On Human-in-the-Loop CPS in Healthcare: A Cloud-Enabled Mobility Assistance Service

Ricardo C. de Mello†*, Mario F. Jimenez¶, Moises R. N. Ribeiro†, Rodrigo Laiola Guimarães‡, and Anselmo Frizera-Neto†

†Electrical Engineering Graduate Program, Federal University of Espírito Santo, Av. Fernando Ferrari, 514, Goiabeiras, Vitoria, 29075-910, Brazil. E-mails: moises@ele.ufes.br, frizera@ieee.org
¶Bioengineering Graduate Program, El Bosque University, Bogotá, Colombia. E-mail: mariof.jimenez@gmail.com
‡Postgraduate Program in Computer Science, Federal University of Espírito Santo, Av. Fernando Ferrari, 514, Goiabeiras, Vitoria, 29075-910, Brazil. E-mail: rlaiola@inf.ufes.br

(Accepted January 01, 2019)

SUMMARY
Despite recent advancements on cloud-enabled and human-in-the-loop cyber-physical systems, there is still a lack of understanding of how infrastructure-related quality of service (QoS) issues affect user-perceived quality of experience (QoE). This work presents a pilot experiment over a cloud-enabled mobility assistive device providing a guidance service and investigates the relationship between QoS and QoE in such a system. In our pilot experiment, we employed the CloudWalker, a system linking smart walkers and cloud platforms, to physically interact with users. Different QoS conditions were emulated to represent an architecture in which control algorithms are performed remotely. Results point out that users report satisfactory interaction with the system even under unfavorable QoS conditions. We also found statistically significant data linking QoE degradation to poor QoS conditions. We finalize discussing the interplay between QoS requirements, the human-in-the-loop effect, and the perceived QoE in healthcare applications.

KEYWORDS: Healthcare; Cyber-physical systems; Assistive robotics; Communication networks; cloud robotics; QoS; QoE.

1. Introduction
Healthcare is one of the areas in which robotics is being applied, and future medical facilities are likely to be equipped with devices such as robotic caretakers and autonomous stretchers, which will be also sharing space with patients and health professionals. In this context, some recently developed assistive robots (e.g., smart walkers and wheelchairs) can be classified into cyber-physical systems (CPS) as they combine sensing, communication, and control/computing to interact with a physical entity. Such class of healthcare robots, apart from most robotic systems, usually interacts at physical and cognitive levels with patients to provide functionalities where reliability and safety are essential.

The exploitation of the cloud robotics paradigm can expand the capabilities and features offered by assistive CPS. Migrating robot’s processing capabilities to the cloud leads to lighter and less expensive robots. Furthermore, through the access of big data sets and massively parallel computing, heavy adaptive control algorithms can be employed to command multiple connected robots in
dynamic environments. In healthcare facilities, mobility assistive devices can also benefit from those concepts, as cooperation and information sharing among robots are enhanced, allowing for improved navigation.

In this context, cloud-enabled CPS for patient mobility assistance will face challenges related to end-to-end (E2E) quality of service (QoS) of a complex system combining communication networks and cloud computing. Healthcare CPS are expected to require from the network uninterrupted wireless connectivity, throughput guarantee, and stringent latency and jitter features.

Another intricate problem arises when objective QoS metrics need to be translated into indexes of quality of experience (QoE) to the end-user. When interacting with a robotic system during mobility assistance, the user is often in-the-loop from the control algorithms’ point of view. As shown by Pons, stability issues arise in such human-in-the-loop systems. Latency onto the control loop greatly impacts the overall stability, affecting both physical and cognitive interaction between human and CPS.

Nevertheless, most of the works conducted upon human-in-the-loop CPS do not consider direct physical interaction between human and CPS. Lee et al. point the challenges and directions for the recent generation of medical CPS, while Shah et al. state that approaches to QoS in healthcare services are scarce in the literature as this is an emerging field. Hammer et al. consider that QoS needs to be carefully controlled in CPS and its impact on QoE needs to be taken into account. The challenges for the control of human-in-the-loop CPS are listed by Munir et al., who state that the understanding of the complete spectrum of human-in-the-loop control is needed as more sophisticated applications appear. Thus, a better understanding into QoS influence over QoE is needed not only in CPS for healthcare but also in other systems heavily influenced by the human-in-the-loop effect such as car driver assistance and wearable exoskeletons.

In sensitive applications, such as patient mobility assistance, cloud services cannot be provided on a best-effort basis, thus QoE is likely to impose the QoS requirements for those future network/cloud infrastructures and not the other way round. However, the interaction of cloud technologies, healthcare robots, and patients is yet to be fully understood, especially in human-in-the-loop CPS that explores the physical interaction between human and robot.

This paper brings those discussions to light by exploring the deployment of a cloud-enabled human-in-the-loop CPS for mobility assistance. To this end, we present a pilot experiment over a healthcare use case in which a cloud-enabled device assists users by guiding their displacement throughout the environment. We aim at growing an understanding on the effects of variable network QoS conditions over perceived QoE in human-in-the-loop CPS. Therefore, in our pilot experiment, 33 volunteers tested and evaluated our system under different QoS conditions, providing valuable insight of how such a CPS can be affected by network constraints. This work reports our findings and discusses the challenges for migrating the control of healthcare CPS to cloud platforms as well as the intricate relation between QoS and QoE in such a class of systems.

