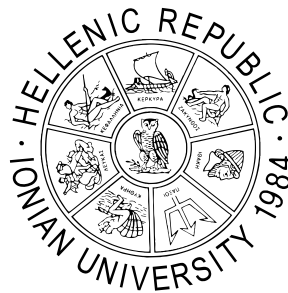


IONIAN UNIVERSITY

**Faculty of Information Science
and Informatics**

Department of Informatics



M.Sc. Thesis

**Development of Clustering Algorithms for
Information Dissemination in Vehicular Networks**

Spiridon Vergis

9 June 2021

Supervisor

Dimitrios Tsoumakos, *Associate Professor,*
Ionian University

Reviewing Committee Members

Dimitrios Tsoumakos, *Associate Professor,*
Ionian University

Konstantinos Oikonomou, *Professor,*
Ionian University

Michael Stefanidakis, *Associate Professor,*
Ionian University

Dedication....

I would like to dedicate this work and manuscript to my family. Their invaluable support and assistance gives me the strength to carry on pursuing my dreams.

Abstract

Over the last few decades many technological advancements of significant importance have been introduced in modern societies. These advancements have made people's everyday lives much easier. With the development of such technological advancements, though, new and more complex issues arise. Some of these issues that modern societies have to cope with, are the efficient management of resources such as fossil fuels or electricity, the phenomenon of overcrowded cities or the mitigation of traffic congestion of the road network. To overcome such issues, the development of advanced and complex applications is required.

With the rise of 5G Networks and the increasing number of proposed solutions derived from the concepts of *Smart Cities* and *Internet of Things (IoT)*, these issues can be handled. The concept of Smart Cities is that of a city that exploits advanced communication technologies to improve the quality of its citizens' lives and allow for better and more advanced administration of its bureaucratic and operational processes. The IoT paradigm is designed to support such applications. It suggests the integration of micro-controllers, sensors, transceivers for digital communication, and suitable protocol stacks into objects of everyday life, and thus making them able to collect data regarding their surroundings and communicate with one another, or with their users. Therefore by making everyday life objects "smart", applications that cope with the aforementioned issues can be deployed.

One of the most challenging environments to deploy such applications is that of the vehicular networks. These networks have unique characteristics, such as high mobility, that make them stand out from other environments. Vehicular networks suffer from issues such as traffic congestion, accidents or road network alterations. Each of these issues has its own requirements. For example, during the traffic congestion, it is crucial for a network to take actions and mitigate the traffic congestion during periods of heavy traffic. In the case of a car accident or a road network

alteration, it is important to spread the information among the vehicles, before they reach the point where the incident takes place. Similarly in the occasion of a sudden deceleration of a vehicle, the rest of the vehicles must be informed within a very small time window, to avoid the cause of an accident.

Vehicular Ad Hoc Networks (VANETs) have the potential to become the viable solution to Intelligent Transportation System (ITS) real-world applications. The high mobility of the vehicles along with the scalability, on the other hand, are the main challenges that the unique environment of vehicular networks has. Mobility-based characteristics can be exploited to overcome this issue, by forming clusters of vehicles with similar attributes within the network. Clustering is a common technique used in many fields, such as statistical data analysis, machine learning or pattern recognition. It is a methodology used to group a set of objects into smaller subsets that contain objects that share common characteristics. The same principle could be utilized in VANETs to divide the vehicular network into smaller groups that share the same interests and attributes.

For many years VANET clustering has been the focus of many researchers. Initial studies considered the clusters to be separate groups of vehicles guided by one vehicle. These studies, though, do not take into account the high mobility of VANETs. Since then, many works have suggested more mobility-resilient approaches, such as multi-hop clustering methodologies or machine-learning clustering algorithms. These approaches aim to increase the duration of a cluster's lifetime. Despite this interest, no one to the best of the author's knowledge has proposed a mobility-based multi-head clustering algorithm for VANETS.

This paper presents a Multi-Head Clustering Algorithm that utilizes mobility-based metrics to form clusters in a vehicular network. The proposed algorithm assumes that a vehicle can belong to more than one cluster. Several simulations are carried out to test the algorithm's effectiveness during different scenarios. For the implementation of the presented algorithm OmnNeT++ is utilized in combination with the Simulation of Urban MObility (SUMO). Finally, the algorithm is compared to a double-head clustering algorithm based on performance evaluation metrics regarding the clusters' lifetime. The results outline that the proposed algorithm achieves a greater cluster lifetime duration. In addition, the average size of the clusters along with the number of clusters each node belonged to, are significantly greater than those extracted by the double-head clustering algorithm.

Περίληψη

Τις τελευταίες δεκαετίες, πολλές τεχνολογικές ανακαλύψεις έχουν καθιερωθεί στην καθημερινότητα των ανθρώπων. Η ενσωμάτωσή τους έχει κάνει την καθημερινότητα των ανθρώπων πολύ πιο εύκολη. Με την καθιέρωση τέτοιων τεχνολογικών ανακαλύψεων, όμως, δημιουργούνται νέες προκλήσεις που χρήζουν επίλυσης. Τα ζητήματα που απασχολούν τις περισσότερες μοντέρνες κοινότητες σήμερα έχουν να κάνουν με μία πληθώρα κατηγοριών, όπως ο διαμοιρασμός των ηλεκτρικών πόρων μιας πόλης ή το κυκλοφοριακό ζήτημα του οδικού δικτύου της. Στην σημερινή πραγματικότητα που χαρακτηρίζεται από πολυπλοκότητα και συνεχής αλλαγής η ανάγκη για μοντέρνες και αποδοτικές λύσεις στα προβλήματα αυτά, είναι αρκετά μεγάλη.

Με την εμφάνιση των 5G Δικτύων και την εμφάνιση πολλαπλών ερευνών βασισμένων στην ιδέα των *Έξυπνων Πόλεων (Smart Cities)* και του *Ίντερνετ των Πραγμάτων (Internet of Things)*, παρατηρείται πως η επίλυση αυτών των ζητημάτων μπορεί να επέλθει από την χρήση αυτών των τεχνολογιών. Οι Έξυπνες Πόλεις, σαν ιδέα, αποτελούν εκσυγχρονισμένα αστικά κέντρα που χρησιμοποιούν την τεχνολογία της πληροφόρησης και της επικοινωνίας για να βελτιώσουν τις συνθήκες διαβίωσης των ατόμων που ζουν σε αυτές, καθώς και να βελτιώσουν την επίβλεψη διάφορων γραφειοκρατικών και λειτουργικών υπηρεσιών. Η ιδέα του Ίντερνετ των Πραγμάτων περιγράφει την ενσωμάτωση μικρο-υπολογιστών, αισθητήρων, κεραιών για επικοινωνία και των αντίστοιχων πρωτοκόλλων επικοινωνίας, σε συσκευές που χρησιμοποιούνται καθημερινά από ανθρώπους. Με αυτό τον τρόπο θα είναι ικανά να συγκεντρώσουν μετρήσεις σχετικές με το περιβάλλον τους και να επικοινωνήσουν με άλλες τέτοιες συσκευές ή με τους χρήστες τους.

Ένα από τα πιο περίπλοκα περιβάλλοντα για την ανάπτυξη τέτοιων εφαρμογών είναι εκείνο των οδικών δικτύων. Τα οδικά δίκτυα διαθέτουν ιδιαίτερα χαρακτηριστικά, όπως η υψηλή κινητικότητα των κόμβων. Επιπρόσθετα, αντιμετωπίζουν μια πληθώρα από ζητήματα, όπως αυτό της κυκλοφοριακής συμφόρησης, των αλλαγών του οδικού

δικτύου, των αυτοκινητιστικών ατυχημάτων και αρκετών ακόμα. Το καθένα από αυτά έχουν και τις αντίστοιχες απαιτήσεις για την επίλυσή τους. Για παράδειγμα, στην περίπτωση της κυκλοφοριακής συμφόρησης, είναι σημαντικό να διαδοθεί η πληροφορία στο δίκτυο πριν φτάσουν στο σημείο που λαμβάνει χώρα το γεγονός αυτό. Αντίστοιχα, στην περίπτωση της απότομης ελάττωσης της ταχύτητας του, ένα όχημα πρέπει να προειδοποιήσει άμεσα τα υπόλοιπα οχήματα που το ακολουθούν για να αποφύγει την πρόκληση αυτοκινητιστικού ατυχήματος.

Τα ad hoc δίκτυα οχημάτων (Vehicular Ad Hoc Networks) έχουν την ικανότητα να αποφέρουν λύσεις σε τέτοια περιβάλλοντα με την ανάπτυξη εφαρμογών έξυπνων συστημάτων μετακίνησης (Intelligent Transportation Systems). Η υψηλή κινητικότητα των οχημάτων, καθώς και το ζήτημα της επεκτασιμότητας τέτοιων εφαρμογών, αντίθετως, είναι δύο από τις βασικές προκλήσεις που τα ιδιαίτερα περιβάλλοντα των οδικών δικτύων πρέπει να επιλύσουν. Η χρήση δεδομένων, σχετικών με την κινητικότητα των οχημάτων, μπορεί να βοηθήσει στην επίλυση των προκλήσεων αυτών, δημιουργώντας συστάδες οχημάτων που διαθέτουν κοινά δεδομένα. Η συσταδοποίηση αποτελεί μία κοινή τεχνική ομαδοποίησης αντικειμένων ή/και δεδομένων σε ομάδες που διαθέτουν τα ίδια ή παρόμοια χαρακτηριστικά. Πολλοί κλάδοι χρησιμοποιούν την τεχνική της συσταδοποίησης, όπως η στατιστική ανάλυση δεδομένων, η αναγνώριση μοτίβων ή/και η μηχανική εκμάθηση. Η ίδια τεχνική μπορεί να εφαρμοστεί στα οδικά δίκτυα για να τα χωρίσει σε μικρότερες ομάδες οχημάτων που μοιράζονται τα ίδια ή/και παρόμοια χαρακτηριστικά.

Η μελέτη της συσταδοποίησης των οδικών δικτύων, αποτελεί το βασικό θέμα έρευνας πολλών επιστημόνων. Οι πρώτες μελέτες, όμως δεν υπολογίζουν και την υψηλή κινητικότητα των οδικών δικτύων. Έκτοτε, πολλές έρευνες έχουν μελετήσει την συσταδοποίηση των οδικών δικτύων, αναπτύσσοντας τεχνικές πιο ανθεκτικές στην υψηλή κινητικότητα των οχημάτων. Αυτές οι έρευνες έχουν ως στόχο την αύξηση χρόνου ζωής κάθε συστάδας του οδικού δικτύου. Παρόλα αυτά, καμία έρευνα, εις γνώση των συγγραφέων, δεν έχει προτείνει την χρήση αλγορίθμου συσταδοποίησης όπου το κάθε όχημα μπορεί να ανήκει σε πολλαπλές συστάδες.

Η παρούσα έρευνα παρουσιάζει έναν Αλγόριθμο Συσταδοποίησης Οχημάτων με Πολλαπλές Συστάδες που χρησιμοποιεί δεδομένα από την κινητικότητα των οχημάτων για την δημιουργία συστάδων στο δίκτυο. Ο αλγόριθμος αυτός προϋποθέτει πως το κάθε όχημα μπορεί να ανήκει σε περισσότερες από μία συστάδες. Για την δοκιμασία της αποδοτικότητας του αλγορίθμου, πολλαπλές προσομοιώσεις διεξάγονται. Για

την υλοποίηση των προσομοιώσεων, χρησιμοποιείται το OmNeT++ σε συνδυασμό με το εργαλείο οδικών προσομοιώσεων SUMO. Τέλος ο προτεινόμενος αλγόριθμος συγκρίνεται με αντίστοιχο αλγόριθμος της πρόσφατης βιβλιογραφίας, βασιζόμενος σε μετρικές σύγκρισης που εκμαιεύονται από την βιβλιογραφική ανασκόπηση. Τα αποτελέσματα δείχνουν πως ο προτεινόμενος αλγόριθμος επιτυγχάνει μεγαλύτερο χρόνο ζωής της συστάδας. Επιπρόσθετα, ο μέσος αριθμός από συστάδες και ο αριθμός των συστάδων που ανήκει το κάθε όχημα, είναι ιδιαίτερα μεγαλύτερα από αυτά που του αλγορίθμου της βιβλιογραφίας

Acknowledgments

This manuscript represents an important milestone of my life, both as academic and personal achievement. For this reason, I find it necessary to thank, for their priceless assistance, all the people that helped me.

First and foremost I would like to express my utmost gratitude and respect to my supervisor Prof. Dimitios Tsoumakos, and Prof. Konstantinos Oikonomou. Their guidance, helpful comments, and feedback are some of the main reasons that helped this work progress and come to fruition. In addition, I would like to thank all the members of the reviewing committee, that took the time to read and evaluate my thesis.

I would also like to thank my tutors for their valuable guidance throughout my studies and for providing me with the tools I needed to successfully complete my thesis. Furthermore, I would like to thank my classmates Konstantinos, Alexandros, Stelios, Korina, Andrianna, and Giannis for their invaluable support that helped me endure many trials and tribulations. I would also like to express my appreciation to Dr. Athanasios Tsipis for the support and useful feedback he provided for the development of this work.

Last but not least, I would like to thank my parents for always being there for me. I could not have completed this dissertation without their support and patience.

