

# Towards High Quality VoIP in 3G Networks an Empirical Approach

Andres ARJONA<sup>1</sup>, Cedric WESTPHAL<sup>2</sup>, Antti YLÄ-JÄÄSKI<sup>3</sup>,  
Martin KRISTENSSON<sup>1</sup>, Jukka MANNER<sup>3</sup>

<sup>1</sup> Nokia Siemens Networks, Finland, <sup>2</sup> DoCoMo Labs, USA, <sup>3</sup> Helsinki University of Technology, Finland

Email: {andres.arjona, martin.kristensson}@nsn.com, cwestphal@docomolabs-usa.com,

{antti.yla-jaaski, jukka.manner}@hut.fi

Received August 22, 2008; revised October 10, 2008; accepted October 12, 2008

## Abstract

Third generation (3G) packet switched WCDMA networks with high-speed downlink packet access (HSDPA) are currently being deployed worldwide to provide wireless broadband connectivity. When introducing HSDPA in 3G networks the end user experience and system capacity with voice over IP applications improve considerably. When later on adding also high-speed packet uplink access (HSUPA), the system capacity and end user experience will improve even further. This paper analyzes with measurements the VoIP quality over current Release 5 HSDPA networks. VoIP is expected to be a widely used application over 3G data services. The results show that even though the introduction of HSDPA significantly reduces the user-to-user voice delay, the performance is satisfactory only for selected devices. Overall, the end user experience is still significantly worse than with circuit switched solutions and is not acceptable. The current limitations with VoIP in HSDPA networks with a too large delay can be improved by using the RLC UNACK mode, potentially decreasing the jitter buffer size and reducing the terminal processing delay. In the longer term, HSUPA and several features in 3GPP Release 7 standards will bring further performance improvements in both user plan latency and system capacity.

**Keywords:** HSDPA, VoIP, WCDMA, Voice Quality, MOS

## 1. Introduction

Voice over IP (VoIP) is becoming a widely deployed service in data networks, and it will penetrate from the fixed network domain into wireless network domain. The characteristics of fixed networks and wireless networks are fundamentally different, which will impact the performance of services. In this article we analyze the VoIP service performance in wireless HSDPA and WCDMA networks.

High Speed Download Packet Access (HSDPA) [1] networks are being intensively deployed to provide broadband connectivity to mobile devices, such as handheld terminals and laptops. This broadband wireless access is able to support voice applications over a packet data connection instead of traditional circuit switched calls. With the introduction of multi-radio devices with HSDPA,

WCDMA, and WiFi capabilities as well as integrated VoIP clients, ubiquitous connectivity across any of these networks is possible using the same mobile terminal. However, while the mobile terminal and client are the same, performance differs depending on the wireless access in use.

Most of the studies of VoIP over 3G network focus on simulation works. However, there is little data on the performance in actual networks. Since VoIP is expected to become a widely used application, and comes pre-configured in many current handsets, it is of great importance to better understand the performance of such application over 3G networks. We set to answer the following question: is VoIP over 3G network commercially viable with the current state of the arts networks?

This paper studies the quality of VoIP in wireless networks with multi-radio mobile devices both in the lab and in live network environment setups by conducting a

methodic performance analysis based on the E-Model [2,3](we will describe the E-Model in more details in Section 2). Likewise, our study will encompass the signaling performance required for VoIP applications.

The key contribution of the paper is to characterize the performance of VoIP over 3G network, and to identify the main differences between HSDPA and WCDMA. We perform a thorough empirical evaluation of VoIP quality and signaling performance with HSDPA and WCDMA. From our evaluation, we will observe that:

- VoIP performance is acceptable in HSDPA networks only for VoIP clients on devices with enough processing power, such as laptops;
- VoIP performance is rarely acceptable in WCDMA networks, even for those high performance clients;
- WCDMA performance can be significantly improved by having retransmissions only at the BTS, not the RNC;
- The delay introduced by the end-user terminal is a critical factor in the performance.

Our study takes into consideration both the performance of the network and also the performance of real embedded VoIP clients. In addition, we validate the results of our study by comparing them to the actual performance in a densely deployed HSDPA network in Finland. Based on the results, we analyze the primary differences in performance between simulations found in the literature, our lab experiences and a live network case study. Finally, we discuss possible features that can improve the performance enough in current and future releases to support VoIP in all handheld devices.

The remainder of the paper is organized as follows. Section 2 describes our research approach, Section 3, 4 and 5 present results from a laboratory setup, a live network scenario, and for VoIP signaling performance respectively. Subsequently, in Section 6 we describe some standardization improvements. In Section 7 we discuss the available related works and finally in Section 8 we draw conclusions.

## 2. Methodology and Test Environments

Our experiments are composed of measurements in a HSDPA and WCDMA testbed, as well as a live HSDPA network of a Finnish operator. We are interested in measuring both the VoIP service audio quality in both laboratory and live setups and SIP signaling latencies for registering users and setting up calls.

### 2.1. VoIP Quality Methodology

The evaluation methodology consisted of multiple VoIP tests carried out in a radio interference free environment. These conditions were achieved in a laboratory setup by using an RF room for the BTS and clients [4]. The tests

included different wireless access technologies and variable combinations of codecs, signal conditions, number of clients and fading profiles among others.

The main evaluation was carried out with two similar tools based on the E-Model [2], which is a ITU-T recommendation for VoIP evaluation. Firstly, with a NSN proprietary tool, which is an implementation similar to the one described in [3], and secondly, with IxChariot, which is a widely used voice evaluation tool [5]. Finally, a third tool based on the PESQ evaluation model was used to determine the average end-to-end delay with real embedded VoIP clients. With such setup, we can evaluate the performance of the different wireless access technologies based on the following test objectives:

- VoIP quality performance with the E-Model;
- Voice quality characterization for different wireless accesses, signal conditions, configurations and fading profiles;
- Benchmark of two voice quality evaluation tools based on the E-Model;
- Estimation of the average end-to-end delay when a real embedded VoIP client is used;
- Effect of simultaneous background traffic during a VoIP call;
- Characterization of delay sources and possible optimizations.

