Soft Resource Reservation for Time-Delayed Teleoperation over Mobile Networks

Massimo Condoluci[1], *Member, IEEE,* Xiao Xu[2], *Student Member, IEEE,* Qian Liu[2], *Member, IEEE,* Toktam Mahmoodi[1], *Senior Member, IEEE,* Eckehard Steinbach[2], *Fellow, IEEE,* Mischa Dohler[1], *Fellow, IEEE*

Abstract

The emerging Tactile Internet (TI) will enable control-oriented networks for remotely accessing or manipulating objects or devices. One major challenge in this context is how to achieve ultra-low-delay communication between the local operator and the remote object/device to guarantee the stability of the global control loop and to maximize the user's quality-of-experience (QoE). Being one of the major human-in-the-loop applications of the TI, haptic teleoperation inherits its delay-sensitive nature and requires the orchestration of communication and control approaches. In this paper, we focus on two teleoperation systems: (1) teleoperation with the time-domain passivity approach (TDPA), which is highly delay-sensitive but supports highly dynamic interaction between the operator and a potentially quickly changing remote environment; (2) model-mediated teleoperation (MMT), which is tolerable to relatively larger communication delays, but unsuitable for quickly changing remote environments. We first study the characteristics of both TDPA and MMT schemes (the human operator's QoE versus communication delays, haptic packet update rate, and haptic packet arrival rate) through subjective tests. Based on these studies, we propose a novel soft resource reservation mechanism for the uplink scheduling of mobile networks which accomplishes a double-fold benefit: (1) It achieves admirable round-trip delay reduction compared with the current legacy scheme; (2) It improves the human operator's QoE by leveraging the traffic characteristics of the control schemes during the scheduling design. The simulation results confirm the efficiency of the proposed scheme.

Index Terms

Teleoperation, QoE, resource allocation, resource reservation, scheduling, Tactile Internet

Massimo Condoluci, Toktam Mahmoodi and Mischa Dohler are with the Dept. of Informatics, King's College London, UK. E-mail: {massimo.condoluci, toktam.mahmoodi, mischa.dohler}@kcl.ac.uk.

Xiao Xu, Qian Liu and Eckehard Steinbach are with the Chair of Media Technology, Technical University of Munich, Germany. Qian Liu is also with the Dept. of Computer Science and Technology, Dalian University of Technology, China. E-mail: {xiao.xu, qian.liu, eckehard.steinbach}@tum.de.

I. Introduction

THE past decade has witnessed a tremendous growth of the Mobile Internet which connects millions of mobile devices on a global scale. More recently, we witnessed the emergence of the Internet of Things [1] which enables the transition from the network of mobile communication devices to the network of billions of physical devices, objects, animals, and human beings. Most recently, these different Internet embodiments embrace the rise of the Tactile Internet (TI) [2], [3], which aims at providing ultra-low-delay and ultra-high-reliable communications enabling a paradigm shift from conventional content-oriented communication to a "*control*" oriented communication. The TI is of particular relevance for the realization of "*human-in-the-loop*" remote teleoperation applications which are highly delay sensitive and require a tight integration of the communication and control mechanisms.

Teleoperation systems with haptic feedback allow a human user to immerse into a distant or inaccessible environment to perform complex tasks. A typical teleoperation system, as illustrated in Fig. 1, comprises a slave and a master device, which exchange haptic information (forces, torques, position, velocity), video signals, and audio signals over a communication network [4]. In particular, the communication of haptic information (position/velocity and force/torque signals) imposes strong demands on the communication network as it closes a global control loop between the operator and the remote robot. As a result, the system stability is highly sensitive to communication delay [5].

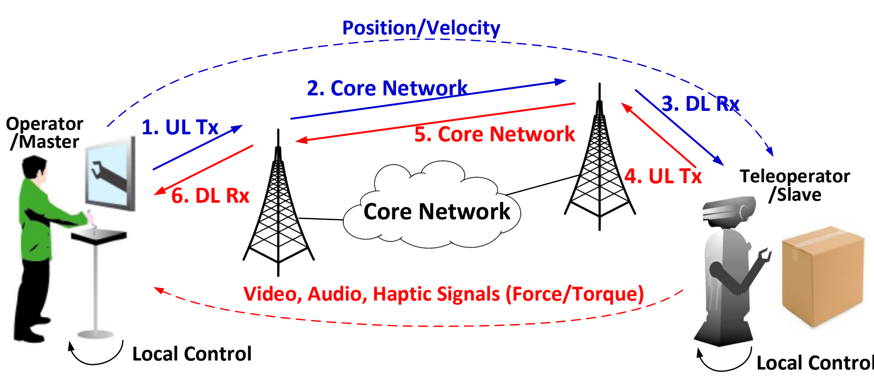


Fig. 1. Illustration of a bilateral haptic teleoperation over LTE networks.

In addition, high-fidelity teleoperation requires a high sampling rate of 1 kHz or even higher [6] to ensure a high quality interaction and system stability. In order to keep the communication delay as small as possible, haptic sensor readings are typically packetized and transmitted once available. Teleoperation systems, hence, require 1000 or more haptic data packets/s to be transmitted between the master and the slave device. For Internet-based communication, such high packet rates are hard to maintain.

State-of-the-art solutions that address the aforementioned teleoperation challenges (low delay and high packet rate) focus on combining different stability-ensuring control schemes with haptic packet rate reduction methods. Several lossy perceptual haptic packet rate reduction schemes [7], [8] were developed in recent years (please refer to Section II for details). On the other hand, since the communication delay can range from a few milliseconds up to several hundred milliseconds, teleoperation systems require a control scheme which stabilizes the teleoperation system in the presence of end-to-end communication delays that are larger than a couple of milliseconds. Fig. 2 presents a qualitative analysis of the trade-off between the level of communication delay and the level of abstraction in control schemes for teleoperation [9]. We can observe from Fig. 2 that passivity-based control, e.g. the time-domain passivity approach (TDPA) described in [10], [11], [12], is suitable for short-distance (low-latency) teleoperation with dynamic scenes and a high level of interaction between the master and the remote environment (i.e. high update rate); the model-mediated teleoperation (MMT) [9], [13], [14] is able to deal with relatively larger communication delays (i.e. for medium or long distance application scenarios), but is unsuitable for quickly changing environments. Teleoperation for very large delay is typically performed using supervisory control and will not be further considered in this paper. It is obvious that different control schemes have different delay tolerance, and require different amounts of resources from the communication network. Therefore, a successful design of the communication scheme should exploit the characteristics of different control schemes and be able to provide robust and reliable support for various teleoperation applications.

Although teleoperation systems could be considered as yet another domain in machine-type communications (MTC), their traffic pattern is strongly different from typical MTC traffic. Generally, MTC traffic is characterized by small packets, short-lived flows and storms of access requests. In contrast, the haptic traffic, while depending on the used control scheme, has a different nature. Our experimentation, for example, shows bursty and irregular patterns of data traffic between the haptic master and slave device (this is further elaborated in the paper). To this end, we thoroughly study the pros and cons of a number of resource request and allocation solutions, such as the scheduling request (SR) [15], [16], the semi-persistent scheduling [17], [18], and the contention-based approach [19], in terms of addressing the specific needs of the

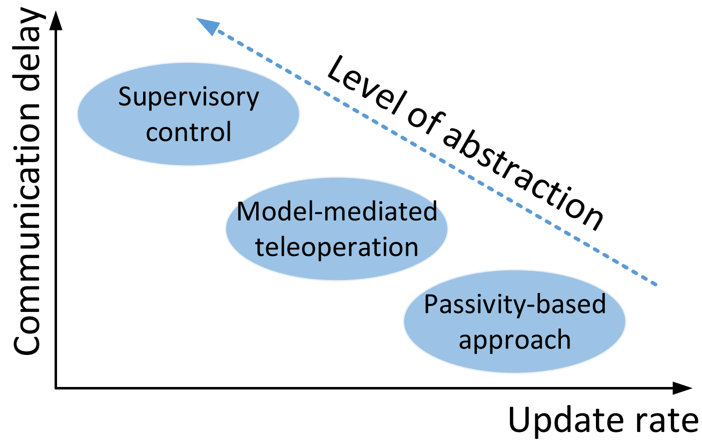


Fig. 2. The level of abstraction and data complexity with update rates and robustness to delays (adapted from [9]).

haptic data traffic, and the corresponding control scheme.

We can further observe from Fig. 1 that as a major component of the global control loop, the human operator has a straight impact on the overall performance of the teleoperation system. Since the quality-of-experience (QoE) of the human-in-the-loop degrades for increasing communication delay, the scheduler design should by all means explore the possibility to reduce the communication delay in order to achieve a satisfactory QoE performance during teleoperation.

In this paper, we propose a *soft resource reservation* in the SR procedure of mobile networks, in which the Uplink (UL) grant from one transmission is soft reserved for the following transmissions, and therefore the latency is significantly reduced while the spectral efficiency is maintained, thanks to the "*soft*" nature of the resource reservation (this is in strong contrast to the semi-persistent scheduler). The key contributions of the proposed scheduler can be summarized as follows:

(1) Through extensive analysis, we show that the proposed soft resource reservation mechanism achieves a noticeable round-trip latency improvement for haptic communications.

(2) The achieved latency improvement, mentioned in (1), allows the teleoperation system to constantly perform with satisfactory QoE using TDPA scheme, in typical *short haul* communications (domestic or regional communications).

(3) With the exploitation of the characteristics of the haptic traffic, the proposed scheduler is able to provide a satisfactory QoE performance for teleoperation with both TDPA and MMT schemes.

The remainder of this paper is organized as follows. In Section II, we will give a brief introduction to the most representative control schemes for time-delayed teleoperation. In Section III, we will demonstrate the characteristics of haptic traffic in teleoperation systems, through experimentation. The key performance indicator for teleoperation is also introduced here. Section IV provides a thorough comparative study on the legacy schedulers and proposes a novel strategy, followed by performance studies in Section V. We also provide an analysis of the performance of teleoperation systems and some recommendations in this section. Finally, this paper concludes with Section VI.

