REVIEW ARTICLE

Issues and Challenges Facing Low Latency in Tactile Internet

Tara Ali Yahiya[1]*, Pinar Kirci[2]

[1]*Department of Computer Science and Engineering, School of Science and Engineering, University of Kurdistan Hewler, Kurdistan Region - F.R. Iraq*
[2]*Department of Engineering Sciences, Faculty of Engineering, Istanbul University - Cerrahpasa, Istanbul, Turkey*

***Corresponding author's email:** t.ibrahim1@ukh.edu.krd

Received: 15 April 2019 **Accepted:** 23 May 2019 **Available online:** 20 June 2019

ABSTRACT

Tactile internet (TI) is considered as the next step toward a revolutionary impact on the society; this is due to the introduction of different types of applications mainly the haptic ones that require strict quality of service guarantee especially in terms of latency. This would be a major challenge toward the design of new communication technologies and protocols to provide ultra-low latency. This article discusses the diverse technologies, communication protocols, and the necessary infrastructure to provide low latency based principally on the fifth generation of the mobile network that is considered as the key enablers of the TI. Furthermore, current research direction, along with future challenges and open issues, is discussed extensively.

Keywords: Fifth-generation, Haptic, Mobile Edge Computing, Tactile Internet

1. INTRODUCTION

The increasing number of connected machines, new vertical services, security schemes, and business models need improved performance criteria just such as high connectivity, coverage of wireless communication, ultra-low latency, security, and throughput. The fifth-generation (5G) of mobile networks (5G) is emerged to propose flexible and high-speed networks to fulfill these criteria. The support of millimeter wave, machine-type communications (MTC), mobile edge computing (MEC), network function virtualization (NFV), software-defined networking (SDN), and narrowband IoT are the most important keys enabling for new applications requiring stringent guarantee for quality of service (QoS) parameters. Such new applications are represented by evolved mobile broadband, ultra-reliable low latency communication, and massive MTC Maier *et al.* (2018), Shafi *et al.* (2017), and Parvez *et al.* (2018). 5G attracted great attention of European policymakers due to its pervasive structure. Furthermore, it is expected to be a milestone in productivity growth and economic improvement with its higher capacity electronic transmission infrastructure Andrews *et al.* (2014) and Aijaz (2016).

Access this article online

DOI: 10.25079/ukhjse.v3n1y2019.pp47-58	E-ISSN: 2520-7792

Recently, the IEEE P1918.1 working group has introduced the tactile internet (TI) as "A network or network of networks for remotely accessing, perceiving, manipulating or controlling real or virtual objects or processes in perceived real time by humans or machines. Zhang *et al.* (2018)" as it started to define the framework, application scenarios, and technical assumptions. Specifically, the new applications supported by the TI are industry X.0, Automotive, E-healthcare that includes a wide range of use cases and may require or not low latency depending on whether if they are time critical or not Zhang *et al.* (2018). Enormously low latency together with a high level of availability, reliability, and security are associated and provided in TI Amodu and Othman (2018), Ateya *et al.* (2017).

In this article, we will provide an overview of the 5G mobile network that provides the most important capabilities supporting the TI, especially those related to low latency. This would include the 5G architecture starting from the radio access network (RAN) to the most evolved functionalities introduced in its core network to support the different applications of TI.

This article is organized as follows: Section 2 describes the TI with haptic communication as a principle application. Section 3 provides an overview of the 5G network as a key enabler for TI, and Section 4 presents the challenges associated with TI applications in terms of latency, Section 5 presents the technology enablers for TI, Section 6 describes the most recent progress in the research area to provide low latency, Section 7 discusses the open issues and challenges to be solved in TI, and finally, Section 8 concludes the article.

2. THE TI

Today, ultra-reliable and ultra-responsive network connection is worked on for providing real-time control and physical tactile experiences remotely. With the latest technological advancements, the new era is proposed as TI, and it will achieve a kind of true paradigm shift from content-delivery to skill-set delivery network types Ali-Yahiya (2018). The TI will utilize the traditional wired internet, the mobile internet, and the IoT as the infrastructure of transport from an end-to-end perspective. It will construct a novel internet structure having improved capabilities however with the current fourth generation mobile communication systems, it is hard to ensure the technical needs of the TI, thus 5G mobile communication systems will be used at the wireless edge and as a key enabler for TI due to its automated core network functionalities. Besides, it is expected that TI brings a novel perspective for human-machine interactions with its low latency in constructing real-time interactive systems. Basically, the TI is presented for ensuring lower transit time, high reliability, lower latency, better security, and sufficient capacity as a communication infrastructure. TI is expected to be utilizedin vital aspects of life, such as health, transportation systems, industrial automation, and education with allowing billions of devices' communication with each other simultaneously and autonomously Aijaz (2016).

The TI will ensure better communication and realistic social interactions in many areas especially the one that needs an end-to-end latency of 1 ms which is not reached by current cellular systems or wireless local area network. In addition, the provided end-to-end latency rate is vital for critical TI applications, especially the haptic ones. Indeed, it is difficult to present the whole list of TI applications, the most remarkable of them are given such as gaming, automation in industry, education, autonomous driving, individualized manufacturing, robotics, unmanned autonomous systems, virtual, and augmented reality and health care.

In health care, with the TI, telerehabilitation, telesurgery, and telediagnosis applications can be utilized by patients with a telerobot or with many other telediagnostic tools. Tele-rehabilitation and telediagnosis applications will be utilized by patients remotely. The location of the doctors will not be important for diagnosing a patient. Furthermore, in telesurgeries, telediagnostic tools will be controlled by doctors remotely away from the patient. High precision and better diagnosis will be ensured by TI based health-care technologies with ever-improving telemedical technology.

Lately, robotics became a significant topic for researchers in many environments; however, it includes many challenges and high complexity. Indeed, remote and real-time controlling, and providing synchronous haptic feedbacks of robots are spectacular features of autonomous robotics. For controlling robots, latency in action and reaction times is vital; this is why the speed of light that is essential for guaranteeing the low latency.

In the virtual reality, a haptic virtual environment is provided through TI through audio-visually and the touch sense. In augmented reality, real and computer-generated content is combined and visualized for the user. Further, providing additional information to users improves many assistance systems in education and maintenance. By the help of the TI, the content in augmented reality will be improved from static to dynamic. Thus, a real-time virtual extension will be provided for the user's field of view, and many mistakes will be notified and prevented Ateya *et al.* (2018).

