

Transcoding free voice transmission in GSM and UMTS networks

Sara Stančin, Grega Jakus, Sašo Tomažič
University of Ljubljana, Faculty of Electrical Engineering

Abstract - Transcoding refers to the conversion between two encoding schemes of a digital signal. It is usually performed where two interfaces do not support the same encoding. Transcoding introduces some undesired effects into the signal, the most important of which are distortions and delays. In this paper we give our attention to possibilities of transcoding free operations in a GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications System) network. Tandem Free Operation (TFO) in GSM networks enables transmitting voice transparently through the core network without transcoding. Although TFO has some advantages, such as improvement in speech quality and reduction of delays, it also has many limitations. Transcoder Free Operation (TrFO) is similar to TFO but is employed in the packet-based core networks, such as UMTS. TrFO overcomes some of the TFO limitations. TrFO reduces bandwidth and voice call costs. It increases network capacity and is more robust than TFO. In a UMTS network, when TrFO is not possible, TFO can still be attempted. Interworking of both mechanisms is necessary for mixed GSM/UMTS networks.

1. INTRODUCTION

Despite the increasing data rates and the amounts of the transferred data, voice calls remain the most important application in the mobile domain. However, in the past, the quality of voice was not a primary concern to the mobile operators. Voice was compromised using aggressive voice compression to save the scarce and costly frequency spectrum. The advent of new radio access technologies, more efficient compression techniques, multifunctional mobile terminals supporting multimedia applications and the consequent high user expectations forces mobile operators to offer higher quality voice applications.

One of the factors that have a negative impact on voice quality in mobile networks is *transcoding*. The term *transcoding* refers to the conversion between two encoding schemes of a digital signal. Transcoding can be performed using the same format, this is known as *self tandeming* or two different formats known as *cross tandeming*. Transcoding is used where two interfaces do not support the same encoding scheme.

Ideally, transcoding of a compressed signal is performed without the prior decompression into some intermediate format (e.g. G. 711 [1] for audio transcoding). However, while such conversion is feasible in the context of video processing, the audio and speech can currently employ merely a 'brute force' approach. This means that before compression in the target format, decompression into G.711 format is necessary.

Transcoding introduces some undesired effects into the signal. The most important are distortions and delays. The distortions are cumulative and are a consequence of:

- loss of audio information
- quantization errors
- algorithmic errors (pre-echo, metal sounds, oscillations etc.)

Other downsides of transcoding are the need for additional DSP (*Digital Signal Processing*) resources, unsupported cryptography between the endpoints and more difficult implementation. Due to all of the above-mentioned negative impacts on voice quality, transcoding should be avoided whenever possible [2].

To reveal the effects of transcoding on user experience, various tests were conducted. Performance of various AMR codecs tandeming with GSM codecs is presented in [3].

In a GSM (*Global System for Mobile Communications*) network, encoding schemes for voice transmitted through the core and radio parts of the network differ. In this paper we present the possibility to serve a voice call in a GSM network without voice transcoding. We examine what are the elements and logic necessary to provide this functionality in a GSM network. We also present the differences when providing transcoding free voice transmission in a UMTS (*Universal Mobile Telecommunications System*) network and we give a brief overview of transcoding free voice calls when GSM and UMTS networks are interworking.

2. GSM/UMTS NETWORK TRANSCODING SCHEMES

During a voice call in a GSM network, both mobile devices perform voice encodings to make user voice suitable for GSM radio network transmission. On the GSM radio interface transmitted voice is encoded using Full Rate (FR) [4], Half Rate (HR) [5], Enhanced Full Rate (EFR) [6] or Adaptive Multi Rate (AMR) [7] codecs. These schemes incorporate voice compression necessary in order to assure better use of the limited-bandwidth radio channel. Voice frames are then typically decompressed and re-encoded for transport over the 64 kbps circuit switched links through the core network. For such transport, the G.711 standard is used that is common in digital switched telephone networks. The reason networks were designed in such a way is simple connections to other networks (e.g. Public Switched Telephone Network, PSTN) and possible additional voice processing in the core network itself, like for example echo cancellation. In the common GSM voice transcoding scenario it is therefore needed to perform transcoding.

