Research Article

B5G Ultrareliable Low Latency Networks for Efficient Secure Autonomous and Smart Internet of Vehicles

Sultan S. Alshamrani 1, Nishant Jha 2, and Deepak Prashar 2

1Department of Information Technology, College of Computer and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
2School of Computer Science and Engineering, Lovely Professional University, Phagwara, India

Correspondence should be addressed to Deepak Prashar; deepak.prashar@lpu.co.in

Received 8 June 2021; Revised 23 July 2021; Accepted 21 August 2021; Published 15 September 2021

Academic Editor: Haruna Chiroma

Copyright © 2021 Sultan S. Alshamrani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, 5G and beyond 5G (B5G) systems, Ultrareliable Low Latency Network (URLLC) represents the key enabler for a range of modern technologies to support Industry 4.0 applications, such as transportation and healthcare. Real-world implementation of URLLC can help in major transformations in industries like autonomous driving, road safety, and efficient traffic management. Furthermore, URLLC contributes to the objective of fully autonomous cars on the road that can respond to dynamic traffic patterns by collaborating with other vehicles and surrounding environments rather than relying solely on local data. For this, the main necessity is that how information is to be transferred among the vehicles in a very small time frame. This requires information to be transferred among the vehicles reliably in extremely short time duration. In this paper, we have implemented and analyzed the Multiaccess Edge Computing- (MEC-) based architecture for 5G autonomous vehicles based on baseband units (BBU). We have performed Monte Carlo simulations and plotted curves of propagation latency, handling latency, and total latency in terms of vehicle density. We have also plotted the reliability curve to double-check our findings. When the RSU density is constant, the propagation latency is directly proportional to the vehicle density, but when the vehicle density is fixed, the propagation latency is inversely proportional. When RSU density is constant, vehicle density and handling latency are strictly proportional, but when vehicle density is fixed, handling latency becomes inversely proportional. Total latency behaves similarly to propagation latency; that is, it is also directly proportional.

1. Introduction

Connecting vehicles with each other and the infrastructure around them could play a significant role in autonomous vehicles’ future and improving safety. Vehicles can communicate with one another, respond to traffic signals, and even see around corners. 5G is on the way to making this a reality, delivering the higher speeds and more space needed for smart traffic control systems and fully autonomous vehicles. However, autonomous vehicles are not the only way to connect cars; they compete with Wi-Fi technologies called Dedicated Short-Range Communication (DSRC) or Cellular Vehicle to Everything (C2VX) that will eventually use the 5G networks. Vehicles connected to cellular networks are becoming accessible and more common due to recent telecommunications industry developments. Indeed, because of critical applications such as autonomous vehicles, automotive monitoring, traffic control, and traffic management, the network’s link will be inseparable from future vehicles communication systems. Vehicles would be able to communicate among themselves or with the network infrastructure via this link to exchange vital data, and accidents will be avoided, and lives can be saved by using them. A crucial piece of information for safety applications is the location or, in general, the vehicles kinematic condition [1]. Together with the evolution of radio and transport
technologies, a profound approach has been established. The management and maintenance of these networks have been active in the revolution, to achieve high versatility, from manually built and programmed infrastructures to systems that can self-manage themselves, equipped with advanced intelligence. Many of these skills have been accepted by all of the new criteria. The so-called 5G architecture is currently the latest revision of the mobile network specifications, based on the 15th release published in 2019 by the 3GPP Consortium [2]. According to this scenario, research activities in the field of mobile networks are an important activity for opening the way to new requirements and satisfying customer and operator requests [3]. Operating a network infrastructure is a vital activity that requires a high degree of accountability when information goes through the device. Over the last 20 years, general network monitoring solutions have been established to obtain this type of information, focusing in particular on big data center network infrastructures. On the other hand, no standard for the retrieval of real-time monitoring data from the mobile network has been defined by 3GPP [3]. In addition, only aggregated counters are available, which provides information at an inadequate frequency for real-time operations.

Automation is one of 5Gs main drivers [4] and through the definition of 5G self-organizing networks are supposed to reduce the life-cycle expense of the infrastructure, as observed by operators who have implemented it in Long-Term Evolution (LTE) networks [5]. Automation in the next-generation mobile networks would also presume that a strategic position requires advanced agents to organize the infrastructure and maintain it. The authors of [6, 7] discussed the importance of Artificial Intelligence (AI) and Machine Learning (ML) for improving vehicles and agents connectivity and quality of service. Ideally, these will be motivated by algorithms from AI [8]. However, AI algorithms need a wide base of data to be trained upon to enable systems to take accurate actions. Thus, the 5G core network’s flow tracking and other metrics are one of the enablers of this technology. The 5G system is an extension of previous systems, but it can be seen as a big technological change that will alter certain paradigms based on conventional mobile networks.

Since 5G wireless networks must be capable of addressing the problems encountered by 4G networks, such as higher bandwidth, lower end-to-end latency, high data rate, large device connectivity, and consistent quality of experience provisioning [9, 10], hence, they must have the potential to address these issues. The proposed general 5G infrastructure is built on the interconnection of numerous new technologies, such as Massive MIMO networks, Cognitive Radio Networks, and Mobile and Static Cell Networks [9]. Traditional performance metrics, such as spectral quality and network bandwidth, must be increased due to the continued advancement of 5G technologies, and a wide range of connectivity modes must be offered to increase customer experience [9]. Ultrareliable Low Latency Communication (URLLC) [11] is one of the most important 5G use cases. URLLC is expected to play a key role in providing networking for emerging technologies like autonomous vehicles, smart factories, and so on [11]. URLLC is a 5G New Technology with stringent latency and reliability [12]. Due to hard latency conditions, URLLC traffic is usually scheduled on top of ongoing enhanced Mobile Broadband (eMBB) transmissions and cannot be queued [12]. Because of the advent of multiservice networking [13], beyond 5G networks (B5G) or 6G systems have significant interest, and the reliability and latency considerations in 6G may be use case specific [13]. B5G is also required to accommodate ultralong battery life, eliminating the need for charging devices [14]. The advent of multiservice technologies will be useful for improving network intelligence due to an improvement in network complexity [13].

