Impact of Handover on VoIP Speech Quality in WiMAX Networks

Zdenek Becvar, Pavel Mach and Robert Bestak
Department of Telecommunication Engineering
Czech Technical University in Prague
Technicka 2, 166 27 Prague, Czech Republic
{becvaz1, machp2, bestar1}@fel.cvut.cz

Abstract— **VoIP is one of the most emerging technologies in the area of speech communications. VoIP is widely deployed in fixed line access networks. However, user's requirements on the mobility within communication and on the quality of the speech communications are increasing. Therefore, VoIP technology is more and more integrated into broadband wireless networks with QoS support such as WiMAX. The latest version of WiMAX standard, based on IEEE 802.16e, has implemented a full mobility support. The mobility is allowed due to a handover procedure. IEEE 802.16e specifies only one mandatory handover procedure: hard handover. This handover type is easy to implement, but it increases end-to-end packet delay that is critical for delay sensitive services such as VoIP. This paper evaluates an impact of the handover on the speech quality in VoIP networks.**

Keywords-VoIP, WiMAX, Handover, Speech Quality

I. INTRODUCTION

VoIP (Voice over Internet Protocol) is widely deployed technology and telecommunication operators seek to profit from it. The main advantage of this technology is a utilization of existing infrastructure in the form of internet connection. The usage of this type of communication is very cost effective. Unfortunately, this advantage brings some weak points that are expectable due to the low quality of internet connection. The Quality of Service (QoS) is mostly monitored issue by telecommunication operators and vendors. The quality of speech can be affected by many factors such as packet loss, packet delay, jitter, echo, noise [1], harmonic and inharmonic distortion [2], etc.

The QoS parameters are very closely connected with user's satisfaction with a received speech quality. The user's satisfaction is expressed with a subjective listening score which is a result of subjective listening tests. The substitutions to the subjective tests are objective methods. The objective methods assess the speech based on a signal processing without need of real listeners. PESQ method (Perceptual Evaluation of Speech Quality) is mostly used objective method for evaluation of the speech quality [3].

According to the standard IEEE 802.16e [4], a handover mechanism implemented in the WiMAX (Worldwide Interoperability Microwave Access) make possible to support a migration of a Mobile Station (MS) from the air interface provided by one Base Station (BS) to the air interface provided by another BS.

A short interruption can occurs in a communication between a MS and BSs due to handover. A short break increases a delay of packets that can cause the decrease in quality when using delay sensitive services (e.g. VoIP). In VoIP, the continuous data flow is needed on the receiving side in order to reconstruct a voice waveform correctly. If a data packet is not available at the given time, it has the same effect as if the packet would be lost. This causes decreasing of the quality of recovered signal on the receiving side.

The rest of the paper is organized as follows. The next section explains an interruption during handovers. Section III describes principles and requirements for a speech quality assessment in VoIP networks. Section IV is focused on analysis of speech quality degradation by handovers in WiMAX. Next section provides results of simulations. Last section presents our conclusions.

II. HANDOVER INTERRUPTION

Standard IEEE 802.16e specifies 3 types of handovers: Hard Handover (HHO), Macro Diversity Handover (MDHO) and Fast Base Station Switching (FBSS). The handover interruption occurs during a movement of the MS from the serving BS to the target BS. Duration of the handover interruption varies according to the type of handover.

A. Duration of handover interruption

Duration of a delay caused by HHO depends on the length of frames used in communication.

In case of MDHO, a MS and BSs have to maintain a diversity set. The diversity set is a list of the BSs, which are involved in the handover procedure. The MS defines an anchor BS among BSs in the diversity set. The MS is synchronized and registered to the anchor BS. The MS communicates (including user traffic) simultaneously with all BSs in the diversity set (including anchor BS). Communication with all BSs in the diversity set is processed by the MS/BSs. Therefore, if the diversity set contains more than one BS, there is no delay of data packets caused by MDHO. The delay is similar as in HHO in case of one BS in the diversity set.

In the case of FBSS, the situation is similar as in MDHO. A MS and BSs also have to maintain the diversity set. The MS transmits/receives data to/from all BSs in the diversity set, but only a communication with the anchor BS is processed. FBSS handover actually means a changing of the anchor BS. No delay is introduced by the anchor BS switching if the diversity set contains 2 or more BSs. If the diversity set includes only one BS, introduced delay is he same as in HHO.