2.1. Haptic Communications

Haptic communication and 5G are essential foundations for the TI. As the haptic devices permit the users to feel, touch and manipulate things over real and virtual fields. So far, the transmission of haptic information has been considered over telepresence systems. Thus improving a standard haptic codecs family is a main challenge in the context of the TI. Including tactile and kinesthetic information, this kind of codec family may be a provider of scalability in the network edge. Furthermore, it addresses a layered structure for haptic data, including multi-modal sensory information Antonakoglou *et al.* (2018).

The TI should consider providing haptic feedback together with visual feedback and audio feedback at the master domain.

Because different sensory modalities need to be combined in the human brain, so increased perceptual performance is gained. Here, the main problem is cross-modal asynchrony. Different modalities such as visual, auditory, and haptic have variable needs (transmission rate, sampling, latency, etc.), as a result of these requirements, cross-modal asynchrony emerges. A multiplexing scheme is used to address the problem with provisioning temporal integration of variable modalities Aijaz *et al.* (2017).

The communication of haptic data provides the ability to interact with remote objects and humans. Haptic data are gained by the human senses and related to the haptic perception. This data type is produced in the master domain by a human system interface, then they are collected in the controlled domain by passing over the network domain. The haptic data are utilized to manage a teleoperator or a haptic rendering algorithm in the controlled domain. In the form of telepresence and teleaction systems, haptic communication systems are transparent to the user [Figure 1].

Haptic actuators and haptic sensors are haptic devices which are needed to present a haptic communication. The tactile information is sensed by haptic sensors by attaching them to a teleoperator end. Furthermore, range, reliability, cost, and response time should be considered while choosing a haptic sensor, regardless of the method used for sensor design. Haptic sensors sense heat, force, moisture, etc., which are experienced by the teleoperator while interacting with its environment. Then, the sensed haptic data are sent to the user as haptic feedback. The haptic feedbacks are provided by haptic actuators, which are known as haptic feedback devices.

Haptic feedback can be categorized as cutaneous touch and kinesthetic touch. Kinesthetic touch is composed of the relative position of neighboring body parts' senses and muscle tension senses. The touch of pertaining to the skin is a kind of cutaneous touch. However, it is very complex to design a haptic feedback system that will provide a feeling of touch that will be similar to what the user will feel in the real world Berg *et al.* (2017).

Physical interaction between real and virtual objects is ensured by TI over large areas. There are two types of haptic interaction scenarios which are passive networked haptic interaction and active networked haptic interaction. In passive interaction, there is a remote exploration and perception of objects without manipulating of the objects in the remote area. In active haptic interaction, a manipulative component is added to the remote physical interaction. Thus, the remote environment can be manipulated and perceived by the user Sachs *et al.* (2019).

3. 5G OVERVIEW

As the technology of choice, the 5G is expected to support and facilitate the usage of TI applications and its data transfer through the user equipment (UE), RAN, and core network. This is due to the flexibility of 5G architecture and its elements which are depicted in Figure 2, according to 3GPP organization technical specification (2017). The 5G architecture is designed to be flat where the control plane is separated from the user plane, this is due to some reasons of practicality and adaptation of functionalities for any kind of service needs, and this would support the scalability easily. The core network of 5G is designed based on modular

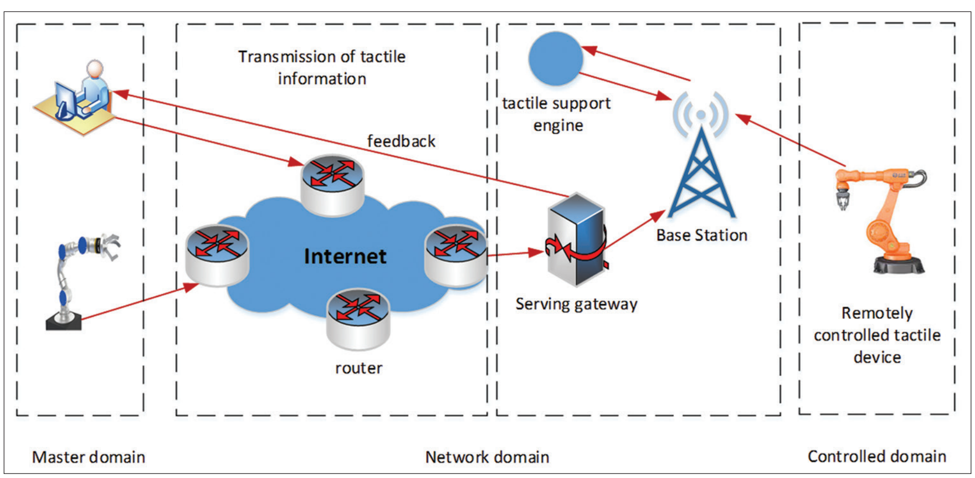

Figure 1. Haptic communication architecture

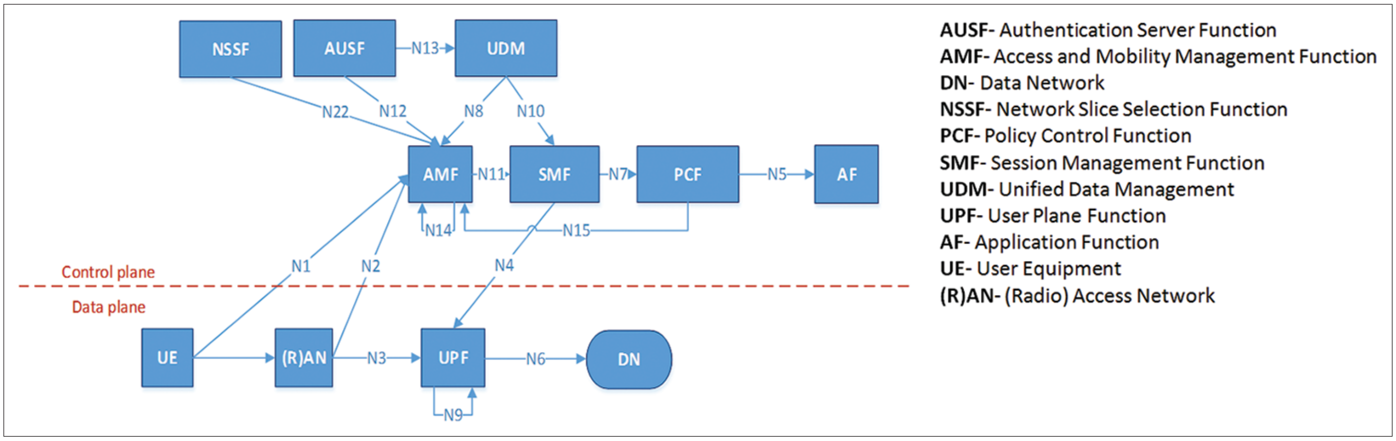

Figure 2. Fifth-generation architecture

functions that generate services; all the modules communicate with each other. This is the reason why 5G core network supports the virtualization of network functionalities and network slicing to be created as soon a new service emerges Chen and Zhao (2014), Technical Specification (2017), and Khodashenas *et al.* (2016).

