Software-Defined Heterogeneous Vehicular Networking: The Architectural Design and Open Challenges

Adnan Mahmood *, Wei Emma Zhang and Quan Z. Sheng *

Department of Computing, Macquarie University, Sydney, NSW 2109, Australia; w.zhang@mq.edu.au
* Correspondence: adnan.mahmood@hdr.mq.edu.au (A.M.); michael.sheng@mq.edu.au (Q.Z.S.)

Received: 3 January 2019; Accepted: 7 March 2019; Published: 11 March 2019

Abstract: The promising advancements in the telecommunications and automotive sectors over the years have empowered drivers with highly innovative communication and sensing capabilities, in turn paving the way for the next-generation connected and autonomous vehicles. Today, vehicles communicate wirelessly with other vehicles and vulnerable pedestrians in their immediate vicinity to share timely safety-critical information primarily for collision mitigation. Furthermore, vehicles connect with the traffic management entities via their supporting network infrastructure to become more aware of any potential hazards on the roads and for guidance pertinent to their current and anticipated speeds and travelling course to ensure more efficient traffic flows. Therefore, a secure and low-latency communication is highly indispensable in order to meet the stringent performance requirements of such safety-critical vehicular applications. However, the heterogeneity of diverse radio access technologies and inflexibility in their deployment results in network fragmentation and inefficient resource utilization, and these, therefore, act as bottlenecks in realizing the aims for a highly efficient vehicular networking architecture. In order to overcome such sorts of bottlenecks, this article brings forth the current state-of-the-art in the context of intelligent transportation systems (ITS) and subsequently proposes a software-defined heterogeneous vehicular networking (SDHVNet) architecture for ensuring a highly agile networking infrastructure to ensure rapid network innovation on-demand. Finally, a number of potential architectural challenges and their probable solutions are discussed.

Keywords: V2X communication; Internet of Things; Internet of Vehicles; heterogeneous networking; software-defined networks; safety-critical vehicular applications

1. Introduction

Over the past few decades, the promising notion of vehicular ad hoc networks (VANETs) has been thoroughly studied and well-researched by researchers in both academia and industry [1]. However, the emerging and promising paradigms of cloud computing, fog and/or edge computing, software-defined networks (SDN) and network functions virtualization have not only completely revolutionized the wireless networking industry, but have further spurred considerable innovative advancements for the transportation sector. This is coupled with other recent significant technological advances pertinent to the evolution of connected and autonomous vehicles and pervasive usage of numerous state-of-the-art sensory devices installed onboard vehicles that facilitate in a diverse range of cooperative vehicular safety applications, i.e., forward collision warnings, emergency vehicle assistance, (vulnerable) pedestrian collision mitigation, blind intersection warnings and hazardous location alerts, amongst many others. These safety applications are not only critical in nature, but further require a low-latency infrastructure with a maximum tolerable delay ranging between 10 ms and 50 ms [2]. Furthermore, modern-day connected vehicles are equipped with on average 100 sensors...
onboard, and this number is anticipated to increase up to 200 towards the end of the year 2020 [3]. These sensors not only generate the bulk amount of data, but also play an indispensable role in creating and sharing of ambient intelligence for vehicular cooperative communication. Furthermore, as per an estimate of Intel [4], an averagely-driven connected vehicle (i.e., a personal vehicle used for day-to-day routine purposes and not for any commercial operations) in the near future would generate around 4000 MB (40 TB) of data for every eight hours of its driving. This is in addition to the vehicular user’s data consumption, which on average stands at 650 MB per day and is expected to reach 1.5 GB per day by 2020.

The questions, therefore, arise as: (a) how to tackle such a flood of data so that the meaningful information could be accumulated, processed and utilized for the above-referred vehicular safety applications; (b) which particular radio access technologies would be able to facilitate the transmission of such sort of a meaningful information with higher data rates and lower end-to-end delay; and finally, (c) where this all processing (i.e., compute and storage) needs to be tackled; as sending these data back to the remote back-end servers would not only require excessive bandwidth, but may also result in excessive load on the backhaul, thereby increasing the network management overhead and compromising the service-level objectives of diverse vehicular safety applications.

The emerging and promising paradigm of software-defined networking (SDN) indicate a possible solution to these vehicular networking challenges. SDN has been conceived and subsequently deployed for wired networks. However, as of late, there is a rapid shift of interest towards deploying SDN for both the wireless and ad hoc domains. This has, in turn, stimulated the interest of the academic community to look into the possibility of designing SDN-based vehicular networks that would not only enable secure and high bandwidth communication services, but may also provide low latency for the safety-critical vehicular applications. SDN de-couples the control plane from the data plane, and the overall management and orchestration of network resources is carried out via a logically-centralized programmable controller. This, therefore, facilitates enabling a vendor-independent control of the entire network for both network carriers and enterprises, in turn considerably simplifying the network design and operations and laying out the foundations for a highly flexible and programmable networking infrastructure. Hence, given a programmable SDN controller, it is easier to configure disparate network devices and to deploy a wide array of new applications instantly. Nevertheless, despite of several advantages that SDN brings to a networking infrastructure, it is also vulnerable to a number of security attacks since malicious entities may launch attack on either the data plane via targeting the network elements from within the network itself and via the southbound application programming interface (API), by directly attacking the control plane as it acts as the centralized point of intelligence for the entire underlying network or on the applications plane by targeting certain specialized applications and via the northbound API. However, ensuring security in a SDN-based network remains beyond the scope of this article.