Contents

A	Introduction	1
A.1	Importance of VANETs	2
A.2	The Main Challenge	4
B	VANETs: Existing Systems and Challenges	7
B.1	Concept and background	7
B.2	VANET Clustering Algorithms	8
B.3	Performance Evaluation Approaches	10
C	Multi-Head Clustering Algorithm	13
C.1	Overall System Description	13
C.2	Principle and Algorithm	16
C.2.1	INIT State	16
C.2.2	ND State	17
C.2.3	Cluster Formation	17
C.2.4	Re-clustering Procedure	18
C.2.5	Cluster's life-cycle	20
C.2.6	Cluster Reforming Overhead Reduction	20
C.2.7	Competence Score	21
D	Performance Evaluation	23
D.1	Simulation Tools and Scenarios	23
D.1.1	SUMO and Veins	23
D.1.2	Omnet++	24
D.1.3	Simulations Scenarios	24

D.2	Evaluation Metrics	26
D.2.1	CH lifetime	26
D.2.2	CM lifetime	26
D.2.3	Alone lifetime	26
D.2.4	Cluster Size	26
D.3	Double-Head Clustering Algorithm Comparison	27
D.4	Simulation Results	27
E	Summary and Conclusions	37
I	References	40
I	Additional Images	51
II	Glossary	53
III	Notations	54
IV	Abbreviations	55

List of Figures

A.1	VANET topology	3
C.1	Vehicle State Diagram.	14
C.2	The types of messages transmitted in the network.	15
C.3	Multi-Head Clustering Algorithm Flowchart.	16
C.4	Message exchange during the INIT state.	19
C.5	Example of the cluster forming.	22
D.1	Omnet++ Simulation.	24
D.2	The average time a vehicle spends time not being part of a cluster.	28
D.3	The average cluster lifetime.	29
D.4	The average time a vehicle operated as a CM.	30
D.5	The average cluster sizes.	32
D.6	The average number of CHs every CM spotted.	33
D.7	The average number of CNs a vehicle had.	34
I.1	Map of Corfu Town. This is the road network utilized in this work.	52

List of Algorithms

1	The Cluster Formation Algorithm.	18
---	--	----

Chapter A

Introduction

AS the number of vehicles increases annually at a vast rate, the risk of serious incidents such as accidents or traffic jams have the potential to manifest easier. Furthermore, the dissemination of information, regarding such incidents, around the vehicular network is crucial for modern societies. Based on the research carried out by INRIX Industries [1], which analyzes global road traffic data, during the year 2018 the annual cost of traffic congestion (per driver) ranges between 350 to 2500 euros. Therefore, to overcome such issues, the need for smarter vehicular applications is grave [2]. With the introduction of 5G networks [3] concepts that could in theory provide viable solutions to serious issues such as that of air pollution and increased traffic congestion [4, 5, 6], have the potential to become real-world applications.

Vehicular Networks and especially Vehicular Ad Hoc Networks (VANETs) [7, 8, 9], are a very promising technology that has the potential to increase traffic safety and efficiency, and to enable numerous other applications in the domain of vehicular communication. VANETs appear to have some similar characteristics with Mobile Ad Hoc Networks (MANETs) [10, 11]. On the other hand, VANETs have specific attributes that make them unique. The high speeds of the vehicles in the networks mean that the topology of VANETs keeps changing rapidly. Furthermore, since these nodes move at such high speeds with wide availability of choices of paths to follow, the connections in the network keep disconnecting [12]. These attributes are the reason that generic applications, designed for MANETs, cannot be applied directly to VANETs. Figure A.1 depicts the architecture of a VANET. VANET architecture

mainly consists of vehicles (V), Road Side Unit (RSU) and Infrastructure Domain (I). Most clustering techniques form clusters that consist of a Cluster Head (CH) and multiple Cluster Members. The role of the CH in most cases is to be the leading node of the cluster, which usually spreads the information to the rest of the CMs. In several cases, Cluster Gateways are introduced in each cluster. These nodes work as communication media between clusters.

The addition of short-range radios to both vehicles and highway infrastructure has the potential to significantly enhance the driving experience [13], providing increased safety and convenience [14]. From a safety perspective, a car that informs other drivers of its sudden deceleration could prevent an accident entirely [15, 16].

A.1 Importance of VANETs

The goal of a VANET architecture is to allow communication among nearby vehicles and between vehicles and fixed roadside equipment leading to the following three possibilities [17]:

- *Vehicle-to-Vehicle (V2V) Communication*: allows the direct vehicular communication without relying on a fixed infrastructure support and can be mainly employed for safety, security, and dissemination applications
- *Vehicle-to-Infrastructure (V2I) Communication*: allows a vehicle to communicate with the roadside infrastructure mainly for information and data gathering applications
- *Hybrid Communication*: combines both Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I). In this scenario, a vehicle can communicate with the roadside infrastructure either in a single hop or multi-hop fashion, depending on the distance, i.e., if it can or not access directly the roadside unit. It enables long distance connection to the Internet or to vehicles that are far away [18].

In VANETs, each vehicle is equipped with the technology that allows the drivers to communicate with each other as well as with the roadside infrastructure, e.g., base stations are also known as Road Side Units (RSUs), located in some critical sections of the road, such as traffic lights, intersections, or stop signs, to improve the

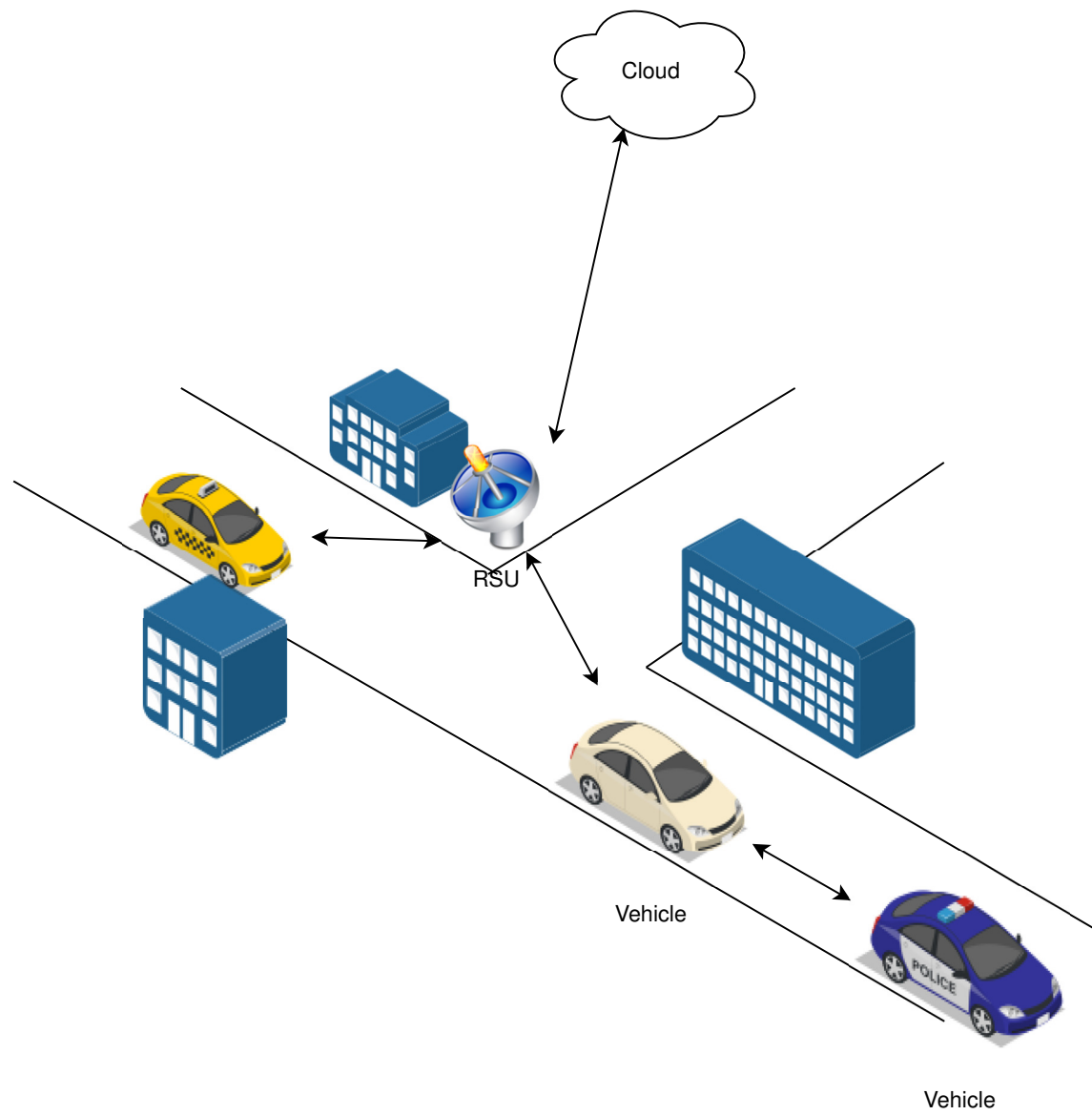


Figure A.1: VANET topology

driving experience and make driving safer. By using such communication devices, also known as On-Board Units (OBUs), vehicles can communicate with each other as well as with RSUs[19].

As mentioned above, VANETs can be utilized for a plethora of applications, whether it is road-related or network-related. When it comes to time-sensitive information such as road blockades or accidents, it is important to transmit the information in time to the nodes that will find it useful. Such applications aim to optimize the transmission of such messages, since receiving one in time is important.

Another important aspect of VANETs is the implementation of a mechanism to guarantee the security of such networks, especially against malicious attacks [20]. These attacks could be coming from malicious nodes that, either aim to disrupt the VANET's integrity by pretending to be a part of it and broadcasting false information or by collecting sensitive data regarding its neighbors.

A.2 The Main Challenge

Based on the aforementioned it is obvious that there are a plethora of matters that need to be sorted out to boost the efficiency of VANET clustering. However one of the main challenges of such networks is the lifetime duration of the clusters.

The main reason behind the clustering of a VANET is the efficient V2V and V2I communication. Therefore, a cluster must stay active long enough for its members to exploit its existence and exchange information. Initial studies that aimed in the proposal of a VANET clustering algorithm overlooked that assumption, proposing clustering algorithms that do not have a long lifetime duration [21, 22].

Since then, many researchers have proposed several methodologies to lengthen the lifetime duration of the clusters. The use of mobility-based parameters provides the simplest and fastest way to form clusters since no further calculations are required.

Another promising methodology is the integration of machine learning techniques in the aforementioned approaches to improve the stability of the clusters. These techniques utilize the mobility-based parameters to estimate the cluster's lifetime or take actions to lengthen it. These approaches on the other hand have higher computational needs and need more time to train the model they use. Furthermore, in the case of false information (e.g. malicious attacks), the trained model might

worsen the algorithm's performance.

Another promising solution is the use of multi-hop clustering techniques. In this case, clusters contain nodes that can be more than one hops away from them. Such applications could also improve the VANET's stability since fewer clusters are formed and each cluster's range can extend further than its one-hop vicinity. Forming such clusters, though, means that the end-to-end delay of the clusters is increased along with the message overhead. Therefore such approaches are not ideal due to these issues.

A novel approach in the field of VANET clustering is the use of multi-head clustering approaches. Few works have suggested the use of multiple CHs in a cluster to preserve the cluster formation even when the main CH loses the connection with the rest of the CMs. However, most of the works that propose such methodologies, suggest the use of multiple CHs within one cluster. So far, to the author's knowledge, no previous works have suggested the integration of a more abstract technique where a vehicle can be part of more than one cluster.

The algorithm presented in this work adopts the multi-head clustering approach presented in the previous paragraph. The main purpose of this work is to boost the cluster's lifetime and its efficiency by distributing information to as many interested vehicles as possible. The presented algorithm is put to the test using several networks and vehicular network simulation tools. The proposed algorithm is able to surpass the performance of other similar algorithms found in the literature.

The remainder of the manuscript is organized as follows: In Chapter B a complete past literature review is presented. This chapter consists of Section B.1 that describes the concept of VANETs, Section B.2 that outlines and categorizes the clustering approaches presented in previous works, and Section B.3 that presents the performance evaluation metrics utilized in most works to test the performance of a clustering algorithm. Chapter C analyzes the proposed Multi-Head Clustering Algorithm. It consists of Section C.1 where the overall system description is presented and Section C.2 which presents a detailed description of the algorithm. In Chapter D, the evaluation of the performance of the proposed algorithm is presented. More specifically, Section D.1 presents the simulation tools utilized in this work, Section D.2 outlines the parameters exploited to measure the proposed methodology's performance, Section D.3 presents the algorithm to which the proposed approach is compared and Section D.4 shows the results of the simulation runs. Finally,

Chapter E draws conclusions regarding the proposed approach and the tests along with guidelines for future works

Chapter B

VANETs: Existing Systems and Challenges

The issue of traffic congestion has increased exponentially over the last decades. The rapid rise of car manufactures annually along with the availability and drop of cost means that this issue will continue to grow. unless actions are taken against it. Concepts such as those of Smart Cities and the Internet of Things (IoT) have the potential to cope with such issues [23, 24].

B.1 Concept and background

Even though the concepts of Smart Cities and IoT applications have been the main research subject of many studies over the last decades, generic solutions derived from them cannot be applied to the unique environment of vehicular networks, meaning that it has its own special challenges and requirements. Vehicular networks display high mobility and therefore connections among the nodes tend to change rapidly.

The concept of Mobile Ad hoc Networks (MANETS) [25] shares some characteristics with the vehicular networks. One of the first proposals to apply mobile ad hoc network fundamentals on vehicular networks has been presented by H Zhu et. al. [26]. Such applications over the years became known as Vehicular ad hoc Networks (VANETs). VANETs create networks in a vehicular network for information dissemination.

B.2 VANET Clustering Algorithms

The development of clustering algorithms for information dissemination in Vehicular Ad hoc Networks (VANETs) started around the first decade of the 21st century with works such as those of T.D.C. Little and A. Agarwal [21]. Using flooding as a message dissemination technique might be simple but increases the message overhead exponentially [27]. Therefore many researchers have proposed the use of grouping techniques so that vehicles with the same interest belong in the same group.

