On Cellular Networks Supporting Healthcare Remote Monitoring in IoT Scenarios

Andrea Petroni[1]*, Pierpaolo Salvo[2] and Francesca Cuomo[1]

[1]Department of Information, Electrical and Telecommunication Engineering (DIET), "Sapienza" University of Rome, Rome, Italy,
[2]Fondazione Ugo Bordoni (FUB), Rome, Italy

In the next few years, fundamental technological transitions are expected both for wireless communications, soon resulting in the 5G era, and for the kind of pervasiveness that will be achieved thanks to the Internet of Things. The implementation of such new communication paradigms is expected to significantly revolutionize people's lives, industry, commerce, and many daily activities. Healthcare applications are considered to be one of the most impacted industries. Sadly, in relation to the COVID-19 pandemic currently afflicting our society, health remote monitoring has become a fundamental and urgent application. The overcrowding of hospitals and medical facilities due to COVID-19, has unavoidably created delays and key issues in providing adequate medical assistance. In several cases, asymptomatic or light symptomatic COVID-19 patients have to be continuously monitored to prevent emergencies, and such an activity does not necessarily require hospitalization. Considering this research direction, this paper investigates the potentiality of cloud-based cellular networks to support remote healthcare monitoring applications implemented in accordance with the IoT paradigm, combined with future cellular systems. The idea is to conveniently replace the physical interaction between patients and doctors with a reliable virtual one, so that hospital services can be reserved for emergencies. Specifically, we investigate the feasibility and effectiveness of remote healthcare monitoring by evaluating its impact on the network performance. Furthermore, we discuss the potentiality of medical data compression and how it can be exploited to reduce the traffic load.

Keywords: remote healthcare, cloud services, LTE, data compression, IoT

1 INTRODUCTION

Over the last few years, we have witnessed how the advent of Internet of Things (IoT) has led the world into a new communications Era. The recent breakthrough in technology, both in terms of hardware and software, has allowed human life to be pervasively supported and influenced by machines and devices. In this context, Machine to Machine (M2M) Communications have emerged as a typical IoT paradigm where the data exchange among devices takes place automatically over both wired and wireless channels (Al-Fuqaha et al., 2015). Having billions of entities simultaneously connected has therefore led to the redefinition of the concept of global networking. Moreover, the coexistence of several activities and applications relying on different traffic patterns has to be properly supported.
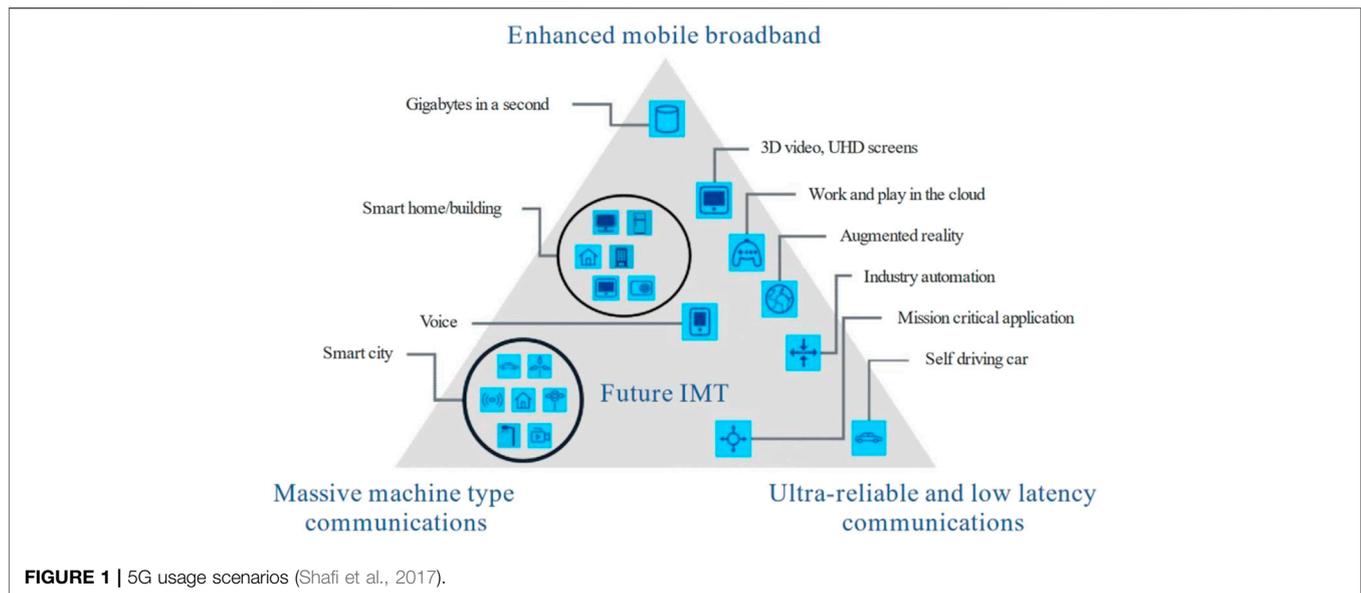
FIGURE 1 | 5G usage scenarios (Shafi et al., 2017).

Such a task represents one of the main challenges of the upcoming 5G paradigm, where the enhancements provided in terms of quality of service (QoS) pass through the smart management of the available network resources.

In this paper we investigate the effectiveness of the use of cellular technologies combined with IoT for a specific application scenario, i.e., remote patient assistance. We foresee the use of these technologies as being one of the components of an efficient response to the COVID-19 pandemic. We contribute under three perspectives: 1) by providing a discussion on how this remote healthcare can be framed in the 5G infrastructures; 2) by indicating the benefits of data compression in cloud-based services designed to collect and analyse IoT data measured at the patient's side; 3) by assessing the behaviour in the current cellular setting, based on LTE. As for this latter point, since 5G infrastructures have not been fully deployed yet, the existing 4G network framework (5G in fact partially relies on 4G) can be fruitfully exploited to support healthcare activities. To this aim the testing has been done in that context, but guidelines can be easily extended to 5G.

1.1 5G and IoT Infrastructures in View of Health Monitoring

It is well known that the IMT 2020 (International Mobile Telecommunications) organized envisioned 5G usage scenarios in a pyramidal structure that summarizes the current trends in IoT applications (**Figure 1**). Specifically, three main categories are placed at the pyramid edges (Shafi et al., 2017): Ultra-reliable and low latency communications (URLLCs), Enhanced mobile broadband (eMBB) and Massive machine type communications (mMTC).

URLLCs, introduced by the 3GPP (Third Generation Partnership Project) Release 15 standards, are oriented towards mission-critical scenarios with stringent constraints in terms of scalability, latency and reliability. Similar requirements

characterize factory, process and power system automation in industry, autonomous driving, road safety and traffic management in intelligent transportation, remote surgery in healthcare, and many other fields of applications (Bennis et al., 2018). High data rate connections over wide areas are supported by eMBB services, developed from the 4G broadband framework and destined to significantly outperform LTE (Long Term Evolution) performance. The enhanced connectivity is beneficial for faster network access and better user experience when dealing with high-quality video streaming, file share/ transfer and device remote control (Wan et al., 2018). Finally, mMTC considers the presence of a very large number of devices sporadically exchanging an extremely variable, but limited, volume of data without any constraint to the delay. Such a scenario is exactly the one related to the IoT, affecting both people daily lives (Palattella et al., 2016) and industry (Bandyopadhyay and Sen, 2011). The features of mMTC are perfectly tailored to the nature of IoT traffic, with benefits also in terms of energy savings for battery supplied sensors. Wearable devices, smartphones, and other portable instruments used in Personal Area Networks (PANs) are typically supported by technologies like Zigbee, Bluetooth Low Energy, and Wi-Fi. On the other hand, the so-called Low Power Wide Area Networks (LPWANs) have emerged as potential technologies that are able to provide coverage in the order of kilometers for IoT devices. In this category we can find Sigfox and LoRaWAN, developed to operate in the ISM (Industrial, Scientific, Medical) unlicensed electromagnetic spectrum band (Centenaro et al., 2016). Alternatively, LTE-M and, more recently, Narrowband IoT (NB-IoT) were defined as cellular based technologies built within the LTE framework (Foubert and Mitton, 2020).

Overall, the services categorization in 5G has been realized by jointly considering the requested QoS, the traffic volume that needs to be supported, and the characteristics of the end devices. However, it is worth noting that some applications span over multiple, heterogeneous, usage scenarios, therefore they cannot

be identified in a single category of 5G services. An example if this is smart healthcare, that, due to its current relevance, has emerged as one the most challenging fields where cellular technology can find a use (Ullah et al., 2019). Patient tracking and monitoring, emergency alarm systems, and remote surgery represent some of the potential use cases, each one characterized by specific performance and constraints. Network architectures for e-Health have already been proposed for private use in hospitals and medical facilities, while remote monitoring has been implemented through dedicated applications. Some issues, however, are still under investigation. The monitoring of hospitalized patients with data stored in a private cloud server improves the information management, but the facilities still have to face the costs of treatment and the potential reduction of available inpatients beds.

Currently, this problem has been highlighted with the sudden spread of COVID-19. The high number of patients requiring assistance has caused the collapse of health systems in many countries because of the lack of hospitalization resources. Therefore, the prioritization of hospital services has become crucial, with patient assistance being necessarily implemented into new settings wherever possible.