To make autonomous vehicles (AVs) work efficiently, the V2I system must be fast, capable of exchanging messages without any delay, and capable of working under low latency. Current techniques are not capable of offering a reliable environment for faster exchange of information between the AVs. The architecture suggested in this paper is based on Multiaccess Edge Computing (MEC) architecture on baseband units (BBUs), which will allow the processing of tasks to be performed by AVs locally without depending upon the remote cloud servers. Current studies based on remote cloud servers are having various research gaps, such as small cell base stations having restricted resources for computation, and they can be overloaded easily. Furthermore, the quality-of-service (QoS), end-to-end latency management is very challenging without effective computing resource management.

The major contributions of this paper are as follows:

(i) Analysis of the literature, identification of challenges of the existing works, and techniques to resolve them
(ii) Examination of the architecture based on mobile edge computing (MEC) running together with Virtualized Radio Access Network (vRAN) services on Edge Servers improving the efficiency and security of autonomous vehicles and enabling a 5G-based URLLC networks for smart Internet of Vehicles (IoVs)
(iii) Testing of the suggested architecture to show how MEC enabled autonomous vehicles is more efficient and secure

The rest of the paper is organized as follows. The related work is covered in Section 2. In Section 3, the challenges of 5G-based autonomous vehicles are discussed, as well as the importance of the work. Section 4 covers the research significance. The combination of MCC and MEC is defined in Section 5. The MEC-based BBU architecture for autonomous vehicles is explained in Section 6. The conclusions are explained in Section 7, limitations are mentioned in Section 8, and the article is concluded in Section 9.

2. Related Work

5G comes with a few new interesting technologies that may be of great use in the remote control and Industrial Internet of Things. Ultrareliable Low Latency Communication
(URLLC), enhanced Mobile Broadband (eMBB), and massive Machine Type Communication (mMTC) are three of these technologies [15]. URLLC stands for UltraReliable Low Latency Communication and is one of the many advantages that 5G can have. URLLC is projected to provide one-digit millisecond cycle times [16] and 99.99999% efficiency of transmission over one square kilometer, while multiple gNB (next-generation node B) provides the signal [17]. Combining with the ability to mount smaller gNBs on rooftops made it possible for urban environments to have these low latencies over wide areas. According to researchers Kim et al., more base stations would also make it possible to reduce access times and enable support for more devices [18]. Earlier on ether-net-based networks, stable low latencies were most prevalent, and even there, they required TSN (Time-Sensitive Networking) for the versatile and linked delay requirements [19]. As it has the capacity to carry large payloads at high transmission rates [19], eMBB is similar to 4G LTE. The eMBB speeds are in Gigabit per second (Gbps) range and are ideal for heavy network use.

Traffic, such as video content handling, is a type of job that suits eMBB well as every day more video data is generated and consumed, which places a large load on today's networks. eMBB can use high-bandwidth channels, resulting in higher usable bitrates and is thus more suitable for handling video traffic [20]. Another 5G functionality that targets massive IoT (Internet of Things) is Massive Machine Style Communication or mMTC for short. The use case for mMTC is the provision of network connectivity for many devices that communicate over long distances through short messages [19, 21]. In general, IoT devices are not reliant on reliability and data rate but rely on the ability to communicate over long ranges instead. According to [22], a single technology for remote operation is not appropriate. For accurate monitoring, low latency from URLLC is needed, eMBB is required to transmit the vehicle's view, and mMTC is suitable for sensor readings and similarly small intermittent data transmissions. Other researchers are investigating if 5G is to be used along with the upcoming V2X [23] (vehicle-to-everything) standard. In terms of latency, reliability, and range, the researchers proposed specifications similar to what URLLC can provide, but with the bit rates that the only eMBB can deliver [24]. Remote operating vehicles have been used in recent years to perform different kinds of work. Three types of vehicle design trends have contributed to these jobs: exploration rovers, Unmanned Ground Vehicles (UGV), or hazardous duty machines [25]. Therefore, it is necessary to research the field of remote vehicles to collect information on suitable vehicle designs and input devices for vehicle control.

In [26], the authors have developed an industry-service classification of autonomous vehicles based on 5G. In order to extract the advantages of the next generation of cellular networks for positioning, several works have been done. The authors of [27] propose a 5G-based radio network architecture that combines various radio access technologies and cloud-based radio access network functionalities to provide a stable and privacy-preserving CAV network. The autonomous driving requirements of high precision and low latency can be met by 5G-based positioning techniques. However, cellular systems are not intended to maintain LOS for the UE all the time [1]. A simplified model for the probability, frequency, and length of blockage in mm-wave cellular systems was proposed by Jain, Kumar, and Panwar [28], and in their paper, they explained that the design of mm-wave networks can often be motivated by blockage rather than capacity requirements. As described above, the use of INS could support full-form GNSS in the failure times and during the outages, it could be able to compensate for high errors of 5G-based positioning systems.