The most popular grouping technique in VANETs is Clustering [28]. In VANETs, clustering algorithms fall into two (2) categories, one-hop, and multi-hop.

One-hop clusters are formed only by the Cluster Head (CH) and the Cluster Members (CMs) that are one-hop away. Many researchers have proposed the use of mobility-based metrics to form clusters in VANETs. In their work C. Shea et al. propose a VANET clustering method that exploits each vehicle's relative mobility to form clusters [22]. Similarly, W. Fan et al. present a VANET clustering algorithm that utilizes the average velocity and average acceleration of each vehicle to find the node suited for CH [29]. Another mobility-based algorithm for VANETs is proposed by Ni et al. using the Doppler effect and predicting the relative speed of the vehicles [30]. A mobility-based clustering method using the moving direction of the vehicles along with the link lifetime estimation is presented by M. Ren et al. [31].

Apart from exploiting only mobility-based parameters of a VANET, past literature has presented the use of machine learning algorithms to form clusters. A fuzzy-logic clustering algorithm is presented by K.A. Hafeez et al. in [32]. Their algorithms exploit each vehicle's relative speed and distance from its neighbors and utilizes a machine learning mechanism to predict the future speed and position of the vehicles and maintain the clusters. Noteworthy are the works of V.E. Ballas and L. Jain and of S. Sharma and A. Kaul [33, 34]. In [33] the authors propose a fuzzy-logic methodology based on the Artificial Bee Colony approach. S. Sharma's and A. Kaul's methodology utilizes a machine-learning algorithm to form clusters and fuzzy logic to boost the stability of the clusters.

K-means algorithms are frequently utilized for Wireless Sensor Networks (WSNs) [35, 36]. Because of the peculiar nature of vehicular networks, these algorithms cannot be applied directly to VANETs. Several studies have proposed k-means-based algorithms to form clusters. In their work I. Hussain proposes a Cluster Head

Selection approach separated into two parts. During the first part the algorithm provides division parts to form vehicle groups using k-means and then it utilizes the Floyd-Warshall Algorithm to calculate the pair of all shortest paths for each vehicle in those vehicle groups [37]. R. Chai et al. propose a k-harmonic means clustering algorithm called Adaptive K-Harmonic Means Clustering Algorithm [38]. Their algorithm utilizes spatial parameters, such as relative distance and velocity between nodes and centroids, and the bandwidth of candidate CHs.

Hierarchical clustering has also been proposed by many researchers as a viable solution to form clusters in VANETs. P Bhosale and A. Vidhate propose a clustering algorithm for VANETs using agglomerative hierarchical clustering [39]. In their work, the CHs are selected based on the amount of time each node acted as a CH in the past.

Many works propose the concept of multi-head clustering in which a CM can have more than one CH. G.H. Alsuhi et al. propose a Double-Head Clustering Algorithm for VANETs [40, 41]. Their method uses mobility-based parameters and each cluster contains a secondary CH in case the connection with the main CH is lost. A similar approach is presented in S.C. Lo's et al. work [42]. The main difference though is that in Lo's et al. work, each cluster can have one main CH and multiple secondary CHs. X. Duan's et al. work proposes an SDN enabled Double-Head Clustering Algorithm in which each cluster consist of the main CH and a backup CH [43, 44].

Multi-hop clustering can reduce the number of clusters since each cluster can cover a wider area, and provide greater stability. The reason behind that is that instead of creating new clusters, vehicles can join clusters that have already formed. On the other hand, the end-to-end delay and packet loss increase [45]. Y. Chen's et al. work presents a distributed multi-hop algorithm [46]. In their work, each vehicle takes into consideration the relationship among the other vehicles in its neighborhood to choose a CH. In this distributed manner, multi-hop clusters are formed. As outlined by E. Dror et al. their multi-hop hierarchical clustering algorithm can create two-hop clusters without the use of GPS [47]. A mobility-based multi-hop clustering algorithm for VANETs is presented in S. Ucar's et al. work [48]. The average relative speed among n-hop neighbors is calculated and the CH is elected based on the mobility.

More recent evidence suggests that the use of a reputation-based algorithm along

with the utilization of the Firefly Algorithm [49] [50]. The reputation of each node is calculated using the vehicles' spatial metrics along with the duration each served as a CH. Another recent work utilizes linear regression to form clusters as a supervised machine learning technique [51].

Based on the aforementioned it is concluded that the majority of the past related works focuses mostly on preserving the CH connectivity with the rest of the CMs, either by adding multiple CHs in one cluster or expanding its range further than one hop away. No works, to the best of the author's knowledge, focus mostly on preserving the CMs of the VANET in their current state by allowing them to be part of multiple clusters.

The work presented in this manuscript proposes a Multi-Head Clustering Algorithm using mobility-based metrics to form clusters. This methodology aims to increase stability and reduce cluster reforming overhead.

B.3 Performance Evaluation Approaches

Based on the aforementioned, it is with no doubt that over the years, a plethora of VANET clustering methodologies has been proposed. In order to compare these works, though, and calculate their performance a set of common evaluation metrics has to be utilized. R Kaur's et al. survey paper presents a comprehensive list of the evaluation parameters utilized in VANET clustering works [52].

Many researchers test the performance of such algorithms using network performance evaluation metrics, such as the packet delivery ratio and message overhead [53]. Similarly, other works have used metrics, such as the end-to-end delay, the average throughput, and the connectivity in addition to the aforementioned metrics [54, 27, 37]. The majority of the proposed methodologies, use cluster-based metrics to evaluate their performance. These metrics consist of i) the CH duration, ii) the CM duration, iii) CH Change rate, iv) the duration a vehicle did not belong to any cluster, and v) the average cluster size [48, 40, 55, 22, 31, 32].

Table B.1: Performance Evaluation Metrics.

Evaluation Metric	Works
Packet Delivery Ratio	[53, 54, 27]
Message Overhead	[53, 54, 27]
End-to-End Delay	[54, 27]
Average Throughput	[54, 27]
CH duration	[48, 40, 55, 22, 31, 32, 37]
CM duration	[48, 40, 55, 22, 32]
CH Change rate	[48, 40, 55, 22, 31]
Duration not in a cluster	[48, 40, 55, 22, 31]
Average Cluster Size	[32]
Connectivity	[37]

A detailed list(CNs of performance evaluation metrics used in some of these works is presented in Table B.1. As shown in the table above, several works focus on the use of network performance metrics to calculate the performance of VANET clustering algorithms. For example, many researchers have used the Packet Delivery Ratio to calculate the effectiveness of their methodology. This represents the number of messages that were transmitted successfully within the cluster. Apart from the effectiveness, these works also measure the Message Overhead in these clusters. That way the efficiency of their methodologies can be measured. Some of these studies have also measured the End-to-End Delay within the clusters. This is very crucial, especially for multi-hop clustering algorithms, that always suffer from worse End-to-End delay.

As for the cluster-based performance evaluation metrics, readings such as the lifetime of the clusters or their size can also provide accurate results for their stability and efficiency. Most works give a lot of emphasis on the duration of a CH's lifetime. That is because the CHs are the main primary unit of each cluster and therefore, their time as a CH also represents the cluster's lifetime. It is also important, though, to take into consideration the time each node operated as a CM since these nodes are also part of the clusters. A clustering algorithm is not effective if the CMs keep losing their connection with the CH, because that means that they are not communicating for long enough to exploit the cluster formation. The average cluster

size will represent the overall impact a cluster had over its neighborhood. The greater the average number, the greater the influence of the cluster on the surrounding nodes. As for the duration, a vehicle did not belong to any clusters, it could be affected by many things, such as the time it takes for an algorithm to start forming clusters, the message overhead and/or the possible lost transmissions in the network, as it might occur in wireless communications.

The performance evaluation metrics used in this work focus on the cluster's stability, by measuring the CH duration, the CM duration, the cluster size, and the time a node is left without a cluster. These parameters are utilized by the majority of the works regarding this particular topic and according to the author they are appropriate to compare the proposed algorithms to similar algorithms found in the literature.

Chapter C

Multi-Head Clustering Algorithm

The algorithm presented in this manuscript considers V2V communications. No Road Side Units (RSUs) or additional infrastructure is utilized for message exchange or transmission. Each vehicle is able to communicate with every other vehicle within its range either through broadcasting a message to all of them or by sending a unicast message to a specific one. Since the system exploits mobility-based parameters to form clusters, it is assumed that each vehicle is equipped with a Global Positioning System (GPS). It is also presumed that each vehicle is equipped with a dedicated-short range communications (DSRC) antenna to communicate with the other vehicles. Carrier access is managed by CSMA/CA. In addition, it is assumed that failed transmissions could take place in the networks since the vehicles are communicating with one another, wirelessly.

C.1 Overall System Description

The proposed methodology assumes that each vehicle in the vehicular network can belong to more than one cluster. That way, the vehicles have better chances of maintaining communication with a cluster and the cluster formations are less prone to constant changes.

Vehicle States Upon the beginning of the algorithm, the vehicle is set to the Initialization State (INIT) . During this state, the vehicle does not have any knowledge of the vehicles around it. The next state is called Neighborhood Discovery State

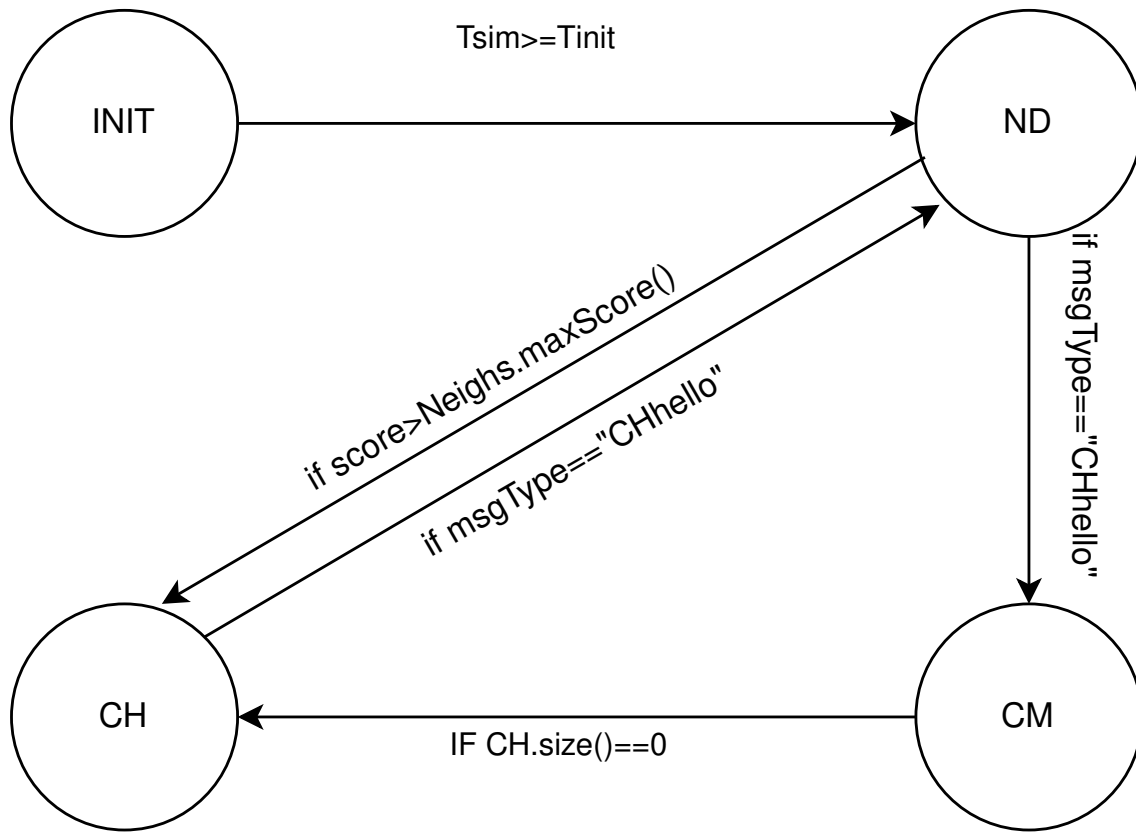


Figure C.1: Vehicle State Diagram.

(ND) during which, the node collects the messages from its neighboring vehicles to form a cluster. The final two states are the Cluster Head State (CH) and the Cluster Member State (CM). The vehicle is set to the former state when it discovers that it is the best candidate in its neighborhood to become a CH and the latter state is for the vehicles that receive a message from a CH and are not a CH themselves. The state diagram of this algorithm along with all possible transitions is depicted in Figure C.1.

Messages and Packets The communication between the vehicles is achieved either by broadcasting messages to one-hop neighbors or by unicasting a message to a specific neighbor. These messages are also used by the vehicles as a way to verify the consistency of a neighbor. There are three types of messages utilized in this work as shown in Figure C.2. The main purpose of the INIT messages is

INIT Message

Node ID	Message Type	Speed	Acceleration	Position (x,y)	Heading Direction
---------	--------------	-------	--------------	----------------	-------------------

CH Hello Acknowledgement Message

Node ID	Message Type	Speed	Acceleration	Position (x,y)	Heading Direction	Receiver Node ID
---------	--------------	-------	--------------	----------------	-------------------	------------------

Neighborhood Discovery Message

Node ID	Message Type	Speed	Acceleration	Position (x,y)	Heading Direction	Score
---------	--------------	-------	--------------	----------------	-------------------	-------

Figure C.2: The types of messages transmitted in the network.

to establish a neighborhood for each vehicle and exchange mobility-based metrics. These metrics consist of i) the speed of the sender, ii) its acceleration, iii) its position (in coordinates) at the time of transmission, and the relative direction in which it is headed. The direction that the vehicle is heading is presumed to be a vector for mathematical reasons that will be further explained in Subsection C.2.7.