Although a number of architectures have been recently proposed for guaranteeing an enhanced network resource management in VANETs (kindly refer to Section 3 for details), most of them have not accounted for the unique VANET-associated features and characteristics in their designs, i.e., frequent changes in network topology owing to the highly dynamic behaviour of vehicles in the data plane, extremely large and distributed network, stringent delay constraints, the need for efficient and smooth handovers, etc. Moreover, a number of these architectures primarily rely on accumulating the centralized intelligence in a single centralized SDN controller, which on the one hand, provides a global view of the entire underlying network, but on the other hand, may become a single point of network failure in case of any unforeseen event. Thus, a re-design of the existing vehicular networking architectures is highly indispensable.

Accordingly, this article is one of the first few research studies to bring forth the notion of a highly reconfigurable software-defined heterogeneous vehicular networking (SDHVNet) architecture to facilitate rapid network innovation for meeting the stringent performance requirements of diverse safety-critical vehicular cooperative applications and services. SDHVNet is a robust and performant
next-generation heterogeneous networking architecture for designing intelligent transportation systems (ITS). In contrast to the existing architectures proposed in the research literature, centralized intelligence is augmented with the localized intelligence to avoid a single point of network failure. The remainder of this article is organized as follows. In Section 2, we outline a brief background of vehicular networks, analyse the key radio access technological candidates for vehicular communication along with their potential and limitations and discuss the need for heterogeneous networking. Section 3 summarizes the current state-of-the-art in the context of ITS. Section 4 depicts our proposed hierarchical and logical architecture for the envisaged SDHVNet. Six key design challenges, together with their probable solutions, in the context of the deployment of safety-critical applications on such SDHV Nets are also deliberated. Finally, conclusions are drawn in Section 5.

2. Background and Motivation

Vehicular networking is one of the key technologies that caters to the realization of a variety of the aforementioned vehicular safety applications, i.e., forward collision warnings, emergency vehicle assistance, vulnerable pedestrian collision mitigation, blind intersection warnings and hazardous location alerts. These applications thus allow for a collection and dissemination of useful contextual information between the vehicles (vehicle-to-vehicle (V2V) communication), among the vehicles and infrastructure (vehicle-to-infrastructure (V2I) communication), among the vehicles and supporting network (vehicle-to-network (V2N) communication) and between the vehicles and vulnerable road pedestrians (vehicle-to-pedestrian (V2P) communication), thereby strengthening the basis for the promising paradigm of vehicle-to-everything (V2X) communication, as depicted in Figure 1. A secure and low-latency communication between the vehicles and among the vehicles and the supporting infrastructure and network is quite critical to the successful implementation of such applications. V2X communication makes vehicles an integral part of the Internet of Things (IoT) landscape [5]. Accordingly, the emerging yet promising paradigm of the Internet of Vehicles (IoV) has also recently started taking its place in the research literature [6,7].

![Figure 1. Towards seamless, ubiquitous vehicle-to-everything (V2X) communication.](image)

Over the past several years, the IEEE 802.11p/DSRC has been considered as the de facto standard for the implementation of numerous vehicular networking applications and services. IEEE 802.11p/DSRC is considered as a short-range wireless technology that originally evolved from the WiFi standard and primarily operates in a 5.9-GHz ITS bandwidth [8]. While DSRC provides a fast two-millisecond over-the-air latency, its standard performance degrades to a significant extent in urban scenarios with abundant high-rise buildings and intersections, leading to considerable blockage in the line-of-sight communication. Other limiting factors include fading, the high mobility of vehicles, and uncoordinated medium access mechanisms [9]. On the contrary, the Third-Generation Partnership Project (3GPP) has recently promulgated the notion of C-V2X, i.e., cellular vehicle-to-everything communication,
a technological paradigm using existing and developing cellular standards for a diverse array of vehicular connectivity applications and use-cases [10]. C-V2X is currently being developed as part of the 3GPP objectives to accelerate the development of cellular systems from 4G to 5G by incorporating enhancements to LTE Broadcast and LTE Direct. LTE Broadcast would facilitate both V2I and V2N communication by leveraging traditional cellular infrastructure, wherein messages can be broadcast from V2X servers to numerous vehicles concurrently, while the individual vehicles can unicast the messages back to the server [11]. Enabling V2I and V2N communications is enormously advantageous for several vehicular applications, i.e., receiving alert messages from the traffic management authorities warning of traffic accidents and conditions several miles ahead up the road or communicating with a smart parking facility to locate and reserve the available parking space automatically. LTE Direct would enable robust V2V communication with a low latency of about one millisecond, at distances of up to hundreds of meters and, more notably, both in-coverage and out-of-coverage of the traditional cellular infrastructure [12,13].