The E-Model is a voice quality evaluation model that is based on network performance metrics. It is based on a mathematical algorithm and provides an “R” performance value based on the sum of four “impairment factors” considered to be cumulative. The algorithm is depicted in Equation (1) where, “ $I_s$ ” is Signal to Noise Ratio, “ $I_d$ ” is delay (ms), “ $I_{ef}$ ” is packet loss (%), and “ $A$ ” is expectation factor.

$$R = 100 - I_s - I_d - I_{ef} + A \quad (1)$$

In practice, ITU-T proposes to use a simplified version of this algorithm. The simplified algorithm considers that noise cancellation is encountered in the network and also dismisses the expectation factor. The expectation variable is supposed to be used to provide a balance for some environments in which the user expects a degraded quality, such as satellite connections. However, since this variable is merely subjective it is recommended to ignore it. The simplified algorithm is depicted in Equation (2).

$$R = 93.2 - I_d - I_{ef} \quad (2)$$

The R value can be associated with the Mean Opinion Score (MOS) values, which is a subjective grade for voice quality based on studies carried out by ITU-T. However, even though the R-value can match a MOS value, it cannot predict the absolute opinion of an individual user.

In this paper we calculate the MOS scores with two tools based on the E-Model: a Nokia proprietary tool and IxChariot, which is a widely used tool. These tools send

dummy packets that resemble VoIP packets. The packet size and transmission intervals are tied to the modeled codec. Based on the received packets, network performance values are calculated and the E-Model algorithm is applied to determine a MOS score. Figure 1 shows an overview of the environment and the E-Model based tools.

In this paper we emphasize the performance of the G.729 codec, which is the only codec supported by all the measuring tools used in this research. G.729 is also similar to AMR-NB. AMR codec is the main building block for a future codec for 3GPP based networks. ITU-T has set out standards for maximum voice quality for several codecs, including G.729 and G.711. However, there is not yet agreement on a standard AMR codec maximum quality definition in relation to the E-Model. Therefore, we can make the assumption that the performance values measured with G.729 codec are representative and are a useful basis for our analysis. In addition, G.711 is not an appropriate codec for wireless networks such as HSDPA due to its high bitrate. However, G.711 is one of the most largely supported codecs, and it is widely used in the Internet. Also, due to legacy equipment it is used in many cases, even though AMR and other lower bitrate codecs (e.g. iLBC) are encouraged. For this reason we study both G.729 and G.711.

## 2.2. VoIP Signaling Performance Methodology

The evaluation methodology consisted of a variety of VoIP calls using Nokia N95 terminals. We chose this terminal due to its widespread penetration in the market and because it includes an embedded VoIP client by default. This client can also be configured to work with other SIP systems (e.g. Gizmo project). We did not use a 3rd party implementation with Skype because there were no suitable clients for the N95 at the time of the study. We captured SIP packet traces directly from the mobile terminal wireless interface [6]. With such variables we evaluated the different wireless networks available from the following test objectives:

- SIP registration delays
- VoIP call signaling delays (post-dial, answer-signal, and call-release delays)

The two main activities in VoIP calls are: first, a registration to the VoIP server which is required to make and receive calls, and second, the voice call setup itself. The packet captures were carried out with a NSN proprie-

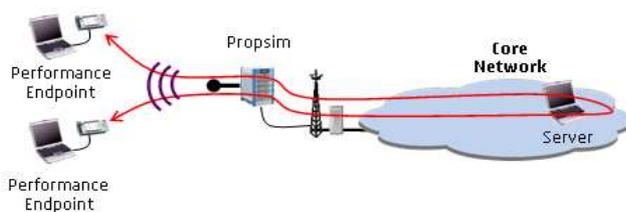


Figure 1. VoIP quality test environment.

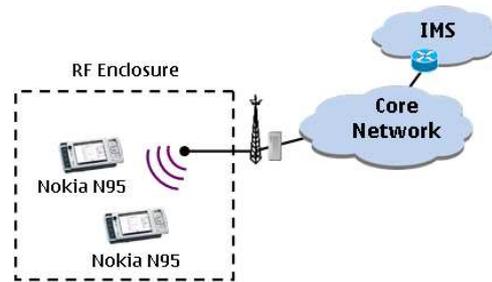


Figure 2. VoIP signaling test environment.

tary tool with a function similar to TCPdump, and analyzed with Wireshark Protocol Analyzer [7].

All the calls were carried out with two identical terminals with exactly the same setups, registered to the same VoIP server in the NSN IP Multimedia Subsystem (IMS) and via the same wireless access in an interference free environment. The measured scenarios were HSDPA-to-HSDPA, WCDMA-to-WCDMA, and WiFi-to-WiFi calls. The maximum transfer bitrates were set in the RNC and HLR configurations to model different wireless access scenarios. For WCDMA, maximum uplink and downlink transfer rates were fixed at 64/64 kbps and 128/128 kbps. For HSDPA downlink was 3.6Mbps and the uplink was fixed at 128 kbps. In the case of WiFi, transfer rates were left with default configuration (802.11g and maximum transfer rate). Figure 2 shows the test environment setup.

The core network and IMS system were privately owned and under very low load. The wireless access systems were based on NSN Release 5 equipment for HSDPA and WCDMA tests with default settings. For WiFi tests, we used a Belkin Pre-N Router with default configuration. The core network and IMS system were based on Nokia equipment. The tests executed consisted of multiple iterations of each of the voice call scenarios and registration to the VoIP server. We provide average result values from the measurements. The measurements took place during February-March 2007.

## 2.3. Live Network Case Study Methodology

The final stage of our study consisted of evaluating voice quality in a live HSDPA network. The network evaluation took place in the Helsinki metropolitan area, and the live network in use was provided by Elisa, Finland's largest 3G operator. The HSDPA coverage in the Helsinki metropolitan area is densely deployed and assumed to be based on NSN equipment similar to the one used in our lab measurements (Release 5 equipment). Therefore, its performance is directly comparable to our previous results. The test objectives for this phase are as follows:

- Characterize the base performance of the network (throughput and round trip time) under different signal conditions
- Evaluate the VoIP quality in different signal conditions

(excellent, medium and poor)

- Evaluate the average VoIP quality in a mobile scenario and determine the signal quality distribution for the test route

Our approach to the live network measurements was modeled in the following way. First, we made a basic network performance evaluation in different radio environments based on signal-to-noise ratio  $E_c/N_0$  levels [4].

$E_c/N_0$  values are an objective figure for quality conditions because they take into account both signal strength and the current interference level encountered in the cell. Based on these basic network performance figures, we can evaluate the average performance in terms of maximum downlink and uplink throughput, as well as average round trip time for a particular  $E_c/N_0$  range. As a result, we are able to define three signal conditions ranges: 1) Good signal 2) Medium signal, and 3) Poor signal. Second, we evaluate the VoIP quality with the same NSN Proprietary tool used in previous tests under the three different signal conditions. This allows us to get a good metric of what is the quality in a static scenario under specific signal conditions. Third, we evaluate the average VoIP quality under a mobile scenario. The test route chosen crossed a major part of the Helsinki metropolitan area from West to East. The tests were carried out along the route in both directions twice. In addition, we measured the signal levels ( $E_c/N_0$ ) along the whole driving route and carry out statistical distributions for the values.