Because of different supported voice encoding schemes when passing from the radio to the core network and reverse, transmitted voice must be transcoded. Figure 1

presents the main network elements activated during call setup where a GSM user initiates a voice call with another GSM user.

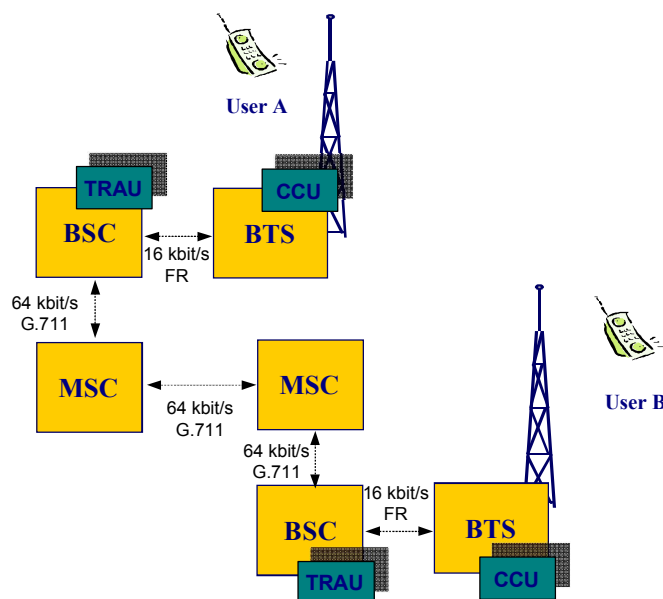


Figure 1: Voice transcoding in a GSM network.

in a GSM network Two units are responsible for voice transcoding: the Transcoder and Rate Adaptation Unit (TRAU) and the Channel Coding Unit (CCU). TRAU is an independent network unit responsible for voice encoding and decoding as well as for data rate adaptation. Between two TRAU units in a mobile network, transmitted voice is encoded to the 64 kbit/s G.711. The TRAU unit is logically a part of the BSC (*Base Station Controller*) while its physical location can be between the BSC and the BTS (*Base Transceiver Station*) or between the BSC and the MSC (*Mobile Switching Centre*). The second possibility enables cost reduction of the leased lines between the MSC and the BSC due to lower bit rates.

On the radio interface, GSM encoders support 16 kbit/s logical channels. 20 ms voice frames are encoded with 260 bits giving a bit rate of 13 kbit/s. The difference between 13 and 16 kbit/s represents 60 bits of voice coding information including coding scheme and rate bits. These bits are transmitted in the so-called TRAU frames.

The CCU is a part of the BTS. It takes care for channel coding and radio network quality measurements. Upon this information, the CCU can determine a suitable encoding scheme. In a GSM network, information about

the selected encoding scheme is sent in-bound, together with transmitted user data.

Figure 2 presents voice transcoding in a UMTS network. In UMTS, the standard voice encoding scheme for transmission over the UMTS radio network is the narrow-band AMR scheme [8]. The scheme consists of 14 modes providing bit rates from 4,75 kbit/s up to 12,2 kbit/s. The selected mode primarily depends on radio channel conditions and on voice content. Beside the standard narrow-band scheme, wide-band AMR (AMR-WB) scheme can also be used consisting of bit rates from 6,60 kbit/s to 23,85 kbit/s and encoding the bandwidth up to 16 kHz.

In general, transcoding operations in a UMTS network are a part of the media gateway (MGW) function set. Other MGW functionalities are: announcement services, echo cancelation, DTMF (*Dual-Tone Multi-Frequency*) detection and generation, support for transport protocols like ATM (*Asynchronous Transfer Mode*), IP (*Internet Protocol*) or TDM (*Time Division Multiplex*), support for *lu* interfaces, bad frame treatment, IP protocol-based functions like RTP/RTCP (*Real-Time Transport Protocol/Real-Time Transport Control Protocol*), encryption and QoS (*Quality of Service*).