However, the aforementioned technologies are not yet capable of supporting a gigabit per second data rate for sharing of onboard raw sensor data (i.e., from visual cameras, radars and LiDARs) between the vehicles and with the infrastructure [3]. Automotive cameras are typically responsible for generating a considerable proportion of sensor data on the vehicles, and the required data rates are typically around 100 Mbps and 700 Mbps for low- and high-resolution raw images, respectively, after significant compression has been applied [14]. Practically, the maximum data rate for DSRC is only around 6–27 Mbps, while 4G cellular systems are still limited to approximately 100 Mbps in high mobility scenarios, though much lower data rates are typical. In this context, millimetre wave communication (mmWave) remains a pivotal approach for realizing the aim of higher bandwidth next-generation connected vehicles. The mmWave band has already been rolled out in the market in the form of the IEEE802.11ad and supports a data rate of 7 Gbps [15]. There are substantial challenges, i.e., lack of accurate mmWave vehicular channel models, insufficient penetration rates and beam alignment overhead, that still prove critical in realizing the full potential of mmWave V2X communication systems. However, it can still prove attractive for a number of powerful vehicular safety applications such as the bird’s eye view and see-through highlighted in the 5G Automotive Vision of the 5G Public Private Partnerships Group (5G-PPP Group) [16]. Hence, a heterogeneous combination of diverse wireless technologies appears to be one of the most viable options for next-generation ITS communication platforms so that the advantages of one technology reasonably offset the disadvantages of the other. Table 1 depicts the salient characteristics of candidate networking technologies that can match the challenging requirements of the diverse vehicular networking applications.

Heterogeneity is also supported in the 5G Vision [17] promulgated by the 5G-PPP Group, which regards the future 5G networks to be a heterogeneous set of air interfaces comprise of both existing and future wireless networking technologies (especially as terahertz communication is currently being explored for vehicular networking [18]). Seamless handovers among heterogeneous technologies (vertical handovers) are also a native feature of the 5G-PPP’s 5G Vision. Hence, heterogeneity can help achieve better network performance guarantees. Nevertheless, heterogeneity itself is an intricate task to handle and leads to network fragmentation and inefficacy in network resource utilization. Furthermore, transitioning from one radio access technology to another and the multi-hop process involved in the routing of the network traffic could add to the overall end-to-end delay and needs to be carefully tackled. Especially in the case of dense vehicular environments where resource demand is particularly high and several network routing paths are available, there is a need to look for optimal paths within the shortest possible time. This could be addressed with the help of intelligent routing algorithms and via efficient network resource management. The emerging paradigm of SDN proposes a possible solution to these networking challenges by providing an intelligent orchestration of the network through its salient characteristics of reprogrammability, agility, scalability, elasticity and flexibility. An illustration of a heterogeneous vehicular networking architecture is depicted in Figure 2.
Table 1. Potential wireless technological candidates for vehicular communication.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>WiFi (802.11)</th>
<th>DSRC (802.11p)</th>
<th>LTE</th>
<th>LTE-Advanced</th>
<th>mmWave (802.11ad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Range</td>
<td>100 m</td>
<td>1 km</td>
<td>100 km</td>
<td>100 km</td>
<td>10 m</td>
</tr>
<tr>
<td>Maximum Bandwidth</td>
<td>20 MHz</td>
<td>10 MHz</td>
<td>1.4, 3, 5, 10, 15, 20 MHz</td>
<td>100 MHz</td>
<td>2.16 GHz</td>
</tr>
<tr>
<td>Connectivity</td>
<td>Intermittent</td>
<td>Intermittent</td>
<td>Pervasive</td>
<td>Pervasive</td>
<td>Intermittent</td>
</tr>
<tr>
<td>Capacity</td>
<td>Average</td>
<td>Average</td>
<td>Very High</td>
<td>Very High</td>
<td>Extremely High</td>
</tr>
<tr>
<td>Frequency Band</td>
<td>2.4 GHz, 5.2 GHz</td>
<td>5.9 GHz</td>
<td>700–2690 MHz</td>
<td>450 MHz–4.99 GHz</td>
<td>30–300 GHz</td>
</tr>
<tr>
<td>Peak Data Rates</td>
<td>54 Mbps</td>
<td>6–27 Mbps</td>
<td>300 Mbps</td>
<td>1 Gbps</td>
<td>7 Gbps</td>
</tr>
<tr>
<td>Support for Mobility</td>
<td>Low</td>
<td>Moderate</td>
<td>Very High (350 km/h)</td>
<td>Very High (350 km/h)</td>
<td>Low</td>
</tr>
<tr>
<td>V2V Connectivity</td>
<td>Ad hoc</td>
<td>Ad hoc</td>
<td>No</td>
<td>D2D</td>
<td>Ad hoc</td>
</tr>
<tr>
<td>V2I Connectivity</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Market Penetration</td>
<td>Very High</td>
<td>Low</td>
<td>Very High</td>
<td>Potentially High</td>
<td>Low</td>
</tr>
</tbody>
</table>

Figure 2. A heterogeneous vehicular networking architecture.
3. The State-of-the-Art in Intelligent Transportation Systems: An Overview

A brief glimpse of the research literature reveals that a number of research studies have surveyed the potential challenges and limitations for devising an efficient ITS. In [19], the authors presented a comprehensive overview of the LTE-based V2X standardization activities in terms of their scope, probable use-cases, and associated service requirements. Challenges of dense vehicular environments and higher mobilities along with numerous technical design considerations have also been addressed. A survey of heterogeneous vehicular networks outlining research issues, challenges and solutions pertinent to heterogeneity at both the medium access control and network layers has been presented in [20]. In [21], the authors outlined a systematic investigation of existing vehicular communication systems in terms of (their) potential benefits, limitations, diverse vehicular applications and system requirements and proposed a layered-5G vehicular networking architecture comprised of a generic cloud layer, a core network cloud layer, a radio access network layer encompassing diverse radio access networks and vehicles and roadside units’ space. Furthermore, a study of automotive sensing technologies employed for active safety measures has been surveyed in [3] and opined 5G mmWave communication as the only viable option for high-bandwidth connected vehicles.