While HHO is mandatory handover in WiMAX systems, MDHO and FBSS are optional ones. Hence this paper is focused primarily on HHO.

B. *Handover interruption time*

The handover interruption in HHO and MDHO/FBSS with only one BS in the diversity set is caused by a switching of a MS from a serving BS to a target BS (Fig. 1).

Before the handover initialization, the MS communicates with the serving BS (phase 1, Fig. 1). When the MS crosses a boarder of cells between the serving BS and the target BS, the MS is disconnected from the serving BS and for a while the MS has no connection to the network (phase 2, Fig. 1). Subsequently, a new connection is established with the target BS (phase 3, Fig. 1). Short delay rises between a time of disconnection from the serving BS and a time of setting up the connection with the target BS. During interruption, packets are routed from the serving BS to the target BS via backbone network. When the connection between the MS and the target BS is finally established, the packets are sent to the MS.

According to [4], the handover procedure can be divided into several phases (Fig. 1). Within Network Topology Advertisement phase, the MS collects information about BSs in its neighbourhood. During Scanning phase, the MS seeks for a suitable handover target BS or BSs that are suitable to be added to the diversity set. Scanning results are reported back to the serving BS.

The results of scanning are used in the next phase of handover procedure – Cell Reselection. Within the cell reselection, a possible target BS is selected based on signal parameters and offered QoS. The cell reselection phase is followed by the Handover Decision and Initiation process. Handover can start after all conditions and requirements for its initialization are fulfilled. The next step of handover procedure is the synchronization to the downlink channel of the new target BS. However, before the synchronization is executed, firstly all connections between the MS and the serving BS should be closed. Henceforth, all data transmissions between the MS and all BSs are cancelled (orange arrow in Fig. 2).

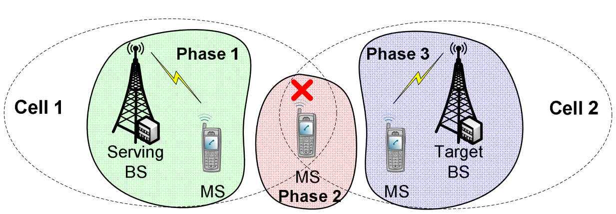

Figure 1. Origination of handover interruption

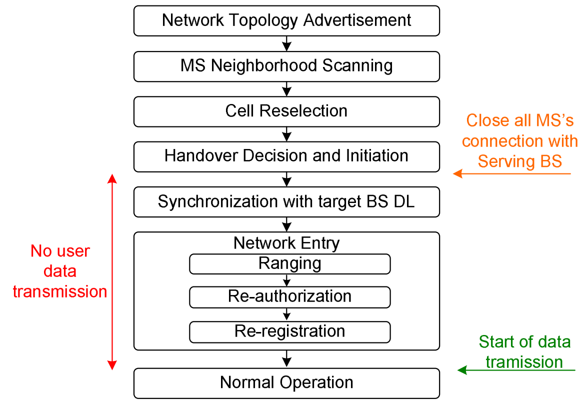

Figure 2. Decomposition of handover procedure into stages

As soon as the synchronization with the downlink channel is done, the MS can begin Network Re-entry procedure. The network re-entry is composed of three subphases: Ranging, Re-authorization and Re-registration. In the ranging process, the MS obtains information about an uplink channel. Ranging is followed by authorization and registration of the MS to the target BS. The MS can begin a communication (normal operation) after successful authorization and registration to the target BS (green arrow in Fig. 2).

III. VOIP SPEECH QUALITY

Several complications connected with VoIP technology such as a packet loss, packet delay or packet delay variation (called jitter) may be recognized. The most of these problems generally lead to the loss of data flow continuity and further to the loss of signal information elements. From the point of human perception system view, this loss is represented as a dropout.

The speech quality can be evaluated either by subjective tests or by objective methods. Both, subjective and objective methods, use a parameter MOS (Mean Opinion Score) [5] to speech quality assessment. MOS scale range used in subjective tests is from 5 to 1 (Excellent = 5; Good = 4; Fair =3; Poor = 2; Bad = 1). In practice, objective methods are usually used due to easier implementation.