The remainder of this work is organized as follows. Section 2 presents an overall view on current challenges and applications for cloud-enabled CPS for mobility assistance. Section 3 describes our case study scenario and Section 4 introduces CloudWalker, our cloud-enabled CPS. Section 5 describes our pilot experiment, the adopted protocols, and results. Section 6 presents a discussion about lessons learned from the pilot experiment. We finalize with closing remarks pointing to the need for broader perspectives when assessing QoE in such class of human-in-the-loop CPS.

2. An Overview on Cloud-Enabled CPS for Mobility Assistance

Placing cloud computing at the center of a connected healthcare system can enable the systematic use of multiple heterogeneous devices at different facilities, allowing for the so-called Healthcare as a Service (HaaS) to arise. Figure 1 depicts the big picture of such system, where robotic devices exploit the cloud to provide services for patients and to feed databases with information that can be used by remote libraries, other robots, adaptive control algorithms, and health reports generators.

Currently, cloud computing has been reported to empower several healthcare CPS in applications such as patient localization, monitoring and status recording, and construction of databases and ontologies. There are also initiatives in social robotics to design cloud-enabled companion and caregiver robots for the elderly, as in Fiorini et al., leveraging the cloud capabilities to expand the current state-of-the-art. Regarding robotic devices employed for assistance and rehabilitation, cloud-enabled applications have been used to collect data from exoskeletons and to allow on-line remote monitoring of therapy.
On human-in-the-loop CPS in healthcare

Fig. 1. Big picture overview of a connected healthcare system.

Fig. 2. Connected healthcare system: (a) healthcare services enabled by communication and computation technologies (black arrows hints a navigation assistance service discussed in this paper); (b) “cloudification” scale of current cloud-enabled solutions.

Assistive mobility devices are beginning to benefit from cloud computing concepts. In Fu et al., smart phone sensors collect navigational information of a wheelchair and transmit them to the cloud for storage. In Salhi et al., robotics concepts are integrated to the cloud to share multiple wheelchair positions and the map of the environment to feed the wheelchairs’ local controllers. Nevertheless, such class of applications do not fully explore the potentials of CPS nor of cloud computing. To better address this topic, we’ll now briefly discuss about potential applications that can be empowered by the cloud and current challenges and limitations.

2.1. Potential cloud-empowered applications for mobility assistance

Assistive robots in general can benefit in multiple ways from this new ecosystem composed of network infrastructure and cloud computing. Figure 2(a) exemplifies this by linking assistive devices’ main features, and communication and computation features as enablers to compose healthcare services. Black arrows highlight the path composing a guiding service that will be latter discussed in this paper: a navigation system exploiting wireless connectivity, network’s low latency, and...
cloud platform’s high processing power to provide a real-time mobility assistance healthcare service. Navigation features can largely benefit from cloud computing as they often require algorithms with high computation cost, such as laser or vision-based simultaneous localization and mapping (SLAM).\textsuperscript{31,32}

Assistive devices’ guiding features might help patients that do not know the facility well or that present cognitive problems (e.g., following instructions). Applications on cloud platforms might leverage multiple robots’ inputs to construct and update global maps dynamically, even pointing to more crowded areas to be avoided by path planning algorithms. The cloud also has the potential to empower human–robot interaction by enabling improved human intent recognition and shared control. This could potentially lead to devices that can continuously cope with users as gait conditions change over the years.

Among robotic devices designed for patient mobility assistance, there are the so-called smart walkers. Such robotic devices have been designed to improve pathological gait using support platforms for upper limbs and the exploitation of the user’s own locomotion capabilities.\textsuperscript{33} Smart walkers usually present some, or even all, of the features illustrated in Fig. 2(a) and recent advancements in control and interaction techniques demand an ever-increasing embedded processing power. Thus, smart walkers evolution into cloud robotics concepts enables another paradigm shift for those devices in order to expand offered features and capabilities.

To the best of our knowledge, the ANG-MED robot\textsuperscript{9} is the only smart walker in the literature to use cloud-based services. It is a passive walker (i.e., without motors on the wheels) that uses cloud-based applications to provide authentication, database, and brake and motion control services for the caregiver. The ANG-MED has been evaluated by care professionals with positive results and Tsardoulas et al.\textsuperscript{9} state that the viability of such applications was verified during experimentation. Nevertheless, to unleash a generation of cloud-enabled mobility assistive devices, some challenges and limitations must be overcome.

2.2. Current limitations and challenges

As can be seen in Fig. 1, communication technologies play an important role in future healthcare data transmission, storage, processing, and eventual feedback in real time. Nevertheless, network and cloud-related constraints have a direct impact on cloud-based applications. Current research points that to mitigate such issues, edge cloud architectures can help providing stringent latency and jitter,\textsuperscript{34} while software-defined networking (SDN) is expected to provide an uniform mechanism for E2E network programmability.\textsuperscript{35} Nevertheless, concerns have been raised about SDN’s ability in low-latency reconfiguration operations toward the core of the networks\textsuperscript{36} and there is still a lack of studies addressing such technologies as enablers of human-in-the-loop CPS applications.

Wireless technologies are also important building blocks in such applications, and 5G technology will be a key enabler for providing connectivity in mobile real-time applications. The goals for critical applications in 5G are to provide latency figures as strict as 1 ms alongside $10^{-5}$ unavailability ratios,\textsuperscript{37} which would meet the stringent requirements set by human-in-the-loop CPS. Moreover, current off-the-shelf technologies such as the WiFi have been demonstrated to support connectivity in mobile applications.\textsuperscript{38} Nevertheless, control for critical applications must be designed taking into account possible reliability issues in wireless networks.