In [22], a brief survey of both academic and industrial advances for realizing the notion of IoV has been presented along with a debate on the potential challenges and research issues in the implementation of the V2X connectivity. Furthermore, a survey deliberating on the state-of-the-art vehicular localization techniques, their performance and applicability to autonomous vehicles has been presented in [23], wherein the authors primarily focused on sensor-based technologies (GPS, inertial motion units, cameras, radars, LiDARs and ultrasonic sensors) to determine the position of vehicles on a specified coordinate system and employed cooperative techniques (i.e., V2V and V2I communication via several wireless communication technologies) in order to enhance the locational accuracy and reliability. In [24], the authors investigated the relationship between big data and IoV within a vehicular context and primarily focused on how the IoV facilitates the big data acquisition, ensures a seamless, ubiquitous big data transmission and enhances the storage and computational abilities for the same. It further deliberated on a big data-enabled IoV and evaluated how big data mining could bring considerable advantages to the IoV development in certain aspects, including, but not limited to, network characterization, protocol design and performance evaluation.

Security is also one of the indispensable components in designing a highly efficacious and cooperative ITS and therefore demands careful consideration. A self-contained and systematic survey encompassing security, trust and privacy-related challenges pertaining to VANETs has been presented in [25], wherein the authors outlined several anonymous authentication mechanisms, location privacy protection schemes, trust management models along with their efficacy and various types of network simulators, mobility simulators and integrated simulation platforms. In [26], the authors presented a comprehensive survey of the recent state-of-the-art VANET security architectures, frameworks, security standards and protocols, classification of several critical vehicular security attacks and their probable solutions and challenges that act as the bottlenecks in the evolution of secure ITS architectures along with future research directions. It is also highly pertinent to mention that the recent research focus has shifted from the conventional cryptography-based security solutions (i.e., based on the certificates and public key infrastructures) to a number of trust management schemes since: (a) vehicles in a vehicular network are highly dynamic in nature and are randomly dispersed throughout the network; (b) the presence of a seamless networking infrastructure cannot be guaranteed at all times; and (c) a cryptography-based solution could be easily compromised due to insider attacks, which are not only one of the most common security attacks, but are also extremely difficult to detect and handle [27].

Furthermore, research studies evaluating the technical feasibilities and performance analyses of wireless networking technologies supporting diverse vehicular applications have been conducted. A study evaluating the performance of heterogeneous vehicular networks (i.e., comprised of DSRC, LTE and WiFi) for both V2V and V2I communication has been delineated in [28]. An application layer handoff scheme has also been envisaged that not only guarantees optimal utilization of available
wireless technologies, but further ensures minimizing of corresponding backhaul communication requirements. In [29], a signalling game mechanism has been proposed for warranting an always best connected service for vehicles traversing a geographical region equipped with heterogeneous networks. A heterogeneous network with aims to satisfy both safety and non-safety communication requirements of autonomous driving has been delineated in [30]. The study presented an enhanced protocol stack and also conceived the communication messages indispensable for supporting autonomous driving vehicles. Furthermore, a multi-tier heterogeneous adaptive vehicular networking architecture so as to ensure reliability and low latency for safety-critical message dissemination in a vehicular networking environment has been presented in [31]. The said architecture integrates LTE and DSRC technologies for balancing the network traffic via offloading the packet forwarding from the cellular networks. The architecture encompasses both high-tier nodes (i.e., public authority-operated vehicles such as buses, taxis or any other recognized authority’s operated vehicles) and low-tier nodes (i.e., private vehicles). The high-tier nodes broadcast beacons with relevant information via the DSRC, whereas low-tier nodes receiving the beacons are registered with the high-tier nodes and communicate with the infrastructure via DSRC, and not LTE. Hence, all V2V communication takes place via the registered high-tier nodes, which primarily act as message relays. In [32], the authors presented a heterogeneous vehicular networking framework in order to meet the communication requirements of numerous ITS applications and services, along with a comparison of different radio access networks’ candidate technologies. The authors opined that in contrast to DSRC, LTE is suitable for V2I communication, whereas DSRC is much more practical than LTE D2D for V2V communication.

Off late, the emerging yet promising paradigm of SDN has been exploited for vehicular networks. In [33], a brief survey of existing and future challenges of SDN-based vehicular networks has been highlighted. In [34], an SDN-based vehicular communication architecture has been envisaged so as to provide a far more agile configuration capability and to enable rapid network innovation on-demand. Use-cases relevant to adaptive protocol deployment and multiple tenants’ isolation have been highlighted to discuss the advantages of the said architecture in detail. In [35], a scalable and responsive SDN-enabled vehicular networking architecture, facilitated with mobile edge computing, has been suggested to minimize the data transmission time and for improving the quality-of-experience (QoE) of the vehicular users for a diverse range of latency-sensitive applications. In [36], the authors suggested a hierarchical SDN-based architecture for vehicular networks and accordingly developed a communication protocol to address the lack of connection/coordination from the centralized SDN controller. Evaluation of the same was carried out on a real urban mobility scenario.

