Performance of a 3G-Based Mobile Telemedicine System

E. A. Viruete Navarro, J. Ruiz Mas, J. Fernández Navajas, C. Peña Alcega
Communication Technologies Group (GTC) – Aragon Institute of Engineering Research (I3A)
University of Zaragoza
{eviruete, jruiz, navajas, tinape}@unizar.es

Abstract

The performance of a multi-collaborative mobile telemedicine system operating over Third-Generation (3G) mobile networks is presented. It is designed to communicate the personnel of an ambulance with medical specialists in a remote hospital through a Universal Mobile Telecommunications System (UMTS) mobile access. The system architecture is based on advanced signalling protocols that allow multimedia multi-collaborative conferences in IPv4/IPv6 3G scenarios. The system offers real-time transmission of medical data and videoconference, together with other non real-time services. It has been optimized specifically to operate over 3G mobile networks using the most appropriate codecs and mechanisms to improve the quality of the services offered. Evaluation results show a reliable performance over IPv4 UMTS accesses (64 Kbps in the uplink).

1. Introduction

Mobile Health (m-Health) is an emerging area of telemedicine in which the recent development in mobile networks and telemedicine applications converge. m-Health involves the exploitation of mobile telecommunication and multimedia technologies, and their integration into new mobile healthcare delivery systems. The wide coverage provided by cellular networks and their capacity to serve moving vehicles, have brought about new possibilities in telemedicine.

One of the first wireless telemedical systems that utilized Second-Generation (2G) Global System for Mobile Communications (GSM) networks addressed the Electrocardiogram (ECG) transmission issues [1]. In recent years, several m-Health and wireless telemedical systems based on GSM have been reported [2], allowing remote diagnosis in mobile and hardly accessible environments. The recent developments in digital mobile telephonic technologies and their impact on mobility issues in different telemedical and telecare applications are clearly reflected in the fast growing commercial domain of mobile telemedical services. A comprehensive review of wireless telemedicine applications and the latest advances on m-Health systems is presented in [3].

However, 2G-based systems lack the necessary resources to transmit bandwidth-demanding real-time medical data. The Third-Generation (3G) Universal Mobile Telecommunications System (UMTS) overcomes this and other limitations in order to support a large variety of services with different Quality of Service (QoS) requirements, thus extending the use of flexible m-Health applications with higher transmission rates. UMTS also introduces the IP Multimedia Subsystem (IMS) [4], an IPv6 network domain designed to provide support for real-time multimedia services. The IMS uses the Session Initiation Protocol (SIP) as signalling and session control protocol. SIP allows operators to integrate real-time multimedia services over multiple access technologies.

This paper presents a 3G-based m-Health system designed for different emergency scenarios. Medical specialists in the hospital take part in a multipoint conference with the personnel of an ambulance, receiving compressed and coded biomedical information from a patient, assisting them in a diagnosis prior to his reception.

2. 3G m-Health system architecture

The 3G m-Health system (Fig. 1) has been built using standard off-the-shelf hardware, instead of developing propriety hardware as in [5], uses free software and commercial 3G wireless UMTS cellular data services. In the first stages of its design, user requirements and functional specifications were established in collaboration with medical specialists.

The details of the 3G system architecture are shown in Figs. 2 and 3. The system comprises of the signalling and session control, medical user services

