, 2021)

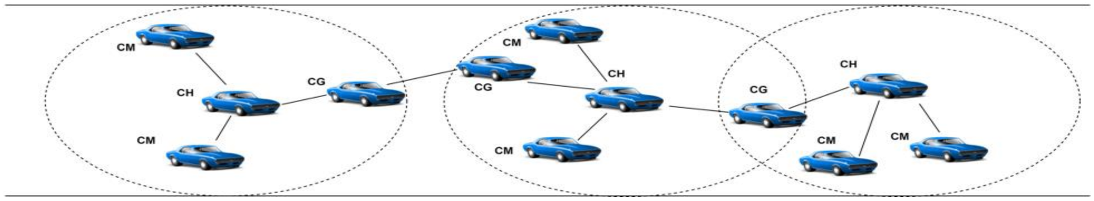
Most clustering algorithms start with the CH selection phase followed by the clusters formation phase. After the CH selection stage according to specific criteria as will be described later, the cluster forms by CH according to predefined criteria such as distance, cluster radius, transmission range, etc. In single-hop topology, cluster formation depends on the information of the vehicles which are at the one-proximity hop distance as shown in Figure 2.2(a).

In this topology, the cluster formation time and overhead are reduced. On the other hand, in dense networks, the single-hop topology causes collisions and a low Packet Delivery Ratio (PDR). When the network is sparse (low density), the probability of CH staying single is high, because no neighbor is affiliated with it. In multi-hop topology, cluster formation depends on the information of the vehicles which are at the K-proximity hop distance as shown in Figure 2.2 (b).

In the last years, multi-hop topology becomes more used in clustering algorithms, because it increases information transmission and decreases both disconnection and re-affiliation (Ren *et al.*, 2021).



(a)



(b)

Figure 2.3: VANET Scenarios
(a) Flat network, (b) clustered network (Katiyar, Singh and Yadav, 2020)

2.3.1 Clustering Role in VANET

Data dissemination protocols among vehicles require flood and relay data in the network. When data flooded to all neighbors, they received a broadcast packet from the source, then they rebroadcast these packets to their nearest neighbors. The iteration of rebroadcasting among vehicles in the dense flat network causes broadcast storm problem which as shown in Figure 2.3 (a) (Latif *et al.,* 2018). In the flooding approach, the chance of receiving data to the destination decreases with flat multi-hop networks (Latif *et al.*, 2018).

The probability of receiving data increases with delay and overhead increasing. A clustering approach is an appropriate solution for such a problem. In a multi-hop model, it is important to find the best route to send data from source to destination (Hadded *et al.*, 2017).

There are three types of routing protocols: Proactive, Reactive, and Geographical (Kumar and Dave, 2011)(Dua, Kumar and Bawa, 2014). All of them have limitations and drawbacks in VANET environments. Proactive is not suitable due to frequent topology changes of VANET.

Reactive has a high delay in network establishment. Geography can't work in a dynamic environment because it needs to detect the location of vehicles using GPS. The limitations of routing protocols increase the requirement to use a routing protocol that improves network reliability and scalability and decreases the delay. In routing protocols based on clustering, the network is divided into subgroups called clusters. The routing protocols which work inside the cluster are called intra-routing protocols (Paul and Islam, 2012).

The routing protocols which work among clusters are called inter-cluster routing protocols. In both types, CH takes all responsibilities of routing instead of all cluster's vehicles (Katiyar, Singh and Yadav, 2020).

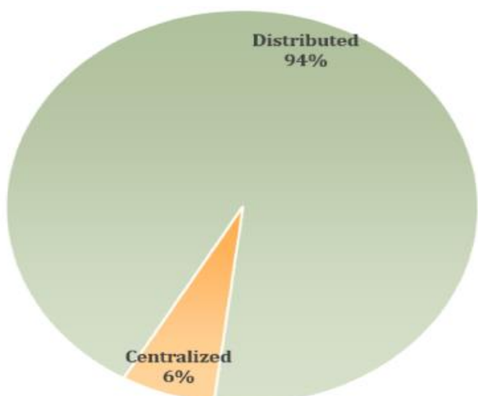


Figure 2.4 Cluster Formation methods ratio
(Senouci, Harous and Aliouat, 2020)

2.3.2 Cluster Formation Methods

According to (Senouci, Harous and Aliouat, 2020), there are two famous strategies in cluster formation: center-based cluster formation and distributed cluster formation. Figure 2.4 shows the ratio of studies that work in each of these two methods. The less percentage of the center-based is due to the reason that this approach is applied in the current years due to the technological developments.

2.3.2.1 Center-Based Cluster Formation

In this strategy, a central node takes the role of cluster formation and node's affiliation to suitable clusters. RSU is always used as a central node which takes the role of an administrator who forms clusters and informs each node about its affiliation. The modern trends in clusters formation algorithms depend on this type (Cheng and Huang, 2019), (Rossi *et al.*, 2017), (Talib *et al.*, 2020), and (Senouci, Harous and Aliouat, 2019). In the beginning, RSUs send a HELLO message to all nodes to inform the network availability, then, each node sends its request to the corresponding RSU, finally, RSUs send to each node its affiliation and role. According to (Abdel-Halim and Fahmy, 2018), there is only 6% of studies was used this strategy center-based cluster formation strategy will be considered and applied in this dissertation.

2.3.2.2 Distributed Cluster Formation

In this strategy (which is not based on RSUs), each node in the network, after receiving the information from its neighbors, seeks to find the CH then sends an affiliation message to this CH. Many studies used this strategy such (Azizian, Cherkaoui and Hafid, 2016), (Alsuhli, Khattab and Fahmy, 2019), (Chen *et al.*, 2015), and (Alioua *et al.*, 2017). There is 96% of studies was used this strategy as indicated in Figure 2.4.

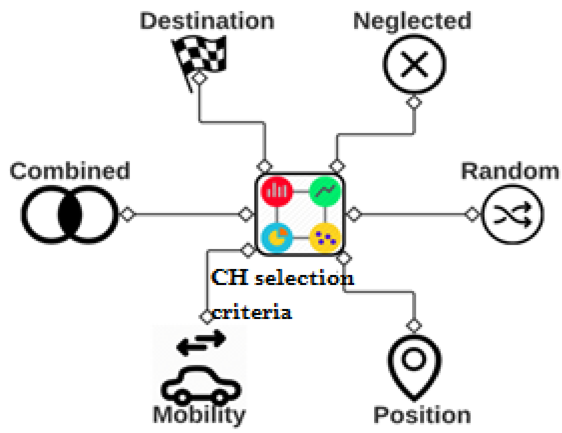


Figure 2.5 CH selection metrics (Senouci, Harous and Aliouat, 2020)

2.3.3 Cluster Head Selection Criteria

There are many criteria to select the CH. These criteria are shown in Figure 2.5. The following sections are describing these criteria :

1) **Neglected metric:** the CH selection in this type is done without any metric. It usually uses Heuristic algorithms in CH elections which have two goals: run time and finding the optimal solution (Maizate and El Kamoun, 2013).
2) **Random metrics**: this type contains early algorithms for MANET clustering. In this type, a random value is used as a metric for CH selection(Nguyen *et al.*, 2015).
3) **Position metrics**: These metrics depends on Euclidian distance and GPS.

4) **Mobility metrics:** this metric depends on the mobility parameters in selecting the CH such: as the signal, acceleration, velocity, and link stability (Fan *et al.*, 2011).
5) **Combined metrics:** In selecting the CH process, many metrics such as velocity, acceleration, and density can be collected in one weighted algorithm to select the set of CHs.
6) **Destination metrics:** the clustering algorithms can use information about mobility metrics related to the destination object like current location, speed, and etc. .

2.3.4 Center-Based and Secure Clustering Algorithm

A certain VANET clustering approach called Center Based and Secure Clustering algorithm (CBSC) algorithm is utilized as a benchmark in this dissertation (Cheng and Huang, 2019). CBSC is analyzed and stated in the following statements. At the beginning, each node sends its information to the near RSU, the information represents by vehicle's ID, acceleration, position, current speed and direction. RSUs take the role of clusters management in this approach. CH represents messenger CMs and RSU to exchange information. Each vehicle is able to communicate using DSRC and LTE. The vehicles communicate as intra-cluster using DSRC, while the vehicles communicate as inter-cluster using LTE. CBSC performs better cluster stability and packet delivery ratio compared to some other algorithms. The CBSC clustering algorithm stages are (Cheng and Huang, 2019):

A: Clusters formation step: in this step, the vehicles send a BEACON message to the nearest RSU. After that, the system analyzes the information of position to detect the center of the high density of vehicles. If the density is not high, the system applies an improved a-highest degree algorithm. this algorithm was developed to sure the distance between every two vehicles is larger than DSRC. The system draws dots in the map then applies the blob detection algorithm (Lindeberg, 1998). The last algorithm uses to detect the centers of the highest density then forms the clusters.

B: CH selection step: The CH selection is an important step in any clustering algorithm. In CBSC, the CH selection criteria are dependent. To elect CH which provides stability to the network, a relatively mobility metric ***M*** is calculated as follow: for a ***cluster*** that contains ***N*** vehicles, there is a vehicle ***k***, the difference in a location distance (D_k) between it and all *N* vehicles in *the cluster* can be calculated using Equation (2.1) (Cheng and Huang, 2019):

$$D_k = \sum_{n=1}^{N} \sqrt{(x_k - x_n)^2 + (y_k - y_n)^2} \quad \text{..........(2.1)}$$

The difference between velocity *v* of vehicle *k* and all other N vehicles in the same cluster known as (V_k) is calculated using equation (2.2) (Cheng and Huang, 2019):

$$V_k = \sum_{n=1}^{N} | v_k - v_n | \quad \text{................(2.2)}$$

The difference between acceleration *a* of vehicle *k* and all other N vehicles in the same cluster known as (A_k) is calculated as in equation (2.3) (Cheng and Huang, 2019):

$$A_k = \sum_{n=1}^{N} | a_k - a_n | \quad \text{................(2.3)}$$

The relative mobility metric M is calculated as in equation (2.4) (Cheng and Huang, 2019):

$$M_k = \propto \times \frac{D_k}{\max\{D_n|\forall n \in Ci\}} + \beta \times \frac{v_k}{max\{v_n|\forall n \in Ci\}} + \partial \times \frac{A_k}{\max\{A_n|\forall n \in Ci\}} \quad \ldots(2.4)$$

where α, β, and ∂ are weighted coefficients. The summation of $\propto$, β, ∂ must be 1. These coefficients are used to control the traffic, velocity, and acceleration parameters, for example, if the traffic condition is good, $\propto$ must be higher than the other coefficients. The smaller value of M indicates the vehicle which has small mobility changes compared with other vehicles in the cluster, it represents the most vehicle suitable to become CH in this algorithm approach.

2.4 Prediction in VANET

The nature of VANET such as its high mobility and large-scale network directly affects the dynamic change in deployment. This causes additional overhead and delays. The most important issue is to predict the future movement to build a reliable system and enhance both the communication in VANET and the ITS. The movement in VANET is more regular than in the Mobile Ad hoc Network (MANET), for example, a vehicle moves and saves its current speed with a certain time to reach the goal.

In VANET, to predict the location, each vehicle is provided with a Global Position System (GPS) to provide continuous future information about the location. However, the GPS information is not accurate in all conditions, as in covered streets it may not be as effective. Predicting the future deployment of VANET may depend on the vehicle's speed, direction, and current location. The clustering approach is an important step to deal with VANET to control its behavior and predict it. Therefore, the proposed system is applied to the clustered network to predicate the stability of clusters, to predict the network stability.

There are many studies about prediction in VANET, (Rashid *et al.*, 2020), the authors' studied prediction in VANET based on clustering to increase coverage area, and study link stability. Using prediction in VANET is performed by several methods, like a linear regressing model, mathematical statistics, neural network, historical trend extrapolation, Kalman filter, and the time series method (Zhang, Boukerche and Pazzi, 2011),(Ahmad *et al.*, 2018).

As for the approach adopted in this dissertation, the birth-death process is used to predict cluster stability in the future as one of the efficient tools in VANET prediction.

2.4.1 Stochastic Process

Mathematical models can be divided into probability models and deterministic models. Through this division, probabilistic models are more appropriate to represent most of the life phenomena. The use of probabilistic models is more used than their specific counterparts, as they are given in form of groups or families of random variables instead of a single value for the random variable. Random variables are grouped using a parameter such as time. These are known as stochastic processes (or random/chance processes (Ge, Gao and Quan, 2018).

The theory of stochastic processes plays an important role in the investigation of the stochastic phenomenon depending on time. The first results were achieved in this field with the research on Brownian motion, telephone calls, and traffic accidents. The basis for the mathematical theory of stochastic processes was given by A. N. Kolmogorov (1931) and since then the theory and practice of stochastic processes in it have not undergone significant development (Vvedenskaya *et al.*, 2018).

The birth-death process is a continuous Markov process in which only two ways are stated: birth, which increases the state variable by one, and death, which decreases the state by one. The name of the model is derived from a population representation where transitions are literal births and deaths. The process of birth and death can be used to observe the number of vehicles in a cluster. When birth occurs, the process goes from state n to state n+1. When a death occurs, the process goes from state n to state n-1 (Rasheed *et al.*, 2017)

2.4.1.1 Continuous Markov model (Birth-Death Process)

Markov chains and processes represent stochastic processes with special property. This special property states that the current value depends only on the previous value and the next future value can be estimated from the current value. A birth-death process (as a Markov process) is commonly used to analyze, describe and evaluate the process of estimating the changing in arriving or departing entities number through time (Steel, Hordijk and Kauffman, 2020).

A birth-death model represents a useful type of continuous-time Markov chain. This model has different applications in modeling most of the queuing problems. When the arrivals to any queuing system follow a Poisson process and its departure distribution (or its service times) follow an exponential distribution, then this system can be modeled and analyzed by utilizing the birth- death process. The important parameters in analyzing and modeling such systems are the creation of the steady-state probabilities for a general birth-death process (Tavaré, 2018).

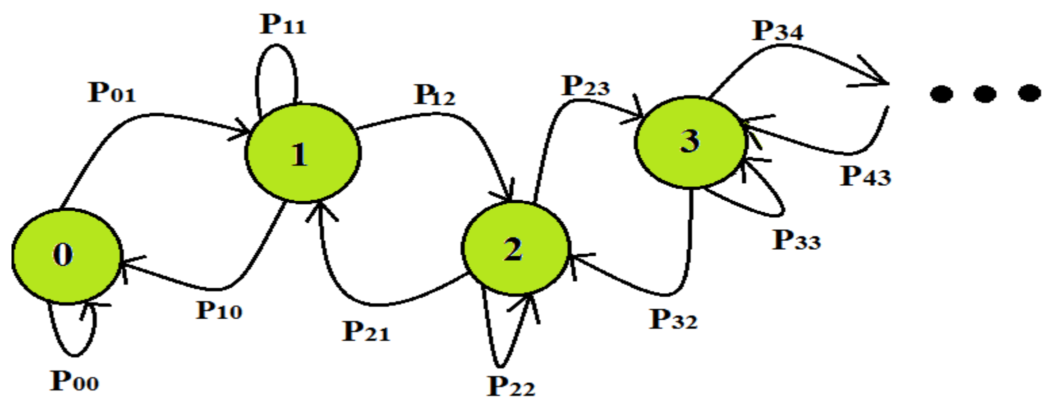


Figure 2.6 Birth-Death process

A continuous-time birth-death process is analyzed based on certain transition property. As shown in Figure 2.6, the probable transitions are to decrease the state by one, increase the state by one or to keep the state as it is without change. Decreasing the state is known as (death process), increasing the state is known as (birth process) and no change is known as no-birth and no-death process(Steel, Hordijk and Kauffman, 2020).

Definition 2.1 (Birth and death process) $\{N(t):t\geq 0\}$ is a birth and death process if there exist $\{\lambda_i\}_{i\geq 0}$, $\{\mu_i\}_{i\geq 0}$ with $\mu_0 = 0$ and μ_i,

$\lambda_i \geq 0$ such that:

$$\left.\begin{aligned} &P\,(N(t+\Delta t)=i+1\,|\,N(t)=i)=\lambda_i\,\Delta t+O(\Delta t)\\ &P\,(N(t+\Delta t)=i-1\,|\,N(t)=i)=\mu_i\,\Delta t+O(\Delta t)\\ &P\,(N(t+\Delta t)=i\,|\,N(t)=i)=1-(\lambda_i+\mu_i)\,\Delta t+O(\Delta t)\\ &P\,(N(t+\Delta t)=i+m\,|\,N(t)=i)=O(\Delta t)\ |m|>1 \end{aligned}\right\} \quad \ldots\ldots\ldots(2.5)$$

Equation (2.5) tells us that if the number of events happened in [0,t][0,t] is i, then how many events will occur in (t,t+Δt] (t,t+Δt]. If we denote to:

$p_j(t) = P(N(t) = j)$, and for $P(N(t + \Delta t) = j)$ we have

$P(N(t + \Delta t) = j) = \sum_i P(N(t + \Delta t) = j\,|\,N(t) = i)\,P(N(t) = i)$

$= P(N(t + \Delta t)=j \,|\, N(t) = j)\,P(N(t) = j) + P(N(t + \Delta t) = j \,|\, N(t) = j-1)\,P(N(t) = j-1) + P(N(t + \Delta t) = j \,|\, N(t) = j+1)\,P(N(t) = j+1) + \sum_{i\neq j,j-1,j+1} P(N(t + \Delta t) = j \,|\, N(t)=i)\,P(N(t) = i)$

It implies that

$$p_j(t+\Delta t)=[1-(\lambda_j+\mu_j)\,\Delta t]\,p_j(t) + [\lambda_{j-1}\Delta t]p_{j-1}(t)+[\mu_{j+1}\Delta t]p_{j+1}(t)+O(\Delta t) \ldots.(2.6)$$

If X_t represents the population size at time t, and the goal is to determine the population after Δt, it is not necessary to know the population in the past, it is sufficient to know the number at time t only. This process is an approximation of the Poisson process which is also a Markov process.

2.5 Data Dissemination

Lots of people lose their lives or infect with deformities due to accidents or sad sudden events on roads, as well as losing fuel and time in the traffic jam. Providing vehicles with timely and frequent information about traffic conditions and unexpected occurrences to dangerous events can improve human safety, ITS efficiency, reduce fuel loss and time (Sutariya and Pradhan, 2010). In VANET, Information can spread among vehicles that contain communication equipment's. Each vehicle plays the role of the sender, router, or receiver. Warning messages can propagate among vehicles to inform about traffic, weather conditions, unexpected events, etc. This process can be done through DSRC which has a low to medium range, which can connect V2V and V2I (Lochert *et al.*, 2007).

2.5.1 Data Dissemination Types

The data dissemination is the message exchange among vehicles and RSUs through V2V, V2I, and I2I. There are many types of data disseminations:

A) V2I/I2V Forwarding (Vehicle to Infrastructure or Road Side Unit)

In V2V, V2I, there are two types of data dissemination: push and pull. In Data dissemination which is based on push, the message is easily transferred from one vehicle to another or RSU. This type is more suitable in a dense network or announcements. In Data dissemination which is based on the pull, each vehicle can request information about the location. This type uses for inquiries about parking or nearby coffee shop etc. in other words the inquiry about information not related to users (Tomar, Chaurasia and Tomar, 2010).

B) V2V Forwarding (Vehicle to vehicle)

In Flooding, the information was made and received in the area. each part takes its role in data spread. This type is useful for low-traffic regions, unlike Relying, this type is more suitable for congestion dense networks (Rashid, Hamdi and Alani, 2020).

C) Geographical Forwarding

In VANET, the frequent changing in topology leads to the fact there is no final_shape in the network, therefore, Geographic dissemination uses to spread data to_the nearest point to the destination.

D) Opportunistic Forwarding

opportunistic routing (information-driven) is one of the data dissemination types. In this type, each vehicle has an ability to overhears the transmission and a neighbor vehicle (closest one) can perform the forwarding process. The other vehicles in the road will ignore and drop the packets. The essential benefits this dissemination type, are to combine many weak links into single strong link and to increases the throughput with improving in the quality of the links (Pophali and Yengantiwar, 2014).

E) Cluster-Based Forwarding

Cluster-based algorithms represent one of the most important applicable algorithms in VANET. It received more consideration by most of the VANET researchers. Cluster-based algorithms are always tries to improves the network performance or at least keep the performance of the VANET in a satisfactory level (Hasson and Abbas, 2021). This routing algorithm is suggested and developed in current and future research’s due to its features as scalable, efficient and distributed. The proposed approaches are focused on certain cluster formation phase, proposed algorithms to select the cluster heads and ensure the cluster stability by maintaining the clustering process phase.

2.5.2 Evaluation Metrics

The effective clustering algorithm is evaluated based on the stability of created clusters, especially the long time of the vehicle when it represents a CH or CM as a measurement called CH duration and CM duration. Also, the average number of clusters must be minimized. The cluster stability is heavily depends on the average number of CHs in each simulation run (Talib *et al.*, 2020).