In [29], the authors surveyed the field of video streaming over wireless technology and explained how to calculate users’ quality of experience (QoE) in both subjective and objective ways. They explained how encoding works and how the outcome of the encoded video can be calculated. When it was streamed over an unreliable mobile network, they also presented the encoded video product. They believe that there is a trade-off between precision and computer power by using QoE metrics and that simplified QoE metrics on cellular networks are less taxing. This paper contributes to the research by sharing the peak signal-to-noise ratio (PSNR) equation as a metric for measuring streamed video signal loss in order to determine which wireless technology is best for remote service. The 5G environment is designed to accommodate multiple use cases [30], and very versatile and intelligent network architecture is needed to serve all of them. The 5G framework is designed to take full advantage of Software Defined Networks (SDN) and network functions virtualization (NFV) to achieve this purpose, incorporating it into a special IP-oriented physical infrastructure system [31]. All modern networks (e.g., longer routes or packet losses) are expected to incorporate certain self-healing systems in order to maximize the resources and to avoid bottlenecks and other network problems, which can track the resource status and act accordingly, directing the implementation of new infrastructures towards the principle of SDN [32, 33].

3. Challenges for 5G-Based Autonomous Vehicles

The future of mobility will benefit greatly from autonomous driving technologies. This allows us to focus on our jobs rather than the stressful job of driving, and it aids in the elimination of human mistakes, enhancing response times, increasing traffic flow quality, and lowering the incidence of road injuries [34]. URLLC is the most recent 5G service tier, targeted at mission-critical communications with a target latency of 1 milliseconds, end-to-end security, and 99 percent reliability [34]. This type of wireless communication technique will be ultrafast and ultrareliable in autonomous driving, which helps enable real-time communication between the vehicles (V2V communication) and its roadside environments (V2I Communication). A brief comparison between 4G LTE and 5G is shown in Table 1.

Although various researches have been done for securing the vehicular networks, some of the major challenges, such as security, privacy, and efficient resource management,
3.1. Security Challenges. V2I and V2V services enable 5G vehicles to communicate with the core network and with other vehicles [35]. Because of the large-scale M2M communications, efficient and reliable mobility management is a major challenge. Several studies have identified general security services for cooperative vehicular systems, but IPv6 integration has not been well performed. In [36], the authors have used Internet Protocol Security (IPSec) and Internet Key Exchange version 2 (IKEv2) for securing Internet Protocol Version 6 Network Mobility (NEMO) in vehicular communication and tried to resolve the challenge for securing the vertical handover condition between 3G and 802.11p. Safe mobility control schemes are currently unable to effectively accommodate group-oriented collaboration scenarios. Cooperative driving is a new technology of 5G vehicular networks that enable autonomous vehicles to travel in platoons to save fuel and reduce the risks associated with driver errors.

Falsification, covert falsification, Sybil assault, emergency braking obstruction, and vehicle location hijacking [37] are just some of the attacks that can damage the V2V service and cause serious road accidents. Message verification methods can be used to counteract these attacks. The batch verification technique is still in use for authentication, but the main problem is determining which signatures are invalid. A highly efficient group testing technique was suggested for the identification of invalid signatures with fewer batch verifications [35]. Forged identity, forged venue, and any forged occurrence that can raise the likelihood of road collisions are all possible attacks [38] that the sender of an update may initiate. To protect the credibility of the communications, necessary countermeasures should be taken to combat these assaults. Potential attacks [38] which the sender of an update launches may include a forged identification, forged location, and any forged event, which may increase the risks of road accidents. Necessary countermeasures should be taken to overcome these attacks to ensure the integrity of the messages.

3.2. Privacy Challenges. Most of the applications for VANETs are dependent on the periodic broadcasting of the beacon messages by vehicles [39]. This message contains the real identity, status of the vehicles, and timestamp. Exchanging information cooperatively between the vehicles and other roadside entities can help in avoiding a collision. However, there’s a major privacy threat for vehicles as their states’ information and location in a broadcasted message could be collected and tampered. If a malicious party has access to the passengers’ records, it is extremely risky. Furthermore, combining IoV and social networks will aid in improving vehicle safety by giving vehicles social attributes [40]. This makes the passengers in the autonomous vehicles anonymous to each other before cooperation connected through wireless connectivity, unlike the traditional online social networks. In this case, the major challenge is the exploration of the efficiency of common attributes for cooperation among autonomous vehicles in proximity. In certain cases, there’s also a risk of disclosing passenger personal details to the general public. As a result, it is important to safeguard passengers’ personal details. Some critical systems, such as autonomous vehicles, must report high precision real-time map updates and face problems, such as the need for information to be validated with the help of a server, which will guarantee the message’s accuracy ahead of time due to computing and storage space limitations. Thus, some of the vehicles’ key information like location is required by the servers for comparing this information for confirming the authenticity of the messages and determining whether the traffic information uploaded by the vehicles in the same area is consistent. The server is unable to acquire detailed vehicle details. As a result, a more reliable and effective multiparty set intersection protocol enabling big data processing is needed for privacy-preserving data sharing.

3.3. VNG Management and Resource Allocation. Due to a large number of autonomous vehicles, new challenges in 5G-SDVN are posed by VNG management [41]. The huge scale of VNG is advantageous for improving services in VNG because it allows for the sharing of newer content while still allowing for a large management overhead. As the size is reduced, all shared content and available capacity are limited, threatening normal services and negatively impacting customer satisfaction. Normally, the network contains a few isolated VNGs [41]. VNGs usually have no direct contact with each other, which makes the passengers in the autonomous vehicles unable to acquire detailed vehicle details. As a result, a more reliable and effective multiparty set intersection protocol enabling big data processing is needed for privacy-preserving data sharing.