A. *Speech quality assessment in VoIP*

The objective method PESQ is used for evaluation of the quality of speeches affected by the handover procedure. PESQ is one of the most spread objective methods developed for end-to-end speech quality assessment in a conversational voice communication. The PESQ method can be used for narrow-band and wide-band networks.

The principle of PESQ is shown in Fig. 3. The PESQ method is based on the comparison of an original (non-degraded) signal X(t) with a degraded signal Y(t). Y(t) is the result of spilling signal X(t) through a communication system. The PESQ method generates a prediction of the quality which would be given to the signal Y(t) in a subjective listening test.

The range of PESQ MOS score (according to ITU-T P.862) is between -0.5 and 4.5. Since this range does not

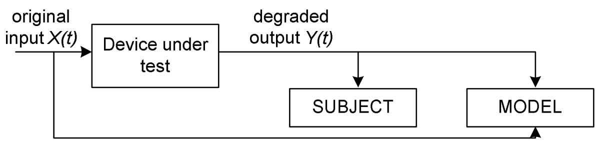

Figure 3. PESQ principle

correspond to a scale used for the subjective test, the ITU-T P.862.1 recommendation [6] recalculates PESQ MOS according to (1) to better match to the subjective test results.

$$y = 0.999 + \frac{4.999 - 0.999}{1 + e^{-1.4945*x + 4.6607}} \tag{1}$$

where x is the objective PESQ MOS score and y is the matching ITU-T P.862.1 MOS score.

B. Requirements for quality assessment of speeches

ITU-T P.862 recommendation describes all requirements on the tested speech signal. The signal shall have a character as a real speech signal carried by a communication network. The tested speech signal shall include speech bursts with duration between 1-3 s. These bursts must be separated by intervals of silence. The speech activity must be between 40% and 80% of the overall speech length. The most of experiments use two or three bursts with overall duration of 8 s. Also, the utterances with duration from 8 to 20 s can be used. Frequency characteristics of the speech signal and signal level alignment must be in accordance with recommendation ITU-T P.830 [7].

Utterances used for speech quality assessment are digital studio recordings. The average length of speeches is between 8 and 12 s. Tested signals are 16-bit linear PCM (Pulse Code Modulation) sampled with 8 kHz sample rate (downsampled from the 48 kHz studio quality).

IV. Degradation of Speech Quality by Handover

HHO and MDHO/FBSS handovers with only one BS in diversity set cause a short break in the communication. Within this interruption, the packets are forwarded from the serving BS to the target BS. Consequently, the packet delay is increased and the speech quality is degraded.

A. Degradation parameters

Generally, 3 types of degradation can occur in VoIP: packet delay, jitter and packet loss. The packet loss can be solved by resending of lost packet using ARQ (Automatic Repeat reQuest) mechanism. On the other side, ARQ increases a packet delay and a jitter. Thus, it is not usually used in VoIP communication. Hence, we assume communication with disabled ARQ. The networks packet losses and jitter are not considered in this paper to separate an impact of handover from other effects.

A packet delay is affected by many factors (e.g. routing, signal propagation...). Overall delay of packets during handover can be defined by the following equation:

$$D_{TOT} = D_{HO} + D_{NET} \tag{2}$$

Overall delay consists of a delay caused by a transport network D_{NET} and a delay caused by the handover phase D_{HO} (denoted as a handover interruption time).

All handover stages following the handover decision and initialization increase the packet delay [8]. The overall delay caused by handover can be expressed by equation:

$$D_{HO} = T_{sync} + T_{cont\_res} + T_{rng} + T_{auth} + T_{reg} \tag{3}$$

where T_{sync} is DL synchronization time, T_{cont_res} is duration of the contention resolution procedure, T_{rng}, T_{auth} and T_{reg} represent times spend by ranging, re-authorization and re-registration respectively. The typical and minimal values of the handover interruption time components are presented in Tab. I.

The length of the re-authorization depends on the number of Security Associations (SA). SA is a set of security parameters describing a connection. More data connections can use same SA.

A length of frame is also considered in handover interruption time evaluation. The WiMAX can use following frame lengths: 2; 2.5; 4; 5; 8; 10; 12.5 and 20 ms. In our paper, two scenarios are specified for an analytical calculation of handover delay (Tab. II). Scenario A corresponds to the "optimal handover", it means the values are selected to achieve a minimal delay caused by handover (Tab. II). The second scenario (scenario B) corresponds to the typical values considered in practice and simulations [8].