II. Overview of Time-delayed Teleoperation with Different Control Schemes

In order to address the two major communication-related challenges (i.e. time delay and high packet rate) of teleoperation systems, various joint communication/control schemes have been developed by combining different stability-ensuring control approaches with haptic data reduction algorithms. The most representative solutions from the literature are the combination of the TDPA from [11] with the perceptual deadband-based (PD) haptic data reduction solution from [20], denoted as TDPA+PD [7], and the combination of the MMT method from [13], [14] with the haptic data reduction solution from [20], denoted as MMT+PD [8]. We will introduce these two approaches in the following two sections.

A. Time-domain Passivity Approach with Perceptual Haptic Data Reduction

The TDPA [10], [11], [12] is a typical passivity-based control scheme for time-delayed teleoperation. The stability arguments are based on the passivity concept, which characterizes the energy exchange over a two-port network and provides a sufficient condition for the input/output stability. The stability of TDPA-based teleoperation systems is guaranteed in the presence of arbitrary communication delays with the help of passivity observers (PO) and passivity controllers (PC). The PO computes the current system energy. The PC adaptively adjust the customized dampers α and β to dissipate energy and thus guarantee the passivity of the system. The haptic data reduction blocks are placed after the POs to irregularly downsample the transmission of haptic packets using perceptual thresholds (see Fig. 3).

The TDPA is a conservative control design. With increasing delay, it leads to larger distortions in the displayed environment properties. For instance, hard objects are displayed softer than they actually are. In addition, the TDPA+PD method can lead to sudden force changes when the PCs

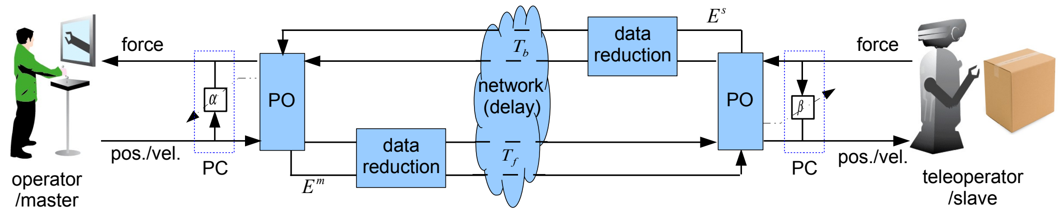


Fig. 3. Overview of the time-domain passivity approach with haptic data reduction (adopted from [7]).

are activated to dissipate the system output energy. This effect becomes stronger with increasing communication delays [7], [11].

Therefore, the TDPA+PD approach is suitable for short-distance teleoperation applications which may operate at the edge of the communication network and meet the low-latency requirements, while maintaining frequent haptic updates between the master and the slave. Thus, it can deal with a high level of dynamics of the objects (motion, deformation, etc.) and interaction patterns.

B. Model-Mediated Teleoperation with Haptic Data Reduction

One of the main issues of the passivity-based control schemes is that the system passivity and transparency are conflicting objectives. This means that the system gains stability at the cost of degraded transparency [5]. In order to avoid this issue, model-mediated teleoperation (MMT) has been proposed to guarantee both stability and transparency in the presence of communication delays [13], [14]. In the MMT approach, a local object model is employed on the master side to approximate the slave environment. The model parameters describing the object in the slave environment are continuously estimated in real time and transmitted back to the master whenever the slave obtains a new model. On the master side, the local model is reconstructed or updated on the basis of the received model parameters, and the haptic feedback is computed on the basis of the local model without noticeable delay. If the estimated model is an accurate approximation of the remote environment, both stable and transparent teleoperation can be achieved [9], [21]. The data reduction scheme is used to irregularly downsample the velocity signals in the forward channel and the model parameters in the backward channel, using perceptual thresholds (see Fig. 4). For the model parameters, these thresholds determine if a model update leads to a perceivable difference in the displayed source. If not, the model change does not have to be transmitted.

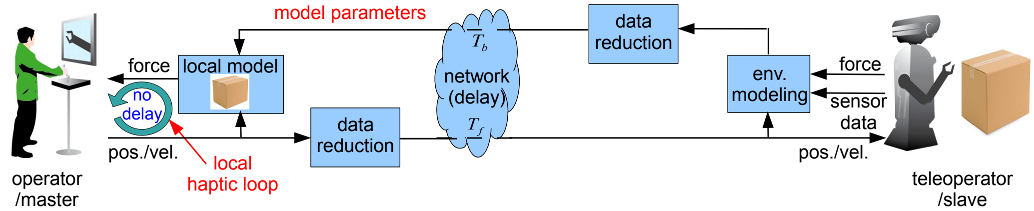


Fig. 4. Overview of the model-mediated teleoperation approach with haptic data reduction (adopted from [8]).

Using the MMT approach, hard objects will not be displayed softer with increasing delay as for the TDPA. Therefore, the MMT approach has better teleoperation quality when the delay is relatively large. Keeping the local model consistent with the environment at the remote side is, however, challenging for dynamic scenes. This way, the MMT approach is favorable for medium/long distance teleoperation applications and scenarios which are characterized by a low level of scene dynamics.

Although the two approaches address different scenarios, it holds for both that the lower the end-to-end delay, the better the system transparency and hence the QoE.

III. Characteristics of Teleoperation with Different Control Schemes: A Case Study

A natural question following the above discussion is which control scheme should be adopted to achieve the best possible system performance. In fact, the choice depends on many factors, for example, the system (hardware) setup and the network communication delay, which is the focus of this paper.

For the system setup, the TDPA needs only a force sensor equipped at the slave to measure the interaction force, while the MMT method requires additional sensors (e.g. a 3D sensor [8]) and computational resources to perform the online environment modeling. The cost for environment modeling in the MMT method is relatively higher than for the TDPA. Teleoperation systems need to determine whether it is worth spending the resources for MMT to gain the potential improvement in the displayed impedance. This is especially challenging when the environment is complex.

The choice between the two control schemes is also affected by the network communication delay. As discussed in II-A and II-B, the performance of the TDPA+PD method is mainly influenced by communication delay. The larger the delay, the softer the displayed impedance

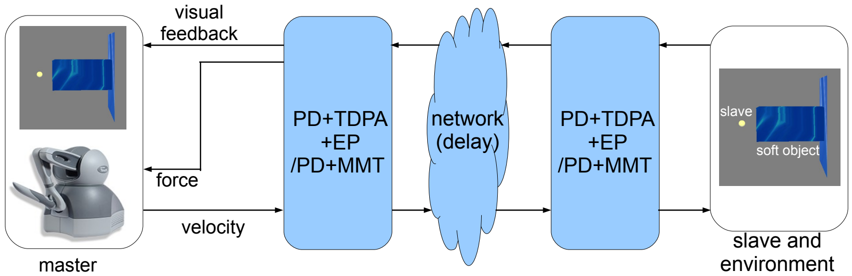


Fig. 5. Experimental setup. The communication network, the slave (represented by a haptic interaction point), and the virtual environment are simulated on a PC using the CHAI3D library.

(stiffness), and the stronger the distortion introduced in the force signals. On the other hand, the performance of the MMT+PD method depends strongly on the model accuracy, which is barely affected by the communication delay for services with low object dynamics. Hence, the TDPA+PD and the MMT+PD methods are the best choice for different delay ranges. In addition, these two methods transmit different types of data in the backward channel, which can lead to very different traffic characteristics over the communication network.

We use a simple 1D spring-damper setup (as shown in Fig. 5) in this section to evaluate and compare the performance of the two control methods in terms of user preference and the generated haptic data traffic. The user's quality-of-experience (QoE) for both control schemes in the presence of difference communication delays is evaluated using subjective tests.

A. Experimental Setup

The experiment was conducted in a virtual environment (VE) developed based on the Chai3D library (www.chai3d.org). The Phantom Omni haptic device was used as the master, while the slave in the VE was designed as a single haptic interaction point (HIP) with negligible mass. The communication network with adjustable delay was simulated in a local PC. The VE contained a 1D non-linear soft object, whose parameters were designed based on the Hunt-Crossley model [22]

$$f_e = \begin{cases} Kx^n + Bx^n\dot{x} + \Delta f & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{1}$$

where f_e is the interaction force (environment force) and Δf is the Gaussian distributed measurement noise with a mean of 0 N and a standard deviation of 0.1 N. x denotes the penetration (compressed displacement). Corresponding parameters were set as: $K = 200$ N/m, $n = 1.5$, and $B = 0.5$ N/ms. For the PD+MMT method, a linear simple spring model ($\hat{f}_e = \hat{K}x$) was employed to approximate the environment. This led to model mismatch and frequent changes in the estimated model stiffness $\hat{K}$ during interaction. The passivity-based model update scheme, introduced in [23], was applied to ensure stable model updates on the master side. All experiments were conducted on a PC with an Intel Core i7 CPU and 8 GB memory. The whole experimental setup is illustrated in Fig. 5.

The round-trip delays were set to 0 ms, 10 ms, 25 ms, 50 ms, 100 ms, and 200 ms, repectively. For each delay, the subjects were tested under three conditions: the TDPA+PD method, the MMT+PD method, and the 0-delay reference without using any control and data reduction schemes. The reference scenario was shown to the subjects before the real experiments. The original environment impedance was displayed and the best performance (uncompressed, non-delayed) for this setup was provided.

The subjects interacted with the virtual object by pressing its surface and slowly varying the applied force. The subjects were asked to give a rating by comparing the interaction quality between each control scheme and the reference scenario. They were asked to take all perceivable distortions (e.g. the force vibration, the force jumps, the perceived impedance variation, etc.) into account when evaluating the interaction quality. The rating scheme was based on Table I. The reference, designated level 5, was considered as the best performance. The reference can be recalled by the subjects at any time during the experiment. Each delay case was repeated four times. The order of the tested delay as well as the order of the tested control methods were randomly selected.

There were 12 participants, i.e. subjects (right handed and ranging from the age of 25 to 33), in the experiment. The participants were asked to wear a headset with active noise cancellation to protect them from the ambient noise. During the experiment, they were first provided with a training session, then started the test as soon as they felt familiar with the system setup and the experimental procedure.