Following are some of the important metrics that can be used in evaluating the stability of the VANET clusters (Rathore, Yadav and Chauhan, 2020)

1) The Clusters Efficiency

The efficiency of clusters can be defined as the percentage of vehicles participating in a clustering procedure during the simulation run (Ren *et al.*, 2017). It is calculated as shown in equation (2.7).

$$\text{Clusters Efficiency} = \frac{number\ of\ vehicles\ in\ cluster - \ number\ of\ out\text{-}layer}{Total\ number\ of\ vehicles} * 100 \ldots\ldots(2.7)$$

2) Average CH Duration

The average CH duration indicates the goodness of the clustering algorithm. It represents the average periods when the node stays as a CH before changing its state to CM. The goal is the highest average time to CH duration, which refers to the less changeable in CHs. It can be calculated as in Equation (2.8) (Ren *et al.*, 2017). .

$$\text{Average CH duration} = \frac{\sum CHs\ periods}{number\ of\ changing\ from\ any\ other\ state\ to\ CH} \quad \ldots\ldots..\ (2.8)$$

3) Average CM duration

This metric is similar to the previous one. It represents the average periods when the node stays as a CM before changing its state to another state. It can be calculated as shown in equation (2.9). This metric refers to the clustering approach stability. The goal is to maximize its duration (Ren *et al.*, 2017).

$$\text{Average CM duration} = \frac{\sum CMs\ periods}{number\ of\ changing\ from\ any\ other\ state\ to\ CM} \quad \ldots\ldots..\ (2.9)$$

4) Average Clusters number:

It represents the number of the total created clusters in each simulation run. In each simulation run different clusters are created initially and they can have merged or destroyed based on the nodes mobility and other road environments. The goal is to minimize it (Ren *et al.*, 2017).

There are many evaluation metrics can be used to evaluate the VANET behavior. Most of these metrics are:

5) **Throughput**: can be measured as average throughput, the total number of packets that transfer divided by total times. This can be estimated by equation (2.10) (Hasson and Abbas, 2021):

$$Throughput = \frac{(number\ of\ received\ packets \times\ size\ of\ packet\ )}{total\ time} \ldots\ldots.(2.10)$$

6) **Average End to End Delay (E2E)**: it is an important factor in data dissemination. It can be calculated by dividing the total time of transfer the message from source to another nodes until it reach to the destination by the summation of received packets. As shown in equation (2.11) (Hasson and Abbas, 2021):

Average End to End Delay =

$$\frac{\sum_{i=1}^{number\ of\ packets}(Time\ of\ recieve - Time\ of\ sending}{number\ of\ packets} \text{……………..(2.11)}$$

7) **Packet Loss**: it can be calculated by subtracting the summation of received packets from the total created packet as shown in equation (2.12).

Packet loss = no. of sended packets – no. of received packets….(2.12)

8) Packet Delivery Ratio (PDR): it can be calculated by dividing the summation of received message by the summation of created message as shown in Equation (2.13)

$$\textit{Packet Delivery Ratio (PDR)} = \frac{\sum number\ of\ recieved\ packets}{number\ of\ sending\ packets} * 100\text{….(2.13)}$$

2.6 VANET Simulation

There are many simulators used to simulate VANET, which provide lots of models, it supports applicability and scalability. SUMO (Simulator for Urban Mobility) is one of the famous VANET simulators, which uses to model traffic mobility, especially in urban. Network Simulator 2 (NS2), Network Simulator 3 (NS3), and Objective Modular Network Testbed in

C++ (OMNeT++)are used in VANET simulation but they aren't used in road traffic simulation. Netlogo is a brilliant tool to simulate agent behavior. Several studies have used it to simulate different works related to vehicular networks. Designing an agent based modeling requires defining three main components agents, environment, and interactions. This process is not all. Describing agent-agent and agent-environment interactions is another need designer should carefully take into account. Besides, there are also two additional components. The first one is Observer who is responsible to ask agents what they should do. The second component is Schedule which is related to what the observer should use to ask agents. In this context, Netlogo provides two buttons in its interface SETUP and GO. The Designer clicks SETUP and then GO to schedule these actions (Ramli and Rawi, 2020).

2.6.1 Netlogo Simulation

Agent Based Modeling is a general framework to model the dynamic systems. In which, many agents are interacting with each other to analyze their behavior. As VANET are systems that contain many vehicles (agents) that behave in self-organized manner, Agent Based Modeling can model such these systems no matter how difficult they are. (Hiroki Sayama, 2015) has defined Agent Based Modeling as “computational simulation models that involve many discrete agents.” In the context of vehicular networks, Agent Based Modeling can be simply defined as a computational model that uses many discrete vehicles. According to this definition, there are two keywords are needed to explain. The first one is “computational.” Indeed Agent Based Modeling is not a mainly mathematical model but a computer model. It means that Agent Based Modeling uses computer to model agents’ behavior instead of mathematical ways. This is strongly advised especially in social sciences as Hiroki claims. Actually agreement with this opinion is fair enough. The second keyword is “discrete” which means that each vehicle has clear parameters that specify the relationship between itself and its environment. Not this only, each vehicle will take an appropriate action basing on its status and the status of surrounding environment. It has been argued that there are certain features should be taken in account when Agent Based Modeling is designed (Ramli and Rawi, 2020).

CHAPTER THREE

THE PROPOSED SYSTEM

3.1 Introduction

VANET is one of the important solutions to Improve the ITS. There are two main challenges in VANET, firstly, the topology of VANET is rapidly changing, secondly, the frequent communication disconnected. To control these challenges, the clustering approach is an appropriate solution. On the other hand, using a data dissemination approach to disseminate timely warning messages in different methods can help in avoiding many accidents and save human lives.

In this dissertation experiments, data dissemination models are applied in VANET based on an improved stable clustering approach. Simulation is a very important tool due to the cost and difficulties of real application in highway scenarios, therefore Netlogo 6.1.1 (as an agent based modeling approach) has been used in implementing the simulation scenarios.

3.2 The Proposed System

The main stages of the proposed system are represented by: designing and building Highway scenarios with different parameters (number of nodes, transmission range, data dissemination type, and acceleration / deceleration of vehicles) using a networks simulator (Netlogo).

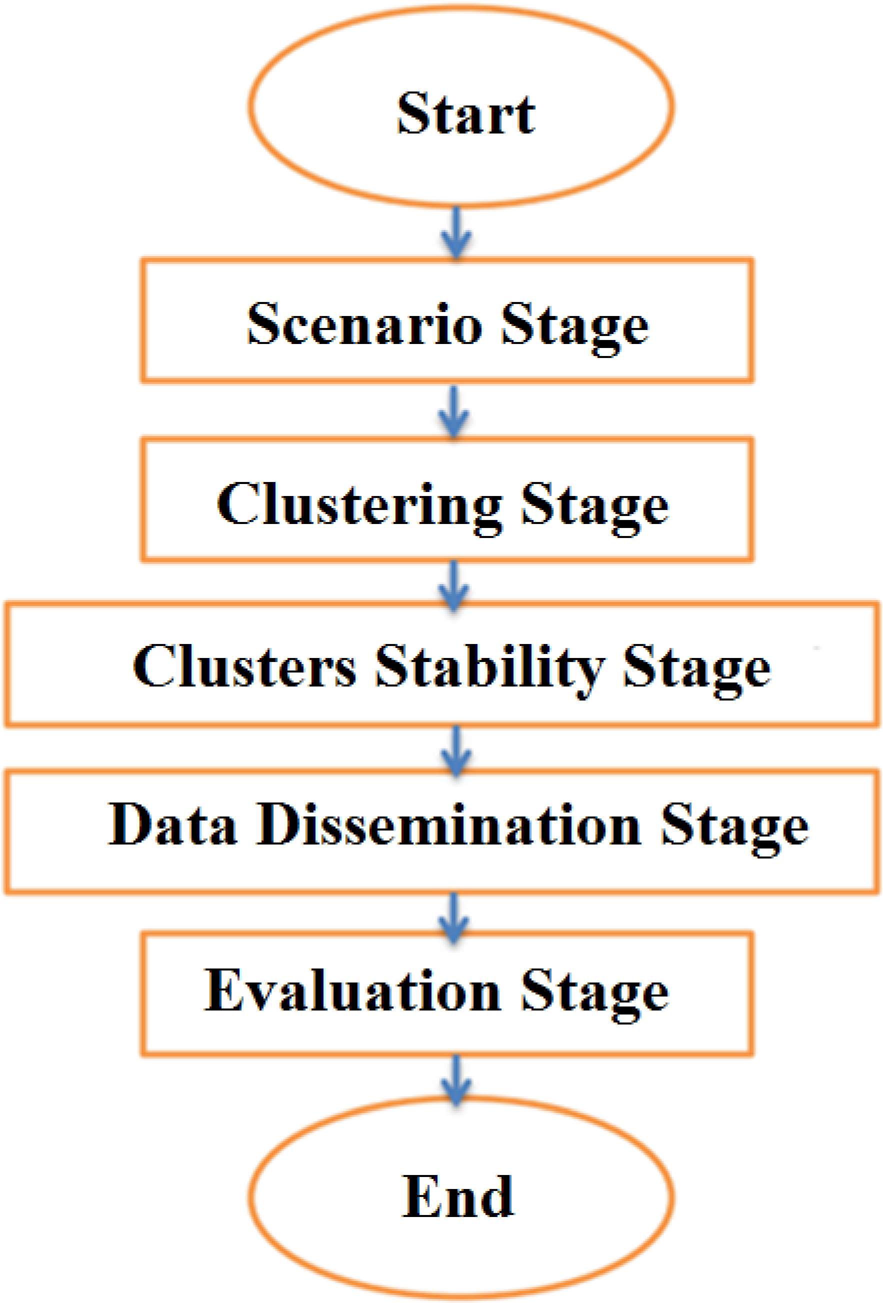


Figure 3.1 The Proposed system

Figure 3.1 refers to the dissertation proposed system. This dissertation applies a developed center-based clustering approach in its main proposed approach depending on a developed center-based clustering approach. The evaluation process of the CH selection method is attempted by comparing it with another known clustering algorithm (CBSC) as a benchmark. The stability prediction of the created clusters.

is analytically proposed based on the Continuous Markov Model (CMM) as a developed model by utilizing the Birth Death Process. Finally, the process of Modeling the data dissemination in VANET is also implemented and evaluated based on the proposed clustering approach in different scenarios.

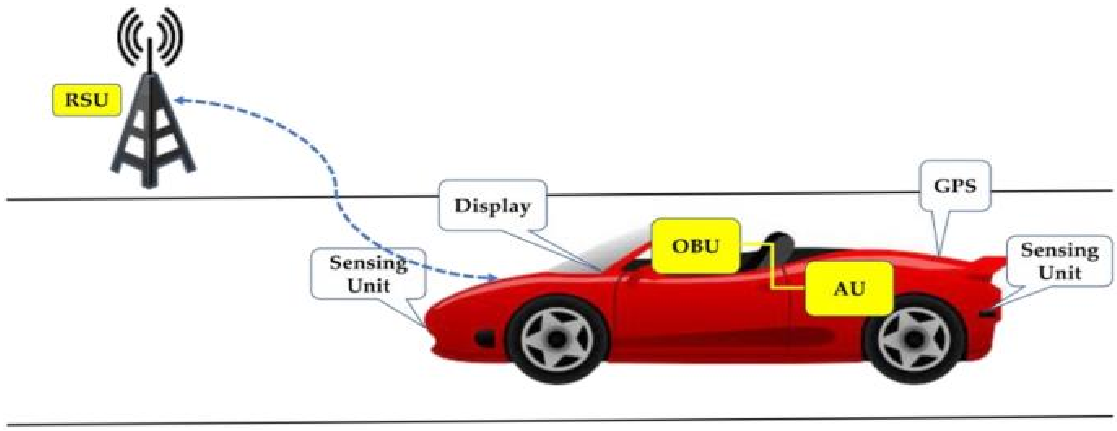


Figure 3.2 Vehicle communication components

3.3 Network Assumptions

As a network model, the following assumptions represent the most significant issues followed in simulation environment. The clustering approach is proposed and applied in different number of vehicles. Figure 3.2 assume the main communication's components in each vehicle.

1. The proposed model is a center-based clustering approach with the assistance of the distributed Road Side Units along the road.

2. Each vehicle contains On-Board Unit (OBU) to be used it in V2V communication. This becomes essential for most Intelligent Transportation System application to offer low cost and availability. The vehicles can also communicate with the RSUs using V2I communication, in addition to the additional facility of I2I when needed.
3. The developed clustering approach is to create non overlapping clusters.
4. Each vehicle, in simulation environment, has an internal parameters such as a unique identifier called (ID), (speed) which refers to the current speed, (xcor and ycor) which refer to the current x coordinate and y coordinate respectively.

3.4 Methodology

The following sections represent the main steps of the proposed system in more details followed in these experiments. Figure 3.3 clarifies:

Preparing the Environment

- Create Road with 3-lanes
- Setup Roadside Units
- Setup vehicles using 4- cases of number of vehicles
- Generate an acceptable random approach to move the vehicles in one directions

Proposed Models

- Suggest and set the developed center-based clustering approach.
- Impliment the CBSC approach.
- Impliment the proposed clustering approach.
- Perform the CH selection and the cluster formation.
- propose the cluster stability predication model using Cotinious Markov model and birth-death process.
- Model the message exchange depending on V2V, V2I and I2I in the suggested clustering approach environment.
- Model the data dissemination as one to one (unicast), one to group (multicast) and one to all (broadcast).

Results Discussions

- Estimate the clustering performance metrics for the CBSC.
- Estimate the significant performance metrics to evaluate the suggested center-based clustering approach.
- Present the comparative results graphically to observe their behaviors.
- Compare and discuss the approach results with CBSC results.
- Estimate the results of the stability prediction Model.
- Calculate the performance metrics of the Data dissemination models
- Results evaluation, discussion and Conclusion.

Figure 3.3 The Proposed system steps

3.4.1 Setup Stage

In this stage, many steps are performed to prepare and set the requirements for the VANET simulation environment.

a. Lanes creation

The environment setup starts with creating a suggested road segment with 10 km length . The road segment is designed and divided into 3-lanes using Netlogo 6.1.1 as an agent based modeling approach. The RSUs are also created and located along the road segment (1 km distance is suggested among them). Therefore, the Equation of RSU's number calculation is suggested according to equation (3.1):

Number of RSU = (L / RSU Coverage Distance)+1………..(3.1)

Where L represents the length of road segment, The RSUs are constructed along the road segment to control the clustering and the data dissemination processes in the suggested environment.

b. Create and Move the Vehicles

In the simulated environment, the vehicles are represented and generated on the road. The entrance of the vehicles to the road segment is also created based on poison random probability distribution with three different mean arrival rate. The suggest three arrival rates are 10 vehicles/hour (for sparse traffic case), 50 vehicles/hour (for medium traffic case), and 100 vehicles/hour (for dense traffic case). The Arrival rate is representing the number of vehicles passing through certain point on the road per unit time. The inter arrival time can be created as the average periods among the arriving vehicles. This inter arrival time is estimated in minutes or second as in Equation 3.2:

$$Inter\ arrival\ time = (1 / Arrival\ rate) * 60 \quad \ldots\ldots\ldots\ldots\ldots (3.2)$$

Vehicles speed is generated randomly based on Uniform probability distribution between ((min and max speed (10- 200 km/hr. in this dissertation) for the vehicles. Vehicles acceleration and deceleration (based on general specifications, are also created based on the uniform distribution (acceleration (0- Max acceleration, and deceleration (0- max deceleration)). Algorithm 3.1 shows these stage main steps.

Algorithm 3. 1 Moving-vehicles

Input: speed, max-speed, min-speed, acceleration, deceleration
Output: Moving vehicles in a road segment

```
1: Vehicles generation on road
2: For each vehicle in road segment
3:     Indicate and save its vehicle ID
4:     For vehicle_i ∈ set of vehicles
5:         If there is other vehicle in front of vehicle_i then
6:             Indicate and save the ID of this front vehicle
7:             speed ← speed of front vehicle – ( deceleration * time)
8:         else
9:             speed← speed + (acceleration * time)
10:        end if
11:    end for
12:end for
13: End
```

3.4.2 Proposed Clustering Approach

In this dissertation, a Highway scenario is assumed in experiments. In a proposed clustering approach, the moving vehicles are organized in non-overlapped clusters. The specific criteria are suggested to select the cluster head. Each vehicle (CM) is sent to (or receive from) its CH only. The message is transferred through the cluster heads to the close RSU that covers this area according to Long Term Evolution (LTE). The communication inside a cluster and between cluster heads are based on DSRC. Each vehicle is equipped with IEEE 802.11 to communicate via LTE and DSRC. Vehicles are also provided with Global Position System (GPS) to provide their locations. The range of LTE is 1.0 Km, and DSRC is 300 M. Broadcast, multicast and unicast are utilized in this dissertation scenarios as three approaches for vehicles data dissemination.

In the broadcast, the source vehicle floods all the vehicles in the network with packets. In the multicast, the source vehicle sends the packets to the specific cluster heads. In unicast, the source vehicle sends the packet to a specific vehicle.

In all the previous traffic types, the source vehicle sends its packet to the adjacent cluster head, then the last one takes the role of sending the packet to the adjacent RSU which is forwarded it to the suitable destination according to the dissemination type.

A random selection of the source and the destination(s) is setting in this dissertation scenarios. The proposed clustering approach assumes that all vehicles driving in one direction (side) on the highway. Each vehicle has the same transmitting capability since they have an equal chance to be

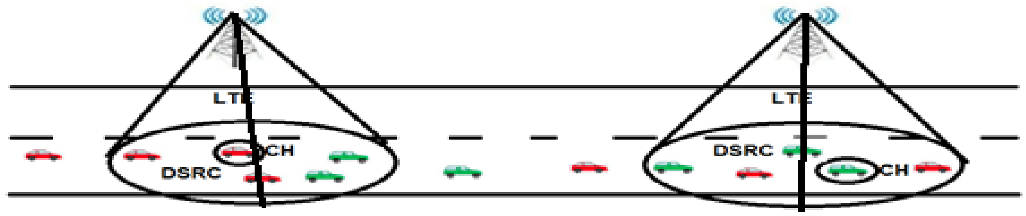


Figure 3.4 Cooperation between RSUs and Vehicles

elected as CHs. The cooperation among the clustered vehicles and the RSUs are shown in Figure 3.4.

Algorithm 3.2 Cluster Formation

```
Input: All Vehicles in road segment V //   all vehicles are founded in V
Output: sets of clusters C1, C2…
P ← V; where P set of all vehicles
While P ≠ ϕ
        p← one vehicle from P          // where p ∈ P
        Cp← ϕ                          //  Initialize node set Cp = ϕ;
        Ve←V
        For each vehicle e ∈ Ve do;    // where e ∈ Ve
            if  Distance (e,p)<= DSRC  then
                Add e to Cp            // Save e to set of cluster
                Remove e  from Ve      // To avoid reprocess it
            else
                Continue
            end if
        end for
        Remove p from P                // To avoid reprocess it
end while
Remove duplicated vehicles from all clusters
end
```

3.4.2.1 Cluster Formation

The set of the collected vehicles in one separated group is called a cluster. Generally, there are two strategies to cluster formation: distributed (Ad hoc) and centralized. In this/ proposed system, the center-based clustering strategy is selected to manage and control the clustering process with the assistance of the RSUs. RSU's can conducts the cluster formation step. Firstly, each node will send the Beacon message to its adjacent RSU, this message contains the position of this vehicle and its ID. Each cluster will be collected by the administrator in the RSU according to the coverage radius of each vehicle. The non-overlapped style is selected in all this dissertation simulation experiments. The relative position is utilized to perform the cluster formation process according to Euclidian distance equation.

When the relative position among the vehicles is less or equal to DSRC value, these vehicles are collected in one cluster. Algorithm 3.2 presents the cluster formation process.

3.4.2.2 Cluster Head Selection

The stability of each cluster cannot be approved unless there is an admin to these nodes in each cluster called CH, the others called CM. The CH is elected by the RSU based on specific parameters. Collecting the vehicles in clusters and selecting a CH will result in reducing the communication network overhead. The ideal method to select CHs will guarantee the increase in clusters lifetime, network reliability and availability.

To select an appropriate CH that can reduce the CH reselection frequently, the focus on vehicles speed as an essential parameters in a highway environment. Usually, the speed is very high in highway environments, therefore this dissertation focused on the relative speed in selecting the cluster heads. Equations 3.3 and 3.4 show the average speed calculation and the CH selection:

$$Ave\_speed^{w} = \frac{\sum_{i=1}^{N^{w}} speed_{v_i}}{N^{w}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

$$V_a^w = (speed_1, speed_2 \ldots\ldots, speed_{N^w}) \quad \ldots\ldots\ldots\ldots (3.4)$$

$$V_{a\_sorting}{}^{w} = \text{Ascending sort of } (V_a) \ldots\ldots\ldots\ldots (3.5)$$

$$ID\ (CH_w) = \begin{cases} Median\ of\ (V_{a\_sorting}w) & if\ number\ of\ vehicles\ is\ odd \\ Average\ (Median\ values) & otherwise \end{cases} \ldots\ldots\ldots (3.6)$$

Where : Ave_speed^{w} represents the average speed of cluster w, V_a^w is an array of the vehicles speed in cluster w, $V_{a\_sorting}$ represents an ascending sorted array of the vehicles speed in cluster w, N^w is the number of CMs in a cluster w, w represents the cluster number, and CH_w is the CH of a cluster w.