The delay of each packet in network, D_{NET} (see [9]), is calculated according to following equation:

$$D_{NET} = T_{EndTr} + T_{EndRc} + 2 * T_{AcNet} + T_{CoreNet} \tag{4}$$

where T_{EndTr} represents delay caused by end-device at transmitting side (incl. signal processing, packetization and serialization). Time T_{EndRc} is delay incurred by receiving end-device and it is composed of processing and jitter buffer delay (we assume no jitter buffer delay since the jitter is neglected). Parameter T_{AcNet} is delay originated in access networks (two access networks are included in telecommunication chain, one access network at each communicating side) and $T_{CoreNet}$ represents a delay introduced in a core network. Parameters T_{AcNet} and $T_{CoreNet}$ include signal propagation, a data serialization and queering. The parameters used for network delay evaluation shown in Tab. III.

TABLE I. Minimal and Typical Values of the Handover Interruption Time Components

Delay	Minimal value	Typical value
T_{sync}	N/A	< 20 ms
T_{cont_res}	0 ms (dedicated ranging slot)	tens ms
T_{rng}	5 frames	6-9 frames
T_{auth}	3 frames + 2 frames per SA	N/A
T_{reg}	2 frames	2 frames

TABLE II. PARAMETERS FOR HANDOVER DURATION CALCULATION

Delay	Duration – Scenario A	Duration – Scenario B
T_{sync}	5ms	20ms
$T_{cont\ res}$	0ms	40ms
T_{rng}	5 frames	7frames
T_{auth}	3 frames + 2 frames per SA	3 frames + 2 frames per SA
T_{reg}	2 frames	2 frames

TABLE III. PARAMETERS FOR NETWORK DELAY CALCULATION

Delay	Duration
T_{EndTr}	60 ms
T_{EndRc}	5 ms
T_{AcNet}	50 ms
$T_{CoreNet}$	80 ms

B. Speech degradation

The speeches for the handover impact evaluation are modified according to the voice signal degradation caused by handover in the real wireless network. Since we investigate an impact of handover, the core network packet losses and jitter are neglected. All degradations are done in MATLAB. The speech processing procedure is presented in Fig. 4.

The first step is a determination of the handover interruptions positions due to HO process itself. This task corresponds to the determination of a time when handover occurs. Based on mobility models defined for handover evaluation according to IEEE 802.16m [10], handover in periodic intervals can be assumed. This simplification has only a minor impact on results as we use a large number of speeches for the speech quality evaluation. This mobility model assumes a direct movement of users with a constant speed among the regular hexagonal cells with constant radius. The calculation of the interruption duration is based on the principle described in section IV.A. Overall delay is converted to the amount of lost VoIP packets:

$$\text{LostPackets} = \frac{D_{TOT}}{PL_{VoIP}} \tag{5}$$

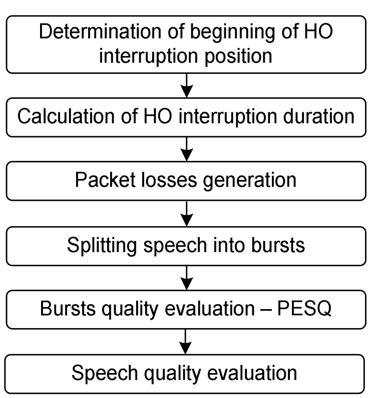

Figure 4. Process of a calculation of handover impact on speech quality

where PL_{VoIP} is the length of VoIP data packet. The parameter PL_{VoIP} is assumed to be 20ms since this value corresponds to the typical packet length used in real VoIP communication.

The packet losses are integrated into the speech by a zeroing of the speech samples at appropriate positions. The whole speech is split into bursts with an approximate duration of 10s to meet the requirements for evaluation by PESQ. The method PESQ results are recalculated according to ITU-T P.862.1 using (1). The quality of whole speech is calculated as an average of the quality of bursts pertaining to the same speech. There are generated 144 speeches for each length of intervals between two handovers and for all handover interruption durations. The results of 144 speeches are averaged to avoid of an impact of a random drops of packet losses into silent parts of the speeches.