The 5G introduced new interface called new radio which is using orthogonal frequency-division multiplexing (OFDM); however, the subcarrier spacing ranges from 15, 30, 60 to 120 kHz which gives a variety of slot length for data transmission. This is very beneficial, especially when dealing with the diverse set of services in TI that require different performance in terms of latency. Hence, when the duration of the slot decreases, the latency of transmission decreases too, choosing this configuration is very useful for the critical real-time application introduced in TI. As for the QoS in 5G, it is a flow-based as the packets are classified according to the QFI od (QoS flow identifier). The 5G QoS flows are mapped in an access network to data radio bearers (DRBs) depending on the implicit requirements for each type of application, as shown in Figure 3.

4. LOW LATENCY AS A CHALLENGE IN TI

The TI has a wide range of application; however, not all of them would require a very low latency like the haptic one. The latency is bounded between 1 and 10 ms for the haptic applications while for video and voice, the latency will range between 10 and 100 ms. The paradigm of the network should be modified or support new infrastructure, especially for haptic communication as it requires ultra-low latency and high reliability. By considering the classical view of the actual mobile network and traditional communication protocols, it is not suitable. They are not suitable for the applications of TI as delay is introduced in every layer of the protocol stack in the end-devices, the master domain, the transport domain for the end-to-end communication.

TI did not introduce only haptic type of data but also audio and video (the traditional Internet traffic) within the same device. The decision to have a standalone codec for each type or a hybrid codec is under study. However, when dealing with QoS, each type of traffic will require its own QoS requirements to be guaranteed through the network. Thus, a multiplexer is interacting actively with each type of traffic depending on how urgent is the need for that type of traffic to be prioritized in terms of resource allocation Sachs *et al.* (2019) and Aijaz *et al.* (2018). The multiplexing technique should take into consideration the high update and packet rate compared to audio and video traffic, as it should also change its behavior not only when multiplexing and sending the data through the network but also for congestion control when it happens, the multiplexer should decide which packet to drop and which has higher priority to be passed through the channel Eid *et al.* (2011), Gokhale *et al.* (2017), and Cabrera *et al.* (2019). The methodology used in the multiplexing algorithm would affect so much the latency, and designing a good multiplexer is a challenging issue due to the variety of the type of internet traffic, including the haptic ones.

The other hurdle in TI is the size of packets that influence the delay of transmission as considering the current size would be a big obstacle towards providing the 1 ms end-to-end delay. This delay includes packet processing in the receiver and transmitter, including encoding and decoding. The OFDM used in the current technology where the symbol duration is long may not be a good option for modulation.

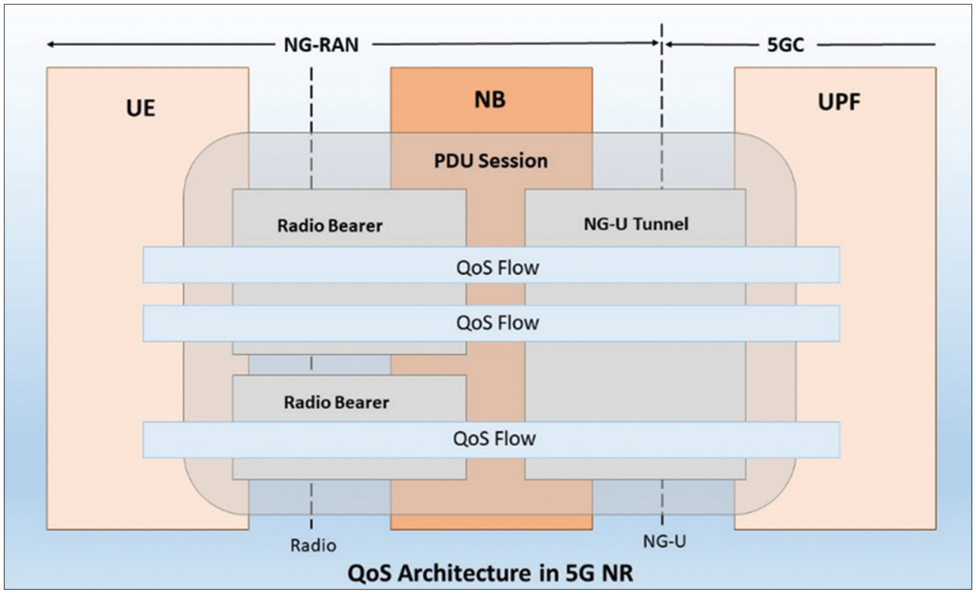

Figure 3. Quality of service architecture in fifth-generation

As for accessing cloud computing capabilities, the delay will increase hugely if the cloud computing is accessed through the traditional internet, this brings a heavy load on the backhaul and core network, as well it will increase the latency, thus violating all the threshold values for latency. Hence, moving the capabilities of the cloud close to the edge will be a good solution, this means that processing and storage should be achieved in RAN and not exceeding it to the core network then internet, consequently reducing the delay to some extent. As the traffic model in TI is not following a regular basis, i.e., the arrival of packets can be sporadic and/or bursty to the schedulers in the medium access control (MAC) layer in the involved entities in the access network, thus the delay experienced by a packet includes the transmission delay and queuing delay in the transmitter side, the decoding and delay in the receiver side. For the applications which have hard latency requirement, the design of the queue delay model should be considered in the whole system.

