Hotspot-Based Cellular Vehicle-to-Everything Networks (C-V2X) for Real-time MEC-Assisted Applications

Olanrewaju Muheeb Ahmed

Doctor of Philosophy

University of York

Physics, Engineering and Technology

September 2023

Abstract

As more connected vehicles pervade our roads, achieving their original objectives which include safety, efficiency and enjoyable experience becomes more important. Achieving CAV (Connected and Autonomous Vehicles) objectives is hinged on reliable and efficient communication between participating road entities. Cellular Vehicle-to-Everything (C-V2X) networks have been touted to have great potential in supporting CAV operations. The development of more Mobile Edge Computing (MEC) assisted applications for CAVs means more vehicles will require access to cellular infrastructure. In high vehicular demand scenarios, cellular infrastructure might struggle to satisfy both traditional and vehicular users. A cluster and relay approach have been suggested to create a hotspot scenario, where strategically located users could stream and relay information to vehicles within its cluster. In highly dynamic driving environments, clusters tend to be unstable leading to poor Vehicle-to-Vehicle (V2V) link performance.

This thesis describes how we approached these problems by answering three key questions; could the re-clustering process be remodelled to improve V2V network performance? At what point in time should re-clustering be initiated to minimize overhead while sustaining performance? Is there an optimal number of clusters that satisfies efficient bandwidth resource utilization for both Vehicle-to-Infrastructure (V2I) and V2V communication? We sought answers to these research questions through simulation and explored two distinct driving environments: the urban and highway environment. By limiting our work to use-cases requiring short-time real-time download of traffic data, we were able to focus our efforts towards improving cognate performance indices such as throughput, jitter and reliability. Our re-clustering remodelling effort and scheduling schemes yielded significant improvement in throughput and jitter and stability performance, while our attempt on bandwidth resource utilization succeeded in improving V2I and V2V user bandwidth.

Table of Contents

List of Tables

List of Figures

Acknowledgements

With deep humility and faith, I express my deepest gratitude to God Almighty for his blessings and divine guidance towards achieving every milestone in my research journey.

I would also like to extend my profound gratitude to the individuals whose unwavering support, guidance, and love have been pivotal in my journey of completing this thesis.

Professor David Grace, your mentorship and scholarly insights have been instrumental in shaping the trajectory of this research. Your patient guidance, dedication to my academic development and your unwavering belief in my potential have been deeply motivating. **Professor Paul Mitchell**, I am immensely grateful for your guidance and expertise throughout this research endeavour. Your keen insights and constructive feedback during TAP meetings have greatly enriched this work, pushing me to strive for excellence. **Dr Kanapathippillai Cumanan**, your valuable input and scholarly wisdom have broadened my perspective and enriched the depth of this thesis.

My Dear Wife, **Mariam Funmilayo Ahmed**, your unwavering support, sacrifices, and enduring love have been my constant source of strength. Your belief in my abilities and your willingness to stand by me during the most challenging times have made all the difference. My dear children, **Abdul-Hameed and Abdul-Adheem Ahmed**. Your smiles and laughter have brightened my days, reminding me of the importance of balance in life. I appreciate your understanding during those moments when I had to prioritize my research.

My precious parents, **Mr S. A. Ahmed, and Mrs W. B. Ahmed**. Thank you! Your unconditional love, encouragement, and the values you instilled in me have been the bedrock of my journey. I am deeply grateful for the educational opportunities you provided and the sacrifices you made to support my dreams. My prized siblings, **Mr Saheed (MBA, MSc), Mrs Hamdalah (PhD) and Mrs Maryam (PhD).** Your support, camaraderie, and shared moments of joy have lightened the load during this academic pursuit. Your laughter and encouragement have been a source of inspiration.

Declaration

I, **Olanrewaju Muheeb Ahmed** hereby declare that I am the sole author of this thesis and the work presented therein is entirely my original work, except where due acknowledgment and appropriate references have been made to the work of others. I also declare that this work has not been submitted in any form to this or any other educational institution for this or any other academic award.

1 Introduction

1.1 Research Background

The ever-growing need for mobility and the consequent development and production of more mobile and fixed transportation infrastructure to meet such needs, necessitates a need to deal with the accompanying safety and traffic congestion issues. Road transportation is particularly prone to losses stemming from safety and traffic inefficiency concerns. According to the World Health Organization (WHO) statistics, road accidents are the leading cause of deaths among young adults. It is recorded that an average of approximately 1.35 million road traffic-related deaths are recorded every year globally [1]. The Sustainable Development Goal (SDG) target-3.6 sought to reduce road accident fatalities by a factor of 0.5 by 2020. This feat is still far from being achieved. In terms of traffic efficiency, the Economists reports that in 2017, traffic jams cost the United States of America (USA), Britain, and Germany a combined sum of $461 billion [2]. Because of the enormous time wasted on roads due to inefficiencies in current road transportation systems, vehicles tend to burn more fuel, releasing emissions that could harm health (air pollutants) and the environment (greenhouse gasses). The work reported in [3] corroborates this position and indicates that traffic inefficiency shows a positive correlation to fossil fuel usage in vehicles. These highlighted transportation issues call for the development of a safer, more efficient and coordinated transport system, hence the advent Intelligent Transportation Systems (ITS).

ITS is defined by the European Union directive 2010/40/EU as a system that enhances ICT based coordination of road entities to improve traffic management and support advanced transport applications [4]. ETSI describes ITS functions as services that support all modes of

transportation and traffic management with the aim of offering users information to make efficient, safe and coordinated use of transport networks [5], [6]. Taking hint from the 2019 transport statistics report published by the UK department for transport which suggests that road transport constitutes 79% of all transportation means [7], it is then safe to assume CAV represents a large component of ITS mobility and focussing our work on solving some of the issues associated with CAV will be worthwhile.

Autonomous and connected driving seeks to reduce road accidents and improve traffic efficiency by reducing human errors in driving and decision making to the barest minimum possible. The separate components of CAV; Autonomous Vehicles (AV) and Connected Vehicles (CV) are individually insufficient but complementary technologies with potential to meet the requirement of a safe, efficient, and enjoyable driving experience. Beyond automation of typical driving function, AVs also offer environment perception and self-navigating functions with different degrees of human intervention [8]. In 2014, the Society of Automotive Engineers, SAE defined the J3016 'Level of Driving Automation' for consumers. SAE defined six different levels of driving autonomy, starting from SAE level-0 that specifies a non-autonomous driving experience in which a driver has full control of driving functions to SAE level-5 which indicates the highest degree of driving autonomy where all driving functions including environment perception and navigation is controlled by the vehicle itself without any human intervention. In the 2021 review of the standard, the level of driving autonomy is defined by the degree of intervention of three defined primary actors: human, driving automation system and other vehicle systems and components. The complete range of the SAE levels of driving autonomy is presented in Figure 1-1 [9], [10].

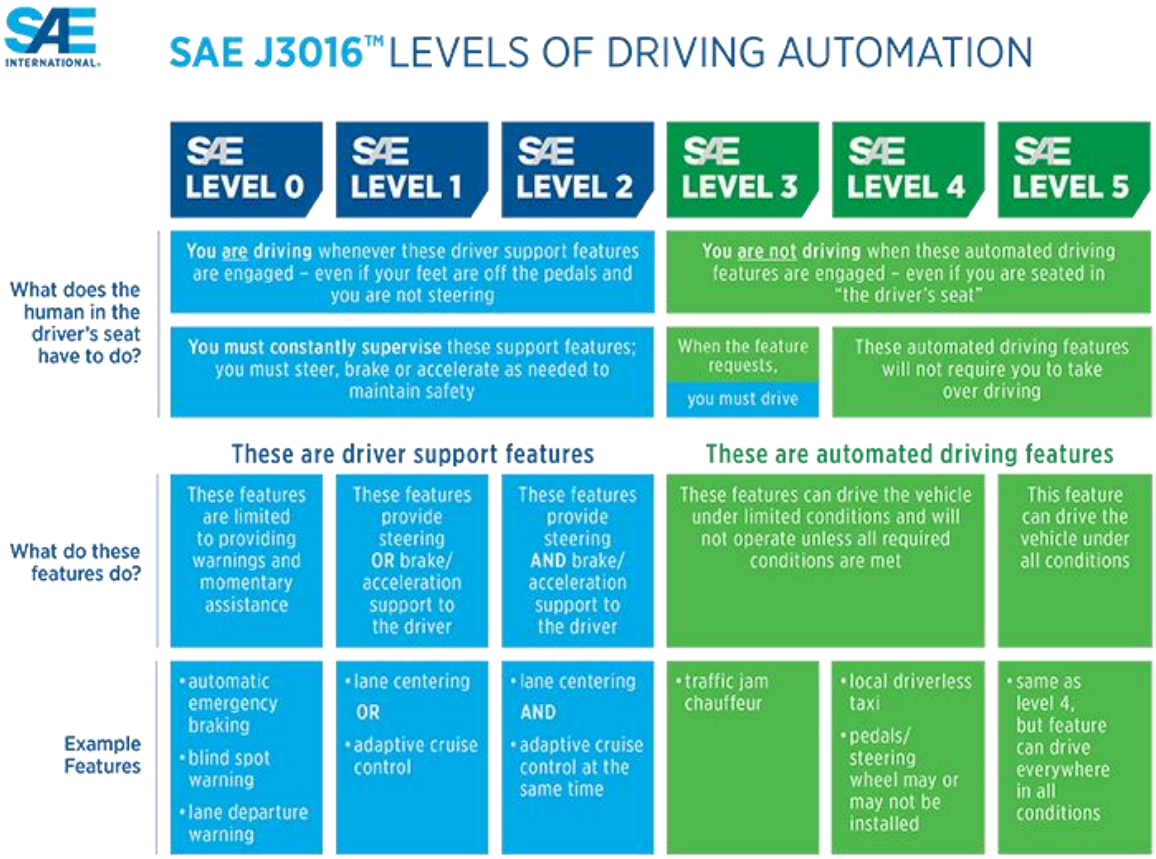

Figure 1-1. The Society of Automotive Engineers, SAE standard definition for level of driving automation [9], [10]

The automation function of AVs is loosely comprised of environmental perception and navigation functions supported by the driving automation system. Sensors are critical to the driving tasks of AVs and are generally in two categories: internal state and external state sensors. While the internal sensors measure vehicle health parameters, the external state sensors are responsible for driving related parameters such as localization, environment perception, navigation, and decision function of AVs. LiDAR, vision cameras, radar, Global Navigation Satellite System (GNSS), vehicle odometer and ultrasonic sensors are example of external state sensors [8]. Despite being equipped with an array of sensors, AVs still have their limitations for intelligent driving purposes. The limitations of AV's became evident following the accidents of Tesla autopilot and Uber self-driving car in 2016 and 2018 respectively. Some limitations include limited environment perception range, poor node coordination, and extreme weather and road conditions [11].

Solving this problem requires information sharing and cooperation amongst all road entities, which in-turn is heavily hinged on the efficiency and reliability of the supporting

communication network. Vehicle-to-Everything (V2X) defines the communication between vehicles and between vehicles and other road entities. V2X communication is a core component of the autonomous driving and Intelligent Transport System (ITS) concept that has gained attention in the automotive and telecommunication industry.

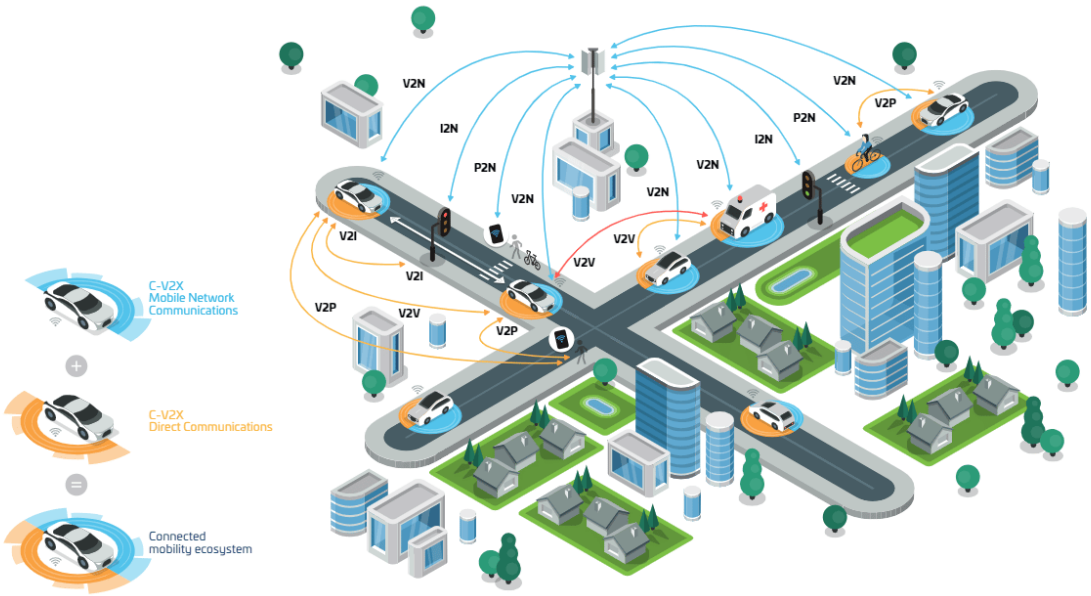

Figure 1-2. C-V2X driving scenario [12]

V2X holds the potential of providing reliable data transmission for both safety and non-safety applications. Different V2X access technologies have been defined by the IEEE (Institute of Electrical and Electronic Engineers) and the 3GPP (Third Generation Partnership Project) for communication between vehicles and other road entities. The IEEE defines a 802.11p standard, which is an evolution of traditional Wi-Fi standard. It focusses on V2V and V2I communication for safety application. The 3GPP on the other specifies the standards for Cellular V2X (C-V2X). Which integrates side-link V2V and V2I communication with traditional cellular networks. Though the IEEE 802.11p specification has gained traction over the years, the 3GPP's C-V2X specification is becoming increasingly popular partly due to its relatively cheaper cost and ease of implementation, since it leverages on existing cellular networks [13]. C-V2X is also believed to have a more sustainable growth potential, since its growth correlates with the development of traditional cellular standards.

Though next-generation V2X communication networks are expected to consist of a variety of communication modes, only four apply to most use-cases; Vehicle to Infrastructure (V2I), Vehicle to Pedestrian (V2P), Vehicle to Network (V2N), and Vehicle to Vehicle (V2V) [14], [15], as shown in Figure 1-2. All these V2X modes, particularly the V2V mode are vulnerable to the harsh communication environment resulting from high mobility and a fast-changing network topology associated with vehicular networks. Clustering has been identified as one way to deal with the issue of the changing vehicular network environment. Extensive research has been done on the implementation of different clustering algorithms in Vehicular Ad-hoc Networks (VANET) to solve issues related to the challenging communication environment such as the hidden node problem and fast changing channel conditions caused by high vehicular mobility [16], [17], [18], [19], [20], [21], [22]. Different types of clustering scheme have also been suggested for the enhancement of cluster stability in C-V2X [13], [23], [24], [25], [26], [27]. However, in cases of download of real time HD or machine maps, high throughput, low jitter, and link stability are key performance indicators. Hence it has become important to develop schemes that in addition to the baseline stability benefits of clustering, improve on these performance indicators with least obtainable effect on parent users of the hosting cellular network. Also, as vehicles move in time relative to one another, the efficacy of clustering algorithms and V2V link performance drop, hence vehicles will need to be re-clustered to maintain acceptable link performance. However, persistent re-clustering could also undermine the stability of clusters. Hence the need to reach a suitable compromise between link performance indices and stability indices.

The primary purpose of the work described in this thesis is to improve the link performance of Cluster Head to Cluster Member (CH-to-CM) links for emerging real time applications by proposing schemes that enhances real-time QoS parameters such as reliability, jitter performance and throughput performance. The core of these contributions is to provide seamless transmission of real-time traffic from MEC servers to clustered vehicles.

1.2 Research Hypothesis and Objectives

The hypothesis upon which the research work presented in this thesis is based is as follows.

"The approach with which and time at which vehicles are clustered and re-clustered in a relay hotspot scenario can be exploited to improve V2X network performance and consequently the Quality of Experience (QoE) for users engaged in real-time and heavy download use-cases."

Popular centroid-based clustering schemes such as k-means start the process with initial selection of arbitrary seeds, an approach that negatively impacts on stability particularly when vehicular nodes are persistently re-clustered. The approach to vehicle clustering and re-clustering proposed replaces the arbitrary seed re-selection method of centroid-based clustering schemes with a more objective memory-based scheme that extrapolates subsequent centroids-in-time from the spatial information of previous clusters. We opined that this approach would enhance throughput, stability, and jitter performance. The number of clusters and cluster range can be carefully selected to minimize the effect of a throughput bottleneck for download applications in a vehicular hotspot network topology. The time at which re-clustering is initiated can be exploited to reduce overhead and improve stability by identifying points-in-time where re-clustering is unnecessary and current cluster posture should be sustained.

Having described our hypothesis, the clear set objectives by which we propose to prove our hypothesis are:

1. Development of a re-clustering algorithm that seeks to keep Cluster Head to Cluster Member CH-to-CM links stable and enhance the link jitter performance without significantly jeopardizing throughput performance.
2. Development of a network-centric re-clustering trigger algorithm that minimizes signalling overhead and jitter performance without significant compromise on throughput and stability performance.
3. Device a method to estimate the optimum number of clusters that maximizes the use of BS bandwidth resources for V2N links as well as minimize link bottleneck along the relay path of a vehicular hotspot network topology.

1.3 Scope of Study

This work specifically looks at improving the V2X link performance for real-time download applications in a vehicular-relay hotspot network topology. The algorithms proposed were

primarily tested for rural-highway driving environment and urban driving environment. The work is entirely simulation based. The traffic trace data for the rural-highway case-study was generated from a SUMO traffic simulator while that of the urban case-study is sourced from an existing real-world traffic trace database. The vehicular traffic density throughout travel time within the simulation is consistent.

1.4 General Assumptions

This section describes the general assumptions upon which the entire work presented in this thesis is built upon.

1. Vehicles can move in different directions relative to one another and at non-zero relative speed.
2. Vehicles are not limited to a constant sped across travel time but are limited to maximum speeds 70 miles per hour and 40 miles per hour in rural-highway and urban environment.
3. The maximum communication range beyond which vehicular nodes cannot send beacons to one another is defined.
4. Each vehicle is equipped with one PC5 V2V and one Uu V2I communication interface.
5. Both V2I and V2V communication interface can work simultaneously
6. All vehicular nodes are of the same height.
7. Topography of the road network is assumed flat.
8. All models and investigations are developed with a focus on single-cell C-V2X network.
9. Frequency is considered near fixed throughout vehicle travel time.

1.5 Research Contribution

The work reported in this thesis makes 4 key contributions to the field of centroid-based cluster maintenance in V2X communication networks. The key contributions include:

1 The development of an SNR-based clustering scheme as compared to the proximity-based schemes used in common centroid schemes like k-means clustering.

2 Development of two different memory-based re-clustering methods which unlike conventional clustering schemes leverage on the stored spatial information of preceding

cluster status to estimate succeeding cluster seeds/centroids and consequently cluster membership.

3 Development of one cluster-centric and one network-centric event triggered re-clustering scheme which evokes re-clustering based on cluster and network performance indices results generated from persistent monitoring respectively. Unlike existing schemes that commonly evokes cluster reformation based on CH membership or frequent vehicular movement in and out of clusters.

4 Development of a resource-aware optimum number of clusters and cluster membership range scheme for V2X network. Unlike existing schemes that are agnostic to cellular bandwidth resources, this approach proposes a more efficient use of both V2N and V2V resources.

1.6 Thesis Structure

This section describes the outline of the research undertaken as presented in succeeding chapters.

Chapter two presents a background study to the rest of the work. It does this by reviewing existing literature and the state of the art of V2X access technologies, use cases, propagation models and resource allocation models. The chapter also discusses existing vehicle mobility models and vehicular clustering models.

Chapter three first describes in detail the foundational hotspot relay topology model and the network-centric clustering model developed upon the hotspot model. Also described are the two memory-based re-clustering models built upon a foundational SNR-based clustering model: The Seed-based Waterfall Scheme (SWS) and Centroid-based Waterfall Scheme (CWS). The performance metric for evaluating the memory-based schemes is discussed and a comparison with baseline schemes is presented.

Chapter 4 describes the proposed cluster quality and the link quality index-based re-clustering event trigger schemes. The rural-highway and the urban driving environment in which the schemes are tested are discussed. The metrics for evaluating their performances are presented and a comparison with two baseline schemes are made.

Chapter 5 discuss the evaluation of the resource-aware optimal number of clusters and cluster range. A discussion and comparison of suggested resource partition schemes for mode-3 resource allocation is presented. A proposal of a detailed resource partition model was also presented. Metrics for evaluating the BS-aware scheme are discussed and are used to compare the resource-aware scheme to existing baseline resource-agnostic schemes.

Chapter 6 draws an overall conclusion to the work described in chapters three, four and five, describing how the set of objectives are met through the methods applied and established by the results presented. This section also describes suggested areas of development for each of our key contributions and proposed areas for future work.

2 Literature Review

2.1 Introduction

In the last two decades, there has been a massive development in the CAV industry. From standard definition to testing and commercialization of CAV technologies and products. Many academic and industry-based research projects have also been carried out in this period. The connectivity component of CAV is primarily enabled by a variation of V2X Radio Access Technologies (RATs). A description of contending V2X RATs to support current and emerging use-cases will be described in section 2.2. In section 2.3, a comprehensive discussion of basic and emerging use case categories, specific use-cases, their Key Performance Indicators (KPIs) and requirements is presented. Designing and evaluating the performance of V2X links to meet specific use-cases requires the use of an appropriate channel or propagation model that represents signal behaviour along or signal interaction with its transmission path from the transmitting road entity to the receiving road entity. Section 2.4 discusses common channel models and scenarios in the 5.9GHZ V2X frequency band. Due to the peculiarity of V2X networks, the 3GPP has suggested different modes of resource allocation to manage the additional spectrum resource demands of mobile and static road entities for V2X side-link transmissions [28]. Section 2.5 describes resource allocation approaches for both base stations covered/controlled nodes and uncovered autonomous node communication. The performance of V2X can either be tested using an experimental testbed in a real-life

environment [29] or software-based simulations [30]. However, field tests tend not to be technically or economically viable due to limitations in flexibility and cost required to quickly test different network topologies and traffic scenarios. Hence in many published works software-based simulation has been adopted to generate realistic mobility and propagation models where performance of V2X networks can be tested. Described in section 2.6 are common mobility models used to reasonably simulate the vehicular mobility in a driving environment. The dynamic nature of the vehicular driving environment and the consequent instability in V2X networks necessitates the need for schemes that mitigates the effect of instability on network performance. Clustering has been considered in different works as a technique to improve stability, reliability, and scalability [31]. In section 2.7, a description of different categories of clustering algorithms used in vehicular networks will be presented, along with a discussion on some of the recent works that have used clustering to improve the performance of vehicular networks.

2.2 V2X Radio Access Technologies

Standardization bodies associated with the vehicular communication project have defined access technology standards to support emerging ITS applications. Many car-manufacturing companies, technology companies and Original Equipment Manufacturers (OEMs) are also discussing and have started the implementation of existing wireless technology standards to support safety and non-safety services in testbeds. For example, the IEEE and ETSI fundamentally standardized the Direct Short-Range Communication (DSRC) technology, while the CellularC-V2X standards are ratified by the 3GPP [32]. These standards define communication between road entities within the context of all V2X communication modes. In 2018, Qualcomm in collaboration with BMW group, Ford Motor Company, 5G Automotive Association (5GAA) and Savari showcased Europe's first cross-vendor demonstration of V2X direct communication for a safety and traffic efficiency use-case [33]. In 2015, Toyota and Lexus became the first automakers to introduce a DSRC system equipped vehicle to the public market and sold over a 100,000 within the first three years [34].

V2X access technologies are designed to support all forms of V2X communication modes defined in the 3rd Generation Partnership Project (3GPP) technical report (TR-Rel 14). According to the report, V2X communication fundamentally consists of three communication

modes: Communication between vehicles and the roadside side network or transport infrastructure, V2I/V2N (Vehicle-to-Infrastructure/Network), communication between vehicles, V2V (Vehicle-to-Vehicle) and communication between vehicles and vulnerable road users, V2P (Vehicle to Pedestrian) modes [35]. The following sub-sections describe in detail candidate RATs that supports V2X communication. The RATs will be described under two broad categories: Cellular based V2X and IEEE 802.11 based V2X.

2.2.1 Cellular Network-Based V2X (C-V2X)

Cellular based vehicular networks exploit features of traditional cellular networks for communication between road entities (on-road and roadside units). Features such as wide network coverage and high capacity sets cellular based vehicular networks apart as compared to their IEEE 802.11 counterparts. C-V2X here is used as a generic name for both LTE-V2X and 5G NR-V2X Cellular-based vehicular networks which can be classified into three broad categories based on base station involvement: relay function, support function or non-functional mode. In the relay perspective, user data and control information are routed through base station nodes. The support function requires base station for signalling, hardware support and/or pre-configuration for direct communication between VUEs and/or between VUEs and other road entities. The non-functional mode involves full autonomy of communication between a vehicle and other road entities with no participation from the base station. While traditional technologies have been exploited for high speed V2N applications with base stations serving as a standalone relay node, it is noteworthy that in the C-V2X context, all these categories are complementary specifically in cooperative driving applications.

C-V2X was first developed as enhancement to existing Cellular networks. Though C-V2X have not been commercially deployed, two variants are already being considered for future deployment: LTE-V2X and 5G NR-V2X.

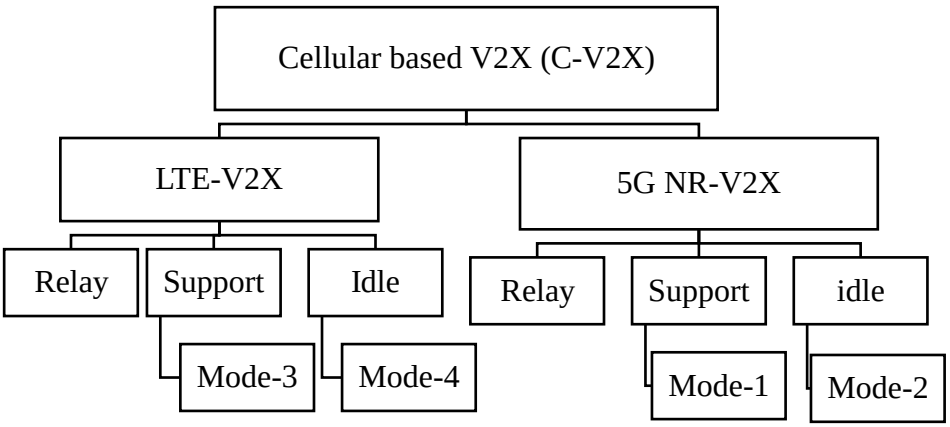

Figure 2-1. C-V2X Tree with Resource Allocation Modes

2.2.1.1 LTE-V2X

LTE V2X is a leading enabler for vehicular communication. LTE-V2X is expected to operate in the 5.9GHz band reserved for ITS services. At the physical channel it uses SC-FDMA (Single-Carrier-Frequency Division Multiple access) with 10MHz and 20MHz channels divided into 50 and 100 Resource Blocks (RBs) respectively. Channels are divided in frequency (RB) and time (subframes) dimensions. They are divided into units of 180 KHz RB and 1ms subframes. A sub-channel is a group of RBs in the same subframe over which user data and control information are transmitted. The size of a sub-channel is dependent on the size of the Transport Block (TB) size and Modulation and Coding Scheme (MCS) used. According to the 3GPP LTE service requirements for delivery of V2X services document, V2X applications consists of four distinct components, V2V (VUE-to-VUE), V2I (VUE-to-Infrastructure, either transmitter UE or receiver UE are UE-type RSU), V2N (VUE-to-Network, either transmitter UE or receiver UE are eNB-type RSU) and V2P (VUE-to-Pedestrian, either transmitter UE or receiver UE are pedestrian UE) [28]. These four application components are supported by 3 operation options as defined in the document: transmission over Proximity-based Communication Interface-5 (PC5), transmission over LTE-Uu interface and transmission over PC5 and LTE-Uu interfaces. However, LTE V2X as specified in the 3GPPP technical specification document focusses on the transmission over the PC5 interface [15].

The first option suggests direct communication over PC5 between road entities, hence supporting V2V, V2I and V2P. In this scenario, a VUE sends either a Cooperative Awareness Message (CAM) or a Decentralized Environmental Notification Message (DENM) to multiple

UEs in proximity via the PC5 side-link interface. According to 3GPP technical specification 22.185, LTE requirements for V2X services includes message transmission being under network control when VUE are served by the eNB of context, a VUE offering V2X services shall be preconfigured by a 3GPP network with message transmission and reception parameters when not served by a E-UTRAN supporting V2X communication [28]. Messages are expected to be transmitted between UEs supporting V2V/I/P at a maximum latency of 100ms. Some use cases such as pre-crash signalling require a lower maximum latency of 20ms. The LTE-V2X transmission supports a relative speed of UEs of 500km/h and maximum absolute speed of 250Km/h regardless of whether the UEs are served by E-UTRAN that supports V2X communication or not. LTE-V2X supports the transmission of CAM and DENM at a maximum transmission rate of 10Hz (10 messages per second) [15], [36]. The Cellular 3GPP network pre-authorization support for V2X communication is provided to UEs not served by a E-UTRAN base station supporting V2X communication [37].

Essentially, communication via the PC5 can either be network assisted (mode-3) or autonomous (mode-4) as specified in 3GPP release 14 [35]. In network assisted direct communication mode, UE resource allocation is centralized and controlled by the EUTRAN eNB. Here, UEs that seek to make direct communication via the LTE side-link PC5 interface, request radio resources and authorization from the eNB. To request radio resources and transmit in this mode, the V2X UE needs to be in the RRC_CONNECTED state. For the eNB to efficiently perform its resource allocation function it would require context information like speed and location from the requesting UE [38]. Mode-3 resource allocation can be approached in two ways: Underlay and overlay approach. In the overlay approach a band of cellular radio resources is reserved for V2X communication while in underlay mode both traditional cellular users and V2X users contend for same resource set. Though the underlay approach is deemed more spectrum efficient, effective resource management could be difficult to implement due mutual interference between V2X users and traditional users [39]. The mode-3 resource allocation schemes are specified to be either dynamic or SPS based. In the dynamic approach, the eNodeB grants side-link resource to UE for every TB (Transport Block) transmission request, while the SPS based approach allows VUEs to retain resources for a duration specified by the eNodeB. In mode 4, UEs autonomously select radio resources for direct communication between UEs without the intervention of eNBs. Though EUTRAN

supporting V2X is not required for radio resource selection, EUTRANs can configure transmission parameters for UEs before going out of EUTRAN-V2X coverage. These preconfigured parameter values are used when UEs are out of coverage [40]. Carrier frequency and number of RBs per sub-channel are two important parameters to be configured [41]. Though specific values have not been defined by the 3GPP, discussions on optimized default values are already ongoing [42]. The resource selection is made by V2X UEs using a Semi-Persistent Scheduling (SPS) Scheme. In this scheme, a pre-selected number of sub-channels are reused for certain number of consecutive transmissions set at the reselection counter. The number is randomly set between 5 and 15 for every new reservation of sub-channels. For every packet transmission the counter value is decremented by one until counter value is equal to zero, then a new set of sub-channels are reserved. New resources or sub-channels are also reserved when the reserved resources are not fit for the packet to be transmitted [40].

The second option supports V2X applications over the LTE-Uu interface. It connects the VUEs to the eNB-type-RSU via the LTE-Uu interface for use cases such as queue warning, V2N traffic flow optimization, V2X road safety services, remote diagnosis, just-in-time repair and V2X direct transmission under MNO (Mobile Network Operator) control [28], [43], [44].

The third option is simply the combination of both the Uu interface and the PC5 interface for specific use-cases. A typical example of such use case as mentioned in 3GPP technical report [35] is road safety service via infrastructure. [45]

2.2.1.2 5G NR (New Radio)-V2X

5G NR V2X is proposed essentially as an improvement to LTE-V2X defined in 3GPP Release 14 [45], offering services to meet requirements of semi and fully autonomous and connected vehicles. 3GPP Release 15 and 16 defines network service requirements to meet advanced V2X application beyond the LTE-V2X requirements defined in Release 14 [36], [46]. Five application areas consisting of vehicle platooning, advanced driving, extended sensors, remote driving, and vehicle quality of service support is defined in ETSI technical specification document TS 122.186 [46]. The 3GPP system supports services expected for all SAE specified level of driving automation [10]. The service requirements specified for the advanced

application areas varies with degree of automation of the VUE in the use case. 5GAA also defined 7 groups of use cases and associated service level requirements in [47] derived from use case classes specified in 3GPP release 14, 15 and 16 [48].

The most relevant network functionalities supporting 5G NR-V2X include Policy Control Function (PCF) and Network Data Analytics Function (NWDAF). As contrast to LTE-V2X, the improvements in 5G NR-V2X as defined in 3GPP technical report Release 16 TR 37.885 include: NR side-link design for V2X; Uu enhancements for advanced V2X application areas; Uu based resource allocation/configuration by LTE and NR; RAT and interface selection; QoS management; and non-cochannel coexistence between NR and LTE side-link [49].

5G system architecture supports V2X communication over two reference points namely: PC5 and Uu interface. The NR side-link supports enhanced V2X communication through the PC5 reference point. The support for enhanced communication over the Uu interface exploits the NR and Multi-Access Edge Computing (MEC) methods to meet the stringent latency requirements of V2X applications. 5G-NR V2X also supports Uu based and autonomous resource allocation for communication over the PC5 reference point, defined as NR mode-1 and NR mode-2 respectively [28], [43], [44] 5G-NR V2X allows the use of multiple 3GPP RATs concurrently for direct communication and permits RAT selection to be carried out by upper layers based on V2X application type. Only LTE and NR RATs are specified in the selection definition [46].

The NR-V2X physical layer is based on the NR-Uu physical layer structure defined in 3GPP rel-15 technical specification document [50]. The two operating Frequency Ranges (FR) for NR-V2X are: FR1 (450MHz – 6000MHz) and FR2 (24250MHz – 52600MHz). Though both frequency ranges are available for use, only FR1 is adopted specifically for NR-V2X [51]. NR-V2X side-link employs Orthogonal Frequency Division Multiplexing (OFDM) waveform with Cyclic Prefix (CP). Multiple numerology is adopted for NR-V2X, with flexible Subcarrier Spacing (SCS) and slots per subframe to support diverse use case requirements. The NR-V2X resource grid organization is derived from NR-Uu resource grid structure. The frame structure is comprehensively defined in 3GPPP technical specification TS 38.211 [50].

Though a considerable amount of work has been done on the mmwave (FR2) V2X band [52], [53], [54], [55], [56] no specification has been defined on how it will be used to support V2X applications. The major drive towards mmwave is the high bandwidth available at higher frequencies, which portends higher data rates. However, the use of mmwave presents certain technical challenges [52], [54]. Some of the open issues identified with the use of mmwave band for V2X communication include channel modelling, channel variation in numerology design, overhead associated with narrow beam alignment, beam tracking and recovery, channel sensing in resource allocation, and vehicle discovery and mobility management.

As specified in the 3GPP technical specification release 16 (TS 22.186 version 16.2.0) general requirements, the 5G NR V2X system shall control the communication range of UEs transmitting messages based on certain message characteristics [36]. The document also specified that the 3GPP system will support identification of V2X application and RAT selection for a V2X application. The system should also support a relative lateral position accuracy of 0.1m. Switching between direct network access and indirect network access via UEs supporting V2X services is also supported. The implementation of vehicle grouping different from platooning is also supported. This has been proposed in many studies reported in the literature [24], [57], [58]. Some improvements are proposed for NR V2X in 3GPP release 17 2019 plenary. Some of the related plenary headlines Include NR side-link enhancement, enhanced V2X services and IoT over Non-Terrestrial Networks (NTN) [59]. The enhancements focus on wider coverage, improved reliability, decreased latency and battery power saving for battery powered UEs. Reliability, latency, and power saving improvements were specifically mentioned in side-link enhancement Work Item (WI), while coverage extension is mentioned in the side-link relaying Study Item (SI). Side-link positioning SI was defined for multiple V2X coexistence and public safety use cases with stringent positioning requirements [60].

2.2.2 IEEE 802.11 Based V2X RATs

IEEE 802.11 based V2X RATs defines all short-range vehicular communication RATs using 802.11 based Physical (PHY) and Media Access Control (MAC) layer protocols [61]. These RAT standards are variations developed from earlier IEEE 802.11 WLAN (Wireless Local Area

Network) standards. Two standards (IEEE 802.11p and IEEE802.11bd) that have evolved from the earlier IEEE 802.11p standard will be further discussed in this section.

2.2.2.1 DSRC (ITS-G5)/IEEE802.11p/WAVE

DSRC is an IEEE 802.11p stand-alone RAT standard that supports direct short distance information exchange between connected road transport entities [61]. The technology is aimed primarily to enable both V2I/I2V and V2V communication mode for safety and traffic efficiency use-cases. DSRC standards were developed to be region-centric and no two DSRC from different geo-telecommunication regions are compatible. In North America, the standard was developed by the IEEE and the American Society for Testing and Materials International (ASTM). The DSRC standard was developed in Europe by ETSI and European Committee for Standardization (CEN) and in Japan by Association of Radio Industries and Businesses (ARIB). In the USA, a 75MHz frequency band between 5.850GHz to 5.925GHz has been allocated for DSRC [61]. The band is divided into seven 10MHz channels (six service channels and one control channel) with a 5MHz channel reserved for future use. In Europe, a 30MHz band has been allocated between 5.875 to 5.905GHz (ITS-G5A) as DSRC spectrum for safety-critical and traffic efficiency applications (one control and two service channels, 10MHz each), 20MHz band is reserved for 5.855-5.875GHz (ITS-G5B) for non-safety applications (two 10MHz service channels), ITS-G5C (between 5.470 to 5.725GHz) is reserved for Wireless Local Area Networks (WLANs), Broadband Radio Area Networks (BRANs) and Radio Local Area Networks (RLANs) while ITS-GD (between 5.905 to 5.925GHz) is reserved for future ITS applications [39]. In Japan, an 80MHz band between 5.770 to 5.850GHz is already in use for electronic toll collection while a 700MHz band between 755.5 to 764.5 MHz is allocated for ITS use-cases [61], [62].

Adopting the DSRC spectrum for vehicular networking requires the development of a specialized protocol stack. Hence, the development of IEEE802.11p (an adaptation of the IEEE802.11 standard) defining the access layer for DSRC and Wireless Access in Vehicular Environment (WAVE) a comprehensive suite of communication standards comprising of IEEE1609.2 and IEEE1609.3 standards defining frameworks for secure and interoperable communication in DSRC respectively [53], [54].

Several advances have been made in the IEEE802.11 standard since the standardization of the IEEE802.11p DSRC standard. The advent of new autonomous vehicular applications makes a case for the development of new standards to meet the requirements of the emerging application use-cases. An IEEE802.11bd task force was established to define the key improvement areas of the IEEE802.11p standard.

2.2.2.2 IEEE 802.11bd

IEEE 802.11p was derived from the PHY and MAC layer of IEEE 802.11a. The IEEE 802.11 Next Generation V2X Study Group was formed in 2018 to exploit the advances in 802.11a/n/ac/ax to improve 802.11p performance, this led to the creation of the 802,11bd task group in 2019. The core design objectives set for IEEE 802.11bd that will set it apart from 802.11p are support for relative speed of up to 500km/h, communication range of up to 2km, up to twice the throughput of 802.11p, coexistence, interoperability, and backward compatibility with IEEE802.11p [63], [64].Some of the techniques being mulled over to achieve these objectives includes, midamble for channel estimate accuracy, congestion sensitive retransmission, alternate Orthogonal Frequency Division Multiplexing (OFDM) numerologies, use of higher order modulation and coding scheme and mmwave spectrum for improved throughput, Dual Carrier Modulation (DCM) for range extension and Block Error Rate (BLER) improvement, interference management for multi-channel transmission, and Next Generation positioning for accurate location detection [61], [63], [64]

2.3 V2X Use cases.

Successful management of vehicular networks in high mobility context requires understanding of communication context and use-case requirements. Context information such as relative speed, distance, location, and trajectory of nodes affects use-case requirements, which in turn affects the QoS requirements. Hence, it is necessary to study these use-cases and their service requirements.

Having mentioned that the primary purpose of V2X communication is safety, non-safety applications such as traffic efficiency, improved commuter experience, and efficient (hence, eco-friendly) fuel consumption in V2X network designs. Several use-cases have been presented in different standards and consortium documents. For example, the 3GPP in its

release-14 technical report document specifies 27 distinct use-cases supported by LTE-V2X with their respective pre-conditions, service flows, post-conditions, and potential requirements [65]. A 3GPP release-15 technical specification further defined advanced service requirements for NR-V2X scenarios under four categories: vehicle platooning, advanced driving, extended sensors, and remote driving [15], [36]. The Fifth Generation Communication and Automotive Research and Institute (5GCAR) introduced five use-case categories for connected and autonomous driving applications: cooperative manoeuvre, cooperative perception, cooperative safety, autonomous navigation, and remote driving [6]. In 2009, European Telecommunication Standard Institute (ETSI) technical report (ETSI TR 102 638 BSA, [ETSI09-2638]) defines a Basic Set of Applications (BSA) which is essentially a classification of mature use-cases supported by mature vehicular communication systems and technologies. The groups include active road safety, cooperative traffic efficiency, cooperative local services, and global internet services [66]. International Telecommunication Union – Radio communication Sector (ITU-R) M.1890 recommendation [ITUR11-1890] also defined eight use-case classes with several use-cases for each [6].

Different use-cases have different topological and QoS requirements. For example, use-cases such as collision awareness and lane merge have different topological and QoS requirements. Lane merge would exploit both V2V and V2I/N while collision-warning use-case are likely to benefit only from V2V communication. In terms of QoS requirements, cooperative safety applications tend to be more delay intolerant, cooperative manoeuvre use-cases are more coverage dependent while cooperative navigation applications are bandwidth-intensive. As specified in 3GPP technical specification (TS 22.186 version 15.3.0 Release 15) different V2X use-cases can be identified using different ITS Application Identifier (ITS-AID) or Provider Service Identifier (PSID) [46].Given the variation in network requirements for use-cases it has become necessary to study the attributes, context and requirements of these use-cases and the different communication modes, network topology, RAT choices, resource allocation modes that are most appropriate to efficiently meet these requirements.