V. RESULTS

The relation between WiMAX physical layer frame duration and handover interruption time is shown in Fig. 5. It is presented for 3 different counts of SAs. The figure shows the linear dependence between frame duration and handover interruption time. The minimal value of handover delay is 29 ms (blue dash line with circle mark in Fig. 5). This delay is caused by optimal handover (Scenario A) with 1 SA. We can also observe a significant impact of number of SA in the packet delay caused by handover.

The packet delays for the speech quality evaluation are determined based on the dependence between handover interruption time and the frame duration. Both scenarios (defined in Tab. II) are considered.

The handover interruption times appertain to the range between 1 and 5 SAs (duration of 50; 75; 100; 150; 200 and 250 ms) were selected for an investigation of dependence between speech quality and intervals between handovers. Besides, the handover interruption lasting for 25 ms was considered in order to fulfill the IEEE 802.16m requirements [10] (orange dash-dot line in Fig. 5).

Exact values of the handover interruption times corresponding to 1 SA in Fig. 5 were selected for an investigation of an impact of the frame duration on the speech quality. The values for both varieties are presented in Tab. IV.

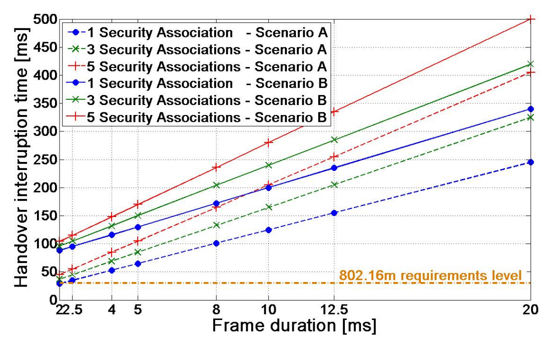

Figure 5. Dependence between the duration of the handover interruption and the duration of the frames

TABLE IV. DEPENDENCE OF THE HANDOVER INTERRUPTION ON THE FRAME DURATION FOR 1 SA

Frame duration (ms)	HO interruption time	
	Variety A	Variety B
2	29	88
2.5	35	95
4	53	116
5	65	130
8	101	172
10	125	200
12.5	155	235
20	245	340

An impact of WiMAX handover on speech quality in VoIP is shown in Fig. 6 and Fig. 7. Fig. 6 presents the dependence of the speech quality on the average interval between handovers. From the figure we can observe nearly constant speech quality for the very short interruption (25ms), the quality is very high for all intervals between handovers. The speech quality degradation is significant for all longer duration of the handover interruption. Especially for fast moving users when the handover is proceeded very often (interval between handovers <30 s; it corresponds to speed of users > 20m/s for cell radius of 600m [10]).

The impact of handover becomes less important as increase the interval between handovers and decrease the handover interruption duration.

Fig. 7 shows an impact of handover on the speech quality. There are considered 4 different average intervals between handovers (10, 20, 40 and 60 s; this corresponds to users speed 60, 30, 15 and 10m/s for 600 m cell radius if we assume the multiple moving MS mobility model defined in [10]) and both abovementioned scenarios. The impact of the frame duration is noticeable only for high frequented handovers. The difference in the quality between 2 ms and 20 ms frame durations is up to approximately 0.4 MOS in both scenarios. In these cases, better speech quality can be achieved by shortening of the frame duration.

In case of longer interval between handovers, the impact of the frame duration on speech quality is nearly not perceptible. The usage of shorter frames does not result in better speech quality in this case. The results of the optimal

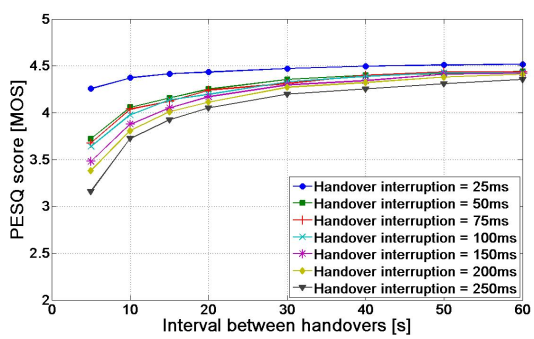

Figure 6. Decrease of VoIP speech quality caused by handover with different duration of the interruption