TFO supports GSM encoding mode adaptation to radio network conditions. When speech is transcoded, the encoding mode is adapted on each of the connection sides separately. When TFO is active and one connection side perceives radio network condition degradation, the TFO must initiate encoding mode change on both connection ends without any negotiation. Both sides must perceive radio network condition improvement in order to adapt the encoding mode.

TrFO (*Transcoder Free Operation*) [11, 12] is similar to TFO but is employed in packet-based core networks which are based on high bandwidth ATM or IP links rather than on 64 kbps TDM links. In such core networks it is therefore possible to transmit voice data streams with other codecs than 64 kbit/s G.711. The MSC can therefore establish a voice connection without activating transcoders as illustrated in Figure 4.

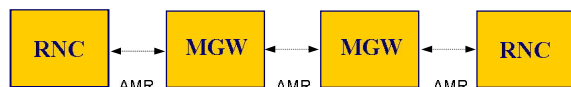


Figure 4: TrFO voice transmission principle in a UMTS network. A voice call can be established without activating the transcoders.

TrFO uses 'out-of-band signalling', which means that messages for transcoder-free operation negotiation and to establishment are not transmitted on the same link as user data. Both mobile terminals report their codec capabilities the corresponding serving MSC before the bearer path is established. It is only when both sides negotiate the common encoding mode that the barrier can be established. TrFO operation uses the Out of Band Transcoder Control (OoBTC) [11] mechanism which is responsible for configuring the call without involving transcoders. It supports encoding mode negotiation and encoding mode list changes/adaptations. Unlike TFO, TrFO is established and controlled before the call is configured. Selected encoding mode can be changed later on during the call.

If avoiding transcoding in a UMTS network can not be fully achieved, *Remote Transcoder Operation* (RTO) [12] can include a single transcoder in the user data path. This does not imply double voice transcoding of user voice. From all possible in-path transcoders, the one used should be the one closest to the user device supporting the higher bit rate encoding scheme. Such a scenario is presented in Figure 5 and is also applicable for establishing voice calls to and from PSTN networks.

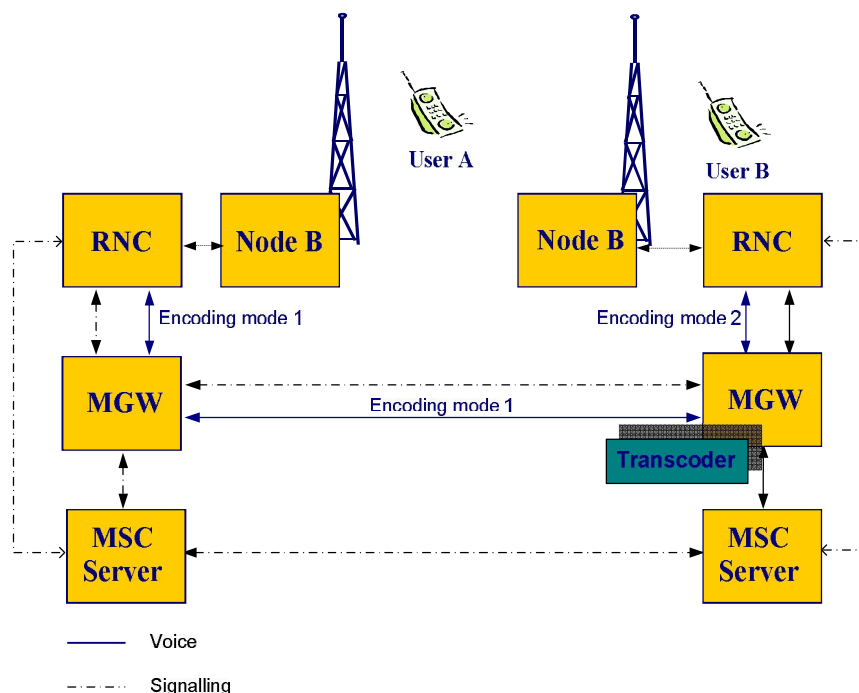


Figure 5: Single voice transcoding principle in a UMTS network.