5. KEY TECHNOLOGIES FOR ENABLING LOW LATENCY IN TI

Despite the fact that there are many challenges that are facing the low latency in the TI, there are some key technologies, frameworks, and communication protocols that attempt to reduce the latency through different methods from an end-to-end network perspective. Those solutions are introduced in the access, core network, and end-user network of the 5G. Along with intelligent artificial methods that are combined with them to guarantee the minimum latency. In the following, we discuss these solutions that are introduced in the different layers of the end-to-end network architecture to achieve low latency:

5.1. SDN and NFV

As key enablers of 5G networks, the SDN and NFV are introduced to automate the network management regardless of the type of hardware used in the core networks. Indeed, all the functionalities of the hardware and network will be abstracted and software to decouple infrastructure from the functions. The SDN is aimed to separate the control plane from the delivery plane so a centralized entity, i.e., the controller can make the decision of the delivery of the traffic according to actions corresponding to rules defined in the flow table of the controller Ateya *et al.* (2018), Matias *et al.* (2015), and Costa-Requena *et al.* (2015). NFV is trying to virtualize or execute the functionalities of the network in a virtualized infrastructure to provide the dynamicity and the adaptive configuration of the network depending on the context that requires certain reconfiguration. The heterogeneous needs of TI in terms of QoS requirements make SDN and NFV the best candidates to satisfy these requirements especially in terms of latency, as they can be adapted to the frequently changing nature of data and traffic since tactile services may coexist with non-TI services Sachs *et al.* (2019). In this case, the traditional QoS mechanisms will fail to guarantee the QoS parameters and some mechanisms based on flow priority should be implemented to respect the end-to-end stringent delay of the TI traffic. Congestion solutions, load balancing, data path, physical location of the NFV and SDN, mobility tracking, etc., all these functionalities will impact on the QoS guarantee and can be flexibly reconfigured to take the 1 ms delay of TI into consideration Giannoulakis *et al.* (2014) and Blanco *et al.* (2017).

5.2. Network Slicing

A common communication infrastructure is one of the points that 5G has introduced in its core network to support different case studies, services, and applications. Its implementation is achieved with the combination of NFV and SDN to create new slices with diverse requirements of any applications in terms of latency, throughput, security, and reliability. Network slicing is an effective method to optimize the resources in 5G by creating a virtual environment with the elements needed by any application. Its flexibility resides in its creation on request and a life cycle duration until the service ends and the slice will be de-allocated. A network slice is a set of network resource along with any functionalities offered by the network and instantiated by the NFV. The network slicing is performing what so-called isolation of resources, this means that each slice will be independent and can request different types of resources, this would guarantee the QoS requirements for the TI as one or several slices will be dedicated to its applications Zhang *et al.* (2017), Rost *et al.* (2017), and Ravindran *et al.* (2017).

5.3. MEC

MEC or multi-access edge computing tries to bring the cloud to the vicinity of end-user, which enables the centralized cloud capabilities to be distributed to the edge of RAN [Figure 4]. This paradigm is assisting in reducing delay, especially for the delay sensitive applications while bringing the computation, processing to be closer to the users themselves Tran *et al.* (2014). This would also assist in the case of TI to enable users to track their real-time information such as behavioral information that can be predictive based on an artificial intelligence algorithm.

The MEC is overcoming the classical cloud by having the edge device close to the user to compute, process and deliver instead of sending the whole data through the backhaul and make congestion in which the latency that TI would like to guarantee would collapse. As well, resource management can be effectively achieved through a collaborative edge computation as the task can be divided into small tasks and be processed through different edge devices. Thus, this would be a good method of offloading the network Mao *et al.* (2017). As well, providing a distributed caching mechanism by extending content delivery network services toward the mobile edge can enhance the users' quality of experience (QoE) while reducing backhaul and core network usage Skorin-Kapov *et al.* (2018). Indeed, there is a tight relationship between the MEC and the QoE; however, since we are dealing with haptic applications, the codec used should reflect the combination between the tactile and kinesthetic. Consequently, a new design for mean opinion score that reflects the best the QoE of users through different codec that combines also video and audio applications emerges.

6. CURRENT RESEARCH WORK DIRECTIONS ON LOW LATENCY IN TI

TI is still in its infancy and new methods should be developed to guarantee the low latency criteria that are one of the essential among the others characterized by ultra-reliability, security, and high QoS and QoE guarantee. However, when referring to the solutions that try to guarantee the minimum latency characterized by 1 ms for the critical haptic

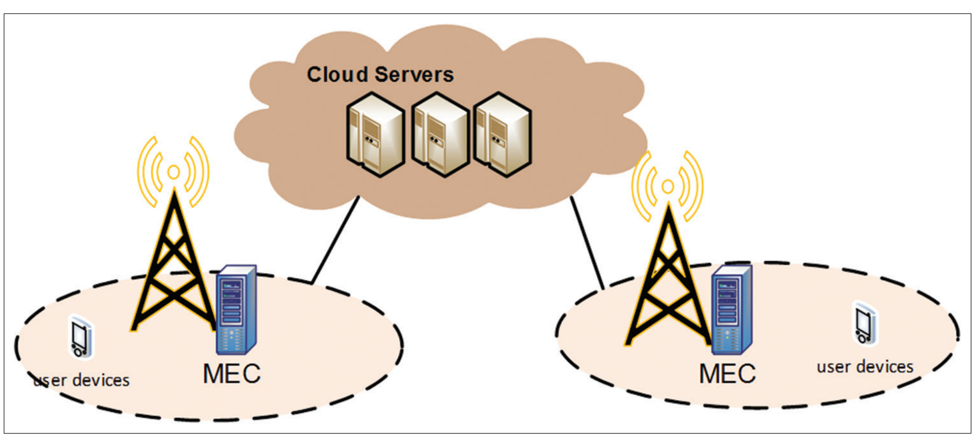

Figure 4. Mobile edge computing architecture

applications to multiple of 100 ms for other type of data combined with haptic one, the adaptive ones can be the best choice. Such kind of solutions should adapt themselves intelligently to the context. Up-to-date, there were many research works trying to deal with the low latency by keeping it as minimum as possible, the method used for achieving this objective can range from the framework, resource allocation, and technological methods such as MEC, network coding, and communication protocols. In the following, we detail some solutions used to reduce the latency explaining the methodology used to achieve that and summarize them in Table 1.

6.1. Resource Allocation

One of the most significant techniques used to reduce the latency is the resource allocation, the main layers involved in the resource allocation and would impact on the end-to-end (E2E) latency in the protocol stack are the physical (PHY) and MAC layers. A cross-layer design involving both layers can be considered as an efficient solution toward obtaining a reduced delay.

One of the characteristics of the TI application is the nature of the traffic generated in the end device is it bustiness. According to the packet arrival process, a burst of packet transmission involves a large amount of data sent in a short time. To exploit the burstiness, authors in Hou *et al.* (2018) classified the packet arrival process into two states: High and low states and they designed different transmission techniques for both states taking into consideration the awareness of the base station (BS) about the quantity of traffic sent by the users. Accordingly, the BS would classify both states based on the Neyman-Pearson method. The behavior of the BS will change according to the amount of traffic; it will reserve dedicated bandwidth for users with a high amount of traffic giving them higher priority while for those of low traffic state a resource pool is shared.