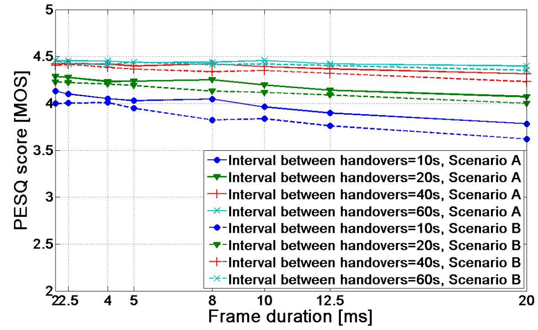

Figure 7. Impact of handovers on the speech quality depending on frame duration

handover (scenario A, solid lines in Fig. 7) evoke insignificant increase (max 0.1 MOS) in speech quality in comparison to the typical handover (scenario B, dash lines in Fig. 7) for longer intervals between handovers. If the high frequent handovers are considered, the optimal handover cause the less decrease (up to 0.25 MOS) of speech quality in comparison with the typical handover.

VI. CONCLUSION

This paper investigates an impact of handover on the speech quality in VoIP communication.

The length of the handover interruption in VoIP packet stream depends linearly on the frame duration. The important factor influencing the packet delay is a number of SA. The packet delay caused by handover increases as increase the number of SA. In the optimal case, the achievement of 802.16m requirements (it means to decrease the handover interruption less than 30 ms) is possible by using the frames with 2 ms duration (handover interruption duration is 29 ms). Delay caused by a typical handover is between 88 ms and 340 ms for one SA and for 2 ms and 20 ms frame duration respectively.

An impact of handover on the speech quality is perceptible for longer durations of the handover interruption or for very frequent handovers (high speed MS). When the 802.16m requirements on handover interruption duration are fulfilled, the impact of handover is negligible even for very fast moving MSs.

A significant increase in the speech quality can be achieved by using shorter frames (e.g. 2 ms frames instead of 20 ms frames bring speech quality increase up to 0.4 MOS for a fast moving MS). The difference in the speech quality is lesser for low frequent handovers (low speed MS). Increase of the quality is up to 0.15 MOS by using 2 ms frames instead of 20 ms.

ACKNOWLEDGMENT

This work has been performed in the framework of the FP7 project ROCKET IST-215282 STP, which is funded by the European Community. The Authors would like to acknowledge the contributions of their colleagues from ROCKET Consortium (http://www.ict-rocket.eu).

References

[1] Z. Becvar, L. Novak, J. Zelenka, M. Brada and P. Slepicka, "Impact of Additional Noise on Subjective and Objective Quality Assessment in VoIP," Proc. 2007 International Workshop on Multimedia Signal Processing, Chania, Greece, 2007, p. 39-42.

[2] Z. Becvar, J. Zelenka and M. Brada, "Impact of Saturation on Speech Quality in VoIP," Proc. IWSSIP 2008. Bratislava, Slovakia. 2008, p. 381-384.

[3] ITU-T Recommendation P.862, "Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to end Speech Quality Assessment of Narrow-band Telephone Networks and Speech Codecs," 2001.

[4] IEEE P802.16e-2005, "Air Interface for Fixed and Mobile Broadband Wireless Access Systems: Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands," New York, February 2006.

[5] ITU-T Recommendation P.800.1, "Mean Opinion Score (MOS) terminology," 2003.

[6] ITU-T Recommendation P.862.1, "Mapping function for transforming P.862 raw result scores to MOS-LQO," 2003.

[7] ITU-T Recommendation P.830, "Subjective Performance Assessment of Telephone-Band and Wideband Digital Codecs," 1996.

[8] Wenhua Jiao, Pin Jiang and Yuanyuan Ma, "Fast Handover Scheme for Real-Time Applications in Mobile WiMAX," Proc. of International Conference on Communication (ICC 2007), Glasgow, Scotland, June 2007.

[9] H. Astrom, "PlayOut Controller – QoS at the IP Edge Points: For networks with packet loss and jitter, Speech Processing," Transmission and Quality Aspect (STQ) Workshop, 2003.

[10] Proposal paper IEEE 802.16m-07/002r4, "IEEE 802.16m System Requirements," October 2007.