Current cloud-based solutions for healthcare devices, in particular the ones used in rehabilitation and assistance, tend to employ small degrees of cloud computing to control these systems. Figure 2(b) illustrates some of the works referenced in this paper on a scale of how much of their control depends on cloud computing. The cloud robotics trend points toward a migration to the cloud, or “cloudification”, of solutions, especially those involving complex algorithms. The case study that will be presented in this work illustrates an extreme case of this control “cloudification”, in which there is no embedded intelligence, and is placed on the rightmost end of Fig. 2(b) spectrum. In any case, there is a timely need for understanding how communication constraints might impact on such applications from an end-user’s point of view.

The concept of QoE arises in a context of telecommunication applications, where both human senses and cognitive features are considered as part of the information processing.\textsuperscript{39} Despite advances in QoE evaluation techniques, there are still challenges in properly weighting ease of use, comfort, and other subjective parameters into objective QoE metrics.\textsuperscript{40} There are also challenges in mapping QoE into QoS requirements to the underlying network for every new application.\textsuperscript{13} Although
cloud-enabled CPS may use previous QoS–QoE studies as stepping-stones, there are specific issues concerning healthcare applications; especially when the end-user is heavily embedded into the system dynamics. Therefore, there might be trade-off solutions balancing network/cloud, CPS dynamics, human physical and cognition features to define, in the end, reasonable network requirements for cloud-enabled human-in-the-loop CPS. To investigate this topic, the next section presents an illustrative case study scenario in which a cloud-enabled CPS provides a mobility assistance service.

3. Healthcare Case Study Scenario: Virtual Trails for Mobility Assistance

To illustrate a mobility assistance service, we present a case study upon a scenario where hospitalized patients and nursing home inhabitants may benefit from the use of cloud-enabled assistive devices. Patients with impaired mobility or cognitive dysfunctions might need help during displacement across clinic, therapy, and accommodations. Particularly in nursing home environments, assistive devices are required on a daily basis by many permanently impaired residents and guiding features can further assist user displacement. Such scenario is used as basis for our pilot experiment, which will be later discussed in this work.

Devices such as smart walkers can exploit emergent technologies in order to safely assist on navigation by guiding user displacement to a desired location. The device should be able to perform localization and to leverage mapping functionality to generate a path toward destination. Such path can be conceived as a virtual trail in which the user can walk along without being able to deviate from it, almost as if in a rail track. Figure 3 illustrates such scenario by showing a person making use of a smart walker inside a facility with corridors and obstacles. The virtual trail established by the device marks the path to be followed, and the walker slowly turns to keep on the trail as the user walks forward.

This guiding feature can be used to assist the navigation of individuals with reduced independence and/or balance, as a consequence of vision and mobility impairments caused by cerebral diseases, cardiopulmonary and musculoskeletal diseases, or even stroke. Furthermore, the haptic feedback indicates the trail to follow, a useful feature for those with orientation and localization problems. As interactions between patient and assistive CPS must be built on bases of navigation applications and coordination with other users, heavy control algorithms may be required, which can be instantiated in a centralized way over the cloud. Implementation details are discussed later in Section 5.1.

Considering the described scenario in the following section, we propose the CloudWalker: a system architecture that enables smart walkers’ paradigm shift into cloud and network-enabled walkers.

4. The CloudWalker Architecture

CloudWalker system architecture allows the use of smart walkers as cloud-enabled CPS, encompassing a system that involves physical and cognitive human–robot interaction with mobility impaired individuals. In such CPS, user’s motion intents are perceived by embedded sensors on a smart walker,
On human-in-the-loop CPS in healthcare

Sensors:
- Movement Intent
- Walker’s speed and orientation

Actuators:
- Wheel Velocities

Network:
- Availability
- Delay
- Packet loss

Control Algorithms:
- Virtual Trail
- Admittance Control
- Odometry

Fig. 4. CloudWalker overview: (a) system architecture, main control loops, and perceived QoE and QoS; (b) UFES Smart Walker in detail.

which communicates with a cloud platform in order to provide information and also receives the appropriate control signals to physically command the robot.

The proposed architecture is described in Fig. 4(a). Physical interaction between user’s upper body and walker is measured by force sensors that capture user’s motion intent. Collected data are sent through the network to a remote cloud platform running the control applications. Data are processed by the control algorithms in the cloud, which sends back to the walker’s embedded computer, the generated control signals to command the actuators, resulting in the motion of the user-walker aggregate.

On our particular system, it would be desirable to implement critical lower-level controllers on the device to maintain usability and stability even under situations where there is no network connectivity, delegating to the cloud only high-level computationally expensive tasks. As the main goal of this study is to access QoS impacts over the QoE on cloud-enabled human-in-the-loop CPS, our architecture does not envision any local intelligence on the device. This inserts communication constraints directly into the control loop, an issue that has been thoroughly addressed in distributed networked control systems literature and that, in our study, maximizes the impacts of QoS on system performance and user-perceived QoE. Therefore, safety measures are necessary during implementation to mitigate possible instability impacts.
CloudWalker architecture is based on an in-house-developed smart walker.\textsuperscript{33} The UFES Smart Walker (detailed on Fig. 4(b)) is a robotic platform equipped with sensors and actuators designed to actively assist on patient’s navigation throughout the environment. It can interact with the user in several ways, such as interpreting the physical interaction forces on the forearms supports and even map the environment by deploying the frontal laser scanner. An embedded computer centralizes acquisition of sensor data, communication with the network, and the forwarding of received control signals to wheel actuators. This application is integrated into a real-time architecture and for the case study, only the force sensors were used to detect patient’s intentions to move. For such application, communication is based on the exchange of 40 B UDP packets between the walker and the remote platform.