In this regard, for asymptomatic patients and those with light symptoms, the use of IoT devices makes their monitoring, outside of the hospital setting, feasible as well. Furthermore, IoT sensors can be programmed and can therefore work automatically without any intervention from the user (this is an advantage especially when dealing with elderly people who may not be familiar with technology). The only drawback of this solution is that IoT supports limited and sporadic traffic, therefore only spot measures of vital parameters can be transmitted efficiently. When the patient monitoring becomes more complex which requires a large amount of data to be transmitted, a different approach must be followed. By referring to the 5G pyramid in **Figure 1**, such a particular scenario can be positioned closer to eMBB than to mMTC, since the pattern of traffic to be supported is quite similar to that envisioned for eMBB, while the number of entities simultaneously connected may be large but not comparable with numbers expected in the IoT.

Furthermore, some pre-optimization of the network traffic can be performed at the application layer where raw data coming from devices are handled. In fact, especially when dealing with cloud-based architectures, the information generated by the end devices represents an updated version of data already stored in the cloud. Therefore, the information to be transmitted may be conveniently reduced only by dealing with the part necessary for the cloud updating. Such a procedure is in general referred to as *data synchronization* and relies on particular data compression algorithms.

1.2 Contributions

In the framework of health monitoring through cellular systems and IoT, we identify a cloud-based cellular network scenario where we implement monitoring applications for people suffering from respiratory disease (and others). We focused on such kind of applications, considering the efforts required currently to remotely manage COVID-19 patients from their home. While

the very latest literature about IoT and cellular technologies has dealt with contact tracing implementation and screening applications (Alsaeedy and Chong, 2020; Chamola et al., 2020), few papers have concentrated on the use of these technologies for monitoring patients at home as was required in many countries during COVID-19 lockdowns. Challenges characterizing the considered use case can be listed as follows:

- The potential presence of a large number of patients concentrated in the same restricted geographical area (a house, a building, a city district) gives rise to scalability issues;
- Many heterogeneous health sensing devices may be used for the same patient, so this implies that the kind of data that must be exchanged may vary in size and QoS requirements;
- Given the specificity of the diseases provided by COVID-19 (respiratory issues) it is fundamental to have punctual and reliable data acquisition.

We then based our study on real medical measurements, exploited to evaluate the network performance considering particular scenarios involving a potential high number of users to be served in a limited area. Furthermore, we discuss the feasibility of data compression on medical information, also evaluating the impact on network traffic load. To the best of our knowledge, such a type of analysis has never been proposed before. In fact, many works dealing with cellular technologies for healthcare applications investigate only those issues related to networking, without any specific consideration of the pattern of data that needs to be handled. On the other hand, algorithms and mechanisms for data compression are typically evaluated in terms of compression ratio and loss, but the impact of data reduction on the network traffic is not discussed at all. In contrast, our contribution addresses both data compression and traffic optimization issues jointly. As stated before, for the performance assessment we used LTE, to have it ready as a practical case, since this is technology has already been assessed. However, the study proposed here paves the way to the 5G evolution since the architectural model as well as the use of cloud computing fit perfectly to that framework.

The rest of the paper is organized as follows. **Section 2** reports the literature review about the use of LTE technology in the healthcare context. In **Section 3** the state-of-the-art of medical data compression is described. The proposed network architecture for remote healthcare monitoring is introduced in **Section 4**. A simulation analysis is performed and discussed in **Section 5**. Finally, **Section 6** draws conclusions.

2 LTE-BASED SOLUTIONS FOR HEALTHCARE

The use of cellular technologies for healthcare applications has been introduced to address two important issues. The first one concerns data accessibility and management, with the aim to improve the quality of interactions between patients and medical personnel. The second one instead regards healthcare services

organization, in terms of both scheduling and prioritizing, so as to minimize latency periods for patients. Therefore, many 4G/5G based healthcare solutions have been developed to support activities in hospitals and other medical facilities where the presence of a large number of individuals is expected. In this direction, network architectures and platforms have been tailored to specific application scenarios (Islam et al., 2015). An example is given in Zhang et al. (2018) where a real-time drug infusion monitoring system based on infrared sensors and NB-IoT is described. The large coverage provided by NB-IoT makes it efficient in smart buildings/areas. Sensors measuring heart rate, blood pressure, and body temperature are typically provided by NB-IoT technology for data transmission, as detailed in Malik et al. (2018). In this regard, it is worth highlighting that NB-IoT belongs to the LPWANs category, so it is perfectly suited to an IoT context, such as the one presented, where small size data are sporadically transmitted. Hence, such technology is not properly tailored to support a close monitoring of activity where a large amount of data is transmitted. Furthermore, since no traffic analysis is provided in Malik et al. (2018), Zhang et al. (2018), it is not possible to evaluate the impact of the proposed solutions on the overall network load, especially when a large number of users is simultaneously connected.

Another potential application scenario for cellular technologies is represented by emergency and alarm systems. In the healthcare context, giving aid to patients out of hospitals as soon as possible is crucial. In this direction, the authors in Cical et al. (2016) propose a framework for supporting ambulance services with remote assistance through video streaming. In that case, LTE Advanced (LTE-A) is considered for data transmission, since real-time video and image delivery requests large bandwidth systems to guarantee high quality and small latency (on the other hand, the IoT paradigm is not suited to provide such performance). More recently, the same use case has been investigated in Rehman et al. (2018) where medical video streaming is proposed in the context of future 5G small cell heterogeneous networks. The paradigm of mobile health is recognized as useful for disaster management as well. In fact, the work in Adibi (2018) shows how affected people can be conveniently supported by means of a vital parameters remote monitoring service. However, such emergency scenarios are quite particular and are rarely expected to occur. Therefore, the impact of these services on the network traffic load is negligible with respect to a continuous remote monitoring application.

In addition to patients support in hospitals and emergency events, which can be conveniently handled by means of cellular networks, the largest part of remote healthcare solutions concerns the monitoring and tracking of individuals outside of medical facilities. In fact, patients are typically provided with body sensors measuring vital parameters that have to periodically be collected and sent to a medical data store or cloud. Such devices can be simple wearables, but also include body implants, so they are designed to be energy efficient and have a guarantee of between 5 and 10 years. Data transmission is performed in a wireless fashion but due to the problem of energy consumption, the device

transmission power is only sufficient enough to cover an area of a few meters. A viable solution to allow long range communication from sensors to the cloud considers the presence of an intermediate node acting as data collector. The work in Adibi (2014) presents an end-to-end network framework for mobile healthcare, referred to as a biomedical sensing analyzer, where data transmission from the medical device to the cloud is performed using a smartphone as a sort of gateway. Specifically, the measurements collected by the sensor are first sent to the smartphone using a short-range communication oriented to PANs. The smartphone is then responsible for data forwarding to the cloud exploiting the LTE-A network. The same approach is exploited in Hindia et al. (2016) where a smartphone application controls and aggregates the measurements coming from several sensors. In that case, the communication between the smartphone and cloud is realized by exploiting LTE-Femtocell networks. Furthermore, a particular scheduling strategy is proposed in order to optimize the data traffic based on the priority of the information that needs to be transmitted. The use of LTE-Femtocells (the dimension of which is in the order of few tens of meters) is envisaged in a IoT context, providing better coverage and more efficient traffic management. For those scenarios where sensors generate limited-size data and where their transmission occurs infrequently, the IoT paradigm is demonstrated to be an effective solution. An example of healthcare services relying on NB-IoT is described in Manatarinat et al. (2019). In this context, one of the most important aspects to consider is that M2M communications are not only efficient from a networking point of view, but they are also user-friendly since the standalone behavior of IoT devices makes their management very easy.

The literature review about the healthcare network has revealed how the proposed solutions are tailored to specific application scenarios and, in particular, to the data traffic that will be handled. However, further optimization can be achieved by resorting to machine learning. An example is given in Hadi et al. (2019), where an algorithm for big data analysis is used to process patient aggregate data coming from biomedical sensors to recognize potentially dangerous events such as a stroke. By doing so, the LTE-A uplink transmissions are scheduled based on the service priority. Concerning healthcare data mining, the authors in Jiang et al. (2021) show how the promising deep learning can be exploited to improve the interaction and data exchange in a cloud-based network. Another solution oriented to the future 5G is proposed in Lloret et al. (2017) where the problem of continuous patient monitoring is tackled. In that case, machine learning is used to efficiently handle the expected large amount of received data, while 5G technology is considered to support a higher number of simultaneous users than 4G.

In summary, the literature review of LTE-based remote healthcare systems highlights how the proposed solutions are mostly presented and discussed in terms of communication architectures and scheduling protocols, with performance evaluation being considered from an applications point of view. On the other hand, no discussion is reported about the LTE framework capability to support such services. That is, LTE

is introduced as part of the system, but the analysis about network traffic is very often neglected.