### Table 1: Comparison between 4G LTE and 5G.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>4G LTE</th>
<th>5G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (mm-waves)</td>
<td>Low (600 MHz–2.5 GHz)</td>
<td>High (24 GHz–52 GHz) and (62 GHz–82 GHz)</td>
</tr>
<tr>
<td>MIMO</td>
<td>High</td>
<td>Small</td>
</tr>
<tr>
<td>Duplex nature</td>
<td>Half duplex</td>
<td>Full duplex</td>
</tr>
</tbody>
</table>
4. Research Significance

For handling a range of resources in VNGs, 5G URLLC has more scalability. Controllers can be used by managers to allocate new management policies to any switch due to the highly efficient reconﬁgurability and programmability of network equipment, which helps to improve network management. An efﬁcient cooperation among the vehicles is encouraged by adopting and is encouraged by adopting global-aware controllers, enabling unprecedented ﬂexibility of the resource scheduling. Resources available are allocated on demand. According to their requirements and resource capabilities, these resources are shared among the vehicles, thereby improving resource optimization. The proposed architecture takes into account vehicle mobility assistance and topology differences, as well as quality of service (QoS) for different services. This type of architecture restricts the development and deployment of new network features by separating the social plane, control plane, and the data plane and making the network strong and centralized contributing to sustainable development.

Vehicles can benefit from feedback obtained from roadside facilities or other vehicles in order to conduct automatic overtaking, cooperative collision avoidance, and high-density platooning. At smart intersections, cars can connect with traffic signals and other networks, allowing emergency responders and buses to be prioritized [34]. Both of these implementations necessitate a high level of redundancy and strict end-to-end latencies, which can only be provided by a URLLC communication network [34]. Furthermore, using either onboard processing capability or cloud storage would not be adequate for storing and processing the massive amounts of data provided by vehicles from their high-resolution cameras and sensors, as well as achieving a higher level of safety than the best human driver by processing real-time traffic conditions within latency of 100 ms [34]. There were limitations on energy and power constraints onboard computing and storage capacities. GPUs used for low latency processing and inference, for example, have high power consumption needs, which are increased by the cooling load to satisfy thermal constraints, reducing the vehicle’s operating range and fuel performance substantially. Local storage units, such as SSDs, can be ﬁlled with sensor data in a matter of hours [34]. Although on-board processing capabilities can be adequate for passenger-vehicle interactions, they may not be adequate for managing workload between vehicles or between vehicles and infrastructure. In the meantime, long latencies and large bottlenecks in data processing as cloud storage are not an adequate solution for the IoV to connect with intelligent vehicles together [34].

5. Combination of MEC and MCC

In order to get better insights of the work done, this section explains why there is a need for MEC. This section will compare MEC and MCC. MEC, together with the increase in popularity of mobile phones, is the natural progression of cloud technologies. In a network infrastructure, mobile edge computing uses mobile base stations to get cloud computing as close to the mobile device as physically possible [43]. Cloud computing is described by the National Institute of Standards and Technology (NIST) as a model for gaining access to a common pool of conﬁgurable computing services that can be conﬁgured and delivered with minimal management effort and service provider involvement [44]. Cloud storage allows more resources to be shared, resulting in improved performance and lower costs. This model has become well-known, and its exceptional simplicity has enabled a wide variety of applications. Cloud computing is a rapidly adapting model that has the potential to become a viable mobile computing approach. One of the most popular applications of cloud computing is to increase the capacity of mobile devices. MCC is the name given to this one-of-a-kind technology. MCC can supplement mobile devices in terms of data capacity, processing power, and mobility [43]. Mobile devices can attach to the Internet in a variety of ways. Mobile networks, Wi-Fi, and satellite connections, for example, can provide access to the Internet through Internet Service Providers (ISPs). ISPs provide the network infrastructure that routes the connections across the appropriate paths on the Internet in order to connect the mobile user to the cloud controller. Cloud controllers manage the incoming requests from mobile clients and distribute them to the relevant cloud providers. Utility computing, virtualization, and service-oriented architecture were used to create these networks [45]. Furthermore, the word "mobile cloud computing" has another meaning. It envisions a set of nearby mobile devices pooling their resources in order to share them. This model is referred to as an "ad hoc mobile cloud." A mobile application task is spread and processed on the computers that belong to the ad hoc mobile cloud in a shared manner in this model. This model was demonstrated in Virtual Cloud Provider [46] by distributing a Map-Reduce architecture across a variety of mobile devices.

The architecture of the MCC still has challenges. Increased latency, device availability sensitivity, operation reliability, and bandwidth constraints are all costs of connecting to cloud servers. These considerations also limited MCC’s ability to support a wide range of applications. For example, augmented reality or assisted cognition rely on sending streams of sensor data and video to a server with enough resources to process them and produce a near-real-time outcome [47]. As a result, a cloudlet, a third mobile cloud computing vision, was proposed. A cloudlet is a compact, resource-rich, self-managed system that can be deployed on a company’s premises. It is decentralized and locally operated, and it uses LAN latency and bandwidth to serve only a few users at a time. In this model, a mobile device also taps into cloud computing space. In contrast to the MCC architectures described earlier, the cloudlet paradigm proposes bringing the cloud closer to the user by placing a device on the ﬁrst hop of the network [47]. This is beneﬁcial to certain actors. Second, the apps would be more responsive to the end-user, allowing for the deferral of critical applications. Additionally, network carriers may use the cloudlets’ location to store media and ﬁles, reducing latency, and energy consumption on the core network. Finally, since their software can be hosted on cloudlets,
application service providers benefit from increased scalability. Cloudlets are still being researched in academia, but commercial implementations based on the same model have only recently become available. The industry has called this paradigm as mobile edge computing (MEC).