As depicted in Fig. 4(a), CloudWalker architecture encompasses three control loops:

- A local control loop, where motor’s dynamics must be taken into account during system’s response.
- A global control loop, with robot’s sensors feeding remotely allocated controllers in order to command actuators accordingly to patient’s intents captured by the force sensor located on the support frame.
- A cognitive loop, where the patient reacts to the smart walker movements by instinctively changing forces applied to the support frame.

The local control loop, related to the walker dynamics, is the faster one, and a time constant is empirically obtained around 100 ms. Thus, the embedded electronics sampling time is set at 10 ms, 1/10 of the verified time constant of the system, as recommended by Dorf.\textsuperscript{44} At every sample interval, data from sensors are sent through the network to the remote platform, which in turn sends back the processed control information containing the target velocities at that moment. Therefore, the round-trip time for network/cloud processing time should ideally be below our sampling window. Control information delayed beyond this sampling window directly impacts not only the global control loop but also the user’s cognitive loop. This leads to control degradation—and even complete loss of controllability in extreme cases—and may mislead the human–CPS interaction at the risk of taking the whole system to an unstable state.

Therefore, to guarantee user safety and comfort while using CloudWalker, latency should be kept as low as possible. For instance, testing exchange of 70 B packets from our location to a major cloud provider on both US’ coasts, an average latency of about 700 ms was observed, decreasing to 50 ms with their edge-like service. Besides latency, jitter also influences performance by inserting information disorder in the feedback loop. Short-term events, such as congestion and errors over wireless links, lead to individual packet loss and then throughput reduction. There are also long-term failures that are related to availability metrics that lead to massive successive packet loss events. In practice, however, actual patient’s dynamics of use might tolerate such real-world cloud performance. In other words, QoS perspective seen by CloudWalker, as depicted in Fig. 4(a), might not directly reflect QoE seen by the human user, as the relationship between QoS requirements, the human-in-the-loop effect, and the perceived QoE is yet to be fully understood.

5. Pilot Experiment

To assess how network/cloud parameters affect CloudWalker, we designed a pilot experiment to investigate the user’s perceived QoE under variable QoS. Therefore, we tested the system under different QoS conditions to subjectively evaluate CloudWalker’s performance through the end-user point of view. A total of 33 individuals participated in our experiments and provided feedback about their perceptions while using the system.

Our pilot experiment is comprised of two phases. In phase 1, we designed an exploratory approach aiming at growing an understanding of how users react to multiple QoS conditions (i.e., E2E latency, network availability, and packet loss) and how the perceived QoE is evaluated. Based on our findings, we moved on to the second phase of the pilot experiment, in which we narrowed down the number of QoS conditions and increased the number of subjects participating in the study on an attempt to gather statistically significant data linking such QoS conditions to user-perceived QoE, feeling of safety, and feeling of control upon the system.
This work proposes a discussion on QoS and QoE on a cloud-enabled CPS, displaying a use case that, despite its deliberate simplicity, is strong enough to illustrate a mobility assistance service without diverting from the focus of the study. The selection of controllers and the fact that there is no computation performed on the device itself is intended to push the system to a QoS-sensitive configuration. The following subsection details the implementation of the healthcare service explained in Section 3, which is valid for both phases of the pilot experiment. The rest of this section describes each phase individually, adopted protocols, and the obtained results. Whenever not stated otherwise, the same methods were applied in both phases of the pilot experiment.

5.1. HaaS: Virtual trails for mobility assistance
We implemented an interaction-based guiding service exploiting communication and computation features highlighted in Fig. 2(b). Such service leads users on healthcare facilities through predetermined virtual trails. This service gives users freedom to walk at own pace and full control over the smart walker’s speed. At the same time, the controller commands curving speed to keep the user on trail. This works on a proportional control strategy: the slower the person walks with the device, the slower the robot changes its orientation. The effect is of a virtual trail, where users can “push” the system along the virtual trail while preventing them deviating from it.

Figure 5 provides a block diagram illustrating the three different control functionalities that compose such mobility assistance service: an odometry algorithm, a path following controller, and an admittance-based controller. Figure 5 also illustrates the control inputs (i.e., the interaction forces as measured by the force sensors) and how the network constraints are inserted in the control loop in our implementation.

The odometry uses embedded sensor signals to estimate walker’s position and relative orientation. This information feeds the path following algorithm, which is based on Monllor et al.,45 to determine the desired orientation of the walker to keep it on trail. This desired orientation is compared with the current orientation and an orientation error is given as the admittance controller input. The admittance-based controller is based on Monllor et al.45 and is given by the following equations:

\[
v_c = \frac{F_{\text{forward}} - m_l \dot{v}}{d_l} \tag{1}
\]

\[
\omega_c = \frac{v_c}{d_\omega} \tanh \tilde{\theta} \tag{2}
\]

where \(v_c\) and \(\omega_c\) are the reference linear and angular velocities, respectively, \(F_{\text{forward}}\) is the total forward interaction force measured by the force sensors, \(\dot{v}\) is the derivative of the current linear velocity, and \(\tilde{\theta}\) is the orientation error. \(m_l\), \(d_\omega\), and \(d_l\) are control parameters empirically adjusted.