# 3 HEALTHCARE DATA COMPRESSION AND SYNCHRONIZATION

The performance of remote healthcare monitoring systems mainly relies on the efficient usage of resources at the network layer. However, the interaction between end device and cloud server is first driven by upper layer protocols responsible for the management of data generated by sensors before being transmitted. In this regard, data compression algorithms are typically implemented on the end devices in order to reduce the amount of information to be sent. Following such an approach returns a twofold optimization. First, energy saving is provided for devices since the time spent in active mode for transmission is minimized (this aspect becomes particularly important when dealing with battery-supplied IoT devices). Then, with a view to networks with thousands of entities simultaneously connected, the traffic load may be significantly reduced. The authors in Stojkoska and Nikolovski (2017) propose a novel coding scheme for delta compression tailored to delay insensitive data transmissions. By considering the time correlation of data, significant performance in terms of data reduction can be achieved, at the expense of a very low computational cost and low device power consumption. Spatio–temporal correlation of data can be fruitfully exploited for big data scenarios, as discussed in Moon et al. (2017). Specifically, by means of Discrete Cosine Transform and Fast Walsh–Hadamard Transform, an efficient lossy data compression can be achieved in terms of both the compression ratio and the error rate. Dealing instead with IoT streaming services, a content-sensitive compression algorithm is introduced in Hsu et al. (2017). Following a lossy approach, video frame reduction is operated with the aim of achieving a convenient trade-off between data reliability and storage reduction.

The same issues concerning image and video compression can also be found in healthcare applications, especially regarding the compression of continuous and burst data measurements. In this direction, a potential use case is represented by remote electrocardiography (ECG) data transmission. The work in Deepu et al. (2017) presents a two-step algorithm for ECG data compression, specifically designed to be energy efficient in a wireless IoT sensors context. First, the size of data to be transmitted is minimized by resorting to lossy high compression ratio techniques. Then, entropy coding is used to minimize potential decompression errors. On-chip ECG data management is also addressed in Joseph et al. (2014) where a Discrete Wavelet Transform based compression algorithm is introduced. With low complexity processing, suited to battery-powered sensors, significant performance in terms of the compression ratio is achieved. Another example of healthcare data compression may concern electroencephalography (EEG). The authors in Nasrallah et al. (2015) present a combined lossy/lossless data compression technique that relies on Discrete Cosine

Transform and adaptive differential pulse coded modulation with the aim of minimizing the data volume to be transmitted while maintaining a high level of information integrity.

The use of data compression is particularly effective when dealing with cloud-based applications where sensors exchange information with a data center. In this scenario, the cloud saves the updated version of data generated by sensors and devices. However, it usually happens that consecutive versions of the same information are not that different from one another; therefore it may be convenient for the sensors to transmit only the new parts of data that are necessary for the cloud to update. In this direction, data compression can be combined with the so-called data synchronization algorithms that rule the interaction between end device and cloud. Timestamp, Bitmap, and RAKE algorithms are exploited in Sari and Riasetiawan (2018) to reduce the amount of traffic between IoT devices and the cloud. The combination of such techniques drives the processing of data in terms of compression, synchronization, and decompression. A multi-layered framework for data compression and storage optimization is described in Hossain and Roy (2018). By conveniently processing the output of the IoT sensors, an optimal network bandwidth usage is achieved during data transmission, furthermore, providing low decompression loss. The work in Petroni et al. (2018) introduces an adaptive data synchronization algorithm, tailored to the IoT context but is also efficient in heterogeneous traffic scenarios. By exploiting a bi-directional communication link, the cloud sends a compressed version of its data to the end device. The end device then operates the comparison between its data and those coming from the cloud to recognize the new parts of information that have to be transmitted for the cloud update.

Overall, compression and synchronization can be achieved through either lossless or lossy approaches. In the first case, data integrity is guaranteed but at the expense of a limited performance in terms of the compression ratio. On the other hand, with the lossy approach the traffic volume can be significantly reduced, but the reliability of data cannot be fully guaranteed. In the healthcare context, data integrity is a crucial issue especially when dealing with applications for emergency scenarios, where the information must be delivered without errors. Therefore, the choice of the most effective method has to be driven by the characteristics of the context under investigation. Finally, another aspect that must be stressed is that most of the mechanisms for healthcare data compression relies on features extraction and pattern recognition algorithms specifically tailored to the data. That is, they are efficient when dealing with specific data, but they may fail if applied to other data patterns (for instance, the method for ECG measurements compression may not provide the same performance when working on EEG data). So, having a data content-insensitive compression mechanism may be convenient so as to guarantee good performance despite the nature of the information that needs to be processed. Such an aspect becomes even more relevant in scenarios where compression is operated on heterogeneous data, potentially coming from different sensors.

FIGURE 2 | Healthcare remote monitoring scenario.
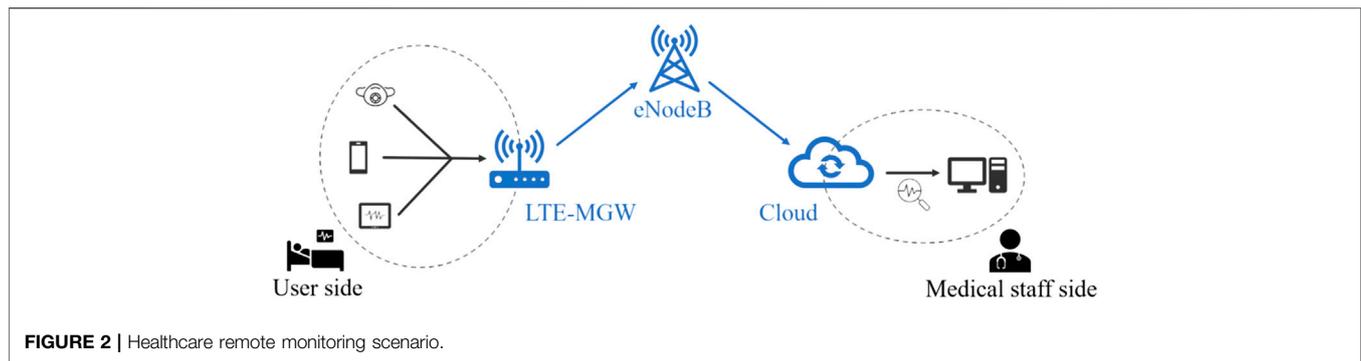
4 ARCHITECTURE MODEL

In this section, we present the network model implementing the envisaged remote healthcare monitoring service. Let us refer to a cloud-based scenario, where data coming from some patient medical instrument/sensor have to be stored in a data center, making them remotely available for medical personnel and doctors. By assuming the final user/patient to be potentially provided with multiple devices, we consider the presence of an LTE medical gateway (LTE–MGW) responsible for patient data collection and transmission over the LTE network (**Figure 2**). Such a solution demonstrates advantages for two main reasons.

First, medical devices may not be provided with an LTE interface and the problem of power consumption may limit the transmission range to only a few meters. Therefore, the LTE–MGW allows entities locally connected in a PAN/LAN context (for instance with Bluetooth or Wi-Fi) to be part of a wider network context. The second benefit concerns the cloud service management, where data processing and transmission considers only LTE–MGW and the data center. So, the computational effort for medical sensors is significantly reduced, since they are no longer directly involved in the cloud network. From an applications point of view, the LTE–MGW implements data compression and synchronization protocols, allowing the optimization of traffic towards the eNodeB and the cloud server. In this direction, the presence of an LTE–MGW can be seen as part of an edge computing paradigm, where the distance between end-devices and the cloud is *virtually* reduced by moving some data processing and storage on an intermediate node (Bangui et al., 2018).

In detail, we present the proposed healthcare remote monitoring system through a two-layered description. The first layer is referred to as the characteristics of the LTE network and is what the service relies on, while the second one concerns the application handling data synchronization between LTE–MGW and the cloud server. So, we do not deepen the interaction between medical sensors and LTE–MGW since this part of the network is not directly involved in the cloud service.

4.1 Network Layer

In this work, we take as a reference model for the non-standalone 5G configuration, where LTE is considered to as access

technology in compliance with the last releases by 3GPP (3GPP, 2019). An example of an access/core network could be obtained by considering two different network segments, namely Evolved Packet Core (EPC) and LTE Radio Access Network (RAN). The EPC, including the S1-U Protocol Stack (GTP over UDP/IP), is composed of the Mobility Management Entity (MME), the Serving Gateway (SGW) and the Packet Data Network Gateway (PGW), and it is responsible for the connection of eNodeBs (eNBs), between them and to external networks (i.e., internet). The LTE RAN segment, including the LTE Radio Protocol Stack (PHY, MAC, RLC, PDCP and RRC), represents the radio access segment for the mobile end-users. The architecture of the LTE cellular network has been realized in ns-3, a discrete event network simulator typically used to evaluate the performance of large-scale systems. Specifically, ns-3 (Baldo et al., 2011, Baldo et al., 2013) provides the modeling of the LTE network data user plane, using two important simulation simplifications. First, it does not implement the control plane, as it is not of interest for most designs and evaluations concerning end transmissions. The other simplification concerns instead the network nodes that connect the LTE system to external networks (for example Internet), i.e., the Packet Gateway (PGW) and the Service Gateway (SGW), which are simulated through a single entity. **Figure 3** shows the network architecture of the LTE-EPC model as defined in the simulator.

The unification of SGW and PGW in a single node has allowed the S5 interfaces/S8, specified by 3GPP, to be removed, thus simplifying the model. On the other hand, the S1-U, S11 and X2 interfaces (necessary for handover procedures) and the entire protocol stack (PHY, MAC, RLC, PDCP and RRC) of the LTE access network have been modeled as specified by 3GPP. **Figure 3** highlights two network segments: the LTE segment, where users (UEs), identified with the LTE-MGW, connect to the eNodeB, and the EPC segment, where eNodeBs are connected between each other and the internet by means of point-to-point optical fiber links.