MEC and cloudlets are similar in that they are both located at the network’s first hop, provide storage and computing to neighboring computers, and are accessible by mobile users using wireless connections. A MEC server can be mounted at an LTE macrobase station’s UMTS Radio Network Controller (RNC) or at a multitechnology cell aggregation site (eNodeB). A multitechnology aggregation site [48] manages a range of local multitechnology access points to have on-site radio coverage. MEC, on the other hand, differs from cloudlets in that it is operated by a mobile network carrier, it contains knowledge specific to network providers, and MEC servers are broadly spread and accessible to all mobile devices. MEC servers also provide access to knowledge about location and mobility [49]. MEC began as an Industry Specification Group (ISG) under the auspices of the European Institute for Telecommunications Standards (ETSI).

According to the ISG’s Introductory Technical White Paper [50], MEC is described by being on site, proximity, reduced latency, location awareness, and network context information. MEC’s first benefit is that it is located on site. This ensures that the MEC server is disconnected from the rest of the network and can run independently. In the case of a link loss to the core network, an application operating on a MEC server will be unaffected and continue to operate normally. Mobile devices, which are the basis of information, are often close to MEC servers. Because of their close proximity, MEC servers can function as data aggregators and gather big data and analytics. Since all data collected from crowd sensing apps and Internet of Things (IoT) sensors can be aggregated and preprocessed on a MEC server before being uploaded to a central repository, this feature gains the most from them. As a result, data flow and mobile networks are limited. Both the provider and the creator of the application benefit from reducing bandwidth usage [50]. The MEC architecture [53] is shown in Figure 1.

In Figure 1, Multiaccess Edge Orchestrator or ME Orchestrator (MEO) manages the mobile edge application packages along with resource orchestration across edge DC and selecting the right mobile edge host for instantiation of application with triggering, termination, and relocation with the help of reference points such as MM1, MM9, MM3, and MV1. The MM1 reference point acts as an instantiation triggering agent between the MEO and operation support system (OSS) along with termination of applications in mobile edge system. The MM9 reference point is used for managing the mobile edge applications requested by the UE application. The MM3 reference point between the MEO and ME platform manager is used for managing the application lifecycle, rules, and requirements for keeping track of available mobile edge services. The MV1 reference point is under research and evaluation. However, a few studies have shown that it acts as a connection agent between MEAO and NFVO, associated with the Os-Ma-nfvo reference point and is also called ETSI-NFV.

Another advantage of the MEC server is the reduction in the latency. Reduced latency enables technologies like augmented reality and cloud gaming to react quickly. MEC servers also exchange information about their location as well as low-level signaling data with applications. This allows for location-based applications, analytics, and distinction in terms of network conditions and location of the content served [50]. A Novel MEC-based framework was developed by Nokia. This framework is called the Redundant Array of Cloud Services (RACS) [51]. It is a MEC solution that covers all the elements needed to develop and build apps packaged as virtual machines, which are managed. This is the only practical implementation of MEC to date. Therefore, an efficient and more secure system is needed to be worked upon. Table 2 summarizes the whole paragraph.

6. MEC-Based BBU Architecture for Autonomous Vehicles

Autonomous vehicles will become one of the main members of 5G in the coming years [54]. Vehicular networks (VANETS) are developing as a significant application for 5G services. In 5G VANETS, autonomous vehicles are more reliant on URLLC than traditional vehicles [53, 54]. Vehicles may use information obtained from roadside units (RSUs) or other vehicles to perform automatic overtaking and crash avoidance in autonomous driving [34]. These applications require a high level of stability and latency, which URLLC can only provide. However, since storage capacities are limited by resource and power constraints, using only cloud computing would not be adequate for processing and storing the vast amount of data provided by autonomous vehicles from numerous sensors and cameras within a latency of 100 ms. The developers of studies have explored some 5G vehicular network infrastructure, and our architecture is based on that. The architecture is given in Figure 2.

In Figure 2, it has been shown that MEC is able to improve the idea of a RSU to higher level and work with no strict deployment of 5G components, such as massive MIMO and beamforming [55]. A fairly recent networking approach is SDN, where the network setup and control take place in an environment that is more cloud-like than standard networking [3]. In this approach, network devices are managed by a central authority responsible for managing the various network devices, rather than relying on a distributed configuration per system. This transformation turns the network into a more modular technology, opening the way for programmable and self-organizing networks. An auxiliary network layer to transport called the control plane to transport signaling and management messages to incorporate this technology. In comparison, most of the data traffic flows through the so-called data plane, which is applied on top of the data plane-configured computers [3].

Although traditional networks rely on a strict relationship between hardware and applications, the various features are not linked to the physical devices in an NFV-based network but rather are implemented in general-purpose commodity servers. Consider a typical router system to understand the idea further. The router features are performed on a dedicated,
specialized computer in a conventional network: the router. Instead, these can be performed with the NFV method in a commodity server or, more commonly, in a virtual environment. With respect to conventional networking, this decoupling provides tremendous flexibility, enabling the tenant to scale the various resources of a network accordingly, to the actual load or to the relevant use-case specifications. Thus, by allocating more resources to the physical framework and creating more instances of the desired functions, it is easy to scale and adapt a network’s implementation. Usually, this strategy is applied in strict conjunction with SDN technology: the network becomes a massive, programmable device that can be easily controlled and tailored to the use-case scenario accordingly. It is worth noting that the definitions of the SDN and the NFV do not depend on each other. Indeed, regular network devices equipped with an SDN approach can be deployed and an NFV system without an SDN can be equipped. However, the two technologies will take advantage of each other and, when implemented together, communicate the best functionalities [3].