An obvious limitation of our study is the fact that due to the nature of a live network, we are not able to know or control the other user traffic that could be taking place at the same time. Therefore, we are not able to pinpoint the sources of e.g. a sudden quality drop or reduced bitrate. However, since we carried out multiple tests, our study provides a realistic view of what is the actual performance that could potentially be achieved in the field. The measurements for the live network study took place during July and August 2007.

### 3. VoIP Quality Analysis and Results

#### 3.1. HSDPA/WCDMA VoIP Performance

The tests to evaluate VoIP quality involved the following variables: signal conditions, wireless access, and fading profiles. Signal conditions were modeled to provide different  $E_c/N_0$  levels by using attenuators. However, the results in this paper show that this variable does not make any sustainable difference and therefore, average result values are given instead. The wireless access technologies used were restricted to HSDPA/128, WCDMA 128/128 and WCDMA 64/64. There was no reason to use higher bitrates in this study since VoIP packets require a low bandwidth. Therefore, we emphasize the limits in which

VoIP can actually be used with an adequate quality level. The bitrates were fixed and therefore, features that adapt bitrate (by increasing or decreasing) during packet switched connections were not used during the tests. Fading was applied with Prosim C8 fading simulator using Pedestrian-A 3km and Vehicular-A 30km fading profiles. The jitter buffer had a depth of 200ms and first packet play delay of 120ms. That is, all packets are delayed at least 120ms to provide a cushion for possible jitter. These are common settings in VoIP clients for wireless cellular systems. According to Wang et al. [8], a conservative jitter buffer playout delay is about 150ms.

Our results are consistent and show that the achieved quality in the HSDPA system is competitive. Based on ITU-T G.107 [2] quality was in average medium for HSDPA with both measurement tools (NSN Proprietary Tool and IxChariot). The average MOS was roughly 3.7 (see Figure 3 and Figure 4). This is a good figure especially considering that typical PSTN systems provide MOS values around 3.5. In the case of WCDMA, quality differed depending on the bitrate used. WCDMA 128/128 provided low quality and WCDMA 64/64 gave a low/poor quality level. The results also show a difference between the measurement tools. Our proprietary tool was able to differentiate more clearly the quality levels between WCDMA 128/128 and 64/64. However, IxChariot does not recognize much difference between these two bitrates. In any case, both tools show that quality in WCDMA is not optimal and is around MOS 3.0 at its best. WCDMA 64/64 MOS varied between 2.25 and 2.7. ITU states that MOS below 2.5 is not recommended for voice services and that nearly all users will be dissatisfied with such a service. Therefore, we can expect that the end user experience with VoIP WCDMA is not stable and will vary.

Table 1 presents the average end-to-end delays in the experiments (including jitter buffer playout delay). The results also show very similar performance regardless of the signal conditions modeled or the fading profile applied. The reason probably relates to fast power control mechanisms which are able to handle such changes in signal conditions in HSDPA and WCDMA. We recommend that further studies would be performed using noise or traffic generators instead of only modeling signal scenarios with attenuators.

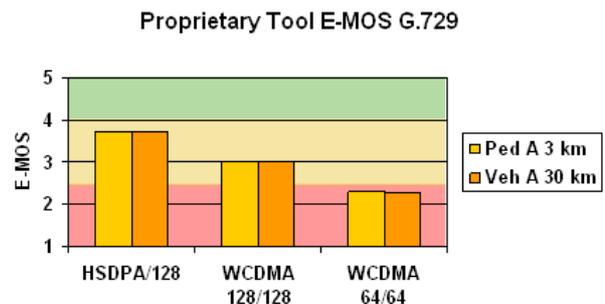


Figure 3. VoIP performance evaluation with proprietary tool.

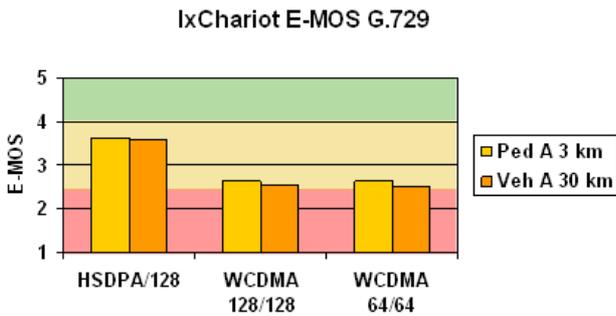


Figure 4. VoIP performance evaluation with IxChariot.

Table 1. Average VoIP end-to-end delays (including jitter buffer).

Access	Proprietary Tool		IxChariot	
	PedA 3km	VehA 30km	PedA 3km	VehA 30km
HSDPA/128	215ms	217ms	223ms	225ms
WCDMA 128/128	295ms	300ms	368ms	381ms
WCDMA 64/64	315ms	355ms	370ms	365ms

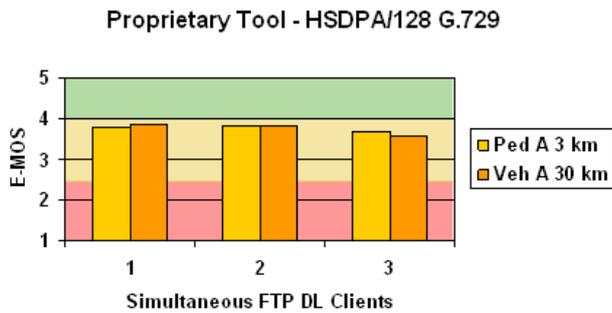


Figure 5. Effect of background traffic.

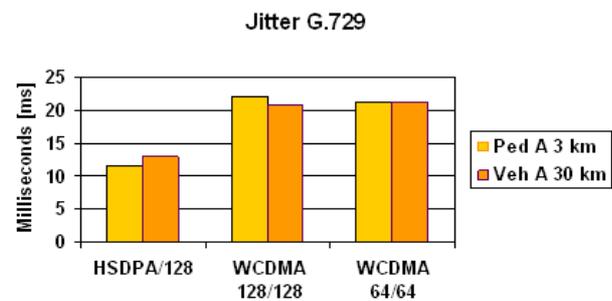


Figure 6. Jitter average.

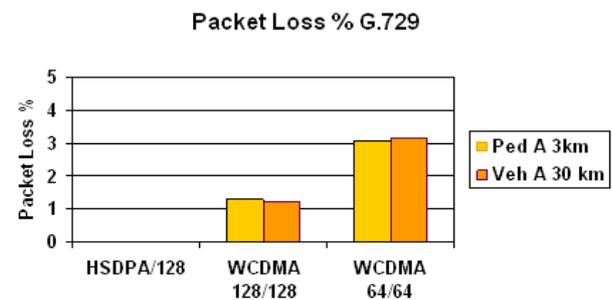


Figure 7. Packet loss percentage.