If no up-to-date control information is received at a given sample interval, the last valid control signal is repeated. As stability issues may arise not only from the inherent characteristics of a human-in-the-loop system but also from communication constraints, we limited forward linear velocity and angular velocity to 0.25 m/s and 0.5 rad/s. Moreover, negative values for linear velocity control signal are set to zero, inhibiting backwards movements. This is done to restrict intense responses on the physical interaction with the human, safeguarding the subjects of our experiment.
In this pilot experiment, the same path was used for all tests with our virtual trail, consisting of straight lines interspersed by smooth curves to both sides (see Fig. 6). Such navigation path emulates common indoor navigation scenarios, as displacement through corridors and halls, which are likely to happen in hospitals and nursing homes.

5.2. Pilot experiment—Phase 1
We designed the first phase of our pilot experiment as an exploratory study toward growing an understanding of the interplay between QoS and QoE in human-in-the-loop CPS. A group of volunteers tested our system under different QoS conditions and provided feedback regarding their experience. We now detail the protocols adopted in phase one and report the observed results.

5.2.1. Phase 1: Experiment protocol. We recruited 12 participants to take part in phase 1 (four women). All subjects can be considered healthy, with no disabilities or gait disorders, and only two of them had prior experience using the smart walker. Participants’ ages ranged from 22 to 33 years old, their weights from 57 to 92 kg, and their heights from 1.61 to 1.85 m. In the beginning of each individual session, participants signed a consent form agreeing to be part of this pilot experiment. There was no compensation involved.

The experiment took place in a 60 m² empty room in the time span of 2 days. A video camera on a tripod recorded the whole activity. The research team, composed of two researchers, informed the participants, in the beginning of their sessions, about the overall objectives of the experiment. No further instructions on how to interact with the device, path to follow, or details of the test conditions were provided. Participants were also encouraged to speak freely about their interaction with the CPS.

All tests were performed over the same predetermined virtual trail, which considered sufficient free space to keep the participants and the smart walker on a safe distance from the walls. Each participant performed five tests under different network/cloud conditions, aware that some conditions could perform better than others.

We specified a total of 10 test conditions, from which 5 were randomly chosen for each participant. Tests ordering was also randomized to avoid parameter-induced bias. We estimated participation time on 20 min to not cause any tiredness to the volunteers. As the user had to figure out by themselves how to interact with CloudWalker, two participants had to be fully removed from our data set as they intentionally pushed CloudWalker beyond the limits of wheels adherence or velocity controllability.

5.2.2. Phase 1: QoS and network parameters. To have full control over network constraints, all processing was performed at the embedded device’s hardware. We have emulated the remote cloud processing by inserting network constraints blocks inside the control loop (see Fig. 5). Three network parameters were considered central to the QoS discussion: E2E latency, packet loss ratio, and network availability (i.e., percentage of uptime over total time).

For each test, only one parameter was changed, whereas the other ones were considered ideal. Moreover, when dealing with variable E2E latency, a normal distribution was applied and jitter was chosen based on a fixed variance mean ratio (VMR) parameter. The 10 test conditions are listed below accompanied by the number of valid participants randomly selected performing each of condition in parenthesis:

- Benchmark: ideal condition (7);
- Deterministic Latency: 10 ms (4);
- Variable Latency (VMR 0.01 ms): 100 ms (2), 300 ms (3), and 500 ms (4);
- Availability: 75 % (4) and 50 % (3);
- Packet Loss: 10 % (2), 20 % (2), and 30 % (6);

5.2.3. Phase 1: QoE metrics. We used the mean opinion score (MOS) to evaluate user-perceived QoE. The MOS is a widespread QoE metric that considers the average of scores given by each user regarding their experience using the system under a specific QoS condition. The concept of QoE attempts to combine user perception, experience, and expectation, and the MOS provides a simple evaluation method, in which the user ranks their experience based on a given scale. Despite being one of the most popular metrics to measure QoE in applications such as media streaming,
On human-in-the-loop CPS in healthcare

Fig. 6. Snapshots from the recorded video illustrating the virtual trail and following.

Fig. 7. Observations from the phase one of the pilot experiment: (a) path traveled by same participant under different test conditions; (b) forward interaction forces and subsequent forward velocities observed, first 10 s of same tests as the ones in the leftmost figure; (c) special case of path deviation, and impacts of connectivity losses larger than 2 s.

During phase 1, at the end of each test under a QoS condition, participants were asked to describe their experience of using the CPS. They had to answer a one-question questionnaire: how do you rate your experience in a scale of 1–5, 1 meaning a negative and 5 a positive experience? We also registered observations and reviewed the recorded material to gather participants’ comments. Moreover, embedded sensor’s data such as forward interaction forces, speed, and trail deviations were also recorded to study its impacts on QoE.

5.2.4. Phase 1: Results and discussion. Figure 6 illustrates four snapshots of a realization for the virtual trail mobility assistance service. Although participants had no previous training, all of them figured out how to interact with the walker and managed to successfully follow the unknown predetermined virtual trail.