Without loss of generality, we refer to a single frequency channel to evaluate the radio coverage and capacity of the LTE RAN, which are expected to be dependent on the number of users (that is, LTE–MGWs) simultaneously connected. Each LTE–MGW is considered as transmitting the data available after compression as fast as possible so as to exploit the whole bandwidth allowed, defined as a function of the
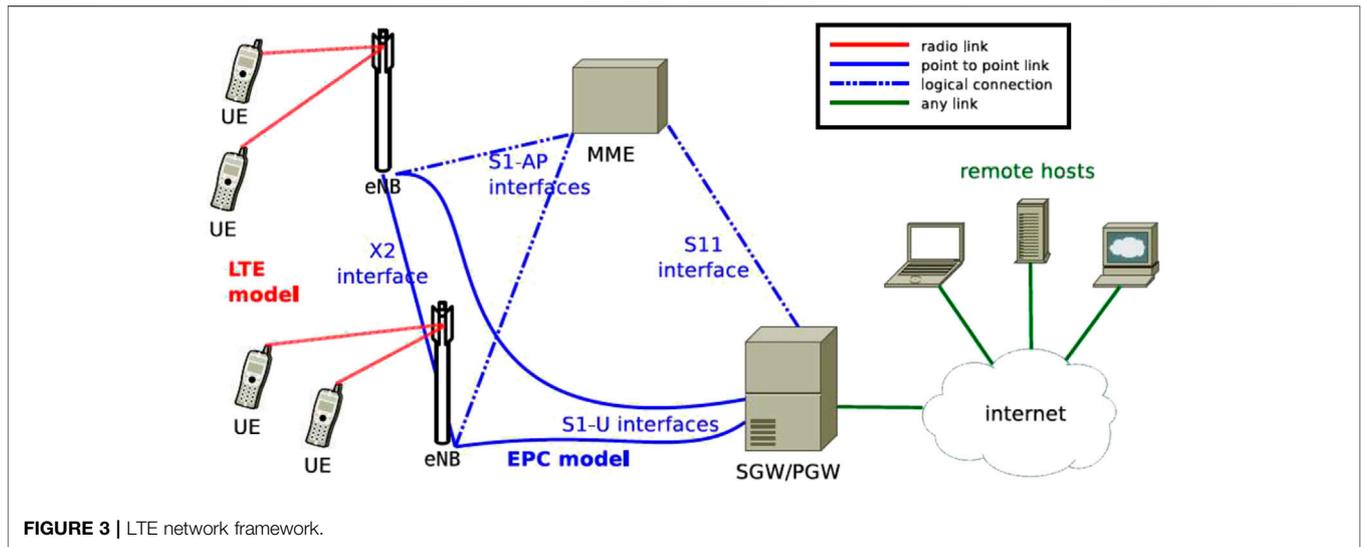
FIGURE 3 | LTE network framework.

signal-to-interference-plus-noise ratio (SINR) and modulation and coding schemes (MCS) characterizing LTE. Further details concerning the simulation scenario are provided in the next section.

4.2 Application Layer: Data Synchronization

The procedure for cloud update is handled at the application layer by means of so-called data synchronization protocols. The aim of such protocols is to make the end device and the cloud cooperate in order to perform effective data sharing, minimizing the amount of overhead information to be transmitted. Therefore, data synchronization results in a sort of information compression, since only part of the data generated by the end device are actually sent to the cloud. For our purposes, we consider the implementation of the synchronization algorithm introduced in Petroni et al. (2018) the essentials of which are recalled below.

Given a client-server scenario, two entities referred to as A and B, respectively, are considered. The client A, represented by the LTE-MGW, stores some data, namely F_A, that have to be transmitted to the server B, which is the cloud. The server B buffers F_B, which is an older version of F_A, which therefore has to be updated. Such a procedure can be accomplished by replacing F_B with the newly received F_A, however this would require F_A to be fully transmitted from A to B. Since the content of F_A and F_B are likely to be unchanged in some parts, it is sufficient for A to send only the new information in F_A which is strictly necessary for B to the update F_B to F_A. Such amount of data is usually referred to as *Delta* (Δ), since it represents the difference between F_B and F_A. By doing so, the dimension of Δ may be significantly smaller than the entire size of F_A, therefore the information overhead to be transmitted is minimized, allowing the network traffic load to be reduced. The computation of Δ is performed by exploiting a bi-directional communication between A and B, described as follows:

1. The server B organizes its data version F_B in non-overlapping blocks, referred to as chunks, of size B_c. Chunks are subject to a double-hashing compression and sent to the client A.
2. Client A scans its current data F_A by means of a moving window exactly equal to B_c. For each portion of data under investigation, the client checks for potential matchings with any chunk received from the server. By doing so, client A is able to compute the differences between F_A and F_B and to store them in a file Δ. Specifically, Δ is composed of an ordered list of literal bytes and tokens. Literal bytes represent the *new* information gathered in F_A but not in F_B, while tokens refer to the index of chunks recognized as present in both F_A and F_B.
3. The file Δ (the size of which is likely to be much smaller than F_A) is sent to the server B and exploited for the update F_B, finally resulting in the exact copy of F_A.

Concerning the procedure for data synchronization described above, two aspects deserve to be highlighted. First, differences in the data computation strictly relies on the chunk size B_c. In fact, if B_c is too large, the probability to find matchings between F_A and F_B may be low, therefore Δ will be mainly composed of literal bytes, with its size becoming comparable with the dimension of F_A (so, the gain in terms of traffic reduction may not be significant). On the other hand, if B_c is small, the computational effort to obtain Δ grows. Typically, data synchronization protocols relying on the delta computation approach are meant to work with static parameters, therefore traffic optimization performance depends also on the characteristics and heterogeneity of the data to be handled. In this direction, in order to make synchronization protocols independent of the nature of data, an adaptive algorithm to calculate the optimal chunk size has been implemented. Specifically, the solution proposed exploits the information about tokens (their number and position within the file Δ) recognized during the current synchronization procedure to predict the optimal value for B_c to be used in the next synchronization event. In fact, having consecutively

FIGURE 4 | Data synchronization mechanism implemented at the LTE-MGW side.
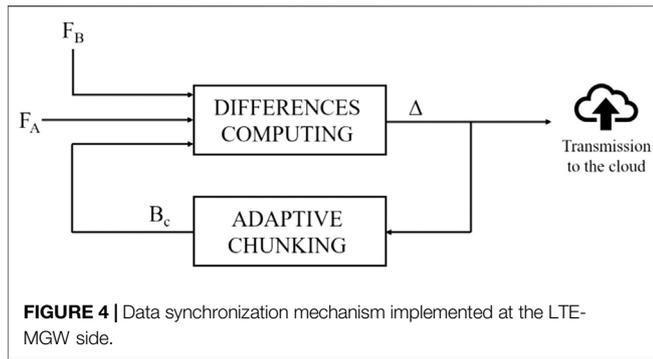
indexed tokens means that there is a portion of data, larger than B_c, that has remained unchanged. Therefore, the chunk size could be conveniently increased so that the number of tokens to be stored in the file Δ is reduced, thus reducing its size. On the other hand, when two tokens are not consecutive, it means that some updates are in between. So, the chunk size should be conveniently decreased to reduce the transmission of redundant literal bytes. By analyzing the sequence of tokens in the file Δ, we are able to calculate several estimates of the new chunk size as follows [see Petroni et al. (2018) for further details]:

$$\begin{cases} B_c^{(m)} = B_c + \mu N_{ac} \\ B_c^{(m)} = B_c - \mu N_{ud} \end{cases} \quad (1)$$

with the first row describing the mth estimate obtained by means of the current chunk size increase, while the second row refers to the mth estimate obtained through the decrease of B_c. In detail, the chunk size estimates are a function of the number of tokens stored consecutively and the number of updated bytes expected between two non-consecutive tokens, respectively N_{ac} and N_{ud}. Finally, μ is a step size ruling the speed of adaptation of the estimates. Depending on tokens™ distribution in Δ, M chunk size estimates are obtained, with the final resulting averaged chunk size given by:

$$B_{c,NEW} = \frac{1}{M} \sum_{m=1}^{M} B_c^{(m)} \quad (2)$$

So, the value $B_{c,NEW}$ represents the new chunk size to be used during the next data synchronization event. The second issue to be discussed is about the computation of Δ and the update of F_B, which requires the client and server to cooperate through a bi-directional communication. Such a paradigm may not be convenient because of the potential latency introduced to implement both a downlink and an uplink transmission flow. However, it is worth noting that devices acting as LTE–MGW are typically provided with large storage resources. Therefore, we can reasonably assume that the LTE–MGW to buffer both the new data F_A and the old one, essentially, corresponds to F_B stored in the cloud. In this scenario, the downlink communication from the server to the client can be avoided since the LTE–MGW has all the information necessary to compute Δ. A pictorial description of data synchronization architecture implemented at the LTE–MGW side is reported in **Figure 4**.

The block responsible for computing Δ receives as input F_A and F_B, both stored in the device. The adaptive chunking algorithm, returning the optimal value B_c (actually referred to as $B_{c,NEW}$ in **Eq. 2**) driving the differences in computation, operates instead in a feedback loop since it exploits the output Δ to predict the input parameter for the other block.