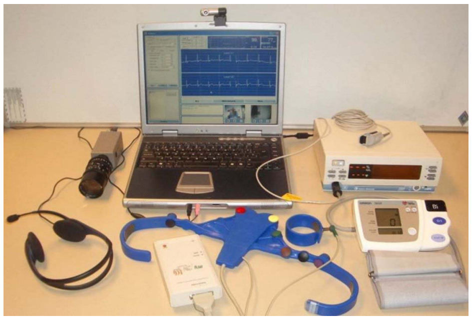

Figure 1. 3G m-Health system

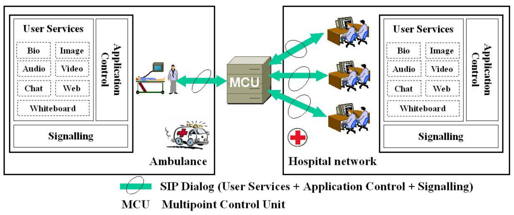

Figure 2. System architecture

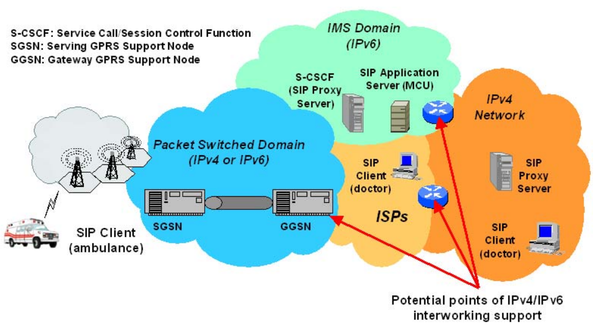

Figure 3. 3G network scenario

and application control subsystems that will be described later together with the QoS control internal subsystem. Several intelligent modules allow the acquisition, treatment, representation and simultaneous media transmission, rather than only one media at a time [6]. Information compression, coding and QoS control (data prioritization, congestion control and dejittering) modules improve transmission efficiency of joint real-time and non real-time data over wireless channels in a more appropriate way than previous systems [7]. In addition, and also unlike [7], this system follows a multi-collaborative design, integrates new real-time multimedia features intended for 3G wireless networks, supports IPv4/IPv6 interworking and uses SIP as the service control protocol, including messages defined specifically for the IMS by the 3rd Generation Partnership Project (3GPP). The IPv4/IPv6 SIP dual stack is the basis to integrate the ambulance and the hospital in any possible 3G scenario (Fig. 3): IPv4 ambulance connecting to an IPv6 hospital (first stages of IPv6 deployment in 3G networks), IPv6 communication through IPv4 islands and, finally, native all-IPv6 communication between them.

Communication between the remote ambulance personnel and medical specialists is established by means of multipoint multi-collaborative sessions through several network environments capable of supporting the different types of multimedia traffic (Fig. 3). The conference model selected is the *tightly coupled conference model* [8], that requires the existence of a Multipoint Control Unit (MCU). System users and the MCU exchange information associated with the different medical user services and its presentation (application control), and also communication and service quality management data (signalling and QoS control).

2.1. Signalling and session control subsystem

The developed signalling allows the exchange of the characteristics associated to the different information flows between system elements and is based on standard protocols that favour interoperability. Signalling tasks, performed by the SIP protocol, begin with the establishment of a SIP dialog with the MCU in which, by means of Session Description Protocol (SDP) messages, the different services are described.

2.2. Wireless medical user services subsystem

The medical user services in the m-Health system are associated with information shared in a multi-collaborative environment. Specifically, the system has services to share audio, ambient video, medical data information (ECG, blood pressure, heart rate and oxygen saturation), high-resolution medical still images, chat, electronic whiteboard and a web service to access clinical information databases. Each kind of information is associated with a medical user service and uses a transport protocol and a codec according to its characteristics (Table 1). Hence, real-time services (audio, video and medical data) use the Real-Time Transport Protocol (RTP), whereas the rest of the services use the Transmission Control Protocol (TCP).

Table 1. Codec operation modes

	Codec	Codec Rate
Medical data	WT[a]	5 10 (Kbps)
Audio	AMR[b]	4.75 5.15 5.9 6.7 7.4 7.95 10.2 12.2 (Kbps)
Video	H.263	5 10 (Frames per second)

a. Wavelet Transform b. Adaptive Multi-Rate

2.3. QoS control subsystem

The QoS in this system is mainly determined by the fluctuant characteristics of the UMTS link, therefore a QoS control process is required in order to obtain a good network performance. This process uses IP packet transfer performance metrics recommended by the International Telecommunication Union (ITU) in its Recommendation Y.1540 [9]. The QoS metrics selected are packet loss ratio, delay variation (jitter) and octect-based IP packet throughput (bandwidth).