6.1. Network Slicing. There is no universal definition for network slicing, even though the principle of considering it as the separation of network traffic is accepted by most authors, through various logical networks, all operating on the same physical infrastructure. Network slicing is the enabler of certain main features of the system in the 5G architecture, enhancing scalability and versatility. A portion of the network can require a different collection of physical nodes, with different functionalities installed in a different network infrastructure location. In this case, the number of nodes used to process user packets can be increased. In the case of an eMBB slice, the URLLC slice would be designed to achieve a lower latency (e.g., by assigning nodes to the edge of the network). This method is also of the utmost importance when using Mobile Edge Computing. The Network Slicing Architecture [56] is given in Figure 3.

As previously mentioned, onboard computing capacities may be adequate for handling passenger-vehicle interactions, but they may not be sufficient for managing workload between V2V and V2I. Cloud storage is therefore insufficient

---

**Table 2: Comparison between MCC and MEC [52].**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>MEC</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency</td>
<td>Shorter (around 1 ms)</td>
<td>Longer (around 30–100 ms)</td>
</tr>
<tr>
<td>Energy savings</td>
<td>Satisfies the latency condition and increases the battery life by 30%–50%</td>
<td>Cannot reduce the consumption of energy of IoT devices simultaneously and thus satisfy the latency requirements</td>
</tr>
<tr>
<td>Awareness of context</td>
<td>High</td>
<td>No awareness of context</td>
</tr>
<tr>
<td>Privacy and security</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

---

**Figure 1: MEC architecture.**
Figure 3: Network slicing framework.
for creating a link between intelligent vehicles due to massive data delivery and long latencies [34]. To address this problem, we must install all computation and storage capabilities at the wireless network edge, like edge caching and edge computing, using a MEC network infrastructure that runs on BBU servers at radio access points along the roadsides. Since it is entirely software-defined and reconfigurable on request, a cloud-native BBU server can be adapted for DRAN and CRAN installation. It can also run virtualized RAN services with network feature virtualization, as well as MEC-based software like self-driving. This architecture is useful for network slicing, which enables the URLLC network to operate on the same physical networks as other 5G networks, saving money on bandwidth, and network running costs.

6.2. Optimal CPU Scheduling. All data processing takes place at the edge server, and only the energy used for transmitting at mobile devices is considered. This energy is usually greater than the total energy expended by the coordination chain’s subunits. Let us consider an OFDMA system [57]. Let \( P(t) \) and \( H(t) \) denote transmission power of the \( mth \) UE over the subcarrier \( n \). As the time interval of transmission is fixed, the total energy consumption [57] is given as follows:

\[
A_m(t) = \sum_{m \in X_m} A_{mn}(t)
\]  

(1)

where \( X_m \) is defined as the set of subcarriers assigned to UE and \( A_{mn} \) is defined as the consumption of energy over the subcarrier \( m \). In URLLC, high reliability is a must. As a result, using the most recent advancements in information theory, the maximum achievable rate for a finite block length can be calculated as [57]

\[
R_{mn}(t) = B_{mn}(t) = \frac{K}{\ln(2)} \frac{y_{mn}(t)}{L} Q^{-1}(\beta_m),
\]

(2)

where \( B_{mn}(t) \) is the formula for Shannon capacity [13] given as

\[
B_{mn}(t) = K \log_2 \left( 1 + \frac{H(t)A_{mn}(t)}{\phi K} \right).
\]

(3)

where \( y \) is defined as the spectral density of the noise power. \( K \) is the channel dispersion, \( Q^{-1} \) is defined as the inverse of the Gaussian Q-function, \( \beta_m \) is defined as the maximum block error rate, and \( L \) is the length of the block, which can also be given as \( L = \phi K \). Let all of the computational tasks be offloaded to the MEH. Consider a local communication queue at each UE, which contains the bits to be transmitted to AP, which further enables the amount of the computations to be performed [57]. The arrival of the data can be given as

\[
Q^m_n(t + 1) = \max(0, Q^m_n(t) - \phi R_m(t)) + \alpha_m(t),
\]

(4)

where \( Q^m_n(t) \) is defined as the local queue at time \( t \) and \( \alpha_m(t) \) is defined as the arrival of the new data, which is available for the transmission from the next time, which is an unknown distribution with random variables. During the process, another queue is called a remote queue, which is given as [58]

\[
Q^p_m(t + 1) = \max(0, Q^p_m(t) - \phi R_m(t) x_m) + \min(Q^p_m(t), \phi R_m(t)),
\]

(5)

where \( R_m(t) \) is defined as the CPU cycles per second, which is assigned by MEH to UE \( m \) during time \( t \), and \( x_m \) shows the number of bits in a CPU cycle. To address the Resource Allocation Challenge, which includes UE’s long-term energy consumption, we must incorporate the idea of virtual queues. This can be given as [58]

\[
X_m(t + 1) = \max\{0, X_m(t) + Q_m^{\text{total}}(t + 1) - Q_m^{\text{average}}\},
\]

(6)

where \( m = 1, 2, 3, \ldots, m \).