To ensure the low E2E delay communication, authors in She *et al.* (2016) considered short frame structure for transmission and took into consideration different parameters that effect of packet loss during the transmission. Since latency is bounded to delay of transmission and queue delay, while reliability is bounded to packet error probability, queueing delay violation probability, and packet dropping probability. Authors worked on a cross-layer design of MAC and PHY layers to optimize these probabilities in relationship with power allocation in the BS. A proactive packet dropping mechanism is proposed to satisfy the QoS requirement with the limited transmit power. Authors in She and Yang (2016) are based on the work of She *et al.* (2016) to use the solution in the context of vehicle collision avoidance; however, they optimized the bandwidth allocation for users based on the queue delay and the power allocation to ensure the reliability.

Again authors in Gholipoor *et al.* (2018) proposed a cross-layer framework that combines traditional and TI data so it can be more realistic. However, instead of using OFDM, they used sparse code multiple access that 5G proposed for transmission paradigm. The aim of the paper is to increase the sum rate of the traditional data while guaranteeing the delay of TI. They proposed a queue in the transmitter to differentiate the TI from traditional ones as well as in the receiver which the BS in their cases proposed different codebook and power allocation for both types of traffic.

The authors in She *et al.* (2016) investigate the impact of spatial diversity and frequency diversity on ensuring the transmission reliability, and the total bandwidth required for a wireless system to support the QoS requirements of massive machine type devices. They employed a two-state transmission model to characterize the transmission reliability constraint based on the achievable rate with finite blocklength channel codes. They assigned multiple subchannels to each active device, from which the device simply selects one subchannel with channel power exceeding a threshold for transmission after channel probing. They optimized the number of subchannels, the bandwidth of each subchannel, and the threshold for each device to minimize the total bandwidth required by the system to ensure the reliability.

In Aijaz (2016) authors used LTE-A network to allocate resources in terms of resource blocks to the haptic devices involved in the communication. Joint uplink (UL) and downlink (DL) scheduling necessitate an information exchange mechanism between UL and DL schedulers. They have investigated the problem of radio resource allocation for haptic communications over 5G LTE-A networks. The radio resource allocation requirements of haptic communications, together with the constraints of UL and DL multiple access schemes, have been translated into a power and RB allocation problem. To reduce the complexity, the problem is first decomposed using an optimal power control policy and then transformed into a binary integer programming problem for RB allocation. Authors proposed a low-complexity heuristic algorithm for joint UL and DL schedule that not only fulfills the utility requirements of haptic devices but also outperforms the classical algorithms.

Table 1. Summary of low latency related research activities in TI

Ref.	Key enabler	Contribution	Methodology	Outcome
Hou *et al*. (2018)	Resource allocation	Neyman-pearson for state classification Bandwidth optimization for different type of traffic	Analytical model	The optimal overall bandwidth required to guarantee the latency and reliability requirements is achieved
She *et al*. (2016)	Resource allocation	Design of system model for a mixture of tactile and nontactile data	Cross-layer model	QoS guarantee for both types of traffic
She and Yang (2016)	Resource allocation	Bandwidth allocation for vehicle TI	Analytical model for queue and transmission delay	End-to-end delay guarantee and optimized bandwidth
Gholipoor *et al*. (2018)	Resource allocation	Resource allocation for a mixture of traffic within SCMA context	Cross-layer model	Increase the sum rate of tactile data and reduction of sum rate or nontactile data
She *et al*. (2016)	Resource Allocation	Uplink transmission optimization for massive machine type devices in TI	Analytical model	Total bandwidth required by the optimized policy is much smaller than reserving bandwidth for each device like human type communications
Aijaz (2016)	Resource Allocation	Resource allocation for haptic communication over LTE-A network	Low-complexity heuristic algorithm for joint UL and DL scheduling	Delay guarantee over a faded channel.
Rimal *et al*. (2017)	MEC	MEC over FiWi networks is proposed for offloading traffic	Network architecture	Delay, response time efficiency, and battery life resulted to be so efficient
Ateya *et al*. (2017)	MEC	Multi-level cloud in the LTE network architecture	Network Architecture	Reduction of round trip delay
El Saddik *et al*. (2011)	MEC	Caching technique for haptic communication	Network architecture	Latency reduction
Swamy *et al*. (2016)	Network Coding	Protocol of communication	Simulation	Latency reduction
Szabo *et al*. (2015)	Network Coding	SDN and network coding in a multi-hop environment	Simulation and modeling	E2E latency reduction
Cen *et al*. (2005), Wirz *et al*. (2008)	Haptic com. protocol	New transport layer for haptic applications: ETP and IPG	Simulation	Latency reduction
Marshall *et al*. (2008), Gokhale *et al*. (2005)	Haptic com. protocol	Modified UDP for supporting haptic data	Simulation	Latency reduction
Al Osman *et al*. (2007), Eid *et al*. (2011), Nasir and Khalil (2012), Clarke *et al*. (2006), Bhardwaj *et al*. (2013), Sakr *et al*. (2011)	Haptic com. protocol	Modified application protocols to support haptic data: Alpahn, AdMux, HoIP	Simulation	Latency reduction

SCMA: Sparse code multiple access, FiWi: Fiber-wireless, SDN: Software defined networking, ETP: Efficient transport protocol, UDP: User datagram protocol, AdMux: Adaptive multiplexer, HoIP: Haptics over internet protocol, QoS: Quality of service, UL: Uplink, DL: Downlink, TI: Tactile internet

6.2. MEC

The MEC is playing a very important role in the reduction of low latency, and it is designed for this purpose. For example, in Ateya *et al.* (2017) authors proposed multilevel hierarchical of cloud units to reduce the round-trip delay in the mobile network especially the LTE network by introducing a new level of cloud units with higher capabilities in the way between core network and eNBs so as the cloud unit (Microcloud) reduces the communication latency. The authors used the concept of MEC to bring the processing function in the vicinity of users by introducing small micro and Minicloud in each level of the network in the way to the core network which in its turns can include the functionality of the cloud. Authors in Rimal *et al.* (2017) proposed a MEC design on an integrated fiber-wireless (FiWi) access networks. They used a hybrid architecture that integrates MEC and the conventional cloud for different types of applications depending on the nature of the application, whether it is latency sensitive or not. The author claimed that using MEC over FiWi is improving network performance in terms of QoS for different types of applications. In El Saddik *et al.* (2011) authors proposed to replay haptic content using caching technique, which is useful for applications relying on haptic feedback. The caching mechanism is very useful in the case of haptic application, as it is movement/behavior repetition, this means that there is no need to reproduce the same movement/behavior all the time since they will be cached and reused whenever there is a need to produce.