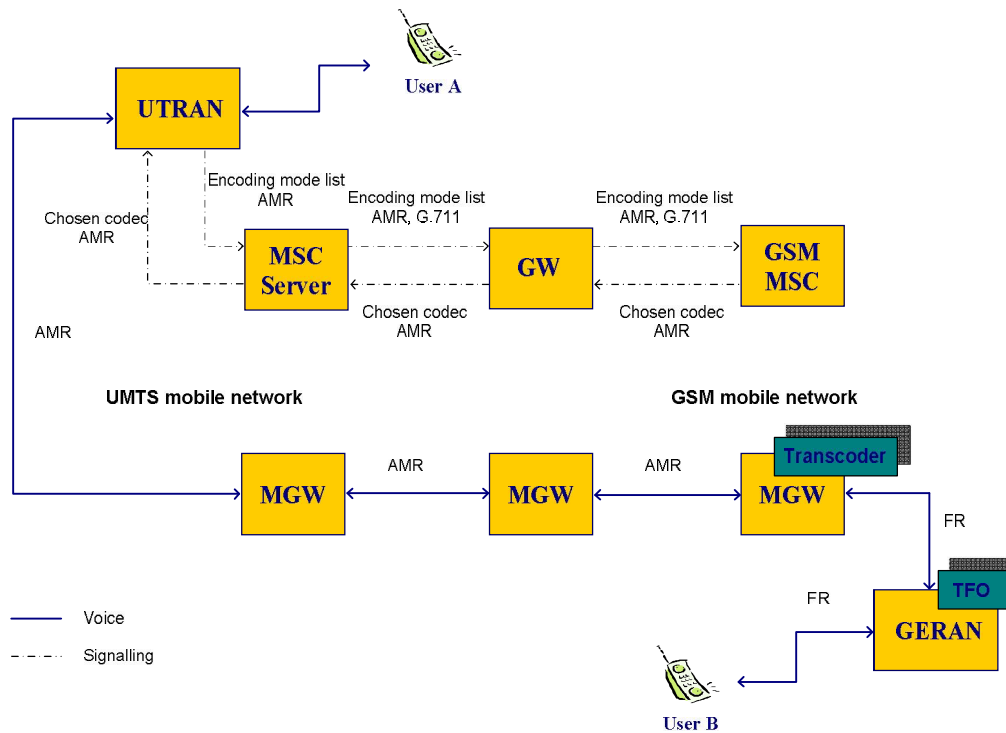


Figure 6: TFO and TrFO interworking in a GSM/UMTS network.

The OoBTC procedure can result with choosing the G.711 codec as the common codec in the UMTS network. In such a case, a transcoder is inserted in the appropriate MGW in order to perform the necessary AMR and G.711 transcoding. The network initiating the call is informed about the selected codec G.711. In such a case, TFO operations in the GSM network are pointless.

4. TFO AND TRFO EFFECTS

Although TFO has some advantages, such as improvement in speech quality and reduction of delays, it also has many limitations:

- Only mobile to mobile calls are supported
- Problems with the so-called digital transparency
- Group and conference calls are not supported
- Problems with hard handovers
- The transmission of DTMF signalization and announcements

The used codecs are negotiated independently during call-setup between mobile terminals and their corresponding TRAU units. Because the TFO procedure is configured after call-setup is completed, if non-compatible codecs are selected during the call-setup, TFO cannot be applied.