In [37], an edge-up SDN-based design has been envisaged for vehicular networks in contrast to the traditional cloud-down design typically conceived for mobile ad hoc networks. Emphasis has been particularly placed on the latency control in order to support a diverse range of vehicular applications. In [38], recent research advances of SDN-based vehicular networks have been investigated, and key requirements for ensuring an efficient network resource management were outlined. A taxonomy was also presented in terms of the salient characteristics of software-defined vehicular networks, i.e., radio access technologies, applications and services, network architectural components, opportunities, system components and operational modes. In [39], an architecture supporting the cohesion of both SDN and named data networking (NDN) has been presented to fetch the requisite content within the vehicular networks. It thus assigns a name to the content (instead of the device, i.e., vehicle or infrastructure), and a pull-based communication approach is then used to retrieve the requisite content, as and when desired. In [40], a collaborative vehicular edge computing architecture has been envisaged for facilitating collaboration between the edge computing anchors to ensure scalable and efficacious vehicular applications and services. An abridged (self-contained) summary of the research challenges surrounding next-generation ITS architectures is depicted in Figure 3.
Seamless and ubiquitous communication in an ‘always best connected’ paradigm

Quality-of-service (QoS), Quality-of-experience (QoE)

Efficacious vertical handover schemes

Efficacious network resource management

Ensuring a highly scalable, re-configurable and reprogrammable agile networking architecture

Fog and/or edge computing (Bringing intelligence to network edge)

Intelligent edge-based caching (IoV caching architectures, content popularity, location-specific content, etc.)

Efficacious network resource management

Software-defined networking

Heterogenous networking

Network traffic prioritization (broadcast storm mitigation, network slicing)

High bandwidth and low-latent architectures for connected vehicles

Secure mechanisms for big data collection in large scale IoV

Heterogeneous nature of IoV big data (spatio-temporal in nature – varies in size, volume, and dimensionality)

Highly dynamic topological structures in vehicular networks poses a constraint in acquiring accurate neighbourhood

Data authentication

Data integrity

Data confidentiality

Access control

Data non-repudiation

Anti-jamming

Figure 3. Research challenges surrounding next-generation ITS architectures.
Although a number of architectures have been recently proposed in the research literature for ensuring an enhanced network resource management in VANETs, most of them did not account for the unique VANET-associated features and characteristics in their designs, i.e., frequent changes in network topology owing to the highly dynamic behaviour of vehicles in the data plane, extremely large and distributed network, stringent delay constraints, the need for efficient and smooth handovers, etc. Moreover, a number of these architectures primarily rely on accumulating the centralized intelligence in a single centralized SDN controller, which undoubtedly provides a global view of the entire underlying network, but may also become a single point of network failure in case of an unfortunate event. Therefore, localized intelligence in addition to centralized intelligence is extremely indispensable for realizing the true potential of SDN-based HetVNets.

4. Towards Software-Defined Heterogeneous Vehicular Networks

4.1. Architecture Design for SDHVNet

The SDN paradigm, although primarily conceived for the management and orchestration of conventional data centres, has recently gained the interest of academia and industry. SDN de-couples the data plane and the control plane, and the network intelligence is forwarded to a centralized SDN controller in a move aimed at making simplified, yet intelligent networking decisions. SDN yields a number of benefits to a network’s management, including, but not limited to, reprogrammability, agility, scalability, elasticity and flexibility [41]. However, it is also pertinent to mention that unlike traditional networks, which are administered via a single point of network management, VANETs tend to possess very high mobilities and are quite distributive in their nature. Since the centralized controller in the SDN-based HetVNets plays a critical coordination role with the highly dynamic vehicles and numerous radio access technologies and should it become unresponsive or unavailable, this may transform into a single point of network failure, thereby seriously undermining the benefits of SDN, and could potentially lead to fatal road incidents. Therefore, a distributed networking architecture in addition to centralized governance is indispensable for ensuring the reliability of SDHVNets. A topological architecture of our proposed SDHVNet is depicted in Figure 4.

It can be observed that vehicles tend to rely primarily on the vehicular clouds or localized intelligence (roadside cloudlets), and centralized intelligence is only invoked once the compute and storage resources at the localized level become inadequate. As is evident from Figure 4, vehicles tend to traverse in the form of vehicular clouds. A vehicular cloud (also referred to as a micro cloud or vehicular platoon) is a group of vehicles, just like the cluster in wireless sensor networks, wherein only the cluster head is responsible for communicating the entire cluster’s status to the localized and/or global management entities. This, in turn, assists in minimizing the excessive network management overhead on the backhaul and ensures that the routing or other similar networking decisions could be carried out in almost real-time with ultra-low end-to-end delay. Furthermore, vehicles that do not wish to become a part of the vehicular cloud traverse independently (although not recommended) and thus directly communicate with the local and/or global management entities.