The Neighborhood Discovery Message contains all these metrics along with the Competence Score (CS) of the sender. The vehicle with the greatest CS becomes the CH. When a node comes to the conclusion that it is the best candidate for CH it starts sending CHhello messages, which are exactly the same as the INIT messages with the only difference being that the Message Type is set to CHhello message.

The CHHello Acknowledgment Message is the unicast message that a vehicle sends to a CH to verify that it received its CHhello message and it will join its cluster. For that reason, the CH Hello Acknowledgment Message contains also the ID of the receiver node.

An acknowledgment message is sent after an INIT or CHhello message is received, to verify the connection between two nodes, for both vehicles. These messages contain the same mobility-based metrics as the INIT message but they are sent as a unicast message.

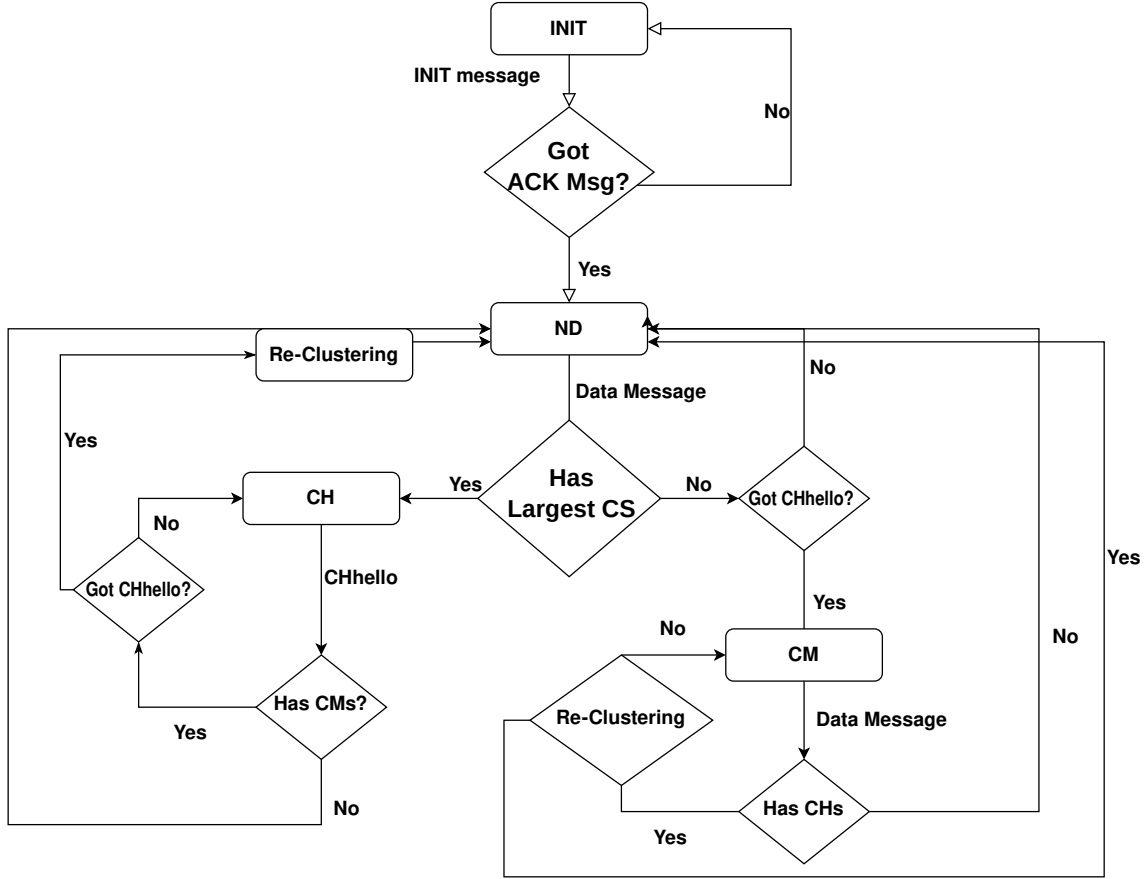


Figure C.3: Multi-Head Clustering Algorithm Flowchart.

C.2 Principle and Algorithm

The algorithm begins immediately once a vehicle starts moving on the road network in a distributed manner, meaning that each vehicle runs the algorithm independently. The flowchart of the algorithm is depicted in Figure C.3.

C.2.1 INIT State

The algorithm is run in a distributed manner. When a vehicle enters the vehicular network it starts to run the algorithm and enters the INIT state.

During the INIT state the vehicles broadcast INIT messages, to make themselves known to their neighbors, with an interval of $T_{INT} = 3s$. When a vehicle receives a

message, regardless of the message type, it checks whether it belongs to one of its CNs. Each vehicle contains a list of all the vehicles that communicate with it on a regular basis. To do that, every vehicle calculates the time difference between the last time it received a message from a consistent neighbor (CN) with the simulation run time. If the time difference is greater than a threshold $T_{CN} = 2T_{INT} + 1s$, then that CN is removed from the list.

C.2.2 ND State

Upon receiving an INIT message, a vehicle adds it in its CNs if the heading direction is the same and sends a unicast acknowledge message to the INIT message sender. As shown in Figure C.4 during the INIT state each vehicle could face different situations. In Figure C.4 the vehicle (a) exchanges messages with vehicles that head towards the same direction as it does and therefore it records them all as CNs. The vehicle (b) on the other hand will not receive any acknowledgment messages and therefore it will not have any CNs. In cases such as those of vehicles (c) and (d), even though they might exchange messages, upon comparing their heading directions, they will dismiss each other and will not be considered CNs.

If a vehicle considers at least one of its neighbors as a CN then it enters the ND state. Each vehicle during this state calculates its CS and checks regularly for a higher CS among its neighbors. A SIGN message identical to the INIT message, containing also the vehicle's CS, is still sent every $T_{INT} = 3s$ and it works as a "sign of life" of a vehicle. No acknowledgment is required when the SIGN message is received. A vehicle will stay in the ND state for only $T_{ND} = 2 \times (T_{INT}) + 1$. The reason behind this statement is that during its previous state, the vehicle gathered enough data to calculate its CS, and therefore during the first T_{INT} it should already have collected enough messages to know whether it is the most suitable candidate to become a CH in its vicinity. After the time T_{ND} , the vehicle which is aware it has the greatest CS, will become a CH.

C.2.3 Cluster Formation

Upon entering the CH state, a vehicle will start broadcasting CHhello messages to inform the other vehicles in its vicinity that a cluster is forming. When a vehicle, that is not a CH itself, receives a CHhello message, it sends a unicast acknowledgment

message to the CH to let it know that it is now a CM of its cluster. If the vehicle that received the CHhello message is in an ND state, it enters the CM state. If the CHhello message was received by a vehicle that was already a CM of another cluster, that vehicle sends a unicast acknowledgment message to the sender of the CHhello message and adds it to its list of CHs. An example of the final result of the clustering algorithm is presented in Figure C.5.

Algorithm 1: The Cluster Formation Algorithm.

```

1 Result:Cluster Formation.
2 initialization;
3 while MsgRcvd == False do
4   | BroadcastInitMsg();
5 end
6 setState(ND);
7 calculateCS(param1,param2,..);
8 if CS == maxCS then
9   | setState(CH);
10  | BroadcastCHHelloMsg();
11 else
12   | if CHHelloRcvd == True then
13     | setState(CM);
14     | setCH(rcvdCH);
15     | addCH(rcvdCH);
16     | sendMsg(DataMsg);
17   end
18 end

```

C.2.4 Re-clustering Procedure

On the occasion that a CH comes within transmission range of another CH and one of them receives a CHhello from the other, the Re-clustering Procedure starts.

The procedure starts with both NDs broadcasting a Re-clustering message and changing their state to ND. Every vehicle that receives a Re-clustering message

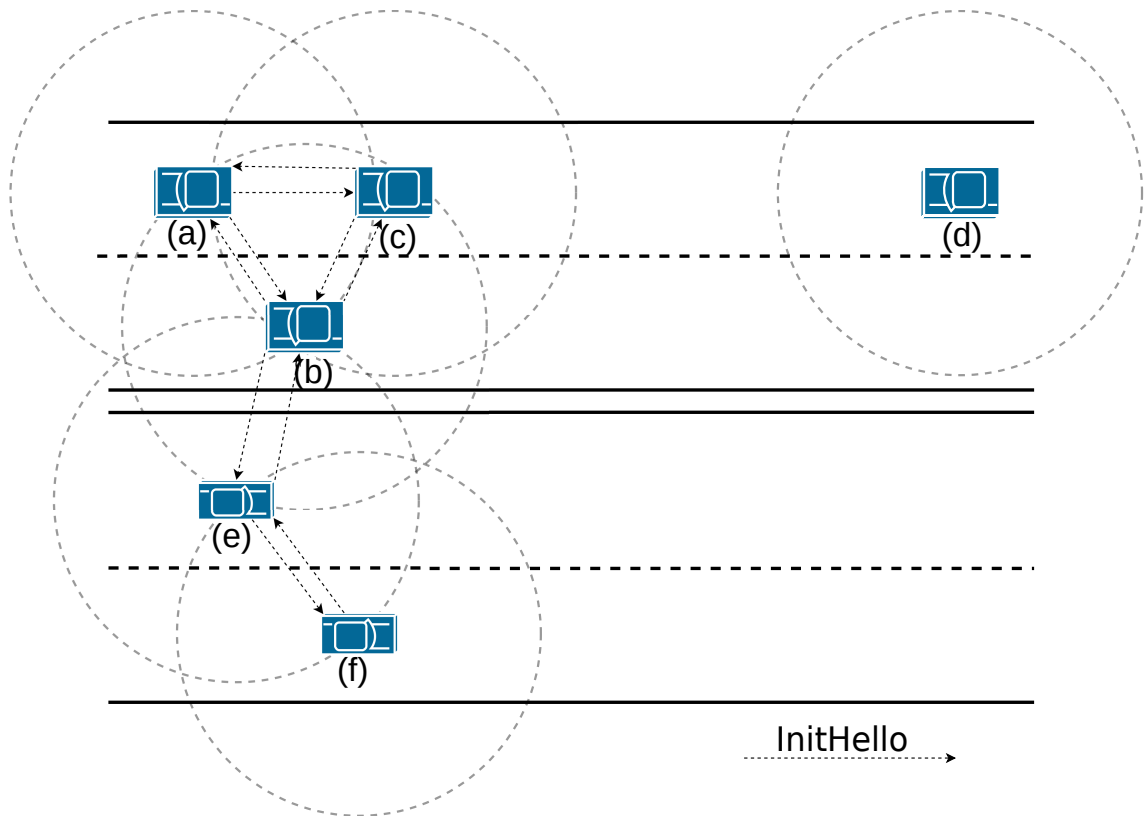


Figure C.4: Message exchange during the INIT state.

checks whether the sender is in its list of CHs. If not, then the message is ignored. Otherwise, that vehicle's state is also changed to ND. With all of the vehicles of both clusters checking for the best CS, a new cluster formation begins, in an attempt to merge the two clusters.

C.2.5 Cluster's life-cycle

In environments such as these of vehicular networks, high mobility among the nodes is one of the fundamental characteristics. Therefore, the fact that eventually, the CMs of a cluster will distance themselves from the cluster, is inevitable.

In this work, both the CHs and CMs check regularly whether they are maintaining communications with each other. Each CH checks its list of CMs for the last time it received a message from each one of them. Respectively each CM checks its list of CHs for the last time it transmitted a message. The evaluation is performed by calculating the difference between the last time a list item transmitted a message and the simulation runtime. If the difference is greater than $2T_{INT}$ it is presumed that the connection between the two nodes is terminated. Due to the algorithm's CM-centralized approach, as long as, a CM belongs to at least one cluster, no actions are needed. When a vehicle ends up with an empty list, meaning that a CH has no CMs and a CM does not belong to any cluster, it enters the ND state to avoid going over the INIT state. As a result, these "orphan" nodes can become CMs immediately once they approach a cluster.

C.2.6 Cluster Reforming Overhead Reduction

The maintenance of a cluster and the achievement of a maximum duration lifecycle is crucial for VANETs. The main reason behind the clustering of VANETs is the dissemination of information effectively in the network, so if a cluster does not live long enough to fulfill that, then there is no point in utilizing it.

To maintain each cluster's existence, the proposed architecture allows for dynamic cluster formation by deleting old CMs and allowing new vehicles to enter the cluster. This is achieved by allowing each vehicle that enters the CH's transmission range to join the cluster as a CM, given the fact the newly arrived vehicle is not a CH itself. This is done because it is presumed that vehicles which are traveling together, are likely interested in the same information regarding the network, e.g.

a road blockage spotted by the CH will probably concern the rest of the cluster as well since all of them are heading towards the same direction. Additionally, a CH always checks its CMs and removes any that so no signs of life.

C.2.7 Competence Score

The election of a CH that will form a cluster is based on the value called Competence Score (CS). SC is calculated using mobility-based metrics.

Based on the aforementioned, every message transmitted by the vehicles contains some specific metrics, such as the acceleration, and speed of the sender node. These values are used by each vehicle to create a list with its CNs. In addition this information is utilized by a vehicle v to calculate the CS of a node using this formula:

$$CS_v = [|u_v - \sum_{i=0}^{\overline{CN.size()}} u_i| + |a_v - \sum_{i=0}^{\overline{CN.size()}} a_i| + |coords_v - \sum_{i=0}^{\overline{CN.size()}} coords_i|] \times (CH.size())$$

where u_i is the speed of the i_{th} vehicle in the CN list, a_i is the acceleration of the i_{th} vehicle in the CN list, and the $coords_i$ the coordinates of the i_{th} vehicle of the CN list.