6.2.1. Rate Allocation Challenge. Since the users in the OFDMA scheme are orthogonal, various issues affect different users. As a result, the problem can be expressed as follows [57] for each user \( m \):

\[
\min \left[ \sum_{m \in X_m} \psi \sum_{n \in X_m} \sum_{m \in X_m} \frac{\phi K Y}{H(t)} \exp \left( \frac{B_{mn} \ln(2)}{K} \right) \right].
\]

(7)

The above equation is subject to

\[
(1) \quad B_{mn} \geq 0, \forall n \in X_m
\]

(2) \quad \sum_{n \in X_m} B_{mn} \leq B_{mn,\text{max}}

(8)

where \( B_{mn} = -2Q^p_m + 2Q^p_m - X_m - \mu_m Y_m \). As \( \sum_{n \in X_m} B_{mn} \) and multiplier \( Y \) are the nondecreasing of \( B_{mn} \). Then, the optimal solution is \( B_{mn} = 0 \forall n \in X_m \) when \( Q_m \geq 0 \). If \( Q_m \) < \( 0 \), then the solution can be given as Lagrangian solution [13]:

\[
\phi = \frac{Q_m \sum_{m \in X_m} \sum_{n \in X_m} \sum_{m \in X_m} \frac{\phi K Y}{H(t)} \exp \left( \frac{B_{mn} \ln(2)}{K} \right) - \sum_{n \in X_m} \eta_{mn} B_{mn} + \lambda_m (B_{mn} - B_{mn,\text{max}})}{\sum_{m \in X_m} \sum_{n \in X_m} \sum_{m \in X_m}}
\]

(9)

where \( \eta_{mn} \) and \( \lambda_m \) are Lagrangian multipliers.

6.2.2. CPU Scheduling at MEH. The second problem deals with the optimization of the scheduling at MEH and can be given as [57]

\[
\min (f_m) = \sum_{m=1}^{M} \phi (X_m + \mu_m Y_m + 2Q^p_m) f_m A_m.
\]

(10)

The above equation is subject to the following conditions:

\[
(1) \quad 0 \leq f_m \leq (Q^p_m/\phi A_m) Y_m
\]

(2) \quad \sum_{m=1}^{M} f_m \leq f_{\text{max}}

There is a linear and optimal solution obtained by the use of simple iterative steps as defined in Algorithm 1. The
algorithm ensures that all virtual queues are mean-rate stable in this situation. The algorithm’s path appears to be as close to the optimal available solution as possible. This is measured in terms of the time it takes to arrive at a stable solution.

In this algorithm, to solve the problem linearly with optimal CPU scheduling, we have adopted the technique of virtual queues [57]. $X_m(t)_m, Y_m(t)_m$ is defined as the virtual queue of UE. $Q_m^\psi(t)_m$ is defined as the nondifferentiability of the maximum function. $A_m$ is defined as the conversion factor, which is used for converting the number of CPU cycles to be processed at MEH into its equivalent bits for adding the length of the two queues of UE [57]. The smaller the values of $A_m$ are, the more computationally intensive the applications would be there, and $F_{max}$ is defined as the computational power of MEHs.

### Algorithm 1: Optimal CPU scheduling.

```
Input: \{X_m(t)_m, Y_m(t)_m, Q_m^\psi(t)_m, A_m\}_m
F_{max}, M
f_{average} = f_{max}, \psi = \{m = 1, \ldots, M\}
While f_{average} > 0 do
(1) m = arg \max_{\psi \in \psi} \{A_m(X_m + \mu_m Y_m + 2Q_m^\psi)\};
(2) f_m = \min \{Qmp/\Phi A_m, f_{average}\};
(3) \psi = \psi - [m];
(4) If \psi = null \rightarrow break;
(5) f_{average} = f_{average} - f_m;
(6) End
```

7. Results and Discussion

For evaluation purposes, we have taken 20 UEs, which are embedded in a wireless framework based on mm-waves with path loss values as given in [66]. We have distributed the users uniformly to an area of 500 m$^2$. We have composed an orthogonal frequency-division multiplexing (OFDM) system of 200 subcarriers/user, with a spacing of 30 kHz. The noise power spectral density is taken as -180dBm/Hz with a transmission time of 20 ms and a block length of 100. The computational power of Mobile Edge Hosts (MEH) is taken as 5.0 × 109 CPU cycles/second. The results obtained are shown in Table 3. A trade-off is plotted, which is found to be increasing along the abscissa from right to left.

Furthermore, the reliability and convergence have also been plotted. The graphs show the boundation imposed on the remote queues on their average long-term lengths. The probability at which the sum of the queue lengths increases is a predefined threshold. The challenges are resolved using a dynamic algorithmic framework that is solved using optimization without having a prerequisite knowledge of the radio channel data arrivals. Through these graphs, we can interpret a fast-converging behaviour and the capacity of the system to adapt in the nonstationary environment. By looking at the transient intervals, in the graphs when the convergence is not achieved, the probability converges quickly to the expected levels. Larger values of $\mu_m$ give a lower convergence time, with a larger variance and vice versa. This can be explained by Figures 4 and 5, respectively.

Now, for analyzing the reliability and latency of the 5G autonomous vehicles, Monte Carlo Simulations [54] were performed with configurations as follows: weight factor is taken as 20 and the length of the road covered by the RSU is taken as 800 meters. The vehicle density [54] is taken as 0.5 vehicles per minute. The message generation exchange rate [54] is taken as 80 messages per second, average service time [54] is taken as 10 milliseconds, and transmission power of the vehicle [54] is taken as 50 dBm. The slot duration [54] is 65 microseconds, noise power density [54] is taken as -180dBm/Hz, and the number of the resource blocks are taken as 20 and we are considering a multiple hop situation [54]. The simulation results are shown in Figure 6–8, respectively.

When the RSU density is constant, propagation latency increases, and when the vehicle density is constant, propagation latency decreases.

From Figure 8, it can be concluded that vehicle density and handling latency are directly proportional to each other, keeping the RSU density fixed, and when the handling latency slightly decreases when the vehicle density becomes fixed for a short time.

Although this graph is somewhat skewed, we may deduce that overall latency rises when vehicle density rises. The overall latency reduces when the vehicle density is fixed.