Finally, we point out that both measurement tools yield quite similar results, with exception of WCDMA 64/64. However, in this case we can observe that our proprietary tool is actually more accurate than IxChariot, especially since IxChariot does not seem to recognize any performance difference between WCDMA 128/128 and 64/64 accesses.

### 3.2. VoIP Performance with Simultaneous FTP Background Traffic

We also conducted some experiments where we added background traffic. The tests included a small number of simultaneous users running FTP downloads in order to evaluate if they had any effect on the VoIP performance.

As we expected, a limited number of users cannot affect VoIP quality (see Figure 5). The reason is tied to the Round Robin Scheduling used in the system, which divides bandwidth equally among users. With only 4 simultaneous users, each user will be given enough bandwidth on a timely basis (every few milliseconds). In order to measure the effect of background traffic we encourage tests with a much larger number of users, e.g. 15-20 would be required. This is out of the scope of this document. Likewise, testing different scheduling techniques such as Weighted Proportional Fair is of interest. However, there are several simulation based studies [9,10] that study VoIP capacity gains for different scheduling schemes including mixed traffic scenarios. However, note that [11] analytically showed that QoS constraints on VoIP reduce the benefit from the Proportional Fair algorithm over Round Robin scheduling.

### 3.3. Effect of Jitter and Packet Loss

The next test included experiments with jitter and packet loss. Jitter and packet loss are presented in Figure 6 and Figure 7. From the results we can see the average jitter and packet loss measures for different access networks.

The results show an increased jitter and packet loss for WCDMA 128/128 and 64/64. Further delay analysis shows that this increase is most likely caused by constant RLC retransmissions. RLC retransmissions have an effect on both jitter and packet loss. Every time a RLC retransmission takes place, it will cause a ~200ms delay peak. This peak can potentially fill the jitter buffer causing an overflow, which results in packet loss. Packet loss also affects voice quality. The frequency of RLC retransmissions is dependent of the access in use. Figure 8 shows an example of the RLC retransmissions (200ms delay peaks) for different wireless access technologies.

The performance of these wireless accesses would improve if RLC retransmissions are avoided as much as possible. One possibility is to use the unacknowledged mode (UNACK) feature in the RNC. The principle of operation in HSDPA [1] is such, that the BTS estimates

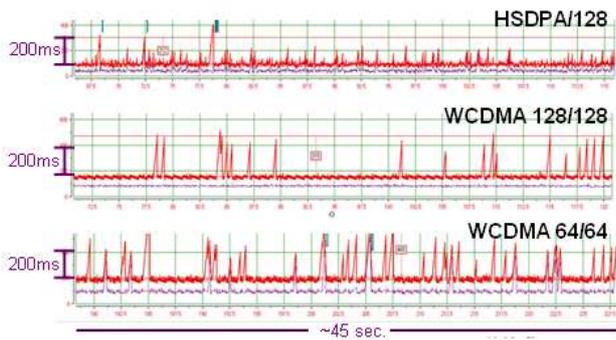


Figure 8. RLC retransmissions.

the channel quality of each user based on the physical layer feedback on the uplink. Subsequently, link adaptation and scheduling takes place at a fast pace. When the packets are first received at the BTS, they are buffered. Then, the BTS transmits the packet; however, it will still keep it in the buffer. The reason being that in case of a failure in the transmission (e.g. decoding failure), a retransmission will take place directly from the BTS without requiring any action from the RNC. This is a powerful advantage since the retransmissions are combined at the terminal. However, if there is a physical layer failure, such as a signaling error, then an RLC retransmission is required, and packets are retransmitted from the RNC (see Figure 9). This obviously results in an increase in delay, which is not beneficial for services like VoIP. While RLC retransmissions are not a very frequent event in HSDPA in static scenarios, they are more likely in mobility scenarios. In contrast, in WCDMA, all retransmissions are RLC retransmissions requiring RNC involvement. In the RLC unacknowledged mode, packets are not retransmitted even if some are lost, for example due to cell change operation [1].

### 3.4. Codec Performance Evaluation

Even though our study focus was on low bit rate codecs (e.g. AMR or G.729), we also evaluated the performance of the G.711 codec. Using G.711 codec in wireless environments is not encouraged due to its higher bit rate. However, since it is one of the most widely supported codecs, there are cases in which it will be used due to other codec incompatibilities. The performance was measured with a proprietary tool. Tests with WCDMA 64/64 using G.711 failed most of the time or resulted in very long delays of several seconds and are therefore excluded.

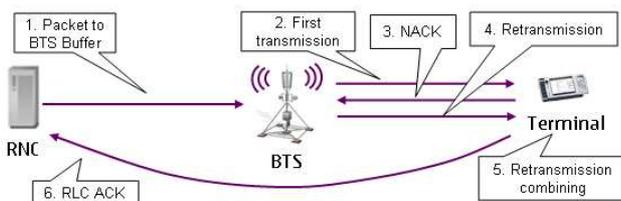


Figure 9. BTS retransmissions handling.

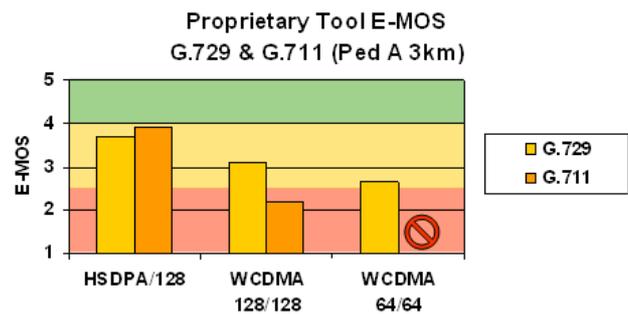


Figure 10. Codec performance evaluation (G.729 and G.711).

Table 2. G.711 codec jitter and packet loss (PedA 3km).

Access	Jitter Average	Packet Loss %
HSDPA/128	13ms	0.34
WCDMA 128/128	19ms	2.78

Figure 10 shows the VoIP quality comparison for both G.729 and G.711 codecs. Table 2 summarizes the jitter average and packet loss encountered when using the G.711 codec.