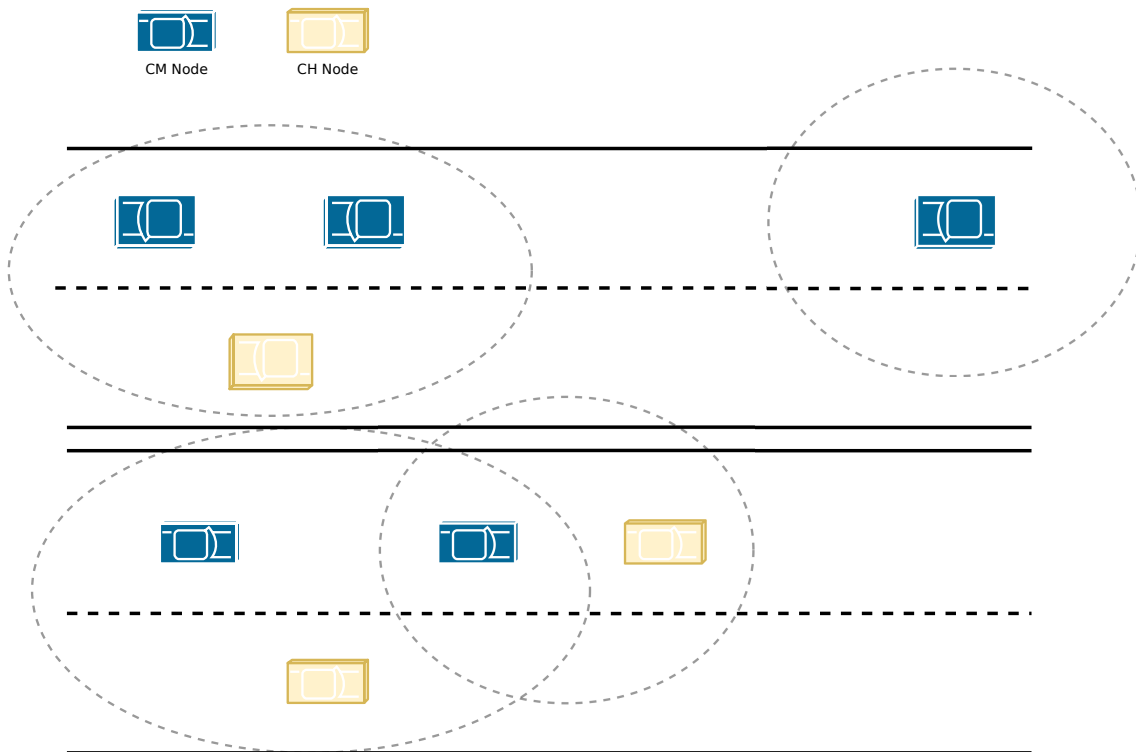


Figure C.5: Example of the cluster forming.

Chapter D

Performance Evaluation

The performance of the presented methodology is put to the test in this section. The tests are performed using vehicular network simulation. Different scenarios are presented in each run to cover a wide range of traffic conditions. The present work is also compared to another multi-head algorithm called Double-Head Clustering [40]. A variety of evaluation metrics are used to outline the algorithm’s effectiveness.

D.1 Simulation Tools and Scenarios

The system utilized in this work consists of different tools and frameworks in order to simulate real-world traffic conditions. Each component of this system is responsible for different functionality in the simulation.

D.1.1 SUMO and Veins

Simulation of Urban MObility (SUMO) [56] is an open-source traffic simulation package created to simulate large vehicular networks. It has the ability to work with many communication network simulation tools with the assistance of a middleware. In this particular work, the middleware utilized is called Veins Framework.

Veins [57] is an open-source framework that allows SUMO to work in combination with a network simulation tool called OMNeT++. A module manager in Veins works as a Traffic Control Interface (TraCI) and connects to SUMO using a TCP connection.

D.1.2 Omnet++

OMNeT++ [58] is a network simulation library and framework designed to simulate the functionality of a real-world network. In this work, OMNeT++ is used to control the communication network of the vehicles in the simulation. In this particular system, the communication is achieved by using IEEE 802.11p as its communication system. This communication standard is designed for dedicated short-range communications and vehicular networks. In OMNeT++, each vehicle is controlled by a mobility module. This module is exploited by Veins to advance the simulation in SUMO at regular time intervals and update the vehicle's mobility information.

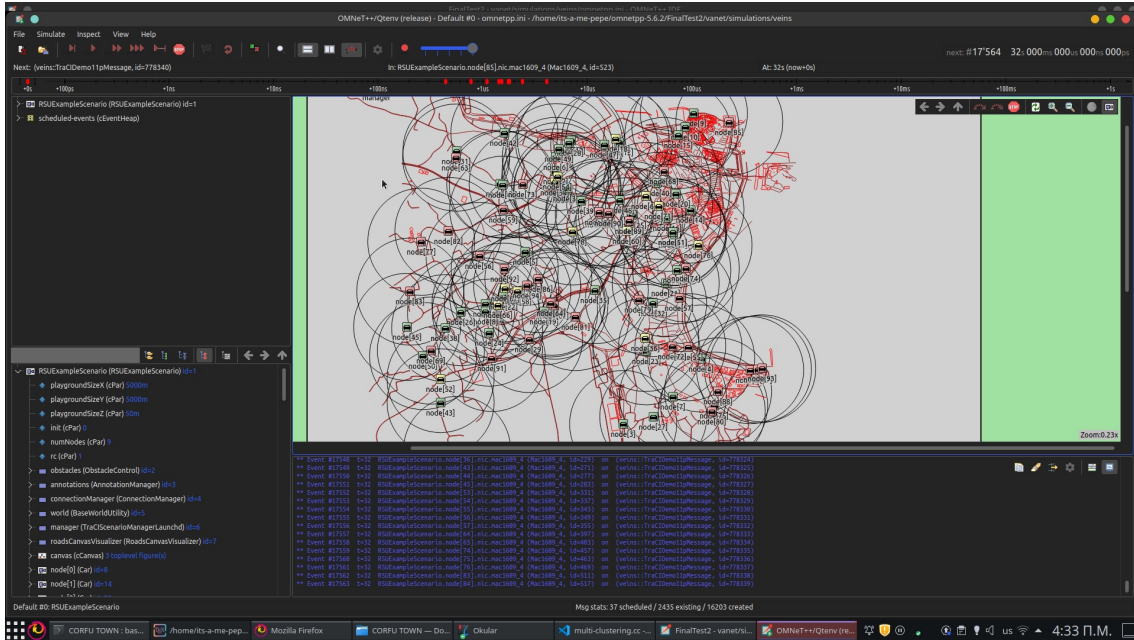


Figure D.1: Omnet++ Simulation.

D.1.3 Simulations Scenarios

To make the simulation scenarios as close to the real-world scenarios as possible, real-world road networks are utilized. That is crucial for methodologies that have the potential to become real-world applications.

With the use OpenStreetMap [59], an area of the map can be extracted in the form of an *xml* file. That file can be used by SUMO to create the traffic simulation.

Table D.1: Simulation Parameter Values

Simulation Parameter	Value
T_{INT}	$3s$
T_{CN}	$2T_{INT} + 1s$
T_{ND}	$2 \times (T_{INT}) + 1s$
Simulation Time	$400s$
Transmission Range	$400m$
MAC protocol	IEEE 802.11p
θ	$\Pi/2$
Transmission Power Level	$90mW$

That file contains also additional topology information, such as buildings, roundabouts, and the lanes for each street, with can be exploited by the simulations. A map of the town of Corfu is extracted in this work since real traffic data was available in [60, 61].

Veins also allows for signal shadowing effects to make the network simulation similar to that of a real-world application. Signal shadowing effects can have a grave influence on radio transmissions. Therefore, this feature along with the building information contained in the *xml* file, simulates accurately the blockage of transmissions caused by obstacles.

The simulation parameter values utilized in this work are depicted in Table D.1.

In the present work, many simulations are run, each with a different traffic density. The simulation runs consist of i) a 100 vehicles network simulation, ii) a 200 vehicles network simulation, iii) a 400 vehicles networks simulation, iv) a 500 vehicles network simulation and v) a 1000 vehicles network simulation. The trips for each vehicle of the network are generated with the assistance of a script contained in the SUMO documents that produce an *xml* file with all the points/coordinates that each vehicle will drive through. For each case of traffic density both the proposed multi-head clustering algorithm and the Double-Head Clustering Algorithm are implemented. In the end, the two algorithms are compared using the evaluation metrics presented in Section D.2.

D.2 Evaluation Metrics

To evaluate the algorithm's performance and stability, a variety of metrics are utilized in this work. These metrics mostly focus on the lifetime of the clusters and their members.

D.2.1 CH lifetime

The first evaluation metric used, is called *CH lifetime*. This metric represents the duration that a node operates as a CH. Since the algorithm does not present a CH replacement method for each cluster, the CH lifetime represents also the lifetime of a cluster.

D.2.2 CM lifetime

CM lifetime is the metric used to represent the duration that a node operates as a CM. This does not represent the time that a CM is part of one cluster, but rather the time that a vehicle served as a CM in general.

D.2.3 Alone lifetime

Another important factor to take into consideration is the duration that a vehicle was left roaming without being a member of any cluster. It is crucial for vehicles to spend the least time possible in these situations since they do not receive or transmit any useful information.

D.2.4 Cluster Size

The final metric is called *Cluster Size*. This metric depicts the average size of members a cluster has during its lifetime. This is utilized to measure the influence on the network each cluster has.

D.3 Double-Head Clustering Algorithm Comparison

The proposed algorithm is compared with the Double-Head Clustering Algorithm presented in [40]. This algorithm proposes the election of a second CH in every cluster, to extend each cluster's lifetime. That way when the CH loses communication with a CM, if the communication with the second CH isn't terminated, the CM will continue to belong in that cluster. A vehicle belongs only in one cluster. Double-Head Clustering Algorithm estimates how long a vehicle will still be part of a cluster using several mobility-based metrics. When that estimation is below a certain threshold, that vehicle will search its neighborhood for any other clusters and will attempt to join one of them.

D.4 Simulation Results

The same VANET simulations are utilized for the performance evaluation of both algorithms. That way, the two algorithms can be compared, purely based on their performance. Firstly, the time-based evaluation metrics, are presented, since the duration of a cluster's lifetime is a subject of grave importance when it comes to such applications. Figure D.2 depicts the average time, a vehicle was left roaming without being part of a cluster. It is outlined that during simulations of the proposed algorithm, the vehicles stayed without a cluster for a longer duration. That can be justified by the fact that the presented algorithm has an extra state (ND state) before it enters a cluster, during which it gathers information about its neighbors and calculates its CS.

On the other hand it is observed that as the size of the vehicles in the network becomes greater, the time a vehicle is left without a cluster, becomes shorter for the proposed algorithm. That is because the vehicle moves to the next state faster since it receives a message faster. On the other hand, the Double-Head Clustering Algorithm's orphan node duration, raises, due to the fact that the vehicles are required to send hello messages more frequently, leading to a high loss of messages due to collisions.

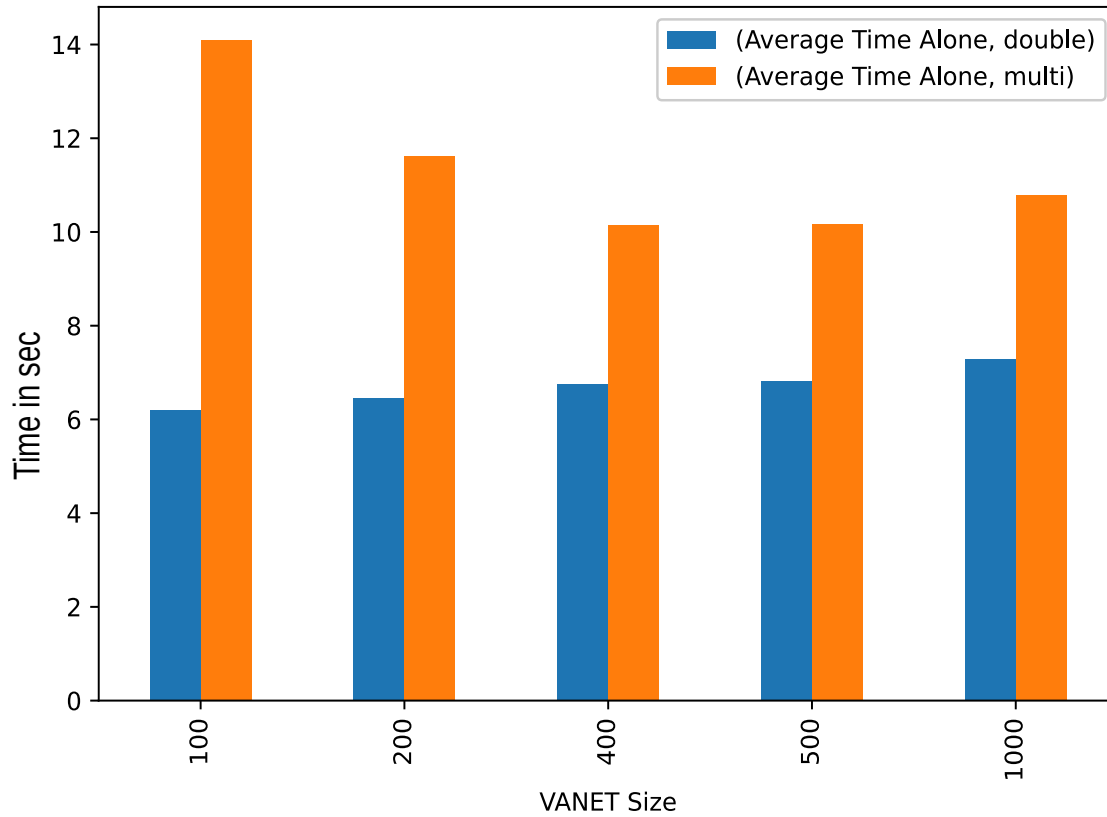


Figure D.2: The average time a vehicle spends time not being part of a cluster.