Some of these vehicles possess selfish behaviour, i.e., they do not interact with other vehicles and even do not relay other vehicles’ messages with the aim to conserve their resources for satisfying personal objectives. Provided that the number of selfish vehicles in a network increases beyond a particular threshold, localized intelligence ensures introducing certain incentive-based mechanisms that could entice selfish vehicles to participate interactively in the network by becoming a part of the optimal vehicular cloud. Such incentives may include, but are not limited to, enhancing the reputation (i.e., a metric for evaluating the trust) of a vehicle within the network, access to higher bandwidths, discounts on certain network services, etc. Nevertheless, such incentives are kept within a reasonable limit so that these vehicles, after enjoying the perks, do not hibernate in the selfish mode once again, but rather strive for more and more incentives over time. Game theory-based approaches are generally utilized for such purposes.
Moreover, a proposed logical architecture of SDHVNet is depicted in Figure 5. As can be seen, the network infrastructure plane (i.e., data plane) encompasses both vehicles and vehicular users, roadside infrastructure, i.e., access points and base stations of diverse heterogeneous radio access technologies, traffic lights and vulnerable road pedestrians. V2V, V2I and V2P communication take place at the data plane. Moreover, the southbound application programming interface (API) facilitates communication between the network infrastructure plane and the control plane. OpenFlow is usually one of the most commonly used southbound APIs. Nevertheless, OpenFlow needs to be considerably enhanced in order to meet the challenges of dynamic vehicular networking environments. The control plane being a software platform is responsible for the management of networking functions virtualized from the network infrastructure plane. It thus collects and maintains the status of all SDN switches, creates and retains an up-to-date networking topology and accommodates an up-to-date frequency manager in order to determine the frequency of a requested vehicular application and service (or any other content) along with the particular duration it has been requested. Subsequently, the cache manager ensures an intelligent edge-based caching by employing dynamic cache management policies and cache eviction strategies. Handover decision manager guarantees that the vehicles remain seamlessly connected to the optimal radio access technology in an always best connected mode for satisfying the bandwidth and stringent latency requirements of safety-critical vehicular applications and concurrently preserves precious network resources by mitigating handover failures and unnecessary handovers.

Trajectory prediction is one of the key components of the control plane and primarily forecasts and updates the anticipated trajectories of the SDN switches. Stationary switches can easily be reached via a reliable connection. However, the real challenge is associated with the mobile switches. Hence, vehicles that are associated with the roadside cloudlets could be easily accessed via the control plane. However, vehicles that are temporarily disconnected need considerable attention: (a) trajectories of public transport are generally fixed, and as such, they could be reached via their fixed time schedules; and (b) connected vehicles are anticipated to be equipped with navigational systems, and drivers are generally expected to traverse along the navigational trajectory, so these navigational routes could be considered as the vehicles’ trajectory for traversing through the network. Lastly, the applications plane...
offers a set of vehicular applications and services indispensable for formulating a next-generation seamless, ubiquitous and undifferentiated ITS platform.

4.2. Architectural Challenges

Apart from the aforementioned architectural features of a SDHVNet, there exist several design challenges that significantly hamper its service performance and implementation in highly dynamic vehicular environments. This subsection highlights some of these design challenges and their proposed solutions.

4.2.1. Seamless, Ubiquitous and Undifferentiated Network Connectivity

The prime intent of heterogeneous vehicular networking is to guarantee a seamless, ubiquitous and undifferentiated network connectivity to ensure the stringent quality-of-service (QoS) and QoE of diverse vehicular safety and non-safety applications and vehicular users, respectively. Undoubtedly, both vehicles and vehicular users should always be connected to the optimal radio access technologies, which could ensure ultra-low end-to-end delay (latency), higher bandwidth, enhanced data rates and that also in a cost-efﬁcacious manner. However, the heterogeneity of diverse radio access technologies is an arduous task to tackle, and therefore, intelligent vertical handover decisions are required to ensure seamless mobility. In a heterogeneous vehicular networking environment, vehicles either traverse along the geographical regions of overlapping radio access technologies or different geographical regions with distinct radio access technologies. Moreover, owing to the highly dynamic behaviour of vehicles, they are expected to perform frequent handovers. Nevertheless, too frequent handovers result in the wastage of precious network resources and need to be mitigated. Furthermore, for any vertical handover to transpire successfully, three vertical handover decision mechanisms should be executed with a minimal possible delay: (a) handover necessity estimation—determining whether a particular handover is essential to an available network (or networks); (b) handover target selection—opting for the best network out of the available networks; and (c) handover-triggering condition estimation—determining the precise time instance to trigger handover to the selected network.

Since the centralized SDN controller in the heterogeneous network possesses a bird’s eye view of the entire underlying architecture, it could help to meet the seamless, ubiquitous and undifferentiated network connectivity. This is illustrated in Figure 4, wherein the abstraction of physical radio resources of diverse radio access technologies in terms of their bandwidth, time and location is suggested.

Figure 5. The logical architecture of SDHVNet.
This abstraction refers to the bandwidth resources that are available for a vehicular application at a particular time and location via each of the radio access technologies in the current and anticipated trajectory of the vehicles. Since the SDN controller has global knowledge of the available physical radio resources of each radio access technology’s BS/AP along the vehicle’s anticipated travelling direction, intelligent vertical handover schemes can be accordingly employed for switching purposes.

4.2.2. Heterogeneous Multi-Hop Routing

Vehicles typically disseminate safety-critical information to other vehicles in their immediate vicinity via multi-hop V2V communication or by relaying the messages through roadside units, each possessing a different radio access technology. Nevertheless, these safety-critical messages may get lost if the serving radio access technology fails or next hop becomes unavailable (especially in sparse traffic conditions), thus resulting in a communication breakdown and wastage of precious network resources, which may lead to fatal accidents on the roads. Furthermore, some of these radio access technologies are comparatively more expensive than the others, i.e., LTE is fairly expensive in contrast to WiFi and DSRC. Thus, a highly reliable and cost-eficacious communication via multi-hop routing is indispensable in a VANET context. It is also crucial in restoring the network connectivity, especially in case the serving radio access technology fails, but an overlapping radio access technology is accessible. Likewise, in the case of vehicles outside the coverage range of serving roadside units, the network packets could still be shared (with them) through multi-hop V2V communication by vehicles associated with the roadside units, and centralized SDN controllers with a global view of the underlying network can map the shortest, yet optimal multi-hop routes for the said purpose.