In this section we will be discussing use-cases and their respective requirements under two broad categories: safety critical applications and non-safety applications (traffic rfficiency and

infotainment). These use-case categories differ in requirement elasticity, with safety applications having stricter requirements.

2.3.1 Safety Critical Use-cases

Safety applications of V2X communication are intended to minimize vehicular accidents, consequently, minimizing vehicular transport-related loss of life and associated cost of accidents [67]. Safety-critical use-cases can also be defined as CAV applications in which V2X information exchange stimulates vehicle control systems [68]. The information exchanged includes the transport entity's velocity, acceleration, position, and trajectory. Safety-critical use-cases also require sensing and the exchange of context information such as road accidents, pedestrians, obstacles, weather, speed limits, road surface indicators, road topology, and other similar information [62]. The common requirements for safety-critical use-cases are ultra-high reliability and ultra-low latency. Each use-case might have additional requirements. These requirements are often very strict. Cooperative safety use-cases are diverse and evolving, hence they are supported exclusively or co-operatively by distinct communication modes. The advancement of wireless access technologies gives room for the development of advanced use-cases and improvement of user Quality of Experience (QoE). Table 2-1 presents common safety-related use-cases with their required communication modes and requirements [6], [15]. It presents the five different cooperative safety use-case categories defined in ETSI technical report [102 638] titled "ITS Vehicular Communication Basic Set of Applications" [66]. For each safety use-case category, we highlight two examples of use-cases there-in along with messaging mode, security reliability, communication, periodic messaging frequency and latency requirements.

2.3.2 Non-Safety Application

Non-safety applications consist primarily of traffic management, infotainment, data downloads, and software update services. These services are aimed at offering an enjoyable and efficient driving experience. They have relatively softer service requirements specifically on reliability and latency.

Table 2-1. Safety use cases and requirements [6], [15]

Cooperative Road Safety Use-case Categories /Application	User Cases	Messaging Modes	Security/ Reliability Requirement	Vehicular Communication Requirement	Minimum Frequency of Periodic Messages	Maximum Latency
Vehicle status warning /Road hazard warning	Emergency electronic brake light	Time limited periodic broadcasting on event	High/High	Ability to receive and transmit broadcast DENM	10 Hz	100ms
	Abnormal function warning	Time limited periodic broadcasting on event	High/High	Ability to receive and transmit broadcast DENM	1 Hz	100ms
Vehicle type warning /Co-operative awareness	Emergency vehicle warning	Periodic triggered by vehicle mode	High/High	Ability to receive & transmit V2X CAM	10 Hz	100ms
	Vulnerable Road User (VRU) warning	V2X co-operative awareness	High/High	Ability to receive & transmit V2I/I2V CAM	1 Hz	100ms
Traffic Hazard Warning /Road Hazard Warning	Wrong way driving alert	Time limited periodic broadcasting on event	High/High	Ability to receive & broadcast V2X DENM	10 Hz	100ms
	Roadwork warning	Temporary message broadcast or geocasting on event	High/High	Ability of RSU to broadcast 12V and VUEs to receive & process I2V	2 Hz	100ms
Dynamic Vehicle Warning/Cooperative Awareness	Overtaking vehicle warning	V2X co-operative awareness.	High/High	Ability to receive and transmit broadcast CAM	10 Hz	100ms
	Lane change assistance	V2X co-operative awareness.	High/High	Ability to receive and transmit broadcast CAM	10 Hz	100ms
Collision risk warning /Co-operative collision avoidance or mitigation.	Co-operative forward collision warning	Unicast V2X co-operative awareness.	High/High	Ability to receive and transmit broadcast CAM	10 Hz	100ms
	Intersection collision warning	Periodic co-operative awareness broadcasting	High/High	Ability to receive, process & broadcast V2X CAM, establish peer2peer unicast session	10 Hz	100ms

Most of the use-cases for non-safety applications are supported by V2I and V2N, with a few supported by V2V. Non-safety applications are dynamic and evolving, so do the requirements of existing use-cases, this prompts the need for improved access schemes that would guarantee that service requirements are met. Table 2-2 presents the common non-safety use case categories [67], along with 3 use-cases in each category, their supporting messaging mode, and service requirements [5], [15], [66], [67]

Table 2-2. Non-safety V2X use-cases and requirements [5], [15], [66], [67]

Non-safety Application Categories	User Cases	Messaging Modes	Security/ Reliability Requirement	Vehicular Communication Requirement	Minimum Frequency of Periodic Messages	Maximum Latency
Traffic efficiency/co-operative traffic Management	Regulatory and Contextual speed limits	Traffic message is prompted by a management entity	Medium/High	Capacity of RSU to receive, process & broadcast I2V CAM	1 to 10 Hz	200ms
	Traffic information and recommended itinerary	Constant periodic traffic message broadcast by management entity	Medium/High	Capacity of RSU to receive, process & broadcast traffic information	1 to 10 Hz	500ms
	Enhanced route guidance & navigation	On demand service	Medium /Medium	Internet access, IPV6 and geonetworking capacity	1 Hz	500ms
Infotainment	Point of interest notification	On demand or periodic promotion message – unicast, broadcast or geocast by ITS provider.	Medium /Medium	RSU/VUE capacity to receive, process and broadcast I2V, establish P2P session and broadcast locally relevant content	I Hz	500ms
	Automatic access control/ parking access	ITS provider triggered unicast full duplex	Medium /Medium	RSU/VUE capacity to receive, process and broadcast and/or unicast I2V CAM. Establishment of P2P session	I Hz	500ms
	Remote diagnosis and in-time repair	M2M internet communication and ITS triggered periodic monitoring (CAM)	Medium /Medium	Internet access with IPV6 and geonetworking protocol capacity. I2V CAM processing	1 Hz	500ms

2.4 V2X Propagation Models

In the USA, Europe, and Asia, the 5.9GHz frequency band has been designated for use by the ITS [5], [69]. Hence, we have restricted our discussion on the V2X propagation models around the 5.9GHz band as presented in the ETSI technical report [103 257] [70]. The use of the 5.9GHz frequency band for V2X applications comes with inherent challenges stemming from its high frequency and low wavelength of about 5cm. The V2X driving environment is characterised by a highly dynamic topology accompanied by a complex propagation environment with features distinct from a typical wireless environment. The distinct properties of the V2X propagation environment ranges from a different combination of link types, different degree of mobility of nodes, vast varying types of nodes and different driving

environment. These factors hugely affect wave propagation, which then makes existing propagation models unsuitable for V2X communication and necessitate the development of distinct channel models for V2X communication.

2.4.1 Driving Environment Scenarios

Many V2X studies, algorithms and solutions are often scenario-based. The scenario selected affects the wave propagation, channel model and consequently the entire network model. There are three common driving scenarios in the communication context: the urban, rural and highway scenario. Another scenario that is gradually gaining attention is the tunnel driving environment [70], [71]. The different scenarios are commonly typified by different levels of vehicular density, roadside structure density, and relative speed of vehicles. These driving environment properties tend to have different effects on wave transmission between vehicles (V2V) and between vehicles and roadside infrastructure and nodes (V2I and V2N). In this section we will be describing the major types of environment scenario as described by ETSI in [70].

2.4.1.1 Urban Scenario

ETSI technical report [103 257] defines an urban scenario as a densely populated area that allows for one-way or two-way vehicular traffic on either a single or multi-lane metropolitan street [70], [72]. They are often characterised with different types of mobile nodes and stationary roadside units ranging from motorcycles, vehicles, bicycles and pedestrians to traffic lights, road signs, streetlights and traffic signals. They are often characterised with roadside structures such as storey buildings. Urban areas are prone to wave scattering effects since there are many stationary and mobile objects with wave scattering potential and most of their antennas (predominantly vehicular antennas) are of the same heights and close to ground. Here, the probability of LOS blockage is higher, and the successful signal reception can be dependent on multipath component(s) [73]. The effect of Doppler is expected to be small or negligible since vehicles generally move at low speeds and the relative speed between vehicles as a result will be low. All types of communication or link modes are often present in an urban environment. A depiction of which is shown in Figure 2-2.

Figure 2-2. A depiction of V2X communication in urban environment [74]

The urban environment has gained early interest from the vehicular communication and the ITS research community in general. Given the complexity of the urban environment, more work still needs to be done to accurately model the perturbation wave transmission experience in the highly dynamic environment. Two recent works done in [75] and [76] investigate propagation characteristics in two different urban environments: Vienna, Austria and Wuhan, China respectively. While the Vienna campaign seeks to prove the workability of an adapted testbed for repeatable and controllable propagation characteristics over varying speeds, the Wuhan campaign seeks to provide a delay profile of the congested urban centre at rush hour.

2.4.1.2 Rural driving environment

Rural driving environments are characterised by open fields, hence have few roadside structures to obstruct V2N transmissions. Due to the expectedly low vehicular traffic flow, the roads are often single lane with bi-directional traffic flow. Since there is generally a lower density of on-road and roadside entities, the probability of line-of-sight blockage will be lower and the scattering effect is also lower, hence the delay spread experienced from multipath will also be minimal. However, the scantiness of vehicles on rural roads might make the rural environment prone to high relative velocity between road entities and potentially generate greater Doppler Effect on transmission. Research specifically looking at characterising rural driving environments as defined by ETSI have not been particularly investigated as most work considers rural motorways and highways synonymous in terms of channel characterisation [72].

2.4.1.3 Highway Driving Environment

ETSI describes highway environments as roads with two or more lanes reserved for unidirectional traffic and with relatively higher maximum allowable speed of between 120 km/h and 140 km/h. Hence, the Doppler spread experienced by V2X channels in highway driving environments can be relatively higher. The density of scatter initiating objects such as guardrails, bridges and overhead road signs are higher in highway scenarios as compared to rural roads. Highways often have long and straight road stretch with relatively unobstructed line-of-sight (LOS), reducing blocking rate and allowing signals to propagate over longer distances. The dynamic topology of highways requires that further work be done in characterising its propagation environment. For example, the work done in [77] investigates the propagation characteristics and stationary distance that satisfies Wide-Sense Stationary Uncorrelated Scattering (WSSUS) in highway for both V2I and V2V link modes. The results shows that the stationary distance is a function of number of multipath components and relative speed. Similar research seeking to evaluate stationary distance at Yeoju expressway in South Korea is reported in [78].

2.4.1.4 Tunnels

Tunnels are defined as a one-way traffic road within a tunnel with two or more lanes. The road is characterised with scattering from the walls, ground, and roof of the tunnel. V2I LOS is expected to be severely blocked by tunnel structure. Understanding the channel characteristics is currently gaining momentum. A measurement campaign in Shanghai, China leveraged a tunnel measurement of V2I and V2V channel presented in 3GPP TR 36.885 [35] to describe the large- and small-scale fading of the channel and then further measured and analysed channel parameters (such as path loss, power delay profile, delay spread) against distance [79]. Another work analysed V2X pathloss under tunnel scenario and found the V2I transmission to be experiencing greater loss with increasing tunnel curvature [80].

2.4.2 Channel Model Types

A reasonably accurate modelling of a V2X wave propagation channel parameter is primarily hinged on carrier frequency and driving environment in which the waves propagates. The inherent complexity of the driving environment necessitates a careful study of vehicular propagation channel parameter characterisation. In the next subsections, we will discuss

different approaches to which propagation channels of driving environments have been modelled. These approaches will be described under two widely used broad categories: Geometric and Non-Geometric Models [72], [81], [82]. The classification approach adopted for this review is presented in Figure 2-3.

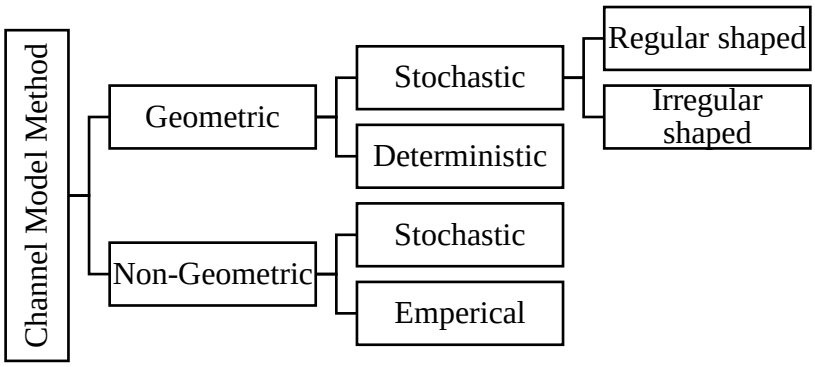

Figure 2-3. Vehicular propagation channel classification

2.4.2.1 Geometric-Based Channel Models

A geometric-based channel model focusses on the physical geometry and layout of the propagation environment. Factors such as shape, position, size and materials of objects within the environment or propagation path are considered when characterising the propagation channel in the environment. The extent to which these factors affect wave properties such as scattering, reflection, dispersion and refraction are also analysed. Geometry-based models are usually either deterministic or stochastic.

Deterministic propagation models based on geometry exploit availability of detailed scenario information. This approach is scenario-specific and computationally demanding. A commonly used geometry-based deterministic model is ray tracing [82]. The ray tracing method exploits geometric optics and numerical methods for solving approximations of Maxwell's equations at high frequencies [83]. Many of the recent works that have used ray tracing to develop propagation models have exploited it for characterising high frequency wave propagation in the mmWave bands. The work in [56] proposed improving V2X mmWave channel model accuracy by exploiting diffuse scattering (DS) as a key propagation mechanism. The results obtained indicate superior performance in terms of received power and delay spread particularly when beam alignment is not in place. An evolutionary ray-tracing propagation paradigm described as Dynamic Ray Tracing (DRT) is presented in [84]. It is a model that offers

multidimensional channel prediction including Doppler spread and time delays for a every ray, by exploiting database of detailed environment attribute at every time instant.

Geometric based stochastic models on the other hand use a pseudo-random distribution of scatterers around transceivers to model the propagation statistics of real driving environments. Geometric based stochastic models are further categorised into irregular shaped and regular-shaped models based on geometry of scatter distribution around transceivers. The work presented in [85] discusses an irregular geometric-based stochastic model for V2V communication in the sub-6GHz frequency range. The model is based on a Single-Input-Single-Output (SISO) system and the results were proven by comparing with a field measurement campaign carried out in Helsinki, Finland. A similar work carried out over MIMO in Lund; Sweden is presented in [86]. Beyond validation of the approach, the result here further proves that the channel demonstrates bib Wide Sense Stationary Uncorrelated Scattering (non-WSSUS) attributes. However, the assumption that all scatterers are on a 2D-plane limits the accuracy of this approach. 3D representation of scatters is suggested to improve accuracy and realism of this approach. A 3D regular-shaped model for V2V communication presented in both [87] and [88] uses an elliptical cylinder model and 2D Von Mises Fisher representation to emulate the distribution of scatterers around transceivers. Both works demonstrated a better representation of a real V2V propagation environment, however with higher computational and complexity burden might be a deterring factor in cases where computing resources are constrained.

2.4.2.2 Non-geometric Channel Model

Unlike geometric models that focus on the geometry of constituent objects in the propagation environment, non-geometric channel models rely on measurements and statistical analysis to characterise the propagation channel. Non-geometric models are developed by extensive measurement campaign of real-world propagation environment data through actual field measurements or the use of channel sounders. Statistical distributions are then used to represent measured channel parameters such as pathloss, fading and shadowing effects. Non-geometric channel models can be further sub-categorised into Empirical and Stochastic models by considering various diffuse scattering models using a ray-tracer tool.

Empirical modelling approach use data gathered from measurement campaigns to identify singular patterns of the propagation environment to be modelled [81]. Hence, they are generally simpler with fewer parameters to characterise the propagation environment. They're commonly used to model propagation in traditional scenarios [89]. Many works have been done to characterise the complexity of wave propagation in driving environments. A Machine Learning (ML) based empirical V2X channel modelling is presented in [90]. In this work, a measurement campaign was conducted, and data was collected from a road network in Dresden, Germany. The aim is to use the labelled data from the campaign to train the deep neural network algorithm to predict RSSI, using a combination of open-source available road information (road type, buildings) and easily obtainable Euclidean distance as output as shown in Figure 2-4. Another empirical model presented in [91] ran sub-6GHz narrow beam measurement campaign for V2I highway communication in Beijing, China. The empirical model is characterised based on measured dispersion parameters such as Root-Mean-Square (RMS), delay spread, Doppler spread and angular spread. The studies analyse the effect of different antenna and beamforming parameters on measured dispersion parameters.

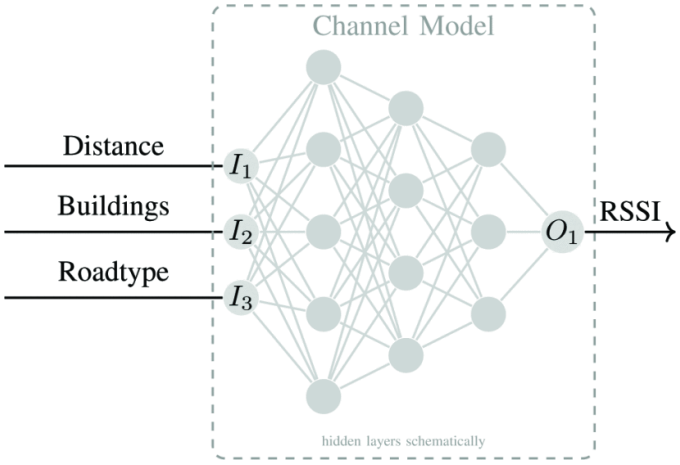

Figure 2-4. Neural network architecture for a non-geometric deterministic model in [90]

Non-geometric stochastic models characterise channel models using stochastics and with no presumption of scattering distribution and constituent object geometry [92], [93]. One of the most popular non-geometric stochastic approaches to channel modelling is the Tapped Delay Line (TDL) model. TDL defines channel response using a Finite Impulse Response (FIR) filter with specific number of taps, each with different dispersion attributes. For example, the channel response can be modelled as superposition of *P* taps as expressed In Equation *(2-1*).

$$h(\tau,t)=\sum_{l=1}^{P}\left(\sum_{n=1}^{N}\alpha e^{j2\pi f_d t}\right)\delta(\tau-\tau_l) \qquad (2\text{-}1)$$

Where each tap has *N* number of unresolvable sub-paths with amplitudes, α and Doppler frequency, f_d. The work in [94] presents six different models based on TDL approach representing three V2V and three V2I link modes measured in a highway in Atlanta, Georgia, USA. The models have been accepted as the standard V2V model for IEEE 802.11p [95].

2.5 C-V2X Radio Resource Allocation Models

As CAVs gain more attention from stakeholders in industry, academics and states the need to develop communication infrastructure to support the growing demands has become necessary. Radio resources are a critical communication capital that suffers stiff contention between increasingly pervasive CAVs. For DSRC, resources are allocated using Carrier Sense Multiple Access/ Collision Avoidance (CSMA/CA) scheme. Detailed discussion on resource management in DSRC can be found in [96]. This section focuses on 3GPP defined resource allocation models for cellular V2X. While detailed physical layer framework of LTE-V2X and NR-V2X side-link can be found in [28], the focus of the rest of this section will be cellular V2X side-link resource allocation in view of coverage availability (resource allocation modes).

2.5.1 C-V2X Resource Allocation Modes

C-V2X broadly offers two modes by which resources are allocated to vehicular users for side-link communication in LTE-V2X and 5G NR V2X respectively. These modes are categorised based on whether the vehicular node requires cellular coverage of the base station (eNodeB or gNB) for resource allocation [28]. The first mode requires cellular coverage and is named LTE-V2X mode-3/ NR-V2X mode-1, while the second mode operates outside cellular coverage and is named LTE-V2X mode-4/NR-V2X mode-2, as illustrated in Figure 2-5 [97]. In LTE-V2X mode-3 /NR-V2X mode-1, the eNB/gNB is responsible for scheduling and channel allocation for V2V and V2I communication either by dynamic allocation/Dynamic Grant (DG) or through Semi Persistent Scheduling (SPS) / Configured Grant (CG) as explained in [28], [48], [98]. Though quite similar resource allocation concepts, NR-V2X has a more developed QoS consideration in its allocation process as compared to LTE-V2X.

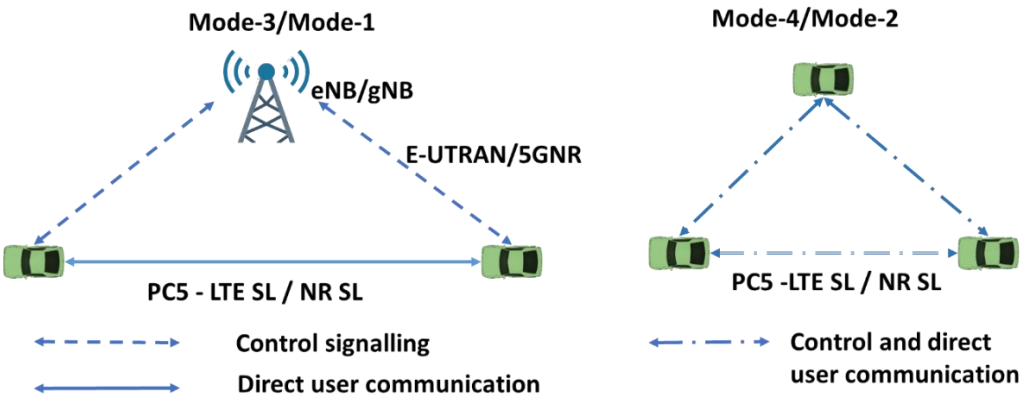

Figure 2-5. Cellular-V2X resource allocation modes

However, it is noteworthy that 3GPP has not defined intricate standards for resource allocation in mode 3 and 1, leaving room for use-case or application specific development. Mode 4 and 2 on the other hand are standardised in greater detail. Here vehicular nodes autonomously select resources using sensing-based Semi Persistent Scheduling (SPS) [99]. A flowchart illustration of working principle of sensing based SPS is presented in Figure 2-6.

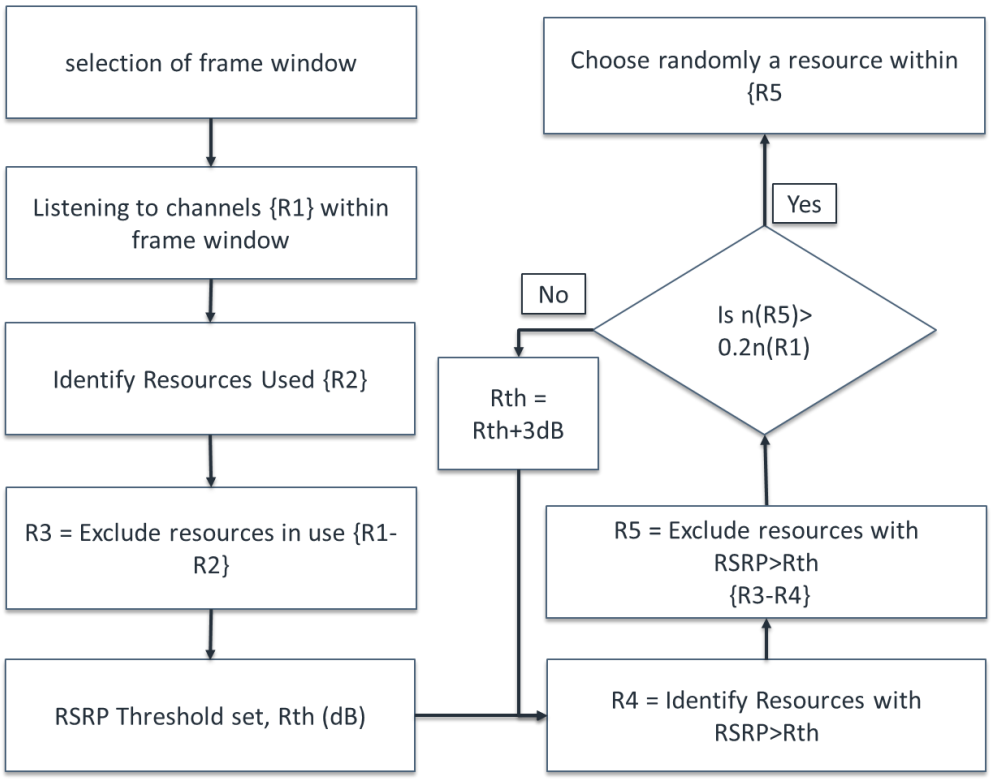

Figure 2-6. Sensing-based SPS algorithm for mode-4 in LTE-V2X [97]

2.5.2 State-of-the-art

Both mode-3/mode-1 and mode-4/mode-2 are still subject to improvement and have attracted huge academic research interests. It is noteworthy that more research effort has been dedicated to mode-4/mode-2 because they are susceptible to greater challenges such as collision, interference, and congestion issues. However, developing resource allocation schemes to optimize available radio resources in mode-3/mode-1 is critical and is open to significant development considering that there is no defined 3GPP standard for mode-3/mode-1 in this regard. Since the work in this thesis is developed upon mode-3 resource allocation, our discussion onward will be focussed on some of the recent developments in the LTE-V2X mode-3 and 5GNR-V2X mode-1 resource allocation category.

The work discussed in [98] proposes an LTE-V2X mode-3 resource allocation scheme that meets the QoS requirements of V2V applications while preventing allocation conflict. QoS in this context is defined as channel capacity per-vehicle. The authors proposed a mathematical framework guided by four conditions: mandating within-cluster subframe reuse, per-vehicle QoS differentiation, prevents inter-cluster sub-channel conflicts and minimizing time dispersion between assigned sub-channels. The approach avoided conflict and improved per-vehicle throughput; however, this approach is hugely dependent on the number of clusters and the clustering algorithm used. The same authors in [100] presented a centralised mode-3 resource allocation scheme using a bipartite graph. Here spectrum allocation and vehicles are denoted by graph vertices while throughput is signified by edges. The approach is used to ensure that the spectrum resources used by vehicles within same clusters are time orthogonal. The simulation results show improvement in throughput but provide no clear indication of how changes in cluster proximity will be handled. Another spatial re-use allocation scheme called DIRAC (aDaptive spatIal Reuse of rAdio resourCes) is presented in [101]. The context-aware scheme solves the problem in [100] by adapting its operation to context condition to ensure all vehicles experience similar interference levels. The results show improvement in overall quality of communication and V2X scalability. While able to handle a level of dynamicity in vehicular environment. It is yet to be demonstrated how this approach copes with different traffic demands and rapidly changing context. In another work, an exchange of resource rank of preference between neighbour vehicles is proposed upon which matching game theory algorithm is built [102]. The aim is to establish fairness and

improved sum-rate in V2I links while guaranteeing reliability of V2V links. The simulation results show the system adapts to vehicular network topology changes and achieves improved reliability and fairness. Adaptive and Maximum Reuse Distance (AMRD) scheme described in [103] seeks to resolve the issues with fixed reuse distance described in [104] and maximum reuse distance (MRD) discussed in [105] by estimating a flexible reuse distance of radio resources that adapts to changes in vehicular density while maximizing the distance between receivers using same spectrum resource. Though AMRD shows superior Packet Error Rate (PER) compared to fixed reuse distance schemes over short distances, but this gain fades with increasing distance and reducing vehicular density. Another improvement on MRD is proposed in [104]. The location based MRD reuse schemes exploits knowledge of the resource querying vehicle to maximise the reuse distance between vehicles using same resource set. The results obtained show improvement in reliability and overall capacity in a high-way scenario.

The discussion on the state-of-the-art of 5G NR V2X mode-1 will be initiated here with the work in [106] that investigates how resource allocation and scheduling schemes based on flexible numerology (sub-carrier spacing, SCS and transmission time interval, TTL) can be used to better meet V2X use-case requirements. The result shows that flexible numerology can be exploited to improve reliability and reduce latency. In another mode-1 scheme presented in [97], Channel State Information (CSI) transmission overhead is aimed to be reduced while maximizing throughput. To realise these objectives, the resource allocation is modelled as a mixed binary integer nonlinear programming (MBINP) problem. Resource allocation schemes presented in [107] propose the use of a Non-Orthogonal Multiple Access (NOMA) power control technique for resource allocation to improve capacity, latency, and reliability of NR V2X side-links. [107] addresses inefficiency issues confronted by LTE-V2X networks in a congestion scenario for unicast and broadcast systems. The simulation results indicate that the NOMA-based system efficiently dropped resource access collision which in-turn reduces latency and improves reliability. A centralised graph-based matching and NOMA resource allocation scheme aimed at reducing latency and improving system capacity is presented in [108] . The results obtained confirms improvement in system capacity through a cooperative game power control for NOMA transmission. In [109] a 2-stage centralised resource allocation and power control scheme is developed to address the inefficiency associated with gathering

location and channel state information in the dynamic network topology. The evaluation of the results obtained from the solution of the optimisation problem shows that the proposed system conserves energy and improves system SL throughput performance.

2.6 Vehicular Mobility Models

This section briefly discusses the framework upon which vehicular mobility can be clearly defined and realistically depicted for further analysis. Available traffic simulator tools are discussed followed by a brief review of cognate research strides.

2.6.1 The framework

Vehicular mobility models are broadly considered to be either microscopic or macroscopic [110]. The macroscopic modelling approach is a high-level approach that is based on fluid dynamics. It models gross quantities such as vehicular density and average velocity. Microscopic modelling on the other hand is more granular. It takes the computationally intensive particle-based modelling approach, considering the vehicles as distinct entities and model traffic properties such as trajectory, inter-vehicular distance, and acceleration.

As presented in Figure 2-7, generating a realistic vehicular traffic model requires consideration of the following components: accurate topological maps, obstacles, attraction/repulsion points, vehicle characteristics, trips, path, acceleration/deceleration, human driving patterns, intersection coordination, time, external interaction. Details of each of the components are discussed in [111].

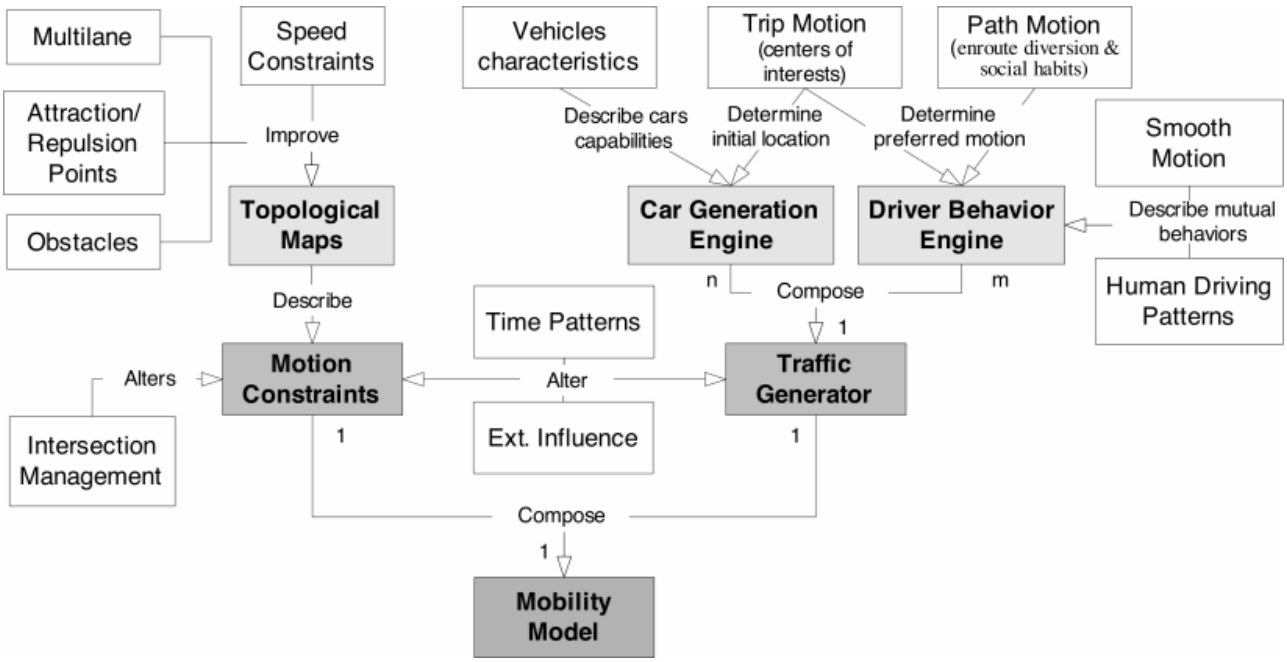

Figure 2-7. Concept guide for generating realistic vehicular mobility models [111]

It is noteworthy that all building blocks in Figure 2-7 are not necessary for all case studies. As every case study have a specific scope an interest. However, the more components a model considers, the more realistic it tends to become. Most models today take a simple approach and only consider components of interest.

2.6.2 Generating Vehicular Mobility Models

The development of vehicular mobility models can generally be categorised to four different classes: Synthetic Modelling approach which is essentially based on mathematical construction, Survey-based Modelling approach which involves model construction from physical measurements, Trace-based Modelling approach that constructs mobility traces from real vehicular traffic and Simulation-based modelling approach that generates vehicular traces from traffic simulators [111]. We have decided to approach our work using the simulation-based modelling approach due to its flexibility and ability to generate complex scenarios with relative ease and accuracy. It is also for this reason that the simulation-based approach will be discussed in greater detail.

2.6.3 Simulation-based Modelling Approach

Simulation-based models are generated from traffic simulators developed from a stringent process of refining of synthetic models and validation using real traces and traffic surveys. Hence, traffic simulators generate reasonably realistic traffic traces for different scenarios. The large set of parameter configuration capability available to users of the simulators makes it suitable to develop a large array of context of varying complexity. Traffic simulators such as PARAMICS [112], SUMO [113], VISSIM [114] and TRANSIMS [115] have the capacity to generate an array of microscopic model parameters.

In vehicular communication networks, two simulation modules are required to form a fully functioning vehicular network simulation: Network simulators and traffic simulators. Network simulators are responsible for constructing a realistic communication channel, routing, and topology model, while the traffic simulator generates vehicular traffic and the mobility environment [116]. In terms of integration with network models, mobility models can be classified into three categories: *Low Integrated Simulators (or Isolated Vehicular Mobility)* in which the mobility simulator module is standalone and not integrated with the network module such as the Random Way-Point model [117], BonnMotion [118], SUMO [113], Street Random Way-point (STRAW) tool [119] etc. Another category is the *High Integrated Simulators (or Embedded Vehicular Mobility)* in which both the mobility module is integrated with the network module. GrooveSim/GrooveNet tool [120] is one of the earliest high integrated simulators [121]. With the objective of increasing model accuracy other simulators have been developed, some of which include City Model [122] to test vehicular network models and MoVes [123] built upon distributed simulation middleware called Artis [123]. The third category of simulators are called federated mobility simulators [111] which defines simulators that develop interfaces that integrate isolated mobility simulators with network simulators such as MATLAB and OPNET, QualNet and NS-2. Examples include TraNS that federates SUMO for mobility model and NS-2 for the network models [124], VGrid Project that integrates SWANS network simulator with a scenario-limited synthetic traffic model [125], MSIE that integrates VISSIM with NS-2 [126], FDK that integrates QualNet with CORSIM traffic generator [127] and open source Vehicles in Network Simulation (VeiNS) that integrates OMNET++ with SUMO [128] etc.

It is noteworthy that most federated and embedded schemes are either not publicly accessible or only commercially available. Most of the open-source simulators, such as the SUMO are isolated, readily available and have parsers that makes them easily integrate-able with network simulators.

2.7 Vehicular Clustering

This section discusses the need for clustering of vehicular networks, the baseline process for most clustering schemes and state-of-the-art of clustering schemes in vehicular networks. Clustering is a grouping technique that enables vehicles of closely aligned intrinsic, operational, spatial or mobility attributes to be grouped and managed efficiently. Usually one or more vehicle nodes is/are elected as Cluster Head(s) (CH) within each cluster while other nodes within clusters are referred to as cluster members (CMs). Vehicles that do not belong to a cluster are called Free Vehicles (FVs) while edge vehicles that belong or serve more than one cluster are called gateways (GWs). Clustering algorithms are generally categorised to be either active or passive. In active clustering, all nodes take part in the clustering process by sharing relevant information with neighbours and taking roles within clusters. Active clustering monitors and adapts to traffic and topology changes, and is therefore suitable for the dynamic network topology of vehicular networks. Passive clustering schemes on the other hand is more static and relies on predefined configuration.

2.7.1 Why clustering?

Clustering in vehicular networks can be exploited to solve issues related to dynamic network topologies, network management, QoS, security and hidden nodes. Addressing these issues are important, given the critical nature and stringent QoS requirements of vehicular network applications. All vehicular clustering schemes can also be classified as either general purpose or purpose built (application-specific). Since application-specific clustering algorithms are designed to meet a specific objective, they tend to perform better in those specific scenarios for which they are designed. It is apparent that application-specific clustering algorithms are inflexible, and application restricted, however, with developments in Software Defined Networking (SDN) and development of more application-specific clustering algorithms, these restrictions can be mitigated while maintaining their application specific performances. Different survey works have classified clustering schemes based on the objectives or key

benefit for which the clustering scheme has been developed [31], [129], [130], [131]. In [31] clustering algorithms are categorised based on objectives as scalability and reliability, stability, delay, and routing overhead. In [129], clustering algorithms are classified based on the following set objectives: collision avoidance, reliability, communication range and stability. The work in [130] classified clustering algorithm-based objectives: cluster stability, load balancing, social awareness, fairness and QoS. [131] [131]also categorised clustering algorithms based on proposed benefits such as stability, data dissemination reliability and QoS. From our observation we regard stability as the most challenging issue with clustering. Beyond categorization, we present a comprehensive list of benefits derivable from clustering algorithm in vehicular networks in Figure 2-8.

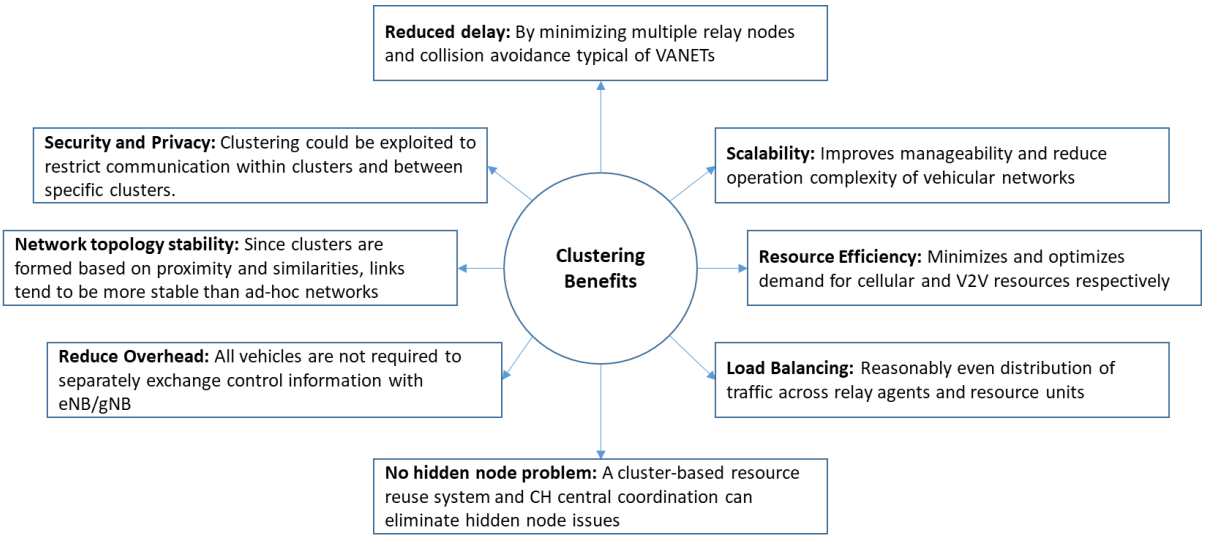

Figure 2-8. Benefits of clustering in vehicular networks [130]

2.7.2 State-of-the-art

From our discussion about objective-based vehicular clustering categorization in 2.7.1, the most targeted clustering objectives are stability, reliability and QoS improvement. Many literatures have studied these performance parameters and proposed new clustering schemes to meet these objectives. In this section we will discuss some of the recent work that sought to achieve these objectives

In [132], a stochastic study of a one-hop vehicular cluster on a single-lane highway with the aim of studying cluster stability is presented. While a stochastic mobility model is adopted, discrete-time Markov chain analysis was used to model cluster membership and cluster

overlap. The result shows that the stability performance observed in the analytical approach adopted is in tandem with MATLAB simulation. Though the work here only validated its analytical approach through simulation and produced useful results that could be exploited for efficient cluster algorithm development, it has not specifically looked at cluster stability improvement.

The works in [133] and [134] present a multi-metric clustering scheme using metrics such as average neighbour distance, relative position, relative speed, and vehicle trajectory to group vehicles into clusters. The multi-metric optimization problem in [134] is optimised using a Non-dominated Sorted Genetic Algorithm version 2 (NSGA-II) while [134] uses YATES algorithm. [134] demonstrates stability improvement over other optimisation methods such as Multi-objective Particle Swarm Optimization (MOPSO) and Multi-objective Differential Evolution (MODE), however the approach assumes re-clustering is inimical to cluster stability and avoids re-clustering in cluster maintenance process. This approach affects intra-cluster links and sustains CHs if they remain a CM even when not best fit for cluster leadership. [134] estimates the stability index of every vehicle before CH selection. This approach shows that the best vehicle is selected as CH for stability but fails to describe how the cluster is maintained over travel time or describe the effect on any QoS parameter.

In more recent work presented [[135], a stability-based clustering scheme proposed for urban environment seeks to improve stability and minimize overhead. The mobility metrics used include, vehicle position, trajectory, and relative speed, while the stability is evaluated based on CH lifetime. A distance threshold is set at which nearing clusters merge. This work only studies how stability vary with traffic flow and maximum lane speed. The results show stability across the range of both variables. Again, the assumptions in [133] is sustained and no result demonstrating stability is enhanced is presented.

[136] presents a Stable Clustering Algorithm for vehicular ad hoc networks (SCalE) that proposes the use of backup CHs to take over when the current CH leaves the cluster to maintain cluster stability and minimize re-clustering. Though stability is improved, how the improvement in stability affects QoS parameters such as throughput and jitter is not evaluated.

Residual Route Time (RRT) is used in [137] to select CH through a navigation-based system. RRT is a function of the duration a vehicle can stay around a cluster neighbourhood during the travel time. Vehicles share their navigation information with other vehicles then each vehicle use the RRT set available to it to either announce itself as CH or recognise one. The scheme demonstrates stability improvement but performs poorly in terms of number of successfully transmitted packets.

