CONCEPTS AND SOLUTIONS FOR LINK ADAPTATION AND INBAND SIGNALING FOR THE GSM AMR SPEECH CODING STANDARD

Stefan Bruhn¹, Peter Blöcher², Karl Hellwig² and Johan Sjöberg²

¹Ericsson Radio Systems AB
S-164 80 Stockholm
SWEDEN

²Ericsson Eurolab Deutschland GmbH
D-90411 Nürnberg
GERMANY

e-mail: Stefan.Bruhn@era.ericsson.se, Peter.Bloecher@eedn.ericsson.se,
Karl.Hellwig@eedn.ericsson.se, Johan.Sjoeberg@eedn.ericsson.se

Abstract – The European Telecommunications Standards Institute (ETSI) has just defined an Adaptive Multi Rate (AMR) speech codec standard for the GSM system with a multitude of source and channel coding rates. The standard aims to provide robust high quality speech together with the flexibility to deliver radio network capacity enhancements by means of low bit-rate operation. The codec rates are dynamically selected with respect to the rapidly changing radio conditions and to local capacity requirements. This paper describes various approaches for link adaptation with respect to varying radio channel conditions and puts a focus on the solution in the AMR standard. Moreover the method of inband signaling that is standardized is discussed and motivated.

I. INTRODUCTION

Speech will remain one of the most important services in the next generation of wireless systems. The challenge is to enhance flexibility and speech quality while at the same time optimizing the spectrum efficiency, particularly in areas where the subscriber density is high [1].

Currently existing speech coding standards in the GSM system are the Full-Rate (FR), Enhanced FR (EFR), and Half-Rate (HR) speech codecs. The full-rate speech codecs deliver a fixed source bit rate of 13 kbps (for EFR some of these bits are redundancy bits, the actual speech coding rate is 12.2 kbps) to the channel encoder. The channel encoder adds redundancy to the source information so that the information is protected against channel errors. The output gross bit rate of the channel encoder which is transmitted over the speech traffic channel is 22.8 kbps. The HR speech codec operates at a source bit rate of 5.6 kbps and a gross bit rate after channel encoding of 11.4 kbps. The FR speech codec was the first GSM speech codec to be standardized. The HR and EFR speech coding standards that followed aimed at providing system capacity and speech quality enhancements, respectively. Each of them was designed to provide a certain level of service quality under some assumed worst case conditions of the channel quality and the traffic load of the radio network. The new AMR standard is based on multi-mode coding with varying source and channel coding rates adapted to the current channel quality. Due to its adaptive nature, AMR overcomes the quality and capacity limitations of the existing GSM speech coders.

Even if AMR was developed to improve the speech operation in the GSM system, due to its flexibility it is also a serious candidate for other systems. The AMR codec has been proposed for the default speech service in UMTS/IMT-2000. The AMR codec also appears to be very suitable for voice-over-internet applications. Depending on the present network load, a higher- or lower-rate speech codec mode may be selected.

This contribution focuses on link adaptation with respect to varying radio conditions and on the inband signaling required for it. Section 2 explains the AMR concept and the dimensions of adaptation in it. Section 3 provides a discussion of link adaptation concepts, which covers different architectural system concepts, conceivable link quality measures and their processing up to the adaptation decision. Section 4 describes the inband transmission of AMR adaptation data and discusses possible alternative approaches. A summary in section 5 concludes the paper.

II. THE AMR CONCEPT

Due to their fixed bit rate operation the existing GSM speech coding standards cannot be optimum for a wide range of transmission conditions. The AMR standard overcomes this problem. In one of its adaptation dimensions it allows for adapting the net or source coding rate. This so-called codec mode adaptation is equivalent...
with adapting the degree of error protection, provided that
the allocated gross bit channel is kept. Codec mode
adaptation is fast, i.e. on speech frame basis, in order to
be able to react on sudden changes of the radio channel
conditions. Codec mode adaptation data has to be
transmitted correspondingly fast. As no sufficiently fast
control channel exists in the GSM system, this kind of
data is transmitted using inband signaling in the speech
traffic channel. The GSM AMR standard also provides
means for adapting the channel mode, i.e. the allocated
gross bit rate. This kind of adaptation allows for trading
speech quality against radio network capacity. Details on
channel mode adaptation are found, e.g., in [2].