The first measure adopted to optimize the use of shared resources among all medical services has been to prioritize them. The most priority services are the medical data, audio and ambient video (in this order), that will take up most of the channel resources. The characteristics of these services are monitored at transmission and reception and are taken into account to increase or decrease codec rates by sending SIP/SDP messages. Non real-time services, on the other hand, will be treated best-effort, adapting to the spare network resources using TCP built-in mechanisms.

One of the main causes of poor QoS is packet loss. As the commercial UMTS 3G wireless cellular service used in this system operates in the UMTS Acknowledged Data Transfer mode at the radio link layer, packet losses are considered to be produced only by network congestion. Part of the QoS control process is based on congestion control, whose signalling can be implicit, activated by packet drops in transmission queues, or explicit, initiated in reception. The congestion control algorithm selected is that presented in [7], but considering real-time service priorities.

Regarding jitter, it can be caused by the variable nature of wireless links or by the joint transmission of all services. Each real-time service has an application dejitter buffer associated with it that has been properly dimensioned to minimize jitter, delay and packets dropped. They are First In-First Out (FIFO) buffers, but considering that if the buffer becomes empty, it does not serve packets to the application until a predefined buffer occupancy threshold is reached. From that point on, the buffer serves packets until it becomes empty. If, on the other hand, the buffer fills completely and a new packet arrives, the first packet stored in the buffer (the oldest one) is dropped.

2.4. Application control subsystem

The MCU forwards the information generated by each medical service according to its associated presentation space defined using the control service. A presentation space defines the way in which the information has to be transferred and its destination. The MCU forwards the information it receives, but first mixes all the audio signals in order to transfer a unique signal to each user. It also forwards only one video signal to each conference participant (selectable by using the control service). Medical data are only generated in the ambulance for the rest of the users.

3. 3G m-Health system performance

The 3G m-Health system has been tested in order to measure its performance and to improve QoS by dimensioning dejitter buffers. Several tests have been carried out in urban scenarios (good coverage level and low speed, as well as static vehicles) using the system over 64/128 Kbps (Uplink/Downlink) IPv4 UMTS accesses (Fig. 4). As the uplink is more restrictive, the results presented correspond to this connection sense.

Packets have been captured in points A to obtain the characteristics of the traffic injected in the uplink (IP-level bandwidth and jitter), whereas measurements in point B allow to obtain network behaviour (packet loss rate and jitter). Several tests have been carried out during various weeks, staggered along the whole day, all days of the week. The duration of these tests has been selected according to the average service time of an ambulance that has been considered to be 10 minutes in a medium-sized Spanish city like Zaragoza.

3.1. Average bandwidth results

Table 2 presents the results about average IP-level bandwidth used by real-time medical services in point A. 48 tests have been carried out every 30 minutes during one day, using isolated user services and varying codec operation modes.

As it can be observed, considering more audio samples per network packet reduces bandwidth use, since transmission efficiency (information carried by each packet to total packet size ratio) is increased. However, there is a limit in the number of audio samples per packet that can be used, because more audio samples per packet yield more audio delay.

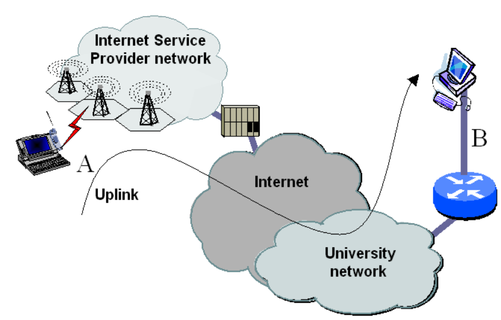

Figure 4. Measurement scenario

Table 2. Average IP-level bandwidth use

	Operation Mode		IP Bandwidth (Kbps)
Audio	*Samples/Packet*	*Codec Rate (Kbps)*	
	1	4,75	21,2
	1	12,2	28,8
	3	4,75	10,5
	3	**12,2**	**18,1**
Ambient video	*Frames Per Second*		
	5		16
	10		**24**
Medical data	*Bit Rate (Kbps)*		
	5		5,3
	10		**10,3**

Moreover, a reduced number of audio samples per packet is more suitable to error-prone environments. Regarding the video service, it is worth noting that its bandwidth use can vary substantially with the movement of the scene captured. Finally, the medical data service adapts well to the codec rate specified because medical data frame sizes are long enough to obtain a good transmission efficiency.