B. QoE vs. Communication Delay

A quantitative evaluation of the subjective ratings (QoE) for the two control methods is illustrated in Fig. 6. The QoE for both control methods degrades with increasing delay. For the MMT+PD method, the QoE is fairly stable, which confirms that the QoE of the MMT+PD approach mainly depends on the model accuracy, rather than the communication delay. In contrast

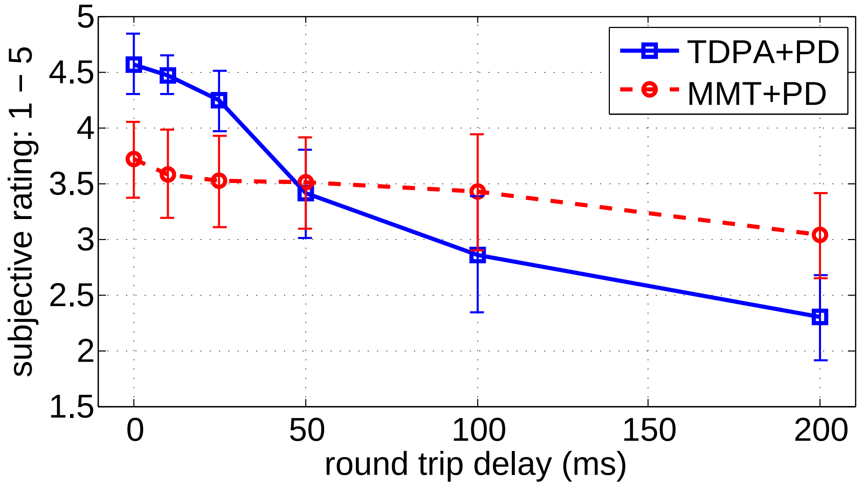


Fig. 6. Subjective rating vs. communication delay for the two control schemes.

TABLE I

RATING SCHEME FOR SUBJECTIVE EVALUATION.

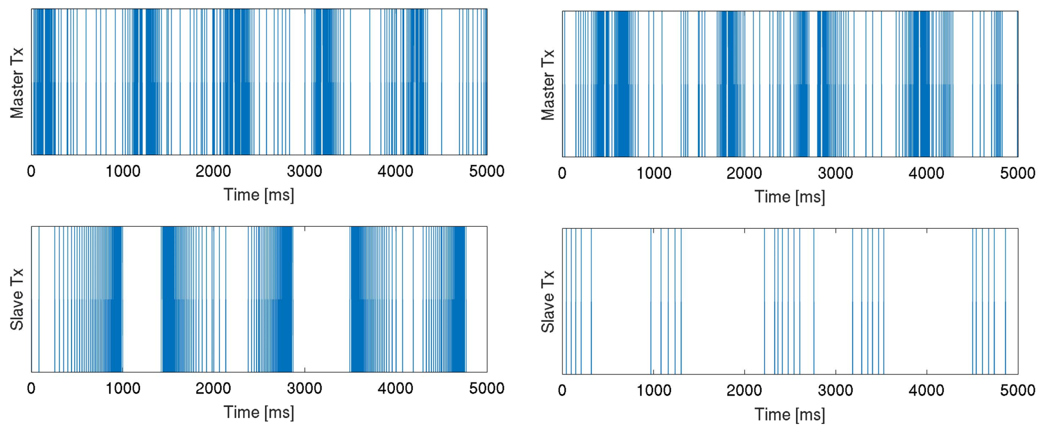
Rating level	Description
5	no difference to the undisturbed signal (perfect)
4	slightly perceptible disturbing (high quality)
3	disturbed (low quality)
2	strongly disturbed (very low quality)
1	completely distorted (unacceptable)

to the MMT+PD method, the QoE of the TDPA+PD method is more sensitive to delay, as discussed in Section II-B.

According to Fig. 6, the TDPA+PD method is able to provide relatively higher QoE than the MMT+PD method when the delay is small. However, the QoE of the TDPA+PD approach decreases quickly with increasing delay. This is because the subjects will perceive more vibrations and force jumps, as well as softer environment impedance when the delay grows. **The overall rating results show that the subjects prefer the TDPA+PD method for small delays, and the MMT+PD method for large delays.**

C. Characteristics of the Haptic Traffic

During the subjective tests, we recorded the haptic traffic for both schemes. From Fig. 7, we can observe the bursty behavior and the irregular updates between the master and the slave for both control schemes. However, the packet update rate of the backward channel (slave-to-



(a) TDPA+PD (b) MMT+PD

Fig. 7. An example of the forward and backward haptic traffic for the two control schemes over a time window of 5s (Subject 1).

master) in the MMT+PD method is significantly smaller than for the TDPA+PD method, which is consistent with the analysis shown in Fig. 2. In addition, the haptic packet size of the two control schemes is different (see Table II). Please note that the packet size shown in Table II hold for the specific experimental setup used in this example. In particular, the packet size for the model update can be much larger for other setups, e.g., if the object geometry is represented by a point cloud or a polygonal mesh.

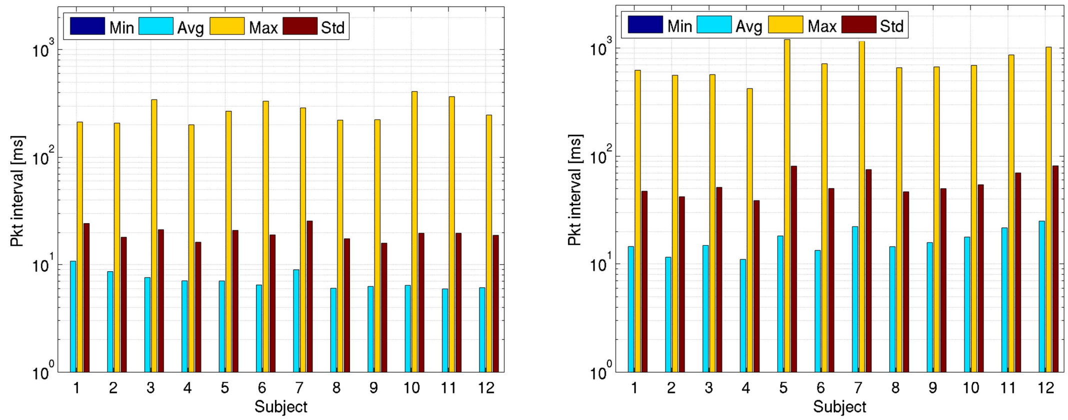
TABLE II

SIZE OF THE HAPTIC DATA PACKETS FOR THE TWO CONTROL SCHEMES.

Packet size	TDPA+PD	MMT+PD
Forward (master-to-slave)	24 bytes	12 bytes
Backward (slave-to-master)	24 bytes	48 bytes

Fig. 8 and Fig. 9[1] extend the analysis in Fig. 7 by focusing on the inter-arrival time of the packets for both the forward and backward channels (all 12 subjects are analyzed). For the TDPA+PD method, we can observe from Fig. 8 that the inter-arrival time is similar for all the subjects and there is no significant difference for the master or the slave device (i.e., a mean inter-arrival time of about 10 ms). The bursty behavior is highlighted by the presence of periods

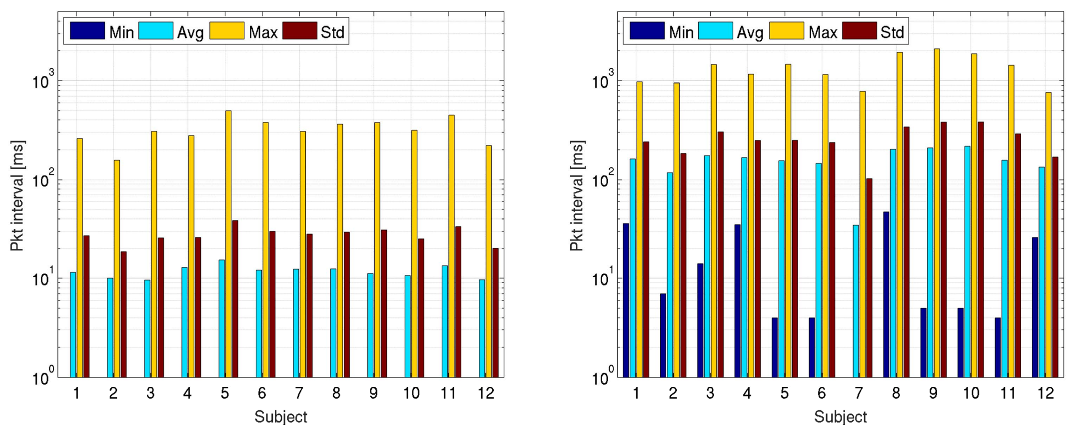
[1]In Fig. 8 and Fig. 9, the minimum inter-arrival time is equal to 1 ms for both master and slave devices. The value is not visible due to the use of a logarithmic scale on the y-axis.



(a) Forward (master-to-slave)

(b) Backward (slave-to-master)

Fig. 8. Packet inter-arrival times for 12 subjects using the TDPA+PD method.



(a) Forward (master-to-slave)

(b) Backward (slave-to-master)

Fig. 9. Packet inter-arrival times for 12 subjects using the MMT+PD method.

of up to 250 ms without any packet transmission from the master side [plot in Fig. 8 (a)], while this off interval reaches up to 1 sec for the slave-to-master communication [plot in Fig. 8 (b)].

For the MMT+PD method, we can observe from Fig. 9 that the inter-arrival times are again similar for all the subjects. However, we can notice a strong difference between the master and slave sides. Indeed, the slave-to-master communication shows a minimum packet inter-arrival time of up to 50ms, an average value of about 100ms and maximum inter-arrival times in the order of 2.5s.

In summary, the subjective experiments confirm the relationship shown in Fig. 2, which indicates that different control schemes have different delay tolerance, and require different amount of resources from the communication network. Also, the experiments have shown that the teleoperation traffic is irregular, bursty and largely unpredictable. **Therefore, in order to achieve high-quality teleoperation, novel communication mechanisms need to be developed with full consideration of the characteristics of different control schemes, especially the influence of communication delay on the user's QoE**.