As a final remark, it is worth noting that the data synchronization algorithm considered here operates a lossless compression and is information insensitive, which means it works independently of the pattern of data to be handled. It can therefore be easily implemented for any kind of application. On the other hand, the literature review reported in **Section 3** has shown that many data compression algorithms are tailored to specific data types such as ECG and EEG, so they may not be suited to different scenarios.
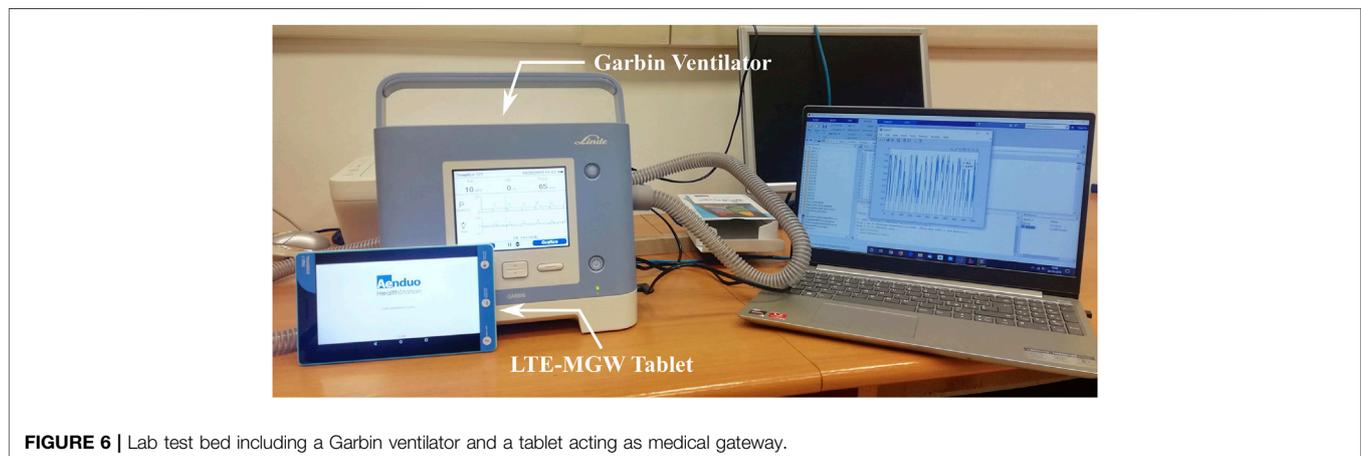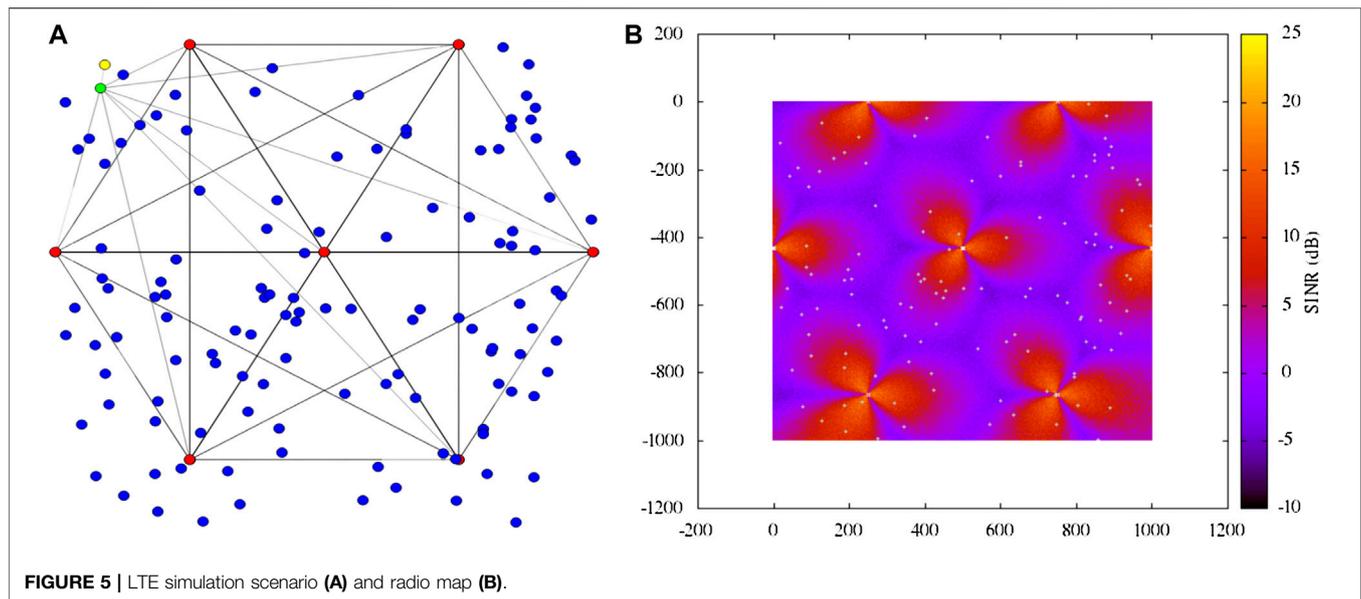
5 SIMULATION RESULTS

The feasibility and performance of the proposed LTE-based remote healthcare monitoring service has been evaluated by means of several network simulations performed with ns-3. The scenario has been designed to represent a typical case of COVID-19 patient monitoring at home. Indeed, some radio access points are organized to provide coverage in a given area and serve groups of users concentrated in some locations (for instance homes, buildings, city districts). In this direction, we consider the LTE scenario depicted in **Figure 5A**, where seven radio access points are organized in a hexagonal spatial layout. Each site is designed with three antennas allowing the coverage area to be divided into three 120° sectors (**Figure 5B**).

According to the proposed architecture, the access points are connected via X2 interfaces (black lines), while the communication PGW/SGW (green node) is realized through dedicated point-to-point links. The presence of the network server (yellow node) has also been included to simulate the end-to-end communication performance. Finally, the blue nodes represent the LTE-MGWs randomly distributed in the considered area.

The data traffic has been modeled considering different patterns of real healthcare parameters that may be measured for COVID-19 patients affected by respiratory issues. Specifically, we divide users into three categories, depending on the service they apply for:

1. Service 1 (S1): data refer to the tracking of respiratory rate and pressure;
2. Service 2 (S2): data on oxygen saturation are provided together with the parameters measured in S1;
3. Service 3 (S3): more complex data about patient health conditions are considered, including respiratory rate, pressure, tidal volume, inspiration and expiration flows.

The services have been simulated by considering oxygen saturation data taken from the available database in Physionet.org (2020), while the other breathing measurements have been obtained from a Garbin ventilator (Linde Medicale,

FIGURE 5 | LTE simulation scenario **(A)** and radio map **(B)**.

FIGURE 6 | Lab test bed including a Garbin ventilator and a tablet acting as medical gateway.

2019) employed in a lab test bed (**Figure 6**). The LTE-MGW has instead been implemented in a tablet, connected to the LTE network via a SIM card and is capable of quite a large data storage (in the order of gigabytes).

It is worth highlighting that S1, S2, and S3 gather measurements potentially coming from different medical devices, therefore we identify each user with an LTE-MGW that is able to perform the data collection as described in **Section 4**. Such a choice is a perfect fit for the home monitoring scenario that we are investigating. However, it is worth noting, for the sake of completeness, that it may also be possible for a single LTE–MGW to serve multiple users in specific scenarios where several patients are placed in a restricted area such as a hospital room or ward. In that case, since data coming from health sensors/instruments are typically provided with a unique ID identifying the corresponding patient, information management between LTE–MGW and the cloud server may be

easily handled. In general, for tablet-like devices acting as LTE–MGW no problems are expected in terms of storage saturation, neither for single nor for multiple users scenarios. Despite the fact that patients may connect to several sensors, the average daily traffic is in order of megabytes, which is far below the LTE–MGW storage capacity. Moreover, note that the LTE–MGW essentially acts as a buffer, storing data only for the time necessary to update the corresponding data in the cloud. Then storage space can be released and reused for future measurements.

Simulations have been performed to evaluate two main aspects of remote monitoring: the impact of the number of users and the amount of data exchanged on the network traffic load. The analysis of the available medical datasets has revealed that daily measurements of users are stored in multiple files, the number of which varies from three to 11 per user, generated at different times. Therefore, we conveniently consider such a

TABLE 1 | Data traffic characteristics for S1, S2, and S3.

Service	Av. daily transmissions	Av. maximum data size (single transmission)	Av. minimum data size (single transmission)	Av. daily data volume
S1	9.35	684.61 (511.31) kB	34.01 (24.32) kB	1887.84 (1,413.45) kB
S2	8.32	704.28 (522.07) kB	42.04 (29.69) kB	1894.41 (1,397.12) kB
S3	6.62	1,625.73 (1,278.34) kB	91.96 (66.12) kB	3,659.12 (2,864.93) kB
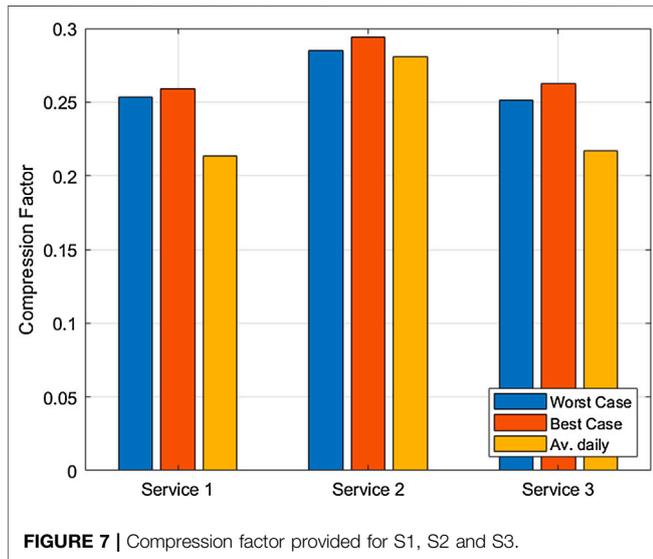
FIGURE 7 | Compression factor provided for S1, S2 and S3.

feature to model the user daily data traffic. Furthermore, the size of the available datasets allowed us to simulate the presence of up to 120 users. The LTE-MGWs are implemented to work both with the data synchronization protocol described in **Section 4** and without any data compression protocol. By doing so, we are able to evaluate the effect of traffic reduction on the network performance. In this regard, **Table 1** reports some reference information about the average data volume generated by S1, S2, and S3. It is worth highlighting that the algorithm for data synchronization introduced in **Section 4** has been implemented in a tablet acting as an LTE-MGW, so the results in **Table 1** actually come from real tests (the network performance analysis has instead been performed by means of simulations).