Figures 7(a) and 7(b) presents the outcomes for predetermined and actually performed paths, interaction forces, and linear velocities from two realizations from the same participant using the CPS under different E2E latency conditions: 10 ms (deterministic) and 500 ms (VRM 0.01 ms). The score
for each realization is also shown so that the complex relations between physical interaction effects, the human-in-the-loop factor, and subjective evaluation could be clearly illustrated. One can observe in Fig. 7(a) that there are surprisingly small deviations from the virtual trail even under 500 ms latency. In fact, the largest deviations occurred under the smallest latency condition, and data from all tests indicate that this is mainly due to user’s gait speed, which tends to increase according to user’s positive perception of the experience.

As it might be expected, the participant better graded the test under smaller E2E latency, and Fig. 7(b) may provide evidences to understand the reason: 500 ms latency degraded the participant’s ability to control the CPS. Cognitive and physical interactions on such human-in-the-loop system are impaired because of temporal mismatches between perception and action. This leads to intense oscillations in the interaction forces and subsequent speed variation in Fig. 7(b).

Figure 7(c) brings two realizations from different participants to illustrate the complexity of assessing QoS effects on user and system overall performance. Although under the same 500 ms (VMR 0.01) latency used in Fig. 7(a), in Fig. 7(c) it caused the CPS to miss out the trail. This leads to safety concerns and may draw limits for QoS degradation. In contrast, Fig. 7(c) also presents a realization under poor network/cloud QoS (only 63.71% availability) and yet successfully performed. However, it is worth noting that no period longer than 2 s of connection loss was observed along the curves of the trail.

MOS results are displayed in Fig. 8(a). We notice that unfavorable conditions clearly deteriorate user’s experience. Among parameters evaluated, E2E latency appears to be of most significance. Network availability also seems to degrade QoE, though the implemented control application showed itself quite resilient to loss of connectivity. Packet loss ratio over the network also seems to affect experience, but given the slow system dynamics and the relatively high sampling rate, no significant effects could be observed.

Statistical limitations from this reduced pool of realizations may be the reason for QoE outliers in Fig. 8(a). Only two tests were performed for 20% packet loss in contrast with six performed under 30%. However, the human-in-the-loop factor also impacts the experiment itself. For example, four participants experienced 500 ms E2E latency, but only one (depicted in Fig. 7(b)) evaluated its QoE in the lowest score, whereas remainder participants had a more pleasant experience without large oscillations and all scored 4.

Service provider’s interest in QoE metrics motivates discussions on the relationship between general quality and service acceptance. Measures such as “good or better” (%GoB) and “poor or worse” (%PoW), which were primarily applied on telecommunications, have been used by providers to estimate percentage of users satisfied (or dissatisfied) in a quantile-based approach, as defined by ETSI’s E-model.39 As a means of statistically extrapolating data from experiments with a limited number of participants, Fig. 8(b) brings results for %GoB and %PoW, and the ratio of users evaluating each test condition from pilot experiment above or below certain MOS thresholds. The task-related thresholds, (i.e., %GoB ≥ 4 and %PoW < 3) were found to be the best fit to the E-model. Given targets for %GoB and/or %PoW for the HAaaS, QoS requirements could be estimated from those results. For instance, by assuming satisfaction level (%GoB) of 80% in Fig. 8(b) would correspond to a 3.8 MOS, and MOS above 3.6 in order to maintain dissatisfaction level (%PoW) below 5%. Therefore meeting both criteria would require a MOS target above 3.8. Therefore, Fig. 8(a) indicates corresponding QoS conditions that meet such MOS target (i.e., E2E latency below 10 ms, availability above 75%, and packet loss below 10%).

5.3. Pilot experiment—Phase 2

During the first phase of the pilot experiment, the volunteers offered valuable feedback regarding their perception of the effects imposed by different QoS conditions into the system. Nevertheless, as an exploratory study, a number of limitations were observed and the MOS results obtained contrasted with some of the comments made by the volunteers. We designed phase 2 to overcome some of those limitations aiming at reaching more solid data linking user-perceived QoE and variable QoS.

In phase 2, we chose to narrow down the number of test conditions and extend the number of volunteers. Each volunteer tested the system under all test conditions and, after each test, they were asked to answer an extended questionnaire designed to extract a more complete set of information regarding the system. We now describe the protocols adopted in phase 2 and report the observed results.
5.3.1. Phase 2: Experiment protocol. We refined the first phase of the pilot experiment and specified that the only independent variable used in phase 2 would be the E2E latency. As dependent variables, on which we wanted to measure the effect of the latency, we considered the users’ QoE, the perception of control, and the perception of safety when using the CloudWalker. In phase 2, we decided to use the following latency conditions: 0, 100, 300, and 500 ms.

We recruited 21 new participants to take part in phase 2 (five women). All subjects are associated with our university and presented no disabilities. Only five participants had prior experience using the smart walker. Participants’ ages ranged from 24 to 56 years old, their weights from 62 to 91 kg, and their heights from 1.63 to 1.88 m. The volunteers received no compensation and we obtained the consent of all of them to collaborate with the study.