The proposed CH selection is based on the Median fundamental. The Median value can be indicated directly if the number of vehicles in a cluster is odd.

The vehicle that has an equal or nearest speed to the average speed from all vehicles is the suitable to be assigned as a CH. Sorting the vehicles speeds according to equations (3.4) and (3.5) respectively, and

Algorithm 3.3 CH selection

Input: NUM_OF_CLU, clusters
Output: CHs
set w←1
Repeat
  For each (cluster $_w$)
    $count_w$ ← number of vehicles in (cluster $_w$)
    Val_w ← $Median$ (speed) $_w$
    If $count_w$ mod 2 = 0 **then**
      CH_w ← ID (Val_w)
    Else
      CH_w ← ID of the vehicle which has nearest speed to Val_w
    Return CH_w
    end if
  end for each
  set w← w+1
Until (w> NUM_OF_CLU)
end

takes the median one which is the nearest one to the average speed value according to Equation (3.6). After applying the cluster formation stage according to the Algorithm 3.2, we have the number of clusters called NUM_OF_CLU . Algorithm 3.3 shows the main steps of CH selection.

3.4.2.3 Cluster Maintenance

The maintenance phase in each clustering process represents the important indication about the cluster stability. Cluster stability is important to improve the cluster efficiency and reliability. The cluster lifetime in VANET is temporary due to the frequent changeable in traffic and mobility. It is important to suggest an acceptable maintaining phase. There are many cases that must be considered such as : the connection loss between CMs and the adjacent cluster head, the connection loss between the RSU and cluster heads, the too-close clusters can be merged or split, etc. This phase has got less attention in VANET's related works. The clustering process steps are indicated in Figure 3.4.

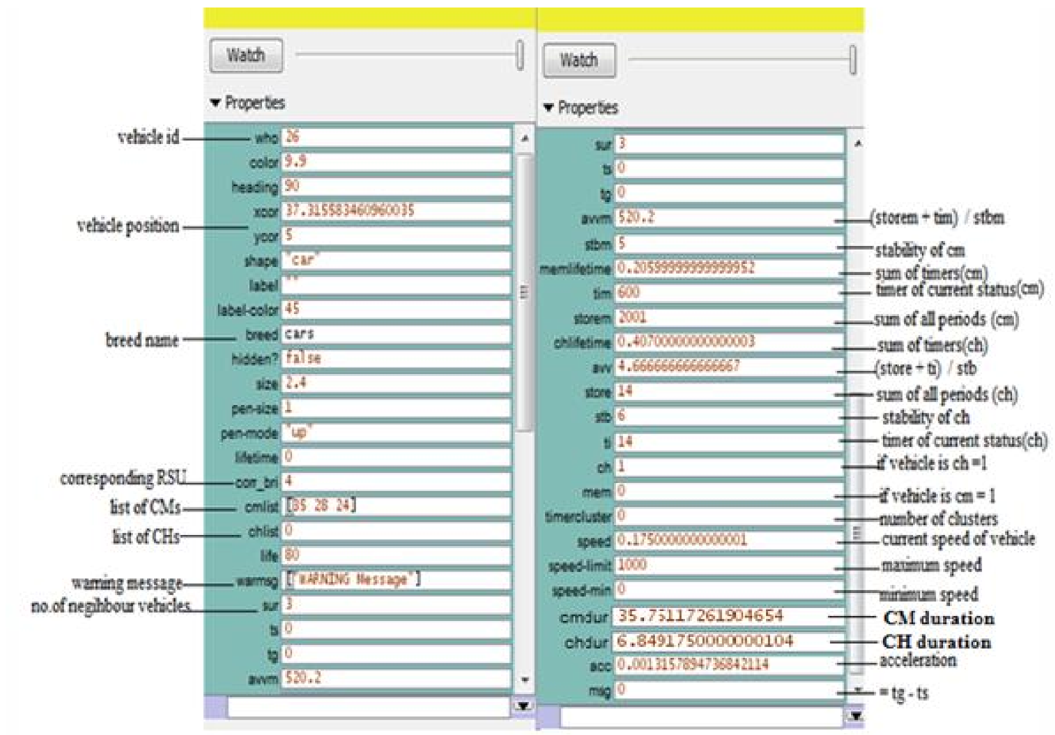


Figure 3.4 Vehicle properties in Net logo

3.5 Environment Setup

In order to evaluate the clustering approach, as it has previously explained, it is important to use the significant clustering parameters such: the average number of created clusters, the average CHs duration, the average CMs duration. To find these metrics, many computations are required. Each vehicle can be suggested as an agent that has many local parameters such as timers, counters, list of neighbor vehicles, list of CMs.

Each vehicle has many variable: (who) refers to the vehicle unique identifier which represent the programming name of a given vehicle. (Color) refers to the vehicle color. (Heading) represents the current direction's angle. (xcor and ycor) represent the coordinate of given vehicle, (Shape) represents the shape of current node, (Label) refers to the visible name, (label color) refers to the color of the visible name. (Breed) refers to the name of breed. (Hidden) is a Boolean value refers to the visible or hidden vehicle. All the previous properties built in the system for each agent. The variable (Cori-bri) refers to the corresponding RSU's In this dissertation, there are four metrics are suggested to insure the clustering stability.

(Cmlist) represents the list of all (CMs) if the given vehicle was a (CH). (Chlist) represents the list of (CHs) if the given node was a RSU. (Life) refers to the number of ticks when the current node is CH or CM. (Warmsg) refers to the received warning message text which received from the source according one type of traffic. (Sur) refers to the number of neighbors in n radius from the current node. (Ch) refers to the statues of the vehicle if it is CH or CM which takes 1 or 0 . The properties avvm, stbm, storem, tim, cmdur and memlifetime associates with CM stability and duration calculation as shown with details in Figure 3.4 .The properties avv, stb, store, ti, chdur and chlifetime associates with CH stability and duration.

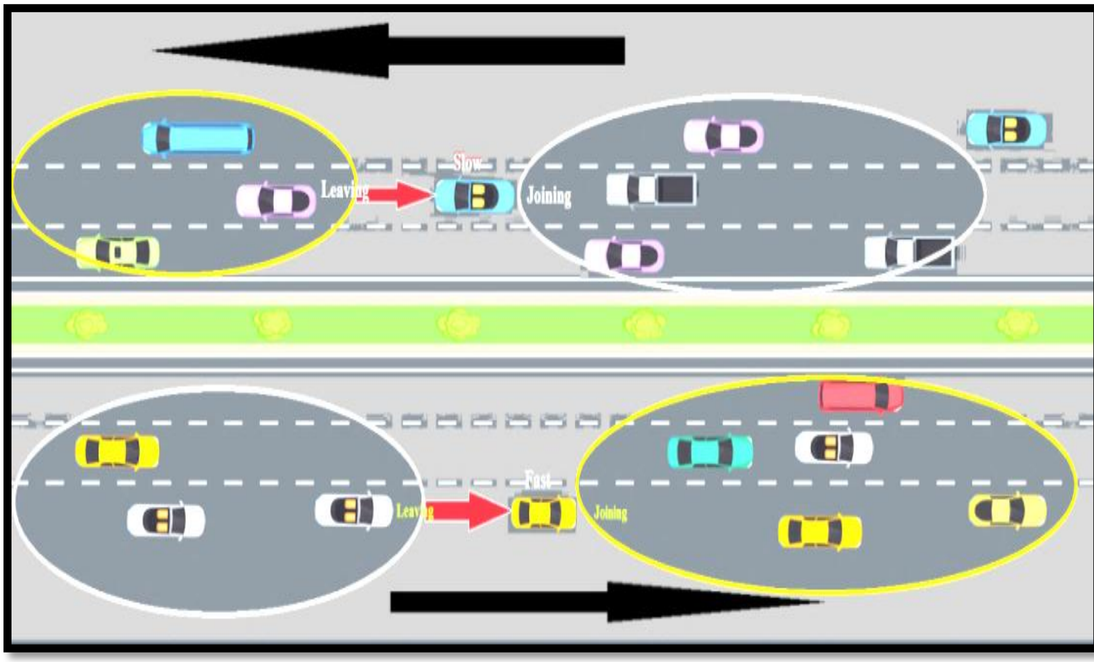


Figure 3.5 Leaving and Joining in road segment

3.6 Cluster Stability Predication

In these scenarios, after the clustering approach has been applied in VANET, each cluster will have composed of many vehicles. For each cluster after a short time interval called Δt_m, there are only three possible expected states: no changes, cluster increasing by one, or cluster decreasing by one only. The cluster increases by one vehicle (called Joining) in two cases, either join from forward or from backward as shown in Figure 3.5.

The forward increment comes from joining a lagged slowly vehicle to the current cluster. The backward increment comes from joining a new hurried vehicle. The cluster decreases one vehicle (Leaving) in two cases, either from forward or from backward. The forward decrement comes from a very fast vehicle in a current cluster which leaving it forward. The backward decrement comes from a slowly vehicle which lagged back from the current cluster. When the consecutive times intervals passed and the number of vehicles remains the same in their clusters, it reflects a good indication of the good stability.

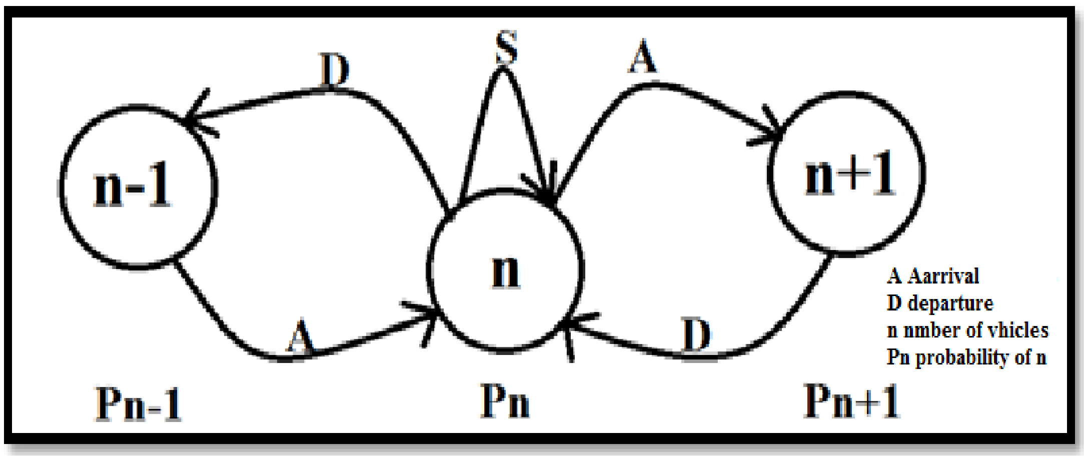


Figure 3.6 Vehicles State diagram in a cluster

The previous three cases are modeled and studied based on Markov fundamentals in all details. Figure 3.6 represents the markov chain approach to vehicle's clustering.

The hall state space for such process shows the possible states. These states probabilities are: no changing (represents that it is in state n and after short time will satay in state n), if a new vehicle joins to the cluster (the state is changed from n to n+1), finally, if the cluster loses one vehicle only (the state is changed from n to state n-1). Each case has a certain probability. The transition case indicated by (A) means either joining from forward or backward, the transition indicated by (D) means leaving either forward or backward.

Table (3.1) List of variables assamptions

Notation	**Description**
Δt_m	Time interval
$L_{forward}$	Losing vehicle due to its high speed
$L_{backward}$	Losing vehicle due to its low speed
$J_{forward}$	Winning vehicle due to its low speed
$J_{backward}$	Winning vehicle due to its high speed
t_m	Current time
N	Number of vehicles at time tm
S	No change
E_n	System state (the event at certain time)
$t_m+\Delta t_m$	The next time
. $P_n(t_m)$	The probability of the cluster contains n vehicles at time t_m
$O(\Delta t_m)$	Zero
$P_n(t_m, t_m+\Delta t_m)$ A	The probability of the cluster contains n vehicles at time $t_m + \Delta t_m$ Arrival
$O(\Delta t_m)$	The error

The modeling assumptions:

1: The joining probability at the interval time ($t_m + \Delta t_m$) in all two cases from forward or backward can be modeled in equations (3.7) (3.8) respectively as:

$P_n (t_m + \Delta t_m) = J_{forward} \Delta t_m + O(\Delta t_m)$……………..(3.7)

$P_n (t_m + \Delta t_m) = J_{backward} \Delta t_m + O(\Delta t_m)$………………(3.8)

2: The leaving probability at the interval time ($t_m + \Delta t_m$) in all two cases forward and backward can be modeled in equations (3.9)(3.10) respectively as:

$P_n (t_m + \Delta t_m) = L_{forward} \Delta t_m + O(\Delta t_m)$…………………..(3.9)

$P_n (t_m + \Delta t_m) = L_{backward} \Delta t_m + O(\Delta t_m)$…………………..(3.10)

3: The probability of more than one event occurrence in the same time is zero as shown in (3.11).

$$P_{n+z\,,\,z>1} (t_m + \Delta t_m) = 0 + O(\Delta t_m)\text{…………}(3.11)$$

Initially, the $P_n(t_m)$ is the probability of the event E_n at time tm, after passed a short period of time ($t_m + \Delta t_m$) the probability becomes $P_n (t_m + \Delta t_m)$ with new event E_n, E_{n+1} or E_{n-1}. to model all the cases , table 2 shows all cases with details:

According to table (3.1) , we can derive the final equation as shown in (3.12), to find the stability of cluster by expected the number of vehicles in any time as below:

$$P_n(t_m + \Delta t_m) = (P_{n-1} ((L_{forward}\,\Delta t_m)\ (1- J_{forward}\,\Delta t_m)\ (1- J_{backward}\,\Delta t_m)\ (1- L_{backward}\,\Delta t_m)$$
$$(L_{backward}\,\Delta t_m)\ (1- J_{forward}\,\Delta t_m)\ (1- J_{backward}\,\Delta t_m)\ (1- L_{forward}\,\Delta t_m\ )\ )$$
$$+\ (P_n\ (1- J_{forward}\,\Delta t_m)\ (1- J_{backward}\,\Delta t_m)\ (1- L_{forward}\,\Delta t_m)(1- L_{backward}\,\Delta t_m))$$
$$+\ (\ P_{n+1}\ (J_{forward}\,\Delta t_m)\ (1- L_{forward}\,\Delta t_m)\ (1- L_{backward}\,\Delta t_m)\ (1- J_{backward}\,\Delta t_m)$$
$$(J_{backward}\,\Delta t_m)\ (1- L_{forward}\,\Delta t_m)\ (1- L_{backward}\,\Delta t_m)\ (1- J_{forward}\,\Delta t_m)\ )$$
$$+\ O(\Delta t_m)\ldots\ldots\ldots\ldots\ldots(3.12)$$

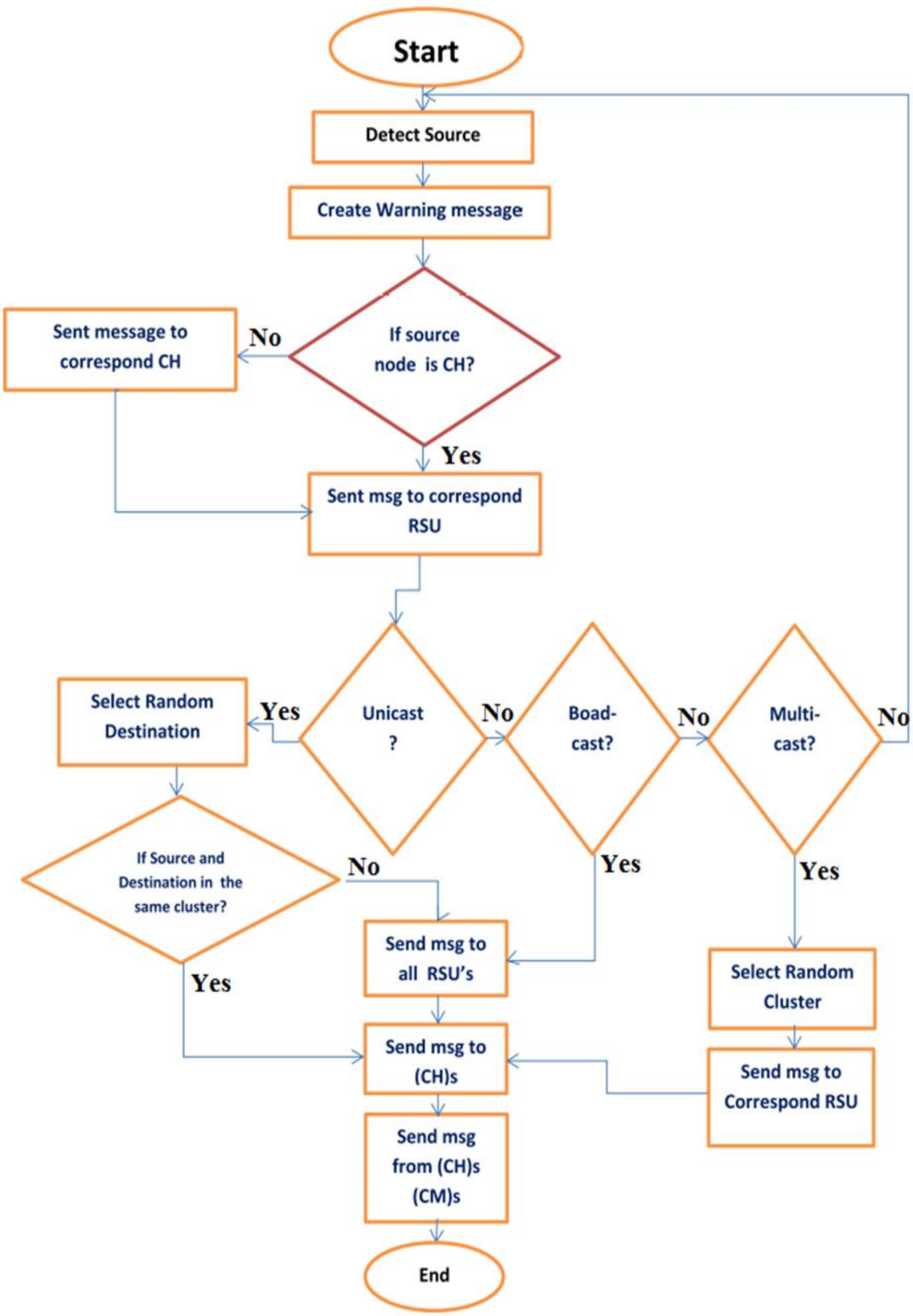
3.7 Data Dissemination Approach

Data Dissemination is implemented (based on the clustering approach) by transmitting a warning message (safety message) according to a specific event from the source vehicle to the recipient(s) in the Region Of Interest (ROI). For example a specific warning message for all police vehicles or ambulance vehicles. In the proposed system, The data transmission is applied with suppose the event in a random place, and the sender is randomly selected. The warning message is transferred in three styles: Broadcast, Multicast, and Unicast regardless of the type of emergent event.

The data dissemination in VANET starts from the vehicle itself or from infrastructure. The safety applications suppose that the data exchange depends on this principle. Data dissemination always refers to the data/information spreading in wireless networks. From the view of networking, the broadcast of the data requires capabilities at the link layer, to allow message spread to all the vehicles in the radio scope.

3.7.1 Select a Random Sender

The message transmits from the source to the destination(s) depending on the type of traffic. If it is broadcast, the message will be transferred



.

Figure 3.7 The block diagram of Data dissemination from the source to the corresponding cluster head, then forward to the corresponding bridge (RSU) which spreads it to all the vehicles behind the source to inform and warn them of the event. If it is multicast the same steps except, the corresponding RSU spreads it to the destination cluster. If it is unicast, the RSU gives it to the goal node through the goal node's RSU. Figure 3.7 represents the flow chart of the data dissemination in all cases.