The second column reports the average number of daily transmissions performed by users belonging to the S1, S2, and S3 service categories, respectively. The third and fourth columns show the average size of a single transmission data. From a network point of view, the maximum data size column corresponds to the worst case, while the minimum data size reference is the best one. Finally, the last column describes the daily average data traffic volume. For table cells containing two values, the first one refers to the full data size, while the value indicated in parentheses is obtained after data compression. Furthermore, based on the reference values in **Table 1**, we measure the performance of the employed data synchronization algorithm in terms of the compression factor, expressed as:

$$CF = 1 - \frac{\Delta}{F_{\text{tot}}} \qquad (3)$$

where F_{tot} is the amount of data collected by the LTE–MGW from medical devices and Δ is the size of information effectively transmitted to the cloud. The value of CF is expected to range from 0 to 1. Specifically, if CF approaches 0, it means that the data reduction gain provided by the synchronization algorithm is low. On the other hand, the higher the CF, the better the data compression achieved. **Figure 7** reports the average compression factor measured for S1, S2, and S3. We refer to the scenario where users concurrently try to transmit the maximum sized data in a single transmission (column three of **Table 1**) as a worst case, while the best case refers to users concurrently transmitting the minimum volume size of data (column four of **Table 1**). The last metrics instead concerns the average daily data traffic.

By observing the results, it is possible to appreciate the quite flat performance provided by the data synchronization algorithm, with values ranging from 0.2 to 0.3. In detail, S2 seems to be the service taking best advantage of data compression. This is due to the fact that measurements referring to oxygen saturation may be quite constant in time; therefore, the resulting data can be conveniently compressed. On the other hand, S1 and S3 measure variable parameters like breath flow, therefore the achievement of a higher compression factor mainly depends on the pattern of data to be handled.

The second part of the analysis provides coverage and capacity evaluations for both the scenarios of compressed and uncompressed data. In particular, we focus on the uplink network direction, considering a certain number of LTE-MGWs (data traffic originator), deployed inside the cellular network, and one destination (gathering server), located outside the cellular network domain. To this aim, during the simulation each traffic originator generates data as fast as possible, trying to fill the bandwidth, up to a maximum quantity of bytes given by the dimension of real data collected in **Table 1**. Once the lower layer transmission buffer is filled, the LTE–MGW waits until space is free to send more data, essentially keeping a constant information flow. Moreover, considering the reliability as the main KPI related to patient monitoring service, we use TCP as the transport protocol.

Network performance is evaluated in terms of the following network metrics:

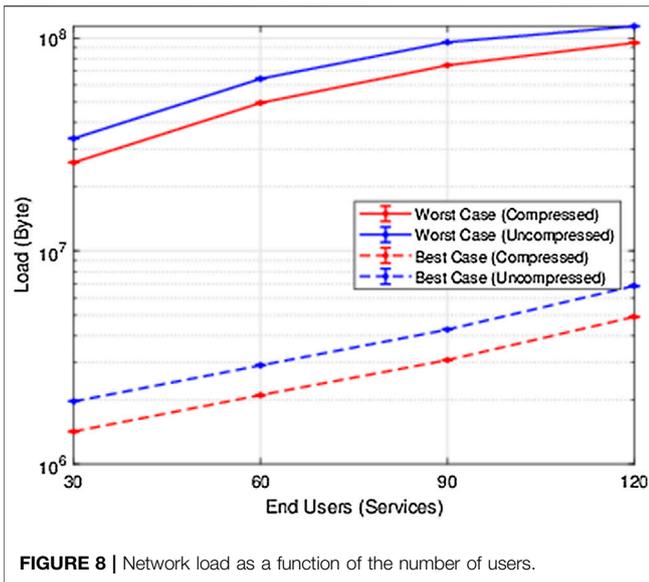- Load: total bytes received by the server during the whole simulation.

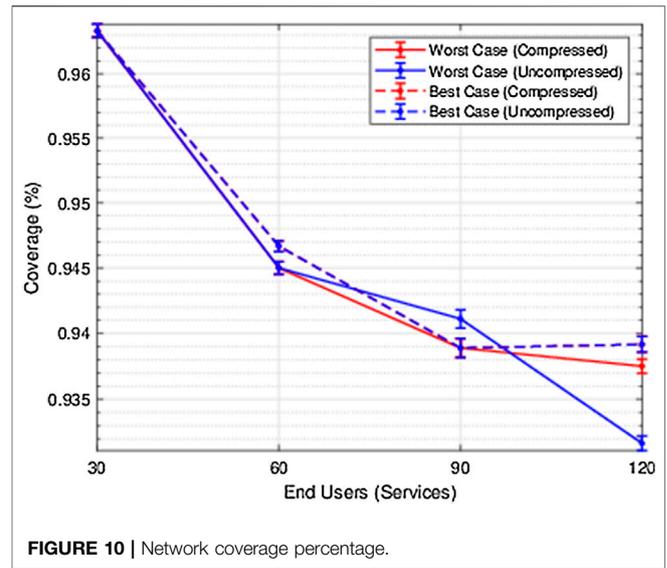**FIGURE 8 |** Network load as a function of the number of users.

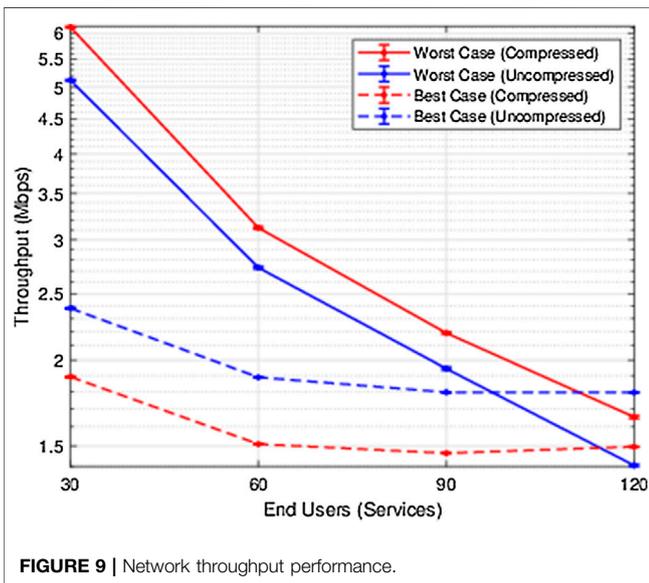**FIGURE 10 |** Network coverage percentage.

FIGURE 9 | Network throughput performance.

- Coverage: ratio between the number of LTE–MGWs able to upload their data on the remote server and the total number of LTE–MGWs deployed in the scenario.

Additionally, service performance is evaluated in terms of the following end-user metrics:

- Throughput: total received bits over to the total receiving time interval.
- Packet Delivery Ratio (PDR): ratio between the number of correctly received packets and the number of related transmitted ones.
- End-to-End Delay (E2E-Delay): end-to-end delay between the transmitting and the receiving of a data packet.

- Jitter: variation of the end-to-end delay between packets belonging to the same data flow.

The performance of the LTE network supporting healthcare remote monitoring has been measured as a function of the number of users to be served, ranging from 30 to 120. These users are uniformly deployed in a squared simulation area of 1 km. Moreover, to statistically validate the simulation results, we perform 10 simulations for each scenario and show the final results in terms of mean value and Confidence Interval of 99% for each performance metric. Users have been distributed uniformly among the services S1, S2, and S3. **Figure 8** depicts the trend of network traffic load as a function of the number of users simultaneously transmitting, hence concerning the connection between users and eNodeBs.

The curves referring to the worst and best case (previously introduced) suggest two important results. First, it can be observed how the effect of data compression allows the network load to be reduced with respect to the case where compression is not considered. In other words, given the number of transmitting users, the network load is reduced when data compression is applied. This is true for any number of considered users, in fact the red lines are always under the corresponding blue lines. As expected, the network load is proportional to the amount of data that needs to be transmitted, therefore the worst case presents higher values than the best one. Second, by having a specific network load as a target, it is possible to appreciate how the use of data compression allows a larger number of users to be served in the network.

The LTE performance analysis also concerned the throughput, expressed in Megabits per second (Mbps), described in **Figure 9**.