The work done in [138] leveraged and extended the advancement of the Optimized Link State Routing (OLSR) with multi-relaying work in [139] to develop an optimized CH and multi-point relay (MPR) selection algorithm that demonstrates superior stability and QoS performance in urban environment using link and street-centric parameters. Though the approach showed improved performance of QoS parameters such as throughput and Packet Delivery Rate (PDR), how this performance is maintained across travel time is not clear.

A fuzzy CH selection in a Cognitive Radio-based vehicular network is presented in [140]. As in [138] the fuzzy system considers street-centric parameters to make decision. It takes in average relative velocity, distance, network connectivity, lane weight and trustworthiness as input to select a stable and reliable CH. Though stability objective was achieved, the work did not describe how the network quality fared over travel time.

Another Fuzzy-based system is introduced in [141]. It combines the hesitant fuzzy with different multi criteria decision making schemes to select CHs that promises stability. Methods such as EVAluation of MIXed data (EVAMIX), technique for order of preference by similarity to ideal solution (TOPSIS) and Vlse Kriterijumska Optimizacija Kompromisno Resenje (VIKOR) methods are studied under a highway driving scenario. The result showed greater cluster stability in EVAMIX as compared to TOPSIS and VIKOR. Like many other papers, [141] did not indicate the QoS performance of the network over time.

The works in [27], [142], [143] make use of K-means and Floyd-Warshall algorithm for cluster formation and CH selection. The Floyd-Warshall algorithm estimates the shortest distance for all vehicle pair. The vehicle with the shortest average distance to all other vehicles will be nominated CH. In all these works improved stability and minimized re-clustering were

achieved without any assessment of the effect on link performance. The only difference with the works here is the context of application.

A QoS focussed scheme that clusters vehicles and elects CHs based on QoS, and vehicle trustworthiness is presented in [144]. This approach demonstrated potential for improved PDR and stability. To maintain stability, it proposes a ranking-based gateway recovery algorithm to replace failed CH. While the key objectives are achieved, the approach do not clearly define what a CH fail indicates but seem to align with the loss of cluster membership approach assumed by the other literatures.

In summary, all the approaches sought to minimize re-clustering and improve stability by sustaining a single CH if it remains active and within the cluster. The downside to this approach is a CH will remain sustained even when it is not the optimal CH choice, and this might affect CH-to-CM link performance and might even affect decision making when inter-CH proximity is used to make re-clustering decisions. Also, when closest CM to an edge-located CH is used as a backup CH as in [145], [146]. The backup CH might also be sub-optimal. Hence a scheme that improves stability while sustaining QoS performance is necessary.

2.8 Conclusion

This section summarises our review of the literature, specifically discussing how we have arrived at narrowing down the scope of our work. From the discussion on V2X access technologies in section 2.2, we inferred that C-V2X has a greater development prospect since its standards are developed along with established cellular access technology standards. We believe that its relative newness, greater coverage, and potential for coordination presents great opportunity for novelty.

The case study we have decided to adopt from our review is an emerging real-time geocast/multicast use-case. Proposals on streaming and sharing of real-time safety and non-safety media content between vehicles over PC5 interface has been discussed in 3GPP technical report [147]. It is upon this discussion we have decided to investigate a non-safety traffic efficiency use-case: real-time download of enhanced route guidance and navigation information, as our case-study of interest.

A standardised geometric-based regular stochastic model presented in [72] has been adopted here. This is based on the believe that a simple pathloss and shadow model will be sufficient to reasonably present the essence of our findings. We have decided to work with models of urban and highway scenarios in our work. We reckon that this scenario widely covers typical areas where vehicular communication will be required.

We have adopted a network assisted mode-3 underlay cluster re-use approach for our resource allocation scheme. We sought this approach to exploit the bird-view oversight of cellular coverage for resource allocation and avoid the interference from mobile cellular users. Similar approaches are discussed in [24], [25], [57], [58]

For the simulation model we chose to adopt the microscopic, simulation-based modelling approach as this tends to properly represent the contribution of each vehicle's mobility to the dynamic topology of the network and tends to be more accurate. Simulation is adopted for flexibility, ease of implementation and cost efficiency. SUMO specifically is adopted because it is open source and is equipped with functions that help to reasonably model realistic scenarios. SUMO is also used in [26], [116].

Since we are focussed mainly on post-formation processes of clusters, we have decided to form our clusters based on simple centroid-based clustering approach upon which we developed maintenance and re-clustering algorithms to sustain stability and link performance.

In the succeeding sections we will be describing in detail how we have implemented each of the highlighted chosen approaches as components of our work.

3 Memory-Based Re-clustering Schemes

3.1 Introduction

Though next-generation V2X communication networks are expected to consist of a variety of communication modes, only four apply to most use-cases; Vehicle to Infrastructure (V2I), Vehicle to Pedestrian (V2P), Vehicle to Network (V2N), and Vehicle to Vehicle (V2V) [15]. All these V2X modes, particularly the V2V mode are subject to a harsh communication environment resulting from high mobility and a fast-changing network topology associated with vehicular networks. Clustering has been identified as one way to deal with the issue of the changing vehicular network environment. Extensive research has been done on the implementation of different clustering algorithms in Vehicular Ad-hoc Networks (VANET) to solve issues related to the challenging communication environment such as the hidden node problem and fast changing channel caused high vehicular mobility [16], [17], [18], [20], [21], [22]. Different types of clustering schemes have also been suggested for the stability of vehicular network topologies as discussed in section 2.7.2. All these schemes employ different CH selection schemes with the aim of selecting a CH that serves as long as possible. However, as vehicles move without CH reselection in time relative to one another, the CH position becomes sub-optimal, the efficacy of these clustering algorithms diminishes, and V2V link performance drops. Hence, vehicles might need to be re-clustered to maintain acceptable link

performance. However, most works described in section 2.7.2 assume that persistent re-clustering negatively impacts stability. Hence developing schemes that mitigate the adverse stability effect of persistent re-clustering or link performance issues associated with stability focussed CH selection scheme is worthwhile.

For download of real time HD or machine map updates, high throughput, low jitter, and link stability are key performance indicators. In a clustered hotspot driving scenario where certain vehicles within the cluster serve as a relay to Base Stations (BS) for other vehicles within the same cluster, the links between other cluster vehicles and the relay cluster head vehicle suffers an average throughput drop as the vehicles move relative to one another in distance and time without re-clustering. Though this comes with a gain in stability, it can become a problem in situations of map data streaming or download of real time updates. Hence, we have resolved to develop a scheme that leverages the throughput gain of persistent CH reselection and SNR-based clustering criteria while exploiting a memory-based re-clustering approach for cluster maintenance to improve jitter and stability performance.

In our work, we showed how SNR can be incorporated into a k-means clustering approach to improve V2V user throughput in a download hotspot scenario. We named this scheme KmSNR and compared its performance with the existing KmFW clustering technique presented in [27]. Having observed a superior throughput performance and comparable stability performance, unlike previous work we show how the cluster information stored in the memory of previous clusters in time can be used to reduce jitter and reduce overhead that could be otherwise introduced by typical persistent re-clustering and improve stability. We achieved this by further building upon our SNR-based scheme, with two different memory-based re-clustering approaches (Centroid-based Waterfall Scheme, CWS and Seed-based Waterfall Scheme, SWS) that seek to sustain the throughput performance of KmSNR while improving stability. These approaches make use of current cluster information to determine succeeding cluster heads and centroid positions in subsequent clustering phases. We compared the results obtained here to our corresponding memoryless scheme as used in our previous work in [26] and the Floyd-Warshall scheme used in [27].

3.2 Vehicular Mobility Traffic Generation

The efficacy of vehicular clustering algorithms can only be tested in a vehicular mobility context. For ease of implementation and flexibility we have decided to approach context generation using simulation based vehicular mobility tools. Due to ease of access and interface readiness for MATLAB integration we have adopted SUMO as our mobility traffic generator of choice.

Though defined as “urban”, SUMO is an open source, microscopic, space-continuous, time discrete and portable traffic simulator capable modelling a large array of different mobility scenarios. In SUMO, road networks can be developed from scratch with specific infrastructure and mobility properties, however our focus here is to demonstrate the performance of our schemes in a realistic scenario that imitates real road networks. Hence, we opted for a simple and efficient SUMO solution called *OSM Web Wizard* which leverages Open Street Map (OSM) using python programming to generate real maps and networks from which our desired target area of analysis or Traffic Analysis Zone (TAZ) is selected and upon which mobility scenarios are generated [113].

The OSM is essentially a collection of python programmes in the directory tools of our SUMO installation root. We invoked the OSM Web Wizard python script from the tools’ directory using the command in

```
Python osmWebWizard.py
```

Figure 3-1. OSM web wizard invocation command

Once the python script started running, the OSM web wizard interface opens via our default browser. The interface opens with a window showing a map overview and the position selection panel as depicted in Figure 3-2. We then searched for our road map of interest, A5012, York, and then selected an area of interest within the map, depicted as the bright region of the entire map.

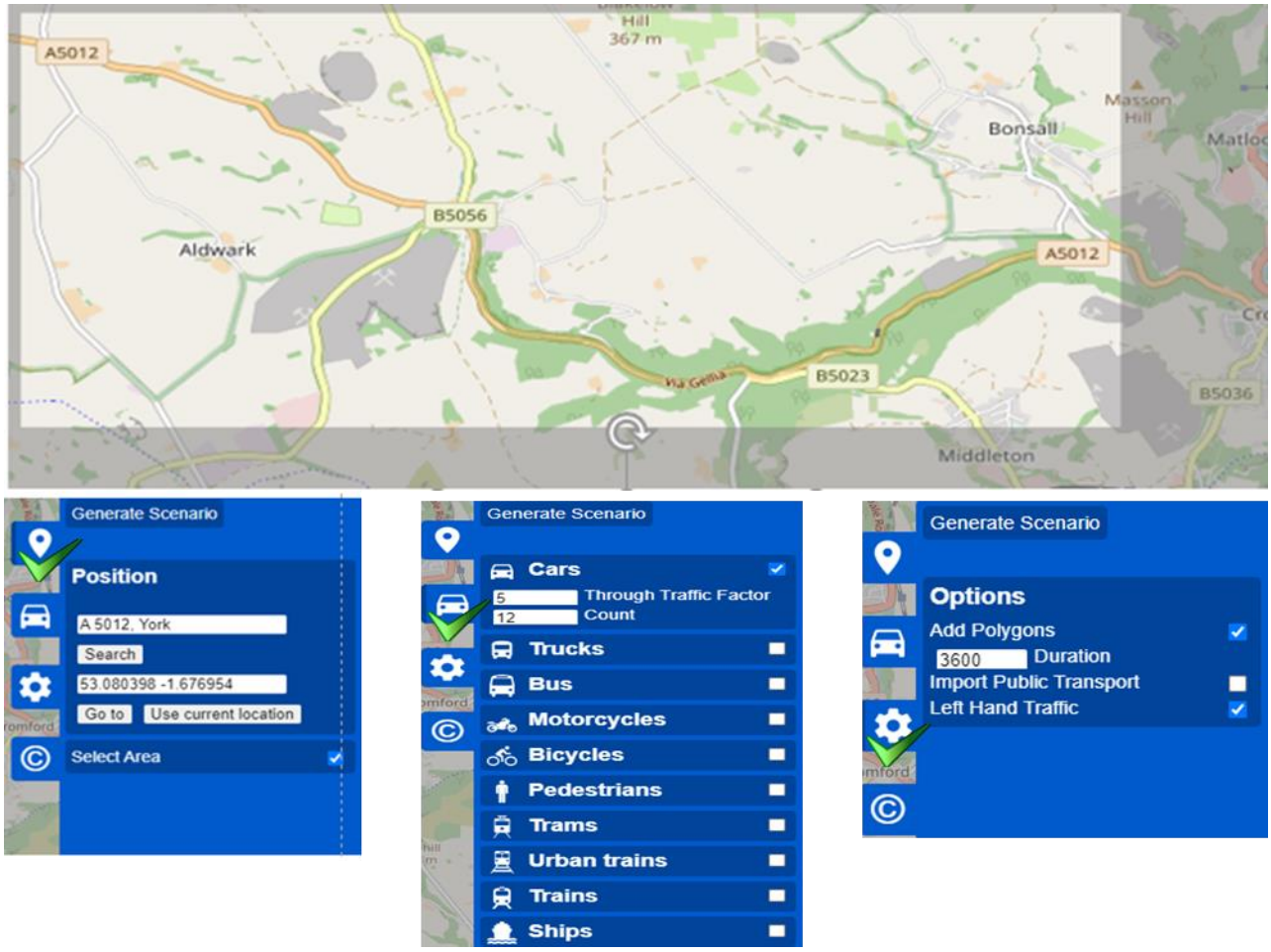

Figure 3-2. SUMO OSM web wizard

After generating the area of interest, we then generated vehicular traffic from the demand generation panel depicted with a car pictogram. As can be seen, the SUMO Web Wizard supports different modes of transportation and mobility nodes which can be activated by ticking the white boxes next to the lists. For simplicity, we have limited our mobility unit type to just cars. For each chosen means of transportation, the OSM Web Wizard generates a random mobility traffic demand based on a probability distribution influenced by two factors:

The through traffic factor, which defines the probability that an edge at the boundary of the simulation will be selected as emergence and departure edge for each vehicle generated, as compared to an edge located within the simulation space. A high value of through traffic factor indicates more vehicles will emerge and depart at the boundary of the simulation space. Specifying a high value is important to retain vehicles within the simulation for as long as possible to study the effect of the mobility of a fixed set of vehicles over time. We selected a very high through traffic factor to maintain a reasonable assurance that at least 100 vehicles

can be studied over time. *The count factor* defines the rate at which vehicles are generated per lane-kilometre. We used a count factor of 65 which means for a two-lane road of 7 km long, 910 vehicles will be generated per hour on the highway. This results in approximately in one car generated every 4 seconds [113].

In the options panel, we chose the simulation to run for an hour after which we generated the scenario which took a few minutes to run. The output of the generated scenario is in the form of a sumo configuration file, *sumo.cfg*. The *sumo.cfg* file opens by default on generation and presents the selected area with a ready-to-run traffic as depicted in Figure 3-3.

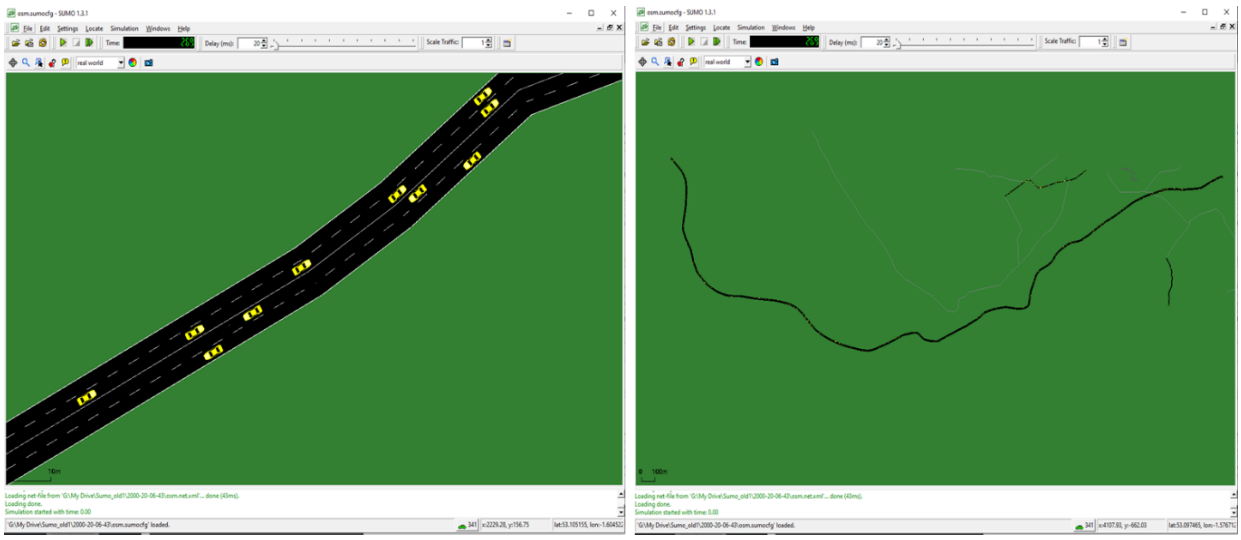

Figure 3-3. A picture of open SUMO configuration file

The diagram on the right of Figure 3-3 clearly shows the map of the A-5012 road, while the diagram on the left is a zoomed-in map of the map after being run to reflect the vehicular mobility.

Since we have decided to integrate SUMO into MATLAB, we had to extract the vehicular traces from *sumo.cfg*. The output of the data extraction is called a Floating Car Data (FCD). The output contains the coordinates of each vehicle at every point-in-time, the speed at every point-in-time, vehicle ID and time stamp. By default, the time stamp is in x-y coordinates, but the extraction process can be adapted to generate a longitude/latitude coordinate. We extracted both the rectangular and geographical coordinates of the vehicular traces for our work and saved the file in a *.csv* format, since this format is MATLAB friendly.

We integrated the trace file into MATLAB and superimposed it on google map of the same road from which it was extracted. We faced some compatibility issues when we tried to reuse Open Street Map (OSM) within MATLAB. A depiction of the superimposed vehicle traces is shown in Figure 3-4.

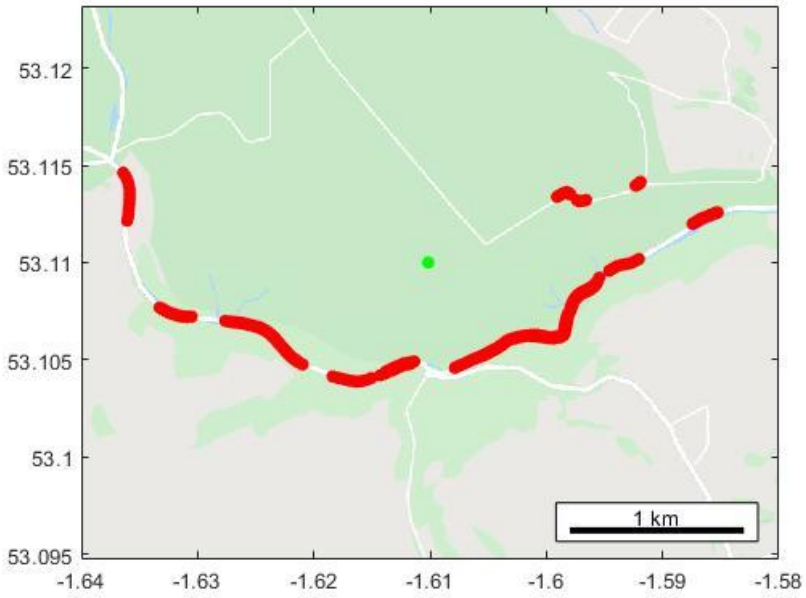

Figure 3-4. MATLAB generated superimposition of vehicle traces on google map

3.3 System Model

The communication model is built as a single cell C-V2X network model. The base station is responsible for initiating clustering of vehicles and resource allocation. The single cell serves multiple vehicle clusters and un-clustered free vehicles directly via the Uu interface. Vehicles can also share information between one another via the PC5 interface. We have modelled only the V2V communication mode into our communication model since this is where the effect of our contribution is realised.

The system assumes every vehicle has both the Uu and PC5 interface, hence capable of both V2V and V2I communication and are each electable as CHs. CMs of same cluster are modelled

to use orthogonal channels and centrally allocated resources, hence effectively avoiding the hidden node problem.

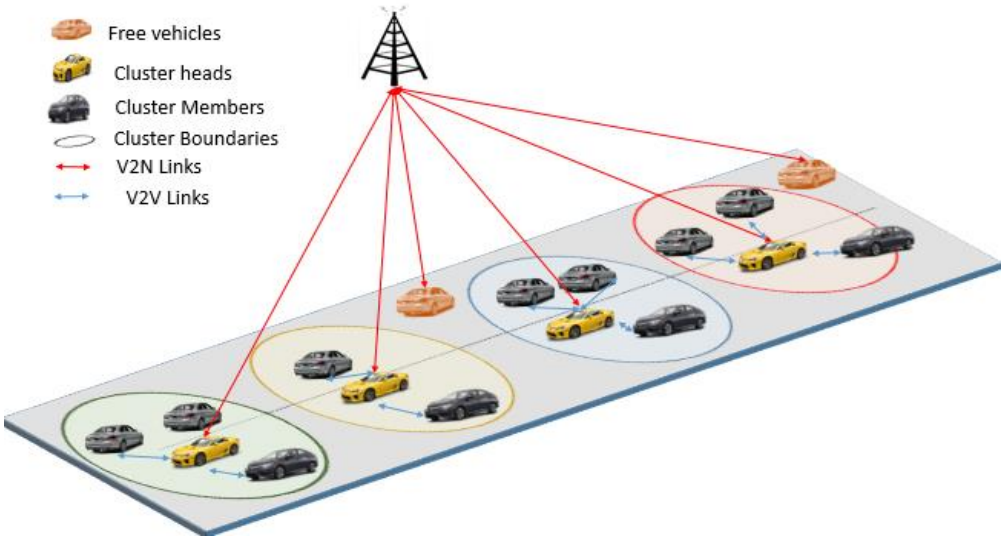

Figure 3-5. System model cluster based c-v2x

We adopted a highway V2V path loss model from [148] for both LOS (Line of Sight) and NLOSv (Non-Line of Sight for vehicles on same street).

$$PL_{los} = 32.4 + 20\log_{10} d_{euc} + 20\log_{10}(f) + X_{\sigma} \tag{3-1}$$

$$PL_{nlosv} = 32.4 + 20\log_{10} d_{euc} + 20\log_{10}(f) + A_{sk} + X_{\sigma} \tag{3-2}$$

$$\begin{gathered} \text{When } v = -0.7 \\ A_{sk} = 6.9 + 20\log_{10}\left[\sqrt{(v-0.1)^2+1} + v - 0.1\right] \\ \text{Else}, A_{sk} = 0 \end{gathered} \tag{3-3}$$

$$v = \sqrt{2}\frac{H_d}{r_f} \tag{3-4}$$

$$r_f = \sqrt{\frac{n_f \lambda d_1 d_2}{d_{euc}}} \tag{3-5}$$

Where X_σ Is the shadowing parameter, A_{sk} is additional blocking attenuation due to vehicular obstruction, H_d is the difference between obstruction height and horizontal link plane, n_f is the Fresnel number, d_{euc} is the Euclidean distance between VUEs or length of V2V links, d_1 is the distance of a specific point to the transmitting antenna, d_2 is the distance of the same specific point to the receiving antenna, r_f is Fresnel ellipsoidal radius, and λ is wavelength.

The probability of LOS and NLOSv is estimated based on a distance threshold, d_{th} of 450 meters. The probability of LOS or NLOS for each V2V link is estimated based on whether the difference, d_δ between the link distance, d_{euc} and distance threshold, d_{th} is positive or negative [148], as described in equations (3-6), (3-7) and (3-8). A link is assumed to be LOS if the probability of LOS, $P(LOS)$ exceeds the probably of NLOS, $P(NLOS)$ and vice versa.

$$d_\delta = d_{euc} - d_{th} \tag{3-6}$$

$$P(LOS) = \begin{cases} \min\left(1, a + {d_{euc}}^2 + b * d_{euc} + c\right) & d_\delta \leq 0 \\ \max(0, 0.54 - 0.001 * d) & d_\delta > 0 \end{cases} \tag{3-7}$$

$$P(NLOS) = 1 - P(LOS) \tag{3-8}$$

$$\text{Where } a = 2.1013 \times 10^{-6}, b = -0.002 \; and \; c = 1.0193$$

3.4 Resource Allocation Model

We have adopted a new cluster-based resource allocation scheme based on 3GPP's LTE-V2X mode-3 resource allocation standard, though the 3GPP has not defined a specific scheduling scheme for LTE-V2X mode-3, they have outlined centralized scheduling procedure requirements. The document states that the VUE needs to be in RRC_CONNECTED mode to transmit in mode 3. In this mode, all communication parameters are known to both the VUE and the base station. The VUE interested in a V2V communication, informs the eNodeB about the impending link to be established by transmitting a side-link UE information message. This message is also used to request resources for side-link communication. Mode-3 supports both dynamic and SPS allocation. SPS however is best suited for the scenario of context, and it is based on the SPS allocation that we have designed our resource allocation algorithm.

We have selected mode-3 for its centralization benefits such as improved resource coordination, avoidance of the hidden node problem and simpler cluster formation process. We have also adopted the overlay as opposed to the underlay approach of resource allocation to eliminate the interference from non-vehicular users served by the base station [160]. We believe the cluster-based approach we have adopted will help efficiently manage the limited access to resource associated with the underlay approach.

The rest of this section describes our variant of the C-V2X mode-3/mode-1 SPS scheme. The resource allocation algorithm dedicates a resource set for V2V side-link communication, because of which interference from cellular base stations and users are non-existent. Resources for V2V communication are allocated to clusters by eNodeBs. However, eNodeBs are not involved in the actual V2V communication. Cluster heads are responsible for the allocation of resources within the cluster based on the Resource Block Group (RBG) interference information received from the eNodeB. Information about the cluster is periodically transmitted to the eNodeB and request for cluster resources is sent via UL-DCCH (uplink Dedicated Control Channel). Vehicle location of CMs can also be reported via the CH to the eNodeB. The eNodeB uses the vehicle information provided by the CH to the eNodeB allocate resources to each cluster. Cluster heads are responsible for allocation of resources within the cluster based on the Resource Block Group (RBG) interference information received from the eNodeB

The physical layer of LTE-V2X is defined to support 10 or 20MHz of channel bandwidth with 50 or 100 Resource Blocks (RB) respectively. A frame is 10ms long and divided into 10 subframes of 1ms each. In our SPS allocation scheme we assumed that resources are retained by each VUE over 100 frames or 1 second. Each RB consists of 12 subcarriers (15 kHz each) and it is 180 kHz wide in frequency. In our scheme, the number of RBs per VUE, N_{rbpu} is dependent on the maximum cluster size under the cell covered by the eNodeB at the scheduling instance. All users are also allocated the same size of resource at every scheduling interval. As shown in equation (3-9) and (3-10), N_{rbpu} is estimated from the total available resource blocks, N_{rb} and maximum cluster size, C_{sz}.

$$N_{rbg} = \lfloor max\{C_{sz}\} \times 1.5 \rfloor \quad (3\text{-}9)$$

$$N_{rbpu} = \left\lceil \frac{N_{rb}}{N_{rbg}} \right\rceil \quad (3\text{-}10)$$

Where N_{rbg} is number of resource block groups (sub-channels) and N_{rbpu} is the size of a resource block group or number, resource blocks in a resource block group or number of resource blocks per user, given that one resource block group is allocated to a user.

The scenario envisages a context where resources are allocated to all CH to CM links for transmission of DENMs such as emergency traffic information, location-based information or imminently required HD maps. First, CMs under a specific cluster request resources via their CHs. Based on periodic status information already reported to the eNodeB, the eNodeB allocates resources to clusters by first checking for free sub-channels able to meet the cluster request. If available, the free sub-channels are randomly allocated to the cluster and eventually to the CMs. If there no free sub-channels available, the algorithm checks for occupied resources and aggregates a set of sub-channels whose current users poses least interference to the CH and allocates them to the cluster. Since the CH is aware of the interference properties of the sub-channels and location of the CM, it can estimate the interference each of the vehicles occupying the allocated resources have on individual CMs and allocate to each CM the resource that poses minimum interference. So, it appears to be a 2-tier resource allocation. Users within same cluster are not allocated same sub-channel. Figure 5 shows a graphical depiction of the allocation. The number of resources allocated to a cluster, N_{rbpg} is given as

$$N_{rbpc} = C_{sz_k} \times N_{rbpu} \quad (3\text{-}11)$$

For a set of clusters, $C = \{C_1 \dots C_k\}$ of size, $C_{sz} = \{C_{sz_1} \dots C_{sz_k}\}$, and CH, $C_h = \{C_{h1} \dots C_{hk}\}$, occupying sets of sub-channels $S_c = \{S_{c1} \dots S_{ck}\}$, if a specific cluster, $C_i \in C$ requires C_{sz_i} sub-channels, assuming all resources have been allocated, the algorithm estimates the sum of power received by C_{hi} from all vehicle exterior to $C_i$that are occupying each sub-channels. Then selects the C_{sz_i}sub-channels that constitutes minimum interference to it for allocation

to corresponding cluster members. If the set of exterior vehicular users occupying a specific sub-channel, p is given as $X_c = \{x_{c1} \dots x_{cg}\}$, and the power received by C_{hi} from each user, $x_{ck}$is $P_r^{x_{ck}}$, the total received power by C_{hi} from exterior vehicles X_c occupying the specific sub-channel is estimated to be

$$P_{C_{hi}}^{X_{c\_p}} = \sum_{k=1}^{g} P_r^{x_{ck}} \tag{3-12}$$

For a specific cluster C_i with cluster head C_{hi} and cluster size C_{szi}. The set of sub-channels allocated to the cluster is given as:

$$\min_{C_{sz_i}} \left\{ P_{C_{hi}}^{X_{c\_p}} \middle| 1 < p \leq N_{rbg} \right\} \rightarrow S_{ci} \tag{3-13}$$

Figure 3-6 shows a simple depiction of the scheduling scheme.

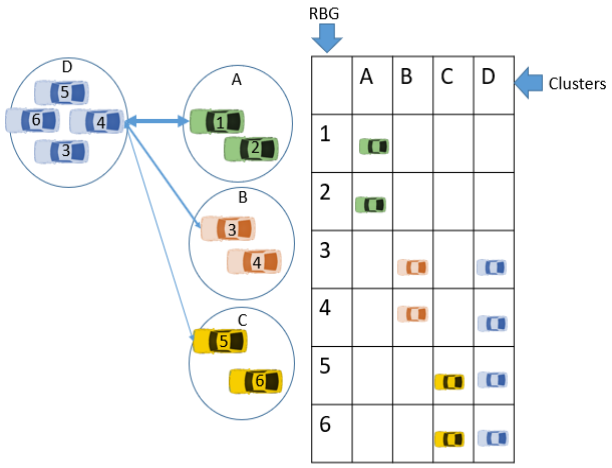

Figure 3-6. Depiction of resource allocation scheme

For a cluster $C_i \in C$ with cluster members, $C_i = \{c_{1\dots\dots\dots c_{C_{sz_i}}}\}$ allocated a set of sub-channels, S_{ci} . C_{hi} allocates these channels to individual CM by estimating SINR, γ to each CM for each sub-channel in S_{ci} , then selects for each CM the sub-channel that presents the maximum SINR. If a CM, $c_f \in C_i$ is to be allocated one of S_{ci} sub-channels, the CH, C_{hi} assigns a sub-channel, $S_{ci\_f} \in S_{ci}$ for which CM experiences maximum SINR, γ_m, given as:

$$max\{\gamma_m | 1 < m \leq C_{sz_i}\} \rightarrow S_{ci\_f} \tag{3-14}$$

3.5 Proposed Clustering Schemes

Efficient and reliable transmission of messages within vehicular networks is crucial to the success of Intelligent Transportation System (ITS) and more specifically to vehicular networks. Management of the fast-changing topology and moving nodes of the network environment is crucial to reliable transmission of data within the network. Logical grouping of vehicles into manageable clusters has been discussed in many literatures as an efficient way to introduce stability [17], [21], [22] and improve bandwidth efficiency [24], [25] in vehicular networks. The grouping is often based on proximity or estimated proximity along travel time. Vehicles in a cluster are identified to be in one of three states, Free Vehicles (FV), Cluster Members (CM) or Cluster Heads (CH). The CH coordinates and facilitates the communication needs of CMs within its cluster. This could be routing, resource allocation or authentication. The FV is a vehicle that is yet to meet the criteria of any cluster, hence, remains outlying. CMs are vehicles who have met a clustering criterion and are now associated with a CH.

3.5.1 K-means Based SNR Clustering (KmSNR)

This work chooses to exploit the simplicity and centroid convergence offered by k-means, while adopting Signal-to-Noise-Ratio (SNR) as our vehicle clustering metric. We envisage that SNR will present a more accurate network performance-oriented grouping as compared to distance metric adopted in traditional clustering schemes. We then compare our result with K-means and Floyd-Warshall Technique (KmFW) presented in [27]. KmFW simply use the traditional k-means to cluster vehicles and then elects CHs by estimating the minimum average shortest distance between vehicles using Floyd-Warshall algorithm.

Unlike KmFW, KmSNR uses a combination of k-means++, k-means and silhouette criterion to generate optimal centroid position. It then elect CH vehicles based on closest proximity to centroid position. This is done to ensure even distribution of centroids and consequently CHs. We then group vehicles into clusters based on their relative SNR levels to CHs. We have described our proposed clustering algorithm in three segments based on functions: evaluating the number of clusters and initial seeds, identifying stable centroids, and identifying CHs and CMs.

3.5.1.1 Kmeans++

The first step in our clustering process is the selection of centroids using kmeans++. Centroids are selected based on the number of required clusters. Here the number of clusters is not predetermined. So kmeans++ will be run for the maximum range of clusters possible, k. The process starts by using k-means++ to select initial centroids called seeds for each k clusters under test. For a given number of clusters, we select a vehicle, $v_1 \in V$ at random. The chosen vehicle, v_1 is identified as the first seed, s_1. The Euclidean distance between all vehicle points and $s_1$is estimated. If the distance between a certain vehicle $v_n \in V$ and the identified seed, s_i is denoted as $d(v_n, s_i)$. The next centroid, $s_2 \in S$ is selected with the probability, $P_{r(d)}$ given in equation (3-15).

$$P_{r(d)} = \frac{d^2(v_n, s_i)}{\sum_{u=1}^{z} d^2(v_n, s_i)} \tag{3-15}$$

Vehicle points with the highest probability,P_r will be selected the next centroid, s_2. To select the next centroid, the distance between all vehicle points and existing centroids are computed. Vehicles are associated with the existing centroid to which they are closest. The distance-based probability, $P_{r(d)}$ is then computed for each assigned vehicle points and the vehicle with the highest probability is selected the next centroid. This process is repeated until the desired number of seeds, s_k are reached.

3.5.1.2 K-means

K-means clustering is then used to obtain a more stable centroid by an iterative data partitioning process. Vehicles are clustered around the k-means++ selected seeds, s_k by computing seed to vehicular point Euclidean distances $d(v_x, s_k)$. Batch update is then used to assign users to closest centroid. Newer and a more central set of centroids are obtained by averaging the positions of vehicular points in each cluster. The iteration is repeated until convergence is reached. The centroid arrived at becomes the theoretical centroid, c_k for each of the number of centroids under test. Vehicles are now attached to centroids to which they are closest.

3.5.1.3 Silhouette Criterion

Having repeated the k-means++ and the k-means procedure on the different numbers of clusters under test, *k*, the silhouette criterion is then used to select the optimum number of clusters to be used, K_{opt}. The silhouette value, $s(u)$ is a measure of proximity-based distinctiveness of resulting clusters. It assigns a value $\{s(u)|-1 \leq s(u) \leq 1\}$ to each CM point indicating the level of compliance to cluster identity relative to neighbouring clusters. CM with value of 1 indicates full compliance to cluster identity, and CMs with value of -1 indicates misplaced cluster assignment. A value of 0 indicates that a CM is at a decision boundary. For a set of possible Number of clusters, Nc, the silhouette criterion is tested across each to determine the optimum number of clusters, K_{opt}. For each number of clusters $k \in Nc$, a silhouette value is estimated, $s(k)$. The k with the highest silhouette criterion value is chosen as the optimum, K_{opt}. Since the silhouette criteria characteristically selects the maximum number of clusters obtainable as optimum, we set a maximum threshold for k ($0.5\ N_{veh}$) for which the silhouette criterion test can be executed. This assumes that on average we will have at least two vehicles per cluster.

$$Nc = \left\{k \middle| k \in \mathbb{N}, 1 < k \leq \frac{Nveh}{2}\right\} \tag{3-16}$$

For a vehicle point $u \in C_u$, let the mean distance between u and all other vehicles within the cluster will be:

$$a(u) = \frac{1}{|C_u| - 1} \sum_{\bar{u} \in C_u} d(u, \bar{u}) \tag{3-17}$$

Where $|C_u|$ is the number of vehicular points in cluster, C_u. To estimate the inter-cluster proximity, we then evaluate the mean of the distance between vehicle point u and all other vehicles in a different cluster C_w as the distance between the point u and the cluster, C_w. For $u \in C_u$, we define the minimum distance between vehicle point u and all other clusters of which vehicle point u is not a part of as:

$$b(u) = \min_{u \neq w} \frac{1}{|C_w|} \sum_{v \in C_w} d(u, v) \tag{3-18}$$

The cluster holding the value $b(u)$ is described as a neighbour cluster of C_u or the next best fit cluster for vehicle point u. We then define the silhouette index of a single vehicle point as:

$$s(u) = \frac{b(u) - a(u)}{\max\{a(u), b(u)\}}, if\ |C_u| > 1 \tag{3-19}$$

The mean $s(u)$ over all points of a cluster is a measure of how closely grouped the CMs of the cluster are relative to CMs of other clusters. The mean $s(u)$ over the entire data set of vehicle points indicates how perfectly clustered the vehicle points are, over a given number of clusters, k. The entire procedure is repeated over a set of number of clusters, Nc.

$$s(k) = \frac{1}{k} \sum_{n=1}^{k} \frac{1}{|C_u| - 1} \sum_{u=1}^{|C_u|} \frac{b(u) - a(u)}{\max\{a(u), b(u)\}} \tag{3-20}$$

If a set of numbers of clusters, Nc, and a set of average silhouette values for all data sets of a given number of cluster, S is given as $Nc = \{k_1 \dots \dots \dots \dots k_n\}$ and $S = \{s_1 \dots\dots\dots\dots s_n\}$, then the optimum number of clusters, K_{opt} will be

$$K_{opt} = \{k \in Nc | \max(s \in S)\} \tag{13}$$

Essentially, K_{opt} defines the number of clusters for which clusters are most distinctively partitioned. A silhouette plot depicting the silhouette indices of all vehicle points for 8 clusters in our test dataset is presented in Figure 3-7 . All vehicular nodes have their silhouette index above 0, indicating that none of the vehicle points are inappropriately placed in a cluster. None of the vehicle points also have an index of 0, indicating no ambiguity in placement. However, we can see that a few clusters have only one member vehicle node, this indicates

which are expectedly perfectly placed. However, this is not ideal and is due to the sparse population of vehicles.

3.5.1.4 SNR Clustering

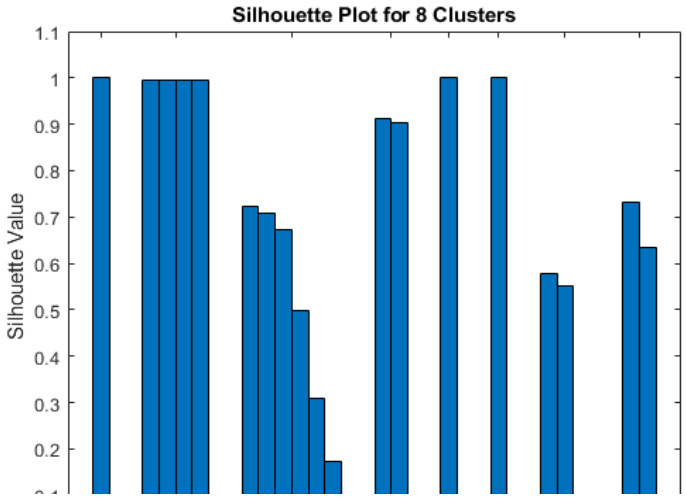

Figure 3-7. Silhouette plot from our pilot dataset depicting the silhouette criteria indices

The k clusters, now have corresponding k centroids defined by set, $C = \{c_1 \dots \dots \dots . c_k\}$. For clarity, we name these centroids the theoretical centroid. We then estimate the Euclidean distance of each vehicle to the theoretical centroids, denoted as $d^2(v_n, c_k)$. The vehicles with the minimum distance to each of the centroids are termed vehicular centroids and are defined by set, $C^v = \{{c_1}^v \dots \dots \dots {c_k}^v\}$. For a theoretical centroid,$c_k \in C^v$ if the set of Euclidean distances of each vehicle in a set, V to c_k is given as $d^2(V, c_k)$, then the corresponding vehicular centroid, ${c_k}^v$ will be that vehicle $v_n \in V$ whose distance to c_k, ${d_k}^{v_n}$ is minimum as expressed in equation (3-21).

$$ {d_k}^{v_n} = min\{d^2(V, c_k)\} \tag{3-21} $$

These vehicular centroids are the chosen CHs around which the logical partitioning of vehicles into clusters are formed. Vehicles are associated with CHs to which they have highest mutual SNR. If the set of Signal-to-Noise Ratio between a vehicle, $v_n \in V$ and each vehicular centroid in C^v is given as $SNR_n = \{{snr_1}^n \dots \dots \dots {snr_k}^n\}$, then the vehicular centroid, ${c_j}^v \in C^v$ to which v_n will associate will be that whose SNR is minimum, ${snr_j}^n \in SNR_n$ as expressed in equation (3-22).

$$snr_j^{\,n} = max\{SNR_n\} \tag{3-22}$$

$$SNR_n = \{[P_R(v_n, C^v)]./[P_N(v_n, C^v)]\} \tag{3-23}$$

$[P_R(v_n, C^v)]$ Is a set of measured received signal strength values by v_n from each vehicular centroid in C^v, while $[P_N(v_n, C^v)]$ is the received noise accompanying the signal. A holistic flowchart of the four steps in the clustering process is presented in Figure 3-8.