IV. Data Transmission and Reception

A master-slave communication is depicted in Fig. 1, depicting different involved parts in a haptic session, i.e., the master sends data to the slave and should receive feedback from the slave in a timely manner to the human operator at the master side. Fig. 1 also highlights different components of the delay in the haptic session, which are: *uplink (UL) transmission*, for the transmission of packets from either the master or the slave to the base station; *downlink (DL) transmission*, for the reception of packets sent from the base station to either master or slave devices; *core network*, exploited to inter-connect the base stations involved in the session.

For the domestic or regional communciations, the core network delay can vary between 1 ms and 20 ms based on the 3GPP study on latency reduction in the LTE [24]. In addition, the concepts of edge cloud and fog computing [25] brings intelligence to the TI, resulting in a further improvement of the communication delay. However, we should point out that a large transmission delay is involved in the inter-continental communication over mobile networks. Compared with the domestic and regional counterparts, this delay is inevitable, but can be considered as a constant. Therefore, in the remainder of this paper, we assume that the transmission delay over the core network is constant, 1 ms [24] and 90 ms [26] for the demostric/regional and inter-continental scenarios, respectively. We will, therefore, focus on how to reduce the latency in the data transmit/receive procedures of the radio access in the network.

Within the radio access, the transmission of packets in the UL direction is performed by means of two mechanisms: random access or schedule request. The random access procedure [27] is performed when the device is not synchronized with or connected to the network. This procedure is, thus, usually performed for the initial attachment of the device to the network. In the haptic session, this procedure happens only once, i.e., during the session setup phase, and it does not affect the data transmission from both master and slave sides. On the contrary, the scheduling request (SR) [15], [16] is performed when a device has an active connection and it needs to request resources for UL transmission to the base station. The SR procedure is depicted in Fig. 10. The device is configured by the base station with a SR period, varying from 1ms

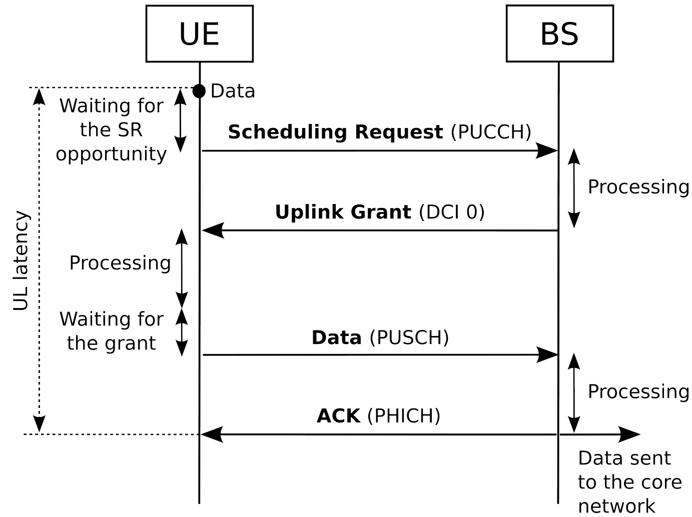


Fig. 10. The scheduling request procedure in LTE.

up to 320ms. In case the device has data to transmit, it should wait until the next available SR opportunity in order to inform the base station about the amount of data in its buffer. To avoid congestion at the base station, the *sr-ProhibitTimer-r9* (whose duration can vary from 0 to 7 SR periods) has been introduced to avoid too many requests to be sent in subsequent SR opportunities. At the reception of the SR, the base station processes the incoming request in order to allocate the resource blocks (RBs) to the device, then it sends an UL grant to the device with the information about the allocated resources. Finally, the device processes the received grant and sends its data on the RBs the base station has allocated to its transmission[2]. As analyzed in [24] and [19], the SR procedure can guaranty (with a given margin) deterministic delays and it is thus capable to guaranty a stable latency for haptic communications. Nevertheless, as studied in [24], the SR procedure introduces a considerably significant delay to the low latency communications due to its own nature, as explained. However, the SR is a device to base station handshake procedure, and can maintain high spectral efficiency at the expense of the latency in data transmission. One possible solution to cut the SR delay is moving to semi-persistent

[2]Another aspect to consider is the HARQ procedure, exploited to inform the device about the effective reception of the transmitted signal and to trigger a re-transmission in case of a failure. From a haptic session point of view, this procedure introduces delays in case of re-transmissions. Nevertheless, solutions such as [28], [29], [30] can be exploited to improve the transmission reliability in haptic session by exploiting the static nature of the involved devices. In this paper, the delay associated with the HARQ procedure is not considered, since the main focus is on analysis of data transmission/reception strategies on the radio segment.

scheduling, which is based on the idea of statically allocating resources, i.e., a given portion of RBs to be exploited with a given periodicity to the device [17], [18]. Therefore, the device can send its data without performing the SR and thus cutting the transmission delay. Semi-persistent scheduling is known as an effective way to reduce delays for traffic with deterministic behavior, such as VoIP, where the device generates one voice packet each 20 ms [31]. On the other hand, haptic traffic can drastically vary during the session, according to the plots in Fig. 8 and Fig. 9. Hence, semi-persistent scheduling would involve wasting of resources during the long off-period of the haptic session, since RBs assigned to the haptic session cannot be exploited by any other devices and this would drastically decrease the spectral efficiency. Alternative solution for avoiding the SR delay is the introduction of a novel contention-based channel, as proposed in [19]. In this case, instead of having resources pre-scheduled for each device, the network assigns resources to be exploited by all haptic devices. These resources, with pre-defined modulation and coding scheme (MCS) and pre-defined packet size, will be shared among the haptic devices that are active in the cell. In other words, any device with packet to be sent, waits until the next available resource and then transmit its packet. The main drawback of this approach, however, is the possibility of collision in case multiple devices select the same RB(s) to send their data. This aspect is exacerbated in haptic session when devices send packets every 1ms during their bursty interval, as observed in Fig. 8 and Fig. 9. **This means that, to effectively avoid collisions and thus retransmissions, the contention-based approach needs to reserve a significant amount of RBs which is spectrally inefficient**.

Despite the explained latency in the UL direction, the DL data communication is considerably shorter and the procedure is less complex than UL. Fig. 11 shows the DL procedure in LTE [24]. Since the base station is already aware of the amount of data to be delivered to the device, it only needs to inform the device about the assigned resources and subsequently send the data. To this end, we will focus on reducing the UL latency, and propose a novel solution for the UL in the remainder of this section.

A. System model

We consider a scenario with K master-slave pairs. For the sake of simplicity, we assume all K masters are co-located and are served by one base station while K slaves are served by another base station. An illustration of our deployment scenario with one master-slave pair can be seen in Fig. 1.

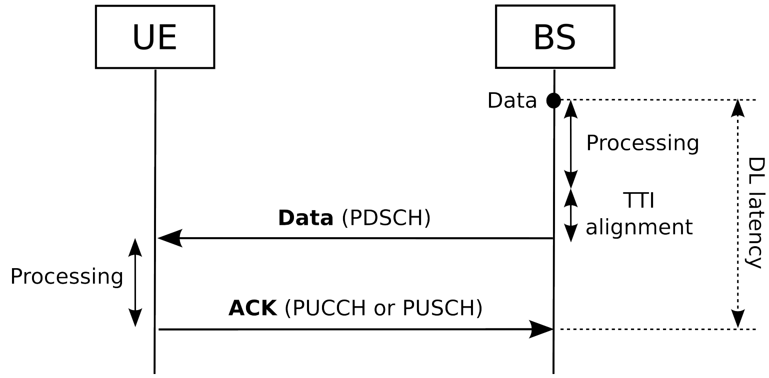


Fig. 11. Downlink reception procedure in LTE.

TABLE III

LIST OF NOTATIONS

Notation	Definition	Value
K	Number of devices	-
$R_D L$	Number of DL RBs	25
$R_U L$	Number of UL RBs	25
T_{SR}	SR period	5ms [24]
T_{DL}	DL scheduling period	5ms [24]
T_{TX}	Transmission time	1ms [24]
T_{PRO}	Processing time	1ms [24]
T_{Al}	TTI alignment	0.5ms [24]
σ_k	SINR device k	-
s_k	Packet size	24B Master/Slave Control Scheme A, 12 B Master Control Scheme B, 48B Slave Control Scheme B
r_k	Number of RBs needed by device k	-
$\mathcal{B}$	Buffer of device k	-

We focus our attention on the transmission and reception procedures within one cell, with R_{DL} and R_{UL} resource blocks reserved, respectively, for data reception in DL and data transmission in UL. The transmission time on the radio interface is denoted with T_{TX}, and the time needed for processing received data is given by T_{PRO}. By considering a generic device k within the cell, s_k and σ_k represents the size of packet to be transmitted/received, and the experienced SINR, consecutively. Finally, the set of packets to be transmitted or received by a device k, is denoted with $\mathcal{B}_k$ (this also indicates the transmission buffer of a device k) and the amount of resource blocks needed to receive the scheduled packets is given by r_k.

B. Legacy Data Transmission

The legacy data transmission is performed by means of the SR procedure, depicted in Fig. 10. This procedure allows the devices to inform the base station about the status of their buffers, and the scheduling requests can be sent from the devices with a periodicity equal to T_{SR}.

The data transmission buffer at device k, i.e., $\mathcal{B}_k$, contains the list of the packets to be sent. We consider an identification for each packet, denoted with p_k, which allow us to trace and measure the data transmission delay for each packet. When a packet reaches the data transmission buffer, it is added to the buffer, i.e., $\mathcal{B}_k = \mathcal{B}_k \cup \{p_k\}$, and hence the packet identification will be increased, i.e., $p_k = p_k + 1$.

In every SR period, the device checks its own buffer and perform the SR in case the buffer is not empty. The binary parameter $\alpha_{k,t}$ indicates if the device k performed a SR at the t-th SR period. The $\alpha_{k,t}$ can be defined as follows:

$$\alpha_{k,t} = \begin{cases} 1, & \text{if } \mathcal{B}_k \neq \emptyset \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

If $\alpha_{k,t} = 1$, i.e., device k sent a SR in the t-th slot, base station needs to take into account the buffer state of device. We define $\mathcal{A}_{k,t}$ to be number of packets of device k taken into account by the base station in the scheduling procedure at the t-th time slot. Clearly, when the SR occurs, $\mathcal{A}_{k,t} = \mathcal{B}_k$.