8. Limitations and Future Scope

In autonomous vehicle technology, vehicles can benefit from the information extracted from their surroundings or roadside units to avoid any accident. To enable a fast message exchange mechanism between the autonomous vehicles and roadside units, we require reliable and low latency techniques which only MEC-based BBU URLLC communication can guarantee. However, there are certain limitations of our study, which would become the base for our future study or for other researchers in this field. AVs work on various sensors, cameras, and techniques, such as LIDAR, and each of the sensors cannot perform fast processing and can cause some delay in exchange of information between the vehicles. Due to the COVID-19 restrictions, our approach cannot be implemented on larger scale and evaluation of image quality using peak signal-to-noise ratio cannot be done to find the best performance of our approach. Our approach can be implemented on small scale and to implement on large scale, it requires high capital investment and time, and the complexity of the system may also increase. Furthermore, some attacks or malicious users can tamper the system, which could lead to accidents. To prevent this, one of the solutions we suggest is integrating 5G URLLC communication with
### Table 3: Output of URLLC.

<table>
<thead>
<tr>
<th>UE</th>
<th>$Q_{\text{average}}$</th>
<th>$\rho_m$</th>
<th>$\Gamma Q_{\text{on}}(t) &gt; Q_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.08 \times 10^7$</td>
<td>0.006</td>
<td>0.0058656</td>
</tr>
<tr>
<td>2</td>
<td>$1.56 \times 10^6$</td>
<td>0.005</td>
<td>0.0038845</td>
</tr>
<tr>
<td>3</td>
<td>$4.01 \times 10^6$</td>
<td>0.004</td>
<td>0.0096214</td>
</tr>
<tr>
<td>4</td>
<td>$7.02 \times 10^6$</td>
<td>0.003</td>
<td>0.0045215</td>
</tr>
<tr>
<td>5</td>
<td>$3.19 \times 10^6$</td>
<td>0.002</td>
<td>0.0054687</td>
</tr>
<tr>
<td>6</td>
<td>$2.22 \times 10^7$</td>
<td>0.001</td>
<td>0.0044568</td>
</tr>
<tr>
<td>7</td>
<td>$2.39 \times 10^7$</td>
<td>0.006</td>
<td>0.0039987</td>
</tr>
<tr>
<td>8</td>
<td>$1.10 \times 10^7$</td>
<td>0.005</td>
<td>0.0028757</td>
</tr>
<tr>
<td>9</td>
<td>$1.00 \times 10^7$</td>
<td>0.004</td>
<td>0.0025647</td>
</tr>
<tr>
<td>10</td>
<td>$1.24 \times 10^7$</td>
<td>0.003</td>
<td>0.0012354</td>
</tr>
<tr>
<td>11</td>
<td>$1.28 \times 10^8$</td>
<td>0.003</td>
<td>0.0042318</td>
</tr>
<tr>
<td>12</td>
<td>$1.12 \times 10^8$</td>
<td>0.007</td>
<td>0.0033320</td>
</tr>
<tr>
<td>13</td>
<td>$2.54 \times 10^8$</td>
<td>0.006</td>
<td>0.0054222</td>
</tr>
<tr>
<td>14</td>
<td>$1.38 \times 10^8$</td>
<td>0.002</td>
<td>0.0036987</td>
</tr>
<tr>
<td>15</td>
<td>$1.55 \times 10^9$</td>
<td>0.007</td>
<td>0.0047785</td>
</tr>
<tr>
<td>16</td>
<td>$1.64 \times 10^7$</td>
<td>0.005</td>
<td>0.0055447</td>
</tr>
<tr>
<td>17</td>
<td>$1.89 \times 10^8$</td>
<td>0.001</td>
<td>0.0042300</td>
</tr>
<tr>
<td>18</td>
<td>$2.01 \times 10^9$</td>
<td>0.002</td>
<td>0.0044798</td>
</tr>
<tr>
<td>19</td>
<td>$1.96 \times 10^9$</td>
<td>0.008</td>
<td>0.0025447</td>
</tr>
<tr>
<td>20</td>
<td>$1.71 \times 10^9$</td>
<td>0.003</td>
<td>0.0039886</td>
</tr>
</tbody>
</table>

**Figure 4: Reliability function curve.**

**Figure 5: Out of service versus time plot.**
Figure 6: Propagation latency with respect to vehicle density.

Figure 7: Handling latency with respect to vehicle density.

Figure 8: Total latency with respect to vehicle density.
blockchain technology, which would secure the cloud infrastructure and prevent system tampering. Further researches are required on improving the system with the help of blockchain technology.

9. Conclusion

Speaking of the device’s latency, all the values calculated in this work were below 1 ms. It could be argued that 1 ms is a short period and, with the general Internet Round Trip Time average, the improvement in this value due to the new features is marginal. Although this is valid in most cases, it is crucial to keep this value as low as possible in certain cases of use envisaged by the 5G. Indeed, in URLLC, the overall Round Trip Time (RTT) of the device must be less than 1 ms, and, thus, even a small increment like the one implemented here can be within the service limits. It is also mandatory to study telemetry’s effect in such situations, studying both optimizations and trade-offs to reduce latency. The tests carried out on the test bed indicate that with appropriate values for the parameters, the core network prototype’s efficiency is not compromised and is therefore a legitimate solution for the implementation of network telemetry in the core network. The rational choices are made regarding the parameters of the system; the assessment should not directly affect the efficiency of the core network services.

Data Availability

Data will be available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported by Taif University Researchers Supporting Project no. TURSP-2020/215, Taif University, Taif, Saudi Arabia.

References

I. K. Jain, R. Kumar, and S. Panwar, “Driven by capacity or...”