Figure D.3 depicts the average cluster lifetime. As seen in this Figure, the proposed algorithm performed better than the Double-Head Clustering Algorithm. That is because, the present work's algorithm treats new vehicles that want to join a cluster, in a less strict manner than that of the Double-Head Clustering Algorithm, in which a new vehicle has to be evaluated based on multiple mobility based metrics. The proposed algorithm, on the other hand, allows for new vehicles to join a cluster, regardless of their speed or acceleration, as long as they are moving in the same direction. As a result, more vehicles are qualified as CMs of clusters lowering their chance to be dismantled. As a result, the proposed algorithms achieve almost three (3) times greater cluster lifetime, compared to the algorithm found in the literature.

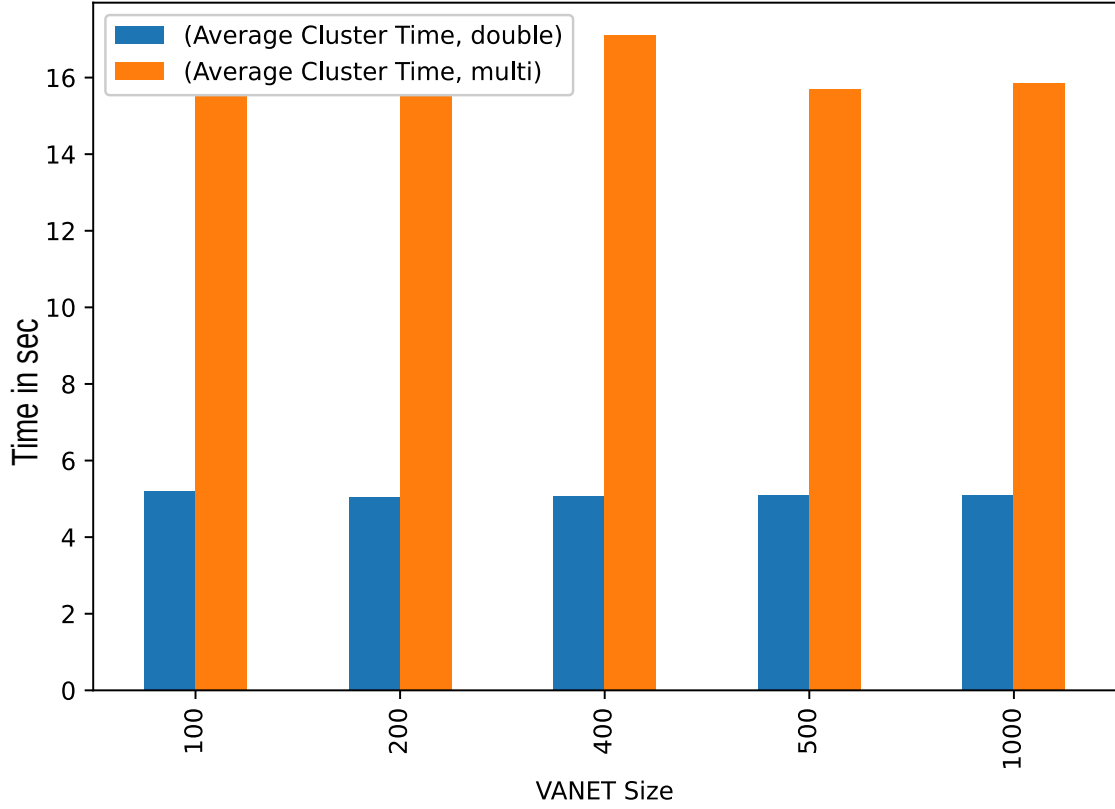


Figure D.3: The average cluster lifetime.

As shown in Figure D.4 the average time a vehicle operated as a CM is greater in the Double-Head Clustering algorithm. It can be seen, though, that the time a vehicle operated as a CM is heavily influenced by the density of the VANET. On the contrary, the present algorithm is not affected by the fluctuations in the density of the VANET and provides a stable average CM time. The algorithm found in the literature, in a VANET with 1000 nodes, drops its CM duration lifetime to less than half of that it achieved with 100 nodes. This means that for a VANET ten (10) times the original, the duration is cut to less than 50% of the initial one. The proposed algorithm, on the other hand, is not affected. That is because each node becomes a CH when it loses communication with all of its CHs. That way, if it eventually approaches other CMs they will become part of its cluster without having to go through the CH re-election process.

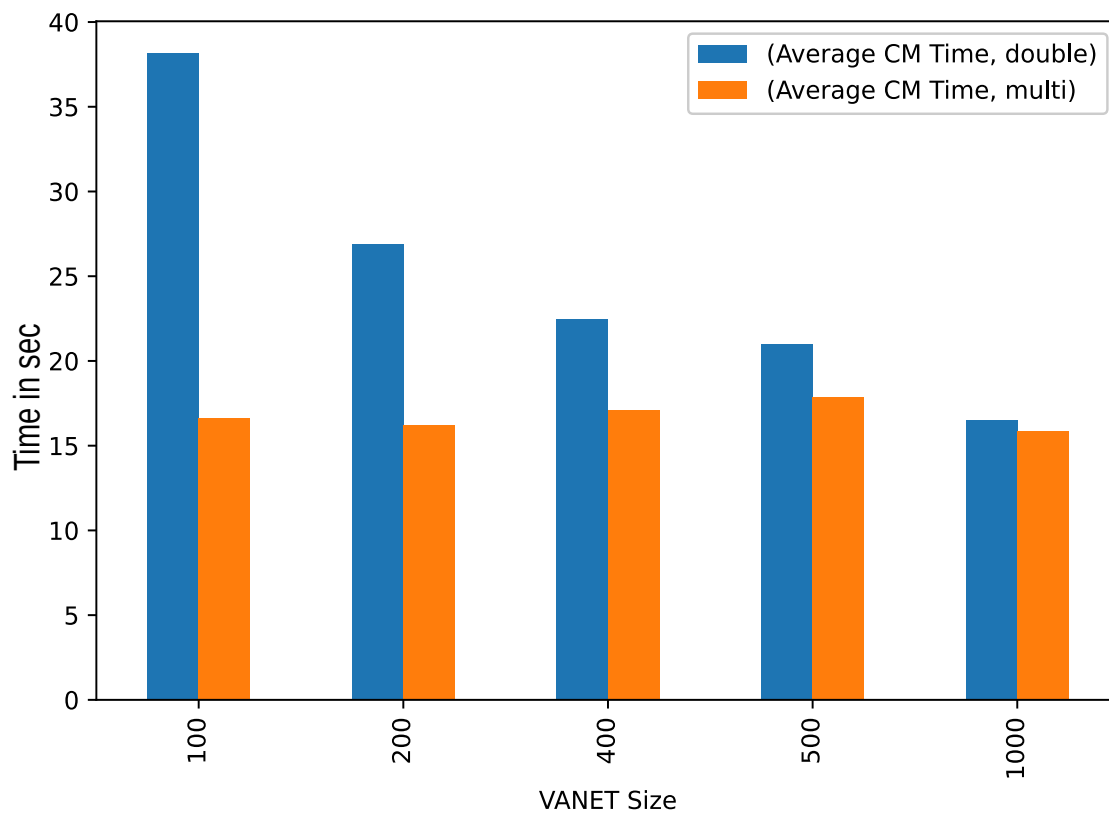


Figure D.4: The average time a vehicle operated as a CM.

Apart from the time-based parameters presented above, parameters regarding the number of neighbors a vehicle can spot and the size of the cluster, are also taken into consideration. These parameters are utilized to measure the influence a cluster or a node can have on the network.

The first parameter of this category is the actual size of the clusters formed in the VANET. As shown in Figure D.5 the average size of a cluster in a VANET with the presented algorithm is greater than that of the Double-Head Clustering Algorithm. That can be attributed to the fact that the proposed algorithm has fewer requirements when it comes to having a vehicle join a cluster as a CM. Therefore, the size of the cluster can easily change at any moment. Furthermore, the presented algorithm assumes that each vehicle can be part of multiple clusters and, therefore, each vehicle is not committed to just one CH. As a result, the size of each cluster can be much greater compared to other approaches, which assume that each CM is bound to one cluster. The average cluster size of the proposed algorithm grows over 50% every time the number of vehicles in the VANET doubles and it is at least double the size of the average cluster sizes observed in the Double-Head Clustering Algorithm.

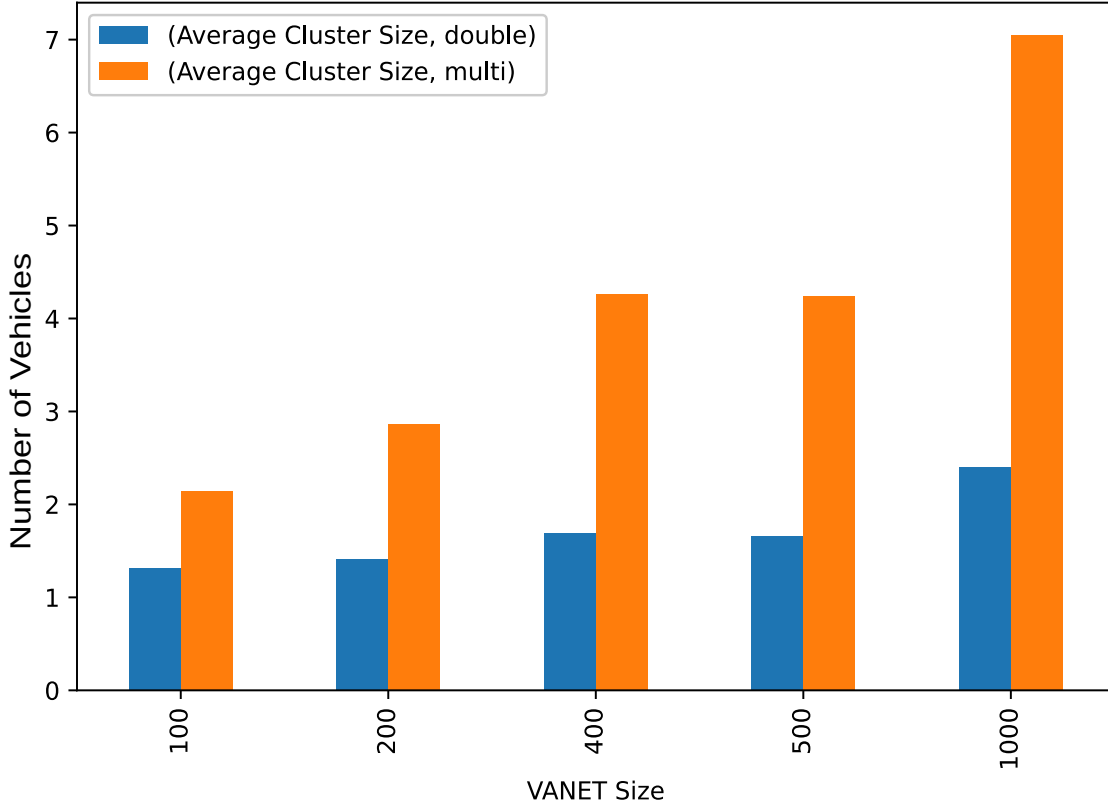


Figure D.5: The average cluster sizes.

The diagram depicted in Figure D.6 shows the average number of CHs a vehicle spotted during a simulation, while it operated as a CM. As shown in this Figure, the CMs in a VANET with the proposed algorithm, have more CHs in their vicinity than those of a Double-Head Clustering Algorithm. That is ascribed to the fact that in the Double-Head Clustering Algorithm, the second CH also works as a CH. In other words, nodes that lost connection with the main CH but still maintain a healthy connection with the second CH will continue to be part of the cluster. In these cases, the cluster also works as a two-hops cluster since it also contains nodes that are two hops away from the main CH. That also means that the number of clusters in the network is smaller than that of the proposed algorithm.

The final performance evaluation diagram is depicted in Figure D.7. This figure shows the difference between what the proposed and Double-Head Clustering algo-

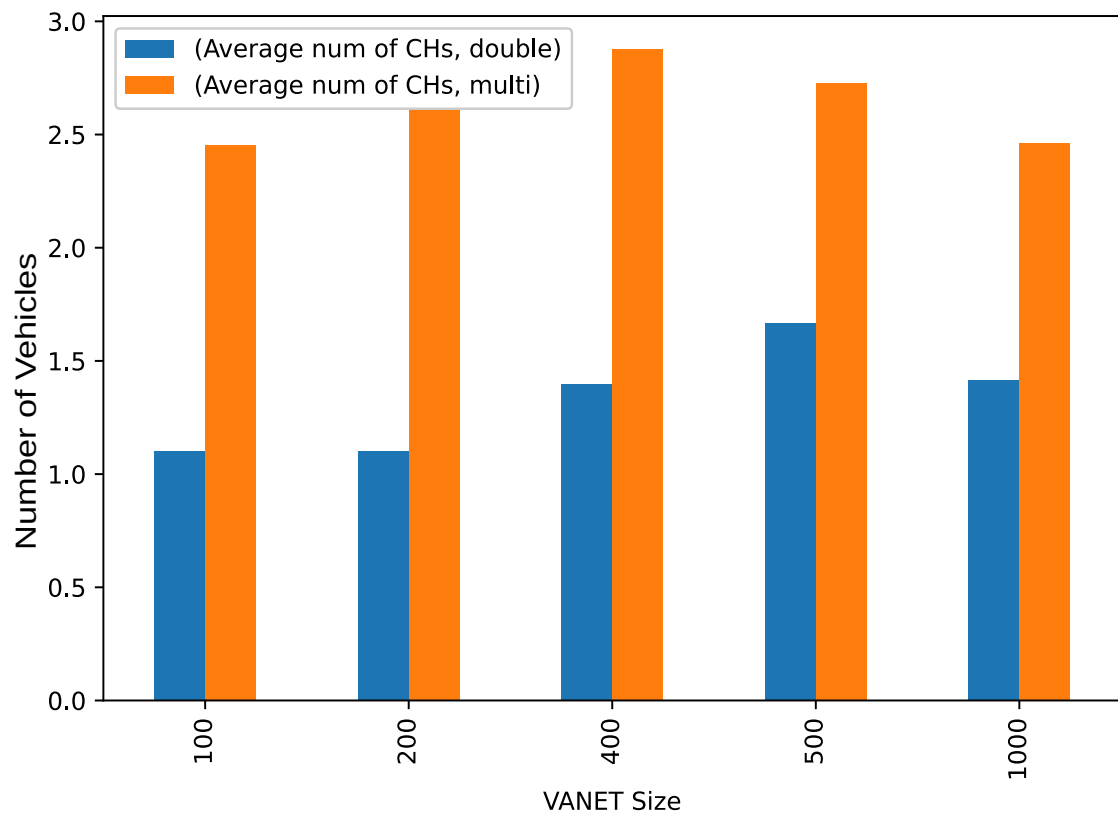


Figure D.6: The average number of CHs every CM spotted.