The total bandwidth consumed by all real-time medical user services fits in a 64 Kbps UMTS channel, even when the most bandwidth-consuming codec rates and the lowest transmission efficiencies are used. Thus, the initial codec operation modes selected have been those highlighted in Table 2, achieving a reasonable trade-off between bandwidth, transmission efficiency, delay and loss ratio. During normal operation, codec modes can vary in response to congestion conditions with the aid of congestion control mechanisms.

The average IP-level bandwidth obtained in point B is very similar to that obtained in A. In addition, no packet losses have been observed in any points, therefore the network does not modify traffic characteristics regarding bandwidth and packet loss.

3.2. Jitter results

48 tests have been carried out every hour during 2 days, with all real-time medical services operating at the same time and at the highest codec rate. These tests are useful for observing the influences between traffics generated by each real-time service.

As all possible jitter effects can be observed in point B, Fig. 5 presents a zoom over 9 seconds of audio interpacket time taken in a test in point B. Audio and medical data packets are generated every 0.06 and 1 second, respectively, therefore jitter effects over the audio service appear uniformly spaced. They cause more than 140 ms of jitter due to the time it takes the UMTS uplink channel to transmit big-sized medical data packets (around 1300 bytes at IP level). Regarding

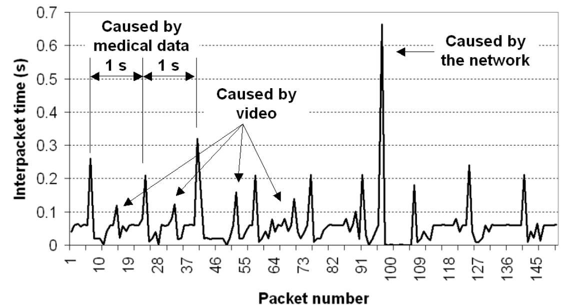

Figure 5. Audio service interpacket time

video, packets are smaller and not uniformly spaced, so the effects of video over audio packets are smaller. In addition, other jitter effects are caused by the network.

The audio and video services do not have a significant influence over the medical data service due to the fact that medical data packets are very spaced between them (1 second, ideally), and a small jitter effect is not noticed in reception with the aid of a minimal dejitter buffer. The last real-time service, the ambient video, also suffers jitter effects caused by the rest of real-time services. However, the most serious effects are produced by the codec used and big reception buffers are recommended. In addition, this service has the lowest priority, so these jitter results are not presented here. In short, a big dejitter buffer is enough to support all possible jitter effects.

3.3. Dejitter buffer dimensioning

Packet captures in point B are useful to obtain the instantaneous application dejitter buffer occupancy for all the real-time services. Using packet timestamps and theoretical packet buffer empty times from the same tests of subsection 3.2., various buffer occupancy calculations varying buffer size and threshold have been carried out in order to dimension dejjiter buffers.

The buffer threshold value must be able to support jitter effects caused by real-time services, whereas the total buffer size value must be able to support jitter effects caused by the network. High values of the threshold cause less situations in which the buffer becomes empty (playback interruptions), but, on the other hand, introduce a bigger fixed delay. In the same manner, bigger buffers allow less packet drops, but also entail bigger delays to enqueued packets.

The most important jitter effects not caused by the network over the audio service are produced by medical data packets. Jitter values rarely exceed 200 ms, so a threshold of 4 packets (240 ms of audio) is enough. Due to the small jitter effects suffered by the medical data service, its buffer threshold can be very small (2 packets). In order to ensure a more robust

Table 3. Average audio packet drop ratio vs. delay (threshold=4)

Buffer size (packets)	Average packet drop ratio (%)	Typical deviation	Fixed delay (ms)	Maximum delay (ms)
4	5.059	0.441	240	240
5	2.113	0.237	240	300
6	1.137	0.198	240	360
7	**0.619**	**0.153**	**240**	**420**
8	0.321	0.098	240	480
9	0.137	0.063	240	540
10	0.060	0.039	240	600
11	0.028	0.019	240	660
12	0.010	0.010	240	720
13	0.004	0.006	240	780
14	0.001	0.003	240	840
15	0.000	0.000	240	900

behaviour, and considering that fixed delay is not relevant, 3 packets have been selected. Finally, as the ambient video service is not critical, the buffer threshold selected has been the minimum (1 packet).