After the transmission of the SR, the device waits for the reception of the UL grant, otherwise the SR is re-transmitted. According to Fig. 10, given t^* the interval when the device sent the SR, the time interval t when the device is expected to receive the UL grant can be computed as:

$$t = t^* + \left\lceil \frac{T_{TX} + T_{PRO} + T_{TX}}{T_{SR}} \right\rceil. \tag{3}$$

This allows to take into account the SR transmission time, the processing and the transmission time at the base station. The binary parameter $\beta_{k,t}$ indicates whether a device k is receiving an UL grant in the t-th time slot. The $\beta_{k,t}$ is defined as follows:

$$\beta_{k,t} = \begin{cases} 1, & \text{if } \sum_{k'=1}^{k} \alpha_{k',t} \leq N \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where N is the maximum number of SRs that base station can handle simultaneously. The $\beta_{k,t} = 1$ means that base station has scheduled the device for data transmission and device k

will receive the UL grant in the t-th slot. Assuming $\mathcal{A}_{k,t} = \mathcal{A}_{k,t^*}$ is the set of device k's packets for transmission, the buffer will be updated as follows: $\mathcal{B}_k = \mathcal{B}_k \setminus \mathcal{A}_{k,t}$. On the other hand $\beta_{k,t} = 0 \wedge \alpha_{k,t^*} = 1$ means the transmitted SR is not granted and device needs to re-schedule a SR procedure at the next SR opportunity.

We assumed that all devices have the same priority and base station does not prioritize the SR of any device. Hence, the SRs received by the base station are handled in a round robin fashion. Similar assumption holds for data scheduling in (6) and (9).

After the reception of the UL grant, the final step is thus the effective data transmission. According to Fig. 10, given t^* the interval when device receives the UL grant, the first time interval t available to send data can be computed as:

$$t = t^* + \left\lceil \frac{T_{PRO} + T_{TX}}{T_{SR}} \right\rceil. \tag{5}$$

The Equation is allows to take into account the processing time at the device side after the reception of the grant plus the time needed to send the data. The amount of resources needed to transmit the data, i.e., r_k, can be computed by considering the amount of data and the SINR experienced by the device: $r_k = f(|\mathcal{A}_{k,t^*}| \cdot s_k, \sigma_k)$. We exploit the binary parameter $\gamma_{k,t}$ to indicate if a device k is transmitting its data in the t-th time slot. The $\gamma_{k,t}$ is defined as follows:

$$\gamma_{k,t} = \begin{cases} 1, & \text{if } \sum_{k'=1}^{k} r_{k'} \cdot \beta_{k',t^*} \leq R_{UL} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

If $\gamma_{k,t} = 1$, the device has been successfully scheduled to transmit its data in the t-th time slot. We save this information through the set $\mathcal{D}_{k,t}$, hence $\mathcal{D}_{k,t} = \mathcal{A}_{k,t^*}$ and $\mathcal{A}_{k,t^*} = \emptyset$. If not enough resources are available in the t-th time slot, base station will not schedule device k and hence $\gamma_{k,t} = 0 \wedge \beta_{k,t^*} = 1$. In this case, the devices will be scheduled in the next available time slot, hence $\beta_{k,t^*+1} = 1$, $\mathcal{A}_{k,t^*+1} = \mathcal{A}_{k,t^*}$, $\mathcal{D}_{k,t} = \emptyset$ and, finally $\mathcal{A}_{k,t^*} = \emptyset$.

When the session ends, p_k represents the identification of the last packet sent by device k. This means that we can build a set of packets sent by device k as: $\mathcal{P}_k = \{1, 2, \ldots, p_k\}$. For each packet $p \in \mathcal{P}_k$, we can compute the transmission delay as follows:

$$\lambda_{k,p} = \frac{T_{SR}}{2} + (t_{TX} - t_{SR}) \cdot T_{SR} + T_{PRO} \tag{7}$$

where $T_{SR}/2$ takes into consideration the average waiting time before sending the SR, while T_{PRO} represents the processing time at the base station after data reception. In (7), t_{SR} indicates the time slot when the device sent the SR relevant to the packet p and can be computed as

$t_{SR} = t | p \in \mathcal{A}_{k,t}$; similarly, t_{TX} indicates the time slot when the device sent the packet p to the base station and can be computed as $t_{TX} = t | p \in \mathcal{D}_{k,t}$.

C. Legacy Data Reception

The legacy data reception is triggered when the base station receives a packet to be delivered to a device within its coverage area. We assume that base station schedules packets in the downlink direction every T_{DL} ms.

The binary parameter $\delta_{k,t}$ indicates if data has been received ($\delta_{k,t} = 1$) or not ($\delta_{k,t} = 1$) by the base station towards device k at the t-th time slot. A packet that is addressed to device k reaching the base station, is represented by p_k. So, if $\delta_{k,t} = 1$ the base station adds this packet to the buffer of data to be delivered to the device, hence, $\mathcal{B}_k = \mathcal{B}_k \cup \{p_k\}$ and then $p_k = p_k + 1$. If $\delta_{k,t} = 1$, we exploit the set $\mathcal{A}_{k,t} = \mathcal{B}_k$ in order to compute the final delivery delay for each packet.

At the reception of the packet, the base station needs to schedule the data reception on the radio channel. By denoting with t^* the interval when the base station received the data addressed to device k, the first available time interval t for data delivery can be computed as:

$$t = t^* + \left\lceil \frac{T_{PRO} + T_{Al} + T_{TX}}{T_{DL}} \right\rceil \tag{8}$$

where t takes into account the processing time at the base station, the time needed for Transmit Time Interval (TTI) alignment (here denoted with T_{Al}) plus the time spent by the base station to send the packet(s). The amount of resources needed to transmit the data, i.e., r_k, can be computed by considering the amount of data and the SINR experienced by the device: $r_k = f(|\mathcal{A}_{k,t^*}| \cdot s_k, \sigma_k)$. Finally, we exploit the binary parameter $\eta_{k,t}$ to indicate if a device k is receiving its data in the t-th time slot; $\eta_{k,t}$ is defined as follows:

$$\eta_{k,t} = \begin{cases} 1, & \text{if } \sum_{k'=1}^{k} r_{k'} \cdot \delta_{k',t^*} \leq R_{DL} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

If $\delta_{k,t} = 1$, the device has been successfully scheduled to receive its data in the t-th time slot; we save this information through the set $\mathcal{D}_{k,t}$. In detail, we set $\mathcal{D}_{k,t} = \mathcal{A}_{k,t^*}$ and then we set $\mathcal{A}_{k,t^*} = \emptyset$. In case the base station did not schedule device k for data reception in the t-th time slot, i.e., $\eta_{k,t} = 0 \wedge \delta_{k,t^*} = 1$, this means that not enough resources are available in this time slot and as a consequence the devices will be scheduled in the next available one. This means that $\delta_{k,t^*+1} = 1$, $\mathcal{A}_{k,t^*+1} = \mathcal{A}_{k,t^*}$, and, finally $\mathcal{A}_{k,t^*} = \emptyset$.

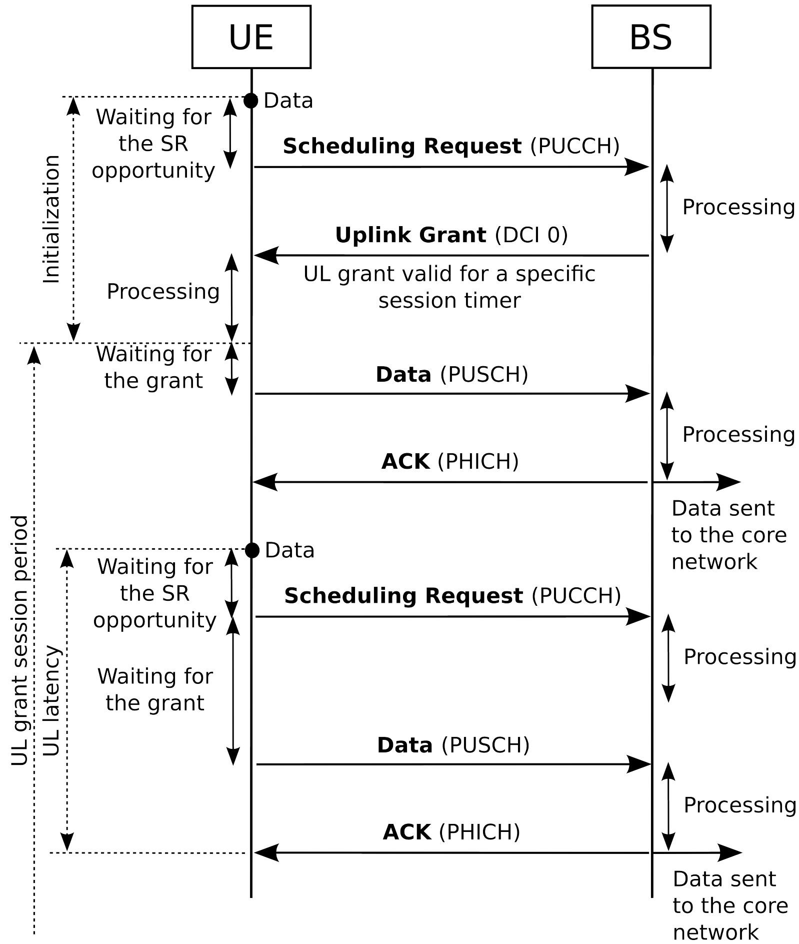


Fig. 12. Our proposed enhanced scheduling request procedure in LTE.