Interestingly, we observe that, when the number of users is sufficiently small, the throughput is higher in the worst case (that is, when the maximum amount of data is transmitted) than in the best case. This means that when the network load is far from the RAN saturation level, users transmitting a large amount of data

FIGURE 11 | Network packet delivery ratio (PDR) performance.
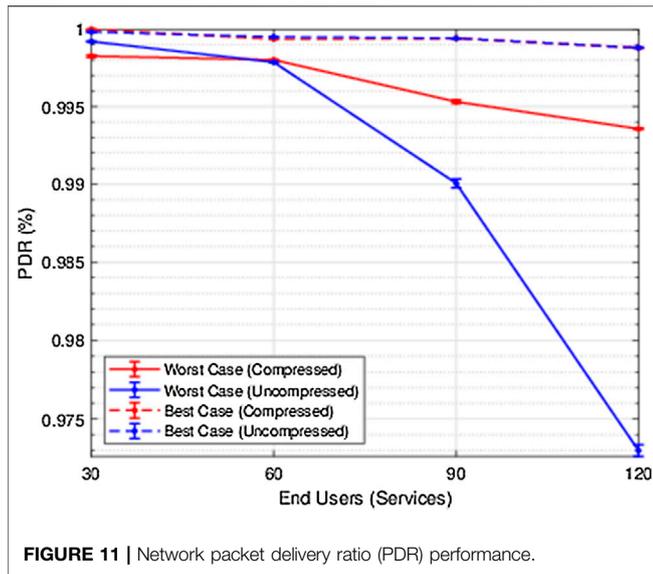
can fully exploit the available network resources and can attain a high throughput. On the other hand, when the amount of data to be handled is limited (that is, in the best case), user performance results are slightly lower because the traffic volume has less impact on the network load. Such behavior is no longer respected when the number of users grows. In fact, as the network load increases, users throughput is necessarily reduced.

From another perspective, QoS provided to users has been measured in terms of coverage and PDR. Specifically, the coverage refers to the percentage of users allowed to access the LTE resources for transmitting, as depicted in **Figure 10**.

In the worst-case traffic scenario, the difference between handling compressed an uncompressed data is not relevant if the number of users is small. On the other hand, a larger presence of active users makes the benefits of data compression more appreciable. When dealing with the best-case scenario instead, the need to transmit a lower amount of information makes the effect of compression insignificant even for a large number of users. Such results

can be confirmed by looking at **Figure 11** which describes the PDR.

Stable trends can be observed for both the worst and best scenarios, with a slight performance decrease only affecting the worst case where uncompressed data are considered. However, the sloping of the blue solid line suggests that, in view of the presence of a larger number of users, data optimization could be crucial in avoiding a significant performance decrease in terms of PDR.

The last aspect investigated is the network latency, evaluated through the E2E-delay and jitter. Overall, by observing both **Figures 12A,B**, how the amount of data to be transmitted impacts on the latency is highlighted.

By recalling the results in **Figure 7**, since the compression factor is quite constant for all the considered scenarios, the trend of blue and red lines are similar except for an amplitude factor. Therefore, the difference in terms of performance is essentially dependent of the average traffic characterizing the worst and the best case.

In summary, the analysis presented here has highlighted how the LTE system, in its current state, can be suitably exploited to support remote healthcare monitoring services. Specifically, we measured the impact of data traffic generated by tens of users on the LTE performance. In this regard, we also investigated the effect of data compression, resulting in fundamental reductions of the network load. To the best of our knowledge, this type of study has not been conducted thus far, so we were not able to perform a comparison with other analyses in the literature. The work in Hindia et al. (2016) reports the throughput analysis for different classes of IoT applications, with the goal of demonstrating that the proposed priority-based data scheduling algorithm is able to efficiently serve up to 50 users, providing the requested QoS. It deals with an IoT scenario where health data are essentially spot measures (blood pressure, temperature, heart rate), therefore the resulting traffic pattern is quite different from what we considered in our study. Moreover, the provided throughput is in the order of 250 Kbps, which is significantly lower than the result we achieved (some Mbps). Data transmission scheduling is also addressed in Adibi (2014), but no performance investigation is provided for the LTE network, which is what the proposed service relies on. The authors

FIGURE 12 | LTE network end-to-end delay **(A)** and jitter **(B)**.
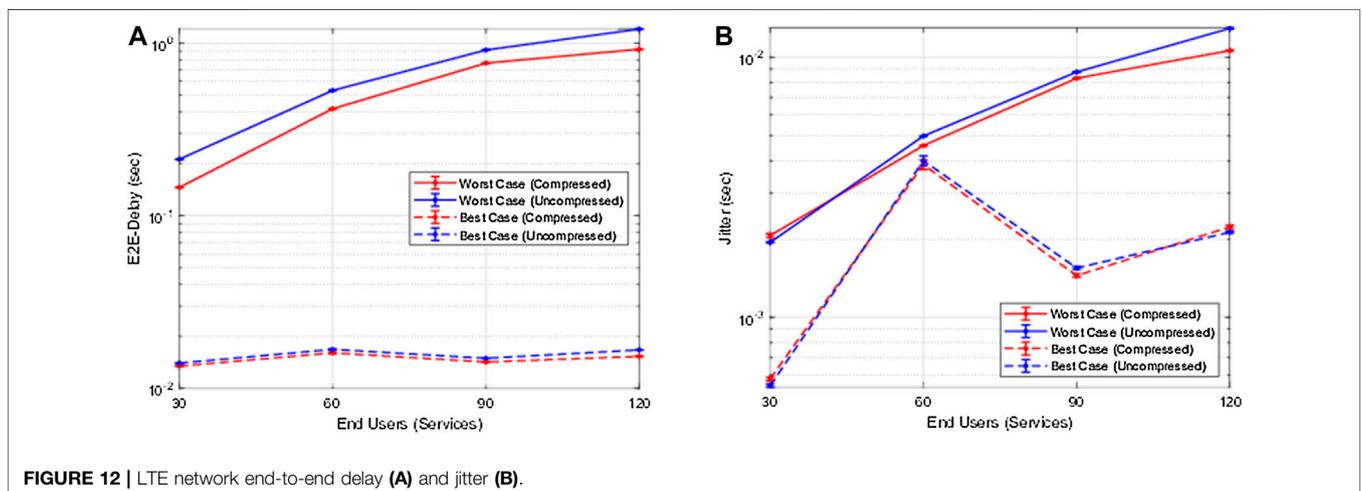
in Hadi et al. (2019) apply the principles of big data analytics in the context of LTE-based out-patients monitoring services. However, the analysis is oriented to the physical layer, measuring the quality of the service in terms of signal-to-interference-plus-noise ratio, while networking issues are not considered.

The lack of network performance analyses also characterizes the literature on 5G-based solutions for healthcare. Some examples are given in Lloret et al. (2017), Zhang et al. (2020) where 5G is considered as the network framework supporting continuous health monitoring activities. However, these studies exclusively concern the scheduling and processing of data, and do not focus on the network performance.

An important aspect to be recalled for 5G is that it considers network slicing, essentially allowing the optimal assignment of resources depending on the QoS requested by the applications. In the field of e-Health, 5G is specifically envisaged to provide a significant boost to services requiring high reliability and low latency (belonging to URLLCs category depicted in **Figure 1**), such as remote surgery (robotic aided) and health alarm systems (Silva et al., 2020). In fact, several works have recently been proposed about the delay of such critical applications (Acemoglu et al., 2020). On the other hand, we expect no particular impact on patients remote monitoring activities (categorized as eMBB service in **Figure 1**), since such kind of service is already supported by LTE. Of course, it is expected that with 5G the network access will be extended to a larger number of users/ devices than in LTE. However, the most important difference between 5G and LTE, in the healthcare field, is related to the support of very low latency applications. Finally, it is worth noting that the architecture considered here shares the same edge computing principles than with future 5G networks, so it is perfectly tailored to upcoming network scenarios.

6 CONCLUSION

In this paper, we investigated the feasibility of an LTE-based remote monitoring system oriented to healthcare applications, so as to allow the physical interaction between patients and medical staff to be substituted, whenever possible, with a reliable, virtual one. The considered monitored data (coming from respiratory medical devices) as well as the kind of patterns from patients represent practical cases that may arise in case of the monitoring of

COVID-19 patients who are at home. To this aim, we evaluated the amount of traffic that can be supported by the LTE to assess scalability issues that may be related to the number of patients to be served. Furthermore, the impact of data compression is investigated, highlighting its benefits when dealing with cloud-based services. The emergency related to the COVID-19 pandemic has highlighted the importance of technology to support the healthcare system. In this context, we showed how remote patient assistance can be effectively realized by exploiting the current LTE framework. In view of the upcoming 5G Era, remote healthcare monitoring applications such as the one investigated here will become part of the eMBB services, allowing a very large number of users to be simultaneously connected, providing faster network access. In this direction, services that already rely on edge computing principles will be a perfect fit for such a new scenario. Therefore, future works will deal with the investigation of such new paradigm, focusing the attention on the data compression for increasing data traffic scenarios.