Moreover, SDHVNets via their characteristics of reconfigurable and reprogrammable networking infrastructure can reduce the service costs either related to downloading of a particular application or upgrading of a particular service. Let us assume that 50 vehicles are traversing through a particular geographical region with only LTE coverage and intend to perform service upgrades. If the packet size is 20 MB with the cost of service upgrade as US$ 0.15 per MB, then the entire service upgrade cost would be $50 \times 20 \times 0.15 = 150$. SDHVNets, on the other hand, could facilitate cooperative sharing by enabling 1/3 of the vehicles to perform service upgrades via LTE and the remaining 2/3 to request the same content via V2V communication, thereby cutting down the cost to $50 \times 20 \times 0.15 \times 1/3 + 0 \times 2/3 = 50$. 

4.2.3. Broadcast Storm Mitigation and Network Slicing

The sheer increase in the availability of onboard sensors and next-generation communication platforms facilitate vehicles in disseminating safety-critical messages to other vehicles and vulnerable road pedestrians in their immediate vicinity. Such safety messages are extremely critical in nature and typically include, but are not limited to, emergency vehicle warnings, forward collision warnings, blind intersection warnings, vulnerable pedestrian warnings, blind curve warnings and queue warnings. However, disseminating of these packets in a concurrent manner or to vehicles not requiring the same results in packet collision, choking of the entire network and hence the packet delivery timelines could be considerably delayed, which within the context of vehicular networks, may lead to a number of grave consequences. This phenomenon is referred to as broadcast storming where a sheer amount of traffic is broadcast, in turn consuming the precious network resources and leaving behind fewer resources to transport normal traffic. Broadcast storm could also be triggered by a malicious entity, and the security of the network is also crucial in order to ensure that such entities are not allowed to penetrate the network in the first instance, but if they do so, appropriate recovery mechanisms should be in place to track and subsequently eradicate them from the network.

However, a broadcast storm could be resolved by appropriately slicing the network. Network slicing is a technique whereby multiple tenants are isolated into distinct groups to improve the network efficiency significantly. This is illustrated via a general scenario depicted in Figure 6. The police vehicle $P$ and ambulance $A$ can broadcast messages to vehicles on both sides of the road. However, it is
unnecessary for the vehicle $P$ to broadcast messages to vehicles $V_3$, $V_4$ and $A$, and the same is true for vehicle $A$, which does not need to broadcast messages to vehicles $P$, $V_1$ and $V_2$. The localized SDN controllers could therefore slice the network with respect to the driving directions of vehicles in the network. Thus, the vehicles would only broadcast packets to other vehicles travelling in their respective direction or their anticipated travelling trajectory. SDN-based network slicing is perhaps easy to configure, primarily owing to the software nature of the SDN controllers, is extremely beneficial within a dynamic vehicular context and ensures ultra-reliable and low-latency communication.

![Diagram](image)

**Figure 6.** Depicting broadcast storm and network slicing in a vehicular networking context.

### 4.2.4. Highly Dynamic and Distributed Behaviour of Vehicles

One of the significant challenges for SDHVNets is the highly dynamic and distributed behaviour of vehicles, which makes it extremely difficult for the SDN control plane to maintain the run-time positions of the vehicles and their anticipated travelling trajectories. This, therefore, leads to network management overhead and a substantial amount of end-to-end delay for safety-critical and non-safety vehicular applications. Undoubtedly, the dynamic mobility and the distributed behaviour of vehicles are in fact the root cause for all the vehicular networking challenges and demands careful attention. Furthermore, it is indispensable to disseminate network packets depending on the geographical position of vehicles instead of their IP addresses, and highly intelligent and efficient localization and trajectory prediction mechanisms need to be designed in order to tackle such a challenge. This requires that the intelligence is passed to the edge of the network. Furthermore, a number of localization techniques have been proposed in the research literature for mobile ad hoc networks, and the same could be explored and subsequently optimized for SDHVNets.

### 4.2.5. Mobility-Aware Edge Caching

Another significant challenge in SDHVNets is to formulate an optimal edge caching policy, i.e., for a given anticipated vehicle (or vehicular user) demand, it is imperative to determine which vehicular applications, services and/or content should be placed in each edge cache so as to minimize the average delay for such requests. If such or similar requests are not being catered to by the edge-based caches, fetching them from the remote back-end servers would induce a significantly larger delay, which becomes extremely critical especially in the context of vehicular safety applications. Since the size of the edge caches is considerably smaller in contrast to the global caches (i.e., local caches at most store $10^{-3}$ or $10^{-4}$ of the global cache content), it is quite essential to cache the right content at the right time and then to flush it out so as to create storage space for the next popular content. One possible way is to devise an age threshold scheme that caches all content requested for more than a particular threshold $\tau$, where $\tau$ is the content age. Content popularity could also be ascertained by the Zipf distribution which implies that the collection frequency decreases quite rapidly with rank and could be mathematically formulated as $c_f \propto 1/i$ or $c_f = ci^k$.