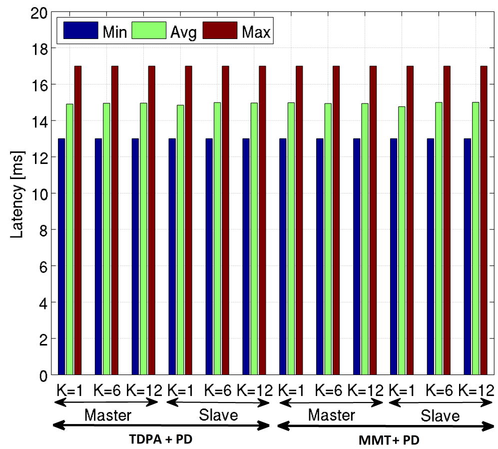
When the session ends, p_k represents the identification of the last packet received by device k. We can thus denote with $\mathcal{P}_k = \{1, 2, \dots, p_k\}$ the set of packets received by device k. For each packet $p \in \mathcal{P}_k$, we can compute the reception delay as follows:

$$\lambda_{k,p} = (t_{RX} - t_{BS}) + T_{PRO} \tag{10}$$

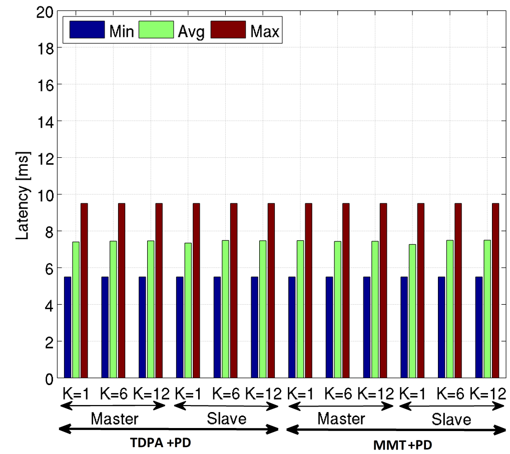
where T_{PRO} represents the processing time at the device side after data reception. In (10), t_{BS} indicates the time slot when the base station receives the packet p to be delivered to device k and can be computed as $t_{BS} = t | p \in \mathcal{A}_{k,t}$. The parameter t_{RX} indicates the time slot when the device receives the packet p from the base station and can be computed as $t_{RX} = t | p \in \mathcal{D}_{k,t}$.

D. Enhanced data transmission

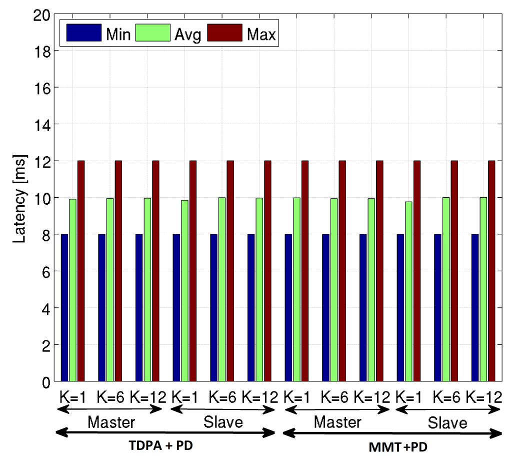
The main novelty of our proposal the introduction of *soft resource reservation* in the SR procedure. As depicted in Fig. 12, the soft resource reservation is composed of two steps. In



(a) Latency in uplink directions by considering the legacy procedures.



(b) Latency downlink directions by considering the legacy procedures.



(c) Latency in uplink direction by considering our proposed procedure.

Fig. 13. Latency in the legacy procedures and in our proposed soft resource reservation.

the first step the device performs the legacy SR procedure for the transmission of the first haptic packet. Therefore, the base station become aware of the packet size relevant to the haptic session. When the base station assigns the UL grant to the device, this grant is *soft reserved* for the device. This means that the device will use this grant for the following transmissions and hence the second step of our proposed procedure is shaped.

In the second step, and when device wants to transmit extra packets, it will send a SR to the base station in order to inform the base station about an incoming data transmission. However, device already knows the UL grant that is softly reserved for its transmission and hence can directly transmit its data in the UL grant assigned during the previous round. Considering the fact that, potentially, the device could generate packets every 1ms, there might be multiple packets for transmission at the device. Therefore, given the SR periodicity T_SR, the reserved resources are allocated in order to accommodate T_SR packets. For example, in the case of one transmission every 1ms, the device will send T_SR packets every T_SR ms. Hence, the soft UL grant will be composed of $r_k = f(|T_SR| \cdot s_k, \sigma_k)$ RBs.

The term "*soft*" in this procedure aims to underline the main difference with respect to the semi-persistent scheduling. In case the base station does not receive any SR from the device, it is aware that the device does not have any data to send and can thus use the resources reserved for this device for other communications within the cell. The UL grant configuration step can be repeated in case of changes such as the exploitation of a different control scheme or changes in the channel quality of the device.

We now present the model of our enhanced uplink transmission procedure, by focusing on the period when the soft UL grant is active. The buffer of each device, i.e., $\mathcal{B}_k$, contains the list of packets to be transmitted. Similar to the above, we consider that each packet has an identification, denoted with p_k. When a packet generated from the application layer reaches the buffer of data transmission, we consider that the novel packet is added to the buffer, i.e., $\mathcal{B}_k = \mathcal{B}_k \cup \{p_k\}$, and then the identification of the next packet to be sent is increased, i.e., $p_k = p_k + 1$. Every SR period, the device checks its own buffer and sends a SR in case it is not empty: in this case, the device basically informs the base station about the number of packets to be sent, in order to make the base station aware of how many resources of the soft UL grant will be used. We exploit the binary parameter $\alpha_{k,t}$ to indicate if the device k performed a SR at the t-th SR period, and $\alpha_{k,t}$ is defined as in (2). If $\alpha_{k,t} = 1$, i.e., the device informed the base station that it is going to transmit $\mathcal{A}_{k,t} = \mathcal{B}_k$ packets in its soft UL grant.

After the transmission of the SR, the device does not need to wait for the UL grant reception, as it already knows the resources allocated to it. By denoting with t^* the interval when device sent the SR, $t = t^* + 1$ represents the instant when the device will be able to send its data by exploiting the resources of the soft UL grant. This means that $\mathcal{D}_{k,t} = \mathcal{A}_{k,t^*}$.

When the session ends, p_k represents the identification of the last packet sent by device k. This means that we can build a set of packets sent by device k as: $\mathcal{P}_k = \{1, 2, \ldots, p_k\}$. For each packet $p \in \mathcal{P}_k$, we can compute the transmission delay as follows:

$$\lambda_{k,p} = \frac{T_{SR}}{2} + (t_{TX} - t_{SR}) \cdot T_{SR} + T_{PRO} \tag{11}$$

In (7), t_{SR} indicates the time slot when the device sent the SR relevant to the packet p and can be computed as $t_{SR} = t|p \in \mathcal{A}_{k,t}$. Similarly, t_{TX} indicates the time slot when the device sent the packet p to the base station and can be computed as $t_{TX} = t|p \in \mathcal{D}_{k,t}$.

V. Simulation Results

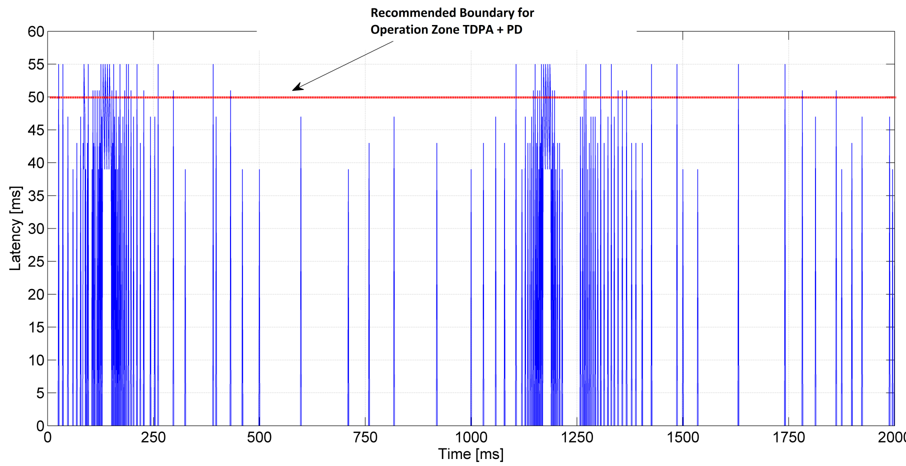
In the simulated scenario, we assume all devices support QPSK, and this means that a device needs two RBs in case of a packet of 48 bytes (slave, the MMT+PD method), and one RB otherwise [15]. We also assume a delay of 1 ms in the core network [24] for regional communications and 90 ms for transatlantic communications [26]. Based on the QoE analysis in Section III-B, we consider 50 ms as a threshold which, if exceeded, involves severe QoE degradation in case of the TDPA+PD scheme. The system could react by switching to the MMT+PD scheme to avoid a severe QoE degradation as this scheme is more robust to high delay.

A. One-Way Communication between Master and Slave Devices

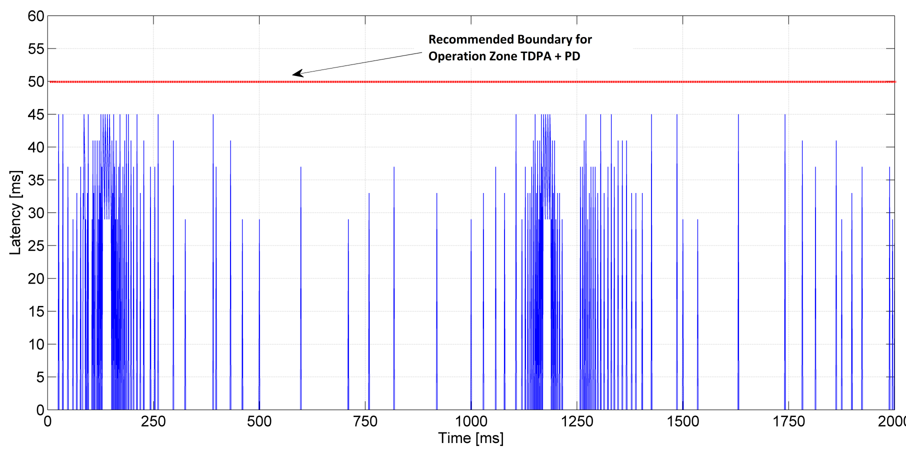
Fig. 13(a) and 13(b) analyze the latencies in the UL and DL directions and for both master and slave sides by considering the legacy procedures. We study three cases, i.e., when the session is composed of 1, 6, and 12 master-slave pairs, and several interesting conclusions can be drawn from this study.