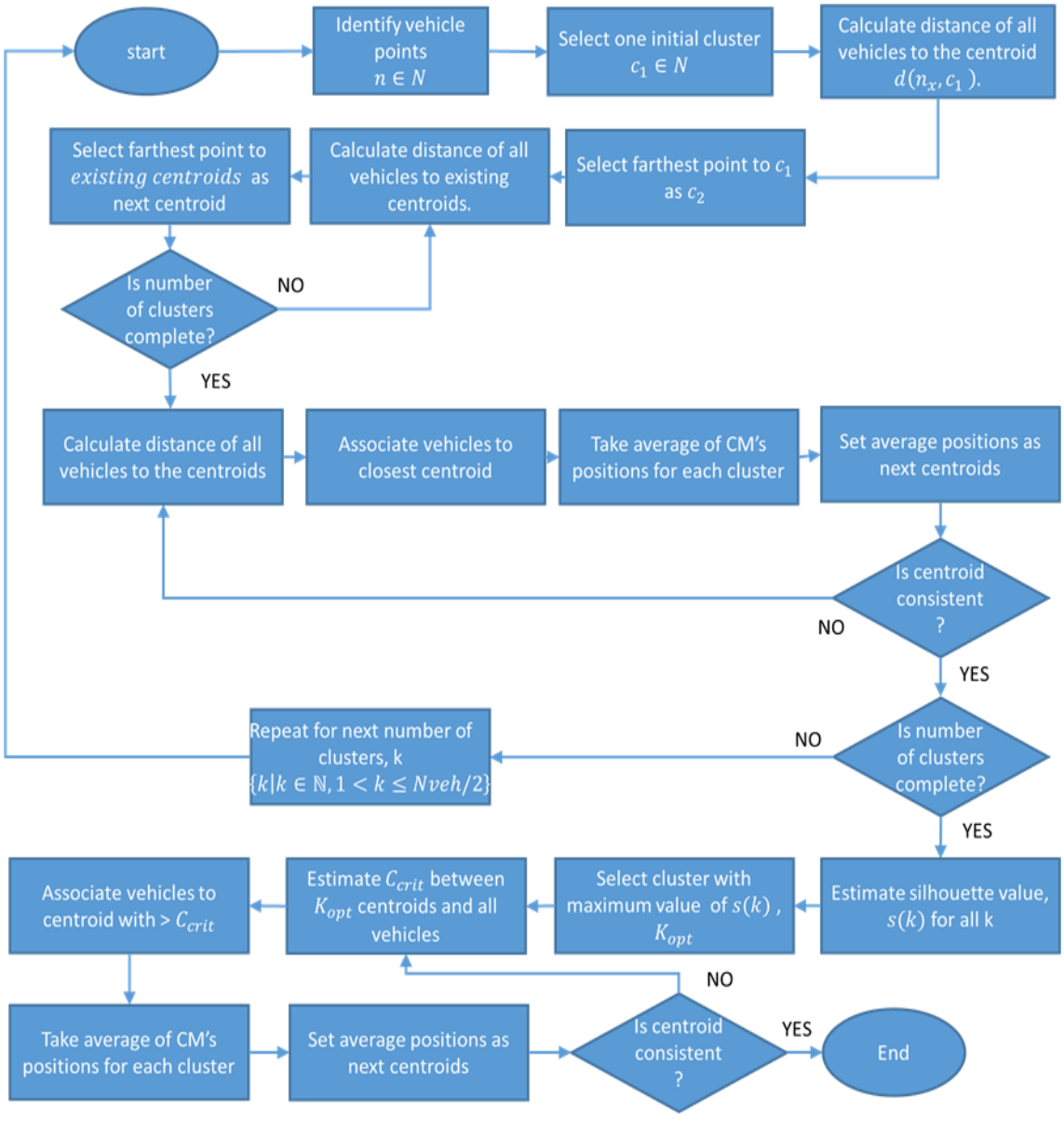

Figure 3-8. A flow chart of the clustering process

3.5.2 Cluster Formation

The formation of clusters proposed for our C-V2X based clustering scheme is initiated by the base station but requires a collaboration of the base station and the vehicles to be completed. Vehicles implanted with C-V2X chips are equipped with the Uu and the PC5 interface. All vehicles are assumed to be capable of transmitting status information to the eNodeB via the Uu interface and to vehicles within range via the PC5 interface. At inception of the clustering process, the eNodeB/gNB is assumed to have knowledge of vehicle positions acquired via the use of hybrid 3GPP and non-3GPP technologies such as Observed Time Difference of Arrival (OTDoA), Global Navigation Satellite System (GNSS) and positioning sensors. The CH status are evaluated and allocated to the appointed vehicles by the base station using the elbow/k-means++/k-means methods. The CH status is communicated to the appointed vehicles via the Uu interface. The CHs then transmits their CH status information, clustering criteria (SNR) and method to neighbouring vehicles via the PC5 interface. Vehicles use the clustering method information to decide based on best SNR the vehicular centroid to which the vehicles are to associate. A flowchart of the process is presented in Figure 3-9.

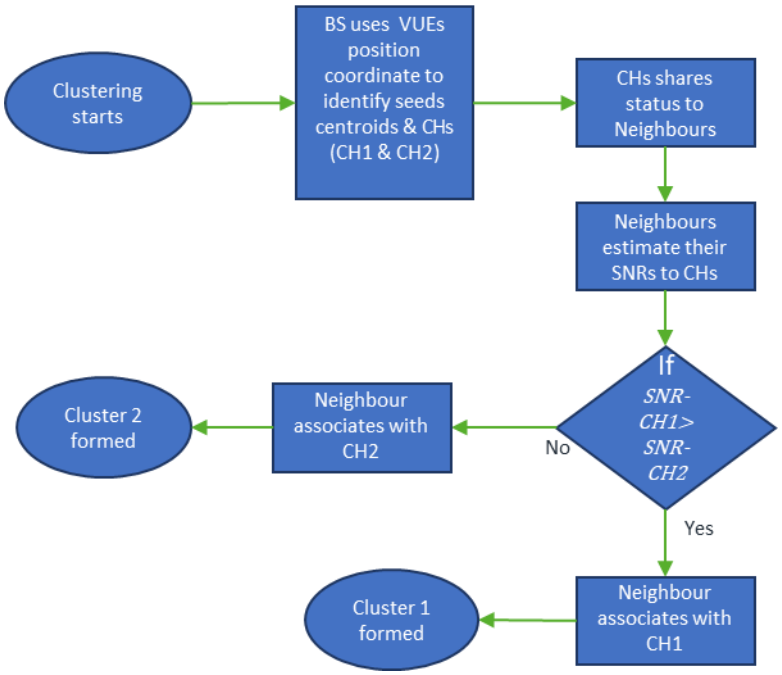

Figure 3-9. Cluster formation process

When a new vehicle is activated in an existing clustering scenario, it shares its status information with other vehicles. It then receives status information from other vehicles and

uses the status information received from CHs to make decision on the CH to which it will associate. If no cluster head is found, it remains a FV and keeps direct Uu connection with the base station for same services for which clustering was intended. In the next clustering instance, the newly activated vehicle along with other vehicles will be considered for CH status appointment.

3.6 Re-clustering Schemes

With some compromise on stability, we have demonstrated a superior throughput performance of our k-means based SNR clustering scheme (KmSNR) over a distance-based clustering scheme (KmFW) as shown in Figure 3-11 and Figure 3-12. Having done this, we introduced two memory-based re-clustering schemes built upon our k-means based SNR clustering scheme. This is intended to sustain the superior throughput performance of CH-to-CM links while improving on stability. The two schemes are waterfall models based on two separate centroid selection techniques, initiated at every re-clustering phase. The idea behind the memory-based scheme is to limit the inherent instability introduced by random initial seed selection of traditional k-means schemes in subsequent clustering phases by by-passing the k-means++ component of the clustering scheme. We have distinctly named the two schemes the Seed-based Waterfall Scheme (SWS) and Centroid-based Waterfall Scheme (CWS).

In SWS, the initial clustering is done in three steps. First the elbow algorithm is used to estimate the number of clusters, k. Then initial seeds are generated using k-means++ algorithm, after which k-means algorithm is used to generate convergent centroids. Finally, CHs are generated, and clusters are formed around them using the KmSNR we have developed. For subsequent clustering phases, the k-means++ component that generates seeds for the k-means algorithm is replaced by a waterfall system that feeds the coordinates of current CHs as seeds for the k-means algorithm which generates convergent centroids for the KmSNR that will generate the new cluster heads and corresponding clusters. As for the implementation, the eNodeB is assumed to have knowledge of the position coordinates of vehicles under its coverage cell. When subsequent clustering is required, the eNodeB uses the position coordinate of the current CHs as the seeds for the clustering process to be initiated. These seeds serve as input to the k–means algorithm, where the new centroids are

generated. After which the new CHs are identified. All of these takes place in eNodeB. The base station informs the new CHs about their new status and the new CHs in-turn shares their new status information with their neighbours. Neighbours then associate with CHs based on a clustering criterion defined in the k-means variant algorithm. The process is repeated in subsequent clustering phases.

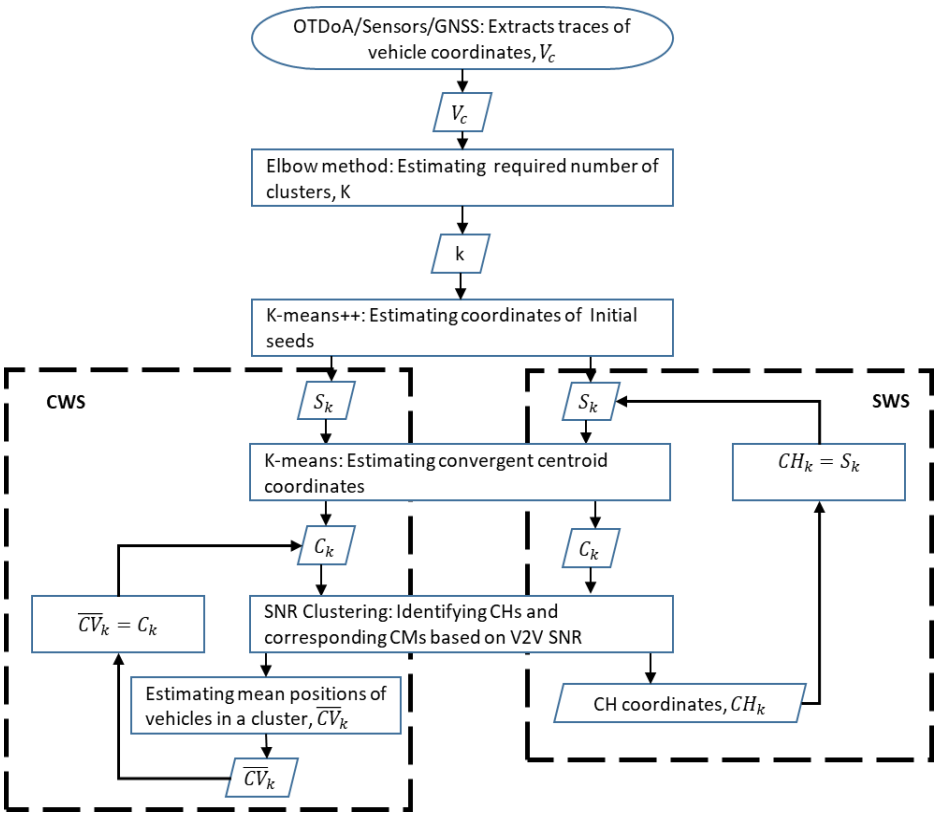

Figure 3-10. A combined high level block diagram of CWS and SWS

For CWS, the initial clustering follows the same process as that of the SWS. However subsequent clustering does not require both the k-means++ and k-means algorithm. CWS does this by a different feedback system that takes the mean position coordinate of vehicles in a cluster and feeds this mean position to the k-means variant algorithm which evaluates the new vehicular CH around which new clusters are formed. The implementation of CWS also involves both the eNodeB and the VUEs. The eNodeB is assumed to be aware of the clusters or will be informed about cluster information by CHs. The eNodeB estimates the mean position coordinate of cluster vehicles for each cluster and uses this information as centroids and input to the k-means variant. The k-means variant estimates and identifies the current CHs. The eNodeB then informs the new CHs about their new statuses. The new CHs

then inform their neighbouring VUEs about their new status. The neighbour VUEs then associate with CHs that meets the clustering criteria specified in the k-means variant algorithm. The process is repeated in subsequent clustering processes. A combined high level block diagram of the re-clustering process of both CWS and SWS is presented in Figure 3-10.

Two baseline approaches are developed for comparison. The first is a total non-re-clustering approach and the other is a complete re-clustering approach from the k-means++ seeding component to the k-means component and then to the k-means variant component.

3.7 Performance Evaluation

This section explains the context and performance parameters by which we have compared the different schemes we have discussed.

The clustering, communication and resource allocation models are built around a scenario of a single carriage highway (A5012, UK) with increased traffic build-up. The vehicle traces are generated from the SUMO (Simulation of Urban Mobility) tool. The vehicle traces depict vehicles moving at different speeds and in both directions along the carriageway. The scenario paints a situation where a large HD/machine map is to be streamed in real time to specific vehicles. The information is transmitted from the base station through CHs (Cluster Heads) to the CMs (Cluster Members).

The performance evaluation focusses on V2V links between individual CHs and corresponding CMs. Each cluster has only one Uu link from the CH to the base station, and side-links corresponding to the number of CMs in a cluster as depicted in Figure 3-5. Hence, the number of vehicles within our simulation, N_{veh} will be a summation of all V2V links in each cluster and the size of CHs. Since we have a single CH per cluster, if the number of clusters is defined as K_{opt} and the set of the size of CMs in each cluster is given as $Sz = \{s_{z1} \dots \dots \dots . s_{zk}\}$ we can estimate the total number of vehicles as:

$$N_{veh} = K_{opt} + \sum_{y=1}^{y=K_{opt}} S_{zy} \tag{3-24}$$

The simulation assumptions are set according to 3GPP TR 37.985 [163]. Table 3-1 shows the simulation parameters used. We considered 100 vehicles moving on a carriage highway over a period 2 minutes and 40 seconds.

Table 3-1. Simulation parameters

Parameters	**Values**
Bandwidth	10MHz
Carrier frequency	5.9GHz
Maximum transmission range	1000m
SINR threshold for successful reception	5dB
Distance threshold for NLOS	475m
Number of V2V vehicles	20
Maximum Transmit power	23dBm
Noise power	-113dBm
Shadowing distribution	Log-normal $\sigma = 3dB$
Resource Block size	180kHz
Range of vehicle speed	[40,120]km/h

The focus here is on the throughput and link reliability performance of the centroid selection models. We evaluated these performance indices by estimating the system content delivery capacity, cluster stability and link stability over all the links between each CH and corresponding CMs.

Our V2V link performance was estimated in terms of content delivery capacity, Cumulative Distribution Function (CDF) of link throughput in time, sum-rate, and jitter.

First the SINR and throughput is estimated as presented in equations **Error! Reference source not found.** and **Error! Reference source not found.**. The throughput performance of V2V links is critical to the efficient transmission of real-time traffic updates in a hotspot network topology. The V2V link throughput CDF performance for the different schemes demonstrates

variation of link throughput values over the performance range and identifies possible outliers that might affect the perception of the throughput performance of schemes. Sum-rate and CDF gives end-end insight into network performance from the overall network capacity to the more granular contribution of individual link throughput to overall performance.

$$\gamma_x = \frac{Pr_{wx}}{N + \sum_{j=1}^{j=n} Pr_{wj}} \tag{3-25}$$

$$\rho_{l_x} = W \log_2(1 + \gamma_x) \tag{3-26}$$

Where Pr_{wx} is downlink received signal strength from CHs to CMs, N is noise and Pr_{wj} is the received interference from transmitting CHs sharing same resource block group, W as link, x.

The content delivery capacity defines the number of vehicles with the link capacity to download a fixed size of urgent traffic data such as traffic or software update within a required time interval. This could be data related to safety or traffic management. This gives a different insight to the perception of how throughput performance affects the safety and efficiency of vehicular networks. Traffic content delivery capacity is estimated by evaluating for each centroid re-selection model the number of vehicles that could successfully complete the download of a fixed size of data, d_{sz} within specific time threshold, t_{th} as indicated in equation **Error! Reference source not found.**.

$$f(t_{th}) = \begin{cases} unsuccesful, & t_{th} - \dfrac{d_{sz}}{\rho_{l_x}} < 0 \\ successful, & t_{th} - \dfrac{d_{sz}}{\rho_{l_x}} \geq 0 \end{cases} \tag{3-27}$$

Sum-rate, ρ_s was estimated to account for the number of vehicles accommodated in each clustering scheme to have a holistic capacity estimation. This is done by summing the throughput per user over entire travel time for each clustering approach.

$$\rho_s = \sum_{t_1}^{t_s} \sum_{n_1}^{n_k} \sum_{m_1}^{m_p} \rho_{l_x} \tag{3-28}$$

Where p, is the total number of CMs, m is the CM index, k is the number of clusters, n is the cluster index and t represents the travel duration, with t_s being the total travel time.

In contexts such as cooperative download of real-time 3D maps or use cases such as cooperative merging requiring real time communication, low jitter offers vehicular nodes synchronous reception of critical traffic data and coordinated responses. Conversely, a poor jitter performance might also generate inconsistencies in received perception information, subsequently prompting uncoordinated responses. Given the total travel time duration of vehicles from the start time to a specific point-in-time, t_k to be$t_k - t_1$, jitter performance (ψ) of the re-clustering decision schemes is estimated by evaluating the variation in the time taken to download a fixed packet size of data, P_{sz} of successive points-in-time. The mean jitter, $\overline{\psi}(secs)$ across the total travel time per user has been estimated to be:

$$\overline{\psi}(secs) = \frac{1}{t_s - t_1}\left(\sum_{t_2}^{t_s} \left| \frac{P_{sz}}{c_{ts}} - \frac{P_{sz}}{c_{ts-1}} \right| \right) \tag{3-29}$$

Reliability of vehicular network links is a critical performance parameter in evaluating vehicular clustering schemes primarily because of the intrinsic dynamic property of vehicular network topology. In vehicular networks reliability is often defined by the stability of vehicular links. Reliability of the link is evaluated here using 3 different stability-metrics: distinct CH stability, CH set stability and CH-to-CM side-link stability. Given a vehicular network instance with N vehicular nodes and k clusters, where the set $C = \left\{ C_{i_1^k} : C_{1,} C_{2,} \dots \dots \dots \dots C_{k,} \right\}$ is a set of clusters with $C_i \in C$ representing individual cluster in time and $c_h = \left\{ c_{hi_1^k} : c_{h1,} c_{h2,} \dots \dots \dots \dots c_{hk} \right\}$ is set of CHs, the number of CHs and number of clusters, *k* are equal. Hence, the maximum number of possible CH changes at any time instant, *t* is *k.* And the total possible CH changes over entire travel time, *T* is *T*k.* Hence, the degree of individual CH status stability, c_{h_stab} is described by the equation in (3-30).

$$c_{h_{stab\%}} = \frac{\sum_{t=1}^{t=T} k - c_{h\Delta\_t}}{kT} \quad (3\text{-}30)$$

Where $c_{h\Delta\_t}$ defines the number of clusters that has changed CH at a specific time, *t.*

Given an array of logical output resulting from observing the CH status of clusters at a given time to be; $L_{c_{h\Delta}} = \left[\bigvee_0^1 L_{c_{h\Delta i_1^k}} : L_{c_{h\Delta 1}}, L_{c_{h\Delta 2}} \dots \dots \dots . . L_{c_{h\Delta k}}\right]$, the metric estimating the number of times all CH set has maintained their CH status is estimated as shown in equation (3-31)

$$c_{h_{all\_stab\%}} = \frac{\sum_{t=1}^{t=T} \bigwedge_1^k L_{c_{h\Delta} i}^t}{T} \quad (3\text{-}31)$$

For the percentage side-link stability measure, $c_{h_{time\_stab\%}}$, we determined the time proportion with which each side-link spanned relative to the total travel time as described in equation (3-32).

$$L_{stab\%} = \frac{\sum_{x=1}^{x=N_l} \frac{t_x}{T}}{N_l} \quad (3\text{-}32)$$

Where t_x is the time duration a side-link, x lasts, while N_l is the number of side-links which can be estimated as;

$$N_l = N - k \quad (3\text{-}33)$$

3.8 Result and Discussion

This section presents the results that have emerged from comparing the clustering and re-clustering models we have developed to baselines and divergent schemes.

The first result presented as shown in Figure 3-11 is the CDF of the throughput performance of the V2V CH-to-CM links for our k-means-based SNR clustering scheme and KmFW. The result indicates that the probability of a vehicle selected at random will have a throughput falling below 4 Mbps if the KmFW scheme is used is 0.9. However, for our SNR clustering scheme, 9 of 10 vehicles can reach up to 6.5Mbps.

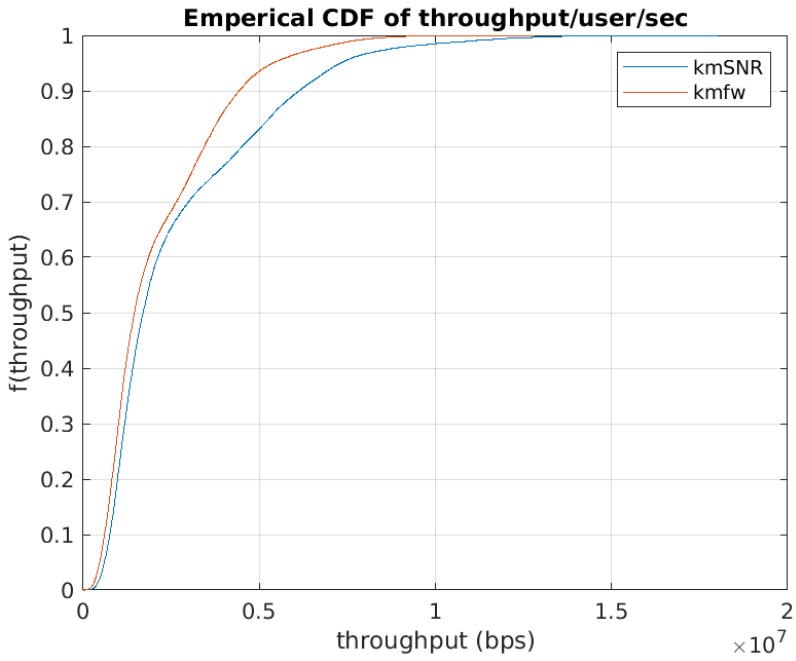

Figure 3-11. CDF of user throughput for K-means clustering algorithm variants

The superior throughput is accounted for by the communication-centric accuracy introduced by using SNR over proximity for vehicle grouping. However, KmFW performs slightly better in terms of cluster and link stability as shown in ***Error! Reference source not found.***.

Figure 3-12. Reliability measure of clustering methods

Having demonstrated the superior throughput performance of KmSNR clustering over KmFW, however with a compromise on stability, we will now be presenting results showing how our new CWS and SWS techniques built upon the k-means-based SNR clustering scheme performed in terms of stability and jitter while sustaining throughput performance.

The result in Figure 3-13 indicates the reliability of V2V CH-CM links for each of the clustering approaches. Two cluster stability and one link stability measures are used to estimate the reliability of the communication system. It can be seen clearly that the introduction of the memory-based waterfall techniques significantly improved the stability of the k-means-based SNR clustering scheme. Both waterfall schemes can be seen to have performed better than the KmSNR and KmFW over all the reliability measures.

We believe this is due to the elimination of the seeding component in subsequent re-clustering phases and the cascading of the clustering phases. The fixed scheme performed best as expected, since their clusters and hence links are kept unchanged, though at the cost of V2V vehicle participation and throughput. The two memory-based schemes can be observed to have comparable values in both the CH stability and link stability evaluation.

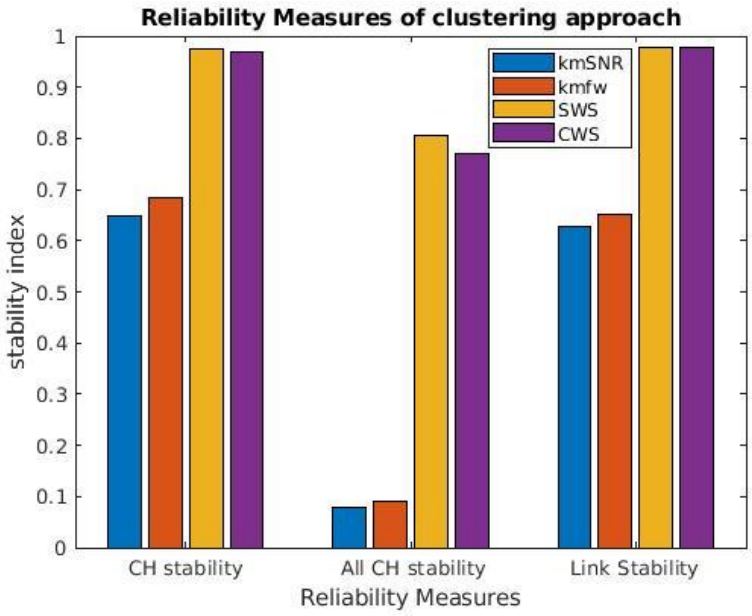

Figure 3-13. Bar chart comparing reliability parameters.

For jitter, it can be clearly seen in Figure 3-14 that the KmSNR clustering schemes outperforms KmFW. However, our memory-based waterfall schemes further outperforms both discrete clustering schemes (KmSNR and KmFW).

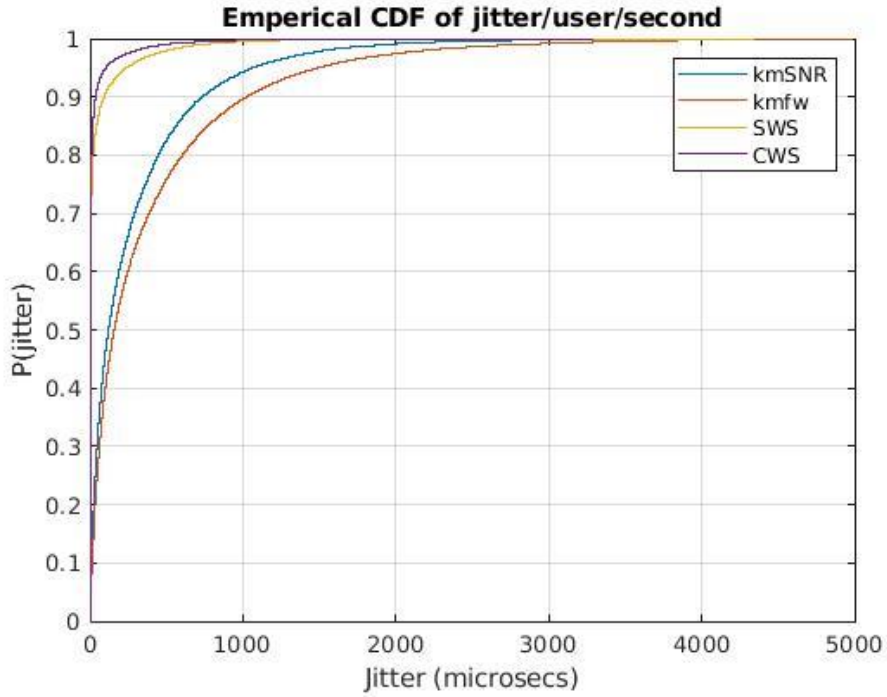

Figure 3-14. CDF plot of jitter performance for clustering approaches

For both memory-based schemes at least 95% of the vehicles has a jitter below 500 microseconds while for the discrete clustering schemes (KmSNR and KmFW) have at most 80% of the vehicles below the 500 microseconds of jitter.

The result in Figure 3-15 reiterates the relatively poor jitter performance of KmFW. We attribute the conspicuously poorer jitter performance of the KmSNR clustering and KmFW clustering to the mutual discreteness between their clustering phases unlike the memory based cascaded clustering phase schemes (SWS and CWS). We attribute the superior jitter performance of CWS to the more direct relationship between the mean cluster position output of a preceding clustering phase and the centroid of the succeeding clustering phase.

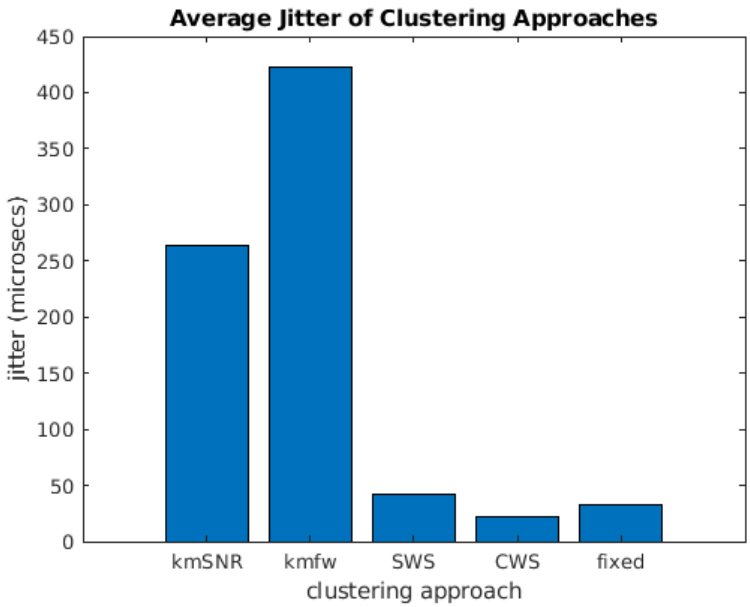

Figure 3-15. Bar chart of average jitter for each centroid re-selection scheme

The SWS scheme demonstrates a 5-fold reduction in average jitter while the CWS shows 10-fold reduction in average jitter compared to the discrete KmSNR scheme. KmFW demonstrated the poorest average jitter performance.

Having demonstrated improvement in jitter and reliability performance of the cascaded clustering phase schemes over KmSNR and KmFW, we now present results demonstrating how the different techniques of cascading clustering phases fared in terms of throughput capacity.

The result in Figure 3-16 and Figure 3-17 show that the V2V throughput capacity for our waterfall schemes remains undiminished despite considerably improved stability and jitter performance.

As expected the number of vehicles able to download the required map data increases with time threshold allowed for download in all the schemes. However, the waterfall schemes sustain the gain in throughput experienced without the waterfall technique and even seem to show slight improvement over other schemes at every time threshold considered. The result further reveals CWS's superior download capacity over SWS, partly due to the averaging approach to centroid selection used in CWS.

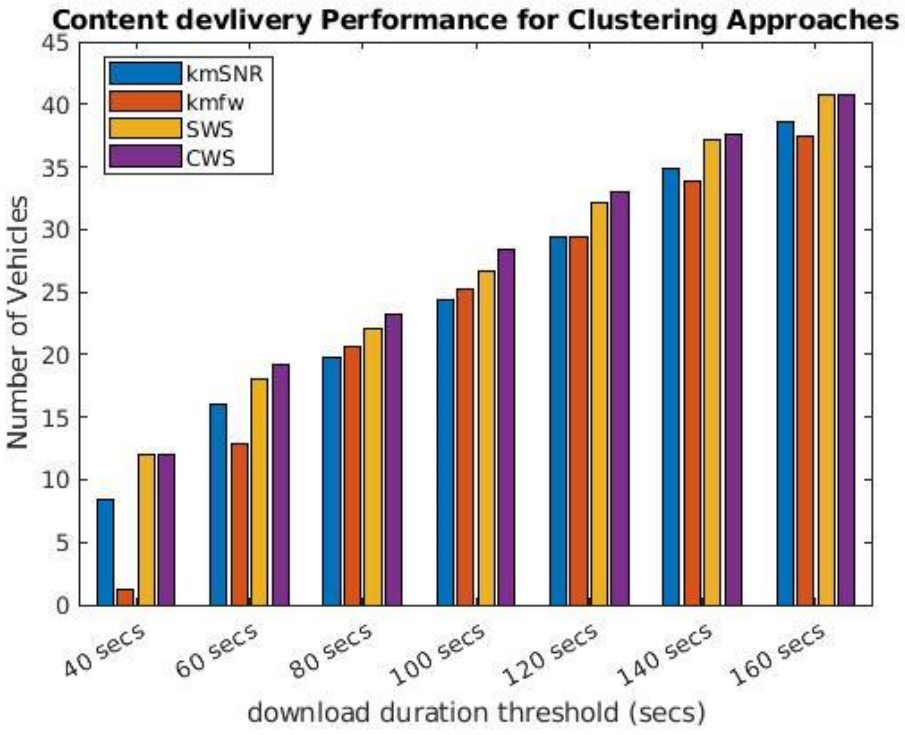

Figure 3-16. Bar chart comparing content delivery capacity.

We have presented the sum-rate for each clustering scheme in Figure 3-17 to consolidate on our claim that the stability and jitter improvement in CWS and SWS has not jeopardised its throughput gain.

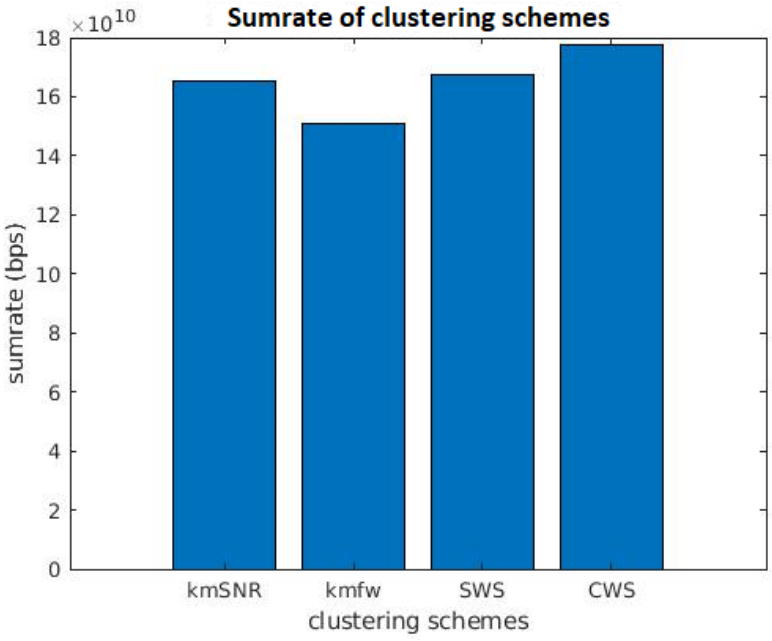

Figure 3-17. Sum-rate of re-clustering schemes

3.9 Conclusion

The work presented in this section describes the design of a k-means based SNR (KmSNR) clustering scheme upon which we built two different memory-based re-clustering schemes in which clustering phases in time are cascaded. The communication resource to which clustered entities operate are allocated based on cluster-oriented LTE-V2X mode-3 SPS scheme overlaid over a single-cell cellular network. We establish that our KmSNR scheme showed an improvement in throughput performance. We also demonstrated that our memory-based schemes showed considerably better stability and resistance to jitter without compromising on throughput gain achieved in our KmSNR approach. We have demonstrated that with our schemes more vehicles are able to completely download a data within a given time constraint and within V2V resources allocated without apparently constituting resource and interference burden on traditional users of the cell. Our scheme is applicable in scenarios where a group of vehicles approach a geo-fenced traffic-region and requires urgent download of critical traffic information. It will also find use in a case where download of real-time maps in a highly dynamic traffic environment is required.

In this work, re-clustering is assumed to be at certain time intervals. For our future work, we propose an event based re-clustering based on cluster and link performance quality, with the aim to further improve V2V throughput and stability.

4 Novel Objective Cluster-Update Scheme

4.1 Introduction

As new real-time and bandwidth intensive vehicular applications such as Augmented Reality (AR) road signs and HD map streaming emerge, developing a new V2X network paradigm to accommodate these applications has become pertinent. Vehicular clustering schemes have been considered as promising solution for reliability and Base Stations (BS) spectrum resource demand issues in a vehicular network environment. However, cluster viability and network performance of cluster based V2X networks drop as vehicles move in time and distance. Hence, the need for clusters to be updated or re-formed at intervals to sustain compliance with use-case requirements. In our work discussed in section 3, re-clustering is initiated at fixed time intervals and is totally agnostic of cluster or network performance. In this work, we developed two context-oriented re-clustering decision schemes, one based on a cluster performance index and the other based on a network performance index. We named the schemes, Cluster Quality Re-clustering Decision (CQRD) scheme and Link Quality Re-clustering Decision (LQRD) scheme respectively. CQRD is inspired by the Calinski-Harabasz index of optimum number of cluster selection. However, unlike the conventional Calinski-Harabasz scheme in which index evaluation criterion is based on Euclidean distance, our scheme's criterion is based on Signal-to-Noise Ratio (SNR). LQRS on the other hand is based on centrally

estimated mean V2V link throughput. In both schemes, the estimated criteria are used to compare current cluster viability against a logically reformed cluster. We compared our schemes with a persistent re-clustering scheme and Cluster Head (CH) membership evaluation scheme. Considering equally weighted throughput, jitter, and stability performance indices, LQRD outperforms the two baseline schemes in both a highway and urban scenario, while CQRD outperforms the CH membership baseline scheme in both scenarios.

4.2 System Model

We adopted the same highway V2V channel model and resource allocation model as used in our previous work reported in chapter 3. However, in this section, we also tested our proposed algorithm in the urban area. This section intends to describe the channel model we have adopted for our urban environment.

We adopted an urban V2V path loss, blocking and shadow model from [148] for both LOS (Line of Sight), Non-Lone of Sight (NLOS) and NLOSv (Non-Line of Sight for vehicles on the same street).

$$PL_{los} = 38.77 + 16.7\log_{10} d_{euc} + 18.2v\log_{10}(f) + X_{\sigma} \tag{4-1}$$

$$PL_{nlos} = 36.85 + 30\log_{10} d_{euc} + 18.9\log_{10}(f) + X_{\sigma}$$

$$PL_{losv} = 38.77 + 16.7\log_{10} d_{euc} + 18.2v\log_{10}(f) + A_{sk} + X_{\sigma}$$

$$\begin{gathered}\text{When } v = -0.7 \\ A_{sk} = 6.9 + 20\log_{10}\left[\sqrt{(v-0.1)^2+1} + v - 0.1\right] \\ \text{Else}, A_{sk} = 0\end{gathered} \tag{4-2}$$

$$v = \sqrt{2} \times \frac{H_d}{r_f} \tag{4-3}$$

$$r_f = d_{hm} \times 0.25 \times \sqrt{\frac{n_f \times \lambda}{d_{euc}}} \quad (4\text{-}4)$$

X_σ Is the shadowing parameter, H_d is the difference between obstruction height and horizontal link plane, n_f is the Fresnel number, d_{euc} is the Euclidean distance between VUEs or length of V2V links and λ is wavelength.

The probability of LOS and NLOSv is estimated based on a distance threshold of 100 meters. The probability of NLOS for each V2V link is estimated individually based on whether their difference, d_δ against the distance threshold is positive or negative [148], as described in equation (3-6). For the urban environment the approach is different, the individual link distance is directly plugged into the LOS probability equation in (4-5). If the probability of LOS is smaller than the probability of NLOS, hen the looked is taken to be LOS and vice versa. More on this can be found in [148].

$$P(LOS) = min\{1,1.05 * e^{-(0.0114d)}\} \quad (4\text{-}5)$$

$$P(NLOS) = 1 - P(LOS) \quad (4\text{-}6)$$

Log-normal distribution with zero mean and a specific value of standard deviation, σ is assumed for shadow fading. The values of σ are adjusted to accurately define specific environment and link types. The values defined for our different link types based on line-of-sight definitions as specified in ETSI technical report 103 257-1 [148] is presented in Table 4-1.

Table 4-1 Shadow fading standard deviation parameter for V2V.

Link Type	Urban	Highway
LOS	5.2 dB	3.3 dB
NLOSv	5.3 dB	3.8 dB
NLOS	6.8 dB	n/a

4.3 Cluster Formation and Optimal K-Selection.

The clustering process employed here is quite similar with that described in section 3.5. It employs the KmSNR approach to clustering the vehicles. However, the approach earlier described made use of heuristics to define the range from which the number of clusters are selected. The definition of the number of clusters or a static input from which number of clusters can be selected is one drawback that limits the performance of kmeans-based clustering schemes in dynamic scenarios. In this chapter, aside from our core work of performance-based determination of re-clustering instances, we have also opted for an approach of determining the optimum number of clusters, K_{opt} using an adaptation of elbow method like the approach described in [149]. Though the elbow method generally uses visual inspection to determine the optimal number of clusters [150], [151], here we employed a quantitative method to obtain the elbow value to address the elbow method reliability issues stemming from the ambiguity associated with the visual inspection of the elbow point.

Our elbow method variant estimates the intra-cluster proximity using the sum of square Euclidean distance estimate. This estimate evaluates the sum of the distance of each vehicle point to their respective cluster centroid. Then averages the sum over the entire set of vehicle points at that instance. This is repeated for different numbers of clusters, *k*. The resulting average similarity index of the entire set of vehicle points are then plotted against *k*, which essentially ranges from 1 to the number of vehicle points, *N*. So, the smaller the average sum of square estimate the better cluster performance. The point on the graph beyond which the mean sum of the square estimate becomes hardly sensitive to changes in number of clusters indicates the optimal number of clusters. Our variant specifically seeks to quantitatively identify that point.