The total buffer size must be able to support jitter effects caused by the network. The first step is to consider an infinite buffer with all the 10-minute tests carried out to calculate the maximum buffer size that would have been needed in order not to drop any packet. The same buffer size has been obtained in all medical data tests (3 packets), therefore a conservative value of 4 packets is a good choice in order to ensure a robust behaviour. Regarding the video service, the maximum buffer occupancy obtained in all the tests has been 25 packets. A more conservative value of 30 packets has been selected to support even worse working conditions. Finally, audio tests present variability in the maximum buffer size. In order to obtain a suitable buffer size, finite buffer sizes have been considered, producing different packet drop ratios. The packet drop ratio has been averaged for all tests and the numeric results are shown in Table 3. Not only packet drop ratio, but also the maximum delay that a particular buffer size causes are the relevant parameters in order to select buffer size. If, for example, a packet drop ratio of less than 1% is desired, 7 packets would be choice.

4. Conclusions

This paper has presented a feasible 3G-based m-Health system targeted specifically for emergency medical scenarios that can also be used in the home telecare area. The system architecture is based on 3G networks and advanced signalling protocols (SIP/SDP) that allow the integration of real-time multimedia services over multiple access channels that support IPv4/IPv6 interworking depending on current commercial UMTS releases. The system has the following features: simultaneous transmission of real-time clinical data, videoconference and other non real-time medical services. The system has been optimized specifically to operate over 3G mobile networks using the most appropriate codecs. Evaluation results show a reliable performance over IPv4 UMTS accesses (64 Kbps in the uplink) and also allow to dimension dejitter buffers to improve the quality of real-time services.

5. Acknowledgment

This work has been supported by Telefónica Móviles de España, projects TEC 2004-04529/TCM and TSI2004-04940-C02-01 from Comisión Interministerial de Ciencia y Tecnología and European Regional Development Fund, project FIS G03/117 from Fondo de Investigación Sanitaria and FPU grant AP-2004-3568 from Secretaría de Estado de Universidades e Investigación.

6. References

[1] R.S.H. Istepanian et al., “Wavelet Compression Methodologies for Efficient Medical Data Transmission in Wireless Telemedicine Systems”, Journal of Telemedicine and Telecare, Vol. 7, No. 1, pp. 14-16, 2001.

[2] R.S.H. Istepanian, B. Woodward, and C.I. Richards, “Advances in telemedicine using mobile communications”, Proc. IEEE Engineering Medicine and Biology Society, Vol. 4, pp. 3556–3558, 2001.

[3] R.S.H. Istepanian, S. Laxminarayan, and C.S. Pattichis, “M-Health: Emerging Mobile Health Systems”, New York, Springer, 2005. To be published.

[4] 3GPP TS 23.228 V6.8.0., “IP Multimedia Subsystem (IMS); Stage 2”, Release 6, 2005.

[5] J. Cullen et al., “Wireless mobile telemedicine: En-route transmission with dynamic quality-of-service management”, National Library of Medicine Symposium on Telemedicine and Telecommunications: Options for the New Century, 2001.

[6] S. Pavlopoulos et al., “A novel emergency telemedicine system based on wireless communication technology – AMBULANCE”, IEEE Trans. Inform. Technol. Biomed, Vol. 2, pp. 261-267, 1998.

[7] Y. Chu, and A. Ganz, “A Mobile Teletrauma System Using 3G Networks”, IEEE Trans. Information Technology in Biomedicine, Vol. 8, No. 4, pp. 456-462, 2004.

[8] J. Rosenberg, “A Framework for Conferencing with the Session Initiation Protocol”, Internet draft, 2004. Work in progress.

[9] ITU-T Rec. Y.1540, “IP Packet Transfer and Availability Performance Parameters”, Dec. 2002.