The functional operation of AMR codec mode adaptation
is as follows. Incoming speech is source and channel
encoded, applying the currently selected codec and
channel modes. The resulting payload gross bits are
transmitted over the air interface together with codec
mode adaptation data. Transmitted codec mode
information consists of an indication of the presently
applied codec mode and of a codec mode request for the
incoming link. The receiving side detects the codec mode
used based on the received codec mode indication and
applies it for channel and source decoding of the received
payload data. The received codec mode request defines
the codec mode to be used on the outgoing link.
Moreover, the receiving side performs measurements on
the incoming link which lead to link measurement data
and, subsequently, to a codec mode request.

III. AMR LINK ADAPTATION

AMR codec mode adaptation means that, at a given fixed
gross bit rate, the partitioning between source bit rate and
channel coding bit rate can be varied. The AMR speech
codec consists of a number of codec modes with different
source bit rates. For each of the source codec modes there
exist corresponding channel codec modes which map the
source bits onto a fixed number of gross bits. Figure 1
displays a principle sketch of speech quality as a function
of the channel conditions for an example AMR codec
with 3 modes operating on a FR channel. In this example
the FR1 codec mode has the highest source bit rate and
thus the highest speech quality under error-free
conditions. The FR2 and FR3 modes have lower source
bit rates and correspondingly lower quality under error-
free conditions. However, due to its relatively low error
protection, the FR1 codec mode is sensitive to channel
errors and breaks down in channel conditions for which
the FR2 and, particularly, FR3 codec modes still exhibit
robust operation. The FR3 mode is most robust and can
still operate under channel conditions where the other
modes have already broken down. The basic idea of
AMR codec mode adaptation is to measure the present
channel conditions and to select the codec mode that
provides the best speech quality for the given conditions.
Ideally, this allows for achieving a speech quality curve
of the AMR codec that corresponds to the envelope of the
quality curves of the individual codec modes.

![Figure 1: Speech quality vs. channel quality.](image)

Architectural AMR system concepts

There are two fundamental architectural AMR system
concepts, namely distributed and centralized control.
Distributed systems are symmetrical between network
and Mobile Station (MS) side. Each receiving side
performs quality measurements on the incoming link and
sends link adaptation commands to the transmitting side.
Advantages of such a concept are the simplicity due to
the symmetry and that the receiver has access to any
conceivable measurement of the incoming link. This
gives the possibility to leave the measures used for link
adaptation unspecified. There is thus room for improving
link adaptation performance when better link quality
estimation methods are available. Moreover, as the
receiver has the power over and the responsibility for its
own link, any improvement in link adaptation
performance will result in a speech quality improvement
on the own local side. Thus, there is a direct market
advantage for systems/MSs with superior link adaptation
performance. In centralized systems, on the other hand,
the adaptation control both for uplink (UL) and downlink
(DL) resides on the network side. The MS assists the
network in the DL adaptation by conveying link quality
measurements to the network. An advantage with this
architecture is that the network has access to DL
measurements, rather than merely adaptation commands
in the symmetric case. This can be used, e.g. for
improved handover control. Furthermore it might enable
the network to compensate for improper adaptation
performance of early versions of AMR MSs. Moreover,
centralized control is also required for scenarios like
Tandem-Free Operation between two MSs in which
adaptation is done over two air interfaces.

A combination of both architectural concepts is most
appropriate and is specified in the AMR standard. The
system has distributed control as described but the
network is given the possibility to override adaptation
commands from the MS.
Link Quality Measures

As the purpose of AMR codec mode adaptation is to select a codec mode which provides the user with maximum speech quality at a given channel condition, link quality measures are needed which reflect the achievable speech quality for the different codec modes. Based on such a measure, the adaptation unit can select the most appropriate codec mode for the present channel condition, see figure 1. Proper adaptation requires prediction of the measure in order to compensate for the adaptation loop delay. An estimate of the measure is needed for that time instant when the adaptation decision becomes effective.