To model the message, this dissertation selects a breed named message (in a simulation tool), then a declaration is made about its variables as local parameters

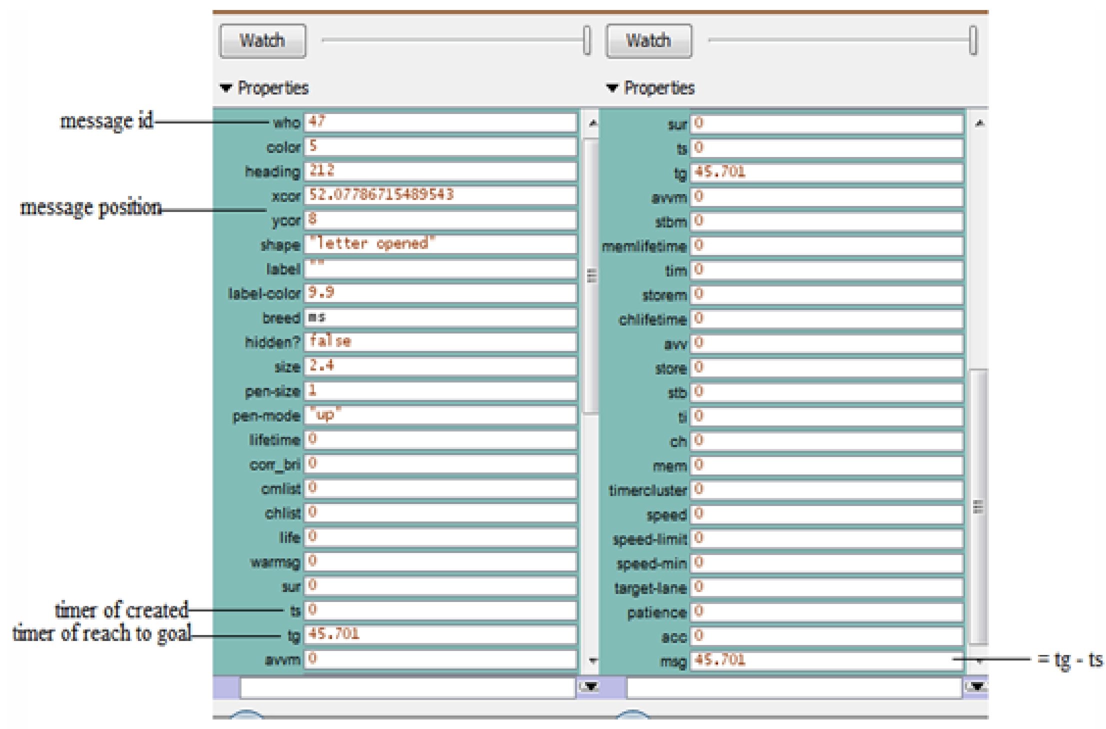


Figure 3.8 Message Properties in Netlogo

3.7.2 Data Dissemination Metrics

To evaluate the data dissemination process in all traffic types, the packet loss, delay, packet delivery ratio and throughput are representing the important metrics. These metrics requires calculations. As in the previous sections, the local parameters can be used to calculate these four metrics. Figure 3.8 shows the message properties in Netlogo.

The received and the total created messages were calculated throw the (broadcast, multicast and unicast) procedures. The throughput is calculated according to Equation (2.10). The packet loss can be calculated according to Equation (2.12) by subtracting the receives messages from the all created messages. The end to end delay is calculated according to Equation (2.11) by divided the summation of times consuming in messages transmission by the received messages. The packet delivery ratio is calculated according to Equation (2.13) by divided received messages by the total messages.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter is to investigate how the results of the proposed system achieve a good performance in terms of many performance metrics. As testing any computational model is required to run either by the real-world environment, simulations. The clustering approach and data dissemination have been designed and implemented using Netlogo 6.1.1 simulation. The obtained results are discussed to show which parameters have an impact on the results of the clustering model and data dissemination. In the context of data dissemination models, the analysis is done to highlight whether the data dissemination model is suitable for real-time applications or for other applications. On the other hand, a comparison between the proposed models and the promising related model is done to emphasize that the proposed models are related to the well-known models. This chapter is organized into six sections including the introduction. The second section contains simulation environment. The third section is dedicated to discussing the results of the clustering approach. The forth section explains the predicated stability using the Markov model results. The fifth section is to show Data dissemination models results. Finally section sixth contains chapter summery.

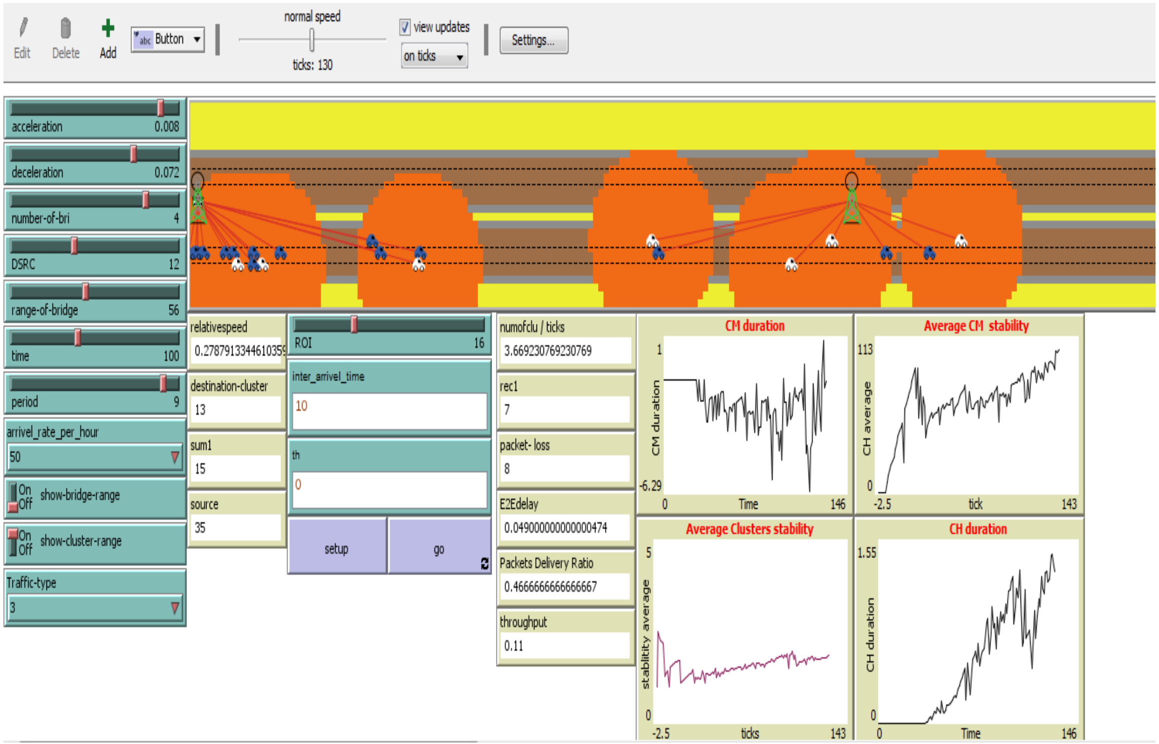


Figure (4.1) a Simulation snapshot

4.2 Simulation Environment

After designing and running the proposed clustering approach according to the simulation environment, the results were collected and presented using Netlogo results viewer. The suggested simulation environment is presented in Table (4.1). Figure (4.1) prsents a simulation snapshot as an impleminatation sample. The orange circuls represent clusters, the white vehicle represents the CH while the blue one represent the CMs. The vehicles are connected to the close RSU according to the transmission radius. If the vehicle is covered by two RSUs, it will be connected to the one with the strongest signal.

TABLE (4.1) Simulation Environment

Simulator	Netlogo 6.1.1
Number of Simulation Run	10
Cluster radius	300 m
RSU coverage area	1 k
Arrival rate	10, 50, 80, 100 Vehicles
Speed distribution	Exponential
Acceleration	0.004 m/s
Deceleration	0.086 m/s
Period of Beacon message	5 s
Simulation Time	100 s
Scenario	Highway
Street length	10 km
Number of lanes	3 per side

4.3 Clustering Stage Results

In this stage, the clustering algorithm has been applied on the road segment. The evaluation process to the clustering proposed algorithm depends on the performance metrics which are related to the clusters stability. Therefore, the CH duration or lifetime, CM duration or lifetime, CH stability, and CM stability have indicate the clustering algorithm efficiency. Different number of vehicles have been selected as indicated in Table (4.1). The results are presented depending on the average of 10 times simulation runs.

4.3.1 Vehicles Intensity

Four cases are tested to simulate the traffic intensity. These cases are sparse, medium, and dense as follows:

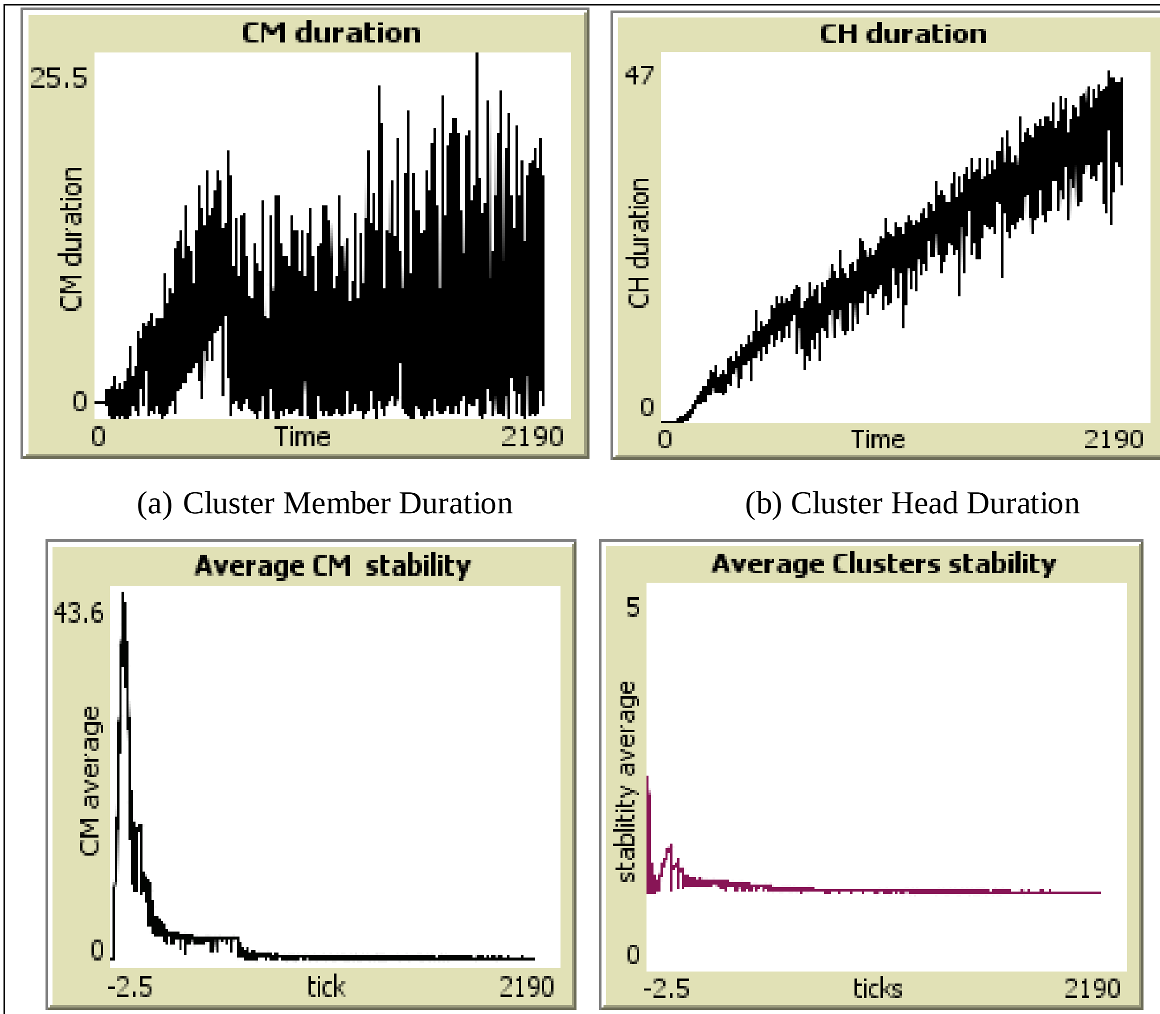
i. Scenario 1:

In this case, the number of vehicles was setting to be 10 with very sparse highway scenario. Figure (4.2) refers to the simulation results of the stability metrics in this case.

In (4.2(a)) the relation between the run time and the cluster member duration is expected and printed, in (4.2(b)) the relation between the simulation run time and the cluster head duration is expected and printed, In (4.2(c)) the average cluster member stability is presented based on changing its state with the simulation run time, and in (4.2(d)) the average cluster head stability is presented based on the CH changing state during the run time.

From Figure (4.2(a)), when the vehicles were convergence, the CM duration was gradually increasing with small fluctuation during simulation time increase , then when the vehicles became sparse, CM duration was very fluctuated. In general, the CM duration is over the average value which was rather good. From the Figure (4.2(b)), the CH duration also incrasing gradually with time increasing, but it was fluctuated.

From Figure (4.2(c)), when the number of vehehicles was 10, the CM stability was suddenly increasing at the beginning then stable for a certain period with small fluctuation then decreased to near zero after that keep stable, in general, the CM stability is a rather good indicator. From Figure (4.2(d)), The CH stability was suddenly increasing at the beginning then stable along time, in general, the CH stability is a very good indicator. From these results, the clustering proposed approach in sparce density is not stable due to the large distance among vehicles and high speed.



(a) Cluster Member Duration (b) Cluster Head Duration

(c) Cluster Member Stability (d) Cluster Head Stability

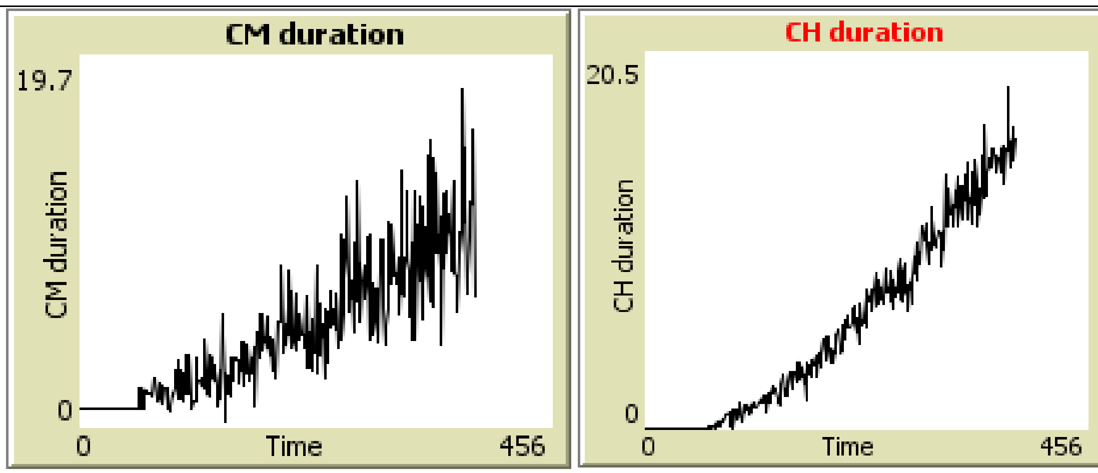
Figure 4.2 Case 10 vehicles

ii. Scenario 2:

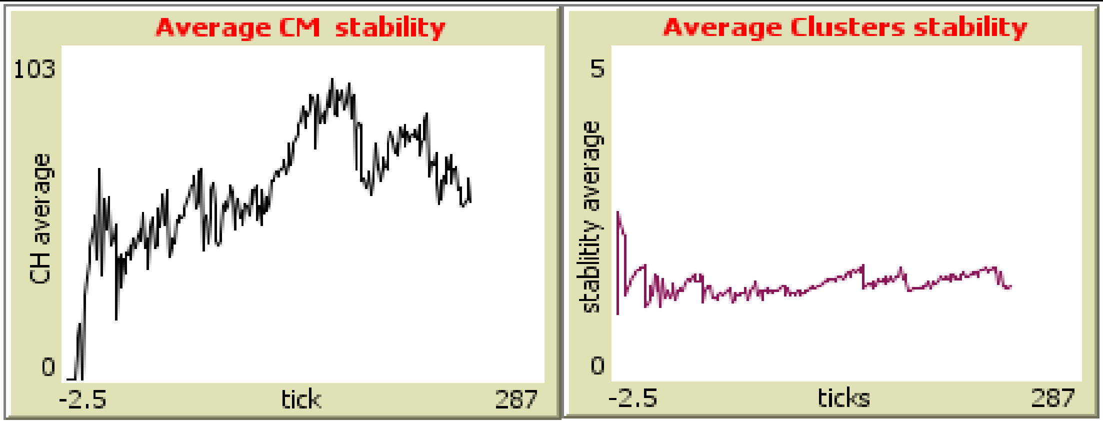
In this case the number of vehicles was 50 with very medium density highway scenario. Figure (4.3) refers to the results of simulation of the stability metrics in this case. The results are indicated in the same way that was followed in scenario1.

From Figure (4.3(a)) when the number of vehehicles was 50, the CM duration decreased and increased then it started fluctuating with the increased vehicles on road. From Figure (4.3(b)), the CH duration also increased gradually with vehicles number and time increasing, also the fluctuation was increased in the same manner.

From Figure (4.3(c)), The CM stability fluctuated increasing at the beginning then stable for a certain period with small fluctuation then decreasing to near zero after that keep stable, in general, the CM stability is a rather good indicator. From Figure (4.3(d)), The CH stability was suddenly increasing at the beginning then changing. From these results, the clustering proposed approach in medium density is stable due to the not large distance among.



(a)Cluster Member Duration (b)Cluster Head Duration



(c) Cluster Member Stability (d) Cluster Head Stability

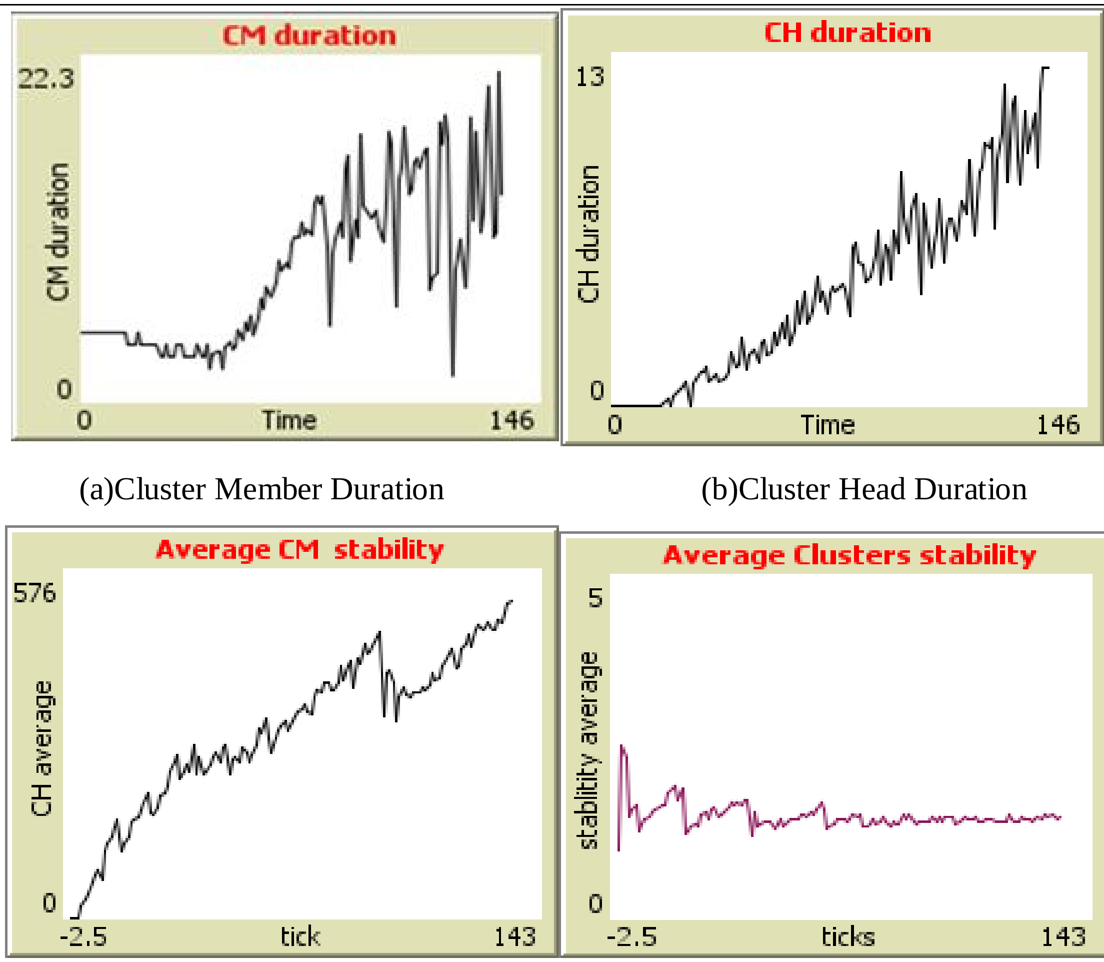
Figure 4.3 Case 50 vehicles

iii. Scenario 3:

In this case the number of vehicles was 80 with very sparse highway scenario. The Figure (4.4) refers to the results of simulation of the stability metrics in this case The results are indicated as in scenario1and scenario2.

From Figure (4.4(a)) when the number of vehicles was 80, the CM duration saved stable through the first quarter of simulation time, then started to increase. In the second half of simulation time, the fluctuation increased. From Figure (4.4(b)), the CH duration increased gradually with entered vehicles number and time increasing, also the fluctuation was almost a little

From Figure (4.4(c)), The CM stability increased gradually with time, and the number of entering vehicles increased. It was larger than the average value, therefore, it was a good indicator. From Figure (4.4(d)), The CH stability was improved with the increased time and number of entered vehicles. This was a good indicator of stability.