First, the increase in the number of devices does not affect the performance, due to the fact that haptic traffic does not represent a large source of traffic load for the cell. The second aspect to remark is that the performance does not vary by considering different control schemes, or in other words, both the legacy procedure as well as the proposed scheme can handle the traffic characteristics of the teleoperation sessions. Observing from Fig. 13(a) and 13(b), the UL latency is between 13 ms and 17 ms, while the DL one varies from 5.5 ms to 9.5 ms.

Fig. 13(c) shows the UL performance achieved using our proposed soft reservation procedure. It can be seen that the latency varies from 8 ms to 12 ms, i.e., a reduction ranging from 30% to 40% w.r.t. the legacy procedure. The improvement is obtained by considering that the device does not need to wait for the UL grant, while it send data directly after the transmission of the SR to effectively reserve the already soft reserved RBs.



(a) Legacy Procedure



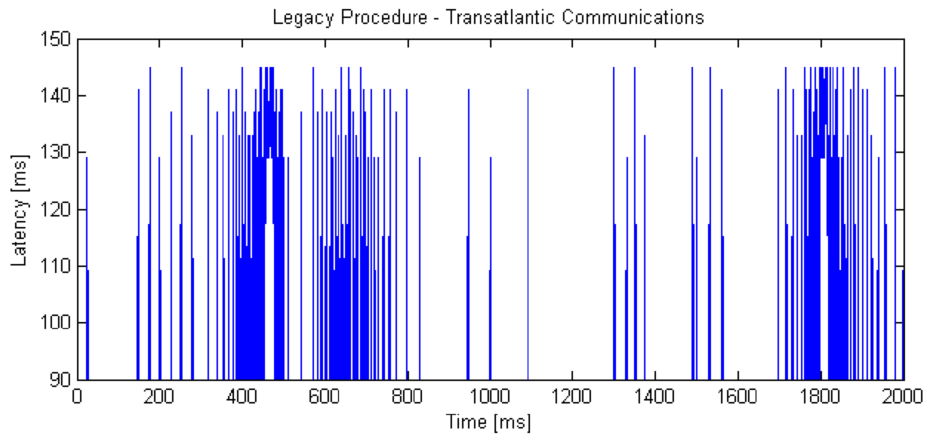
(b) Proposed soft reservation strategy

Fig. 14. Round trip time for the case of one subject, using the TDPA+PD scheme.

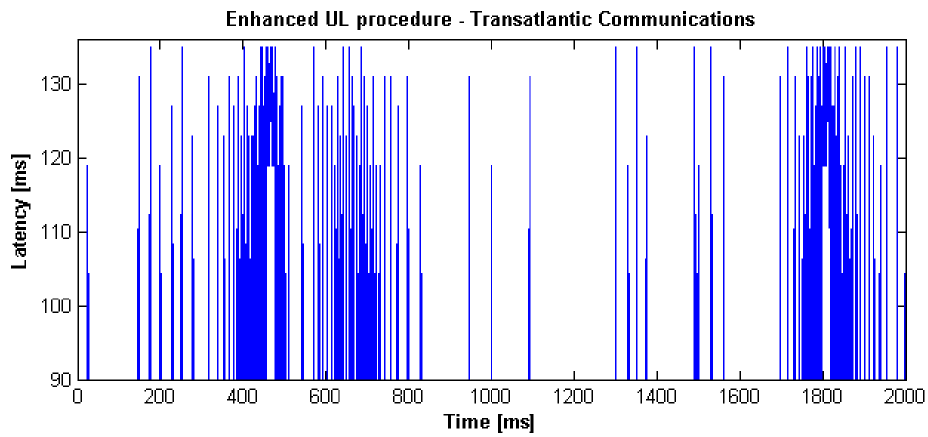
B. Round-Trip Communication Path, Single Haptic Device

After the analysis of the single directions of the haptic session, we now focus our attention on the master-slave-master path, by analyzing the round trip time (RTT) experienced during the session. In this case we look into both control schemes. It is worthwhile reminding that the TDPA+PD scheme is suitable for low latency (with operation boundary of below 50 ms) and the MMT+PD scheme with higher latency threshold.

Fig. 14 shows the RTT obtained with the legacy transmission procedures as well as with our proposed UL strategy, in the case of one subject. Observing from this figure, the legacy procedure experiences RTT increases by up to 55 ms during the bursty period. This means that the legacy



(a) Legacy procedure



(b) Proposed soft reservation strategy

Fig. 15. Round trip time for the case of one subject, using the MMT+PD scheme, over transatlantic communications.

procedures would involve a severe QoE degradation during the bursty interval in case of the TDPA+PD method. Therefore, using the legacy procedures, when a master device detects the beginning of a bursty period, it could potentially trigger a change in the control scheme (from the TDPA+PD method to the MMT+PD method) in order to avoid QoE degradation. Furthermore, Fig. 14 shows the RTT obtained by exploiting our proposed UL transmission strategy, where the RTT delay is reduced to a maximum value of approximately 45 ms, i.e., lower than the considered QoE degradation threshold of approximately 50 ms. Therefore, using the proposed soft resource reservation, the teleoperation can remain within the TDPA+PD method and operate with a high QoE.

Using the same simulation setup, and operating with the MMT+PD scheme, similar RTT delay can be observed both by the legacy and the proposed reservation strategy. Moreover, using the enhanced UL strategy, the RTT can be bounded to acceptable threshold by the TDPA+PD method. We can therefore conclude that the TDPA+PD method is sufficient for connectivity over short distances (within a country for example), as long as the soft resource reservation is deployed on the wireless edge of the connection. On the other hand, the MMT+PD method can be used over global long distance connectivity, where the imposed latency by the Internet is significant compared with the latency experienced at the wireless edge. For example, assuming a maximum of 90 ms for transatlantic communication between London and New York [26], traffic generated by the MMT+PD method and sent over the wireless connection, experience the RTT of up to 145 ms and 135 ms for the legacy procedure and the proposed UL enhanced procedure respectively (also seen in Fig. 15).

C. Round-Trip Communication Path, Multiple Haptic Devices

Fig. 16 shows the ECDF of the RTT delay for the TDPA+PD scheme for all 12 subjects. This plot shows how effectively our proposed strategy maintains a round trip delay below 50 ms even in the case of multiple haptic sessions active in the cell, while in the legacy procedure, more than 35% of the packets experience RTT delay higher than 50 ms, that refers to the low QoE by human operator (the master side) for almost 35% of the whole duration of the session. With a similar simulation setup, the results for the MMT+PD scheme are the same, and therefore the plot is not presented. As explained earlier, the MMT+PD scheme can be considered relevant for long distance communications with longer latency on the backbone network.

D. Analysis of 5G and Shorter TTI

Another analysis we present takes into consideration the exploitation of shorter Transmit Time Interval (TTI), envisioned in 5G systems to be reduced to support low latency services. Hence, we consider a TTI duration of 0.2 ms. Using this TTI, Tables IV and V analyze the average UL and DL latency consecutively by elaborating different sources of latency. Observing from these tables, the legacy procedure can achieve a latency of approximately 9.2 ms and 6.3 ms in UL and DL directions, respectively. Our proposed procedure can, however, reduce the UL latency further down to 4.5 ms.

E. Analysis of QoE Improvement

Let us recall the QoE performance for different control schemes with respect to communication delay, as shown in Fig. 17. Since the proposed soft reservation strategy can reduce the round-

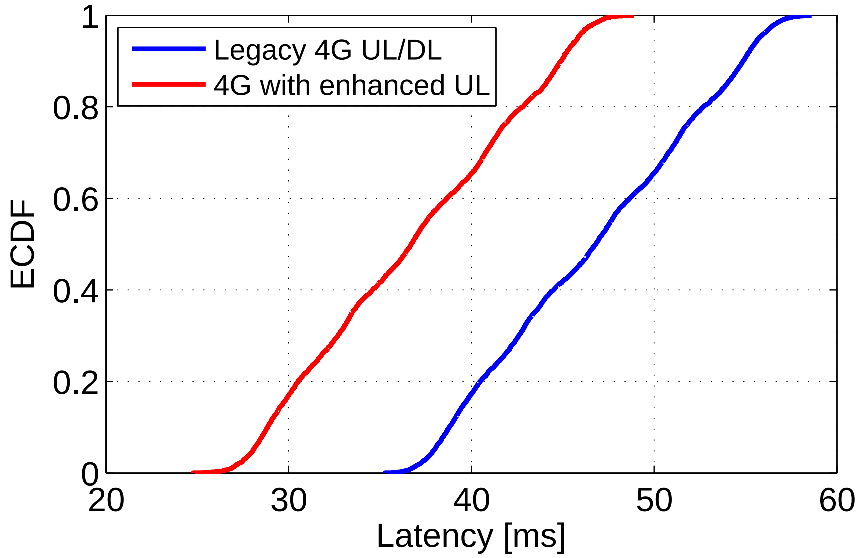


Fig. 16. Empirical cumulative distribution function (ECDF) of the round trip (master-slave-master) time delay for the case of 12 subjects.

TABLE IV
ANALYSIS OF UL LATENCY WITH SHORTER TTI

Step	Legacy TTI 1ms	Legacy TTI 0.2ms	Proposed TTI 1ms	Proposed TTI 0.2ms
Waiting for SR	2.5ms	0.1ms	2.5ms	0.1ms
SR transmission	1ms	0.2ms	1ms	0.2ms
BS Processing	3ms	3ms	-	-
UL grant	1ms	0.2ms	-	-
UE Processing	3ms	3ms	2.5ms	1ms
Data transmission	1ms	0.2ms	1ms	0.2ms
BS Processing	3ms	3ms	3ms	3ms
Total	14.5ms	9.7ms	10ms	4.5ms

trip delay by 10 ms compared with the legacy procedure, QoE ratings of the TDPA+PD and the MMT+PD schemes will achieve a maximal improvement of 0.6 and 0.1, respectively, by using the proposed strategy. We can observe that the QoE improvement of the TDPA+PD method is higher than for MMT+PD method. This is because the performance of the TDPA+PD method is mainly influenced by communication delay, while the performance of the MMT+PD method depends strongly on the model accuracy, which is barely affected by the communication delay for services with low object dynamics.