Another problem is associated with the so-called *digital transparency*. The digital transparency refers to the case when the digital content is not altered in any way by any of the network elements on the path between the TRAUs (IPE, In-Path Equipment). Any intervention by the IPE

would cause the corruption of TFO messages and consequently the failure of TFO transmission. To enable the TFO, the IPE must be disabled or properly configured.

Another event which interrupts TFO is the inter-BSC handover when one of the TRAUs is replaced by the new one corresponding to the new BSC. In this case, the TFO is temporarily interrupted but is later renegotiated if the new TRAU is TFO capable. The intra- and inter- BTS handovers are generally not problematic since the TRAUs do not alter.

TFO is also temporarily interrupted when DTMF tones or announcements must be inserted by a MSC. Since MSC is not aware of the TFO, it can overwrite the TFO signalling and the compressed voice information. The distortion is immediately detected by one of the TRAUs and the tandem transcoding is temporarily re-established.

Group and conference calls are also problematic in the context of TFO. The conference bridges, namely, most often function by mixing voice signals of involved parties encoded using G.711 codec. The mixing of compressed voice signals and TFO messages would again distort the voice signal and the TRAU units would again have to reinsert the tandem transcoding. If a multi-party call turns into a normal call between two parties, TFO can be reconfigured.

Finally, one of the major drawbacks of TFO is its overall effect on network capacity. Even though TFO improves voice quality and decreases delays, it does not improve

the overall capacity of the network. Uncompressed voice is, namely, still transmitted in parallel with TFO traffic.

On the other hand, TrFO reduces bandwidth and voice call costs and increases network capacity (despite its superior quality, AMR-WB codec also requires just a third of the bit rate of the G.711 codec). TrFO is also more robust than TFO as it supports sudden reconfigurations (e.g. because of handovers) via out-of-band signalling. Furthermore, it supports the use of wideband codecs (e.g. AMR-WB) that are not compatible with the G.711 (which is a narrow-band codec) and can therefore not be used with TFO. As AMR-WB can encode twice the frequency range as older GSM and G.711 codecs voice quality is improved.

When TrFO is not possible, TFO can still be attempted.

5. CONCLUSION

In this paper, we presented the TFO and TrFO approaches for avoiding undesired voice transcoding in mobile networks. Both approaches enable better voice quality. TrFO also has many other advantages, such as lower delays and reduced processing requirements. The latter also reduces the cost of voice transmission.

There are some open questions regarding the TFO/TrFO interworking such as mobile call handover from a UMTS to a GSM network and increased signalling.

REFERENCES

- [1] ITU-T. "G.711.0: Lossless compression of G.711 pulse code modulation"
- [2] TIA TSB-116-A Telecommunications - IP Telephony Equipment – Voice Quality Recommendations for IP Telephony
- [3] 3GPP TS 26.090 - Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions"
- [4] Digital cellular telecommunications system (Phase 2+) (GSM); Full rate speech; Transcoding (GSM 06.10 version 8.1.1 Release 1999), ETSI
- [5] Digital cellular telecommunications system (Phase 2+) (GSM); Half rate speech; Half rate speech transcoding (GSM 06.20 version 8.0.1 Release 1999)
- [6] Digital cellular telecommunications system (Phase 2+) (GSM); Enhanced Full Rate (EFR) speech transcoding (GSM 06.60 version 8.0.1 Release 1999)
- [7] ETSI TR 126 976 V6.0.0 (2004-12), "Performance characterization of the Adaptive Multi-Rate Wideband (AMR-WB) speech codec"
- [8] 3GPP TS 26.103 Speech codec list for GSM and UMTS
- [9] 3GPP TS 23.053 Tandem Free Operation (TFO); Service description; Stage 2
- [10] 3GPP TS 28.062 Inband Tandem Free Operation (TFO) of speech codecs; Service description;
- [11] 3GPP TS 23.153 Out of band transcoder control; Stage 2
- [12] 3GPP2 Transcoder Free Operation; Stage 1 - Requirements