Several researchers in academia and industry have already proposed a number of cache management policies, including, but not limited to, leave copy down, leave copy elsewhere, randomly copy one,
probabilistic caching, latency-aware caching, congestion-aware caching and search and PopCache. Furthermore, cache evictions strategies are indispensable as they determine which entries need to be evicted from the cache in order to create sufficient space for the new entries and typically include first-in-first-out, least recently used, least frequently used and RANDOM schemes. Nevertheless, since vehicles traverse at very high speeds and content popularity subsequently changes at a dynamic pace as well, it is quite difficult to estimate the content popularity at any given time and location on an instantaneous basis, and therefore, intelligent dynamic edge-based caching algorithms need to be devised in this regard. In SDHV Nets, the SDN controller possesses sufficient knowledge of a vehicle’s serving edge node (EN) and the time it takes to traverse through the coverage of a respective EN along with the anticipated trajectories and the sequence of ENs in the anticipated trajectories of the vehicles. This thus facilitates the SDN controller to deploy the requisite content in advance on ENs in the anticipated trajectories of the vehicles. Furthermore, it is recommended that the content should be stored on the ENs where the vehicles have a real chance to acquire them, i.e., on slow-moving traffic regions or congested road intersections. Thus, intelligent mobility-aware edge caching architectures could ensure ultra-reliable and low-latency communication, in turn not only meeting the stringent QoS of diverse vehicular safety-critical and non-safety applications, but also guaranteeing the QoE of the vehicular users.

4.2.6. Security

Security is one of the critical concerns in a vehicular networking environment. Over the years, a number of security solutions have been envisaged for the VANETs, which primarily relied on the conventional cryptographic schemes utilizing public key infrastructures and certificates. Nevertheless, cryptographic-based solutions are not feasible for vehicular networks since vehicles are highly dynamic in nature and are distributed throughout the network, the availability of a networking infrastructure cannot be guaranteed at all times, and traditional cryptography-based solutions are also vulnerable to insider attacks. Hence, trust has been recently introduced as an alternative for ensuring security in vehicular networks. In trust-based schemes, vehicles communicate and disseminate safety-critical messages with other vehicles based on trust (i.e., the confidence of one vehicle on the other) and constitute both direct trust and indirect trust. Direct trust is a vehicle’s direct observation about the target vehicle, whereas indirect trust is computed by seeking recommendations from the one-hop neighbouring vehicles in the vicinity of the target vehicle. It is pertinent to mention that each one-hop neighbour that furnishes its recommendation has a different context (both conditions and capabilities), and thus, its recommendation segment should be weighted accordingly. Since the localized controllers have precise knowledge of the vehicles in their coverage area, they can act as a local trust management authority to weigh the individual recommendations and then aggregate them to obtain a trust segment. Moreover, once the trust value of a particular vehicle falls below a particular threshold, the localized controller can term it as malicious and ensure its elimination from the network. It can also intimate the same to the globalized controller for broadcasting this message to other roadside cloudlets and vehicular clouds to ensure that such malicious vehicles do not later become part of any network.

5. Conclusions and Future Directions

Conventional vehicular ad hoc networks (VANETs) are capable of facilitating vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Nevertheless, such VANETs have several inherent shortcomings, including, but not limited to, lower bandwidths, higher end-to-end delays and unbalanced traffic flows. To overcome such issues, in this article, we propose a highly intelligent, robust and performant next-generation heterogeneous networking architecture for ensuring rapid network innovation to meet the stringent performance requirements of diverse safety-critical vehicular applications. Our proposed SDHVNet architecture offers an abstraction of network entities as SDN switches, thus mitigating the inflexibility in the deployment of heterogeneous radio access technologies, and facilitates an efficient orchestration of the network resources. This, in turn, addresses key issues
pertaining to vehicular communication, i.e., guaranteeing ultra-low end-to-end delay for safety-critical vehicular applications, intelligent caching at the network edge, broadcast storm mitigation via efficient slicing of the network, etc.

A considerable number of architectural design issues still need to be investigated. The sheer number of sensors onboard connected vehicles generates a massive amount of data, whose real-time analysis is indispensable in order to ensure a reliable analysis of the traffic conditions on the road, precise behaviour of vehicles and prediction of traffic vis-à-vis its density and throughput per hour, per lane. Therefore, vehicle-to-cloud communication should augment the conventional V2V and V2I communication for such a highly dynamic and distributed networking environment. Furthermore, the deployment of SDN controllers needs to be handled with caution since passing all intelligence to one centralized controller would not only result in a significant amount of network management overhead, but could also result in a single point of network failure. Therefore, intelligence needs to be passed within the network, and especially at the network edge. However, it is pertinent to mention that placing an SDN controller in every roadside cloudlet would result in frequent handovers, which subsequently would lead to wastage of precious network resources. Hence, appropriate SDN placement schemes need to be investigated. Scalability is also a concern in SDN-based heterogeneous vehicular networks since there are a huge number of vehicles in dense vehicular networking environments, and tracking their run-time positions is not only an arduous task to tackle, but also results in a massive amount of management overhead. Structured vs. non-structured and clustered vs. non-clustered routing protocols should be devised in order to address the scalability issues.


**Funding:** The corresponding author sincerely acknowledges the generous support of the Government of the Commonwealth of Australia for funding the research at hand via its ‘International Research Training Program (Allocation No. 2017560)’.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).