(a)Cluster Member Duration (b)Cluster Head Duration

(c) Cluster Member Stability (d) Cluster Head Stability

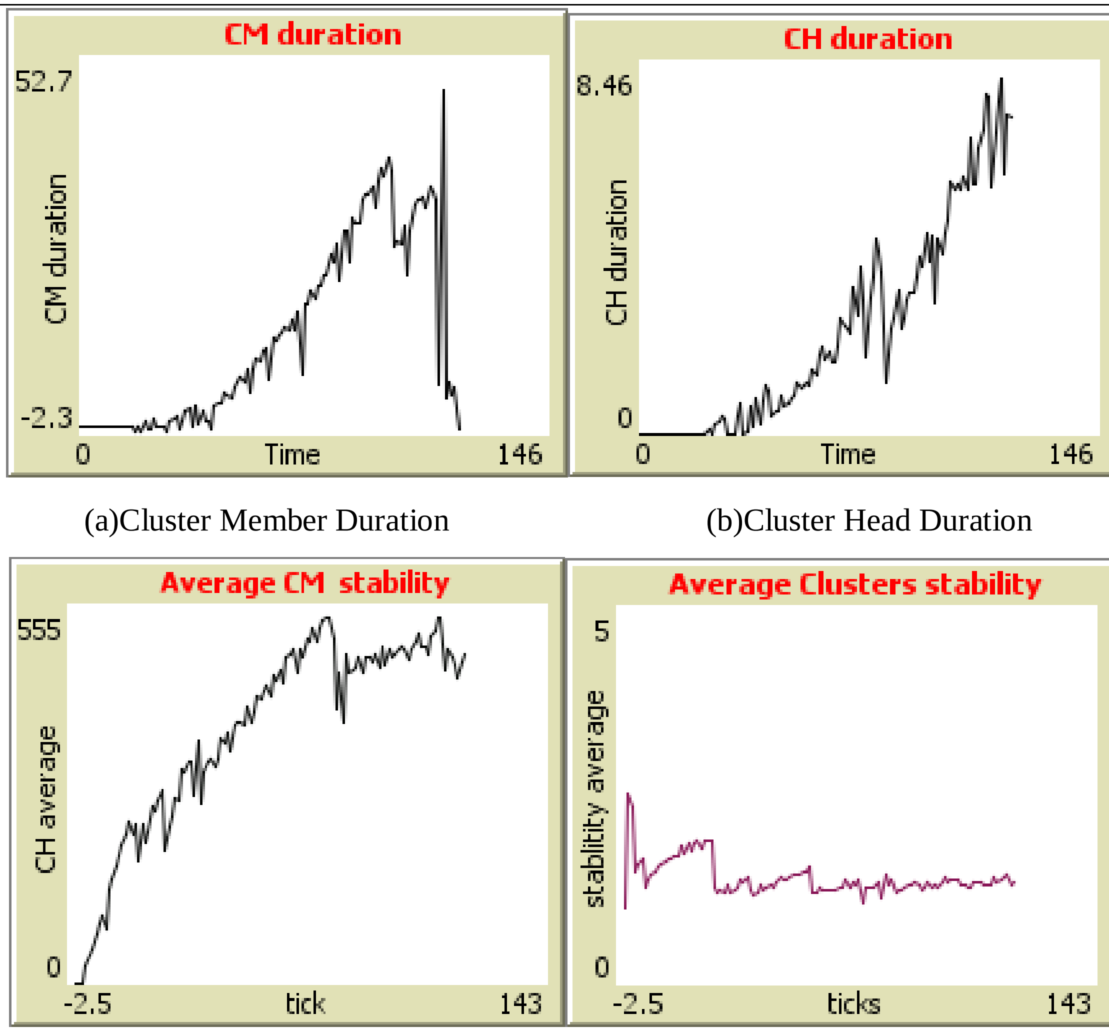
Figure 4.4 Case 80 vehicles

iv. Scenario 4:

In this case the number of vehicles was 100 with very denese highway scenario. Figure (4.5) refers to the results of simulation of the stability metrics in this case.

From Figure (4.5(a)) when the number of vehicles was 100, the CM duration increased with the time and entered vehicles increased. It was a good indicator about stability. From Figure (4.5(b)), the CH duration increased gradually with entered vehicles number and time increasing, also the fluctuation was almost a little. This was a good indicator about stability.

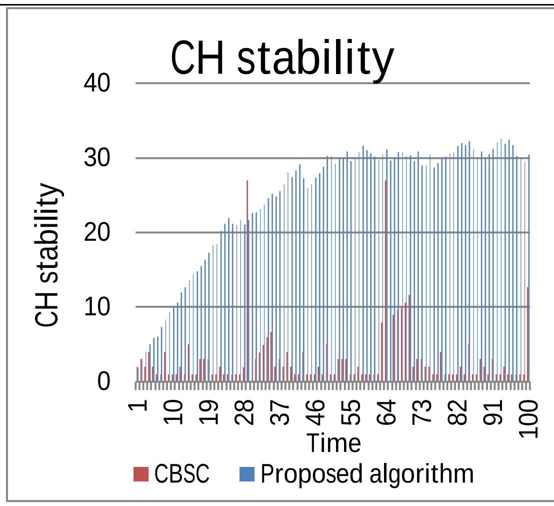
From Figure (4.5(c)), The CM stability increased gradually with time, and the number of entering vehicles increased. It was larger than the average value, therefore, it was a good indicator. From Figure (4.5(d)), The CH stability was stable with the increased time and number of entered vehicles. This was a rather good indicator of stability.



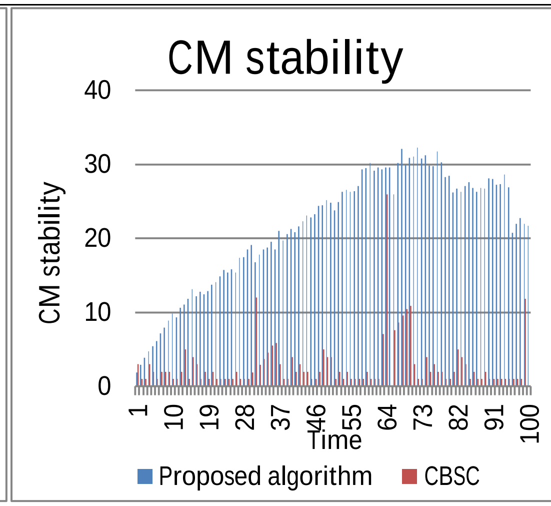
(a)Cluster Member Duration (b)Cluster Head Duration

(c) Cluster Member Stability (d) Cluster Head Stability

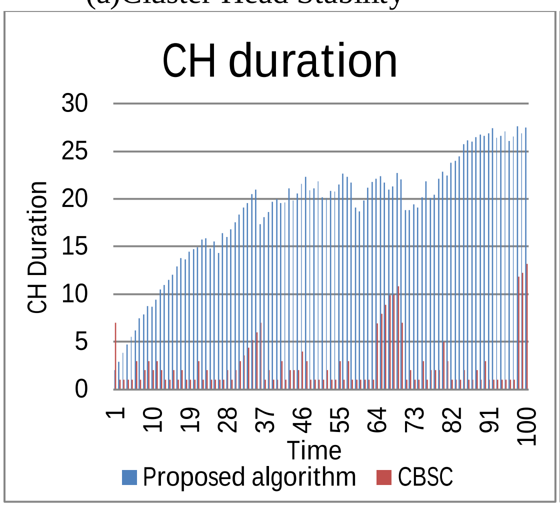
Figure 4.5 Case 100 vehicles



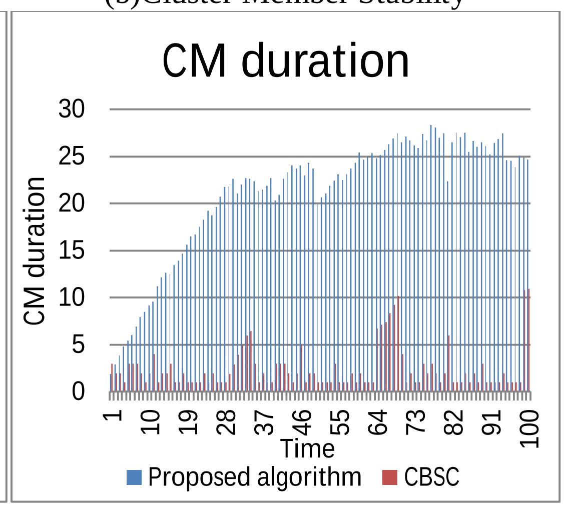
(a)Cluster Head Stability



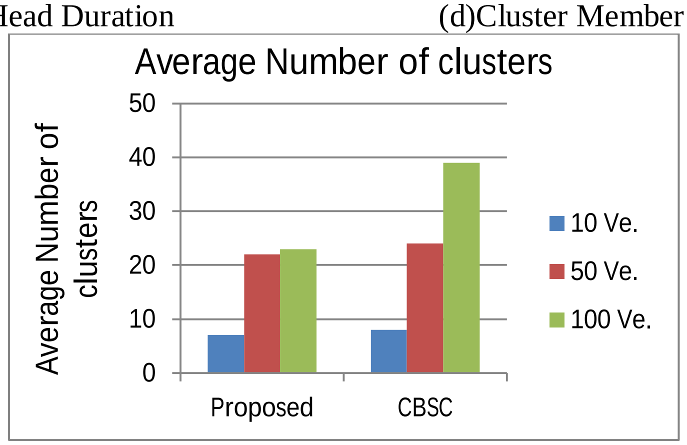
(b)Cluster Member Stability



(c)Cluster Head Duration



(d)Cluster Member Duration



(e) Average number of clusters

Figure 4.6 comparison between the proposed system and CBSC

4.3.2 Evaluation the Proposed Clustering Approach

In order to prove the goodness and efficiency of a proposed clustering approach in this dissertation, it is important to compare it with another system which is dependent as a baseline to the proposed clustering algorithm. The performance metrics for CBSC (Cheng and Huang, 2019) which represents the other system are implemented and calculated for the suggested environments in this dissertation environment. These results are compared with this dissertation proposed clustering approach metrics. The used CBSC algorithm was already compared with other related works such as K-means algorithm and Stable Clustering Algorithm for vehicular ad hoc networks (SCalE). Their previous comparison showed that CBSC is more stable than those others. Figure (4.6) shows the comparision between each two similar performance metrics for the proposed approach in this dissertation and the CBSC algorithm metrics.

From Figure (4.6), all the results indicate that the proposed clustering approach are superior all the performance metrics measured for CBSC. These proposed approach results are more stable than the CBSC results because the proposed system focuses on the speed parameter in CH selection which causes a notices improve in results also in CBSC, the security computation effects in the cluster stability in CBSC. The proposed approach indicate the utilization of the optimization fundamentals in creating the clustering approach.

4.4 Cluster Stability Predicting Results

There are three network sizes selected, namely 50, 80, and 100. Also, the range of clusters was set as a variable. The efficiency of the clustering algorithm depends on the mathematical approach called the stochastic process. The joining and leaving of vehicles to the clusters can be predicted mathematically using a Markov model based on the birth-death process as discussed in chapter three. But in this study, practical implementation is applied to practically fit the mathematical approach of the Markov concept. The number of vehicles in each cluster was observed in different sequential times (i.e. at t_m, $t_m + \Delta t_m$, $t_m + 4\Delta t_m$, $t_m + 8\Delta t_m$, $t_m + 12\Delta t_m$, $t_m + 16\Delta t_m$, $t_m + 20\Delta t_m$), to observe the leaving and joining vehicles. Also, the average and standard deviation (STD) are calculated to indicate central tendency and the data spread. The stable cluster is the cluster that has the number of vehicles equal to or close to the average number of vehicles in each unit of time.

4.4.1 Case 1:

Many different experiments were implemented and analyzed to observe the number of created clusters in different simulation runs for different highway VANET environments. The number of vehicles in each cluster at different successive times was also observed and indicated. The variation in cluster vehicles number at successive times is used as an indication about the behavior of the clustering process, whether or not it can be modeled as the birth-death Markov process.

Table 4.2 Number of vehicles in 12 cluster when Δtm=4 and network size=50

Random clusters	*Number of nodes at t_m*	*Number of nodes at $t_m+\Delta t_m$*	*Number of nodes at $t_m+4\Delta t_m$*	*Number of nodes at $t_m+8\Delta t_m$*	*Number of nodes at $t_m+12\Delta t_m$*	*Number of nodes at $t_m+16\Delta t_m$*	*Number of nodes at $t_m+20\Delta t_m$*	*Average*	*STD*
cluster 1	4	4	4	4	4	4	4	4	0
cluster 2	1	1	1	1	1	1	1	1	0
cluster 3	3	3	3	3	3	3	2	2.857142	0.377964473
cluster 4	3	3	3	3	3	2	2	2.714285	0.487950036
cluster 5	5	6	6	7	7	7	8	6.571428	0.975900073
cluster 6	4	4	4	4	4	4	4	4	0
cluster 7	2	2	2	2	2	2	2	2	0
cluster 8	2	2	2	2	2	2	2	2	0
cluster 9	4	4	4	4	3	3	3	3.571428	0.534522484
cluster 10	3	3	3	3	3	3	2	2.857142	0.377964473
cluster 11	2	2	2	2	2	2	1	1.857142	0.377964473
cluster 12	2	2	2	1	1	1	2	1.571428	0.534522484

Table 4.2 represents the collected data about the 12 created clusters for the case of a highway low traffic (50 vehicles) at different successive times (t_m, $t_m+\Delta t_m$, $t_m+4\Delta t_m$, $t_m+8\Delta t_m$, $t_m+12\Delta t_m$, $t_m+16\Delta t_m$, $t_m+20\Delta t_m$). It is taken into consideration that Δt_m is used as very small time periods equal to (4 millisecond) in this study. The observed change in the vehicle's number is either ±1 or 0. This concept is identical to the Markov process. The expected and standard deviation values for each cluster data indicate high stability with respect to the numbers in each cluster.

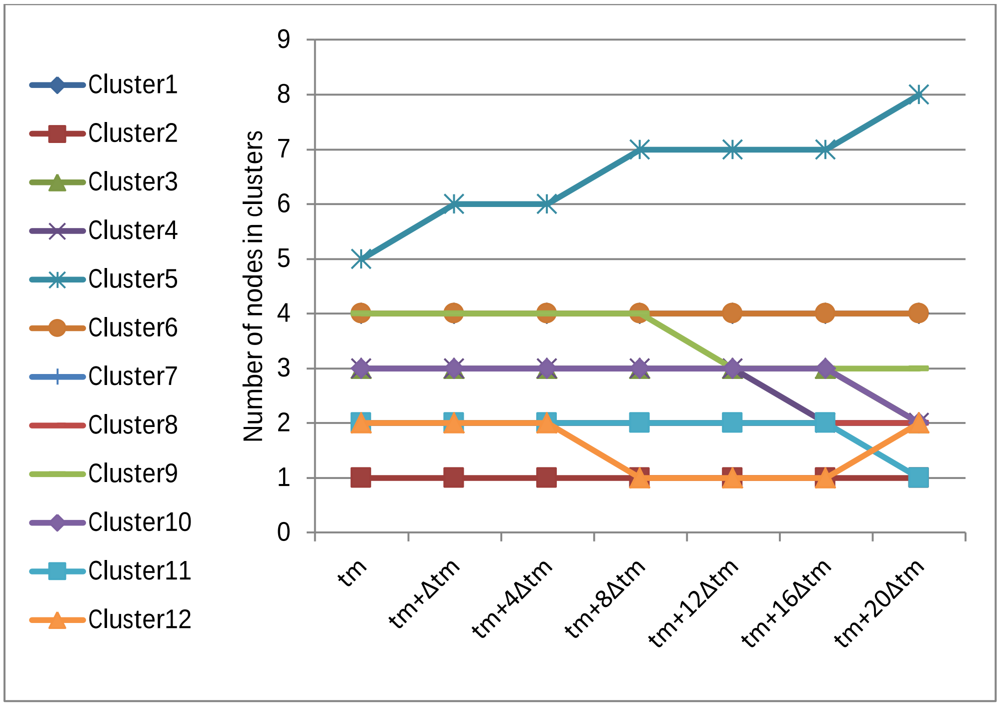


Figure 4.7 Number of vehicles in 12 cluster when Δt_m=4 and network size=50

Figure 4.7 represents the graphical representation of the created data from the case in Table 4.2. A clear indication is observed in cluster 5, as the increment in cluster members is due to the joining process. On the other hand, in clusters 3, 9, and 10, the decrement in cluster members happened due to leaving process. The others either remained stable (S) without change or contained consecutive leaving and joining with sequential times.

4.4.2 Case 2:

In this case, the number of vehicles was 80 vehicles . As shown in table 4.3, the results represents a values of s 12 random samples of these clusters to evaluate the stability when the traffic is medium (80 nodes) with different consecutive times. A clear indication can be observed for clusters 3, 4, 6, and 8, the increment in cluster members is due to the joining process.

Table 4.3 Number of vehicles in 12 cluster when Δt_m=4 and network size=80

Random clusters	*Number of nodes at t_m*	*Number of nodes at $t_m+\Delta t_m$*	*Number of nodes at $t_m+4\Delta t_m$*	*Number of nodes at $t_m+8\Delta t_m$*	*Number of nodes at $t_m+12\Delta t_m$*	*Number of nodes at $t_m+16\Delta t_m$*	*Number of nodes at $t_m+20\Delta t_m$*	*Average*	*STD*
cluster 1	3	3	3	3	3	3	3	3	0
cluster 2	4	4	4	4	4	4	4	4	0
cluster 3	2	2	2	3	3	3	3	2.571428	0.534522484
cluster 4	2	2	2	3	4	4	4	3	1
cluster 5	5	5	5	4	4	4	4	4.428571	0.534522484
cluster 6	1	2	2	2	2	2	2	1.857142	0.377964473
cluster 7	2	1	1	1	1	1	1	1.142857	0.377964473
cluster 8	2	3	4	4	4	4	4	3.571428	0.786795792
cluster 9	3	3	3	5	5	5	4	4	1
cluster 10	3	3	4	3	3	3	2	3	0.577350269
cluster 11	2	3	2	2	2	2	1	2	0.577350269
cluster 12	3	2	2	1	1	1	1	1.571428	0.786795792

On the other hand, in clusters 5, 7, and 12, the decrement in cluster members happened due to the leaving process. The others either remained stable (S) without change or contained consecutive leaving and joining with sequential times.

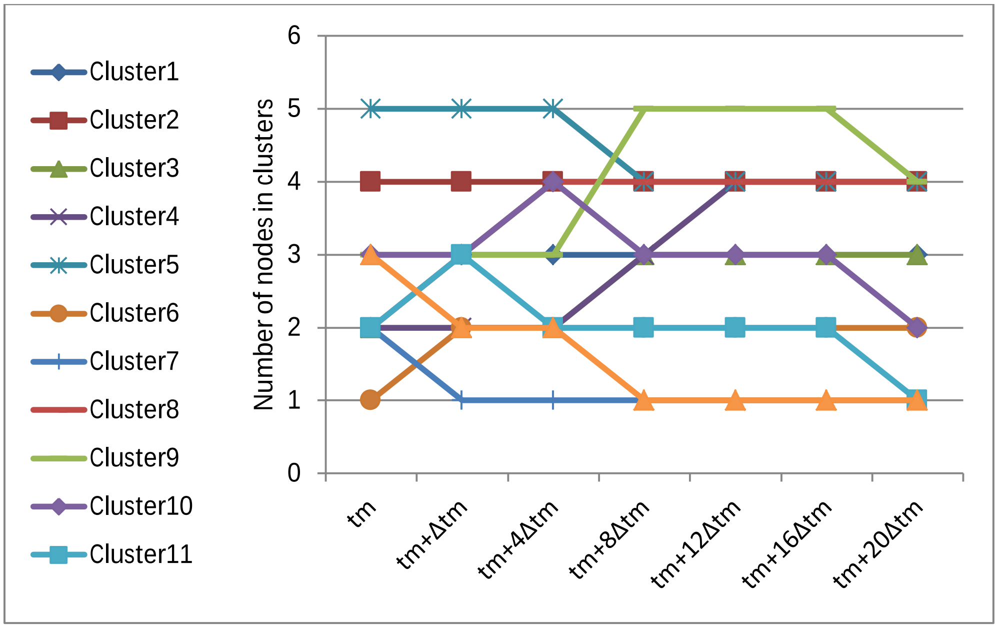


Figure 4.8 Number of vehicles in 12 cluster when Δt_m=4 and network size=80

Figure (4.8) represents the graph of Table (4.3)

Table 4.4 Number of vehicles in 12 cluster when Δt_m=4 and network size=100

Random clusters	*Number of nodes at t_m*	*Number of nodes at $t_m+\Delta t_m$*	*Number of nodes at $t_m+4\Delta t_m$*	*Number of nodes at $t_m+8\Delta t_m$*	*Number of nodes at $t_m+12\Delta t_m$*	*Average*	*STD*
cluster 1	6	6	6	6	6	6	0
cluster 2	5	5	4	4	4	4.4	0.547722
cluster 3	2	2	3	3	3	2.6	0.547722
cluster 4	4	4	4	4	4	4	0
cluster 5	5	4	5	5	5	4.8	0.447213
cluster 6	3	3	3	2	2	2.6	0.5477225
cluster 7	1	2	1	1	2	1.4	0.5477225
cluster 8	2	2	2	3	2	2.2	0.4472135
cluster 9	2	2	2	2	3	2.2	0.4472135
cluster 10	2	2	1	2	3	2	0.7071067
cluster 11	2	2	2	2	1	1.8	0.4472135
cluster 12	1	1	2	3	1	1.6	0.8944271

4.4.3 Case 3:

When the traffic is high, as shown in Table 4.4, a clear indication is observed in cluster 5, as the increment in cluster members is due to the joining process. On the other hand, in clusters 2, 3 and 6, the decrement in cluster members happened due to the leaving process.