For a vehicular network instance X with N vehicular points and k clusters, where $k = \{1,2, \dots \dots . . N\}$ and $C = \left\{C_{i_1^k}: C_{1,} C_{2,} \dots \dots \dots . . C_{k,}\right\}$. Here, C_i represents individual cluster of a vehicular network instance in time. The centroids of the clusters, $c = \left\{c_{i_1^k}: c_{1,} c_{2,} \dots \dots \dots . . c_{k,}\right\}$ corresponding to the individual clusters in set *C* at the time instance X is generated from the converged k-means update as described in sub-section 3.5.1.2. For a given number of clusters, *k* and a set of vehicle points in a cluster C_i at a time instance X given as; $V =$

$\left\{v_{u_1^{n\leq N}}^{i\leq k}: v_1^i, v_2^i \ldots\ldots\ldots\ldots v_n^i\right\}$ and with coordinates *(x, y),* we define the sum of square of Euclidean distance, *SSED* as described in equation (4-8), the mean (over the total number of vehicle points, *N*) of which, we call the mean distortion, *MD* and describe in equation (4-8).

$$SSED_k = \sum_{i=1}^{i=k} \sum_{u=1}^{u=n} (v_{u_x}^i - c_{1_x})^2 + (v_{u_y}^i - c_{1_y})^2 \tag{4-7}$$

$$MD_k = \frac{SSED_k}{N} \tag{4-8}$$

Having estimated the MD_k for each *k* number of clusters, we then normalised the each resulting value using the MinMaxScaler. The normalised value is then scaled between the ranges of 0 to 10, to obtain a normalised scaled value of MD_k, N_{md}^k as described in equation (4-9).

$$N_{md}^k = \frac{MD_k - MD_{(min)}}{MD_{(\min)} - MD_{(\min)}} * 10 \tag{4-9}$$

Considering two adjacent points *i* and *j* on the 2-dimensional plot of N_{md}^k against *K* in Figure 4-1. The points *i* and *j* can be described by coordinates (md_i, k_i) and (md_j, k_j) with each pair of the coordinates representing a value along the N_{md}^k and *K* axis respectively. To form the elbow curve we joined adjacent points on the plot and estimated the Euclidean distance between them. The Euclidean distance between the points *i* and *j* is given as:

$$D_{ij} = \sqrt{\left(md_i - md_j\right)^2 + \left(k_i - k_j\right)^2} \tag{4-10}$$

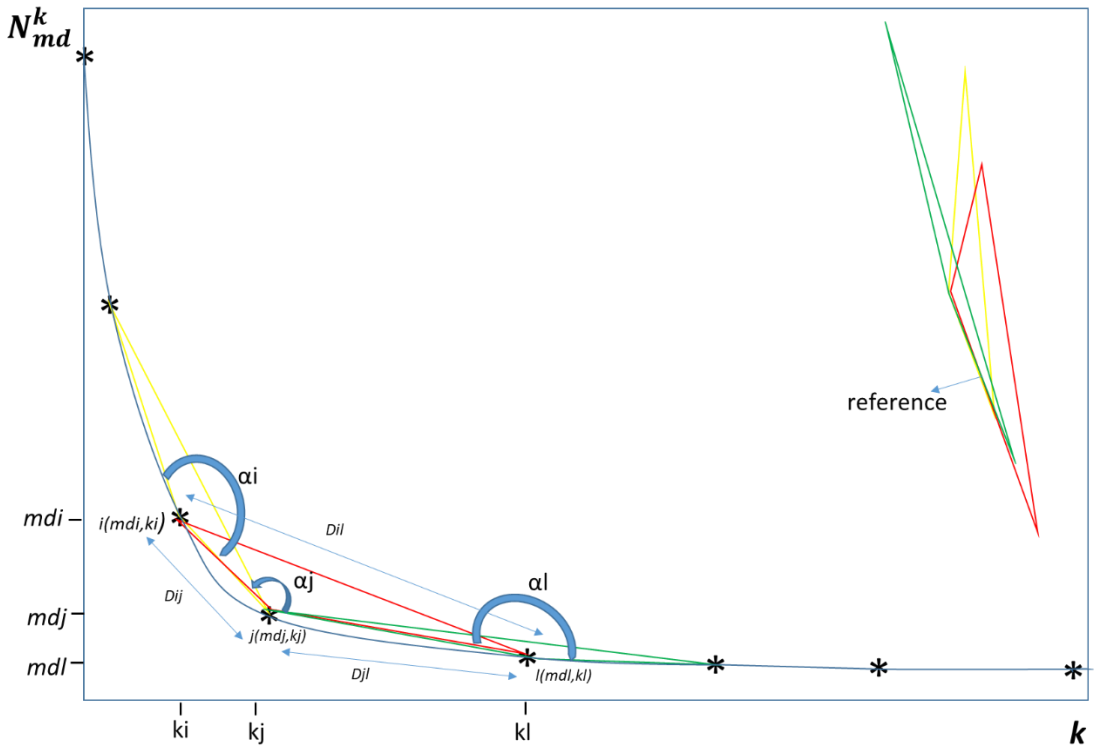

Figure 4-1. Elbow plot of number of clusters against to normalised mean distortion

Now considering 3 points, *i, j* and *l,* forming a triangle with adjacent points on either side themselves. We can estimate the angles, α_i, α_j and α_l that subtends at the points *i, j* and *l* using the cosine trigonometric rule as described in equation (4-11).

$$\alpha_j = \cos^{-1} \frac{D_{ij}^2 + D_{jl}^2 - D_{il}^2}{2D_{ij}D_{jl}} \tag{4-11}$$

As can be observed in the plot, α_j subtending at *j* is the most acute angle and represents the point beyond which mean distortion becomes minimally sensitive to changes in number of clusters and corresponds to the number of clusters k_j.

After evaluating the optimum number of clusters, K_{opt}, we then cluster the vehicles using the SNR-based node association component of the waterfall schemes, KmSNR as described in section 3.5.

4.4 Re-clustering Decision Architecture

After clusters have been formed based on the our KmSNR approach described in section 3.5.1 and re-clustering is based on CWS described in section 3.6, since it demonstrated relative superior performance in the performance metrics considered, here we describe the general architecture upon which our re-clustering decision schemes are developed. We have developed a block representation of our decision architecture and presented in Figure 4-2.

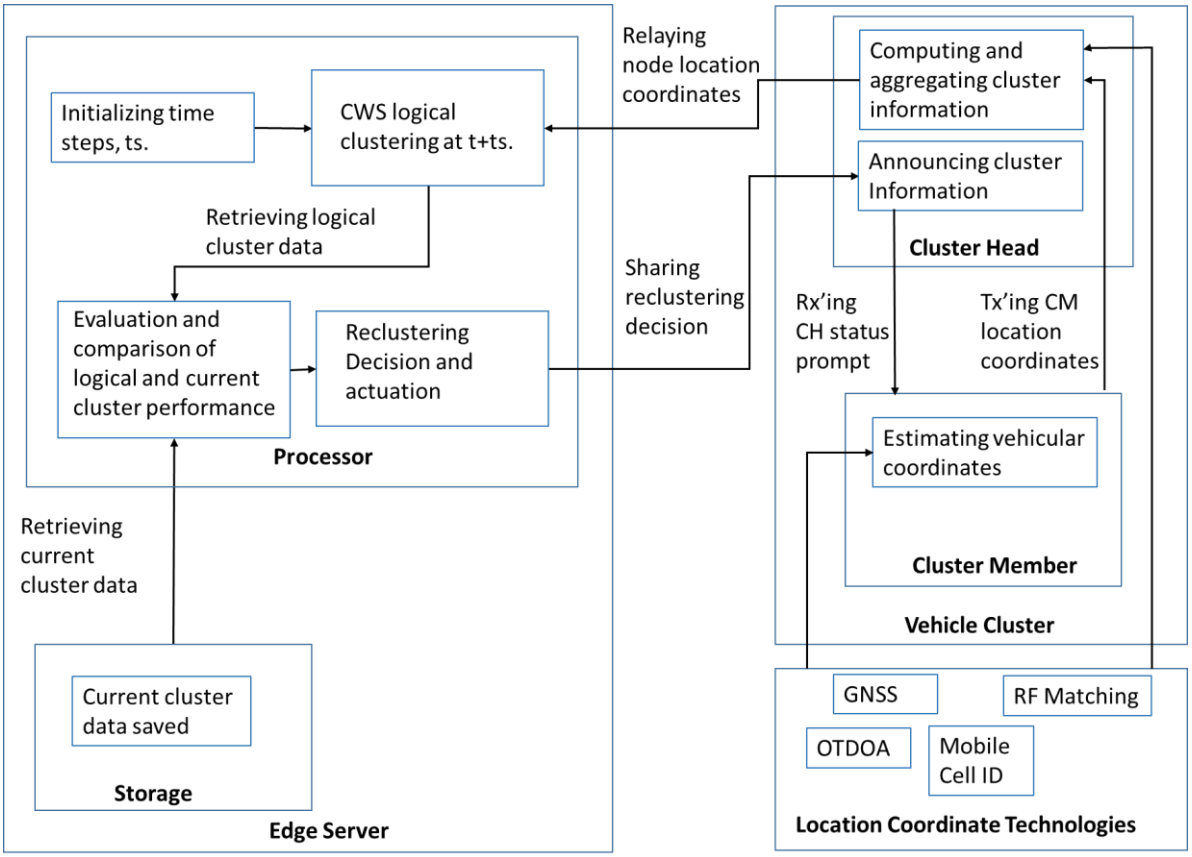

Figure 4-2. Re-clustering decision architecture

After the initial clustering is evaluated, implemented and cluster configuration stored in the edge server, a time interval, t_s, is defined at which the cluster performance is evaluated. As the clusters lifespan/travel duration approaches t_s, the edge server via the BS requests for location coordinates of all vehicular nodes within and outside current clusters. Vehicle nodes could use different location technologies to calculate their position coordinates. For vehicles outside the cluster their location coordinate is sent directly to the edge server via the BS and Uu interface, while vehicles within the cluster have their coordinate information sent via the PC5 interface to the CH and then relayed to the edge server via the BS and the Uu interface.

After the vehicle node position data has been received by the edge server, a logical re-clustering process is initiated based on the collected current vehicle coordinates and our CWS re-clustering scheme. After the logical clusters are formed, the cluster configuration performance is evaluated based on performance parameters of interest. The current cluster configuration is retrieved from edge storage and its performance is also evaluated based on the new position coordinates and same performance parameters by which the logical clusters are estimated. The performance of the logical and current cluster configuration is compared and the cluster configuration with the better performance determines if the current clustering configuration be sustained or the logical cluster be implemented. In essence if the logical cluster demonstrates superior performance, a new clustering configuration based on the logical cluster gets implemented, else the current cluster is sustained. If changes are required, the new CHs are notified, and they'll share their status with neighbouring vehicular nodes for association.

Section 4.5 and 4.6 describes the performance parameters we have employed to decide on when to re-cluster before discussing how these decision algorithms and parameters impacts on certain QoS parameters in section 4.7.

4.5 Cluster Quality Re-clustering Decision (CQRD) Scheme

Our CQRD scheme is inspired by the Calinski-Harabasz Criterion also known as the Variance Ratio Criterion (VRC). It leverages the C-V2X's BS and edge computing central coordination function to execute initial cluster formation and estimation of VRC index to make re-clustering decisions. We have defined the Calinski-Harabasz or *VRC* index for the entire vehicular network is defined as

$$VRC = \frac{SNR_{var\_in}}{SNR_{var\_out}} \times \frac{(k-1)}{(N-k)} \tag{4-12}$$

Where SNR_{var_out} is overall inter- cluster SNR variance, SNR_{var_in} is the overall intra-cluster SNR-based variance and *N* is the number of vehicles and *k* is the number of clusters. The similitude of SNR_{out} in traditional Calinsky-Harabasz estimate defines proximity between

clusters and central node position, however we used it here to estimate the strength of links between CHs and theoretical central node positions.

$$d_{i\_out} = d_i(m_i, m_m) = \sqrt{(m_{ix} - m_{my})^2 + (m_{iy} - m_{my})^2} \tag{4-13}$$

$$snr_{i\_out} = f(d_{i\_out}) \tag{4-14}$$

$$SNR_{var\_out} = \sum_{i=1}^{k} n_i * snr_{i\_out} \tag{4-15}$$

Where n_i is the number of vehicles in cluster *i,* m_i is the geographical coordinate of the vehicular centroid of cluster *i,* $m_m$is the overall mean position of all vehicles, and $d_{i_out}$is the distance between m_i and m_m. The theoretical SNR, snr_{i_out} between m_i and m_m is estimated as function of distance, $f(d_{i\_out})$ adopting the approach used in section 3.3. snr_{i_out} is then summed over the number of clusters, k and defined as the inter-cluster SNR variance, SNR_{var_out}.

Overall intra-cluster SNR-based variance, SNR_{var_in} essentially defines link strength between CHs and their corresponding CMs. It measures how closely clustered vehicle nodes are in terms of signal strength.

$$snr_{i\_x\_in} = f(d_i(m_i, m_{i\_x})) \tag{4-16}$$

$$SNR_{var\_in} = \sum_{i=1}^{k} \sum_{x=1}^{x=n} snr_{i\_x\_in} \tag{4-17}$$

Where m_i remains the coordinate of vehicular centroid and CH, m_{i_x} is the coordinate of cluster member x in cluster C_i, $snr_{i\_x\_in}$ represents the SNR between CH and CM in cluster C_i. A procedural description of how CQRD fits into the re-clustering decision architecture described in section 4.4 is presented in the pseudocode in Algorithm 1. Hence a well-defined

cluster will have a larger SNR_{var_in} and a relatively smaller SNR_{var_out}. In essence the larger the *VRC* the better the cluster quality of the vehicles.

Algorithm 1: Cluster Quality Re-clustering Decision (CQRD) Algorithm

i/o
Input: vehicular node coordinates, CQI & current cluster configuration, C_i
Output: new or affirmed cluster configuration, , C_o
Initialization
V: Set of vehicle nodes
VN: vehicle nodes
N: Number of VN, |V|
K: Number of clusters
SRS: sounding reference signal
T: time instant
Td: total time duration
ts: Initialize re-clustering assessment interval
Cl: Logical cluster configuration
Re-clustering Process
For t=1:td
    i= ts:ts:td
    If t≠ i
        sustain C_i
    Else
        BS requests CM-to-CH side-link SRS & all vehicular coordinates, m
        Current Cluster Variance Evaluation
        For all clusters at i
            BS extrapolates CM-to-CH side-link SNR_{in}
        End
        BS estimates the mean position of all VN, m_m
        For all clusters at i
            BS estimates the SNR_{out} between centroids to the mean vehicular position, m_m
        End
        Estimate Intra-and inter-cluster SNR_{var}
        Estimate VRC_{3c} for current cluster configuration is estimated
        Logical Cluster Variance Evaluation
        BS evaluates new logical cluster configuration, Cl using CWS based on current vehicular coordinates, m.
        For all clusters
            BS estimates the SNR_{out_l} between centroids to the mean vehicular position, m_m
        End
        Estimate Intra-and inter-cluster SNR_{var_l}
        Estimate VRC_{l2c} for current cluster configuration
        Comparator
        If $VRC_{3c} \geq VRC_{l2c}$
            sustain C_i ($C_i = C_o$)
        Else
            adopt C_l as C_o
            $C_l = C_o$
        End
    End
End

CQRD at specific intervals initiate a logical re-clustering, then estimates the *VRC* index of logical clusters and compares to the actual current cluster status. If the logical cluster exhibit a higher *VRC* than the current cluster state, an actual re-clustering process is initiated. However, if the logical re-clustering exhibits lower *VRC* index as compared to the current cluster state, then the current cluster partition is maintained.

The input for the CQRD algorithm's is current cluster configuration, vehicular node coordinates and CQI, while the output is the new or affirmed cluster configuration. At every time interval, t_s the BS requests vehicular cluster configuration information, node coordinates and side-link CQI information. From this information, the BS extrapolates intra and inter-cluster channel condition information SNR_{in} and SNR_{out} and then computes SNR_{var_out} and SNR_{var_in} as described in equations (4-15) and (4-17) respectively, and subsequently these values are used to estimate the SNR-based variance ratio criterion for the current cluster configuration, VRC_{3c}. This represents a measure of the cluster quality of the current cluster configuration. Also, at the same instance, vehicular nodes are logically re-clustered based on Centroid-based Waterfall Scheme (SWS) described in section 3.6 of chapter 3. Again, the values of $SNR_{var_{out}}$ and $SNR_{var_{in}}$ are extrapolated and used to estimate the cluster quality parameter of the new logical cluster configuration, VRC_{l2c}. The BS then compares the VRC values of the current cluster, VRC_{3c} to that of the logical cluster, VRC_{l2c}. If $VRC_{3c} \geq VRC_{l2c}$ then the current cluster configuration is sustained, else the logical cluster configuration is adopted, published to vehicular nodes and subsequently implemented.

4.6 Link Quality Re-clustering Decision (LQRD) Scheme

We designed LQRD to make re-clustering decision based average throughput of clusters. The aim is to minimize re-clustering overhead while improving link throughput, c_l between CH and CM, defined as:

$$\rho_{l_x} = W \log_2(1 + \gamma_x) \tag{4-18}$$

Where W is the bandwidth allocated to each CH-to-CM side-link and γ_x is signal to interference and noise ratio of the side-link

Algorithm 2: Link Quality Re-clustering Decision (LQRD) Algorithm

i/o
Input: vehicular node coordinate, current cluster configuration, Ci & CQI
Output: new or affirmed cluster configuration, C_o
Initialization
V: Set of vehicle nodes
VN: vehicle nodes
N: Number of VN, |V|
K: Number of clusters
SRS: Sounding Reference Signal
CQI: Channel Quality Indicator
SINR: signal-to-Interference-and-noise-ratio
T: time instant
Td: total time duration
ts: Initialize re-clustering assessment interval
Cl: Logical cluster configuration
Re-clustering Process
For t=1:td
  i= ts:ts:td
  If t≠ i
    sustain C_i
  Else
    BS requests CM-to-CH side-link C & all vehicular coordinates, m
    Current Cluster Throughput
    For all clusters at i
      BS extrapolates CM-to-CH side-link $SINR_{x\_cc}$ from CQI
      Estimate side-link throughputs, $\rho_l{}^{i}_{x\_cc}$
      Estimate mean cluster throughput $\overline{\rho_{l\_cc}^{l}}$
    End
    BS estimates the overall mean throughput, $\overline{\rho_1}$
    Logical Cluster Variance Evaluation
    BS evaluates new logical cluster configuration, C_l using CWS based on current vehicular coordinates, m.
    For all clusters
      BS estimates the $SINR_{x\_lc}$ between CH to CMs, m_m
      Estimate side-link throughputs, $\rho_l{}^{i}_{x\_lc}$
      Estimate mean cluster throughput $\overline{\rho_{l\_lc}^{l}}$
    End
    BS estimates the overall mean throughput, $\overline{\rho_2}$
    Comparator
    If $\overline{\rho_1} \geq \overline{\rho_2}$
      return to 17
    Else
      adopt C_l as C_o
      $C_l = C_o$
    End
  End
End

Like CQRD, LQRD is proposed as a non-threshold scheme that compares the mean throughput, $\overline{\rho_1}$ of CH-to-CM side-links in the current cluster to the mean throughput, $\overline{\rho_2}$ of CH-to-CM side-links in the logically re-clustered cluster configuration of same vehicles. The decision to re-cluster is only made when the mean overall throughput of links in the logical cluster configuration exceeds that of the current cluster configuration. At this point the logical clustering information is now communicated to the selected CHs for action and onward propagation and compliance. The average overall cluster, $\overline{\rho_o}$ is defined as

$$\overline{\rho_o} = \frac{1}{\sum_{i=1}^{i=k} n_i} \sum_{i=1}^{i=k} \sum_{x=1}^{x=n} {\rho_l}_x^i \tag{4-19}$$

$$where\ \overline{\rho_o} = \bigvee_{o=1}^{o=2} \overline{\rho_o}$$

$$\bar{\rho} = \max(\overline{\rho_1}, \overline{\rho_2}) \tag{4-20}$$

Both $\overline{\rho_1}$ and $\overline{\rho_2}$ are estimated and the maximum of which decides what action is initiated. Re-clustering is only initiated when $\overline{\rho_1} \geq \overline{\rho_2}$.

Where ${\rho_l}_x^i$ represents throughput of vehicle side-link x in cluster i, k represents the number of clusters and n represents the number of vehicles in a cluster. Algorithm 2 describes the step-by-step approach in which the LQRD is implemented within the re-clustering decision framework discussed in section 4.4.

Just like the CQRD algorithm, the input for the LQRD algorithm are vehicular node coordinates, current cluster information, c_i and CQI, while the output is the new or affirmed cluster configuration. At every time interval, t_s the BS requests node coordinates and CQI. From this information, the BS extrapolates side-link signal-to-interference-and-noise-ratio for the current clusters, $SINR_{x\_cc}$ from CQI information and computes mean side-link throughput of the network, $\overline{\rho_{l\_cc}^l}$ as described in equation (4-19). Also, at the same instance, vehicular nodes are logically re-clustered based on Centroid-based Waterfall Scheme (CWS) described in section 3.6 of chapter 3. In this case, the value of side-link signal-to-interference-and-noise-ratio for the logical clusters, $SINR_{x\_lc}$ are estimated and used to compute mean

side-link throughput of the network, $\overline{\rho^{i}_{l\_lc}}$. The BS then compares the mean throughput values of the current cluster, $\rho^{i}_{l\_cc}$ to that of the logical cluster, $\rho^{i}_{l\_lc}$. If $\rho^{i}_{l\_cc} \geq \rho^{i}_{l\_lc}$ then the current cluster configuration is sustained, else the logical cluster configuration is adopted, published to vehicular nodes, and subsequently implemented.

4.7 Performance Evaluation

This section presents the performance criteria by which we have evaluated the CQRD and LQRD. We compared both clustering schemes with a persistent re-clustering scheme and another commonly used scheme based on CH membership validity.

The performance of our decision algorithms is tested for both urban (ring road, Bologna, Italy) and highway-rural (A64, Yorkshire, UK) scenario. The vehicles move at varying speed and in different directions. The vehicle traces for Bologna ring road were collected from a real-life traces of the traffic along the road while a random trip traffic scenario was generated from SUMO OSM (Open Street Map) Web Wizard for the vehicular traces of A64 highway rural highway. Detailed information on scenario generation from the OSM web wizard has been in section 3.2 and more information can be found in [113].

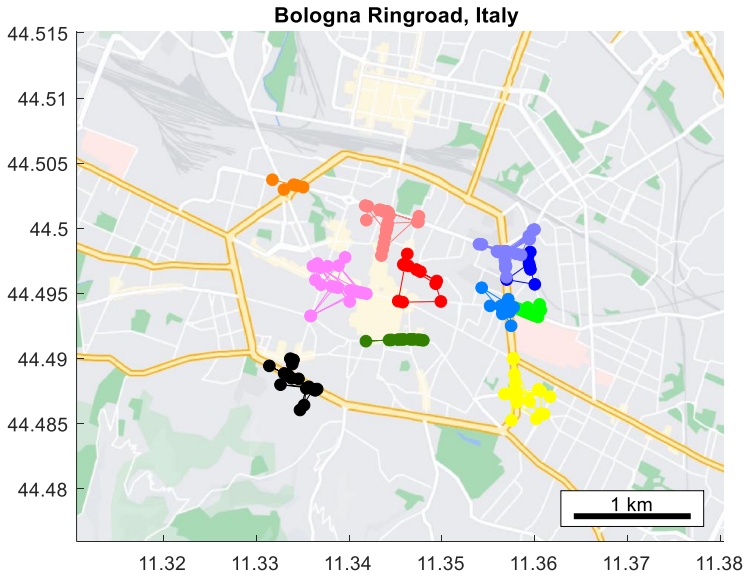

Figure 4-3. Snapshot of clustered vehicle traces in Bologna ring road

A snapshot depiction of the vehicle clusters on the bologna ring-road network is depicted in Figure 4-3, while that depicting vehicle clusters on the A64, York road is shown in Figure 4-4. The maps were generated from google maps, read into MATLAB and vehicle coordinates were overlaid upon them.

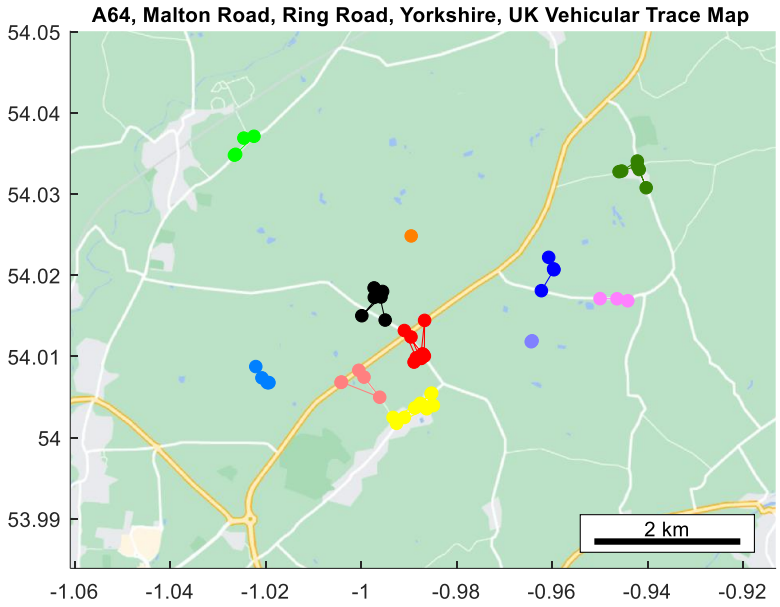

Figure 4-4. Snapshot of clustered vehicle traces in A64, Malton Road, Ring Road, Yorkshire, UK

The coloured dots on the map represents the position of vehicles on the map. The different colours of vehicles represent different clusters. Vehicles of same colours belong to the same clusters while vehicles of different colours belong to different clusters. The x-axis and the y-axis are labelled in longitude and latitude respectively to depict the geographical position of the vehicles. The grey pattern by the roadside in the urban Bologna map symbolises buildings and other urban structures, while the green pattern by the roadside in the rural highway, in Yorkshire, depicts vegetation. The yellow lines depict the main roads.

It is note-worthy that the work to be presented does not compare the performances in urban driving environment against the highway-rural environment but seeks to understand how the different re-clustering decision schemes perform relative to one another in both environments.

The simulation assumptions are set according to 3GPP TR 37.985 [163]. Table 1 shows the simulation parameters used.

Table 4-2. Simulation parameters

Parameters	**Values**
Bandwidth	10MHz
Carrier frequency	5.9GHz
Maximum transmission range	1000m
Distance threshold for NLOS	350m
Maximum Transmit power	23dBm
Noise power	-113dBm
Shadowing distribution	Log-normal $\sigma = 3dB$
Resource Block size	180kHz
Range of vehicle speed	[40,120]km/h
Highway model	ETSI TR 103 257 [72]
Urban Model	ETSI TR 103 257 [72]

The focus here is to evaluate the performance of the CH-to-CM V2V side-links for real-time download applications. We assume enough bandwidth access for CHs on the BS and only V2V can constitute a bottleneck in the 2-hop transmission.

For the V2V side-links, performance is estimated in terms of Cumulative Distribution Function (CDF) of V2V link throughput, and jitter over the entire travel time. We also estimated through normalization context-local sum-rate index, stability index, re-clustering event frequency index and equal-weight QoS index estimated from an equal weight combination of throughput, jitter, and stability.

We estimated the throughput, ρ_{l_x} of each V2V side-link, *x* as defined in equation (4-18). *W* is bandwidth which is modelled to be the same for all vehicles at every instant. So, essentially interference, γ_i is the factor that varies throughput. We have presented the equation for ease of reading, since we will be discussing other performance parameters that are based on the output therein.

$$\rho_{l_x} = W \log_2(1 + \gamma_x) \tag{4-21}$$

The aim is to evaluate and analyse the distribution of side-link throughputs and make reasonable inferences and conclusion on how the decision schemes have influenced the throughput performance of a side-link.

Sum-rate, ρ_o was estimated to account for the total throughput capacity to accommodate all the V2V links in each clustering scheme to have a holistic capacity estimation of the network. This is done by summing the throughput per CH-to-CM V2V link per time over entire travel time for each clustering approach. We then normalised and rescaled the sum-rate over [0 1] or [a b] as presented in equation (4-23).

$$\rho_o = \sum_{t=1}^{t=T} \sum_{i=1}^{i=k} \sum_{x=1}^{x=n} \rho l_{x\_t}^{i} \tag{4-22}$$

$$\rho_{o(norm)} = a + \left[\frac{\rho_o - \rho_{o(\max)}}{\rho_{o(\max)} - \rho_{o(\min)}}\right](b-a) \tag{4-23}$$

Where *T, k* and *n* are travelling time, number of clusters and number of CH-to-CM V2V side-links within a cluster C_i respectively.

Beyond throughput related performance evaluation, the performance of the re-clustering decision schemes will be evaluated using jitter and reliability metrics as presented in the equations between equation (3-29) to equation (3-32). These performance evaluation parameters are being re-applied because we believe that the chosen criteria which triggers re-clustering and the frequency of the trigger influences cluster stability and packet reception consistency.

Finally, we evaluated the decision algorithms using a performance metric we call Quality of Service (QoS) index. This is based on a post-normalisation weighted combination of throughput (ρ), jitter (ψ) and stability ($L_{stab\%}$) as presented in equation (4-24). The essence of this is to have a holistic insight of the best performing model based on all the considered criteria.

$$QoS_{idx} = \omega_1 \bar{\rho}_{o(norm)} - \omega_2 \bar{\psi}_{(norm)} + \omega_3 \bar{L}_{stab\,(norm)} \quad (4\text{-}24)$$

Where ω_1, ω_2 and ω_3 represents the respective weight coefficient in each of the QoS parameter considered. The weights might vary depending on the design focus and use case considered. $\bar{\rho}_{o(norm)}$, $\bar{\psi}_{(norm)}$ and $\bar{L}_{stab\,(norm)}$ are the normalised mean values of throughput, jitter, and stability index.

4.8 Results and Discussion

This section presents the results obtained from the performance evaluation and comparison made between the re-clustering decision models that we have developed (CQRD and LQRD) with key baselines (the persistent re-clustering scheme codenamed as PRS, and CH membership scheme codenamed as CHmem2). We first present the results demonstrating the performance of all the schemes in the urban vehicular traffic environment (Bologna Ring Road) and then present a result comparing the results of the urban environment with a highway rural vehicular environment (A64 Malton Road, Yorkshire).

The result presented in Figure 4-5 is a cumulative distribution of link throughput over the whole travel time for the different re-clustering decision schemes for the urban environment (Bologna ring road). It particularly demonstrates that LQRD and CQRD has superior throughput performance over CHmem and showed that LQRD has comparable performance as compared to PRS, despite relative overhead savings in LQRD and CQRD as shown in Figure 4-6. The relatively superior performance of CQRD over CHmem is attributed to its SNR-centric cluster quality-based re-clustering decision metric compared to the distance-based approach used in CHmem. Also, the consistent monitoring and comparison of the quality index of the current cluster configuration against the index of the newly logically re-clustered configuration constantly keeps the CH-CM side-link in proximity relative to straying interference links. Since the re-clustering process is initiated based on the CH-to-CM SNR, upon which throughput is also remotely dependent. We expect CQRD to show improved throughput performance over CHmem whose CH election is only based on node membership. This is because CQRD inherently improves SNR between cluster members. Unlike CQRD which uses cluster parameter, LQRD uses throughput, the actual network parameter upon which performance is evaluated to decide when clusters need to be reformed. The direct use of

overall average throughput of clusters for decision, directly influences the throughput performance of individual CH-to-CM links.

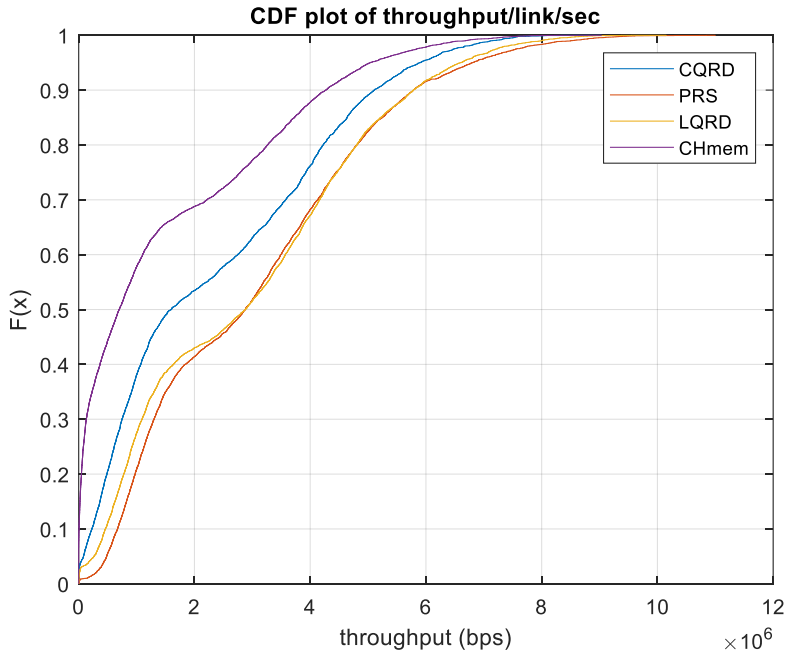

Figure 4-5. CDF plot of throughput per V2V link in time

It can be observed in Figure 4-5 that there is a lower proportion of PRS V2V links with throughput below the 2Mbps mark compared to LQRS V2V links, this is because PRS re-clustering is persistent, consequently persistently recentralizing the CH which keeps the throughput level above a specific limit. Unlike LQRD whose re-clustering is based on mean throughput, making the scheme susceptible to lower throughput before re-clustering.

Figure 4-6 shows the percentage number of clustering instances across the entire travel time. We use this re-clustering frequency phenomenon as a measure for communication overhead, since the re-clustering phases are accompanied by some signalling and transmissions. As can be observed, the PRS as expected shows 100% re-clustering instances over total evaluation instances, meaning at every point of evaluation, re-clustering is initiated. Both CQRD and LQRD have demonstrated lower overhead relative to PRS. While CHmem2 demonstrates even lower re-clustering frequency though at the cost of throughput performance.

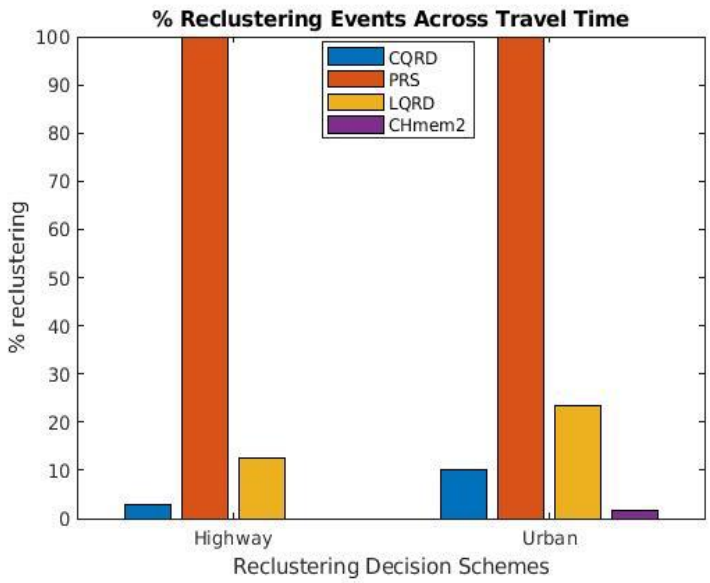

Figure 4-6.Communication overhead Indicated by percentage re-clustering frequency

In Figure 4-7, the throughput capacity performance of the different schemes is compared using sum-rate in both the urban and the highway-rural driving environment. Similar result patterns can be seen in both environments with our LQRD scheme demonstrating marginally larger V2V throughput capacity compared to the computational and overhead intensive PRS scheme. Our CQRD and CHmem both were inferior to the former schemes in this regard.

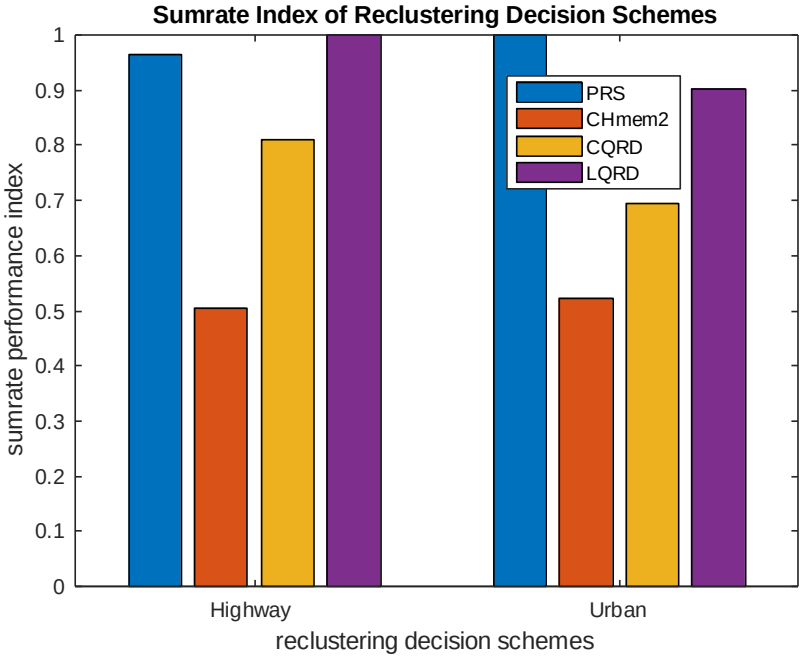

Figure 4-7. Sum-rate comparison for rural-highway and urban environment

However, our CQRD scheme demonstrates between 17 to 60% greater capacity as compared to CHmem which only exhibits a capacity of 72.18Gps and 12.47Gbps in both rural-highway and urban environment respectively. Given equal bandwidth, the capacity in the urban environment is relatively poorer to that exhibited by the rural-highway environment, and this is due to the lower probability of line of sight (LOS) in urban driving environments emanating from roadside structures and vehicular density, which then leads to a low participation of vehicles in V2V communication as shown in Figure 4-8 and consequently a low sum-rate.

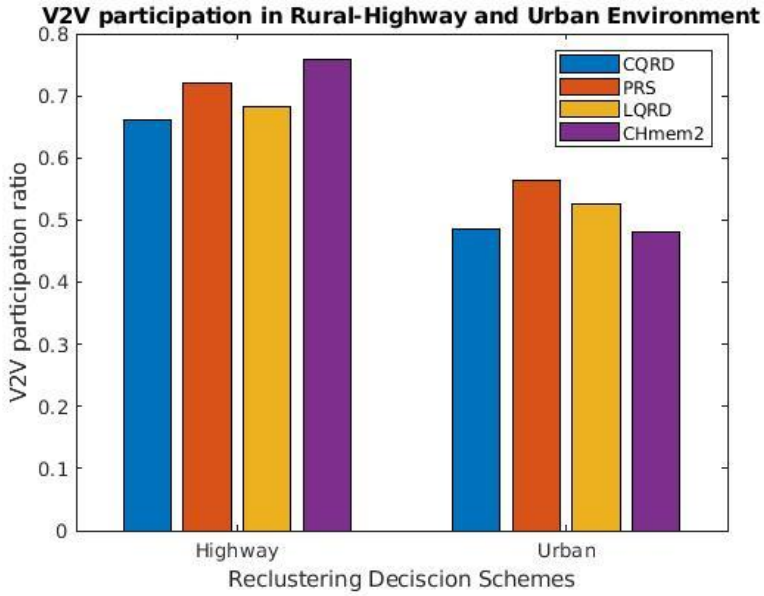

Figure 4-8. V2V participation ratio

The results in Figure 4-9 presents the CDF plot of the jitter experienced on each V2V side-link over the whole travel time for all the decision schemes considered. Here, it can be observed that the persistent approach, PRS demonstrates the poorest jitter while our LQRD scheme is best performing, while the CQRD and CHmem schemes have comparable performances with only marginal difference between them.

The poor performance of PRS can be attributed to changes in CH and cluster membership accompanying the persistent clustering in this approach. These changes tend to abruptly change side-link throughput as they change connection point while changing clusters or as they leave clusters to become free vehicles. Also, despite the seemingly stable CH of CHmem as will be seen in Figure 4-11, the freedom with which vehicles leave and re-enter clusters affects jitter performance. Considering a specific jitter value 500µs, it can be observed that our LQRD schemes guarantees close to 80% of side-links exhibiting jitter of less than 500µs, while PRS

could only afford 60% of side-links exhibiting less than 500µs in our urban environment driving scenario. Though CHmem has similar performance as our CQRD scheme, guaranteeing between 60 to 80% of vehicles having jitter less than 500µs, but this jitter performance comes at a throughput cost for CHmem as presented in Figure 4-5.

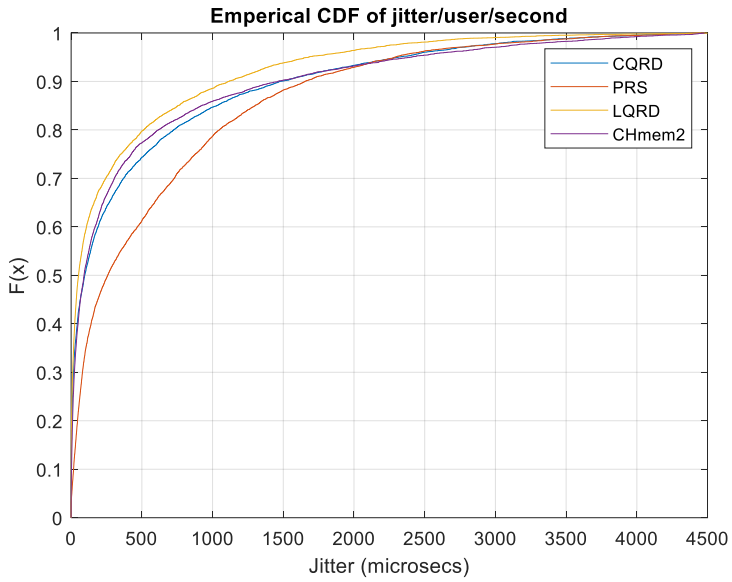

Figure 4-9. Comparing Jitter distribution of re-clustering decision schemes

Our schemes performed better in the rural-highway environment, exhibiting average jitter performance of less than 40µs as compared to jitter of more than 335µs exhibited in urban environment. This is also true for our baseline schemes, however, none of them outperforming our QRD. In fact, both baseline schemes exhibit relatively poorer performances in the rural-highway environment as compared to our QRD schemes as demonstrated in Figure 4-10.

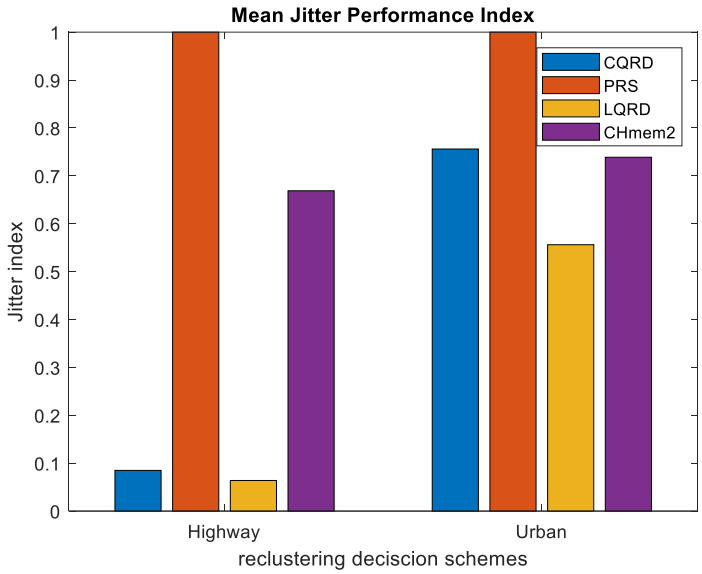

Figure 4-10. Comparison of mean jitter performances of re-clustering schemes in rural-highway and urban environment

In Figure 4-11, reliability of the side-links in the urban environment for all the decision schemes is compared and presented using 3 different stability metrics. The CH stability chart presents the stability of individual CH over travel time, the all-CH stability chart presents the stability of a CH set, while link stability chart presents the stability of a CH-CM side-link over time. The key factors affecting cluster stability includes, CH changes, vehicle speed, the degree membership flexibility, degree of LOS and NLOS changes and the dynamicity road patterns.