### 3.5. Embedded VoIP Client Evaluation

These tests aimed at determining the additional delay resulting from real embedded VoIP clients, such as the one included with the N95. The test setup consisted of establishing a VoIP call using an IMS system with the G.729 codec. Subsequently, we measured the offset delay, that is, the delay between the moment when the original audio sample occurs to the moment the audio sample is reproduced in the other calling end. The tool used for offset measurements was Malden DSLA [12]. Figure 11 shows the measurement environment. The results show that the total offset delay including the VoIP client processing delay is rather high (see Table 3). ITU-T recommends 400ms as the maximum delay for voice services with a reasonable quality. With delay above this limit, conversations are not interactive anymore and result in talker overlaps. Therefore, a voice service with very high delays results in a situation in which most, if not all users are dissatisfied. As a comparison, current circuit switched voice services have a delay of roughly 230–250ms.

With the results we can estimate the client processing delay by subtracting the average end-to-end delay from our tests based on the E-Model, 215ms, 295ms, and 350ms respectively. The result is roughly 210ms additional processing delay when using a real embedded VoIP client. This value differs considerably from the more optimistic processing delay estimations of 50-75ms available in research from [13,14].

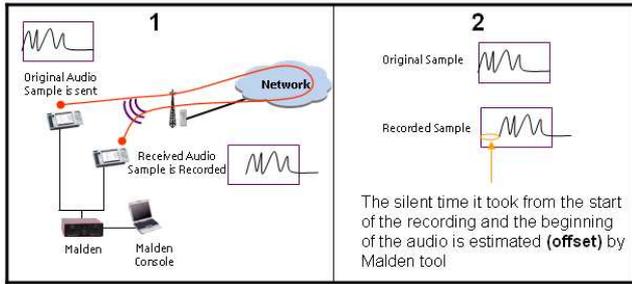


Figure 11. Offset delay measurement environment.

Table 3. Sources of delay (G.729 codec).

	HSDPA /128	WCDMA 128/128	WCDMA 64/64
RTT Delay	85ms	170ms	225ms
Jitter Buffer (100-200ms)	130ms	125ms	125ms
Total E2E Delay	215ms	295ms	350ms
Total E2E Delay, including embedded client delay	425ms	505ms	560ms

### 3.6. HSDPA/WCDMA Overall Effect on VoIP Performance

End-to-End delay is the main reason for low voice quality. With the total end-to-end total delay average values we can extend the analysis by dividing the sources of delay (Table 3).

With this estimation it is quite clear to understand why VoIP does not perform well in current systems with handheld terminals, and particularly live networks, even when the round trip time (RTT) is low. The final end-to-end delay is just too high. We finalize our VoIP quality analysis by modeling the resulting VoIP quality MOS with the additional embedded VoIP client processing delay based on the E-Model (see formula 2). Figure 12 shows this estimation. The results represent a case of a laptop client versus using an embedded client in a handheld device such as the N95 VoIP client. The figure considers both delay and packet loss impairment factors. It must be noted though, that in a laptop client there will also be an additional processing delay. However, such delay is considerably lower, ~50ms in a worst case scenario [15]. Thus, still ~160ms lower than with the mobile device tested.

Future features such as HSUPA in further 3GPP releases will slightly improve performance. For example, the expected average RTT for HSUPA networks is roughly 65ms (a reduction of 20ms compared with HSDPA). This reduction however does not improve the VoIP quality when using a laptop. That is, the average MOS with a laptop will still be the same. Contrastingly, the expected quality improvement for an embedded client is about 0.2

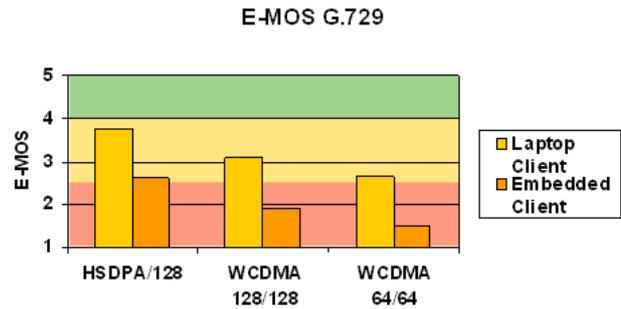


Figure 12. Overall VoIP quality with laptop and embedded handheld clients.

points in the MOS score. The main reason for the limited quality improvement is that the major sources of delay, and therefore, main impairment factors reducing VoIP quality are not directly related only to the wireless access, but to the VoIP client implementation. However, as we described previously, if some HSUPA features like UNACK mode are enabled in the wireless network, it will be possible to reduce the size of the jitter buffer implementation without compromising the VoIP quality. Furthermore, a reduction in the client processing delay is extremely important in order to seriously improve the VoIP quality in the mobile environment.

## 4. VoIP Signaling Analysis and Results

In this section we analyze the latencies for VoIP using the Session Initiation Protocol (SIP). This is an important metric because long delays in the call setup seriously harm the overall VoIP experience; people have certain expectations based on the current circuit switched services, and it is crucial to meet those.

### 4.1. SIP Registration Setup

The signaling [16] and delay measurements for SIP Registration to the VoIP server in the IMS system are depicted in Figure 13. The measurements show that the registration times with HSDPA and WCDMA are about 30% and 50% higher than with WiFi. While this might not seem much, we should remember that SIP registration requires a very limited number of messages. Therefore, as more messages are required, such as with 3GPP based registration, delays will increase.

### 4.2. VoIP Call Signaling

ITU E.721 [17] recommends values for call setup delays in circuit switched calls. The recommended values for call setup (post-dial delay) are 3s for local, 5s for toll and 8s for international connections, with 6s, 8s, and 11s as 95% values. The “call answer” (answer-signal) delay reflects

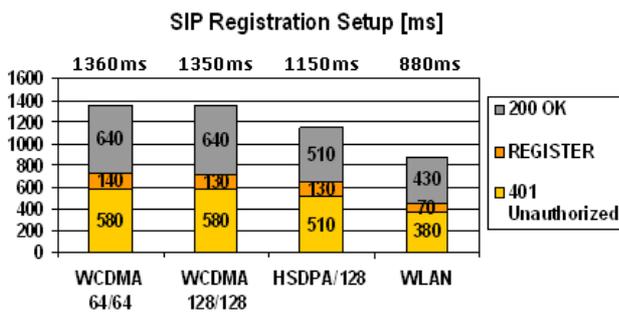


Figure 13. SIP registration signaling delays.