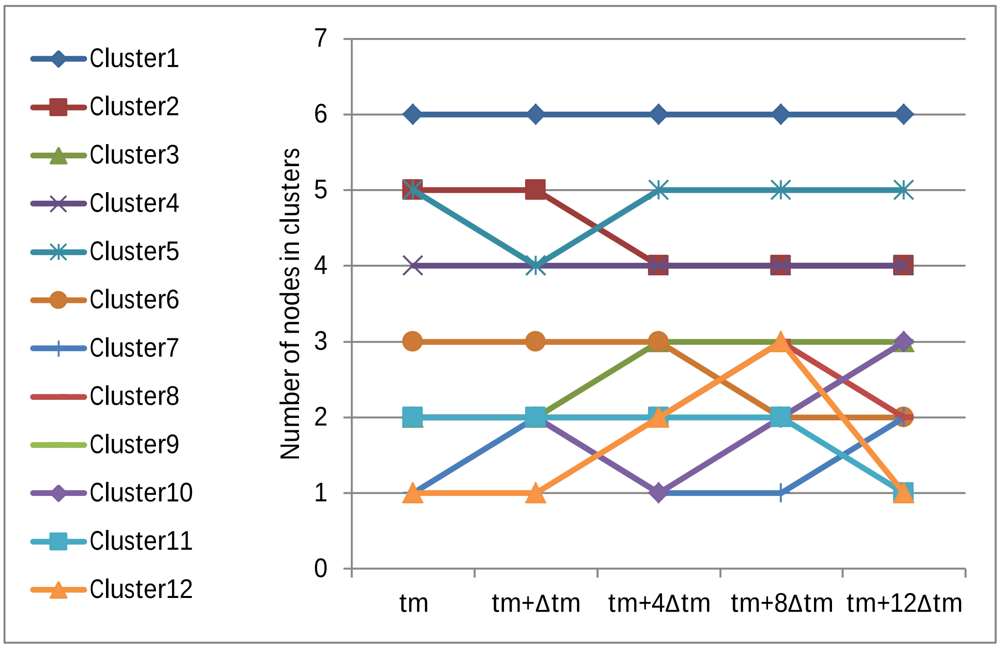


Figure 4.9 Number of vehicles in 12 cluster when Δt_m=4 and network size=100

The others either remained stable (S) without change or contained consecutive leaving and joining with sequential times. Figure (4.9) represents the graph of table (4.4).

4.4.4 Stability Prediction Modeling

A mathematical model based propability theory, Markov and Birth Death process is suggested to predict the stability of the created clusters in this dissertation proposed approach as indicated in 3.6.

The developed eqation (3.12) which can be used to estimate the stability of a cluster based on its average number of vehicles at any time. The following example (4.1) presents a practical implementation to show the propability of finding the designated number of vehicles at certain time during the prediction process.

Example (4.1):

Suppose $L_{forward}$ = 0.3, $L_{backward}$ =0.7, $J_{forward}$ = 0.4, $J_{backward}$ = 0.6, n= 5 (the number of vehicles in the current cluster), Δt_m= 0.01 second and t = 1s . Predicate the stability of the current cluster after Δt_m.

Solution: According to equation (3.12) in chapter 3:

$$
\begin{aligned}
P_n(t_m, t_m + \Delta t_m) = & [\, 0.003 * (1 - 0.004) * (1 - 0.006) * (1 - 0.007) * \\
& (0.007) * (1 - 0.004) * (1 - 0.006) * (1 - 0.003)] \\
& + [(1- 0.004) * (1- 0.006) * (1 - 0.003) * (1 - 0.007)\,] \\
& + [(\, 0.004 * (1- 0.003) * (1 - 0.007) * (1 - 0.006) * \\
& (0.006) * (1 - 0.003) * (1 - 0.007) * (1 - 0.004)] \\
& + 0
\end{aligned}
$$

Pn(1 + 0.01) = 0.00002329 + 0.98014455 + 0.0000203 + 0

= **0.98018814**

From the previous example, the value of the probability is very high (0.98) which gives an indication about the stability of the created clusters in this dissertation.

4.5 Data Dissemination Results

When the performance metrics are calculated with different parameters, the average throughput, delay, packet loss, and packet delivery ratio

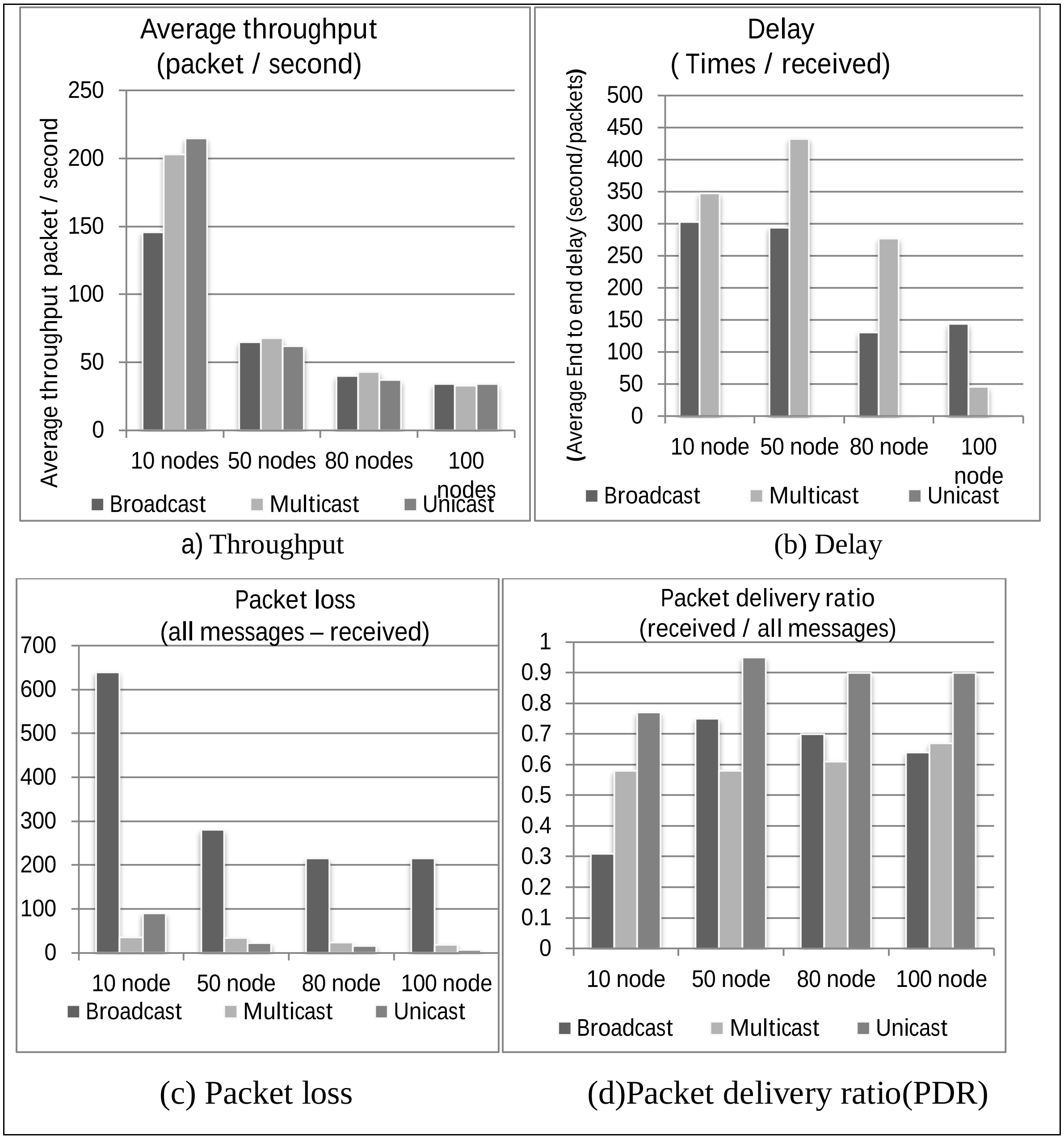


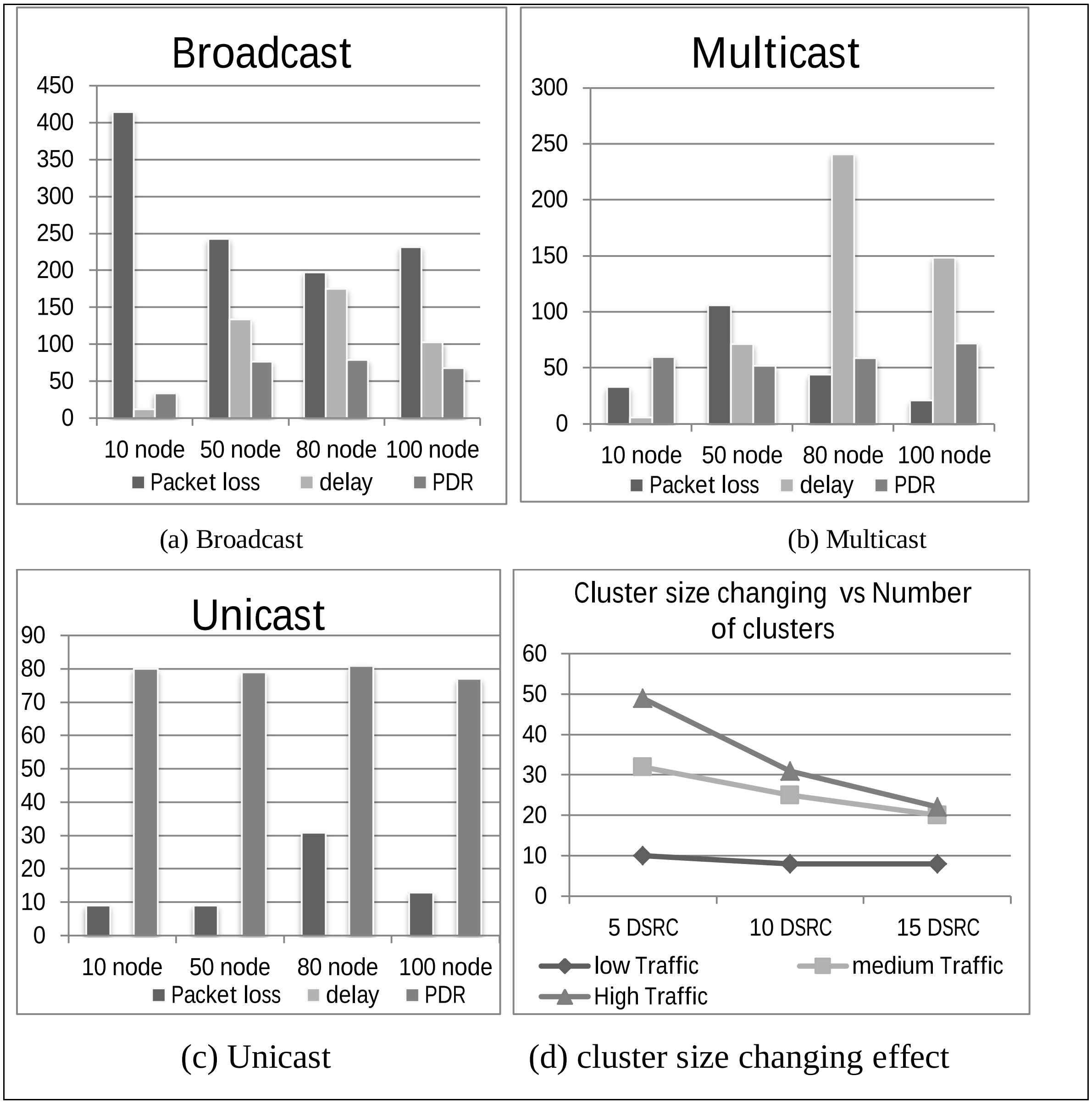
Figure (4.10) Data dissemination Performance metrics

(according to Equations (2.10), (2.11), (2.12), and (2.13) respectively) are obtained as part of the evaluation of the proposed system. When the number of vehicles equals 10, 50, 80, and 100 the results are as shown in Figure (4.10).

The throughput as shown in Figure (4.10(a)), decrease with number of vehicles increasing. The delay in broadcast and multi cast (as shown in Figure(4.10(b)), seems to follow the Poisson distribution. Utilizing Poisson distribution will help in predicting the expected values and their standard deviation in a possible manner. These two values are usually equal. In Unicast, the delay was the lowest value compared with Broadcast and Multicast.

The packet loss in the three communication methods seem to follow the Exponential distribution. Utilizing the Exponential distribution will help in predicting the expected values and their standard deviation in a possible manner. As shown in Figure(4.10(c)), the packet loss in Unicast and Multicast increases along with the increase of vehicles. The packet loss in a Broadcast is higher than the other two communication methods.

In Figure(4.10(d)), the Packet delivery ration in the 3 communication methods seems to follow the Uniform distribution. Utilizing the Uniform distribution will help in predicting and easing the process of estimating the expected values in a possible manner. From Figure (4.10), the optimal size of cluster from the view of data dissemination is 100 nodes in terms of delay and packet loss.



(a) Broadcast (b) Multicast

(c) Unicast (d) cluster size changing effect

Figure (4.11) Data dissemination Performance metrics comparisons

From Figure (4.11(a)), in the Broadcast, the Packet loss was the largest when the number of vehicles was 10 and increased gradually with the number of vehicles increased. While the other two performance metrics were opposite to packet loss. From Figure (4.11(b)), in the Multicast, the performance metrics took the Poisson distribution. From Figure (4.11(c)), in the Unicast, the performance metrics was very good in all number of vehicles. All the results are clarified in Appendix B. The increasing in cluster radius causes reduction in number of clusters due to merging as shown in Figure (4.11(d)).

Table (4.5) represent the numerical results, when the number of vehicles was 10 vehicles, of the data dissemination performance metrics. In addition to throughput, PDR, and delay, there are other metrics like average number of clusters, received message, and packet loss. In the same arrangement,

Table (4.6), Table (4.7), and Table (4.8) represent the numerical results when the number of vehicles 50, 80 and 100 vehicles respectively.

Table 4.5 Data dissemination performance metrics in 10 vehicles

10 Vehicles	Created messages	Received message	Average number of clusters	Packet loss	End 2 end delay	Throughput	Packet delivery ratio
Broadcast	664	249	5	415	12.556	146	34%
Multicast	84	51	5	33	5.611	203	60%
Unicast	63	50	4	13	0.126	215	80%

Table 4.6 Data dissemination performance metrics in 50 vehicles

50 Vehicles	Created messages	Received message	Average number of clusters	Packet loss	End 2 end delay	Throughput	Packet delivery ratio
Broadcast	1068	825	10	243	134.02	45	77%
Multicast	225	119	4.96	106	71.62	46	52%
Unicast	150	119	3.98	31	0.107	37	79%

\

Table 4.7 Data dissemination performance metrics in 80 vehicles

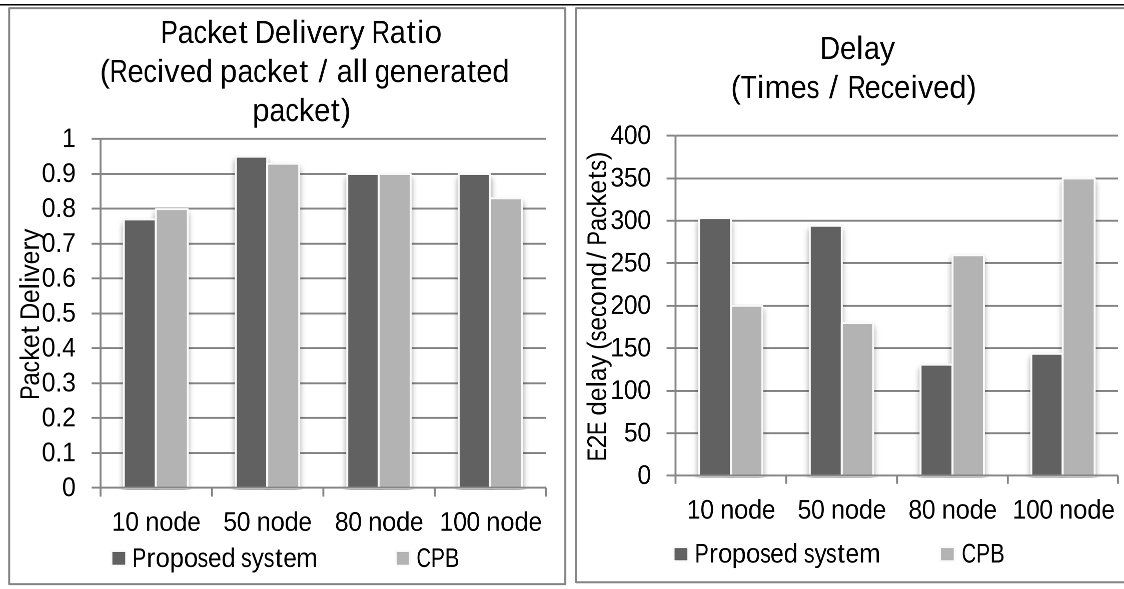
80 Vehicles	Created messages	Received message	Average number of clusters	Packet loss	End 2 end delay	Throughput	Packet delivery ratio
Broadcast	945	747	9	198	175.1	36	79%
Multicast	109	65	9	44	241.0	39	59%
Unicast	49	40	10	9	0.094	30	81%

Table 4.8 Data dissemination performance metrics in 100 vehicles

100 Vehicles	Created messages	Received message	Average number of clusters	Packet loss	End 2 end delay	Throughput	Packet delivery ratio
Broadcast	733	501	9	232	103.089	110.88	68%
Multicast	77	56	9	21	148.64	110.88	72%
Unicast	40	31	10	9	0.05	110.88	77%

4.5.1 Evaluation Data Dissemination Results

To prove the goodness of the data dissemination approach based on the proposed clustering approach, it is important to compare our results with (Liu *et al.*, 2018). In (Liu *et al.*, 2018), data are disseminated based on the distributed clustering approach without RSUs. The Packet Delivery Ratio (PDR) and Delay are compared with Clustering and Probabilistic Broadcasting (CPB) algorithm.



(a) Packet Delivery Ratio (b) Delay

Figure (4.11) Data dissemination Performance metrics comparisons

From Figure (4.11(a)) the PDR in both CPB and the proposed system is close in the case of the sparse network but it seems to be increased in the proposed system with the number of vehicles increasing better than CPB. In Figure (4.11(b)), the Delay in the proposed system decreased with the number of vehicles increasing unlike to CPB. The justification of these results is, in the proposed system, RSUs reduce: congestion, network overhead, and broadcast storm occurrence, due to their role in clusters management and data disseminate from source to the CH then to the RSU which take the role of sending data to the destination(s). While in CPB, the distributed clusters without RSUs expose the network as an exhibition to broadcast storms, congestion, increased overhead, etc.

CHAPTER FIVE

CONCLUSIONS AND FUTURE WORKS

5.1 Introduction

This chapter is to preview the conclusions of the proposed models which are inferred from the performance metrics. These conclusions are taken based on the simulation experiments in a virtual VANET environment. The results have been designed and implemented using Netlogo 6.1.1 simulation. Also, this chapter shows the hardware requirements suitable to perform the experiments.

5.2 Clustering Stage Conclusions

The results of clustering algorithms performance metrics have been collected and discussed clustering , the conclusion of each case shown as following:

1) When the number of compounds is small, the CM duration improved as the number of vehicles increased. Also, the fluctuation begins to decrease. Therefore, choosing Dense network guarantees the quality of CM.
2) The CH duration in different number of vehicles is similar but the fluctuation is decrease as the number of vehicles increased. Therefore the large number of vehicles improver the CH duration.
3) The CM stability in 10, 50, and 80 vehicles were bad indicator, while in 100 vehicle is better because when the number of vehicles in scenario increased, the distance among vehicles becomes small which helps in cluster stability.