6.3. Network Coding

Network coding is a method that is used to reduce the latency in the network and increase communication efficiency and especially the error probability when sending data on an unreliable channel. This is done through network coding based on algebraic algorithms on each node that receives the data and then recode it and send it to the destination to be decoded. This is why both nodes should be synchronized, and the behavior of the network will depend on one hop communication rather than end-to-end communication used in the traditional networks in which each node will store and forward but not process the data Swamy *et al.* (2016).

With the introduction of SDN in 5G, the network coding can reduce more the latency through multihop networks by including its functionality in the controller. This is why authors in Szabo *et al.* (2015) used network coding especially the random linear network coding integrated with SDN to reduce latency in a multi-hop environment. They introduced the functionality of compute-and-forward to the coding through the use of SDN instead that routers use only the functionality of store-and-forward. This would improve latency and reduce packet re-transmission with respect to other traditional approaches.

6.4. Haptic Communication Protocols

To support haptic communication, either new protocols should be designed to the variable nature of the haptic traffic as it is a hybrid one, or the traditional network protocol should be modified to be adapted to it. Talking about higher layers of the protocol stack just like the transport layer, there are two classical options; namely user datagram protocol (UDP) and transmission control protocol (TCP). However, both protocols have their advantages and disadvantage when supporting haptic data transport. For example, TCP can be considered as a heavy protocol and needs that the connection is established between the peers for a transmission happens. In point of view of reliability, the TCP can be considered as a reliable protocol for haptic communication however in terms of providing low latency; it would introduce extra latency which cannot be suitable in a tactile environment where low latency is required. UDP is a light protocol and most suitable for haptic communication but it does not meet the reliability requirement Aijaz *et al.* (2017) and Berg *et al.* (2017).

A protocol called SuperMedia transport for teleoperations over overlay networks (STRON) El Saddik *et al.* (2011) was created to operate over overlay networks transmitting data using different network paths. STRON was compared against TCP and stream control transmission protocol, showing that it performs significantly better in the case of a network that includes paths with heavy packet loss. Timely execution of the protocol handler tasks with real-time interrupts allows for more immediate transmission of haptic data packets. Furthermore, the efficient transport protocol Shafi *et al.* (2017) aims to reduce round-trip delay time which is related to the inter-packet gap (IPG). By monitoring the transfer rate, it is possible to optimize IPG by setting it to a minimum value to maintain stability and maximum performance of the haptic application Cen *et al.* (2005) and Wirz *et al.* (2008).

As for the application layer, temporal management is important since data can be haptic mixed with audio and video, the challenge is to aggregate all these traffic to be transported in one single data stream. This is why synchronization between both ends are important and should be guaranteed. Almost solutions proposed to multiplex these traffic using different algorithms in literature would count on the UDP protocol as a light protocol Marshall *et al.* (2008) and Gokhale *et al.* (2015). In general, the application layer protocols supporting audio, video, and haptic modalities and according to Gokhale *et al.* (2017) can be classified into (1) constant bit rate based telehaptic protocols in which CBR data streams were inject into the network at a steady rate. Some example of this protocol that adopted this method but modified can include the application layer protocol for haptic networking Al Osman *et al.* (2007), adaptive multiplexer Eid *et al.* (2011), and haptics over internet protocol Nasir and Khalil (2012). (2) Adaptive sampling based telehaptic protocols in which haptic signals samples are injected to the network with partially taking the whole samples into consideration this is due to the high bandwidth that they require, this is why and to avoid this problem, some samples will be identified and will be sent Clarke *et al.* (2006), Bhardwaj *et al.* (2013), and Sakr *et al.* (2011).

7. OPEN ISSUES

The TI can be considered as a revolution or an evolution of the type of communications, applications, services as well as business models. It is expected that TI will be an economy booster for different use cases and scenarios. However, this evolution is still in progress and would need a mature framework that can include all the functionalities described in the previous sections to guarantee the low latency which is a key feature of the haptic communications. There are still challenges to be addressed in this regard, even this article is not an extensive review that gathers all functionalities but indeed it gives an idea about the progress in research in this matter which is a good start towards further elaboration in each mechanism described for reducing the latency. To this aim, the following are some open issues that can be investigated further to achieve this goal:

Proactive resource allocation would be the best solution for critical applications. The tactile applications can be classified according to their requirements into low, high, and adaptive delay applications. The resource allocation mechanism should change its behavior according to the type of application with a dedicated mechanism suitable for each one. In general, the proactive end-to-end resource allocation stringent latency requirement is the best way to guarantee the 1 ms latency criteria. Mechanism of scheduling, QoS provisioning, admission control, etc., should be re-designed to react as fast as possible to process haptic traffic.

Mobile core based cloud native is an interesting idea to bring all the functionalities, capabilities and service of the traditional cloud to the core network; this would reduce lots of hops that the haptic traffic should cross through. As well, this can bring the cloud closer to the users in terms of physical location and can reduce the latency. However, having an overloaded core network would bring some problems of congestion should be solved through SDN and orchestration service provided by the VNF.

Mobility would add an additional dimension of complexity that impacts the QoS and especially the latency when changing the point of attachment from one to another especially when the case of vehicular over TI is considered. The delay of handover would influence on the type latency of communication, position based solution can be considered as best solution of such kind of problem.

Artificial intelligence algorithms can be considered in each step of the design of any resource allocation, or orchestration in the core network. The main interested domain to be combined with is the enforcement learning or the process of learning, since the haptic behavior is repetitive and does not need to generate new action to deal with the motion, therefore, solution based on learning would be an interesting area to investigate.

Anticipatory network is also an option to decrease the latency using some principle such as context, prediction, and optimization. By anticipatory network, the context information will be studied then through past and present information, a prediction of the behavior will be achieved, and then an optimization is carried out to meet the requirement of any application. This would be useful in the case of TI especially improving the QoE of users who are involved in the haptic communications Bui *et al.* (2017).

8. CONCLUSION

The TI can be considered as the next step toward new paradigm of communications, technologies, applications, and services. However, new challenges in deploying its use cases and scenarios will be raised in terms of reliability, low latency, security, and QoS. The main issue that we are interested in as a min subject of this article is the low latency that is one of the most important requirements for TI's application. We are more focused on the critical real-time one which needs a fast reaction and feedback from the other end of the network. Thus ensuring an ultra-low latency for the TI is imperative. The framework, technologies, methods that enable a guarantee for low latency are explained in details and a brief direction of research work has been introduced to give an overview on the most recent progress in this area.