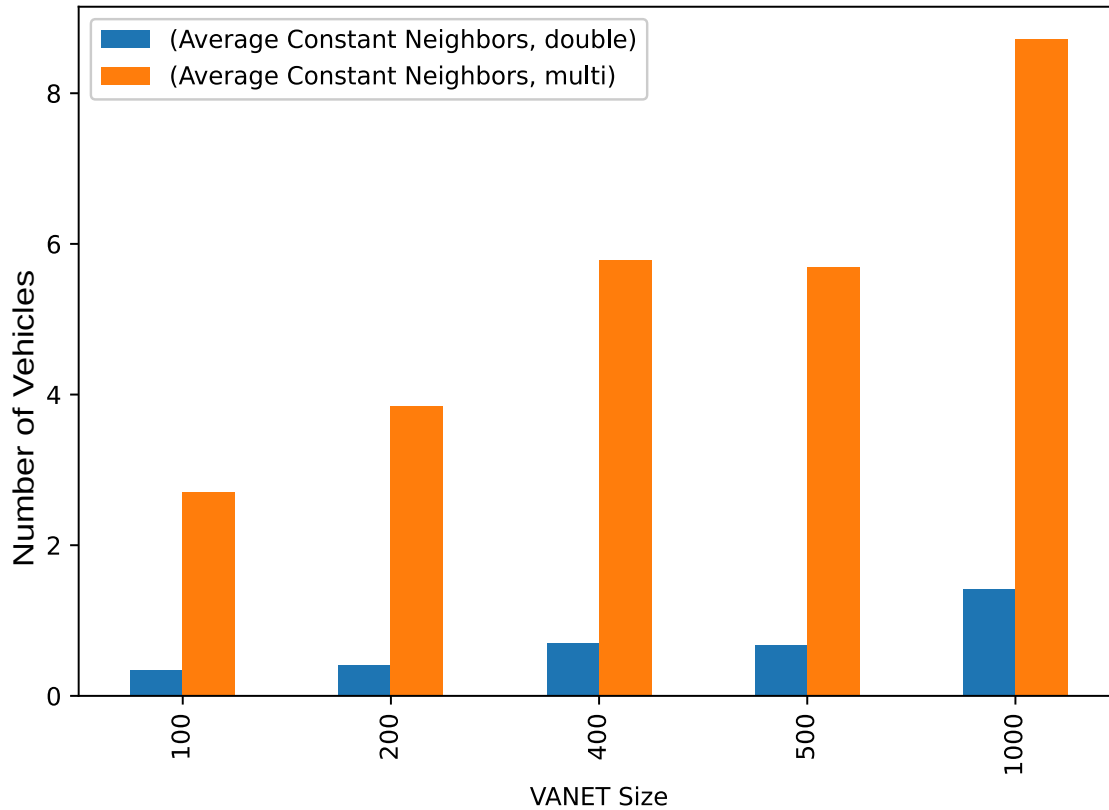


Figure D.7: The average number of CNs a vehicle had.

gorithms perceive as a constant neighbor. To further outline the differences between the two approaches, the average number of CNs is almost the same as the average cluster size. This is because the proposed algorithm uses the timestamp of the last message it received from a CN to check whether it should remove it from the list or keep it. Additionally, the only requirement for a neighbor to become a CN of a node is to move in the same direction as the node. On the other hand, a node in the Double-Head Clustering Algorithm allows only vehicles that move in the same direction and have a speed close to the one it has. Therefore it has a stricter CH policy that allows fewer vehicles to become CNs of a node.

Based on the results derived from the simulations, it is obvious that the proposed algorithm outperforms the Double-Head Clustering Algorithm. This is shown by the fact that the proposed algorithm has a greater cluster lifetime, which is also what

this work aims to achieve. Furthermore, it is shown that the presented algorithm is more resilient to the scalability of such networks since it shows that not only its CMs lifetime duration is not affected by it, but also that the average cluster size becomes greater the denser the VANET becomes. In addition, the average time a vehicle is left without a cluster becomes smaller as the density of the network grows, whereas the exact opposite happens with the Double-Head Clustering Algorithm.

As a result, it is observed that the proposed algorithm achieves both higher cluster lifetime duration and greater cluster size than the algorithm found in the literature.

Chapter E

Summary and Conclusions

THE complex and interdependent nature of modern societies arises a plethora of significant issues. These issues have been introduced in many forms, such as social, environmental, or economical issues.

Over the last few decades, and with the rise of the 5G Networks, many concepts revolving around the solutions to these issues have been presented. One of these concepts is that of *Smart Cities*. The term of *Smart Cities* refers to a society that utilizes technology-based applications to provide solutions to different issues of modern societies and improve the quality of life for the citizens. Another concept that became the focal point of many pieces of research, is that of the IoT. This concept proposes the use of self-configuring, *smart*, devices that can be integrated into objects used in the everyday lives of citizens, to observe several parameters of their environment and take actions to improve the citizens' lives.

In the field of vehicular road networks, the concept of VANETs has attracted a lot of attention from the scientific community. VANETs are considered to be a unique category of MANETs due to the peculiar environment of vehicular networks. The main reason behind the development and deployment of VANETs is the reduction and prevention of accidents, traffic management, and road optimization. However, one of the most important challenges VANETs face is the high mobility of the nodes.

In order to overcome this problem, many works have suggested the utilization of *Clustering*. Clustering is the technique of separating a set of objects into smaller groups that contain the same or similar attributes. Past literature shows that VANET clustering could be either based on the mobility-based parameters of

the vehicles or the network parameters.

Many works have proposed the use of one-hop or multi-hop clustering algorithms, even the exploitation of machine learning techniques to form more stable clusters. When referring to VANET cluster stability, most works talk about the duration of a cluster's lifetime or the rate at which a node changes states. These metrics are commonly utilized in past literature to measure the performance of clustering algorithms.

It is observed that while one-hop mobility-based algorithms are simpler and require less computational power, they tend to be less stable and reliable. On the contrary, algorithms that adopt a fuzzy-logic or machine learning-based approach tend to form more stable clusters since they can learn from the movement of the nodes and use that knowledge. These algorithms, though, tend to have greater computational requirements. Similarly, multi-hop clustering algorithms tend to form more stable clusters in addition to lowering the number of clusters in the network. Although, the end-to-end delay and the message overhead in such applications are greater than those of one-hop applications.

In recent literature, most works focus on developing more complex clustering algorithms that either utilize machine learning methods or rely on a multi-hop basis. On the other hand, to the author's knowledge, no work has suggested the use of a multi-head mobility-based clustering algorithm for VANETs.

The main purpose of this work is to develop a VANET clustering algorithm, for message dissemination. The proposed Multi-Head Clustering Algorithm exploits mobility-based metrics of the VANET to form clusters. No RSUs are utilized in this work.

The proposed algorithm assumes that each vehicle can be a CM of multiple clusters, to maintain the cluster formations further. In addition, it exploits mobility-based metrics, such as the speed and acceleration of each vehicle, along with network-based metrics, such as the number of constant neighbors, to elect a CH. The elected CH announces that it has formed a cluster by broadcasting a CHHello message at a regular interval. When a vehicle receives a CHHello message, it becomes a CM of that cluster and adds the CH to its list of CHs. On the occasion that two or more CHs are within each other's ranges, then they broadcast a re-clustering message, to alert their CMs that a re-clustering process will begin to merge the clusters.

The simulation runs were carried out with the exploitation of SUMO, Veins,

and OMNeT++. Sumo is a vehicular network simulator, that can be used, in collaboration with the Veins framework, to run VANET algorithms and tests in OMNeT++. SUMO supports also the use of a real-world network, which can be created using OpenStreetMaps. The map of the town of Corfu is used in this work. Different vehicular network scenarios are deployed during these runs to test the algorithm's performance, during different states.

The algorithm's performance evaluation is carried out by comparing it with Double-Head Clustering Algorithm. During the performance evaluation, different parameters are utilized to measure the algorithms' performance. The time-based metrics utilized are i) the average cluster lifetime, ii) the average time without a cluster, and iii) the average time a vehicle performed as a CH. The network-based parameters utilized in this work are i) the average size of a cluster, ii) the average number of CNs, and iii) the average number of CHs spotted by a CM.

The results show that the proposed algorithm outperforms the Double-Head Clustering Algorithm in both categories. Even though the proposed algorithm shows a higher average time of the duration a node was left without a cluster, the results show a greater cluster lifetime duration for the Multi-Head Clustering Algorithm. As for the network-based parameters, the proposed algorithms excelled in all three, by having a greater average cluster size, number of CHs per CM, and number of CNs per vehicle.

To conclude, with the development of the proposed algorithm, new challenges arise. Even though the idea of a Multi-Head approach improves the cluster's lifetime, the issue of which CH is the leading one arises. In other words, each CM that has multiple CHs must decide on which CH will be the one whose, information is more valuable and useful for the CM.

Furthermore, to further improve the cluster's lifetime, a mobility-based estimate could be utilized, during which the system will make an estimation on if and when it will lose communication with a cluster and consequently either adjust its mobility metrics or change the leading CH.

To allow the VANET to create clusters not just based on the one-hop neighborhood, but also to expand this concept to a large-scale clustering approach, a hierarchical clustering algorithm could be adopted. That way, even CHs will be able to become CMs of other clusters with CHs that have better CSs, thus creating clusters that can even communicate with each other easier.

Part I

References

Bibliography

- [1] “Home - inrix industries (<http://inrix.com/>).” <http://inrix.com/>, journal=INRIX, author=Inrix.
- [2] V. Kumar, S. Mishra, N. Chand, *et al.*, “Applications of vanets: present & future,” *Communications and Network*, vol. 5, no. 01, p. 12, 2013.
- [3] H. Hui, Y. Ding, Q. Shi, F. Li, Y. Song, and J. Yan, “5g network-based internet of things for demand response in smart grid: A survey on application potential,” *Applied Energy*, vol. 257, p. 113972, 2020.
- [4] B. Gurjar, T. M. Butler, M. G. Lawrence, and J. Lelieveld, “Evaluation of emissions and air quality in megacities,” *Atmospheric Environment*, vol. 42, no. 7, pp. 1593–1606, 2008.
- [5] M. G. Stewart, X. Wang, and M. N. Nguyen, “Climate change impact and risks of concrete infrastructure deterioration,” *Engineering Structures*, vol. 33, no. 4, pp. 1326–1337, 2011.
- [6] B. Gurjar, A. Jain, A. Sharma, A. Agarwal, P. Gupta, A. Nagpure, and J. Lelieveld, “Human health risks in megacities due to air pollution,” *Atmospheric Environment*, vol. 44, no. 36, pp. 4606–4613, 2010.
- [7] S. K. Bhoi and P. M. Khilar, “Vehicular communication: a survey,” *IET networks*, vol. 3, no. 3, pp. 204–217, 2014.
- [8] P. Papadimitratos, A. De La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, “Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation,” *IEEE communications magazine*, vol. 47, no. 11, pp. 84–95, 2009.

- [9] F. Kargl, P. Papadimitratos, L. Buttyan, M. Müter, E. Schoch, B. Wiedersheim, T.-V. Thong, G. Calandriello, A. Held, A. Kung, *et al.*, “Secure vehicular communication systems: implementation, performance, and research challenges,” *IEEE Communications magazine*, vol. 46, no. 11, pp. 110–118, 2008.
- [10] S. Giordano *et al.*, “Mobile ad hoc networks,” *Handbook of wireless networks and mobile computing*, pp. 325–346, 2002.
- [11] B. Williams and T. Camp, “Comparison of broadcasting techniques for mobile ad hoc networks,” in *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, pp. 194–205, 2002.
- [12] F. D. Da Cunha, A. Boukerche, L. Villas, A. C. Viana, and A. A. Loureiro, *Data communication in VANETs: a survey, challenges and applications*. PhD thesis, INRIA Saclay; INRIA, 2014.
- [13] X. Shen, X. Cheng, L. Yang, R. Zhang, and B. Jiao, “Data dissemination in vanets: A scheduling approach,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2213–2223, 2014.
- [14] L. Yang and F.-Y. Wang, “Driving into intelligent spaces with pervasive communications,” *IEEE Intelligent Systems*, vol. 22, no. 1, pp. 12–15, 2007.
- [15] N. Cheng, F. Lyu, J. Chen, W. Xu, H. Zhou, S. Zhang, and X. Shen, “Big data driven vehicular networks,” *IEEE Network*, vol. 32, no. 6, pp. 160–167, 2018.
- [16] B. Parno and A. Perrig, “Challenges in securing vehicular networks,” in *Workshop on hot topics in networks (HotNets-IV)*, pp. 1–6, Maryland, USA, 2005.
- [17] K. C. Dey, A. Rayamajhi, M. Chowdhury, P. Bhavsar, and J. Martin, “Vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) communication in a heterogeneous wireless network—performance evaluation,” *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 168–184, 2016.
- [18] B. K. Chaurasia and S. Verma, “Infrastructure based authentication in vanets,” *International Journal of Multimedia and Ubiquitous Engineering*, vol. 6, no. 2, pp. 41–54, 2011.