Two researchers conducted the experiment in our university over 2 weeks. Each session lasted, on average, about 30 min and consisted of a briefing session, testing using the CloudWalker, and completion of post-experiment questionnaires after each test. We delineated that after using the system under a given latency condition, each participant would answer 5-point Likert scale questions for the dependent variables. The three questions composing the questionnaire are listed:

1. How do you rate the experience of using the smart walker?
2. How do you rate the feeling of control upon the smart walker?
3. How do you rate the feeling of safety when using the smart walker?

Due to the number of participants, we chose a within-group design, in which each participant was exposed to all test conditions (i.e., use the CloudWalker under all the latency conditions).
This approach allows us to effectively isolate individual differences and reduce the noise level among participants; however, it is hard to control the learning effect and impact of fatigue. In order to minimize the learning effect, we randomly determined the order of the test conditions using a “Latin Square Design” and to mitigate the effect of fatigue, we kept the short path used during the first phase of the pilot experiment.

5.3.2. Phase 2: Results and statistical analysis. Before starting the statistical analysis, the data collected through the questionnaires were carefully processed, coded, and consolidated in the CSV (Comma-separated values) format. From there, we could test statistical hypotheses regarding the impact of the latency on the dependent variables (QoE, perception of control and perception of safety). In this process, we extensively used some specialized statistical packages, such as RStudio v1.1.3831 and G*Power2 v3.1.9.3.

In the statistical analysis of the data, we initially considered the repeated measures analysis of variance (ANOVA). The ANOVA is used to compare if there are significant differences among means of a dependent variable as a function of a single independent variable (called a factor) with 2 or more categories or levels. As previously stated, the only independent variable is the latency of the network.

Given that our questionnaires consisted of Likert scale questions, in which the distance between two adjacent points can be unequal, we used the Friedman test, considered the non-parametric equivalent of the repeated measures ANOVA to analyze our data. For the Friedman test, the prerequisites can be considered more relaxed. For example, a group or category needs to be measured three or more times, the measurements need to be independent of each other, the dependent variables must be measured at the continuous or ordinal level, and the samples do not need be normally distributed. It is noteworthy that our data set satisfied all these requirements.

We used G*Power to investigate whether we could detect statistically significant differences with our sample size. We used the ANOVA with repeated measures and specified the effect size $0.253$ (mean effect), $p = 0.05$ (significance level $\alpha$ or Type I error), power = $0.95$, number of groups equal to 1 (analysis within the group) and number of measures per participant equal to 4 (number of latency conditions). We found that a total of 35 participants were required to obtain a statistically significant result with the parameters provided. In our case ($N = 21$), there was an approximately $77\%$ chance of correctly rejecting the Null Hypothesis $H_0$, which can be stated as:

$H_0$: There were no statistically significant differences between the means of users reported QoE when using the CloudWalker under different E2E latency conditions ($\mu_{QoE_{\_latency\_0\_ms}} = \mu_{QoE_{\_latency\_100\_ms}} = \mu_{QoE_{\_latency\_300\_ms}} = \mu_{QoE_{\_latency\_500\_ms}}$).

In the evaluation of the Null Hypothesis $H_0$, the independent variable included four categorical, independent groups: 0 ms ($\mu_{QoE_{\_latency\_0\_ms}} = 4.29, \sigma = 0.56, \text{Mdn} = 4, n = 21$), 100 ms ($\mu_{QoE_{\_latency\_100\_ms}} = 3.86, \sigma = 0.96, \text{Mdn} = 4, n = 21$), 300 ms ($\mu_{QoE_{\_latency\_300\_ms}} = 4.05, \sigma = 0.67, \text{Mdn} = 4, n = 21$), and 500 ms ($\mu_{QoE_{\_latency\_500\_ms}} = 3.67, \sigma = 0.73, \text{Mdn} = 4, n = 21$). Figure 9(a) summarizes the measurements and the observations. We performed the Friedman test in RStudio and found a statistically significant result, $\chi^2(3) = 12.45$ and $p = 0.0060$.

We then performed a post-hoc analysis using the Wilcoxon signed-rank tests with Bonferroni adjustment, resulting in an adjusted significance level of $p < 0.008$. We found that there was no statistically significant difference between the latency values of 0 and 100 ms ($Z = 1.59$ and $p = 0.14$), of 0 and 300 ms ($Z = 1.64$ and $p = 0.16$), of 100 and 300 ms ($Z = -0.57$ and $p = 0.56$), of 100 and 500 ms ($Z = 1.31$ and $p = 0.18$) and of 300 and 500 ms ($Z = 2.21$ and $p = 0.03$). However, we did find a statistically significant difference between the latency values of 0 and 500 ms ($Z = 3.22$ and $p = 0.001$). Therefore, in this case, we reject the Null Hypothesis $H_0$ ($\mu_{QoE_{\_latency\_0\_ms}} \neq \mu_{QoE_{\_latency\_500\_ms}}$) and state there are differences in the mean value of user’s reported QoE under those two test conditions. Such result points to QoE degradation under the worst QoS condition.

We followed the same procedure described above to analyze the effect of the latency factor on the other dependent variables such as perception of control and perception of safety when using the cloud walker (see Figs. 9(b) and 9(c)), respectively. In our analysis, we did not find a statistically significant result for none of these variables (perception of control: $\chi^2(3) = 2.20, p = 0.53$ and perception of safety: $\chi^2(3) = 1.53, p = 0.68 > 0.05$ significance level $\alpha$). In other words, this means that we cannot reject any of these Null Hypotheses based on our sample of participants and...
6. Lessons Learned and Future Directions

Most users followed the assigned virtual trail correctly with no prior information and under QoS conditions far worse than the ones envisioned by 5G networks. Our outcomes unveil trends for relating QoS indexes and human-in-the-loop physical interactions with control loops affecting QoE degradation. Thus, this work should evolve to derive a new QoE model for healthcare in order to estimate distribution of scores for different QoS conditions.