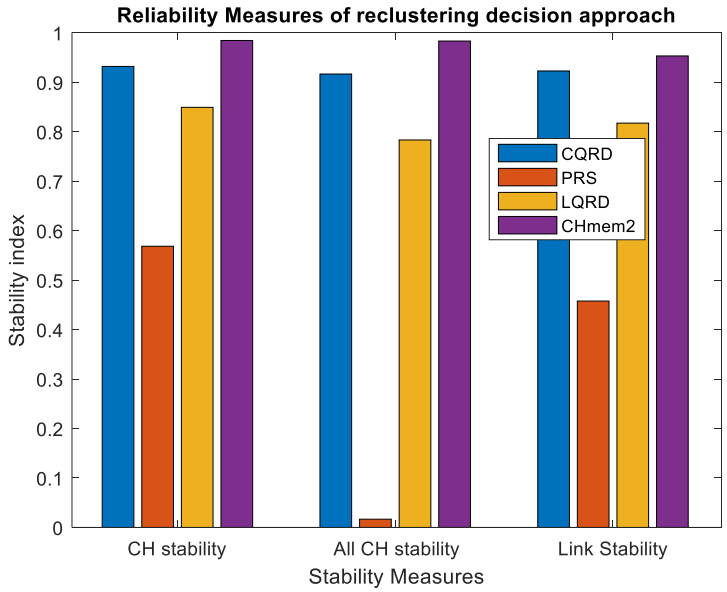

Figure 4-11. Comparing reliability indices for re-clustering decision schemes

For all the reliability metrics evaluated, CHmem was best performing, this is attributed to the low frequency of re-clustering in CHmem, since CH remains unchanged when no re-clustering takes place. The CHmem performance is followed closely by our CQRD scheme, then the LQRD scheme and the finally poorest performing PRS. It can be observed there is significant improvement in stability performance of our QRD schemes over the PRS scheme upon which they are built. Though CH stability is a widely used metric in estimating the performance of clusters, we consider link stability to be more relevant since it reflects the actual steadiness of communication paths. It can be observed that CHmem have a relatively significant low link stability ratio as compared to the ratio observed in other metrics, this is particularly due to the flexibility of cluster membership around its boarders in addition to CH changes. As with jitter performance, our schemes performed relatively better than the baseline schemes in the rural-highway driving environment across the three reliability metrics with a minimum of 0.95 stability index in each of the metrics. CHmem also performs relatively marginally better in the rural-highway environment. We attribute the improved performance of these schemes to the dominance of the effect of LOS consistency and consistent speed which in effect reduces re-clustering frequency and hence CH stability. PRS stability performance in the rural-highway environment remains similar to the performance observed in urban environment.

The QoS index estimate as presented in equation (4-24) is used here to estimate an equal weight index of throughput, jitter, and link stability. Having individually used these metrics to evaluate the performance of the decision schemes, estimating a QoS index that indicates the performance of each scheme depending on the case study of interest. This is done by giving a degree of importance to each QoS parameter or performance metric. Here we have decided to give equal degree of importance to each performance metric. The normalised mean value of throughput and link stability are summed while the normalised mean value of jitter is subtracted, since the former metrics are considered positive, and the latter is considered adverse.

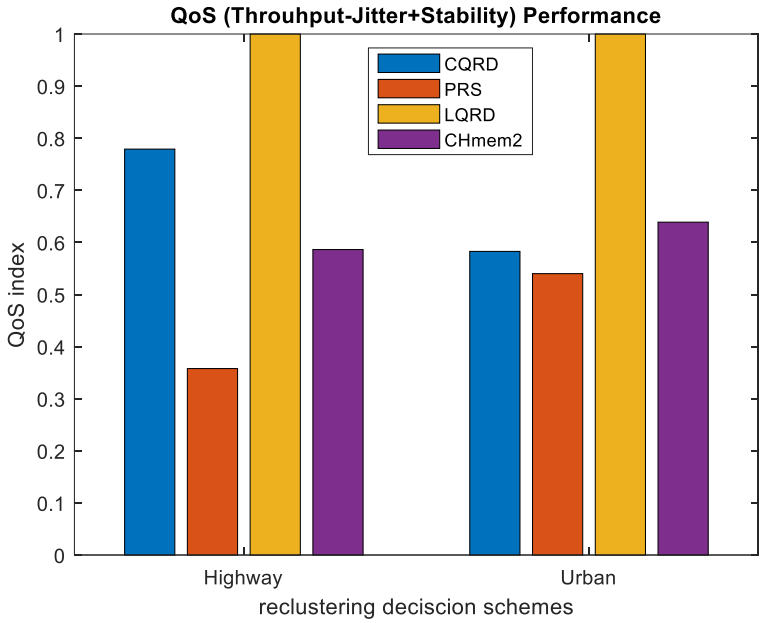

Figure 4-12. Equal weight QoS Index estimate for re-clustering decision schemes in urban environment.

From the bar plot in Figure 4-12, it can be observed that given equal degree of importance of the 3 metrics, the LQRD scheme demonstrates superior performance over CQRD and the baseline schemes. Despite the intensive computation requirement of PRS it appears to perform poorest of all the schemes as it has only excelled in terms of throughput performance. Hence, if throughput weight was to be raised, it has a potential to outperform the CQRD and possibly CHmem in the urban environment. CQRD shows greater promise relative to other decision schemes in the rural-highway environment as compared to the urban environment and would even show greater promise in scenarios where stability is of the essence. However, in urban areas, even when summing weight is skewed towards reliability/stability metrics, CHmem will likely outperform CQRD as it demonstrates greater stability over CQRD.

4.9 Conclusion

In this chapter we have described how we used a threshold-less approach that uses a persistent estimation of cluster and link quality to determine the point-in-time of re-clustering with the aim of achieving a comparable throughput to the persistent re-clustering approach presented in chapter 3 while minimizing jitter, improving stability, and reducing re-clustering overhead.

At the end, a comparable throughput performance to the PRS scheme was achieved by our LQRD scheme while reducing re-clustering overhead by 85%. This in turn resulted in considerably improved stability and jitter performance. LQRD demonstrated this relative superior performance more clearly in the rural-highway environment compared to the urban environment. A superior link stability performance improvement of 70% and 60% in rural-highway and urban environment was established respectively while in terms of jitter a reduction of 95% and 50% was achieved in the rural and urban scenarios respectively. Though LQRD also clearly outperforms the popular CHmem re-clustering decision method in terms of throughput and jitter performances, CHmem still edged it out in all three stability metrics. This is also the case with the performance of LQRD in comparison with CQRD. However, in the overall equal-weight QoS index estimate, LQRD edged all other decision methods in the rural-highway scenario by over 20% margin and in the urban scenario by over 30% margin. CQRD on the other hand outperforms CHmem in terms of throughput in both rural and urban scenarios. In terms of jitter CQRD outperforms both PRS and CHmem in rural-highway environment but was edged out by CHmem in urban scenario. CQRD showed considerable promise in terms of stability performance and only marginally edged out my CHmem in urban and rural scenario. Conclusively, LQRD demonstrates the greatest performance in terms of real time traffic delivery over side-link interface in both urban and rural-highway environments, while CQRD demonstrates the next best performance in the rural-highway environment. Both schemes showed greater promise in rural-highway environment as compared to the urban environment. LQRD will demonstrate greater robustness with varying weight of performance metric as compared to CQRD.

5 Resource Aware Optimal K-value for C-V2X Networks

5.1 Introduction

Our work so far has largely explored improving the performance of V2V links in a hotspot scenario, by exploring how and when re-clustering should be done. Here we have decided to take a birds-eye view of the network and explore the optimal use of both V2V side-link and V2I backhaul resources. In chapter 3 and 4 we approached the selection of number of clusters in two different ways. In chapter 3 we employed the silhouette method while in chapter 4 we used a quantitative variant of the elbow method. These approaches have both considered the proximity between CMs of same cluster, contextually known as within-the-cluster-distance to decide the number of clusters. They attempt to minimize within-the-cluster-distance, which could potentially improve link throughput. However, these methods do not inherently consider the resource available to V2I links on the BS or V2V side-link resources. They have also not considered different cluster size threshold and corresponding Free Vehicles (FV) that will require direct download access from the BS. The consequence of which is a stiff contention for V2I resources between FV and relay CHs, which may lead to the relay link being subjected to lean or insufficient resources which will in-turn become a bottleneck along the downlink path, minimizing the throughput obtainable to side-links despite bandwidth resource availability.

Our approach considers the cluster boundary threshold, number of free vehicles, available resources, and the output of the elbow method in selecting the number of clusters. The goal

is to maximize the use of both V2I and V2V bandwidth resources. We approach the conundrum by first analysing the relationship between the variables considered, then find a solution that minimizes the number of V2I links and maximum cluster size, which in turn maximizes the resources available to both V2I and V2V users. The analysis is first done in the context of the bandwidth resource allocation used in chapter 3 and 4, where separate dedicated resources are allocated for V2I and V2V communication. The results are then compared to our previous method for the number of clusters selection. We then explore 3 more dynamic resource allocation schemes. We compare the results with the conventional schemes and observed a significant throughput performance improvement.

5.2 Communication System Model

A birds-eye view of the V2X necessitates consideration of the communication model of the cluster backhauls in the context of a 2-hop downlink transmission path proposed for download of urgently needed traffic data. To briefly recap our model context as described in sections 3.3 and 3.4 of chapter 3, vehicles are grouped into clusters where the CHs serve as a relay and a download hotspot for the rest of the CMs. The CH-to-CM V2V side-links are modelled in an urban environment as described in chapter 4, with similar interference properties. However, the focus of our work here is the cluster backhaul which is essentially a V2I/N link. The V2I/N cluster backhaul is modelled as an urban environment link between BS and CH vehicles and BS and FV. The pathloss and shadowing model we employed for the backhaul is based on the channel model defined in [152]. The description of our pathloss and shadowing model for urban LOS links between the BS and CHs and FVs is presented in equations (5-1) to (5-4). The LOS pathloss used is dependent on the horizontal distance, d_{2D} between the BS and the vehicular node, VN as depicted in the pictorial representation in Figure 5-1. Whether the pathloss of individual links is a function of break point distance, d_{BP} and node heights (h_{BS} and h_{VN}) as in PL_{los2}in equation (5-3) or not as in PL_{los1} as in equation (5-2) is dependent on the d_{2D} range as described in (5-1). The height used in our model is 1.8m for VN height and 25m for BS height in line with ETSI technical report 138 901 specifications and outlined in the inequality and equation in (5-4)

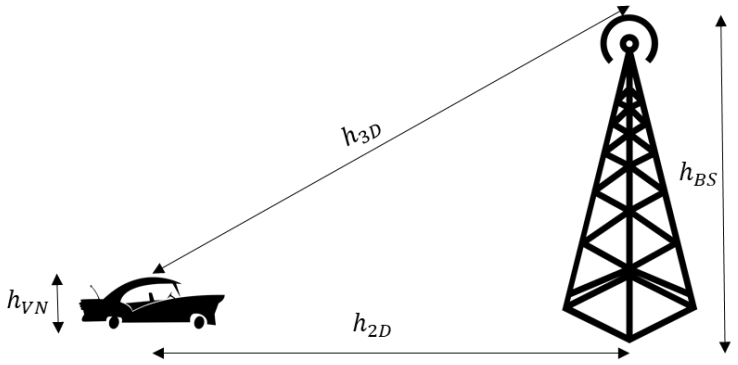

Figure 5-1. Depiction of Backhaul and FV Link Geometric Parameters

$$\begin{cases} PL_{los1} & 10m \le d_{2D} \le d_{BP} \\ PL_{los2} & d_{BP} \le d_{2D} \le 5km \end{cases} \quad (5\text{-}1)$$

$$PL_{los1} = 28.0 + 22\log_{10} d_{3D} + 20\log_{10} f_c + X_\sigma \quad (5\text{-}2)$$

$$PL_{los2} = 28.0 + 40\log_{10} d_{3D} + 20\log_{10} f_c - 9\log_{10}(d_{BP})^2 + (h_{BS} - h_{VN})^2 + X_\sigma \quad (5\text{-}3)$$

$$1.5m \le h_{VN} \le 22.5m, h_{BS} = 25m \quad (5\text{-}4)$$

The break point distance, d_{BP} is a function of effective BS and VN heights (h'_{BS} and h'_{VN}) as expressed in equation (5-5). Where h_{VN} and h_{BS} are actual antenna heights of VN and BS respectively, f_c is carrier frequency and d_{3D} is the 3D distance, d_{3D} between VN and BS as indicated in the depiction in Figure 5-1.

$$d_{BP} = \frac{4\, h'_{BS}\ h'_{VN}\, f_c}{c} \quad (5\text{-}5)$$

The relationship between effective antenna heights and actual antenna heights are described in equations (5-6) and (5-7), where h_E is effective environmental height and in our scenario have set it to 1 meter.

$$h'_{BS} = h_{BS} - h_E \tag{5-6}$$

$$h'_{VN} = h_{VN} - h_E \tag{5-7}$$

The distribution of the shadow fading in this model is log-normal and the standard deviation, σ for the LOS urban scenario is given as 4.

The NLOS link pathloss and shadowing is modelled as described in equations (5-8) and (5-9) in line with the channel models defined in ETSI technical report 138 901. The shadow fading is also log-normal and the standard deviation, σ specified as 6.

$$PL'_{nlos} = 13.54 + 39.08 \log_{10} d_{3D} + 20 \log_{10} f_c - 0.6(h_{VN} - 1.5) + X_\sigma \tag{5-8}$$

$$PL_{nlos} = \max (PL_{los}, PL'_{nlos}) \tag{5-9}$$

$$for\ 10m \leq d_{2D} \leq 5km$$

The pathloss model of the NLOS link, PL'_{nlos} is only adopted when the value exceeds that of LOS, PL_{los} for 2D link distances, d_{2D} ranging from 10m to 5km. It is also noteworthy that the model works for vehicular node antenna heights, h_{VN} between 1.5m and 22.5m. Our VN's are of uniform heights of 1.8m. Base station heights are specified at 25m.

Concerning interference, distinct interference schemes were used for V2I/N links and V2V side-links. For the V2I/N links, we adopted a frequency reuse factor of 1 and hexagonal cell coverage with the BS at the centre of the hexagon. We assume a variation of Fractional Frequency Reuse (FFR) is used by the base station to allocate resources to the V2I/N nodes/links. We have decided to limit our discussion about the type or implementation of the FFR scheme, since the type of scheme adopted has no impact on the downlink interference considered in our scenario. A worst-case scenario of this approach is depicted in Figure 5-2. Where all the six base stations around the serving BS are all transmitting at the same frequency at which the vehicle of reference is receiving from its serving BS.

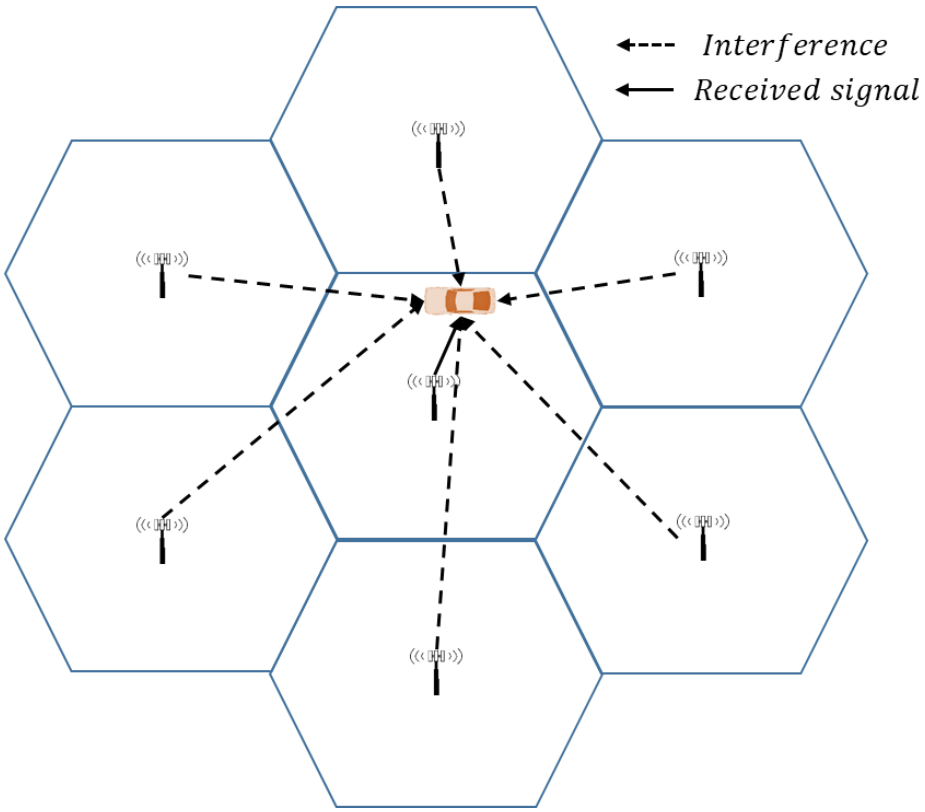

Figure 5-2. A depiction of a case of downlink V2I/N interference

For the V2V side-links the frequency allocation and interference are cluster-based. A typical VN uses a different channel to the channels used by fellow cluster members and same bandwidth channel to the VN's of other clusters that poses least interference. A simple depiction of the interference and resource allocation approach used is depicted in Figure 5-3. The coloured bar at the top represents the entire resource allocated for V2V side-link communication, while each colour represents the equal bandwidth or number of resource blocks allocated to each V2V CH-to-CM side-link. The dotted lines represent interference while the continuous lines represent the received signal link.

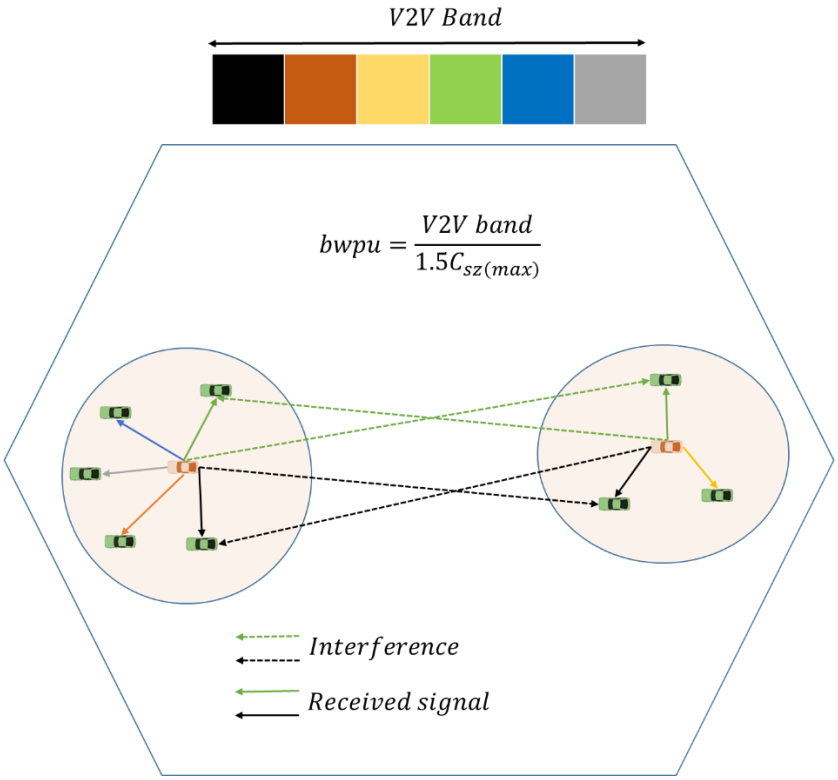

Figure 5-3. A Depiction of interference coordination and resource allocation in for V2V side-link communication

5.3 Cluster Analysis

In an effort to maximize V2I bandwidth, thereby minimizing the potential bottleneck that the V2I could pose across download transmission path, we have sought to understand how vehicular designation and C-V2X bandwidth resources vary with clustering parameters.

In previous chapters the resource assignment assumes a dedicated resource slice for C-V2X, with a further dedicated and distinct bandwidth resource to V2I and V2V links. Our approach to optimizing the use of the V2I bandwidth resource per user link and reducing the cluster backhaul bottleneck is by minimizing the number of V2I user links contending for the resource. Likewise for maximizing the V2V side-link resource per user link, we approach this by minimizing the maximum cluster size at each clustering instance, building on our V2V resource allocation scheme described in Chapter 3, section 3.4. To do this, a study of the relationship between the number of different vehicle components (CH, CM and FV) and cluster parameters such as the number of clusters and cluster distance thresholds needs to be explored.

Our study shows how the number of V2I links, and maximum cluster size varies with different cluster radius threshold and different number of clusters. In Figure 5-4, we present a variation of the average number of clusters across number of clusters against a varying distance threshold between 300m to 1000m.

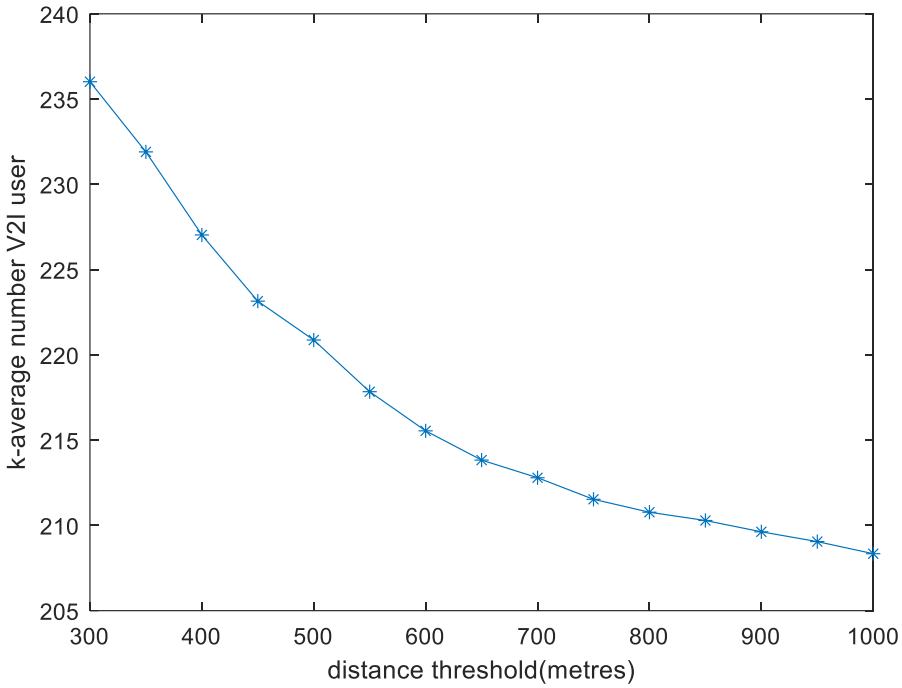

Figure 5-4. Average Number of V2I User Links across Number of Clusters Vs Distance Threshold

The figure indicates that the number of V2I users drops with increasing distance threshold range being considered. With a total variation of 27 users across the threshold range. The minimum number of V2I users is observed at the maximum threshold. The implication of this is that the maximum distance threshold potentially meets the requirement of minimizing the number of users contending for V2I bandwidth resources.

The plot in Figure 5-5 shows the variation of average number of V2I users across distance against the number of clusters. The number of clusters considered ranges across the total number of vehicles, from 1 to 400. It is observed that with just one cluster, the average number of V2I vehicles across the distance threshold is approximately 375, which is essentially the total number of vehicles less the number of cluster members in the cluster. This means we have an average of 25 cluster members in the first clusters across distance threshold and

the total number of free vehicles is around 374, which represents the total number of V2I vehicles less the cluster head of the single cluster.

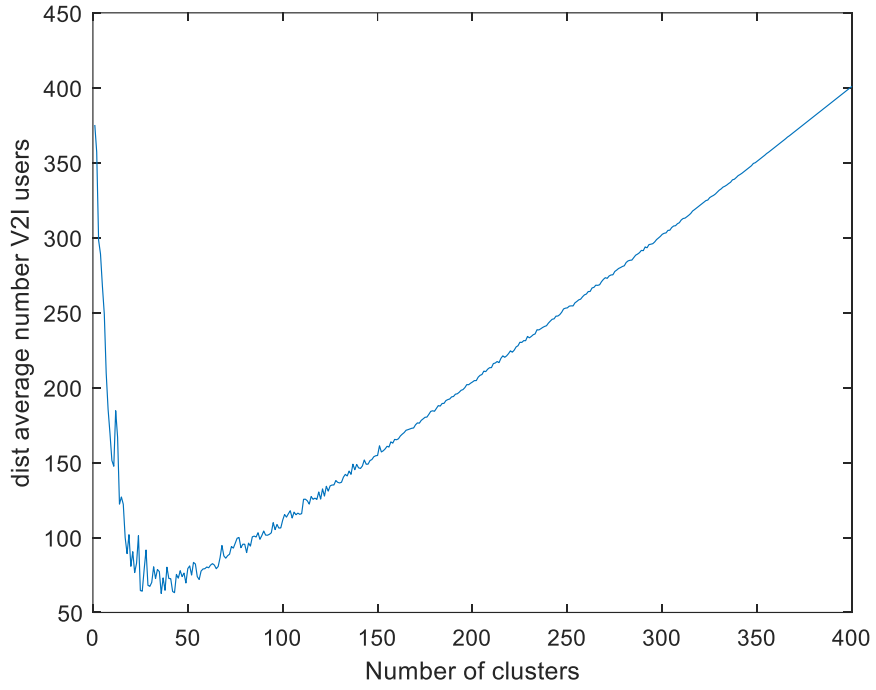

Figure 5-5. Plot of Average Number of V2I User links Across Distance Threshold Vs Number of Clusters

However, as the number of clusters increase, the number of V2I vehicles and links drops until a point is reached where a further increase in the number of clusters increase the number of V2I vehicles or links. This points (number of clusters, number of V2I user links) is reached at approximately (36, 66). From this point onwards there is an almost linear increase I the number of V2I user links with number of clusters, until a point where every individual vehicle is a CH of its own cluster at (400,400).

A 3D-plot showing a comprehensive variation of V2I along distance threshold and number of clusters in presented in Figure 5-6. It also specifically shows the variation of number free vehicles links within the number V2I user links. The number of FV links continues to decrease across increasing number of clusters and increasing distance threshold.

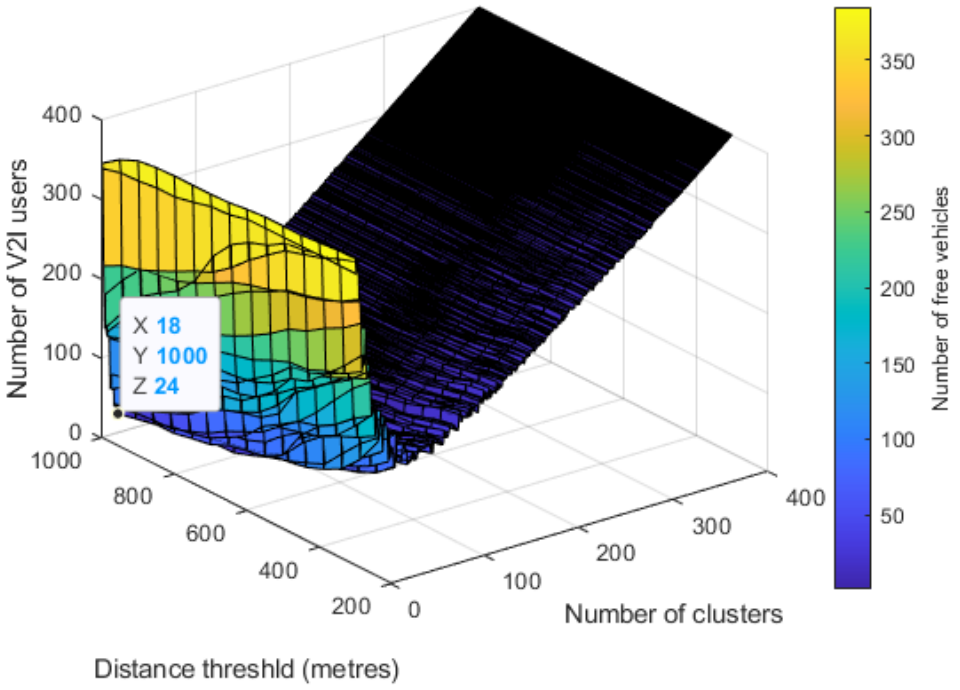

Figure 5-6. Surface plot showing variation of V2I and FV user links across distance threshold and number of clusters.

The minimum number of V2I users is obtained at 18 clusters and 1000m of cluster distance threshold. At this point the number of free vehicles, 6 is the total number of V2I user links, 24 less the number of clusters, 18.

For the CH-to-CM V2V side-links, it is understandable that the number of side-links is the total number of vehicles less the CHs and FVs. But one important parameter in the V2V side-link context is the maximum cluster size, which defines the number of side-links or the number of CMs in the most populated cluster. The importance is particularly related to how resource allocation is done in our V2V resource reuse scheme described in section 3.4. The resources allocated to each CH-to-CM side-link is directly determined by and inversely proportional to maximum cluster size. Figure 5-7 shows how the average maximum cluster size across distance threshold behave in response to changes in the number of clusters. As the number of clusters increases from the minimum, the number of clusters increases until it reaches a point (in this case 8 clusters) where an increase in number of clusters could lead to some clusters splitting. In which case the maximum cluster size starts to drop in general.

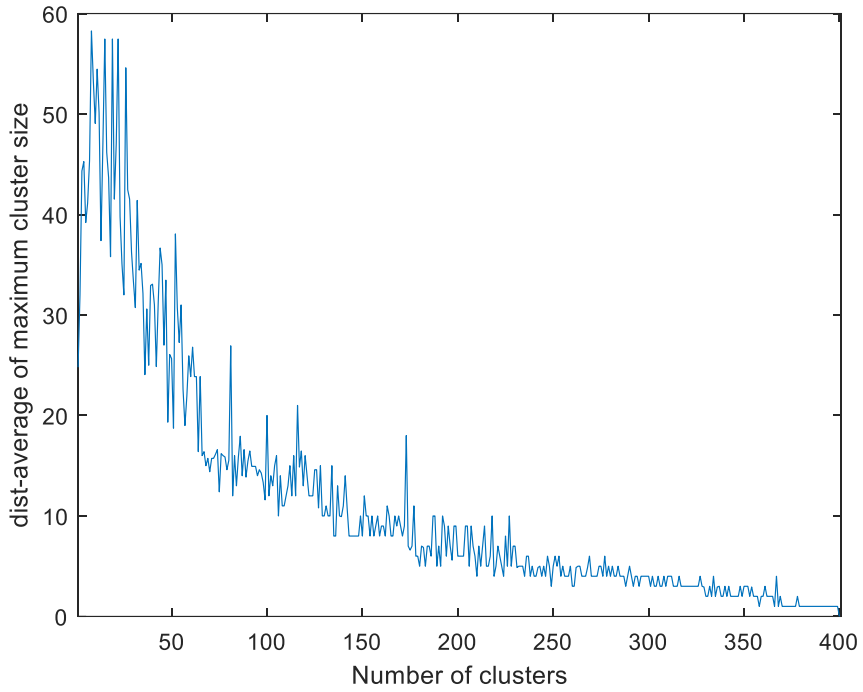

Figure 5-7. Plot of average maximum cluster size across Distance threshold Vs number of clusters

Figure 5-8 on the other hand presents the variation of average maximum cluster size across number of clusters against distance threshold.

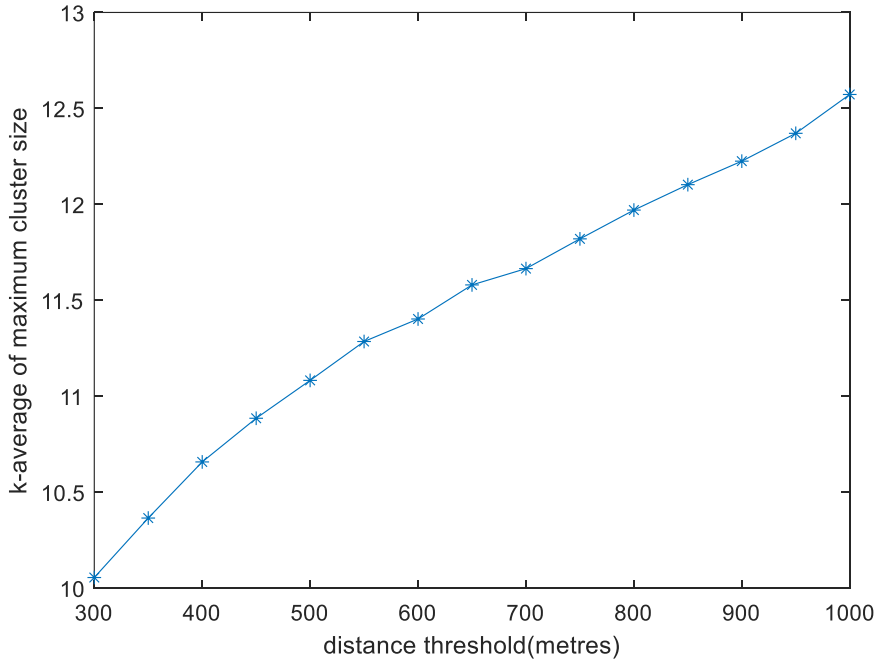

Figure 5-8. Plot of average maximum cluster size across number of clusters vs distance threshold

The plot clearly indicates that the maximum cluster size increases with increasing distance threshold, however compared to the variation across number of clusters, it is observed that the changes in maximum cluster size here is relatively small, with a total variation of less than 3 vehicles as compared to a maximum cluster size variation of approximately 57 vehicles observed in Figure 5-8.

Beyond just understanding the behaviour of the various vehicular node components with changes in distance threshold and number of clusters, the results observed from these analyses have a great impact on the optimizing the use of network resources and mitigating cluster backhaul bottleneck.

5.4 Problem definition

As discussed in sections 5.1 and 5.3, our aim is to maximize the bandwidth available to both V2I and V2V links and potentially minimize the throughput bottleneck along the relay transmission path. Unlike traditional k-selection schemes, we seek to understand the relationship between the k-value, number of un-clustered vehicles and how they affect the bandwidth resources at the disposal of V2I and V2V links.

For each number of clusters, k_x ranging across the entire number of vehicles as described by the set, K in equation (5-10) , the corresponding centroid positions are evaluated using k-means and k-means++.

$$K = \{k_x : 1 \leq x \leq n\} \tag{5-10}$$

For each value of k_x, considering a superset, Z comprising of a set of SNR values, ζ_t. Each set, ζ_t comprises of evaluated SNR values between each vehicle $V_t \in V$ and all centroids, i as defined in equations (5-11), (5-12) and (5-13) and each cluster has a corresponding cluster head, Ch_i as in equation (5-14)

$$Z \sqsupseteq \{\zeta_t : 1 \leq t \leq n\} \tag{5-11}$$

$$V = \{V_t : 1 \leq t \leq n\} \tag{5-12}$$

$$\zeta_t = \{\zeta_i : 1 \leq i \leq k_x\} \tag{5-13}$$

$$C = \{c_i : 1 \leq i \leq k_x\} \tag{5-14}$$

Where n is number of vehicles, V is the set of all vehicles, k_x represents the number of centroids, t represents the index of specific vehicle and i is the index of a specific cluster, CH or centroid.

Also, we consider a specific distance threshold from a set of distance thresholds used to define the radius within which clusters are bounded to be d_{th}. We define our SNR threshold, SNR_{th} as a function of d_{th} as described in Equation (5-15), where the function is based on side-link pathloss, Received Signal Strength (RSS) and noise.

$$f(d_{th}) \rightarrow \zeta_{th} \tag{5-15}$$

Having estimated ζ_{th}, we associate each vehicle to centroids with which the vehicle has maximum SNR, ζ_t and whose ζ_t is below the threshold. For every value of k_x number of centroids, we have cluster identities ranging from 1 to k_t and mapped to each vehicle and saved as a set of vehicle cluster identity, C_L as presented in equation (5-16) with the size of each cluster defined in equation (5-16). The maximum cluster size, C_{max} is identified and the number of free vehicles, F_v is evaluated as presented in equations (5-18) and (5-19).

$$C_L = \underset{1 \leq t \leq n}{\forall \zeta_t} \left\{ \underset{1 \leq i \leq k_x}{argmax} (\zeta_t \geq \zeta_{th}, c_i) \right\} \tag{5-16}$$

$$C_L = \{C_t : 1 \leq t \leq n\}$$

$$C_s = \{C_{si}, 1 \leq i \leq k_x : n(c_i \in C_L)\} \tag{5-17}$$

We then exploit variation in number of CHs, number of FVs and maximum cluster size, C_{max} with the distance threshold and number of clusters to maximize the bandwidth available per V2I and V2V link. Both C_{max} and F_v are estimated as presented in equations (5-18) and (5-19) respectively.

$$C_{max} = max\{C_s\} \tag{5-18}$$

$$F_v = n - \sum_{i=1}^{k_x} C_{si} \tag{5-19}$$

Adopting our existing resource allocation approach, where V2V and V2I links are allocated distinct dedicated frequency band and V2I bands are dedicated and separate from bands used by other BS users, we have decided to approach k-selection in a way that maximizes usage of both V2V and V2I bandwidth resources per link. This approach seeks to keep the bandwidth allocated to V2I and V2V as close as possible with the bandwidth allocated to V2I links greater than the bandwidth allocated to V2V side-links. The optimization problem is defined in equations (5-20) to (5-26), with the multi-objective functions are presented in equations (5-20) and (5-21) , while the constraints are presented in equations (5-23) to (5-26).

$$\frac{B_{v2i}}{k_x + F_v} + \frac{B_{v2v}}{C_{max}} \ \ Maximize \tag{5-20}$$

$$\frac{B_{v2i}}{k_x + F_v} - \frac{B_{v2v}}{C_{max}} \ Minimize \tag{5-21}$$

The first objective function presented in the optimization expression in (5-20) seeks to maximize the combined bandwidth per V2I and V2V link, consequently seeking to reach a compromise between the number of V2I links and cluster size. While maximizing the bandwidth per user link, the second objective function presented in the optimization expression in (5-21) seeks to minimize the difference between V2I and V2V bandwidth per user link. The aim is to prevent excessive skewing of bandwidth towards V2I, which could in turn portend redundant throughput at the backhaul.

We then combine the objective functions to a single super objective function which when maximized, its optimal solution is used to find the maximum combine V2I and V2V bandwidth per user-link. The super objective function is expressed in (5-22).

$$\frac{B_{v2i}}{k_x + F_v} + \frac{B_{v2v}}{C_{max}} - \left(\frac{B_{v2i}}{k_x + F_v} - \frac{B_{v2v}}{C_{max}}\right) \; Maximize \tag{5-22}$$

The objective functions are constrained by the conditions expressed in the inequalities between (5-23) and (5-26). The first inequality presented in (5-23) limits the V2V and V2I optimal bandwidth pair to a pair that where V2I bandwidth is greater than V2V bandwidth. The reason for this is to guarantee some performance reliability for CH's V2I links that shoulders relaying responsibility. A performance issue for CH V2I links has a multiplier effect on CMs. The constraint in (5-24) limits the k-selection solution to number of clusters that satisfies the condition that V2V bandwidth per user link can only be as big as V2I link bandwidth per user. This is useful to keep the number of clusters within the range that sustains the proximity advantage defined by the elbow method. The inequality in (5-25) and (5-26) constrains the objective functions to values where k_x and C_{max} is greater than 1 and to values where $B_{v2i}, B_{v2v}, k_{elb}$ and F_v is non-zero. This is to exclude extremities from solution options.

$$\frac{B_{v2i}}{k_x + F_v} \geq \frac{B_{v2v}}{C_{max}} \tag{5-23}$$

$$k_{elb} + 2 \geq \; k_x \geq k_{elb} \tag{5-24}$$

$$k_x, C_{max} > 1 \tag{5-25}$$

$$B_{v2i}, B_{v2v}, k_{elb}, F_v > 0 \tag{5-26}$$

Here B_{v2i} is the total bandwidth resource allocated for V2I communication, B_{v2v} is the total bandwidth resource allocated for V2V communication and k_{elb} is the optimal k-value as estimated using the quantitative elbow method described in chapter 4.

5.5 Alternate Bandwidth Resource Allocation

The previous chapters and our problem definition in section 5.4 have assumed dedicated bandwidth resource allocation between V2I/N links and V2V side-links. In this approach, the

fixed set of resource slices allocated for V2X communication is split into 2 equal resource set with each resource set allotted and dedicated to V2I/N links and V2V side-links respectively.

Beyond the dedicated resource allocation, we now introduce a shared resource approach where the entire resource slice allocated for V2X communication is equally shared amongst all types of nodes regardless of function and communication mode. Unlike the dedicated allocation approach the optimization is not a multi-objective function since the bandwidth is not portioned. The objective function for the shared resource allocation is presented in (5-27)

$$\frac{B_{v2x}}{k_x + F_v + C_{max}} \quad Maximize \tag{5-27}$$

Again, unlike the dedicated approach, there are fewer constraints. The constraint in (5-23) is irrelevant since there is no resource partition. However, other constraints in the dedicated approach still holds. Also, there is no need for multiple objective functions since all bandwidth per user-link are equal irrespective of user designation.

5.6 Performance Evaluation

Having analysed the clustering process and how the various vehicular node components of the clustering process vary with number of clusters and cluster distance threshold in section 5.3, we will now exploit this analysis to evaluate and compare the performances of the two resource allocation approaches described in section 5.5, in terms the of V2I backhaul bandwidth per user-link across distance threshold and number of clusters, to see which approach offers better mitigation for backhaul bottleneck. We then exploit the optimization approach to observe the performances of the resource allocation approaches while minimizing bandwidth gap between the V2V side-links and V2I/N links. We then exploit this analysis to identify the number of clusters and distance threshold that optimizes the bandwidth per V2I/N backhaul link for both the dedicated and shared bandwidth resource allocation approach.