4) The CH stability was save it case with increasingly fluctuation as the number of vehicles as increased. In sparse network, the CH stability is a good indicator.
5) The comparison of the proposed clustering algorithm with CBSC reflects a good indicator to the optimized results in stability. As it refer previously, the reasons related to the focusing in effective parameter in highway scenario like speed in CH selection.

5.3 Cluster Stability Prediction Conclusions

The numerical results, which related to data dissemination stability prediction in the previous chapter, reflected the probability of joining, leaving or there is no change in each cluster. The number of vehicles in each of the created clusters in different environments are been either unchanged or slightly changed by one only. This indication can be used to represent the mathematical concept of Markov. The average number of vehicles and their standard deviation in each cluster at different successive times are also utilized as a statistical test with central tendency and dispersion for the stability of each cluster. Also these results can be deduce from the Equation (3.12) by apply the probability of leaving and joining.

5.4 Data Dissemination Stage Conclusions

The results showed that the sparse scenario with unicast data dissemination achieves a good performance metrics. In addition, the delay and packet loss were found to be reduced as compared to the other methods. The comparison between the proposed data dissemination approach and CPB algorithm reflects the difference between centralized and distributed cluster formation. In centralized when the number of vehicles increased , the chance of broadcast storm occurrence is little as it happens in the proposed data dissemination approach. In distributed the increasing in number of vehicles threats the scalability and causes broadcast storms.

5.5 Future Works

- The stability can be predicted in two sides of a road segment. In addition, the method of center-based cluster formation can be replaced by the distributed method to reduce the cost of build and setup of the RSUs. As well, the system can be developed by using overlapped clustering method.
- The proposed model can be improving depending on the k-hop clustering approach and selecting two CHs instead of one CH.
- Using probability in data dissemination approach as an effective parameter in sending warning message.

REFERENCES

REFERENCES

Abboud, K., Omar, H.A. and Zhuang, W. (2016) ‘Interworking of DSRC and cellular network technologies for V2X communications: A survey’, *IEEE transactions on vehicular technology*, 65(12), pp. 9457–9470.

Abdel-Halim, I.T. and Fahmy, H.M.A. (2018) ‘Prediction-based protocols for vehicular Ad Hoc Networks: Survey and taxonomy’, *Computer Networks*, 130, pp. 34–50.

Ahmad, I. *et al.* (2018) ‘VANET–LTE based heterogeneous vehicular clustering for driving assistance and route planning applications’, *Computer Networks*, 145, pp. 128–140. doi:https://doi.org/10.1016/j.comnet.2018.08.018.

Al-Sultan, S. *et al.* (2014) ‘A comprehensive survey on vehicular ad hoc network’, *Journal of network and computer applications*, 37, pp. 380–392.

Alioua, A. *et al.* (2017) ‘A distributed multi-hop clustering algorithm for infrastructure-less vehicular ad-hoc networks’, in *International Conference on Information and Communication Technology for Develoment for Africa*. Springer, pp. 68–81.

Allani, S. (2018) ‘Data Dissemination and Aggregation in Vehicular Adhoc Network’. University Pauetdes Pays Adour; University de Tunis El Manar.

Almheiri, S.M. and Alqamzi, H.S. (2015) ‘MANETs and VANETs clustering algorithms: A survey’, in *2015 IEEE 8th GCC Conference and Exhibition, GCCCE 2015*, pp. 1–6. doi:10.1109/IEEEGCC.2015.7060048.

REFERENCES

Alsuhli, G.H., Khattab, A. and Fahmy, Y.A. (2019) 'Double-head clustering for resilient VANETs', *Wireless communications and mobile computing*, 2019:1-17.

Arshad, M. *et al.* (2018) 'A survey of local/cooperative-based malicious information detection techniques in VANETs', *EURASIP Journal on Wireless Communications and Networking*, 2018(1), pp. 1–17.

Azizian, M., Cherkaoui, S. and Hafid, A.S. (2016) 'A distributed D-hop cluster formation for VANET', in *2016 IEEE wireless communications and networking conference*. IEEE, pp. 1–6.

Bao, X. *et al.* (2020) 'Efficient clustering V2V routing based on PSO in VANETs', *Measurement*, 152, p. 107306.

Chen, Y. *et al.* (2015) 'Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow', *Eurasip journal on Wireless communications and networking*, 2015(1), pp. 1–12.

Cheng, X. and Huang, B. (2019) 'A center-based secure and stable clustering algorithm for VANETs on highways', *Wireless communications and mobile computing*, 2019.pp 10

Contreras, M. and Gamess, E. (2020) 'An algorithm based on VANET technology to count vehicles stopped at a traffic light', *International Journal of Intelligent Transportation Systems Research*, 18(1), pp. 122–139.

Cooper, C. *et al.* (2016) 'A comparative survey of VANET clustering techniques', *IEEE Communications Surveys & Tutorials*, 19(1), pp. 657–

681.

Domingo, M.C. and Reyes, A. (2012) ‘A clean slate architecture design for VANETs’, *Wireless Personal Communications*, 67(2), pp. 315–333.

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

Fan, W. *et al.* (2011) ‘A mobility metrics based dynamic clustering algorithm for VANETs’. p 752 – 756

Fernandes, P. and Nunes, U. (2007) ‘Vehicle communications: A short survey’, *IADIS Telecommunications, Networks and Systems, Lisboa, Portugal*, 1556.

Ge, X., Gao, Q. and Quan, X. (2018) ‘A novel clustering algorithm based on mobility for VANET’, in *2018 IEEE 18th International Conference on Communication Technology (ICCT)*. IEEE, pp. 473–477.

Hadded, M. *et al.* (2017) ‘TDMA-aware routing protocol for multi-hop communications in vehicular ad hoc networks’, in *2017 IEEE wireless communications and networking conference (WCNC)*. IEEE, pp. 1–6.

Hamdi, M.M. *et al.* (2020) ‘A review of applications, characteristics and challenges in vehicular ad hoc networks (VANETs)’, in *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. IEEE, pp. 1–7.

Hasan, M. and Sarker, O. (2020) ‘Routing protocol selection for intelligent transport system (ITS) of VANET in high mobility areas of Bangladesh', in *Proceedings of international joint conference on computational intelligence*. Springer, pp. 123–135.

Hasson, S.T. and Abbas, A.T. (2021) 'A Clustering Approach to Model the Data Dissemination in VANETs', in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*. IEEE, pp. 337–342.

Hasson, S.T. and Hasan, Z.Y. (2017) 'Roads clustering approach's in VANET models', in *2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*. IEEE, pp. 316–321.

Jarupan, B. and Ekici, E. (2011) 'A survey of cross-layer design for VANETs', *Ad Hoc Networks*, 9(5), pp. 966–983.

Kannekanti, S.Y. and Nunna, G.S.P. (2017) 'An Efficient Clustering Scheme in Vehicular Ad-Hoc Networks', in *Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), 2017 IEEE 8th Annual*, pp. 282–287.

Karthikeyan, H. (2021) 'Adaptive Clustering Algorithm for Stable Communication in Vanet', *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(9), pp. 1778–1785.

Katiyar, A., Singh, D. and Yadav, R.S. (2020) 'State-of-the-art approach to clustering protocols in VANET: a survey', *Wireless Networks*, 26(7), pp. 5307–5336.

Khayat, G. *et al.* (2020) 'VANET clustering based on weighted trusted cluster head selection', in *2020 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, pp. 623–628.

Kumar, R. and Dave, M. (2011) 'A comparative study of various routing protocols in VANET', *arXiv preprint arXiv:1108.2094* [Preprint]. 6 pages

Latif, S. *et al.* (2018) 'A comparative study of scenario-driven multi-hop broadcast protocols for VANETs', *Vehicular Communications*, 12, pp. 88–109.

Lee, M. and Atkison, T. (2021) 'Vanet applications: Past, present, and future', *Vehicular Communications*, 28, p. 100310.

Li, M., Zeng, K. and Lou, W. (2011) 'Opportunistic broadcast of event-driven warning messages in vehicular ad hoc networks with lossy links', *Computer Networks*, 55(10), pp. 2443–2464.

Li, X. and Li, H. (2014) 'A survey on data dissemination in VANETs', *Chinese Science Bulletin*, 59(32), pp. 4190–4200.

Lin, C.R. and Gerla, M. (1997) 'Adaptive clustering for mobile wireless networks', *IEEE Journal on Selected areas in Communications*, 15(7), pp. 1265–1275.

Lindeberg, T. (1998) 'Feature detection with automatic scale selection', *International journal of computer vision*, 30(2), pp. 79–116.

Liu, L. *et al.* (2018) 'A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs', *Vehicular Communications*, 13, pp. 78–88.

Lochert, C. *et al.* (2007) 'The feasibility of information dissemination in vehicular ad-hoc networks', in *2007 Fourth Annual Conference on Wireless on Demand Network Systems and Services*. IEEE, pp. 92–99.

Mahdi, M.A. and Hasson, S.T. (2017) 'Grouping vehicles in vehicular social networks', *Kurdistan Journal of Applied Research*, 2(3), pp. 218–225.

Maizate, A. and El Kamoun, N. (2013) 'A new metric based cluster head selection technique for prolonged lifetime in wireless sensor networks', *International Review on Computers and Software*, 8(6), pp. 1346–1355.

Maslekar, N. *et al.* (2011) 'Modified C-DRIVE: Clustering based on direction in vehicular environment', in *2011 IEEE intelligent vehicles symposium (IV)*. IEEE, pp. 845–850.

Misener, J.A., Biswas, S. and Larson, G. (2011) 'Development of V-to-X systems in North America: The promise, the pitfalls and the prognosis', *Computer Networks*, 55(14), pp. 3120–3133.

Najafzadeh, S., Ithnin, N.B. and Abd Razak, S. (2014) 'Broadcasting in connected and fragmented vehicular ad hoc networks', *International Journal of Vehicular Technology*, 2014. 15 pages

Nguyen, V. *et al.* (2015) 'Application of the lowest-ID algorithm in cluster-based TDMA system for VANETs', in *2015 International Conference on Information Networking (ICOIN)*. IEEE, pp. 25–30.

Patel, A. and Kaushik, P. (2018) ‘Improving QoS of VANET Using Adaptive CCA Range and Transmission Range both for Intelligent Transportation System’, *Wireless Personal Communications*, 100(3), pp. 1063–1098. doi:10.1007/s11277-018-5609-5.

Paul, B. and Islam, M.J. (2012) ‘Survey over VANET routing protocols for vehicle to vehicle communication’, *IOSR Journal of Computer Engineering (IOSRJCE)*, 7(5), pp. 1–9.

Pophali, M. and Yengantiwar, T.S. (2014) ‘Trusted Opportunistic Forwarding Model in VANET’, *International Journal of Computer Applications*, 975, p. 8887.

Ramalingam, M. and Thangarajan, R. (2020) ‘Mutated k-means algorithm for dynamic clustering to perform effective and intelligent broadcasting in medical surveillance using selective reliable broadcast protocol in VANET’, *Computer Communications*, 150, pp. 563–568. doi:10.1016/j.comcom.2019.11.023.

Ramli, N.I. and Rawi, M.I.M. (2020) ‘An overview of traffic congestion detection and classification techniques in VANET’, *Indonesian Journal of Electrical Engineering and Computer Science*, 20(1), pp. 437–444.

A. Rasheed, S. Gillani, S. Ajmal, and A. Qayyum, "Vehicular ad hoc network (VANET): A survey, challenges, and applications," in Advances in Intelligent Systems and Computing, pp. 39--51, Springer Singapore, 2017.

Rashid, S.A. *et al.* (2020) ‘Prediction Based Efficient Multi-hop Clustering

Approach with Adaptive Relay Node Selection for VANET.', *J. Commun.*, 15(4), pp. 332–344.

Rashid, S.A., Hamdi, M.M. and Alani, S. (2020) 'An overview on quality of service and data dissemination in VANETs', in *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. IEEE, pp. 1–5.

Rathore, N.P.S., Yadav, A. and Chauhan, A. (2020) 'Multi-Agent based Stable Clustering and Collision Detection In VANET', in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, pp. 1147–1150.

Regin, R. and Menakadevi, T. (2021) 'A novel clustering technique to stop congestion occur vehicular ad-hoc networks using node density based on received signal strength', *Peer-to-Peer Networking and Applications*, 14(4), pp. 2361–2371.

Ren, M. *et al.* (2017) 'A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (VANETs)', *Vehicular Communications*, 9, pp. 233–241.

Ren, M. *et al.* (2021) 'A review of clustering algorithms in VANETs', *Annals of Telecommunications*, 76(9), pp. 581–603.

Rossi, G. V *et al.* (2017) 'Stable clustering for ad-hoc vehicle networking', in *2017 ieee wireless communications and networking conference (wcnc)*. IEEE, pp. 1–6.

Saleem, M.A., Shijie, Z., Sarwar, M.U., Ahmad, T., Maqbool, A., Shivachi, C.S. and Tariq, M., 2021. Deep learning-based dynamic stable cluster head selection in VANET. *Journal of Advanced Transportation*, *2021*.

Saleh, H. and Hasson Aljebori, S. (2019) 'A Survey on VANETs: Challenges and Solutions', *International Journal of Engineering and Technology*, 8, pp. 266–274.

Senouci, O., Aliouat, Z. and Harous, S. (2019) 'MCA-V2I: A multi-hop clustering approach over vehicle-to-internet communication for improving VANETs performances', *Future Generation Computer Systems*, 96, pp. 309–323.

Senouci, O., Harous, S. and Aliouat, Z. (2019) 'A new heuristic clustering algorithm based on RSU for internet of vehicles', *Arabian Journal for Science and Engineering*, 44(11), pp. 9735–9753.

Senouci, O., Harous, S. and Aliouat, Z. (2020) 'Survey on vehicular ad hoc networks clustering algorithms: Overview, taxonomy, challenges, and open research issues', *International Journal of Communication Systems*, 33(11), p. e4402.

Seo, M.S., Lee, S.G. and Lee, S. (2019) 'Clustering-based data dissemination protocol using the path similarity for autonomous vehicles', *Symmetry*, 11(2). doi:10.3390/SYM11020260.

Shahid, M.A. (2019) 'Fixed cluster based cluster head selection algorithm in vehicular adhoc network'. University of Windsor (Canada).

Shanmuga Priya.S, E.V.D.. (2014) 'A Novel Approach for Data Acquisition and Handover Scheme in VANET', IJCSIT International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1443-1446

Shen, X. *et al.* (2014) 'Data dissemination in VANETs: A scheduling approach', *IEEE Transactions on Intelligent Transportation Systems*, 15(5), pp. 2213–2223.

Soleymani, S.A. *et al.* (2020) 'A trust model using edge nodes and a cuckoo filter for securing VANET under the NLoS condition', *Symmetry*, 12(4), p. 609.

Steel, M., Hordijk, W. and Kauffman, S.A. (2020) 'Dynamics of a birth–death process based on combinatorial innovation', *Journal of theoretical biology*, 491, p. 110187.

Sutariya, D. and Pradhan, S.N. (2010) 'Data dissemination techniques in vehicular ad hoc network', *International Journal of Computer Applications*, 8(10), pp. 35–39.

Talib, M.S. *et al.* (2019) 'Clustering based affinity propagation in VANETs: Taxonomy and opportunity of research', *Int. J. Recent Technol. Eng*, 7(6S5). p 672-679

Talib, M.S. *et al.* (2020) 'A Center-Based Stable Evolving Clustering Algorithm With Grid Partitioning and Extended Mobility Features for VANETs', *IEEE Access*, 8, pp. 169908–169921.

Tanuja, K., Sushma, T. M., Bharathi, M., & Arun, K. H. (2015). A Survey on VANET Technologies. International Journal of Computer Applications, 121(18), 1-9.

Tavaré, S. (2018) 'The linear birth–death process: an inferential retrospective', *Advances in Applied Probability*, 50(A), pp. 253–269.

Tomar, P., Chaurasia, B.K. and Tomar, G.S. (2010) 'State of the art of data dissemination in VANETs', *International journal of computer theory and engineering*, 2(6), p. 957.

Vvedenskaya, N.D. *et al.* (2018) 'A Local Large Deviation Principle for Inhomogeneous Birth–Death Processes', *Problems of Information Transmission*, 54(3), pp. 263–280.

Wang, Z. *et al.* (2008) 'A position-based clustering technique for ad hoc intervehicle communication', *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2), pp. 201–208.

Yeferny, T. and Arour, K. (2012) 'Efficient routing method in p2p systems based upon training knowledge', in *2012 26th International Conference on Advanced Information Networking and Applications Workshops*. IEEE, pp. 300–305.

Yeferny, T., Arour, K. and Bouzeghoub, A. (2013) 'An efficient peer-to-peer semantic overlay network for learning query routing', in *2013 IEEE 27th International Conference on Advanced Information Networking and Applications (AINA)*. IEEE, pp. 1025–1032.

Yeferny, T. and Hamad, S. (2020) 'Vehicular Ad-hoc Networks :

Architecture , Applications and Challenges', *International Journal of Computer Science and Network Security*, 20(2), pp. 1–7.

Yeferny, T. and Hamad, S. (2021) 'Vehicular Ad-hoc Networks: Architecture, Applications and Challenges', *arXiv preprint arXiv:2101.04539* [Preprint].

Zhang, D. *et al.* (2018) 'New multi-hop clustering algorithm for vehicular ad hoc networks', *IEEE Transactions on Intelligent Transportation Systems*, 20(4), pp. 1517–1530.

Zhang, Z., Boukerche, A. and Pazzi, R. (2011) *A novel multi-hop clustering scheme for vehicular ad-hoc networks*, *Proc. of the 9th ACM international symposium on Mobility management and wireless access*. doi:10.1145/2069131.2069135.

Appendix A

Published Papers

1: Published paper

Conferences > 2021 1st Babylon Internationa...

A developed Clustering Approach to Model the Data Dissemination Types in a Highway

Publisher: IEEE Cite This PDF

Sura J. Mohammed ; Saad T. Hasson All Authors

Abstract:
One of the special subclasses of ad hoc networks is a Vehicular Ad hoc Network (VANET). VANET aims to enhance and save road traffics. Providing information about traffic, accident, possible deviation, weather information, and dangers are important to the vehicle's driver. VANETs are characterized by high mobility and dynamic topology changing which resulted in frequent communication link failures. One of the more famous methods to solve this problem is the stable clustering approach to reduce network topology changes. This paper introduces a developed clustering algorithm, that aims to build a stable cluster in a Highway environment in terms of many parameters such as traffic type, speed, number of nodes, and transmission range of each vehicle. This paper's clustering method is center base depends on Road Side Units (RSUs) as a supervised approach. A simulation is considered to evaluate the performance of the developed algorithm in terms of three types of data dissemination (i.e., broadcast, multicast and unicast) using Net logo programming as an agent-based modeling approach. In this paper we compared the proposed algorithm with the framework in [1] results after apply its parameters in our scenarios according to the performance metrics, the analytical results show increasing the number of vehicles improves the stability and the staying duration, as well as, the average number of clusters improves to be decreasingly.

Published in: 2021 1st Babylon International Conference on Information Technology and Science (BICITS)

Date of Conference: 28-29 April 2021

INSPEC Accession Number: 21079264

Date Added to IEEE *Xplore*: 12 August 2021

DOI: 10.1109/BICITS51482.2021.9509917

ISBN Information:

Publisher: IEEE

Conference Location: Babil, Iraq

I. INTRODUCTION

One of the special subclasses of ad hoc networks is Vehicular Ad hoc Network (VANET) that aims to enhance and save road traffic, this can be done by providing information a ... ation, weather information, and dangers

BICITS'21

1st Babylon International Conference on Information Technology and Science

CERTIFICATE OF ACCEPTANCE

This certificate is granted to

Sura Jasim Mohammed and Saad Talib Hasson

Certifies the acceptance of the research paper entitled:

A DEVELOPED CLUSTERING APPROACH TO MODEL THE DATA DISSEMINATION TYPES IN A HIGHWAY

in

BICITS'21

Which will be held on 28-29 April, 2021 in Babylon, IRAQ, by College of Information Technology, University of Babylon and Technically Sponsored by IEEE represented by IEEE Iraq Section.

Prof. Dr. Hussain Attia
Dean of IT College, University of Babylon

Prof. Dr. Sattar B. Sadkhan
IT College, University of Babylon, IEEE Iraq Section

A developed Clustering Approach to Model the Data Dissemination Types in a Highway

1st Sura Jasim Mohammed
University of Babylon-Iraq
https://orcid.org/0000-0002-8376-2822

2nd Saad Talib Hasson
University of Babylon-Iraq
saad_aljebori@itnet.uobabylon.edu.iq

Abstract: **One of the special subclasses of ad hoc networks is a Vehicular Ad hoc Network (VANET). VANET aims to enhance and save road traffics. Providing information about traffic, accident, possible deviation, weather information, and dangers are important to the vehicle's driver. VANETs are characterized by high mobility and dynamic topology changing which resulted in frequent communication link failures. One of the more famous methods to solve this problem is the stable clustering approach to reduce network topology changes. This paper introduces a developed clustering algorithm, that aims to build a stable cluster in a Highway environment in terms of many parameters such as traffic type, speed, number of nodes, and transmission range of each vehicle. This paper's clustering method is center base depends on Road Side Units (RSUs) as a supervised approach. A simulation is considered to evaluate the performance of the developed algorithm in terms of three types of data dissemination (i.e. broadcast, multicast and unicast) using Net logo programming as an agent-based modeling approach. In this paper we compared the proposed algorithm with the framework in [1] results after apply its parameters in our scenarios according to the performance metrics, the analytical results show increasing the number of vehicles improves the stability and the staying duration, as well as, the average number of clusters improves to be decreasingly.**

Keywords__VANET; V2V, V2I, clustering, CH election, Data dissemination.