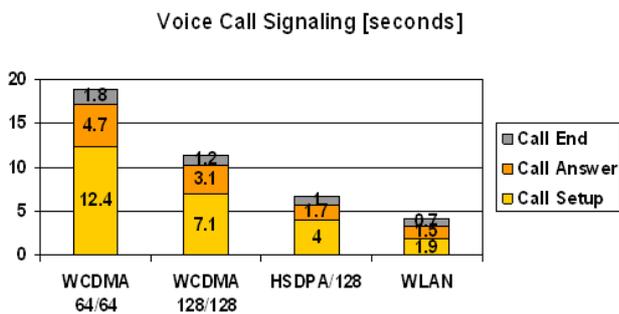


Figure 14. VoIP call signaling delays.

the time it takes from the moment the receiving end accepts the call until the call is actually established. E.721 recommendation is 0.75s for local, 1.5s for toll, and 2.0s for international connections, with 1.5s, 3.0s, and 5.0s as 95% values. Finally, “call end” (call-release delay) is the time it takes for the call to be terminated [18,19]. The signaling [16] and delay measurements for voice call setups when a PDP context is active and the terminal is registered to the IMS system are depicted in Figure 14. The results show an expected increment in the call signaling delays depending on the access used. Since all of the network elements were located in a private network, the environment could be thought of as providing local calls. Our results also show that an embedded mobile VoIP client experiences an increased delay compared to a PC client, such as the one measured by Curcio and Lunden [18] with a WCDMA network.

The setup delays for VoIP calls might be impacted with additional delays in a cellular system in cases where there is no active PDP context, and also due to a required registration to the IMS. The PDP context activation delay was ~3 seconds in our tests. Simulations by Pous et al. [20] propose 2.24 seconds. Based on these values, the always-on enabled calls can be in line with E.721 recommendations. However, when the PDP context is not active, the delay with WCDMA can vary between 11 to 17 seconds, and thus, exceed the recommended values. HSDPA delay in this case is around 8 seconds, which is similar to the recommendation for international calls. However, additional delays from e.g. traversed networks, gateways, and proxies could result in larger total delays than those recommended.

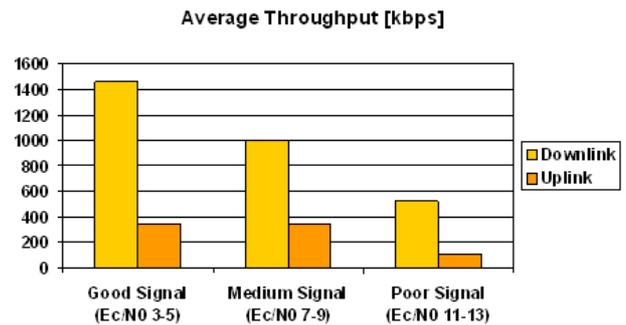


Figure 15. Average throughput in Elisa HSDPA network.

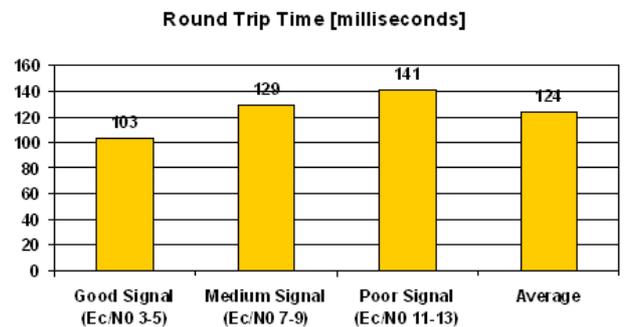


Figure 16. Average round trip time in Elisa HSDPA network.

## 5. Live Network Case Study

### 5.1. Generic HSDPA Performance

In this section we describe the generic evaluation of the live HSDPA (Release 5 equipment) network performance in Helsinki. The results for throughput and round trip time are depicted in Figure 15 and Figure 16. The measurement results show an increase in round trip time delay when compared to the average values measured in the lab environment (85ms). This means that the VoIP quality (MOS) will be worse than our results in Section 3, and therefore VoIP support will be even more difficult. Throughput was measured via multiple file downloads and uploads from a local server in Finland; while RTT was measured with 32Byte ICMP Echo Request and Reply (ping) packets to the same server.

### 5.2. VoIP Quality

The VoIP quality in the Elisa HSDPA network is likewise slightly lower than our lab measurements (see Table 4). The mean opinion score was 3.5, 3.5 and 3.3 for good, medium and poor signal conditions. However, we have to consider that once again, the VoIP quality was measured for laptop based VoIP communication. That is, it does not account for the additional processing delay for the terminal VoIP client implementation previously described.

The results in the live case still indicate that VoIP support for a handheld embedded client will be poor. However, these values do take into account the jitter buffer play out delay. The most noticeable difference between the three scenarios is the packet loss ratio, which increases as the signal quality decreases.

### 5.3. VoIP Quality in Mobility Scenarios

The mobile environment tests were measured from a van driving through a test route at average speeds of 60-80 km/h without stopping. The selected test route crosses the Helsinki metropolitan area from East to West and it is entirely covered by Elisa HSDPA network according to their publicly available coverage map. The test route was about 18.5km and it took approximately 15min to travel. The test route was driven several times to validate the results.

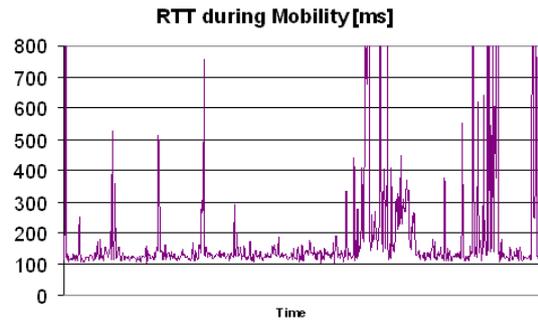
The results show that the average performance is lower than in static scenarios. A mobile scenario obviously brings several additional challenges due to the different cell changes along the test route. The number of cell changes along the route was 28 and were characterized via the changes in scrambling codes used. Table 4 summarizes the VoIP quality results.

Furthermore, in mobility scenarios the amount of RLC retransmissions required is very noticeable. To characterize the retransmissions, we conducted an additional test along the test route in which we sent continuous ping packets of 32B (see Figure 17). The results show a large amount of delay peaks resulting from these retransmissions. Therefore, it further supports our lab measurements and emphasizes the importance of the unacknowledged mode feature. We expect that this mode would potentially take the majority of large delay peaks, and thus, improve VoIP quality. However, if this mode is used, there is a possibility that the packet loss ratio will increase, and for that reason, it is very important to validate future results as well even if the feature is enabled.