An important characteristic of link quality measures is the independence of the codec mode. Codec mode independent measures are those which do not require the results of mode dependent processing stages of the receiver, namely of the channel decoder or any subsequent stages. Mode dependent measures, on the other hand, have the potential advantage of being more related to the actual speech quality since they are taken closer to the speech synthesizer. However, as, by definition, they are obtained only for one mode, it is difficult to predict what quality would have been measured in any other mode, in order to make the correct adaptation decision. Furthermore, they are less suitable for bad channels, for which they become unreliable, since likely wrong codec modes are applied for decoding. Examples for mode dependent measures are estimates of the raw bit error rate (BER), the residual BER, or of the frame error rate (FER). Raw BER estimates can be obtained by re-encoding the decoded sequence and comparing with the received gross bit sequence. FER estimates are derived from the CRC of the most important speech bits. The FER measure, which is based on frame error statistics over a number of recently received frames, is very granular which makes it less suitable for prediction. Examples for mode independent measures are Carrier to Interference (CI) ratio estimates or measures derived from the Viterbi metric of the equalizer. They can be obtained burst-by-burst (TDMA burst) and thus provide a better time resolution than mode dependent measures. Independent measures principally allow for each mode to lookup the expected speech quality in tables and to make the adaptation decision on that basis. These measures are suitable also for bad channel conditions since they do not suffer from decoding in a wrong codec mode.

A problem with link quality measures is the possible variance with the channel type. A given value of the measure does not necessarily imply the same degree of speech quality, regardless of the channel type. E.g., at the same CI or raw BER the channel decoder might still work properly for fast fading channels for which some of the bits have strong soft values. In contrast to that, for slow fading channels, for which the received soft bits are equally weak, the channel decoder might already fail. As a consequence, link adaptation should be based not only on the plain link quality measure but also on estimates of the present channel type.

Figure 2: Actual and filtered C/I trajectories for a frequency-hopping channel with shadowing

Link Adaptation Based on C/I Estimates

The AMR standard leaves the link quality estimation open. However, it provides an example solution, which is based on burst-wise C/I estimates [3]. C and I estimates are obtained after the estimation of the channel impulse response prior to equalization. A carrier signal estimate is calculated by convoluting the training sequence with the channel impulse response. The difference of the proper part of the received signal and the carrier signal estimate gives the noise signal estimate. The ratio in dB of the energies of both signals gives a C/I estimate for the present burst.

AMR codec mode adaptation is able to cope with relatively slow channel quality variations as shadowing causes them. Figure 2 shows a typical example of a C/I trajectory of an ideal frequency-hopping channel with shadowing. It becomes obvious that for fast fading channels and particularly for frequency-hopping channels, the instantaneous C/I estimates are highly fluctuating, meaning that they cannot be used for adaptation as they are. Rather, the fluctuations originating from fading need to be removed such that only the slowly varying component resulting from shadowing is preserved. One suitable way for removing the fast fluctuations and performing the prediction of the C/I measure is to apply linear filtering. The filter is designed as a Wiener filter with the design constraint to minimize the estimation error of the shadowing profile at that future
time instant when the adaptation decision becomes effective. A suitable filter length is found to be in the order of 500 ms. As can be seen, the filter output is close to the original shadowing profile and thus suitable to base the adaptation on it.

**Codec Mode Selection**

For codec mode adaptation, the channel quality measure has to be mapped to codec modes. This is in principle done by quantizing the measurements where the levels of the quantizer used represent the different codec modes. However, hysteresis in this mapping is appropriate, as otherwise the selected codec mode might either undesirably fast provided that the channel measurement is close to the decision boundary between two codec modes. Figure 3 shows an example of the mapping from C/I measurements to 4 different codec modes.

![Codec Mode Selection Diagram](image)

Figure 3: Mapping from C/I to Codec Mode

**IV. INBAND TRANSMISSION OF CODEC MODE INFORMATION**

AMR speech codec performance is sensitive to transmission errors of codec mode information. Transmission errors on codec mode indications directly lead to complete frame losses, as the decoder is not able to apply the correct mode for decoding of the received bit stream. Transmission errors on codec mode requests lead to the selection of an inappropriate codec mode and, thus, to non-optimum quality of the decoded speech. On the other hand, as transmission resource for codec mode information is "stolen" from the speech traffic channel, the robustness of speech data transmission is compromised. Hence, bit rate efficient and robust transmission methods for codec mode information are required.