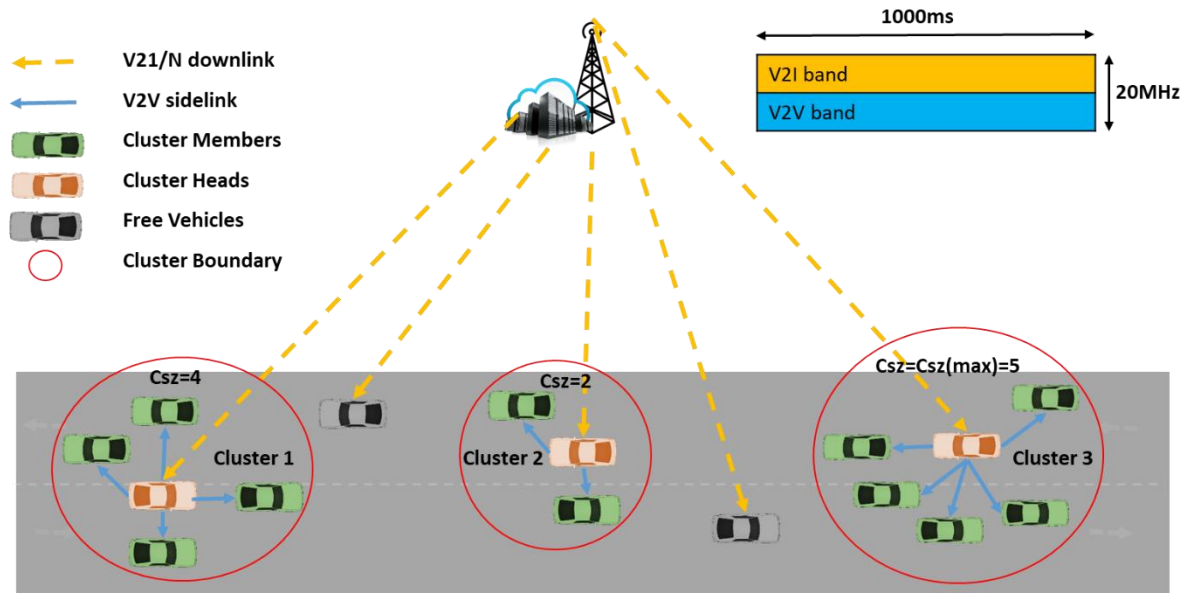

Figure 5-9. A schematic of the network topology upon which the resource-aware algorithm is proposed.

The network topology depicted in Figure 5-9 describes the network topology to which our resource-aware k-selection algorithm is proposed. The yellow link lines labelled V2I/N indicates the backhaul link from the CH to the BS. The red cars are the CHs while the green cars are the CM with a blue link line representing the CH-to-CM side-links. The grey cars are the FVs. On the right top corner, we depicted the resource available to all the V2X links. This is demonstrative of both the shared and dedicated resource allocation schemes.

The simulation assumptions for the network model here are set in according to 3GPP TR 138 901 [72], [152]. Table 5-1 shows the simulation parameters used. In our simulation, we made available a bandwidth of 20MHz available for both V2I and V2V users at 5.9GHz carrier frequency. The driving environment that is envisaged is an urban environment, and this is because we believe that the need for the optimisation of V2I bandwidth usage is greater in the urban environment, where we believe the chances of V2I NLOS is slightly higher due to building and structures such as bridges.

Table 5-1. Simulation parameters for V2I and V2V in the backhaul scenario

Parameters	Values
Available V2X Bandwidth	20MHz
Carrier frequency	5.9GHz
Maximum transmission range	30Om to 1000m, interval: 100m
BS transmission height	25m
Vehicle height	1.8m
BS type	Macro-sites (UMa)
Maximum Transmit power	49dBm
Noise Figure	9dBm
Shadowing distribution (V2V)	Log-normal $\sigma = 3dB$
Shadowing distribution (V2I) LOS	Log-normal $\sigma = 4dB$
Shadowing distribution (V2I) NLOS	Log-normal $\sigma = 6dB$
Range of vehicle speed	[40,120]km/h
Urban V2I model	ETSI TR 138 901/103 257 [152]
Urban V2V Model	ETSI TR 103 257 [72]

The different approaches

5.7 Result and Discussion

This section presents results of our study of how the bandwidth per user-links vary with cluster distance threshold and number of clusters. We also present here the result showing the optimal solution of distance threshold and number of clusters that maximizes backhaul user bandwidth and minimizes backhaul bottleneck.

The result in Figure 5-10 indicates that the alternative shared resource allocation approach shows superior bandwidth per user-link across varying number of clusters compared to the dedicated bandwidth approach. The difference in BW/user-link magnitude between the two approaches increases with the number of clusters. This is expected because in the shared approach the bandwidth proportion accessible to V2I/N user-links increases with number of V2I/N user-links while in the dedicated approach these resources are fixed.

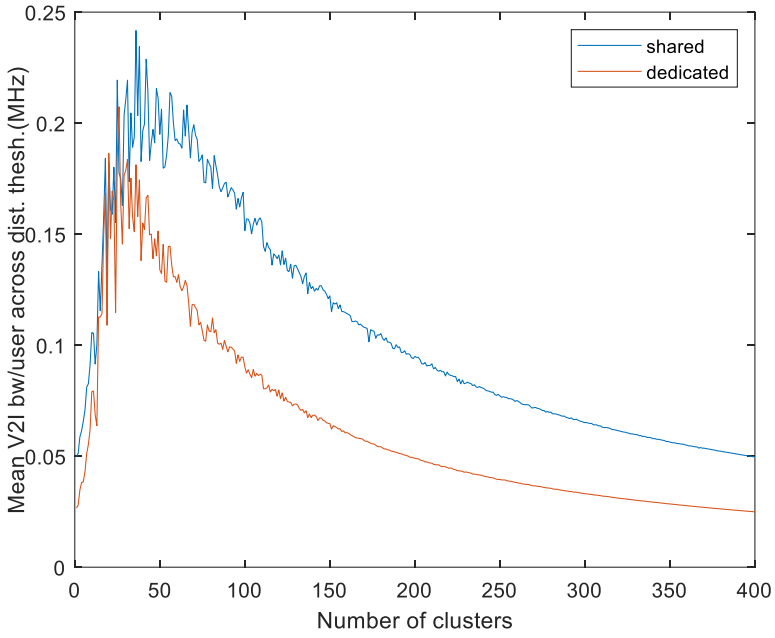

Figure 5-10.Mean bandwidth per user across distance threshold vs number of clusters, K

The peaks observed in each approach varies in magnitude and at what number of clusters it occurs. However it can be observed from the plot showing the variation of V2I/N users with number of clusters in Figure 5-5 that these peak points occur at the troughs of the plot in Figure 5-5, where we have the minimum number of V2I/N users.

Though incremental in absolute value, it can be observed that the average bandwidth per user across the number of clusters increases with distance threshold as established in Figure 5-11, this is in agreement with the observation in Figure 5-4, where the average number of V2I users across number of clusters increase with increasing distance threshold. Hence, it is safe to assume that for all the distance considered (300:100:1000), a distance threshold of 1km offers the maximum bandwidth per user.

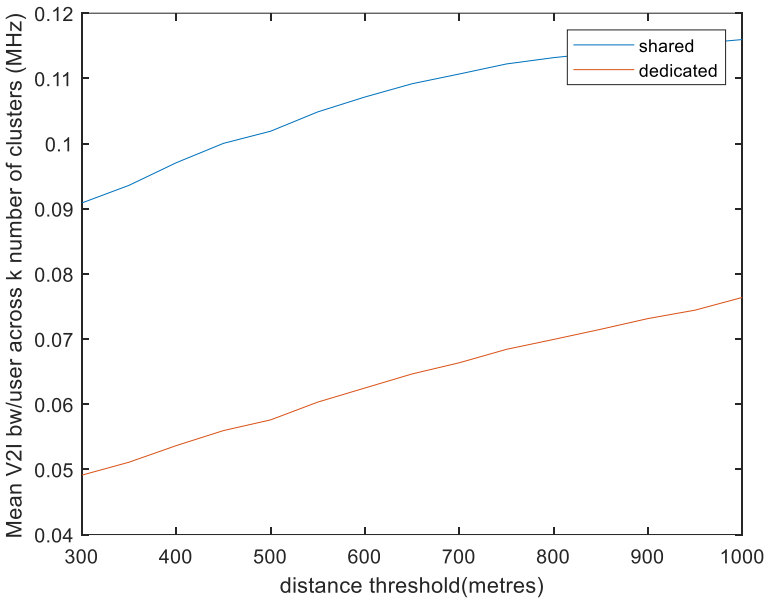

Figure 5-11. A plot showing the variation of mean bandwidth per user across number of clusters with distance threshold.

It can again be observed that the shared approach showed clear superiority with more than 70% performance improvement as compared to the dedicated bandwidth approach. This performance is consistent across distance thresholds.

Having observed V2I bandwidth per user-link performance across the number of clusters and distance threshold, it is also important to understand the bandwidth per user-link performance along V2V path using both dedicated and shared resource allocation schemes.

The result presented in Figure 5-12 shows how the mean bandwidth per V2V user-link across distance thresholds varies with the number of clusters. For the dedicated resource allocation approach, the bandwidth per user generally increases with increasing number of clusters. This is because the maximum cluster size, C_{max} varies inversely with number of clusters and the bandwidth per V2V user link as described in equation (5-20). On the other hand, the shared approach, the BW/user for the V2V path takes the similar shape as the V2I BW/user curve. This is because the allocated bandwidth is not independent but affected by the number of V2I/N users. It is also observed that it reaches its peak at the point where V2V users are at minimum value and due to the approach of allocation both V2V and V2I users for the shared resource approach are allocated equal bandwidth per user-link.

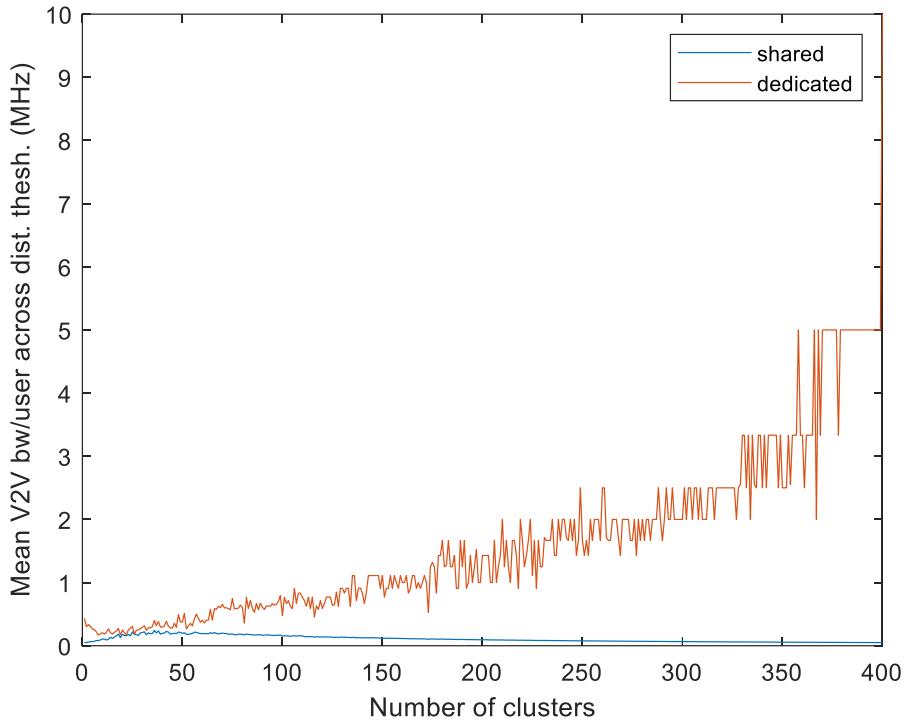

Figure 5-12. Plot showing mean V2V bandwidth per user link across distance threshold against number of clusters

The plots in Figure 5-13 shows mean V2V BW/user-link performance across number of clusters against threshold distance. In the shared approach, it is observed that there is only an incremental increase in BW/user-link with distance threshold as observed in Figure 5-13 with the V2I/N user-links. However, with the dedicated resource allocation approach the BW/user-link drops with increasing distance threshold, which is a consequence of increasing maximum cluster sizes, C_{max} with increasing distance threshold. However, it can be observed that the distance threshold has little influence on the BW/user-link, particularly in the dedicated resource allocation method. Again, it is also observed that the V2V bandwidth experienced in dedicated resource allocation scheme is significantly larger than in the shared allocation approach.

Recall, however, that for the dedicated resource allocation scheme, the cluster backhaul BW/user-link is significantly smaller than the V2V BW/user-link. This situation portends a potential throughput bottleneck at the cluster backhaul in the uplink and throughput redundancy in the downlink.

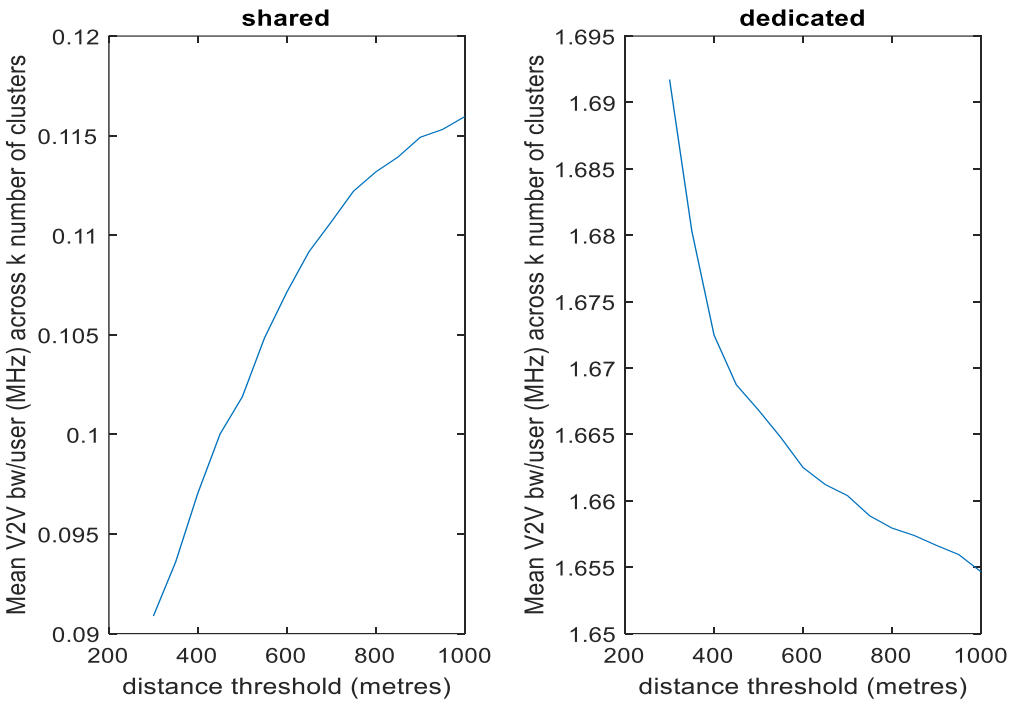

Figure 5-13. Variation of mean V2V bandwidth per user variation across number of clusters against distances threshold

For simplicity and to minimize computation load, we believe we can minimize the effect of the redundancy conundrum between the cluster backhaul links and the CH-to-CM V2V links by selecting an optimal number of clusters and distance threshold solution that offers the maximum V2I/N backhaul and V2V bandwidth sum for which V2I/N backhaul has equal or minimum bandwidth edge over V2V. Our belief is based upon reasonable heuristic assumption after considering for key factors upon which throughput is dependent as described in Table 5-2: interference, link blockage, link length and intra-cell reuse factor.

Table 5-2. Table showing throughput bottleneck contributing factors.

Factors (ς)	BS-CH (V2I)	CH-CM (V2V)
Link blockage	LOS	NLOS
Interference	Moderate	High
Re-use (bandwidth)	None	Yes
Link Length	Long	Short
$\boldsymbol{\rho_i = \psi/n(\varsigma)}$	$\boldsymbol{2/4 = 0.5}$	$\boldsymbol{2/4 = 0.5}$

Our heuristics assumes each factor, ς listed in Table 5-2 have equal weight of contribution to SINR. The green cells in the table indicates a positive contribution to SINR while the red cells represent a negative contribution to SINR. Where $\boldsymbol{\rho_i}$ is described as total impact ratio and ψ is called link positive contribution index. Based on the impact ratio, we assume an equal SINR impact on throughput. Hence, we will henceforth focus on optimizing bandwidth along the entire signal path.

The result displayed in Figure 5-14 depicts the optimal solution $(d_{th} = 1000m, k = 18)$ derived from using the dedicated resource allocation scheme. It demonstrates the maximum V2I+V2V BW/user-link value satisfying the objective function defined in equations (5-20) and (5-21), and constraints defined in the inequalities between (5-23) and (5-26) inclusive. The optimal value 0.65MHz is reached with 18 clusters and at 1000 meters. It can be seen that the elbow method k-selection approach that suggests 13 clusters at 1000 meters does not fall within our optimization constraints as the V2V link is allocated greater bandwidth per user link (0.21MHz) as compared to V2I/N backhaul bandwidth per user-link at (0.10MHz). This portends potential bottleneck at the cluster backhaul which will affect all CMs within the cluster.

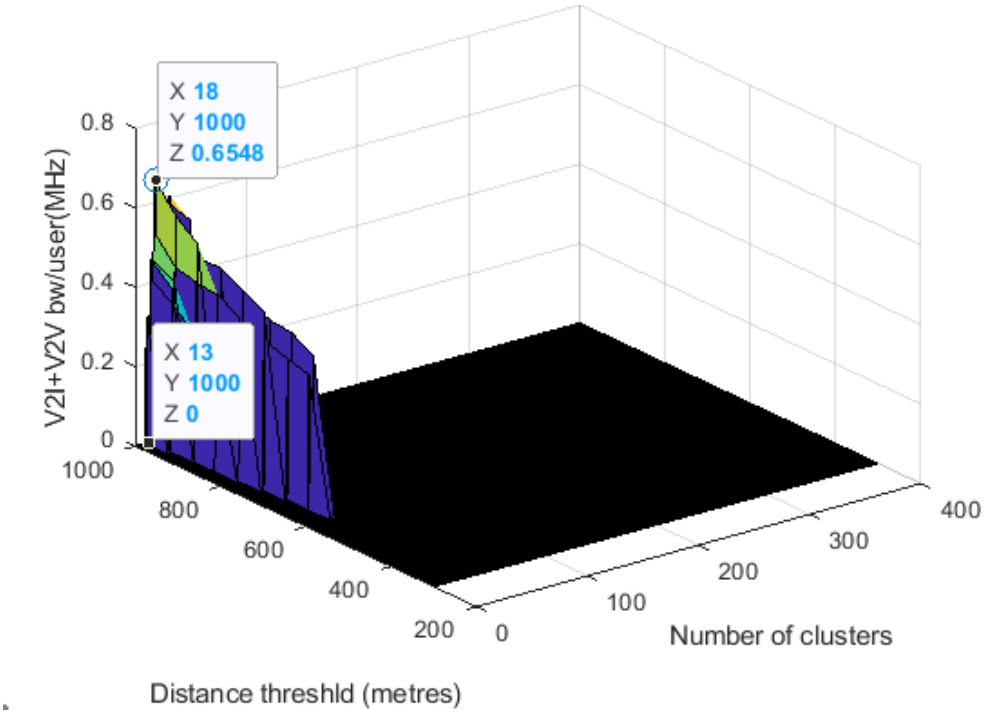

Figure 5-14. A 3D Visualisation of the Optimal Solution of Dedicated Resource Allocation Scheme

For the shared resource allocation method, the results obtained as displayed in Figure 5-15 shows that all possible solutions meet the constraints we have defined. However, the optimal solution was estimated to be at 36 clusters and 1000 meters with a bandwidth per user-link value of 0.65MHz. In this case our silhouette-based elbow solution is apparently within the defined constraints but not the optimal solution.

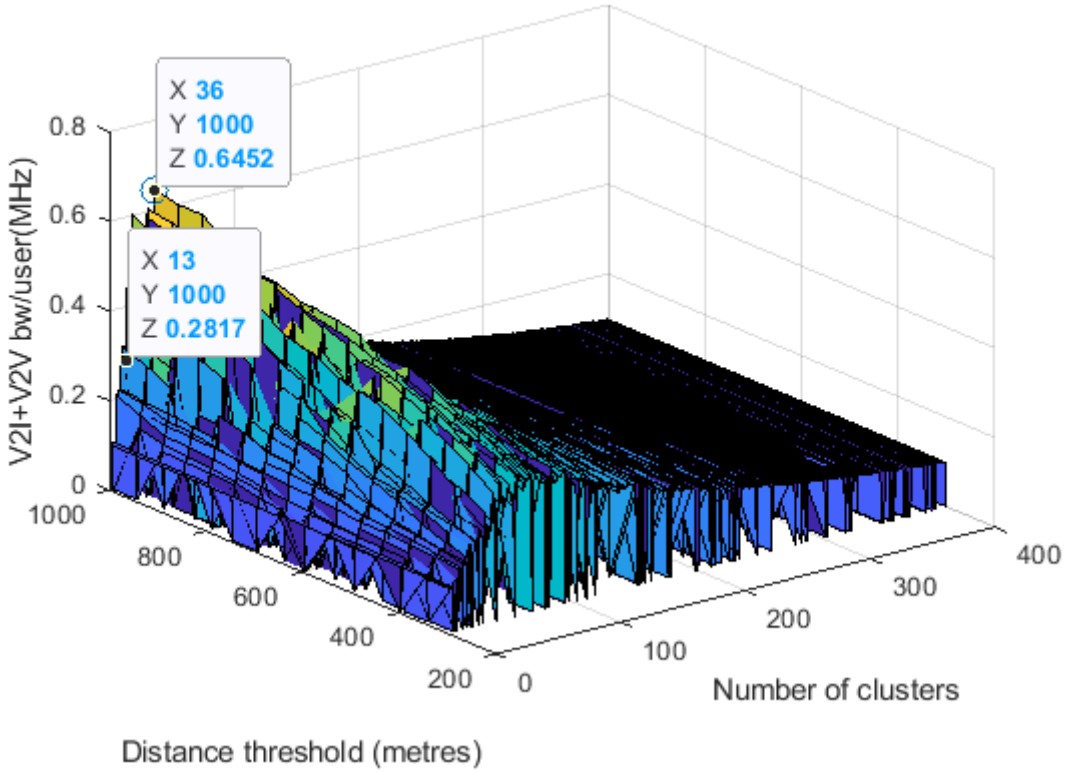

Figure 5-15. A 3D visualisation of the optimal solution of shared resource allocation scheme

It is observed that the dedicated approach seem to have a higher optimal bandwidth per user-link value as compared to the shared approach, however, it is further observed that the difference between V2I and V2V in the dedicated approach is quite significant as indicated in the bar chart presented in Figure 5-16, which portends a potential throughput bottleneck along the downlink path. Unlike the shared approach where both the V2I/N and V2V bandwidth are equal.

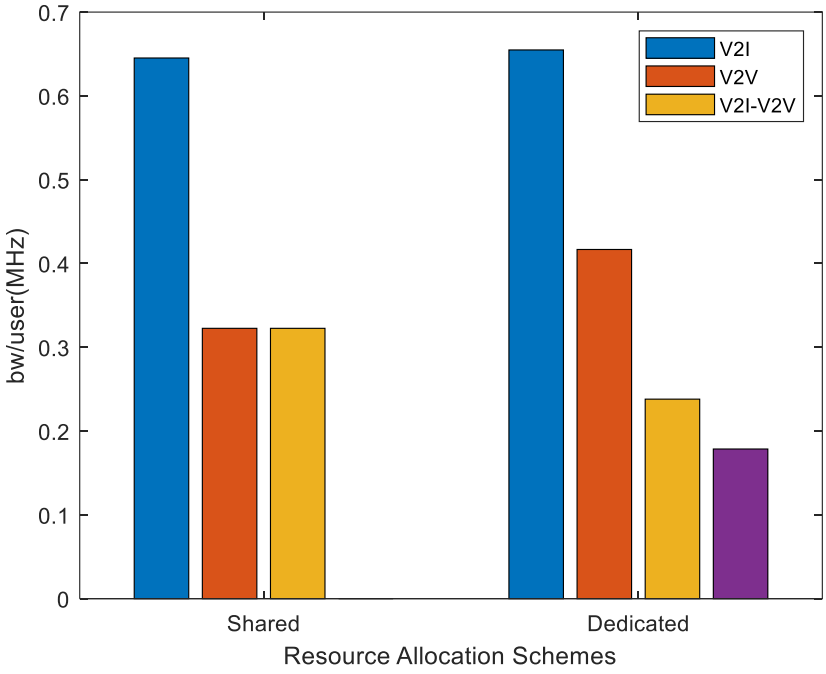

Figure 5-16. Optimum BW/user indicators for both shared and dedicated resource allocation schemes

Having considered an isolated scenario of 400 vehicles. We took a step further to observe how well the k-selection scheme performs against the silhouette elbow method across different vehicle densities for the shared and dedicated resource allocation types.

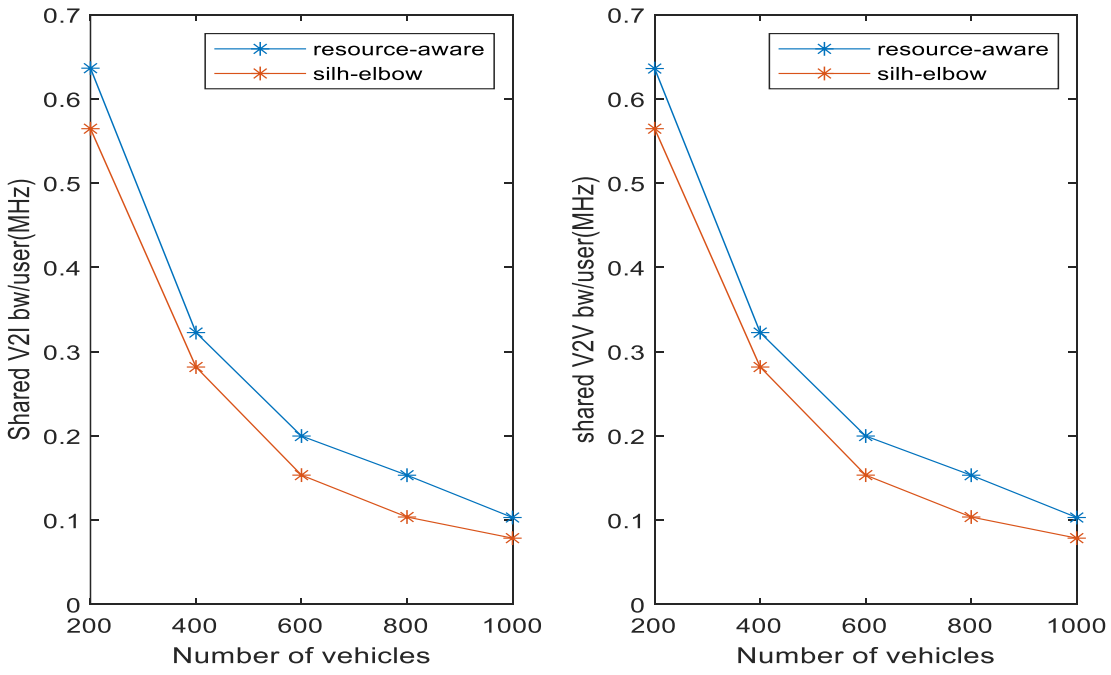

Figure 5-17. Performance of k-selection schemes in shared resource allocation method

The result in Figure 5-17 shows the V2V and V2I bandwidth performance of both the resource-aware and silhouette-based elbow scheme in the context of the shared resource allocation scheme. For the V2I links, we observed that our resource-aware scheme clearly outperforms the baseline. For both schemes the performance drops with increasing vehicular density, given a fixed overall bandwidth size available for access. The same pattern is observed for the V2V bandwidth per user performance.

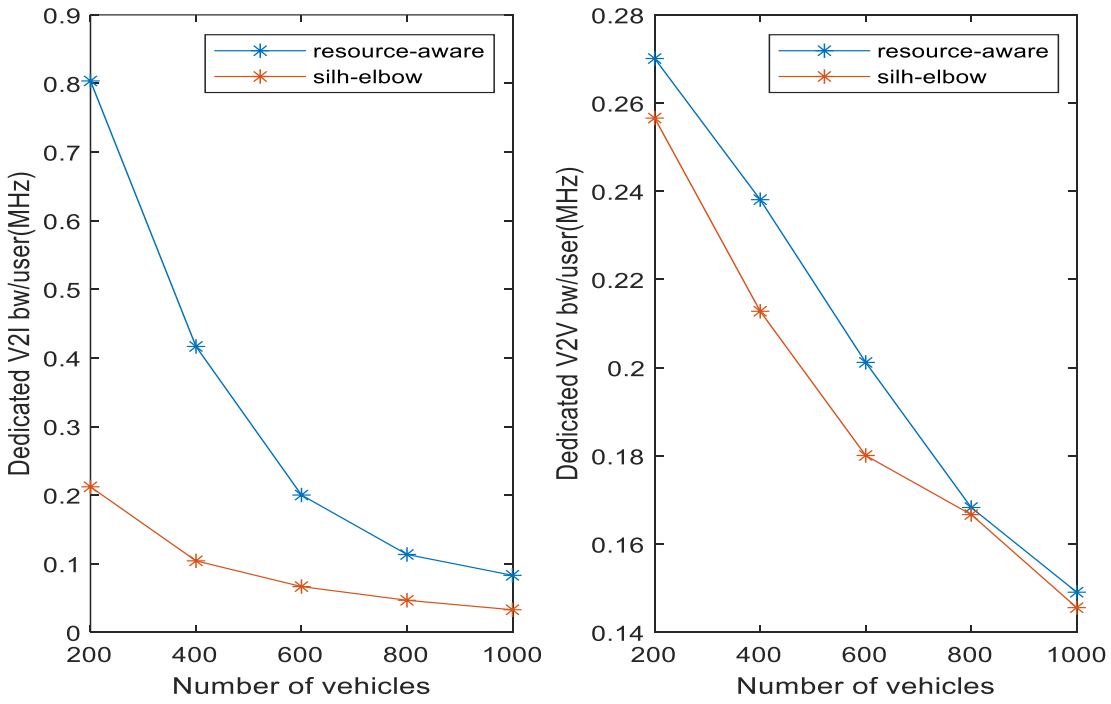

Figure 5-18. Performance of k-selection schemes in dedicated resource allocation method

The result in Figure 5-18 shows the bandwidth per user performance of our resource aware scheme in comparison with the elbow method in the context of the dedicated resource allocation scheme and for both V2I and V2V link-users. The results here also showed our resource-aware approach again outperforming the elbow method across the different number of vehicles considered for both V2I and V2V. We observed for both V2I and V2V and in both schemes that bandwidth per user drops with increasing number of vehicles. This is expected, as the demand for bandwidth becomes more crunch with increasing number of users and user-links accordingly. For V2I, we observed an initial sharp drop in the bandwidth

per user for our resource-aware schemes that slows down with increasing number of vehicles. For the V2V we noticed that both schemes' bandwidth per user dropped with increasing number of vehicles, however the drop percentage is relatively small as compared to V2I. This is because for V2V the factor responsible for the change is a change in maximum cluster size which is limited by the cluster radius distance threshold.

5.8 Conclusion

In this chapter we described our development of a bandwidth-resource-aware method of selecting the optimum number of clusters and distance threshold in a clustered vehicular network. In the previous chapters we have used the silhouette method and a silhouette-based elbow method which essentially is resource and case-study agnostic. Here we developed a mechanism that takes into consideration the case study of context (hotspot-based real-time traffic data download) and the resources at the disposal of the V2X network to carry out the download. The aim is specifically to minimize V2I/N-to-V2V download throughput bottleneck and redundancy.

We approached this by designing a V2I communication model in addition to existing our V2V model to simulate the backhaul path of the cluster. Unlike most studies [20], [26], [153], [154], where outliers are generally not considered in cluster analysis, we followed the model development with a detailed study of the variation of the size of cluster vehicular components with number of clusters and cluster radius distance threshold. Having got an understanding of the behaviour of the various vehicular components (CH, CM and FV), we then developed an optimization problem and used the greedy search approach to identify the number of clusters and cluster radius distance threshold that maximizes resource usage with minimal biases to V2I/N user-links. The resource-aware k-selection scheme was put to test in both a dedicated resource allocation and shared resource allocation scheme.

The results observed clearly demonstrates a better utilization of available bandwidth using our optimized resource-aware k-selectin scheme in both the dedicated and shared resource allocation methods scenario and in repetitive demonstrations across different vehicle population densities. The implication of this is that the cluster backhaul bandwidth resources

can be optimized to minimize the potential bottleneck or throughput redundancy in either the backhaul or side-link layer of the network.

6 Conclusion and Future Work

6.1 Conclusion

This section gives a concise summary of the entire work in this thesis and explanation of the transition across different chapters. It summarises the results and their respective implication, highlights the contribution to knowledge and the limitations therein. This section also revisits our main hypothesis, explaining how our work supports the claim.

6.1.1 Work summary.

This section presents the summary of every chapter presented in this thesis so far. It gives a concise presentation of the essence of the work done, the studies that supports the framework, methods used, and results obtained.

In chapter 1 of this thesis we gave a broad introduction into the world of V2X communication highlighting its importance for safety and traffic efficiency. Though traffic efficiency stands at the core of the real-world motivation for this research, the work has significant implication in emerging safety use-cases. The Tesla autopilot incidence of 2011 shows that driving automation is not enough for safety but improved coordination between all road entities is needed to meet expected safety goals. Beyond applicability, the research motivation for this work stems from the need to solve the stability, reliability and performance issues in vehicular networks resulting from the high mobility and changing network topology of vehicular nodes. The entire work in this thesis was done in the context of the different use-cases that requires real-time download of urgent traffic information. Hence, we looked in the direction of clustering algorithm, which has proved to be an efficient method to manage topology changes and minimize cell resource clogging. However, we reckoned that the most stability enhancing clustering algorithms that have been developed adversely affect throughput performance or at least is without consideration of real time QoS parameters like throughput and jitter.

Hence, this thesis presents approaches to improve stability without impairing QoS parameters such as throughput and jitter performance. Also, unlike most previous work that have approached stability improvement by different CH selection approach, our approach has explored cluster maintenance and re-clustering methods. We explored how and when should re-clustering be initiated and later took a bird-view of the network to see how improved V2V performance could be hindered by the V2I backhaul and took a pre-emptive approach to optimize V2I bandwidth resource.

Having presented a background to the thesis, chapter 2 presents a comprehensive background to the theories, concepts, technologies, and research works upon which the framework of this research has been built. We reviewed V2X RATs, uses cases, propagation models, radio resource allocation, vehicular mobility model and clustering approaches. After a comprehensive review, cellular-V2X is adopted due to maturity of its foundational technologies and network coordination capacity. Realtime geo-cast/multicast use-case for download of enhanced route guidance, traffic and navigation information has been adopted as our use-case of choice due to the great safety and efficiency potential it holds. The series of work presented in this thesis are tested in both urban and rural-highway context. The propagation model we used in the urban and rural driving environments is the standardised geometry-based stochastic model, a decision based on the ease of use and simplicity and the level of abstraction necessary for observing network parameters of interest. A review of V2X resource allocation schemes informs the exploitation of mode-3 underlay intra-cellular cluster-based reuse approach for resource allocation. A microscopic simulation-based modelling approach was adopted for our vehicular traffic. This approach accurately represent the low-level vehicular traffic attributes that is required to represent our desired model effectively and reasonably. We opted for SUMO traffic simulation tool and SUMO web-wizard to easily and cost efficiency model diverse scenarios with reasonable accuracy.

Building upon the theoretical framework developed in chapter 2, chapter 3, presents an approach that answers the questions of how stability can be improved without compromising on V2V link performance. We developed a k-means based SNR (KmSNR) clustering scheme upon which we built two different memory-based re-clustering schemes in which clustering phases in time are cascaded. The aim of the SNR based cluster formation approach as

opposed to the distance-based approach is to improve throughput performance. The memory-based approach was developed to minimize the adverse stability effect associated with the initial random selection of seeds by cutting out the seed selection in subsequent re-clustering phases and smoothening cluster reformation process. The two memory-based schemes CWS and SWS is based on feeding the current CH coordinates into different stages of our clustering process, specifically as centroids and seeds respectively. The performance of these schemes is tested in a highway scenario, where we compared the SNR based scheme with a recent Floyd-Warshall clustering algorithm for stability improvement in VANETs. Our SNR based schemes showed a better throughput performance but marginal poorer stability performance. However, after the implementation of the memory based schemes, we observed a significant stability improvement without any consequence on throughput performance. A significant jitter performance improvement was also observed. It should be noted that all the schemes discussed here is based on persistent re-clustering.

The frequency of the re-clustering techniques presented in chapter 3 is time-based and not based on specific performance objective. Also, from the discussion on previous re-clustering decision approaches in chapter 2, it was deduced that a common approach to CH election and/or re-clustering decision is based on CH membership status and not performance. Building on this, the work presented in chapter 4, seeks to answer the question of when to initiate re-clustering events to improve network performance or at least minimize overhead. Two non-threshold schemes (CQRD and LQRD) were developed. The schemes initiate re-clustering by comparing current cluster topology against a logically re-clustered topology based on performance criteria such as cluster quality and link quality (throughput) respectively. The results observed indicate a superior throughput performance for both CQRD and LQRD over the common CH membership approach, with LQRD marginally outperforming PRS despite a considerably lower re-clustering instances across time as compared to the persistent time-based approach (PRS). LQRD also demonstrated better jitter performance over all other schemes with PRS showing the poorest jitter performance. In terms of stability though, the CH membership approach still performed better as compared to our schemes. We developed a QoS index, giving equal importance to throughput, jitter and stability and found LQRD to have outperformed all other schemes with CQRD coming second in the rural-highway scenario and third in the urban scenario.

The work presented in chapter 3 and 4, we have worked solely on CH-to-CM V2V side-links and have demonstrated that some throughput performance improvement therein. In chapter 5 we decided to have an aerial look of the clustered network to investigate the cluster backhaul and the resources available to it. We decided to develop a resource-aware k-selection method that minimises the number of vehicular nodes requiring access to the BS, thereby increasing the bandwidth available to each V2I user. We started by analysing the relationship between the number of the different vehicular node types (CH, FV, and CM) across number of clusters and distance threshold variations. We then apply this analysis in two different resource allocation schemes. The first one is an allocation approach where a fixed set of equal bandwidth resources are allocated to V2I and V2V vehicular users and the second approach is one in which equal resources are allocated per users regardless of vehicular node type. WE named them the dedicated and shared approach respectively. We observed an optimal number of clusters, k and distance threshold values at the point where bandwidth per V2I+V2V users are maximum and the difference between V2I and V2V users are minimum for both allocation schemes and compared the results to our elbow method of selecting the number of users. In both schemes we observed that our approach realised a better bandwidth per V2I+V2V user-link value and minimum bandwidth difference between V2V and V2I user-links as compared to the elbow method values. We can infer from this that our approach portends lower bandwidth induced throughput bottleneck and redundancy in the downlink.

6.1.2 Revisiting our Hypothesis and Objectives

In this section, we will be revisiting our hypothesis to establish if our work and the results obtained therefrom supports or disproves our hypothesis. In section 1.2 of this thesis we defined our hypothesis as follows:

"The approach with which and time at which vehicles are clustered and re-clustered in a relay hotspot scenario can be exploited to improve V2X network performance and consequently the Quality of Experience (QoE) for users engaged in real-time and heavy download use-cases."

The means by which we propose to validate the hypothesis is set under three objectives under which we will be explaining how the hypothesis has been validate.

The first objective involved demonstrating how re-clustering can be approached to improve V2V side-link performance and reliability. In chapter 3 we proved that we could improve throughput, stability, and jitter performances of side-links by employing an SNR based clustering criteria over distance-based approach in the cluster formation process and by the use of two different memory-based techniques that feeds back CH coordinates as seeds and centroid into different stages of succeeding re-clustering process. In this part of our work, it was proved that while exploiting the benefits of persistent re-clustering and SNR based cluster formation, stability can also be improved using the newly introduced memory-based re-clustering approach.

The second objective involved exploiting how the instance in time of initiating the re-clustering process can be exploited to improve V2V side-link performance. Here, rather than the use of time intervals or CH membership to determine instance of re-clustering, we approached two performance-based metrics; cluster quality index (CQRD) and average throughput (LQRD) to determine when re-clustering should be initiated. The result yet again validated our hypothesis, as the results obtained indicates both schemes to have better throughput performance as compared to the common CH membership scheme, with the LQRD also marginally outperforming the persistent re-clustering approach used in our previous validation attempt. Though a marginal throughput improvement was recorded, a significant improvement in Jitter and stability performance was recorded over PRS.

The third objective attempts to resolve a potential cluster backhaul bottleneck that could ensue from improvements in CH-to-CM V2V side-link throughput performance. Regardless of any throughput improvement that could be made by achieving the first two objectives, a poor throughput performance on the backhaul V2I/N link will limit the performance of the side-link. We improved V2I performance by exploiting a new approach to selecting the number of vehicular clusters. The new approach selects k-number of clusters that maximises the use of bandwidth resources available to V2I users. Hence, maximizing backhaul throughput and essentially expanding the limits within which our hypothesis is realisable.

6.2 Review of Limitation and Recommendation for future Work

This section reviews the limits of the framework within which the work we have reported in this thesis has been implemented and as an offshoot describes the recommended areas of exploration, we have identified during the work and other areas that can be improved upon. We will be describing these limits and future work in four categories; the topology of context, the use-case under consideration, and clustering.

The topology upon which this work is built is a single cell hotspot topology, wherein intra-cellular vehicular clusters are formed, and no detailed consideration is given to inter-cell interference. CMs' access to traffic data at edge server fare limited to their respective CHs that serve as the only relay points to the edge server located at the BS. The topology in our framework also assumes a single interface (V2I or V2V) is usable at every point-in-time and that each vehicle is equipped with one of each. These topological assumptions however realistic, limits the array of relay opportunities that individual vehicles can avail for download of traffic data. For future work, we propose the exploration of scenarios and algorithms that enables CMs to exploit the option of selecting between multiple interfaces and links and/or a coordination of multiple interfaces and links. This could be a combination of relay and direct points or combination of multiple relay or multiple direct points. This suggests the adoption of heterogeneous vehicular networks where vehicles could be equipped with multiple V2X RATs of same or different evolutional lines such as IEEE 802.11p and IEEE 802.11bd or IEEE 802.11bd and 5G-V2X respectively. It is also worth investigating how our proposed schemes work in a multi-cellular context. Particularly how the cluster-based resource allocation works within different frequency reuse schemes and how interference and handover affect the outcome of the schemes.