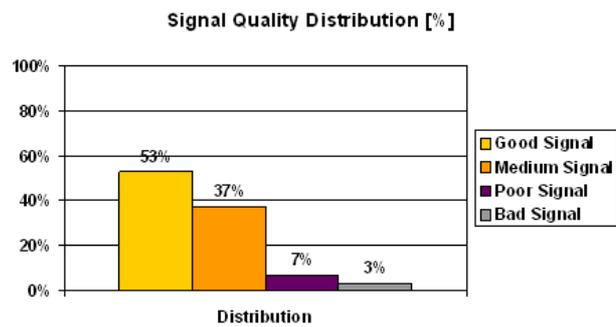
In addition, during the mobile tests, we recorded the signal conditions to characterize the signal quality distribution

**Table 4. VoIP quality in Elisa HSDPA network (including jitter buffer).**

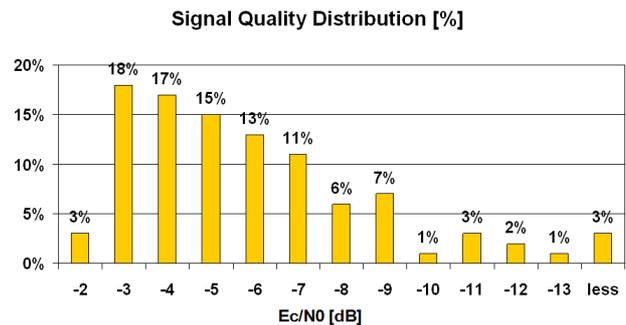
Scenario	Delay Avg.	Jitter Avg.	Packet Loss %	MOS
Good Signal (Ec/N0 -3 to -5)	288ms	19ms	0.4	3.5
Medium Signal (Ec/N0 -7 to -9)	283ms	19ms	1.0	3.5
Poor Signal (Ec/N0 -11 to -13)	266ms	14ms	2.6	3.3
Mobile Environment	331ms	22ms	1.9	3.2



**Figure 17. Round trip time during mobility tests.**



**Figure 18. Signal quality distribution.**



**Figure 19. Detailed signal quality distribution.**

along the test route. The measurements show that in general, it is highly probable to get a good signal level and that the coverage is well deployed (see Figure 18 and Figure 19).

## 6. Future Directions

At the current moment, the performance of VoIP in 3G networks is far from optimal. However, with some of the features and improvements in further 3GPP releases, the performance will improve. For instance, Release 6 equipment reduces RTT to roughly 65ms, and even lower with Release 7. Likewise, with Release 7 operators have other choices for deployment prior to full VoIP rollouts. For instance, advances such as Circuit Switched voice over HSPA (CS over HSPA) can improve capacity to similar levels as with VoIP. In this case, traditional voice

is carried over packet data. Hence, since VoIP does not provide any significantly better capacity figures over CS over HSPA, operators can delay VoIP deployment can be delayed until adequately performing terminals and networks are available. This however, is only possible if several features are upgraded in several network elements. These improvements occurred while this manuscript was under review. CS over HSPA is expected to be included in 3GPP Release 7 [21].

## 7. Related Work

Although, there is prior work investigating the VoIP performance in WCDMA and HSDPA systems, it is not very extensive and mostly based on simulations. For instance, some papers [22,23,24] study VoIP performance in WCDMA and provide some baseline results. In addition, other works [25,26] provide some estimated values for processing delays. In these studies, the assumption for the estimations is based on whether the call is towards a landline or a mobile end. Some performance simulations are also available [8,10,13,14,27–29]. However, the simulations only provide a delay budget rather than a description of the end user experience. Contrastingly, our study focuses on end user experience and VoIP quality rather than delay budgets alone. The delay budget values used in simulations vary from 80-150ms for studies ignoring encoding/processing delays and jitter buffer implementations [9,27,28,30,31], to 250-300ms for studies that assume such delays to some extent [8,10,13,14,29]. In addition, the estimations used in simulations are in general overly optimistic in regards to, e.g. client processing delay. Kim [14] considers the processing/encoding delay to be 50ms, while Ericson [13] assumes around 75ms. These delay values include the jitter buffer playout delay as well. Therefore, it is noticeable they are too optimistic, especially when compared to our experiment results with actual handsets and VoIP jitter buffer client implementations.

Even though, it is understandable that the exact encoding/processing delays and jitter buffer playout delays are client specific, unless they are modeled accordingly, or at least, to some extent, the differences in performance between simulations and actual deployments will remain very visible. Therefore, simulations results are only comparable to laptop based performance at its best and not to actual handheld performance, which in the end is the primary use case for VoIP services. Other simulation study [28], notices the importance of reducing RLC retransmissions to improve performance in FTP and HTTP browsing. However, the study does not address its importance for VoIP services. Finally, Wager and Sandlund [32] conduct simulations to determine the amount of possible lost frames of VoIP speech in HSDPA mobility scenarios.

In regards to VoIP signaling, SIP call setup delays and signaling performance have been studied previously mostly for Internet scenarios. ITU E.721 [17] recommendation and an IETF Internet Draft [33], provide call setup delays recommendations for circuit switched and Internet Telephony systems respectively. Additionally, Eysers and Schulzrinne [19] provide guidelines for Internet Telephony call setup and signaling transfer delays. In regards to 3GPP based wireless accesses, Kist and Harris [34] provide simulations for transfer delays with 3GPP signaling, while Fathi et al. [35] and Pous et al. [20] modeled signaling performance. Further, Curcio and Lunden [9] provide measurements for a WCDMA setup using laptop clients for local, international and overseas calls. Most of the mentioned research focuses on simulations, and does not consider some end user cases such as calls in wireless environments starting from different states. Additionally, performance with different wireless radio accesses and configurations under the same conditions is not available. Also, the available works do not use an embedded VoIP client in a handheld mobile terminal, which yields different delay values compared to a PC. HSDPA signaling performance has not been evaluated either. Our research aims at covering these items. The importance of evaluating a mobile terminal relies in the fact that the eventual substitution of circuit switched calls in 3GPP networks (HSDPA and WCDMA) for VoIP calls will take place with a handheld mobile device and not with a PC or laptop. Likewise, multi-radio devices can provide ubiquitous access via different wireless access technologies with distinct performance characteristics.

The lack of actual measurement performance values in literature could be mainly due to the unavailability of integrated VoIP clients in the terminals and available HSDPA networks. However, with the introduction of some multi-radio devices with VoIP capabilities (e.g. Nokia N95, Nokia 6110), it is possible to use VoIP applications without a PC.

## 8. Conclusions

Multiple measurements were carried out to evaluate and characterize the VoIP quality and VoIP signaling performance in HSDPA and WCDMA wireless accesses. The results show that HSDPA access is capable of providing a competitive VoIP quality compared to circuit switched voice. However, WCDMA in 128/128 and 64/64 bitrate configurations can only provide low and poor qualities, the main issues are long delays and packet losses, which occur often due to RLC retransmissions that overflow the jitter buffer capacity. However, the main issue with HSDPA is not only tied directly to the wireless access performance, but to the mobile device capabilities. Our results show that embedded mobile VoIP clients can introduce an increased delay due to processing when compared to laptop performance. This processing includes

e.g. encoding/decoding, and other operating system tasks. The additional delay has a considerable voice quality reduction effect. Further, the results from the test cases experimented in a live network resulted in lower performance when compared with similar laboratory measurements. Also, the effect of mobility in regards to VoIP quality degradation is quite noticeable. The degradation is due to handovers during the test route that increase the ratio of RLC retransmissions.