TABLE V

ANALYSIS OF DL LATENCY WITH SHORTER TTI

Step	TTI 1ms	TTI 0.2ms
Processing	3ms	3ms
TTI alignment	0.5ms	0.1ms
DL data transmission	1ms	0.2ms
Data encoding	3ms	3ms
Total	7.5ms	6.3ms

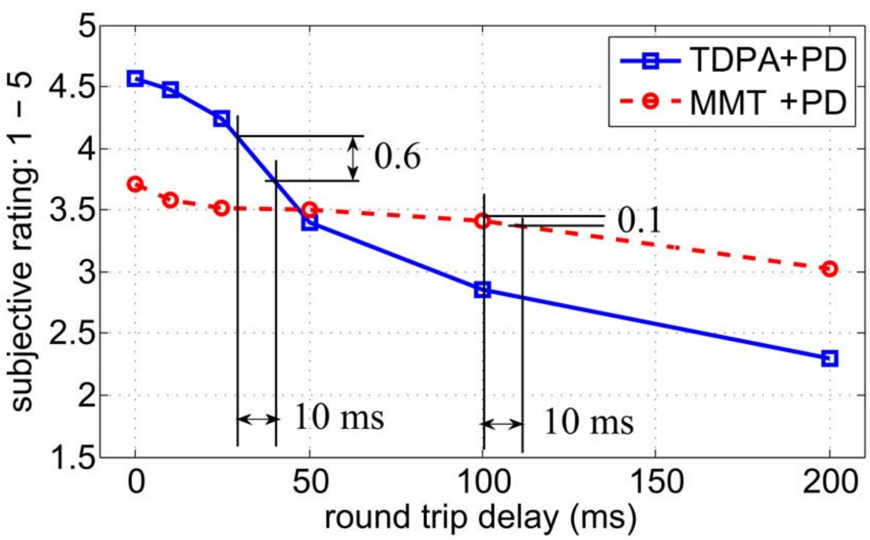


Fig. 17. QoE enhancement by using the proposed soft reservation strategy.

F. Delay-aware Control Scheme Recommendations

As discussed in Section II and III, the TDPA+PD and the MMT+PD methods should be adopted at different delay ranges in order to achieve a high teleoperation performance. We can conclude from Fig. 17 that the TDPA+PD method is recommended for teleoperation with a delay less than 50 ms, while the MMT+PD method is favorable for teleoperation larger than 50 ms. This way, the teleoperation with TDPA+PD scheme should be performed at the edge of the communication network for short-distance (e.g. domestic or regional) teleoperation applications, while the MMT+PD scheme is recommended for medium/long distance (e.g. international or intercontinental) teleoperation applications.

More importantly, we can observe from Fig. 16 that the proposed soft reservation strategy reduces the round-trip delay by 10 ms compared with the legacy scheme. This means that if the estimated round-trip delay of the legacy approach is 60 ms, the round-trip delay will become

50 ms when using the proposed strategy. Since the QoE performance of the TDPA+PD approach is higher than for the MMT+PD method for low latency scenarios, when the proposed algorithm is adopted, we will recommend the TDPA+PD scheme for teleoperation with a round-trip delay of less than 60 ms, and the MMT+PD scheme for the rest cases.

TABLE VI
CONTROL SCHEME RECOMMENDATION WITH RESPECT TO THE ROUND-TRIP DELAY

Estimated round-trip delay when using the legacy approach (ms)	Legacy	Proposed
$<$50	TDPA+PD	TDPA+PD
50$\sim$60	MMT+PD	TDPA+PD
$>$60	MMT+PD	MMT+PD

VI. CONCLUSION

In this paper, we proposed a soft reservation strategy for the UL scheduling of LTE networks aiming at providing ultra-low-delay services to various teleoperation scenarios. The development of the proposed strategy highly depends on the exploration of the characteristics of the TDPA+PD and MMT+PD schemes (e.g. the tolerable delays, the busty behavior and irregular update rate of haptic traffic). The simulation results illustrated the efficiency of the proposed soft reservation strategy which reduces the round-trip delay by an average of 10 ms compared with the legacy solution. Because of the delay-sensitive nature of teleoperation systems, this achievement will bring admirable QoE improvements to teleoperation under different application scenarios.

In addition, we also concluded from the simulation results that the TDPA+PD scheme was suitable for low-latency (short-distance) teleoperation with dynamic objects and interaction between the master and the slave, while the MMT+PD scheme was able to deal with relatively larger communication delays (and hence can be applied to medium or long distance application scenarios). These results can be considered as a valuable guidance to appropriately select control schemes for teleoperation under different environment dynamics and communication delays.

REFERENCES

[1] F. Xia, L. T. Yang, L. Wang, and A. Vinel, "Internet of things," *International Journal of Communication Systems*, vol. 25, no. 9, p. 1101, 2012.

[2] G. P. Fettweis, “The tactile internet: applications and challenges,” *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, 2014.

[3] M. Dohler, “The tactile internet iot, 5g and cloud on steroids,” in *IET Conference Proceedings*, The Institution of Engineering & Technology, 2015.

[4] W. R. Ferrell and T. B. Sheridan, “Supervisory control of remote manipulation,” *IEEE Spectrum*, vol. 4, no. 10, pp. 81–88, 1967.

[5] D. A. Lawrence, “Stability and transparency in bilateral teleoperation,” *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 624–637, 1993.

[6] J. E. Colgate and G. G. Schenkel, “Passivity of a class of sampled-data systems: Application to haptic interfaces,” *Journal of Robotic Systems*, vol. 14, no. 1, pp. 37–47, 1997.

[7] X. Xu, C. Schuwerk, B. Cizmeci, and E. Steinbach, “Energy prediction for teleoperation systems that combine the time domain passivity approach with perceptual deadband-based haptic data reduction,” *IEEE Transactions on Haptics*, accepted for publication.

[8] X. Xu, B. Cizmeci, A. Al-Nuaimi, and E. Steinbach, “Point cloud-based model-mediated teleoperation with dynamic and perception-based model updating,” *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 11, pp. 2558–2569, 2014.

[9] P. Mitra and G. Niemeyer, “Model-mediated telemanipulation,” *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 253–262, 2008.

[10] J. Ryu, “Bilateral control with time domain passivity approach under time-varying communication delay,” in *16th IEEE International Conference on Robot and Human Interactive Communication*, (Jeju, Korea), August 2007.

[11] J. Ryu, J. Artigas, and C. Preusche, “A passive bilateral control scheme for a teleoperator with time-varying communication delay,” *Elsevier Journal of Mechatronics*, vol. 20, no. 7, pp. 812-823, October 2010.

[12] J. Artigas, J.-H. Ryu, C. Preusche, and G. Hirzinger, “Network representation and passivity of delayed teleoperation systems,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, (San Francisco, CA, USA), September 2011.

[13] B. Hannaford, “A design framework for teleoperators with kinesthetic feedback,” *IEEE Transaction on Robotics and Automation*, vol. 5, no. 4, pp. 426-434, August 1989.

[14] B. Willaert, J. Bohg, H. Brussel, and G. Niemeyer, “Towards multi-dof model mediated teleoperation: using vision to augment feedback,” in *IEEE International Workshop on Haptic Audio Visual Environments and Games*, (Munich, Germany), October 2012.

[15] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures.” TR 36.213.

[16] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification.” TR 36.321.

[17] H. Jin, C. Cho, N. O. Song, and D. K. Sung, “Optimal Rate Selection for Persistent Scheduling with HARQ in Time-Correlated Nakagami-m Fading Channels,” *IEEE Transactions on Wireless Communications*, vol. 10, pp. 637–647, February 2011.

[18] J. B. Seo and V. C. M. Leung, “Performance Modeling and Stability of Semi-Persistent Scheduling with Initial Random Access in LTE,” *IEEE Transactions on Wireless Communications*, vol. 11, pp. 4446–4456, December 2012.

[19] S. Andreev, A. Larmo, M. Gerasimenko, V. Petrov, O. Galinina, T. Tirronen, J. Torsner, and Y. Koucheryavy, “Efficient small data access for machine-type communications in LTE,” in *2013 IEEE International Conference on Communications (ICC)*, pp. 3569–3574, June 2013.

[20] P. Hinterseer, S. Hirche, S. Chaudhuri, E. Steinbach, and M. Buss, "Perception-based data reduction and transmission of haptic data in telepresence and teleaction systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 588–597, 2008.

[21] C. Passenberg, A. Peer, and M. Buss, "A survey of environment-, operator-, and task-adapted controllers for teleoperation systems," *Mechatronics*, vol. 20, no. 7, pp. 787-801, October 2010.

[22] K. Hunt and F. Crossley, "Coefficient of restitution interpreted as damping in vibroimpact," *ASME Journal of Applied Mechanics*, vol. 42, no. 2, pp. 440-445, June 1975.

[23] X. Xu, C. Schuwerk, and E. Steinbach, "Passivity-based model updating for model-mediated teleoperation," in *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on*, pp. 1–6, IEEE, 2015.

[24] 3GPP, "Study on latency reduction techniques for LTE." TR 36.881.

[25] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," *IEEE Network*, vol. 30, pp. 46–53, July 2016.

[26] Verizon Enterprise Solutions, "IP Latency Statistics ." http://www.verizonenterprise.com/about/network/latency/.

[27] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives," *IEEE Communications Surveys Tutorials*, vol. 16, pp. 4–16, First 2014.

[28] R. Kwan and J. C. Ikuno, "Effective HARQ code rate modelling for LTE," *Electronics Letters*, vol. 49, pp. 462–464, March 2013.

[29] M. R. Raghavendra, S. Nagaraj, K. V. Pradap, and P. Fleming, "Robust Channel Estimation and Detection for Uplink Control Channel in 3GPP-LTE," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pp. 1–5, Nov 2009.

[30] M. Centenaro and L. Vangelista, "HARQ in LTE uplink: A simple and effective modification suitable for low mobility users," in *2015 IEEE International Conference on Communications (ICC)*, pp. 3763–3769, June 2015.

[31] D. Jiang, H. Wang, E. Malkamaki, and E. Tuomaala, "Principle and Performance of Semi-Persistent Scheduling for VoIP in LTE System," in *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 2861–2864, Sept 2007.