The use-case in consideration in our scheme particularly looks to a short-term real time download of traffic data and urgent software updates. However emerging use cases that might require a long continuous download of high precision and data intensive data will require studies over a longer period of time and across multiple cells. The implication of this would be that the approach where we have used a fixed number of clusters across time might be unsustainable after some time in this context. A newer memory-based approach where data from current cluster structure will still be carried over in successive re-clustering phases,

however, with the number of clusters adapting to changing topologies and increasing number of vehicular nodes can be conceptualized for further research.

The clustering approach throughout this work reported in this thesis is based on CH selection from within vehicular nodes. However, vehicles nodes move in different directions, make different mobility decisions, and are restricted by road networks. These make them often unpredictable and incapable of assuming the best coordinates to serve CMs. This in turn limits the stability and throughput performance obtainable from the CH-to-CM V2V side-links. For future work, a consideration of dedicated cluster heads whose mobility will be predictable and less prone to issues of line-of-sight is worthy of some research effort. We propose the use of dedicated Unmanned Aerial Vehicles (UAVs) as CHs which has the potential to persistently compute and assume the best CH position to serve CMs. UAVs also has the advantage of better aerial coordination, coverage, and LOS.

The re-clustering decision making algorithm that we have introduced in chapter 4 of this thesis is based on instantaneous decision making, where vehicles make computations and instant decision based on the information currently available to them. This limits the speed with which the network responds to changes in the highly dynamic network. Though the schemes presented in chapter 4 are still a reasonable re-clustering approach for semi-autonomous vehicles, however for fully autonomous driving scenarios, we propose pre-acquisition of vehicular trajectory by the BS-based MEC, to make pre-emptive network and traffic decisions to improve network and traffic efficiency, while minimizing response times. Also, the use of ML prediction algorithms, such as Recurrent Neural Network (RNN) or Long Short-Term Memory (LSTM) Networks can be employed to make traffic flow prediction and improve traffic management to prevent congestion by pre-emptively scheduling vehicular updates and make intelligent re-clustering decisions.

In chapter 5, we discussed two different V2X resource allocation schemes that essentially gives same weight of resource entitlement to all V2I user-links, comprising of CH-to-BS and FV-to-BS. However, both links have different level of importance and should be allocated different levels of bandwidth resource. The CH has a relay responsibility and any throughput bottleneck on its relay path will affect all other vehicles within its cluster. For future work we propose a deterministic approach to resource allocation where the throughput requirement

of end users is ascertained. Then bandwidth allocation to each user-link will be weighted based the level of relay responsibility each user bears and the link channel condition (SINR). We suppose this approach will allocate a fairer bandwidth resource to CHs and further reduce the risk of potential downlink bottlenecks.

Bibliography

[1] "Road traffic injuries." Accessed: May 29, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

[2] "Guest Article: Saving Young Lives, Protecting the Planet, and Growing the Economy: Road Safety for 2030 | SDG Knowledge Hub | IISD." Accessed: Oct. 22, 2021. [Online]. Available: https://sdg.iisd.org/commentary/guest-articles/saving-young-lives-protecting-the-planet-and-growing-the-economy-road-safety-for-2030/

[3] "Chapter 1 Summary for Policymakers", doi: 10.1017/CBO9781107415324.004.

[4] S. F. Hasan, N. Siddique, and S. Chakraborty, "Intelligent transportation systems: 802.11-based vehicular communications," *Intelligent Transportation Systems: 802.11-based Vehicular Communications*, pp. 1–183, Sep. 2017, doi: 10.1007/978-3-319-64057-0.

[5] A. Eduardo Fernandez *et al.*, "Deliverable D2.1 5GCAR Scenarios, Use Cases, Requirements and KPIs," 2017.

[6] A. Eduardo Fernandez *et al.*, "Deliverable D2.1 5GCAR Scenarios, Use Cases, Requirements and KPIs," 2019.

[7] D. for Transport, "Transport Statistics Great Britain: 2019," 2019, Accessed: Apr. 11, 2023. [Online]. Available: www.gov.uk/dft

[8] D. J. Yeong, G. Velasco-hernandez, J. Barry, and J. Walsh, "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review," *Sensors 2021, Vol. 21, Page 2140*, vol. 21, no. 6, p. 2140, Mar. 2021, doi: 10.3390/S21062140.

[9] "J3016_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International." Accessed: Apr. 17, 2023. [Online]. Available: https://www.sae.org/standards/content/j3016_202104/

[10] "SAE International Releases Updated Visual Chart for Its 'Levels of Driving Automation' Standard for Self-Driving Vehicles." Accessed: May 29, 2023. [Online]. Available: https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles

[11] H. Jianhua *et al.*, "Cooperative Connected Autonomous Vehicles (CAV): Research, Applications and Challenges," in *2019 IEEE 27th International Conference on Network Protocols (ICNP)*, Chicago: IEEE, Sep. 2019. Accessed: Apr. 18, 2023. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8888126&tag=1

[12] "Autonomous Vehicles & the Role of C-V2X Cellular Technology." Accessed: Sep. 30, 2023. [Online]. Available: https://www.iotforall.com/autonomous-vehicles-and-the-role-of-5g-cellular-technology

[13] P. Luoto, M. Bennis, P. Pirinen, S. Samarakoon, K. Horneman, and M. Latva-Aho, "Vehicle clustering for improving enhanced LTE-V2X network performance," *EuCNC 2017 - European Conference on Networks and Communications*, Jul. 2017, doi: 10.1109/EUCNC.2017.7980735.

[14] "C-V2X explained - 5GAA." Accessed: Apr. 16, 2023. [Online]. Available: https://5gaa.org/c-v2x-explained/

[15] TSGS, "TS 122 185 - V14.3.0 - LTE; Service requirements for V2X services (3GPP TS 22.185 version 14.3.0 Release 14)," 2017, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[16] Y. Luo, W. Zhang, and Y. Hu, "A new cluster based routing protocol for VANET," in *NSWCTC 2010 - The 2nd International Conference on Networks Security, Wireless Communications and Trusted Computing*, 2010, pp. 176–180. doi: 10.1109/NSWCTC.2010.48.

[17] X. Cheng, B. Huang, and W. Cheng, “Stable clustering for VANETs on highways,” *Proceedings - 2018 3rd ACM/IEEE Symposium on Edge Computing, SEC 2018*, pp. 399–403, Dec. 2018, doi: 10.1109/SEC.2018.00053.

[18] G. H. Alsuhli, A. Khattab, and Y. A. Fahmy, “An evolutionary approach for optimized VANET clustering,” *Proceedings of the International Conference on Microelectronics, ICM*, vol. 2019-December, pp. 70–73, Dec. 2019, doi: 10.1109/ICM48031.2019.9021941.

[19] A. A. Khan, M. Abolhasan, and W. Ni, “An evolutionary game theoretic approach for stable and optimized clustering in vanets,” *IEEE Trans Veh Technol*, vol. 67, no. 5, pp. 4501–4513, May 2018, doi: 10.1109/TVT.2018.2790391.

[20] P. Fan, J. G. Haran, J. Dillenburg, and P. C. Nelson, “Cluster-Based Framework in Vehicular Ad-Hoc Networks,” 2005.

[21] A. Bavalatti and A. V. Sutagundar, “Multi-Agent based stable clustering in VANET,” *Proceedings of the 2017 International Conference On Smart Technology for Smart Nation, SmartTechCon 2017*, pp. 1033–1038, May 2018, doi: 10.1109/SMARTTECHCON.2017.8358527.

[22] G. Liu, N. Qi, J. Chen, C. Dong, and Z. Huang, “Enhancing clustering stability in VANET: A spectral clustering based approach,” *China Communications*, vol. 17, no. 4, pp. 140–151, Apr. 2020, doi: 10.23919/JCC.2020.04.013.

[23] T. Wang, X. Cao, and S. Wang, “Self-adaptive clustering and load-bandwidth management for uplink enhancement in heterogeneous vehicular networks,” *IEEE Internet Things J*, vol. 6, no. 3, pp. 5607–5617, Jun. 2019, doi: 10.1109/JIOT.2019.2904036.

[24] J. Zhao *et al.*, “Cluster-based resource selection scheme for 5G V2X,” *IEEE Vehicular Technology Conference*, vol. 2019-April, Apr. 2019, doi: 10.1109/VTCSPRING.2019.8746637.

[25] F. Abbas, G. Liu, P. Fan, and Z. Khan, “An Efficient Cluster Based Resource Management Scheme and its Performance Analysis for V2X Networks,” *IEEE Access*, vol. 8, pp. 87071–87082, 2020, doi: 10.1109/ACCESS.2020.2992591.

[26] K. Popoola, D. Grace, T. Clarke, and M. Ahmed, “An Iterative k-means++ and Ant Colony Clustering Scheme for Vehicular Networks; An Iterative k-means++ and Ant Colony Clustering Scheme for Vehicular Networks,” 2022, doi: 10.1109/NIGERCON54645.2022.9803091.

[27] 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,” *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 8, no. 12, 2017, Accessed: Jun. 10, 2023. [Online]. Available: www.ijacsa.thesai.org

[28] TSGR, “TR 137 985 - V16.0.0 - LTE; 5G; Overall description of Radio Access Network (RAN) aspects for Vehicle-to-everything (V2X) based on LTE and NR (3GPP TR 37.985 version 16.0.0 Release 16),” 2020, Accessed: May 22, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[29] H. Cao, S. Gangakhedkar, A. R. Ali, M. Gharba, and J. Eichinger, “A 5G V2X testbed for cooperative automated driving,” *IEEE Vehicular Networking Conference, VNC*, vol. 0, Jul. 2016, doi: 10.1109/VNC.2016.7835939.

[30] A. Choudhury, T. Maszczyk, C. B. Math, H. Li, and J. Dauwels, “An Integrated Simulation Environment for Testing V2X Protocols and Applications,” *Procedia Comput Sci*, vol. 80, pp. 2042–2052, Jan. 2016, doi: 10.1016/J.PROCS.2016.05.524.

[31] M. Ayyub, A. Oracevic, R. Hussain, A. A. Khan, and Z. Zhang, “A comprehensive survey on clustering in vehicular networks: Current solutions and future challenges,” *Ad Hoc Networks*, vol. 124, p. 102729, Jan. 2022, doi: 10.1016/J.ADHOC.2021.102729.

[32] Z. H. Mir, J. Toutouh, F. Filali, and Y. B. Ko, “Enabling DSRC and c-V2X integrated hybrid vehicular networks: Architecture and protocol,” *IEEE Access*, vol. 8, pp. 180909–180927, 2020, doi: 10.1109/ACCESS.2020.3027074.

[33] "5GAA, BMW Group, Ford and Groupe PSA Exhibit First European C-V2X Direct Communication Interoperability Between Multiple Automakers | Ford of Europe | Ford Media Center." Accessed: May 02, 2023. [Online]. Available: https://media.ford.com/content/fordmedia/feu/en/news/2018/07/11/-5gaa--bmw-group--ford-and-groupe-psa--exhibit-first-european-c-.html

[34] "Toyota and Lexus to launch DSRC technology to connect vehicles and infrastructure in the U.S. in 2021 -." Accessed: May 02, 2023. [Online]. Available: https://site.ieee.org/connected-vehicles/2018/04/16/toyota-and-lexus-to-launch-dsrc-technology-to-connect-vehicles-and-infrastructure-in-the-u-s-in-2021/

[35] "TS 133 185 - V14.1.0 - 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Security aspect for LTE support of Vehicle-to-Everything (V2X) services (Release 14)," 2017, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[36] TSGS, "TS 122 186 - V15.4.0 - 5G; Service requirements for enhanced V2X scenarios (3GPP TS 22.186 version 15.4.0 Release 15)," 2018, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[37] U. N. Kar and D. K. Sanyal, "An overview of device-to-device communication in cellular networks," *ICT Express*, vol. 4, no. 4. Korean Institute of Communications Information Sciences, pp. 203–208, Dec. 01, 2018. doi: 10.1016/j.icte.2017.08.002.

[38] K. Katsaros and M. Dianati, "Evolution of Vehicular Communications within the Context of 5G Systems," *Enabling 5G Communication Systems to Support Vertical Industries*, pp. 103–126, Jul. 2019, doi: 10.1002/9781119515579.CH5.

[39] F. Abbas, P. Fan, and Z. Khan, "A Novel Low-Latency V2V Resource Allocation Scheme Based on Cellular V2X Communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2185–2197, Jun. 2019, doi: 10.1109/TITS.2018.2865173.

[40] R. Molina-Masegosa, J. Gozalvez, and M. Sepulcre, "Configuration of the C-V2X Mode 4 Sidelink PC5 Interface for Vehicular Communication," *Proceedings - 14th International Conference on Mobile Ad-Hoc and Sensor Networks, MSN 2018*, pp. 43–48, Jul. 2018, doi: 10.1109/MSN.2018.00014.

[41] TSGR, "TS 136 331 - V15.3.0 - LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification (3GPP TS 36.331 version 15.3.0 Release 15)," 2018, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[42] E. Standard, "Intelligent Transport Systems (ITS); Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band," 2013.

[43] TSGR, "TS 138 300 - V15.3.1 - 5G; NR; Overall description; Stage-2 (3GPP TS 38.300 version 15.3.1 Release 15)," 2018, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[44] S. Zeadally, M. A. Javed, and E. Ben Hamida, "Vehicular Communications for ITS: Standardization and Challenges," *IEEE Communications Standards Magazine*, vol. 4, no. 1, pp. 11–17, Mar. 2020, doi: 10.1109/MCOMSTD.001.1900044.

[45] TSGS, "TS 122 186 - V16.2.0 - 5G; Service requirements for enhanced V2X scenarios (3GPP TS 22.186 version 16.2.0 Release 16)," 2020, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[46] TSGR, "TS 136 302 - V14.2.0 - LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Services provided by the physical layer (3GPP TS 36.302 version 14.2.0 Release 14)," 2017, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[47] "5GAA Releases White Paper on C-V2X Use Cases: Methodology, Examples and Service Level Requirements - 5GAA." Accessed: May 29, 2023. [Online]. Available: https://5gaa.org/5gaa-releases-white-paper-on-c-v2x-use-cases-methodology-examples-and-service-level-requirements/

[48] M. H. C. Garcia *et al.*, "A Tutorial on 5G NR V2X Communications," *IEEE Communications Surveys and Tutorials*, 2021, doi: 10.1109/COMST.2021.3057017.

[49] "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Study on Vehicle-to-Everything (Release 16)," 2019.

[50] TSGR, "TS 138 211 - V15.3.0 - 5G; NR; Physical channels and modulation (3GPP TS 38.211 version 15.3.0 Release 15)," 2018, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[51] TSGR, "TS 138 101-2 - V15.8.0 - 5G; NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone (3GPP TS 38.101-2 version 15.8.0 Release 15)," 2020, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[52] K. Sakaguchi *et al.*, "Towards mmWave V2X in 5G and Beyond to Support Automated Driving," no. 6, 2021, doi: 10.1587/transcom.2020EBI0001.

[53] G. Yang, M. Xiao, M. Alam, and Y. Huang, "Low-Latency Heterogeneous Networks with Millimeter-Wave Communications," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 124–129, Jun. 2018, doi: 10.1109/MCOM.2018.1700874.

[54] T. Shimizu, V. Va, G. Bansal, and R. W. Heath, "Millimeter wave V2X communications: Use cases and design considerations of beam management," *Asia-Pacific Microwave Conference Proceedings, APMC*, vol. 2018-November, pp. 183–185, Jan. 2019, doi: 10.23919/APMC.2018.8617303.

[55] B. Antonescu, M. T. Moayyed, and S. Basagni, "mmWave channel propagation modeling for V2X communication systems," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, vol. 2017-October, pp. 1–6, Feb. 2018, doi: 10.1109/PIMRC.2017.8292718.

[56] B. Antonescu, M. T. Moayyed, and S. Basagni, "Diffuse Scattering Models for mmWave V2X Communications in Urban Scenarios," *2019 International Conference on*

Computing, Networking and Communications, ICNC 2019, pp. 923–929, Apr. 2019, doi: 10.1109/ICCNC.2019.8685661.

[57] F. Abbas, G. Liu, Z. Khan, K. Zheng, and P. Fan, "Clustering based resource management scheme for latency and sum rate optimization in V2X networks," *IEEE Vehicular Technology Conference*, vol. 2019-April, Apr. 2019, doi: 10.1109/VTCSPRING.2019.8746297.

[58] K. Sehla, T. M. T. Nguyen, G. Pujolle, and P. B. Velloso, "A New Clustering-based Radio Resource Allocation Scheme for C-V2X," *IFIP Wireless Days*, vol. 2021-June, 2021, doi: 10.1109/WD52248.2021.9508289.

[59] "Release 17." Accessed: May 29, 2023. [Online]. Available: https://www.3gpp.org/specifications-technologies/releases/release-17

[60] M. Harounabadi, D. M. Soleymani, S. Bhadauria, M. Leyh, and E. Roth-Mandutz, "V2X in 3GPP Standardization: NR Sidelink in Release-16 and beyond," *IEEE Communications Standards Magazine*, vol. 5, no. 1, pp. 12–21, Mar. 2021, doi: 10.1109/MCOMSTD.001.2000070.

[61] Z. Machardy, A. Khan, K. Obana, and S. Iwashina, "V2X access technologies: Regulation, research, and remaining challenges," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 3, pp. 1858–1877, Jul. 2018, doi: 10.1109/COMST.2018.2808444.

[62] P. K. Singh, S. K. Nandi, and S. Nandi, "A tutorial survey on vehicular communication state of the art, and future research directions," *Vehicular Communications*, vol. 18, Aug. 2019, doi: 10.1016/j.vehcom.2019.100164.

[63] W. Anwar, N. Franchi, and G. Fettweis, "Physical layer evaluation of V2X communications technologies: 5G NR-V2X, LTE-V2X, IEEE 802.11bd, and IEEE 802.11p," *IEEE Vehicular Technology Conference*, vol. 2019-September, Sep. 2019, doi: 10.1109/VTCFALL.2019.8891313.

[64] R. Molina-Masegosa, J. Gozalvez, and M. Sepulcre, “Comparison of IEEE 802.11p and LTE-V2X: An Evaluation with Periodic and Aperiodic Messages of Constant and Variable Size,” *IEEE Access*, vol. 8, pp. 121526–121548, 2020, doi: 10.1109/ACCESS.2020.3007115.

[65] “3 rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on LTE support for Vehicle to Everything (V2X) services (Release 14),” 2015. [Online]. Available: http://www.3gpp.org

[66] “Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions,” 2009.

[67] K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, and Y. Zhou, “Heterogeneous Vehicular Networking: A Survey on Architecture, Challenges, and Solutions,” *IEEE Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2377–2396, Oct. 2015, doi: 10.1109/COMST.2015.2440103.

[68] R. Jacob, N. Franchi, and G. Fettweis, “Hybrid V2X Communications: Multi-RAT as Enabler for Connected Autonomous Driving,” *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, vol. 2018-September, pp. 1370–1376, Dec. 2018, doi: 10.1109/PIMRC.2018.8580953.

[69] “C-V2X in action - 5GAA.” Accessed: May 11, 2023. [Online]. Available: https://5gaa.org/c-v2x-in-action/

[70] E. Standard, “EN 302 663 - V1.2.0 - Intelligent Transport Systems (ITS); Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band,” 2012.

[71] J. Brož, T. Tichý, V. Angelakis, and Z. Bělinová, “Usage of V2X Applications in Road Tunnels,” *Applied Sciences 2022, Vol. 12, Page 4624*, vol. 12, no. 9, p. 4624, May 2022, doi: 10.3390/APP12094624.

[72] MTR, “ETSI TR 103 257-1 V1.1.1,” 2019, Accessed: May 29, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[73] T. Abbas and F. Tufvesson, “Line-of-sight obstruction analysis for vehicle-to-vehicle network simulations in a two-lane highway scenario,” *Int J Antennas Propag*, vol. 2013, 2013, doi: 10.1155/2013/459323.

[74] C. Campolo, A. Molinaro, A. Iera, R. R. Fontes, and C. E. Rothenberg, “Towards 5G Network Slicing for the V2X Ecosystem,” *2018 4th IEEE Conference on Network Softwarization and Workshops, NetSoft 2018*, pp. 303–307, Sep. 2018, doi: 10.1109/NETSOFT.2018.8459911.

[75] F. Pasic, S. Pratschner, R. Langwieser, and C. F. Mecklenbrauker, “High-Mobility Wireless Channel Measurements at 5.9 GHz in an Urban Environment,” *2022 International Balkan Conference on Communications and Networking, BalkanCom 2022*, pp. 100–104, 2022, doi: 10.1109/BALKANCOM55633.2022.9900633.

[76] K. Yang *et al.*, “Vehicle-to-vehicle radio channel characteristics for a long-distance urban congestion scenario at 5.9 GHz,” *2021 IEEE 4th International Conference on Electronic Information and Communication Technology, ICEICT 2021*, pp. 577–581, Aug. 2021, doi: 10.1109/ICEICT53123.2021.9531229.

[77] C. S. Kim, J. S. Kim, J. Y. Hong, J. S. Lim, and Y. J. Chong, “Propagation characteristics of Urban and Highway Vehicle-to-Everything(V2X) Channels at 5.9 GHz,” *International Conference on ICT Convergence*, vol. 2021-October, pp. 872–876, 2021, doi: 10.1109/ICTC52510.2021.9620982.

[78] C. S. Kim, H. J. Kim, J. S. Lim, J. Y. Hong, and Y. J. Chong, “Stationary Distance Measurement Using Various Methods for Expressway Environment at 5.9 GHz,” *ICTC 2019 - 10th International Conference on ICT Convergence: ICT Convergence Leading the Autonomous Future*, pp. 1062–1065, Oct. 2019, doi: 10.1109/ICTC46691.2019.8939815.

[79] Z. Su *et al.*, "Measurements and Characteristics for Vehicle to Everything Channel in Tunnel Scenario at 5.9GHZ," *2021 7th International Conference on Computer and Communications, ICCC 2021*, pp. 230–234, 2021, doi: 10.1109/ICCC54389.2021.9674537.

[80] H. Jin, C. Li, and W. Wu, "Path Loss Analysis for VANET in Tunnel Environment," *International Conference on ICT Convergence*, vol. 2021-October, pp. 241–244, 2021, doi: 10.1109/ICTC52510.2021.9621031.

[81] F. A. Rodriguez-Corbo, L. Azpilicueta, M. Celaya-Echarri, A. V. Alejos, and F. Falcone, "Propagation Models in Vehicular Communications," *IEEE Access*, vol. 9, pp. 15902–15913, 2021, doi: 10.1109/ACCESS.2021.3049884.

[82] W. Viriyasitavat, M. Boban, H. M. Tsai, and A. Vasilakos, "Vehicular communications: Survey and challenges of channel and propagation models," *IEEE Vehicular Technology Magazine*, vol. 10, no. 2, pp. 55–66, Jun. 2015, doi: 10.1109/MVT.2015.2410341.

[83] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015, doi: 10.1109/ACCESS.2015.2453991.

[84] D. Bilibashi, E. M. Vitucci, and V. Degli-Esposti, "Dynamic Ray Tracing: Introduction and Concept," *14th European Conference on Antennas and Propagation, EuCAP 2020*, Mar. 2020, doi: 10.23919/EUCAP48036.2020.9135577.

[85] O. Renaudin, V. M. Kolmonen, P. Vainikainen, and C. Oestges, "Wideband measurement-based modeling of inter-vehicle channels in the 5-GHz band," *IEEE Trans Veh Technol*, vol. 62, no. 8, pp. 3531–3540, 2013, doi: 10.1109/TVT.2013.2257905.

[86] J. Karedal *et al.*, "A geometry-based stochastic MIMO model for vehicle-to-vehicle communications," *IEEE Trans Wirel Commun*, vol. 8, no. 7, pp. 3646–3657, Jul. 2009, doi: 10.1109/TWC.2009.080753.

[87] Y. Yuan, C. X. Wang, Y. He, M. M. Alwakeel, and E. H. M. Aggoune, "3D Wideband Non-Stationary Geometry-Based Stochastic Models for Non-Isotropic MIMO Vehicle-to-Vehicle Channels," *IEEE Trans Wirel Commun*, vol. 14, no. 12, pp. 6883–6895, Dec. 2015, doi: 10.1109/TWC.2015.2461679.

[88] Q. Zhu *et al.*, "A novel 3d non-stationary vehicle-to-vehicle channel model and its spatial-temporal correlation properties," *IEEE Access*, vol. 6, pp. 43633–43643, Jul. 2018, doi: 10.1109/ACCESS.2018.2859782.

[89] M. F. Iskander and Z. Yun, "Propagation prediction models for wireless communication systems," *IEEE Trans Microw Theory Tech*, vol. 50, no. 3, pp. 662–673, Mar. 2002, doi: 10.1109/22.989951.

[90] F. Pinzel, J. Holfeld, A. Olunczek, P. Balzer, and O. Michler, "V2V-and V2X-Communication data within a distributed computing platform for adaptive radio channel modelling," *MT-ITS 2019 - 6th International Conference on Models and Technologies for Intelligent Transportation Systems*, Jun. 2019, doi: 10.1109/MTITS.2019.8883347.

[91] W. Zhang, T. Zhou, and L. Liu, "Measurements and Modeling of Narrow-Beam Channel Dispersion Characteristics in Vehicle-to-Infrastructure Scenarios," *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Mar. 2023, doi: 10.1109/WCNC55385.2023.10118782.

[92] Z. Huang, C. Xiang, and Z. Nan, "An Improved Non-Geometrical Stochastic Model for An Improved Non-Geometrical Stochastic Model for Non-WSSUS Vehicle-to-Vehicle Channels Non-WSSUS Vehicle-to-Vehicle Channels," vol. 17, no. 4, 2019, doi: 10.12142/ZTECOM.201904009.

[93] C. X. Wang, X. Cheng, and D. Laurenson, "Vehicle-to-vehicle channel modeling and measurements: Recent advances and future challenges," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 96–103, Nov. 2009, doi: 10.1109/MCOM.2009.5307472.

[94] G. Acosta-Marum and M. A. Ingram, “Six time-and frequency-selective empirical channel models for vehicular wireless LANs,” *IEEE Vehicular Technology Conference*, pp. 2134–2138, 2007, doi: 10.1109/VETECF.2007.448.

[95] “IEEE Standard for Information technology-- Local and metropolitan area networks-- Specific requirements-- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments,” pp. 1–51, Jul. 2010, doi: 10.1109/IEEESTD.2010.5514475.

[96] M. Noor-A-Rahim, Z. Liu, H. Lee, G. G. M. N. Ali, D. Pesch, and P. Xiao, “A Survey on Resource Allocation in Vehicular Networks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 701–721, Feb. 2022, doi: 10.1109/TITS.2020.3019322.

[97] S. Xiaoqin, M. Juanjuan, L. Lei, and Z. Tianchen, “Maximum-Throughput Sidelink Resource Allocation for NR-V2X Networks with the Energy-Efficient CSI Transmission,” *IEEE Access*, vol. 8, pp. 73164–73172, 2020, doi: 10.1109/ACCESS.2020.2983715.

[98] L. F. Abanto-Leon, A. Koppelaar, and S. H. De Groot, “Subchannel allocation for vehicle-to-vehicle broadcast communications in mode-3,” *IEEE Wireless Communications and Networking Conference, WCNC*, vol. 2018-April, pp. 1–6, Jun. 2018, doi: 10.1109/WCNC.2018.8377360.

[99] A. Nabil, K. Kaur, C. DIetrich, and V. Marojevic, “Performance Analysis of Sensing-Based Semi-Persistent Scheduling in C-V2X Networks,” *IEEE Vehicular Technology Conference*, vol. 2018-August, Jul. 2018, doi: 10.1109/VTCFALL.2018.8690600.

[100] L. F. Abanto-Leon, A. Koppelaar, and S. H. De Groot, “Graph-Based resource allocation with conflict avoidance for V2V broadcast communications,” *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, vol. 2017-October, pp. 1–7, Feb. 2018, doi: 10.1109/PIMRC.2017.8292606.

[101] D. Sempere-García, M. Sepulcre, and J. Gozalvez, "LTE-V2X Mode 3 scheduling based on adaptive spatial reuse of radio resources," *Ad Hoc Networks*, vol. 113, p. 102351, Mar. 2021, doi: 10.1016/J.ADHOC.2020.102351.

[102] D. ; Yuan, D. ; Hu, X. Chen, D. Yuan, D. Hu, and X. Chen, "Resource Allocation in C-V2X Mode 3 Based on the Exchanged Preference Profiles," *Electronics 2023, Vol. 12, Page 1071*, vol. 12, no. 5, p. 1071, Feb. 2023, doi: 10.3390/ELECTRONICS12051071.

[103] D. Yasuda *et al.*, "Radio resource allocation based on adaptive and maximum reuse distance for LTE-V2X sidelink mode 3," *IEICE Communications Express*, vol. 10, no. 10, pp. 792–797, Oct. 2021, doi: 10.1587/COMEX.2021XBL0127.

[104] K. Kim, D. Kim, S. Choi, D. Kwon, and J. W. Choi, "Location-Based Maximum Reuse Distance Resource Scheduling for LTE Cellular-V2X Sidelink Mode 3," *2022 International Conference on Electronics, Information, and Communication, ICEIC 2022*, 2022, doi: 10.1109/ICEIC54506.2022.9748563.

[105] G. Cecchini, A. Bazzi, M. Menarini, B. M. Masini, and A. Zanella, "Maximum Reuse Distance Scheduling for Cellular-V2X Sidelink Mode 3," *2018 IEEE Globecom Workshops, GC Wkshps 2018 - Proceedings*, Dec. 2018, doi: 10.1109/GLOCOMW.2018.8644360.

[106] C. Campolo, A. Molinaro, F. Romeo, A. Bazzi, and A. O. Berthet, "5G NR V2X: On the impact of a flexible numerology on the autonomous sidelink mode," *IEEE 5G World Forum, 5GWF 2019 - Conference Proceedings*, pp. 102–107, Sep. 2019, doi: 10.1109/5GWF.2019.8911694.

[107] B. Di, L. Song, Y. Li, and Z. Han, "V2x meets noma: Non-orthogonal multiple access for 5g-enabled vehicular networks," *IEEE Wirel Commun*, vol. 24, no. 6, pp. 14–21, Dec. 2017, doi: 10.1109/MWC.2017.1600414.

[108] F. Zhang, M. M. Wang, S. S. Chen, L. Shan, M. Xiao, and M. Xu, "Applying NOMA to NR V2X: A Graph-based Matching and Cooperative Game Approach," *IEEE Vehicular*

Technology Conference, vol. 2021-April, Apr. 2021, doi: 10.1109/VTC2021-SPRING51267.2021.9449051.

[109] F. Zhang, M. M. Wang, X. Bao, and W. Liu, "Centralized Resource Allocation and Distributed Power Control for NOMA-Integrated NR V2X," *IEEE Internet Things J*, vol. 8, no. 22, pp. 16522–16534, Nov. 2021, doi: 10.1109/JIOT.2021.3075250.

[110] D. Helbing, "Traffic and related self-driven many-particle systems," *Rev Mod Phys*, vol. 73, no. 4, p. 1067, Dec. 2001, doi: 10.1103/RevModPhys.73.1067.

[111] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: A survey and taxonomy," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 4, pp. 19–41, 2009, doi: 10.1109/SURV.2009.090403.

[112] "Paramics 3D Traffic Microsimulation & Modelling Software for Transport - SYSTRA." Accessed: Jun. 02, 2023. [Online]. Available: https://www.paramics.co.uk/en/

[113] "Eclipse SUMO - Simulation of Urban MObility." Accessed: Jun. 02, 2023. [Online]. Available: https://www.eclipse.org/sumo/

[114] "Download and test PTV Vissim now | MyPTV." Accessed: Jun. 02, 2023. [Online]. Available: https://www.myptv.com/en/mobility-software/ptv-vissim/demo

[115] L. Smith, R. Beckman, and K. Baggerly, "TRANSIMS: Transportation analysis and simulation system," Jul. 1995, doi: 10.2172/88648.

[116] W. Liu, X. Wang, W. Zhang, L. Yang, and C. Peng, "Coordinative simulation with SUMO and NS3 for Vehicular Ad Hoc Networks," *Proceedings - Asia-Pacific Conference on Communications, APCC 2016*, pp. 337–341, Oct. 2016, doi: 10.1109/APCC.2016.7581471.

[117] S. Mao, "Fundamentals of communication networks," *Cognitive Radio Communications and Networks*, pp. 201–234, 2010, doi: 10.1016/B978-0-12-374715-0.00008-3.

[118] “Informatik 4: BonnMotion.” Accessed: Jun. 02, 2023. [Online]. Available: https://net.cs.uni-bonn.de/wg/cs/applications/bonnmotion/

[119] “STRAW - STreet RAndom Waypoint - vehicular mobility model for network simulations.” Accessed: Jun. 02, 2023. [Online]. Available: http://oldaqualab.cs.northwestern.edu/projects/144-straw-street-random-waypoint-vehicular-mobility-model-for-network-simulations-e-g-car-networks

[120] “PRECISE | Vehicular Network Virtualization Platform.” Accessed: Jun. 02, 2023. [Online]. Available: https://precise.seas.upenn.edu/research-projects/groovenet-vehicular-network-simulator

[121] F. Benabdallah, A. Hamza, and M. Bechrif, “Simulation and analysis of VANETS performances based on the choice of mobility model,” *Proceedings of Computing Conference 2017*, vol. 2018-January, pp. 1238–1242, Jan. 2018, doi: 10.1109/SAI.2017.8252248.

[122] R. Supervisor, “Evaluation of Routing Protocols for Vehicular Ad hoc Networks (VANETs) in Connected Transportation Systems (135),” 2018, Accessed: Jun. 02, 2023. [Online]. Available: http://www.ntis.gov

[123] L. Bononi, M. Di Felice, M. Bertini, and E. Croci, “Parallel and distributed simulation of wireless vehicular ad hoc networks,” *ACM MSWiM 2006 - Proceedings of the 9th ACM Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, vol. 2006, pp. 28–35, 2006, doi: 10.1145/1164717.1164725.

[124] M. Piórkowski, M. Raya, A. L. Lugo, P. Papadimitratos, M. Grossglauser, and J.-P. Hubaux, “TraNS,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 12, no. 1, pp. 31–33, Jan. 2008, doi: 10.1145/1374512.1374522.

[125] J. Anda, J. Lebrun, D. Ghosal, C.-N. Chuah, and M. Zhang, “VGrid: Vehicular AdHoc Networking and Computing Grid for Intelligent Traffic Control”.

[126] C. Lochert, B. Scheuermann, A. Barthels, A. Cervantes, M. Mauve, and M. Caliskan, “Multiple Simulator Interlinking Environment for IVC”, Accessed: Jun. 03, 2023. [Online]. Available: http://www.network-on-wheels.de.

[127] “FDK - Federated Simulations Development Kit.” Accessed: Jun. 03, 2023. [Online]. Available: https://www.cc.gatech.edu/computing/pads/fdk.html

[128] “Veins.” Accessed: Jun. 14, 2023. [Online]. Available: https://veins.car2x.org/

[129] P. Yang, J. Wang, Y. Zhang, Z. Tang, and S. Song, “Clustering algorithm in VANETs: A survey,” *Proceedings of the International Conference on Anti-Counterfeiting, Security and Identification, ASID*, vol. 2016-February, pp. 166–170, Feb. 2016, doi: 10.1109/ICASID.2015.7405685.

[130] M. F. Khan, K. L. A. Yau, R. M. Noor, and M. A. Imran, “Survey and taxonomy of clustering algorithms in 5G,” *Journal of Network and Computer Applications*, vol. 154, Mar. 2020, doi: 10.1016/j.jnca.2020.102539.

[131] M. Mukhtaruzzaman and M. Atiquzzaman, “Clustering in vehicular ad hoc network: Algorithms and challenges,” *Computers and Electrical Engineering*, vol. 88, Dec. 2020, doi: 10.1016/j.compeleceng.2020.106851.

[132] K. Abboud and W. Zhuang, “Stochastic modeling of single-hop cluster stability in vehicular ad hoc networks,” *IEEE Trans Veh Technol*, vol. 65, no. 1, pp. 226–240, Jan. 2016, doi: 10.1109/TVT.2015.2396298.

[133] M. Hadded, R. Zagrouba, A. Laouiti, P. Muhlethaler, and L. A. Saidane, “A multi-objective genetic algorithm-based Adaptive Weighted Clustering Protocol in VANET,” *2015 IEEE Congress on Evolutionary Computation, CEC 2015 - Proceedings*, pp. 994–1002, Sep. 2015, doi: 10.1109/CEC.2015.7256998.

[134] B. Marzak, H. Toumi, M. Talea, and E. Benlahmar, “Cluster head selection algorithm in vehicular Ad Hoc networks,” *Proceedings of 2015 International Conference on Cloud*

Computing Technologies and Applications, CloudTech 2015, Nov. 2015, doi: 10.1109/CLOUDTECH.2015.7336994.

[135] M. Ren, L. Khoukhi, H. Labiod, J. Zhang, and V. Veque, "A new mobility-based clustering algorithm for vehicular ad hoc networks (VANETs)," *Proceedings of the NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 1203–1208, Jun. 2016, doi: 10.1109/NOMS.2016.7502988.

[136] G. V. Rossi, Z. Fan, W. H. Chin, and K. K. Leung, "Stable clustering for Ad-Hoc vehicle networking," *IEEE Wireless Communications and Networking Conference, WCNC*, May 2017, doi: 10.1109/WCNC.2017.7925786.

[137] Z. Yang, W. Wu, Y. Chen, X. Lin, and X. Chen, "Navigation route based stable clustering for vehicular ad hoc networks," *China Communications*, vol. 15, no. 3, pp. 42–56, Mar. 2018, doi: 10.1109/CC.2018.8331990.

[138] M. Kadadha, H. Otrok, H. Barada, M. Al-Qutayri, and Y. Al-Hammadi, "A Cluster-Based QoS-OLSR Protocol for Urban Vehicular Ad Hoc Networks," *2018 14th International Wireless Communications and Mobile Computing Conference, IWCMC 2018*, pp. 554–559, Aug. 2018, doi: 10.1109/IWCMC.2018.8450405.

[139] P. Jacquet, P. Mühlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, "Optimized link state routing protocol for ad hoc networks," *Proceedings - IEEE lnternational Multi Topic Conference 2001: Technology for the 21st Century, IEEE INMIC 2001*, pp. 62–68, 2001, doi: 10.1109/INMIC.2001.995315.

[140] M. A. Saleem *et al.*, "Expansion of Cluster Head Stability Using Fuzzy in Cognitive Radio CR-VANET," *IEEE Access*, vol. 7, pp. 173185–173195, 2019, doi: 10.1109/ACCESS.2019.2956478.

[141] S. Chettibi, "Combination of HF set and MCDM for stable clustering in VANETs," *IET Intelligent Transport Systems*, vol. 14, no. 3, pp. 190–195, Mar. 2020, doi: 10.1049/IET-ITS.2019.0283.

[142] Z. Khan and P. Fan, “A Novel Triple Cluster Based Routing Protocol (TCRP) for VANETs,” *IEEE Vehicular Technology Conference*, vol. 2016-July, Jul. 2016, doi: 10.1109/VTCSPRING.2016.7504092.

[143] M. S. Talib *et al.*, “A center-based stable evolving clustering algorithm with grid partitioning and extended mobility features for VANETs,” *IEEE Access*, vol. 8, pp. 169908–169921, 2020, doi: 10.1109/ACCESS.2020.3020510.

[144] H. Fatemidokht and M. Kuchaki Rafsanjani, “QMM-VANET: An efficient clustering algorithm based on QoS and monitoring of malicious vehicles in vehicular ad hoc networks,” *Journal of Systems and Software*, vol. 165, p. 110561, Jul. 2020, doi: 10.1016/J.JSS.2020.110561.

[145] J. Pereira, M. Diaz-Cacho, S. Sargento, A. Zuquete, L. Guardalben, and M. Luis, “Vehicle-to-vehicle real-time video transmission through IEEE 802.11p for assisted-driving,” *IEEE Vehicular Technology Conference*, vol. 2018-June, pp. 1–6, Jul. 2018, doi: 10.1109/VTCSPRING.2018.8417766.

[146] H. Khan, S. Samarakoon, and M. Bennis, “Enhancing Video Streaming in Vehicular Networks via Resource Slicing,” *IEEE Trans Veh Technol*, vol. 69, no. 4, pp. 3513–3522, Apr. 2020, doi: 10.1109/TVT.2020.2975068.

[147] TSGS, “TR 126 985 - V16.0.0 - 5G; Vehicle-to-everything (V2X); Media handling and interaction (3GPP TR 26.985 version 16.0.0 Release 16),” 2020, Accessed: Jun. 14, 2023. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[148] ETSI, “Intelligent Transport Systems (ITS) Access Layer; Part 1: Channel Models for the 5,9 GHz frequency band,” *ETSI Technical Report*, vol. 1.1.1, no. ETSI TR 103 257-1, 2019, Accessed: Feb. 29, 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/103200_103299/10325701/01.01.01_60/tr_10325701v010101p.pdf

[149] C. Shi, B. Wei, S. Wei, W. Wang, H. Liu, and J. Liu, “A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm,” *EURASIP J*

Wirel Commun Netw, vol. 2021, no. 1, pp. 1–16, Dec. 2021, doi: 10.1186/S13638-021-01910-W/FIGURES/6.

[150] "Stop Using Elbow Method in K-means Clustering, Instead, Use this! | by Anmol Tomar | Towards Data Science." Accessed: Jul. 23, 2023. [Online]. Available: https://towardsdatascience.com/elbow-method-is-not-sufficient-to-find-best-k-in-k-means-clustering-fc820da0631d

[151] D. Marutho, S. Hendra Handaka, E. Wijaya, and Muljono, "The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News," *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, iSemantic 2018*, pp. 533–538, Nov. 2018, doi: 10.1109/ISEMANTIC.2018.8549751.

[152] TSGR, "TR 138 901 - V14.3.0 - 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 14.3.0 Release 14)," 2018. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

[153] S. Ucar, S. C. Ergen, and O. Ozkasap, "VeSCA: Vehicular stable cluster-based data aggregation," *2014 International Conference on Connected Vehicles and Expo, ICCVE 2014 - Proceedings*, pp. 1080–1085, 2014, doi: 10.1109/ICCVE.2014.7297517.

[154] R. Singh, D. Saluja, and S. Kumar, "Reliability improvement in clustering-based vehicular Ad-Hoc network," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1351–1355, Jun. 2020, doi: 10.1109/LCOMM.2020.2980819.