I. INTRODUCTION

One of the special subclasses of ad hoc networks is Vehcular Ad hoc Network (VANET) that aims to enhance and save road traffic, this can be done by providing information about traffic, accident, possible deviation, weather information, and dangers. There are urban, rural, and highway scenarios in VANET. In highway scenarios, the topology is sparser and the connectivity is more intermittent. In urban, the topology is dense [2].

Notable attention had been taken by both industry and academic fields in the issues of considering the fluidity and road traffic because it can provide safety and comfort to the drivers and passengers [3]. The connection between two vehicles is called V2V, the connection between vehicle and roadside unit is called V2I, in addition to the hybrid connection that contains V2V and V2I together. These types are worked based on Wireless Local Area Network (WLAN) technology. There are many important issues in VANET like network stability, scalability, efficient data dissemination, etc. The high speed of vehicles causes a periodic change in network topology, this is the main challenge faced by data dissemination among vehicles.

To deal with all these issues, the clustering technique is one of the most famous solutions to send messages between source and destination in the method that achieves saving the Qos and making performance metrics near to the optimum. A cluster is a set of connected nodes doing together closely for the same reason and has a position with the same topological structure. The clustering algorithms are different from one to the other in clusters formations and selecting the cluster head method, there are many criteria to choose the cluster head that save the performance metrics without degraded [3, 2, 9]

As we mention in VANET there are many communication techniques V2V, V2I, and I2I the infrastructure such as Road Side Unit (RSU) in addition to the application server, it export and import information for entertainment or safe driving or navigation. communication in VANET is categorized into two classes, dedicated short-range communication (DSRC) and long-term evolution (LTE). Due to the DSRC is about 300 meters, it used to communicate the V2V in the range of one cluster. LTE uses to communicate V2I and I2I because the range is larger than DSRC. The combination between DSRC and LTE is useful to decrease overhead via DSRC and congestion via LTE because the connection is between cluster heads and RSU instead of all the nodes and this drops the overhead. In our approach, we combine DSRC in V2V and LTE in V2I together due to the previous reasons. Therefore in the designed scenario, each vehicle has two interfaces one for DSRC and the other for LTE. This architecture can improve the performance metrics but it requires a high stable cluster and a small number of cluster heads to ensure maximum bandwidth utilization, it is a challenge to solve these problems, therefore the main task of this paper is designing an efficient clustering algorithm. [1, 7, 8].

In this paper, we concern with the highway environment due to its challenges and because there is only 10 percent of researchers work in this environment [9]. The rest of this paper is organized as follows: Section 2 reviews related works. In section 3, the clustering theory and process are discussed in detail. In section 4, a methodology is discussed in detail. Section 5 gives the performance metrics and simulation environment. Section 6 summarizes the experimental results and conclusion.

2: Published paper


DUHOK CSASE 2022 International Conference

IEEE

UNIVERSITY OF DUHOK

Certificate of Participation

This is to certify that

Sura Jasim Mohammed
Saad Talib Hasson

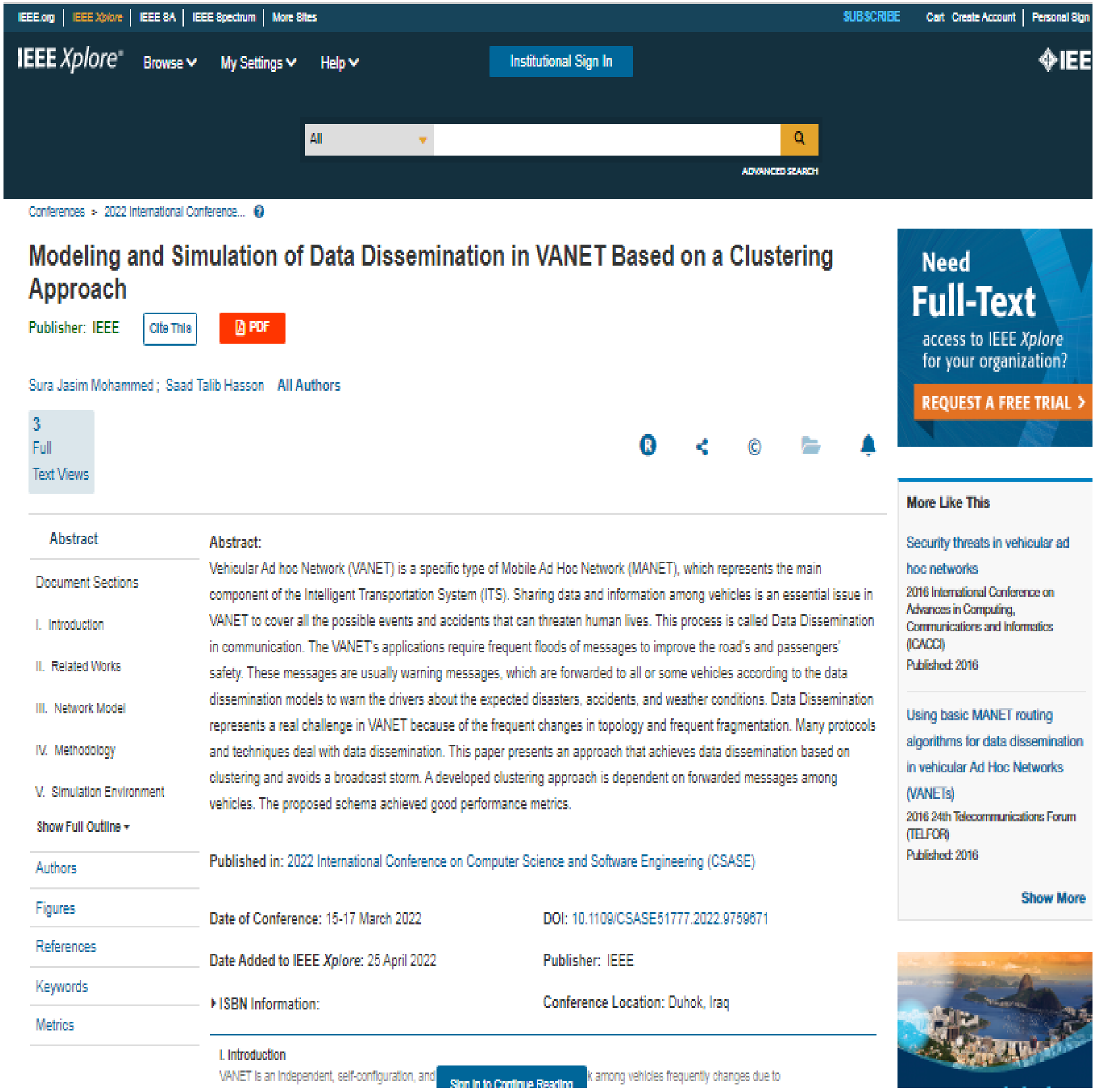
Have participated with a paper entitled

Modeling and Simulation of Data Dissemination in VANET Based on a Clustering Approach

at the 2nd International Conference on Computer Science and Software Engineering **(2022 CSASE)**. This conference is organised by University of Duhok and technically sponsored by IEEE Computer Society.

Dr.Dawood Sulaiman Atroshi
Conference Chairman
President of the University of Duhok

15th - 17th March 2022



Conferences > 2022 International Conference...

Modeling and Simulation of Data Dissemination in VANET Based on a Clustering Approach

Publisher: IEEE Cite This PDF

Sura Jasim Mohammed ; Saad Talib Hasson All Authors

Abstract:
Vehicular Ad hoc Network (VANET) is a specific type of Mobile Ad Hoc Network (MANET), which represents the main component of the Intelligent Transportation System (ITS). Sharing data and information among vehicles is an essential issue in VANET to cover all the possible events and accidents that can threaten human lives. This process is called Data Dissemination in communication. The VANET's applications require frequent floods of messages to improve the road's and passengers' safety. These messages are usually warning messages, which are forwarded to all or some vehicles according to the data dissemination models to warn the drivers about the expected disasters, accidents, and weather conditions. Data Dissemination represents a real challenge in VANET because of the frequent changes in topology and frequent fragmentation. Many protocols and techniques deal with data dissemination. This paper presents an approach that achieves data dissemination based on clustering and avoids a broadcast storm. A developed clustering approach is dependent on forwarded messages among vehicles. The proposed schema achieved good performance metrics.

Published in: 2022 International Conference on Computer Science and Software Engineering (CSASE)

Date of Conference: 15-17 March 2022
DOI: 10.1109/CSASE51777.2022.9759671
Date Added to IEEE *Xplore*: 25 April 2022
Publisher: IEEE
ISBN Information:
Conference Location: Duhok, Iraq

I. Introduction
VANET is an independent, self-configuration, and [Sign In to Continue Reading] k among vehicles frequently changes due to

Modeling and Simulation of Data Dissemination in VANET Based on a Clustering Approach

1st Sura Jasim Mohammed
College of Information Technology
University of Babylon
Iraq-Babylon
sura.phd2018@uobabylon.edu.iq

2ndSaad Talib Hasson
College of Information Technology
University of Babylon
Iraq-Babylon
saad_aljebori@itnet.uobabylon.edu.iq

Abstract **- Vehicular Ad hoc Network (VANET) is a specific type of Mobile Ad Hoc Network (MANET), which represents the main component of the Intelligent Transportation System (ITS). Sharing data and information among vehicles is an important issue in VANET in order to cover all the possible events and accidents which can threaten human lives. This process is called Data Dissemination in communication. The VANET's applications require frequent floods of messages in order to improve the road's and passengers' safety. These messages are usually warning messages, which are forwarded to all or some vehicles according to the data dissemination models to warn the drivers about the expected disasters, accidents, and weather conditions. Data Dissemination represents a real challenge in VANET because of the frequent changes in topology and frequent fragmentation. There are many protocols and techniques that deal with data dissemination. This paper presents an approach that achieves data dissemination based on clustering, and avoids a broadcast storm. A developed clustering approach is dependent on forwarded messages among vehicles. The proposed schema achieved good performance metrics.**

Keywords - VANET, Clustering, Data Dissemination, Broadcast, Unicast, Multicast.

I. Introduction

VANET is an independent, self-configuration, and temporary Ad hoc network. The link among vehicles changes frequently due to them leaving and joining the VANET [1].

In VANET, there are many embedded components in vehicles that help in data exchange to help drivers in making the suitable decision. Spreading data over wireless networks requires capabilities in the link layer to send the messages across the entire network. The forms of communication used for data dissemination in VANET are Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), and Infrastructure to Infrastructure (I2I) [2]. The broadcasting is applied using many routing protocols to find paths from source to destination [3]. Each node needs to rebroadcast the packet to the other nodes, which causes the blind flooding [4]. In dense networks, the overhead and broadcast storms are increased [5][6] .

The motivations behind the data dissemination in VANET are to flood the vehicles with all information about traffic, weather conditions, and probable disasters to save the humans' lives and avoid accidents [7].

On the other hand, the motivation behind using clustering as a base for data dissemination is to avoid a broadcast storm and enhance the performance metrics. The performance metrics are used to prove how the proposed system can enhance the data dissemination in a highway scenario[8].

Most of the published papers in data dissemination based on VANET depend on the protocols in data dissemination and how to face the broadcast storm problem. As for the approach used in this paper, the principle is different, because the data dissemination depends on clustering as a first stage in order to reduce the overhead and delay in addition to avoiding a broadcast storm problem [8].

In this paper, the proposed system supposes data dissemination in VANET based on clustering approach. The data dissemination is presented in three models: broadcast, multicast, and unicast. The clustering approach organizes the data dissemination with a good performance metric, and the broadcast storm was alleviated in the proposed system. The content of this paper is organized in the following way: Section II states the related works, section III reviews the network model, section IV the Methodology, and section V contains the simulation environment. Section VI contains the results and discussions, and section VII refers to the conclusions and future works.

II. Related Works

The most up-to-date papers that deal with Vehicular Data Dissemination are summarized in the following:

In [1], the proposed system uses the cross-layer to get information about the channel, and they proposed a Signal to Noise Ratio (SNR) to find the effective transmission range to send data. They reduced the broadcast storm problem by minimizing the number of nodes which retransmit the messages. The authors used the path loss model and physical layer as their performance metrics.

In [9], a data dissemination strategy was proposed. The main challenge is to avoid the collision and delay with maximum dissemination when the data was transmitted to the nodes. Their proposed system depends on a scheduling perspective to improve the data dissemination by sharing data with neighbor nodes. The authors used the average downloading delay as a performance metric to evaluate their study.

The authors in [10] proposed a Multi-Criteria-based Relay Election Protocol for Data Dissemination in urban VANETs (MCRE-DDP). The relay node depends on a number of parameters in its selection, such as signal-to-noise ratio, speed, the distance between senders and receivers. The utilized performance metrics are redundancy ratio, link load, and speed.

3: Published Paper

International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 14 (2022) pp. 010-018

Received: 1 August, 2021; Accepted: 15 December, 2021; Publish: 9 April, 2022

A New Method for Predicting Cluster Stability in VANET Based on the Birth-Death Process

Sura Jasim Mohammed[1] and Saad Talib Hasson[2]

[1] *Software Department, College of Information Technology, University of Babylon*
Iraq-Babylon
Phd_sura2018@uobabylon.edu.iq

[2] *Network Department, College of Information Technology, University of Babylon*
Iraq-Babylon
saad_aljebori@itnet.uobabylon.edu.iq

Abstract: **One of the main characteristics of a Vehicular Ad-hoc Networks (VANET) is high mobility. Therefore, the high overhead and end-to-end delay are natural results. The prediction of future movements for vehicles provides traffic observations to improve and enhance the network performance and to serve the VANET applications. The clustering approach is a popular technique used in most studies to deal with the network's dynamic topology. In most studies, the clustering algorithms are tested and evaluated according to the cluster head (CH) selection approach. Selecting a suitable CH will lead to the prolonging of the cluster lifetime. A long life time represents an indication for cluster stability, also known as cluster maintenance. Cluster maintenance has got less consideration by the previous VANET studies. The stability prediction represents a challenging task in the clustering approach. This paper highlights the stability prediction in the VANET's Highway scenario. It proposes a new approach to evaluate the vehicle's cluster stability prediction by utilizing the stochastic process. The modeling of joining and leaving takes place using the birth-death process as a stochastic process for evaluating the stability of the clustering algorithm. The main advantage of this approach is to predict and evaluate the vehicle cluster stability with time, using a new mathematical model.**

Keywords: VANET, clustering, birth-death, stochastic process, stability, cluster maintenance, Markov.

I. Introduction

Vehicle Ad-hoc Networks (VANETs) are networks that can supply information on roads to the drivers in order to make driving safe and easier. It will become widespread in the very near future, creating a notable change to human lives [1]. The enormous safety, suitability, and trade possibilities of VANETs are the causes of VANET's spread. VANETs are simply multi-agent wireless networks that are constructed to solve traffic issues by allowing vehicles to communicate together [2]. It provides wireless communication among vehicles by using a dedicated short-range communication (DSRC), which is essentially improved by IEEE 802.11a, and Long Term Evolution (LTE) [3]. Each vehicle contains an on-board unit (OBU) sensor to connect with the nearby vehicles or Road Side Unit (RSU) if it is the CH according to its coverage area and function. There are many models built to implement, support, and improve VANET [4], [5].

However, there are many issues in VANET, such as efficient data dissemination, network scalability, and stable network topology. In order to study VANET's behavior and evaluate its performance, the clustering technique must be the first step to control its changeable topology and to build a meaningful group of vehicles, as shown in Figure 1. The clustering technique contains three phases: cluster creation, cluster head (CH) selection, and cluster maintenance. Most of the related studies have focused on the second phase (CH selection) of any clustering process, while very few studies considered the last phase (cluster maintenance). The stability of vehicle clusters has a high impact on VANET performance [3]. The motivations behind clusters stability prediction are related to the nature of VANET, such as dynamic changes in topology, high speed, and frequent disconnection in networks. In addition, the cluster head (CH) stability plays a crucial role in network scalability and robustness. The stable clusters refer to good performance metrics and QoS. Most of the researchers focused on CH selection methods to improve the stability of the clustering process. The researchers used the performance metrics to evaluate the cluster stability depending on CHs selection process, CHs duration, and CHs lifetime [4]. In this study, the cluster stability is measured and evaluated depending on the stochastic process (Birth-Death process), using the probability of losing or winning one vehicle in each cluster. The process is modeled using a proposed mathematical model to estimate the stability of clusters by means of a set of assumptions derived from probability [5].

In this new approach, the problem is modeled by estimating the joining and leaving vehicles to/from the clusters. Joining and leaving happen through the third phase (maintenance phase) of each clustering process. In a highway scenario, the frequent disparity among vehicles' speeds causes

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل
كلية تكنلوجيا المعلومات

طريقه عنقدة جديدة لتحسين نشر البيانات في VANET بالاعتماد على معادلة ماركوف

أطروحة

مقدمة إلى مجلس كلية تكنولوجيا المعلومات في جامعة بابل كجزء من متطلبات الحصول على درجة الدكتوراه فلسفة في تكنولوجيا المعلومات – البرمجيات

مقدمة من قبل
سرى جاسم محمد عبد الامير

باشراف
أ.د.سعد طالب حسون الجبوري

هـ ١٤٤٣ ٢٠٢٢م

الخلاصة

في السنوات القليلة الماضية ، تفتح الشبكة المخصصة للسيارات (VANET) اتجاهات بحثية جديدة واسعة. يمثل جوهر نظام النقل الذكي (ITS). تساعد VANET في تزويد المستخدمين بالسلامة والمعلومات الترفيهية. ومع ذلك ، هناك العديد من المشكلات في VANET ، مثل النشر الفعال ل لبيانات ، وقابلية توسيع الشبكة ، وطوبولوجيا الشبكة المستقرة. لذلك أولاً ، لدراسة سلوك VANETوتقييم أدائها ، يجب أن تكون تقنية التجميع هي الخطوة الأولى للتحكم في هيكلها القابل للتغيير وبناء مجموعة ذات مغزى من المركبات. تم افتراض سيناريو الطريق السريع في هذه الرسالة. في نهج التجميع المقترح ، يتم تنظيم المركبات المتحركة في مجموعات غير متداخلة. يتم اقتراح المعايير المحددة لاختيار رئيس الكتلة. بشكل عام ، هناك استراتيجيتان لتشكيل الكتلة: الموزعة (المخصصة) والمركزية. في النظام المقترح لهذه الرسالة ، يتم اختيار استراتيجية التجميع على أساس المركز لإدارة عملية التجميع والتحكم فيها بمساعدة وحدات جانب الطريق .(RSUs) ثانيًا ، تتمثل القضية الأكثر أهمية في التنبؤ بالحركة المستقبلية لبناء نظام موثوق به وتعزيز كل من الاتصال في VANETونظام النقل الذكي. تعتمد كفاءة خوارزمية التجميع على النهج الرياضي الذي يسمى العملية العشوائية. يمكن التنبؤ رياضيًا بانضمام المركبات وتركها إلى المجموعات باستخدام عملية ماركوف. ولكن في هذه الدراسة ، يتم تطبيق التطبيق العملي لملاءمة النهج الرياضي لمفهوم ماركوف عمليًا. لوحظ عدد المركبات في كل مجموعة في أوقات متسلسلة مختلفة لمراقبة مغادرة المركبات والانضمام إليها. أخيرًا ، في VANET ، يمكن نشر المعلومات بين المركبات التي تحتوي على معدات اتصال. تلعب كل مركبة دور المرسل أو الموجه أو المتلقي. يمكن أن تنتشر رسائل التحذير بين المركبات. لذلك ، يتم أيضًا تنفيذ عملية نمذجة نشر البيانات في VANET في ثلاثة أنواع: البث ، والبث المتعدد ، والبث الأحادي ، ويتم تقييمها بناءً على نهج التجميع المقترح في سيناريوهات مختلفة.

أظهرت نتائج التجربة أن أداء خوارزمية التجميع المقترحة كان أفضل من نظام التجميع الآمن المستند إلى المركز (CBSC) من حيث استقرار الكتلة وعدد الكتل التي تم إنشاؤها ، كما تشير المقاييس إلى التحسن في النتائج مع زيادة حجم الشبكة. في نماذج نشر البيانات ، يحقق السيناريو المتناثر مع نشر بيانات البث الأحادي مقاييس أداء جيدة بدلاً من البث والبث المتعدد. كان أداء نموذج نشر البيانات المقترح أفضل من خوارزمية التجميع والبث الاحتمالي. من حيث التأخير ونسبة تسليم الحزمة.