E2E latency is the major cause of lower QoE and also raised safety concerns as controllability might degrade to unstable conditions. There are need for studies with larger VMR for more realistic jitter assessment. In that case, we recommend long-lasting realizations and extra instructions to participants due to potential safety concerns. Prior experience over benchmark conditions can also be tried for a comparative QoE evaluation.

In phase 1, we also noticed that the analysis of MOS scores alone did not reflect entirely participants’ experiences. For instance, one said “this test is better than the previous one” and then gave the same score in both tests. In another session, a test in which the participant said “the walker seems to be failing” received a 3 and it was not the lowest score given by that particular user. Comments like “this was the best one,” “smooth,” and “comfortable” were usually linked to highest scores, but negative comments were not necessarily followed by the lowest score.

Positive comments could also be linked to higher experience scores in phase 2. We also observed that user’s comments in fact reflected their scores regarding the sensation of safety and of being in
control. Most users stated that the different QoS conditions did not impact on such perceptions and thus graded most of the test conditions similarly.

In phase 2, we identified a statistically significant difference between the means of reported QoE for different latency groups. As expected, among the latency conditions we used, the 500 ms was considered the worst ($\mu_{QoE_{\text{latency 500 ms}}} = 3.67$). However, it is important to observe that as the latency of the network increases, the perception of control and the perception of safety do not seem to degrade as fast as the QoE, given that we were quite far from identifying a statistically significant difference for these two dependent variables.

Although phase 2 results indicate interesting findings, it is important to note that despite the effect of the independent variable (i.e., latency) on the dependent variables, it does not mean that we can generalize the conclusions to any latency condition and any population. This is because we fixed beforehand what would be the latency values used in the experiment. The study also had some limitations. The group of participants can be considered homogeneous in terms of age, level of education, and physical condition. Therefore, we should keep in mind that our results should not be immediately generalized and the impact of other factors should be further investigated in future work.

Other assistive devices could be conceived as cloud-enabled CPS, each of them presenting particular characteristics and needs. Our study was performed over a guiding feature that can be employed by devices such as smart wheelchairs and canes, which are also affected by the human-in-the-loop effect. Efforts are needed in the direction of finding more suitable ways for the evaluation of cloud-enabled CPS healthcare applications and also in establishing trustworthy means for relating QoS to QoE. Future studies should address some of the limitations found in our pilot experiment and HaaS providers will need to establish clear relationships between QoS and end-user acceptability.

User acceptance could be strongly affected whether the effects of QoS over QoE are not properly understood and mitigated in any human-in-the-loop CPS that rely on cloud-based services. As different service configurations are developed, each of them delegating part or even the totality of computational tasks to the cloud, more than addressing safety risks, there is a need to holistically evaluate the system including the end-user’s point of view. The results from our pilot study corroborate with this view, also pointing that user opinion alone may not reflect the actual experience. Despite our efforts to understand how safety and the feeling of control can be affected by QoS, there is still a need for deeper investigation in the direction of such a holistic evaluation.

We envision a combination of physician–patient–therapist assessment not only for QoE but also including “acceptability” metrics. In future works, we should also include alternative tools to measure long-term as well as short-term QoE measurements. A better QoE assessment may be directly obtained from physiological data. For this end, galvanic skin response (GSR) sensors are highly accurate for indicating the user internal state.

It is especially difficult to compare our results to prior works regarding smart walkers or even other assistive devices due to the lack of standardization in methodology. Works such as Wachaja et al. and Werner et al. presented objective and subjective metrics to evaluate guidance features in smart walkers, locally executing the control algorithms and focusing on the end-user point of view. Upcoming studies involving human-in-the-loop CPS could benefit from our results as stepping stones for new implementations.

7. Conclusions
This article investigated the potential of emerging communication/computation technologies to unleash a new generation of healthcare assistive devices. To motivate discussions about human-in-the-loop effect over a concrete case, we envisaged and implemented a pilot experiment for a healthcare service based on a virtual trail mobility assistance through CloudWalker, a system linking smart walkers to cloud platforms. Upon this illustrative HaaS use case, challenges for migrating the control of a smart walker to cloud computing platforms could be discussed regarding QoS requirements, the human-in-the-loop effect, and the perceived QoE.

Although the CPS system proved to be resilient to poor QoS, in our pilot test latency could be linked with lower QoE. There are also clear safety concerns to be addressed in the future, especially regarding jitter. Moreover, QoE is yet to be better understood and assessed in CPS with human-in-the-loop effects. Closer involvement of health professionals is necessary for composing trustworthy QoE and acceptability metrics. Only then consistent QoS requirements for network/cloud can be derived for HaaS by extensive and comprehensive clinical tests.
Acknowledgements
The research leading to these results received funding from the European Commission H2020 program under grant agreement no. 688941 (FUTE Bol), as well as from the Brazilian Ministry of Science, Technology, Innovation, and Communication (MCTIC) through RNP and CTIC. This work was partially supported by FAPES (80709036 & 72982608), and CNPq (304192/2016-3). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Finance Code 001.

References