Therefore, the main aspects that can potentially improve VoIP quality performance with the current systems are mainly to reduce the number of RLC retransmissions by using unacknowledged mode, potentially use smaller jitter buffer sizes, and reduce the embedded VoIP client processing delays. High quality VoIP in 3G networks will be possible. However, it is tied to improvements in several areas such as wireless network delay, client implementation, and client processing delay. Finally, a main improvement developed while this manuscript was in process is CS over HSPA, which improves capacity and thus, can allow operators to delay VoIP deployment projects until networks and terminals have better performance.

## 9. References

- [1] H. Holma, and A. Toskala, "HSDPA/HSUPA for UMTS," John Wiley, 2006.
- [2] ITU-T, Recommendation G.107 "The E-model, a computational model for use in transmission planning," 2003.
- [3] R. Cole and J. Rosenbluth, "Voice over IP performance monitoring," ACM SIGCOMM'01.
- [4] A. Arjona, C. Westphal, A. Ylä-Jääski, and M. Kristensson, "Towards high quality VoIP in 3G networks: An empirical study," In Proceedings IEEE AICT'08, Athens Greece, June 8–13, 2008.
- [5] IxChariot, <http://www.ixiacom.com>.
- [6] A. Arjona and A. Ylä-Jääski, "VoIP call signaling performance and always-on battery consumption in HSDPA, WCDMA and WiFi," in Proceedings IEEE WiCOM'07, Shanghai China, September 21–23, 2007.
- [7] Wireshark Protocol Analyzer, <http://www.wireshark.org>.
- [8] B. Wang, K. Pedersen, T. Kolding, and P. Morgensen, "Performance of VoIP on HSDPA," IEEE Vehicular Technology Conference VTC'05 Spring, Vol. 4, pp. 2335–2339, May 30–June 1, 2005.
- [9] P. Lundén, and M. Kuusela, "Enhancing Performance of VoIP over HSDPA", In Proceedings IEEE VTC'07 Spring, April 21-24, 2007.
- [10] A. Braga, E. Rodriguez, and F. Cavalcanti, "Packet scheduling for VoIP over HSDPA in mixed traffic scenarios," in Proceedings 17th IEEE PIMRC'06, September 2006.
- [11] H. Kim, "Loosing Opportunism: Evaluating Service Integration in an Opportunistic Wireless System," IEEE INFOCOM'07, May, 2007.
- [12] Malden DSLA, <http://www.malden.co.uk/dsla>.
- [13] M. Ericson and S. Wänstedt, "Mixed traffic HSDPA scheduling–impact on VoIP capacity," in Proceedings IEEE VTC'07 Spring, April 21–24, 2007.
- [14] Y. Kim, "VoIP Service on HSDPA in Mixed Traffic Scenarios," in Proceedings IEEE CIT'06, September 2006.
- [15] Cisco, "Understanding delay in packet networks," Document ID: 5125, March 2007.
- [16] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session initiation protocol," RFC 3261, IETF June 2002.
- [17] ITU, Recommendation E.721, "Network grade service parameters and target values for circuit switched services in the evolving ISDN," 1999.
- [18] I. Curcio and M. Lundan, "SIP call setup delay in 3G networks," ISCC'02, Taormina Italy, July 1–4, 2002.
- [19] T. Eyers and H. Schulzrinne, "Predicting Internet telephony call setup delay," IPTel2000, Berlin, April 2000.
- [20] M. Pous, D. Pesch, G. Foster, and A. Sesmum, "Performance evaluation of a SIP based presence and instant messaging service," 3G 2003, June 25–27, 2003.
- [21] Nokia Siemens Networks and Nokia, "Supporting CS over HSPA," 3GPP R2-073487, August 2007.
- [22] F. Poppe, D. Vleeschauwer, and G. Petit, "Choosing the UMTS air interface parameters, the voice packet size and the de-jittering delay for a voice-over-IP call between a UMTS and a PSTN party," IEEE INFOCOM, Vol. 2, pp. 805–814, April 2001.
- [23] F. Poppe, D. Vleeschauwer, and G. Petit, "Guaranteeing quality of service to packetized voice over the UMTS air interface," 8th International Workshop on Quality of Service, pp. 85–91, June 2000.
- [24] R. Cuny and A. Lakaniemi, "VoIP in 3G networks: An end-to-end quality of service analysis," in Proceedings IEEE VTC'03 Spring, April 2003.
- [25] ITU-T, Recommendation G.114, "One-way transmission time," 2003.
- [26] TIATR-41.1.2, "VoIP end to end delay budget planning for private networks," Cisco 2000.
- [27] G. Rittenhouse and H. Zheng, "Providing VoIP service in UMTS-HSDPA with frame aggregation," in Proceedings IEEE ICASSP'05, March 18–23, 2005.
- [28] L. Bajzik, L. Korössy, K. Veijalainen, and C. Vulkán, "Cross-layer backpressure to improve HSDPA performance," in Proceedings IEEE PIMRC'06, Helsinki Finland, June 2006.
- [29] P. Hosein, "Scheduling of VoIP traffic over a time-shared wireless packet data channel," in Proceedings IEEE ICPWC'05, January 23–25, 2005.
- [30] Y. Seo and D. Sung, "Performance of VoIP in HSDPA based on an adaptive power allocation scheme," in Proceedings IEEE WCNC'06.
- [31] P. Hosein, "Capacity of packetized voice services over time-shared wireless packet data channels," in Proceedings IEEE INFOCOM'05, March 13–17, 2005.
- [32] S. Wager and K. Sandlund, "Performance evaluation of HSDPA mobility for voice over IP," in Proceedings IEEE

Vehicular Technology Conference VTC'07 Spring, April 22–25, 2007.

- [33] H. Lin, T. Seth, A. Broscius and C. Huitema, “VoIP Signaling performance Requirements and Expectations,” IETF Draft, June 1999.
- [34] A. Kist and R. Harris, “SIP Signaling Delay in 3GPP”, IFIP Interworking'02, October 13–16, 2002.
- [35] H. Fahti, S. Chakraborty, and R. Prasad, “Optimization of SIP session setup delay for VoIP in 3G wireless networks,” IEEE Trans. on Mobile Computing, Vol. 5, No. 9, September 2006.