- [19] Y. Qian and N. Moayeri, “Design of secure and application-oriented vanets,” in *VTC Spring 2008-IEEE Vehicular Technology Conference*, pp. 2794–2799, IEEE, 2008.
- [20] F. Qu, Z. Wu, F.-Y. Wang, and W. Cho, “A security and privacy review of vanets,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2985–2996, 2015.
- [21] T. D. Little, A. Agarwal, *et al.*, “An information propagation scheme for vanets,” in *Proc. IEEE Intelligent Transportation Systems*, pp. 155–160, 2005.
- [22] C. Shea, B. Hassanabadi, and S. Valaee, “Mobility-based clustering in vanets using affinity propagation,” in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pp. 1–6, IEEE, 2009.
- [23] V. Albino, U. Berardi, and R. M. Dangelico, “Smart cities: Definitions, dimensions, performance, and initiatives,” *Journal of urban technology*, vol. 22, no. 1, pp. 3–21, 2015.
- [24] S. Li, L. Da Xu, and S. Zhao, “The internet of things: a survey,” *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.
- [25] T. Rahman, I. Ullah, A. U. Rehman, and R. A. Naqvi, “Clustering schemes in manets: Performance evaluation, open challenges, and proposed solutions,” *IEEE Access*, vol. 8, pp. 25135–25158, 2020.
- [26] H. Zhu, M. Li, L. Fu, G. Xue, Y. Zhu, and L. Ni, “Impact of traffic influxes: Revealing exponential inter-contact time in urban vanets,” *IEEE Transactions on Communications*, vol. 39, no. 12, p. 1729, 1991.
- [27] N. Ababneh and H. Labiod, “Safety message dissemination in vanets: Flooding or trajectory-based?,” in *2010 The 9th IFIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, pp. 1–8, 2010.
- [28] R. Xu and D. Wunsch, *Clustering*, vol. 10. John Wiley & Sons, 2008.
- [29] W. Fan, Y. Shi, S. Chen, and L. Zou, “A mobility metrics based dynamic clustering algorithm for vanets,” 2011.

- [30] M. Ni, Z. Zhong, and D. Zhao, "Mpbcc: A mobility prediction-based clustering scheme for ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4549–4559, 2011.
- [31] M. Ren, L. Khoukhi, H. Labiod, J. Zhang, and V. Veque, "A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (vanets)," *Vehicular Communications*, vol. 9, pp. 233–241, 2017.
- [32] K. A. Hafeez, L. Zhao, Z. Liao, and B. N.-W. Ma, "A fuzzy-logic-based cluster head selection algorithm in vanets," in *2012 IEEE international conference on communications (ICC)*, pp. 203–207, IEEE, 2012.
- [33] M. Arif, G. Wang, T. Peng, V. E. Balas, O. Geman, and J. Chen, "Optimization of communication in vanets using fuzzy logic and artificial bee colony," *Journal of Intelligent & Fuzzy Systems*, no. Preprint, pp. 1–13, 2020.
- [34] S. Sharma and A. Kaul, "Hybrid fuzzy multi-criteria decision making based multi cluster head dolphin swarm optimized ids for vanet," *Vehicular Communications*, vol. 12, pp. 23–38, 2018.
- [35] P. Sasikumar and S. Khara, "K-means clustering in wireless sensor networks," in *2012 Fourth international conference on computational intelligence and communication networks*, pp. 140–144, IEEE, 2012.
- [36] S. Dhiviya, A. Sariga, and P. Sujatha, "Survey on wsn using clustering," in *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, pp. 121–125, IEEE, 2017.
- [37] I. Hussain and C. Bingcai, "Cluster formation and cluster head selection approach for vehicle ad-hoc network (vanets) using k-means and floyd-warshall technique," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 12, pp. 11–15, 2017.
- [38] R. Chai, X. Ge, and Q. Chen, "Adaptive k-harmonic means clustering algorithm for vanets," in *2014 14th International Symposium on Communications and Information Technologies (ISCIT)*, pp. 233–237, IEEE, 2014.

- [39] P. Bhosale and A. Vidhate, “An agglomerative approach to elect the cluster head in vanet,” in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, pp. 1340–1344, IEEE, 2016.
- [40] G. H. Alsuhli, A. Khattab, and Y. A. Fahmy, “Double-head clustering for resilient vanets,” *Wireless communications and mobile computing*, vol. 2019, 2019.
- [41] G. H. Alsuhli, A. Khattab, and Y. A. Fahmy, “A mobility-based double-head clustering algorithm for dynamic vanet,” in *2018 International Japan-Africa Conference on Electronics, Communications and Computations (JAC-ECC)*, pp. 91–94, IEEE, 2018.
- [42] S.-C. Lo, Y.-J. Lin, and J.-S. Gao, “A multi-head clustering algorithm in vehicular ad hoc networks,” *International journal of computer theory and engineering*, vol. 5, no. 2, p. 242, 2013.
- [43] X. Duan, Y. Liu, and X. Wang, “Sdn enabled 5g-vanet: Adaptive vehicle clustering and beamformed transmission for aggregated traffic,” *IEEE Communications Magazine*, vol. 55, no. 7, pp. 120–127, 2017.
- [44] X. Duan, X. Wang, Y. Liu, and K. Zheng, “Sdn enabled dual cluster head selection and adaptive clustering in 5g-vanet,” in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, IEEE, 2016.
- [45] S. A. Rashid, L. Audah, M. M. Hamdi, and S. Alani, “Prediction based efficient multi-hop clustering approach with adaptive relay node selection for vanet,” *Journal of Communications*, vol. 15, no. 4, pp. 332–344, 2020.
- [46] Y. Chen, M. Fang, S. Shi, W. Guo, and X. Zheng, “Distributed multi-hop clustering algorithm for vanets based on neighborhood follow,” *Eurasip journal on Wireless communications and networking*, vol. 2015, no. 1, pp. 1–12, 2015.
- [47] E. Dror, C. Avin, and Z. Lotker, “Fast randomized algorithm for 2-hops clustering in vehicular ad-hoc networks,” *Ad Hoc Networks*, vol. 11, no. 7, pp. 2002–2015, 2013.

- [48] S. Ucar, S. C. Ergen, and O. Ozkasap, “Multihop-cluster-based ieee 802.11 p and lte hybrid architecture for vanet safety message dissemination,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2621–2636, 2015.
- [49] X.-S. Yang and A. Slowik, “Firefly algorithm,” in *Swarm Intelligence Algorithms*, pp. 163–174, CRC Press, 2020.
- [50] R. Kumar and S. Barani, “Reputation based clustering system in vehicular adhoc networks,” *International Journal of System Assurance Engineering and Management*, pp. 1–8, 2021.
- [51] L. Sellami and B. Alaya, “Samnet: Self-adaptative multi-kernel clustering algorithm for urban vanets,” *Vehicular Communications*, vol. 29, p. 100332, 2021.
- [52] R. Kaur, R. K. Ramachandran, R. Doss, and L. Pan, “The importance of selecting clustering parameters in vanets: A survey,” *Computer Science Review*, vol. 40, p. 100392, 2021.
- [53] M. Alowish, Y. Takano, Y. Shiraishi, and M. Morii, “Performance evaluation of a cluster based routing protocol for vanets,” *Journal of communications*, vol. 12, no. 2, pp. 137–144, 2017.
- [54] L. Nkenyereye, L. Nkenyereye, S. M. R. Islam, C. A. Kerrache, M. Abdullah-Al-Wadud, and A. Alamri, “Software defined network-based multi-access edge framework for vehicular networks,” *IEEE Access*, vol. 8, pp. 4220–4234, 2020.
- [55] K. A. Awan, I. Ud Din, A. Almogren, M. Guizani, and S. Khan, “Stabtrustb•” a stable and centralized trust-based clustering mechanism for iot enabled vehicular ad-hoc networks,” *IEEE Access*, vol. 8, pp. 21159–21177, 2020.
- [56] “Simulation of urban mobility.”
- [57] “Vehicles in network simulation.”
- [58] Andras, “Omnet.”
- [59] “Openstreetmap.”

- [60] S. Vergis, V. Komianos, G. Tsoumanis, A. Spiggos, and K. Oikonomou, “Implementation of a low-cost vehicle traffic monitoring system in the town of corfu,” in *2019 4th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, pp. 1–6, IEEE, 2019.
- [61] S. Vergis, V. Komianos, G. Tsoumanis, A. Tsipis, and K. Oikonomou, “A low-cost vehicular traffic monitoring system using fog computing,” *Smart Cities*, vol. 3, no. 1, pp. 138–156, 2020.

Appendices

Appendix I

Additional Images

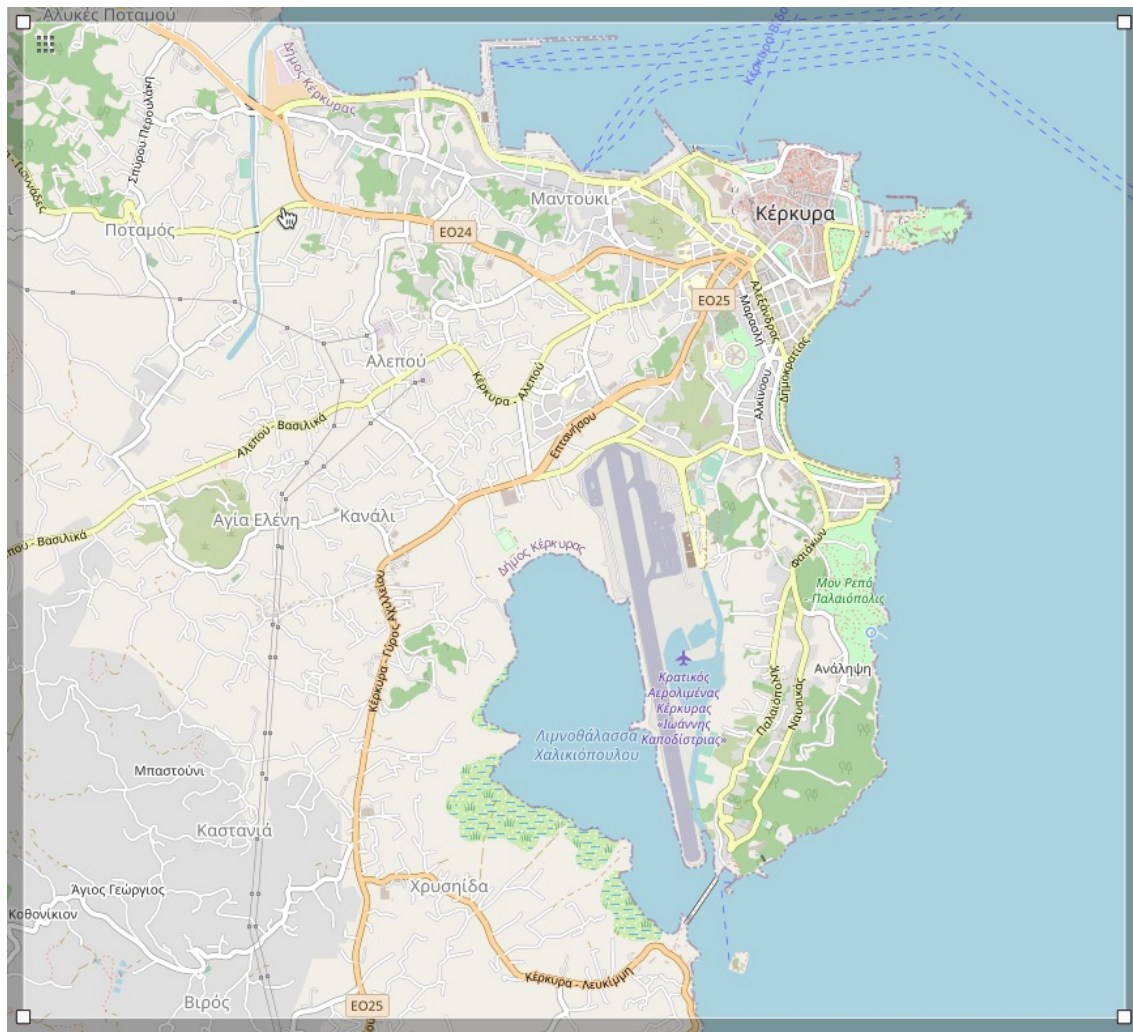


Figure I.1: Map of Corfu Town. This is the road network utilized in this work.

Appendix II

Glossary

Sink node	The node of a wireless sensor network that accumulates all sensed data.
Cluster Head	The node of a cluster that is responsible of disseminating information.
Cluster Member	The node of a cluster that collects the information sent by the Cluster Head
Cluster	A group of objects that share the same or similar characteristics.

Appendix III

Notations

V	Set of network nodes
CS_v	Competense Score of vehicle v
u	Speed of vehicle
a	Acceleration of vehicle
$coords_v0$	The exact coordinates of the vehicle v at a specific time

Appendix IV

Abbreviations

V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
RSU	Roadside Unit
IoT	Internet of Things
VANETs	Vehicular Ad-Hoc Networks
MANETs	Mobile Ad-Hoc Networks
GPS	Global Positioning System
CH	Cluster Head
CM	Cluster Member
INIT	Initialization State
ND	Neighborhood Discovery
CS	Competence Score