The most recent technology which will enable TI especially its haptic applications is the 5G network which is expected to be standardized in 2020. This is due to the functionalities that automate and utilize the resources in the core network in an efficient way just like SDN and NFV along with its combination with multi-access edge computing to bring the cloud near to the users to reduce the latency. The automation is also represented by adding on the top of the functionalities of 5G some methods of artificial intelligence especially the learning one since almost all haptic applications are repetitive ones.

TI is still in its beginning and research works just started to address the main issues facing its deployment using the current technologies, communication, and protocols. The main conclusion of this paper is that, to enable TI, modifications should be achieved in the RAN, core network, communication protocols in the different layers of the protocol stack. In our future work, we will tackle all the aspects related to the ultra-reliability which is another critical requirement to be fulfilled in TI.

REFERENCES

Aijaz, A. (2016). *Towards 5G-enabled Tactile Internet: Radio Resource Allocation for Haptic Communications*. Doha: IEEE Wireless Communications and Networking Conference Workshops. p. 145-150.

Aijaz, A., Dawy, Z., Pappas, N., Oteafy, S., Simsek, M., & Holland, O. (2018). Toward a tactile internet reference architecture: Vision and progress of the IEEE P1918.1 standard. *IEEE Communications Standards Maganize*, 54(9), 112-118.

Aijaz, A., Dohler, M., Aghvami, A. H., Friderikos, V., & Frodigh, M. (2017). Realizing the tactile internet: Haptic communications over next generation 5G cellular networks. *IEEE Wireless Communications*, 24(2), 82-89.

Al Osman, H., Eid, M., Iglesias, R., & El Saddik, A. (2007). Alphan: Application Layer Protocol for Haptic Networking. In: *Haptic, Audio and Visual Environments and Games*. 2007 IEEE International Workshop IEEE. p. 96-101.

Ali-Yahiya, T. (2018). Towards society revolution. *UKH Journal of Science and Engineering*, 2(2), 37-38.

Amodu, O. A., & Othman, M. (2018). Machine-to-machine communication: An overview of opportunities. *Computer Networks*, 145, 255-276.

Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, C. K. A., & Zhang, J. C. (2014). What Will 5G Be? *IEEE Journal on Selected Areas in Communications*, 32(6), 1065-1082.

Antonakoglou, K., Xu, X., Steinbach, E., Mahmoodi, T., & Dohler, M. (2018). Toward haptic communications over the 5G tactile internet. *IEEE Communications Surveys and Tutorial*s, 20(4), 3034-3059.

Ateya, A. A., Muthanna, A., Gudkova, I., Abuarqoub, A., Vybornova, A., & Koucheryavy, A. (2018). Development of intelligent core network for tactile internet and future smart systems. *Journal of Sensor Actuator Networks*, 7(1), 1.

Ateya, A. A., Vybornova, A., Kirichek, R., & Koucheryavy, A. (2017). *Multilevel Cloud Based Tactile Internet System.* Bongpyeong: 19th International Conference on Advanced Communication Technology. p. 105-110.

Berg, D. V. D., Glans, R., Koning, D. D., Kuipers, F. A., Lugtenburg, J., Polachan, K., Venkata, P. T., Singh, C., Turkovic, B., & Wijk, B. V. (2017). Challenges in haptic communications over the tactile internet. *IEEE Access*, 5, 23502-23518.

Bhardwaj, A., Dabeer, O., & Chaudhuri, S. (2013). *Can we improve over Weber Sampling of Haptic Signals*? 2013 Information Theory and Applications Workshop IEEE. p. 1-6.

Blanco, B., Fajardo, J. O., Giannoulakis, I., Kafetzakis, E., Peng, S., PérezRomero, J., Trajkovska, I., Khodashenas, P. S., Goratti, L., Paolino, M., & Sfakianakis, E. (2017). Technology pillars in the architecture of future 5G mobile networks: NFV, MEC and SDN. *Computer Standards and Interfaces*, 54(4), 216-228.

Bui, N., Cesana, M., Hosseini, S. A., Liao, Q., Malanchini, I., & Widmer, J. (2017). A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques. *IEEE Communications Surveys Tutorials*, 19(3), 1790-1821.

Cabrera, J. A., Schmoll, R., Nguyen, G. T., Pandi, S., & Fitzek, F. H. P. (2019). *Softwarization and Network Coding in the Mobile Edge Cloud for the Tactile Internet*. Proceedings of the IEEE.

Cen, Z., Mutka, M. W., Zhu, D., & Xi, N. (2005). Supermedia transport for teleoperations over overlay networks. In: Boutaba, R., Almeroth, K., Puigjaner, R., Shen, S., & Black, J. P., editors. *Networking 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 1409-1412.

Chen, S., & Zhao, J. (2014). The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication. *IEEE Communications Magazine*, 52(5), 36-43.

Clarke, S., Schillhuber, G., Zaeh, M. F., & Ulbrich, H. (2006) Telepresence Across Delayed Networks: A Combined Prediction and Compression Approach. In: *Haptic Audio Visual Environments and their Applications*. 2006 IEEE International Workshop IEEE. p. 171-175.

Costa-Requena, J., Santos, J. L., Ahokas, K., Premsankar, G., Luukkainen, S., Pérez, O. L., Itzazelaia, M. U., Ahmad, I., Liyanage, M., Ylianttila, M., & de Oca, E. M. (2015). *SDN and NFV Integration in Generalized Mobile Network Architecture*. Paris: European Conference on Networks and Communications. p. 154-158.

Eid, M., Cha, J., & El Saddik, A. (2011). Admux: An adaptive multiplexer for haptic audio visual data communication. *IEEE Transactions on Instrumentation and Measurement*, 60(1), 21-31.

El Saddik, A., Orozco, M., Eid, M., & Cha, J. (2011). *Haptics Technology: Bringing Touch to Multimedia*. New York: Springer-Verlag.

Gholipoor, N., Saeedi, H., & Mokari, N. (2018). *Cross-Layer Resource Allocation for Mixed Tactile Internet and Traditional Data in SCMA Based Wireless Networks*. Barcelona: IEEE Wireless Communications and Networking Conference Workshops. p. 356-361.

Giannoulakis, I., Kafetzakis, E., Xylouris, G., Gardikis, G., & Kourtis, A. (2014). *On the Applications of Efficient NFV Management Towards 5G Networking*. Akaslompolo: 1st International Conference on 5G for Ubiquitous Connectivity. p. 1-5.