DATA AVAILABILITY STATEMENT

AUTHOR CONTRIBUTIONS

FC, AP, and PS contributed to conception and design of the study. AP wrote the first draft of the manuscript. PS organized the database. FC supervised the simulation analysis and manuscript writing. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

REFERENCES

3 GPP, 2019 3 GPP. (2019). *Release 15 specifications*.

Acemoglu, A., Krieglstein, J., Caldwell, D. G., Mora, F., Guastini, L., Trimarchi, M., et al. (2020). 5g robotic telesurgery: remote transoral laser microsurgeries on a cadaver. *IEEE Trans. Med. Robot. Bionics* 2, 1. doi:10.1109/TMRB.2020. 3033007

Adibi, S. (2014). Biomedical sensing analyzer (BSA) for mobile-health (mhealth)-LTE. *IEEE J. Biomed. Health Inform.* 18, 345–351. doi:10.1109/JBHI.2013. 2262076

Adibi, S. (2018). Mobile health personal-to-wide area network disaster management paradigm. *IEEE Sensors J.* 18, 9874–9881. doi:10.1109/JSEN. 2018.2872418

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., and Ayyash, M. (2015). Internet of things: a survey on enabling technologies, protocols, and applications. *IEEE Commun. Surv. Tutorials* 17, 2347–2376. doi:10.1109/comst. 2015.2444095

Alsaeedy, A. A. R., and Chong, E. K. P. (2020). Detecting regions at risk for spreading covid-19 using existing cellular wireless network functionalities. *IEEE Open J. Eng. Med. Biol.* 1, 187–189. doi:10.1109/ojemb.2020.3002447

Baldo, N., Miozzo, M., Requena-Esteso, M., and Nin-Guerrero, J. (2011). "An open source product-oriented lte network simulator based on ns-3", in *Proceedings of the 14th ACM international conference on modeling, analysis and simulation of wireless and mobile systems*. (New York, NY: Association for Computing Machinery), 293–298. doi:10.1145/2068897.2068948

Baldo, N., Requena-Esteso, M., Miozzo, M., and Kwan, R. (2013). "An open source model for the simulation of lte handover scenarios and algorithms in ns-3", in

Proceedings of the 16th ACM international conference on modeling, analysis & simulation of wireless and mobile systems. New York, NY: Association for Computing Machinery, MSWiM '13, 289–298. doi:10.1145/2507924.2507940

Bandyopadhyay, D., and Sen, J. (2011). Internet of things: applications and challenges in technology and standardization. *Wirel. Pers. Commun.* 58, 49–69. doi:10.1007/s11277-011-0288-5

Bangui, H., Rakrak, S., Raghay, S., and Buhnova, B. (2018). Moving to the edge-cloud-of-things: recent advances and future research directions. *Electronics* 7, 309. doi:10.3390/electronics7110309

Bennis, M., Debbah, M., and Poor, H. V. (2018). Ultrareliable and low-latency wireless communication: tail, risk, and scale. *Proc. IEEE* 106, 1834–1853. doi:10.1109/jproc.2018.2867029

Centenaro, M., Vangelista, L., Zanella, A., and Zorzi, M. (2016). Long-range communications in unlicensed bands: the rising stars in the iot and smart city scenarios. *IEEE Wirel. Commun.* 23, 60–67. doi:10.1109/mwc.2016.7721743

Chamola, V., Hassija, V., Gupta, V., and Guizani, M. (2020). A comprehensive review of the covid-19 pandemic and the role of IoT, Drones, AI, Blockchain, and 5G in managing its impact. *IEEE Access* 8, 90225–90265. doi:10.1109/access.2020.2992341

Cical, S., Mazzotti, M., Moretti, S., Tralli, V., and Chiani, M. (2016). Multiple video delivery in m-health emergency applications. *IEEE Trans. Multimed.* 18, 1988–2001.

Deepu, C. J., Heng, C., and Lian, Y. (2017). A hybrid data compression scheme for power reduction in wireless sensors for IoT. *IEEE Trans. Biomed. Circuits Syst.* 11, 245–254. doi:10.1109/TBCAS.2016.2591923

Foubert, B., and Mitton, N. (2020). Long-range wireless radio technologies: a survey. *Future Internet* 12, 13. doi:10.3390/fi12010013

Hadi, M. S., Lawey, A. Q., El-Gorashi, T. E. H., and Elmirghani, J. M. H. (2019). Patient-centric cellular networks optimization using big data analytics. *IEEE Access* 7, 49279–49296. doi:10.1109/access.2019.2910224

Hindia, M. N., Rahman, T. A., Ojukwu, H., Hanafi, E. B., and Fattouh, A. (2016). Enabling remote health-caring utilizing IoT concept over LTE-femtocell networks. *PloS One* 11, e0155077–17. doi:10.1371/journal.pone.0155077

Hossain, K., and Roy, S. (2018). "A data compression and storage optimization framework for iot sensor data in cloud storage", in 21st international conference of computer and information technology (ICCIT). 1–6.

Hsu, C., Fang, Y., and Yu, F. (2017). "Content-sensitive data compression for IoT streaming services", in IEEE international congress on Internet of Things (ICIOT), 147–150.

Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., and Kwak, K. (2015). The internet of things for health care: a comprehensive survey. *IEEE Access* 3, 678–708.

Jiang, H., Starkman, J., Lee, Y.-J., Chen, H., Qian, X., and Huang, M.-C. (2021). Distributed deep learning optimized system over the cloud and smart phone devices. *IEEE Trans. Mobile Comput.* 20, 147–161. doi:10.1109/TMC.2019.2941492

Joseph, B., Acharyya, A., and Rajalakshmi, P. (2014). "A low complexity on-chip ecg data compression methodology targeting remote health-care applications", in 2014 36th annual international conference of the IEEE engineering in medicine and biology society, 5944–5947. doi:10.1109/EMBC.2014.6944982

Linde Medicale. (2019). *Garbin ventilator respiratory data [Dataset]*.

Lloret, J., Parra, L., Taha, M., and Toms, J. (2017). An architecture and protocol for smart continuous ehealth monitoring using 5G networks for IoT and body sensors, *Comput. Netw.* 129, 340–351. doi:10.1016/j.comnet.2017.05.018

Malik, H., Alam, M. M., Moullec, Y. L., and Kuusik, A. (2018). Narrowband-iot performance analysis for healthcare applications. *Proc. Comput. Sci.* 130, 1077–1083. doi:10.1016/j.procs.2018.04.156

Manatarinat, W., Poomrittigul, S., and Tantatsanawong, P. (2019). "Narrowband-internet of things (nb-iot) system for elderly healthcare services", in 2019 5th international conference on engineering, Applied Sciences and Technology (ICEAST).

Moon, A., Kim, J., Zhang, J., and Son, S. W. (2017). Lossy compression on iot big data by exploiting spatiotemporal correlation, in IEEE high performance extreme computing conference (HPEC). 1–7.

Nasrallah, M., El-Hajj, A. M., and Dawy, Z. (2015). "On EEG lossy data compression for data-intensive neurological mobile health solutions", in International conference on advances in biomedical engineering, ICABME, 309–312.

Palattella, M. R., Dohler, M., Grieco, A., Rizzo, G., Torsner, J., Engel, T., et al. (2016). Internet of things in the 5g era: enablers, architecture, and business models. *IEEE J. Sel. Area. Commun.* 34, 510–527. doi:10.1109/jsac.2016.2525418

Petroni, A., Cuomo, F., Schepis, L., Biagi, M., Listanti, M., and Scarano, G. (2018). Adaptive data synchronization algorithm for iot-oriented low-power wide-area networks. *Sensors* 18, 4053. doi:10.3390/s18114053

Physionet.org. (2020). *Pattern analysis of oxygen saturation variability [Dataset]*.

Rehman, I. U., Nasralla, M. M., Ali, A., and Philip, N. (2018). "Small cell-based ambulance scenario for medical video streaming: a 5g-health use case", in *2018 15th international conference on smart cities: improving quality of life using ICT IoT, HONET-ICT*, 29–32.

Sari, K., and Riasetiawan, M. (2018). "The implementation of timestamp, bitmap and rake algorithm on data compression and data transmission from iot to cloud", in 4th international conference on science and technology (ICST), 1–6.

Shafi, M., Molisch, A. F., Smith, P. J., Haustein, T., Zhu, P., De Silva, P., et al. (2017). 5g: a tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE J. Sel. Area. Commun.* 35, 1201–1221. doi:10.1109/jsac.2017.2692307

Silva, R. D., Siriwardhana, Y., Samarasinghe, T., Yliantilla, M., and Liyanage, M. (2020). "Local 5g operator architecture for delay critical telehealth applications", in IEEE 3rd 5G World Forum (5GWF), 257–262. doi:10.1109/5GWF49715.2020.9221292

Stojkoska, B. R., and Nikolovski, Z. (2017). "Data compression for energy efficient ioT solutions", in 2017 25th telecommunication forum (TELFOR), 1–4.

Ullah, H., Gopalakrishnan Nair, N., Moore, A., Nugent, C., Muschamp, P., and Cuevas, M. (2019). 5G communication: an overview of vehicle-to-everything, drones, and healthcare use-cases. *IEEE Access* 7, 37251–37268. doi:10.1109/access.2019.2905347

Wan, L., Guo, Z., Wu, Y., Bi, W., Yuan, J., Elkashlan, M., et al. (2018). 4G\/5G spectrum sharing: efficient 5G deployment to serve enhanced mobile broadband and internet of things applications. *IEEE Veh. Technol. Mag.* 13, 28–39. doi:10.1109/mvt.2018.2865830

Zhang, H., Li, J., Wen, B., Xun, Y., and Liu, J. (2018). Connecting intelligent things in smart hospitals using nb-iot. *IEEE Internet Things J.* 5, 1550–1560. doi:10.1109/jiot.2018.2792423

Zhang, Y., Chen, G., Du, H., Yuan, X., Kadoch, M., and Cheriet, M. (2020). Real-time remote health monitoring system driven by 5g mec-ioT. *Electronics* 9, 1753. doi:10.3390/electronics9111753