Gokhale, V. Nair, J., & Chaudhuri, S. (2017). *Teleoperation over a Shared Network: When Does it Work*? 2017 IEEE International

Symposium on Haptic, Audio and Visual Environments and Games, Abu Dhabi. p. 1-6.

Gokhale, V., Chaudhuri, S., & Dabeer, O. (2015). HoIP: *A Point-to-point Haptic Data Communication Protocol and its Evaluation*. 21st National Conference on Communications, Mumbai. p. 1-6.

Gokhale, V., Nair, J., & Chaudhuri, S. (2017). Congestion control for network-aware telehaptic communication. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 13(2), 17.

Hou, Z., She, C., Li, Y., Quek T. Q. S., & Vucetic, B. (2018). *Burstiness Aware Bandwidth Reservation for Uplink Transmission in Tactile Internet*. Kansas City, MO: IEEE International Conference on Communications Workshops. p. 1-6.

Khodashenas, P. S., Aznar, J., Legarrea, A., Ruiz, C., Siddiqui, M.S., Escalona, E., & Figuerola, S. (2016). *5G Network Challenges and Realization Insights*. Trento; 18th International Conference on Transparent Optical Networks. p. 1-4.

Maier, M., Ebrahimzadeh, A., & Chowdhury, M. (2018). The tactile internet: Automation or augmentation of the human? *IEEE Access*, 6, 41607-41618.

Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). *Mobile Edge Computing: Survey and Research Outlook*. IEEE Communications Surveys. Available from: https://www.arxiv.org/abs/1701.01090. [Last accessed on 2018 Nov 12].

Marshall, A., Yap, K. M., & Yu, W. (2008). Providing QoS for networked peers in distributed haptic virtual environments. *Advances in Multimedia*, 2008, 14.

Matias, J., Garay, J., Toledo, N., Unzilla, J., & Jacob, E. (2015). Toward an SDN-enabled NFV architecture. *IEEE Communications Magazine*, 53(4), 187-193.

Nasir, Q., & Khalil, E. (2012). Perception Based Adaptive Haptic Communication Protocol (Pahcp). In: *Computer Systems and Industrial Informatics*. 2012 International Conference IEEE. p. 1-6.

Parvez, I., Rahmati, A., Guvenc, I., Sarwat A. I., & Dai, H. (2018). A survey on low latency towards 5G: RAN, core network and caching solutions. *IEEE Communications Surveys and Tutorials*, 20(4), 3098-3130.

Ravindran, R., Chakraborti, A., Amin, S. O., Azgin, A., & Wang, G. (2017). 5G-ICN: Delivering ICN services over 5G using network slicing. *IEEE Communications Magazine*, 55(5), 101-107.

Rimal, B. P., Dung Pham V., & Maier, M. (2017). Mobile edge computing empowered fiber-wireless access networks in the 5G era. *IEEE Communications Magazine*, 55(2), 192-200.

Rost, P., Mannweiler, C., Michalopoulos, D. S., Sartori, C., Sciancalepore, V., Sastry, N., Holland, O., Tayade, S., Han, B., Bega, D., Aziz, D., & Bakker, H. (2017). Network slicing to enable scalability and flexibility in 5G mobile networks. *IEEE Communications Magazine*, 55(5), 72-79.

Sachs, J., Andersson, L. A. A., Araújo, J., Curescu, C., Lundsjö, J., Rune, G., Steinbach, E., & Wikström, G. (2019). *Adaptive 5G Low-Latency Communication for Tactile Internet Services*. Proceedings of the IEEE.

Sakr, N., Georganas, N. D., & Zhao, J. (2011). Human perception-based data reduction for haptic communication in six-dof telepresence systems. *IEEE Transactions on Instrumentation and Measurement*, 60(11), 3534-3546.

Shafi, M., Molisch, A. F., Smith, P. J., Haustein, T., Zhu, P., De Silva, P., Tufvesson, F., Benjebbour, A., & Wunder, G. (2017). 5G: A tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE Journal on Selected Areas in Communications*, 35(6), 1201-1221.

She, C., & Yang, C. (2016). *Ensuring the Quality-of-Service of Tactile Internet*. Nanjing: IEEE 83rd Vehicular Technology Conference. p. 1-5.

She, C., Yang, C., & Quek, T. Q. S. (2016). *Cross-Layer Transmission Design for Tactile Internet*. Washington, DC: IEEE Global Communications Conference. p. 1-6.

She, C., Yang, C., & Quek, T. Q. S. (2016). *Uplink Transmission Design with Massive Machine Type Devices in Tactile Internet*. Washington, DC: IEEE Globecom Workshops. p. 1-6.

Skorin-Kapov, L., Varela, M., Hoßfeld, T., & Chen, K. T. (2018). A survey of emerging concepts and challenges for qoe management of multimedia services. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 14(2), 1-29.

Swamy, V. N., Rigge, P., Ranade, G., Sahai, A., & Nikolić, B. (2016). *Network Coding for High-Reliability Low-Latency Wireless Control*. Doha: IEEE Wireless Communications and Networking Conference. p. 1-7.

Szabo, D., Gulyas, A., Fitzek, F. H. P., & Lucani, D. E. (2015). *Towards the Tactile Internet: Decreasing Communication Latency with Network Coding and Software Defined Networking*. Budapest, Hungary: Proceedings of European Wireless, 21th European Wireless Conference. p. 1-6.

Technical Specification. (2017). *System Architecture for the 5G System, Under Change Control 3GPP*. Technical Specification, No. 23.501.

Technical Specification. (2017). *Technical Specification Group Services and System Aspects: Service Requirements for the 5G System, 3GPP*. Technical Specification, Document, No. TS 22.261.

Tran, T. X., Hajisami, A., Pandey, P., & Pompili, D. (2017). Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges. *IEEE Communications Magazine*, 55(4), 54-61.

Wirz, R., Ferre, M., Marn, R., Barrio, J., Claver, J., & Ortego, J. (2008). Efficient transport protocol for networked haptics applications. In: Ferre, M., editor. *Haptics: Perception, Devices and Scenarios*. Vol. 5024. Berlin Heidelberg: Lecture Notes in Computer Science, Springer. p. 3-12.

Zhang, H., Liu, N., Chu, X., Long, K., Aghvami, A., & Leung, V. C. M. (2017). Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges. *IEEE Communications Magazine*, 55(8), 138-145.

Zhang, Q., Liu, J., & Zhao, G. (2018). Towards 5G Enabled Tactile Robotic Telesurgery. Available from: https://www.arxiv.org/abs/1803.03586. [Last accessed on 2019 Dec 09].