Looking at typical trajectories of the link quality, as shown in figure 2, it is found that, in general, it varies only slowly. Consequently, codec mode changes occur relatively seldom. The entropy of codec mode information is thus low. A suitable transmission method needs to take this characteristic into account.

Two different basic approaches for transmission of codec mode information are continuous and discontinuous transmission. While for continuous methods transmission resource is constantly allocated, discontinuous methods allocate the transmission resource only when required, i.e., in order to signal a codec mode change. A disadvantage with discontinuous transmission is that it needs to be signaled when codec mode information replaces regular payload data. A problem with this technique is the risk of error propagation in the case of transmission errors. Continuous transmission seems thus to be more appropriate.

In order to reduce the amount of codec mode data to be transmitted continuously, codec mode indications and requests are not sent simultaneously but in alternating frames. Even if this constrains the adaptation such that codec mode changes may not occur in every frame, it does not significantly compromise the adaptation performance. One drawback with this technique is, however, that any loss of codec mode information due to transmission errors will affect two consecutive frames.

Following conventional concepts, source coding would be applied prior to channel coding in order to reduce the data rate required for codec mode transmission. However, according to the AMR standard, codec mode adaptation is done with at maximum 4 codec modes. Codec mode information, i.e., codec mode indications or requests are thus represented with 2 bits. Hence, as the source bit rate is that low, there exist no efficient low delay source coding schemes, which keep the adaptation loop delay low. As a consequence, a conventional channel coding scheme would have to protect highly redundant data and thus waste limited transmission capacity. The next section describes a technique, which overcomes this problem.

For channel encoding itself basically two different approaches exist, namely block and convolutional coding. Convolutional coding can provide better protection than simple block coding if the codec mode information is encoded together with the payload speech bits. However, as the codec mode to be used for decoding need to be known prior to decoding, codec mode indications are problematic but not excluded to be convolutionally encoded together with the speech bits. At the cost of higher complexity and the constraint that the codec mode indication is encoded first and with the same code rate for all codec modes, partial decoding could be used in order to figure out the codec mode prior to the final decoding. For simplicity reasons, however, in the AMR standard both codec mode indications and requests are encoded using block codes.
A Priori Knowledge Based Decoding of Codec Mode Information

There exists no suitable source coding scheme which can remove the redundancy of the codec mode information prior to transmission. However, the redundancy can also be exploited at the receiving side. This makes it possible to protect the codec mode information with only weak codes, and consequently, only little transmission resource is required. The redundancy within the received codec mode information can be used as a priori knowledge in a maximum likelihood channel decoder. A priori likelihoods are obtained by applying a simple first order Markov model for the codec mode information. The model is defined by a matrix \( T \) containing the transition probabilities for switching from a given mode at time instant \( n-1 \) to a certain codec mode at time instant \( n \). If, at time instant \( n-1 \), the maximum likelihood decoder has produced a vector \( p(n-1) \) of likelihoods for each code word, then a vector of a priori likelihoods \( p_{\text{prior}}(n) \) for time instant \( n \) for the code words is found by multiplication of \( T \) with \( p(n-1) \). Soft output channel decoding of the received symbols provides a vector \( p_{\text{predicted}}(n) \) of likelihood parameters for each possible code word. Multiplying this vector element-wise with the vector \( p_{\text{prior}}(n) \) leads to the resulting likelihood vector \( p(n) \) containing the likelihoods for each code word. As final decoding result that code word is selected with maximum likelihood.

The Markov probability model can incorporate all statistical knowledge of the codec mode information. Usable knowledge is, e.g., that mode changes occur seldom. The probabilities to switch from one mode to another are small compared to the probability for keeping the mode. Moreover, the adaptation unit is constraint to switch at maximum to neighboring modes. Thus, transition probabilities to non-neighboring modes will be set to 0. The entity sending a codec mode request has further obviously the knowledge of the requested mode. Thus, with taking into account the delay it takes until the request is granted, the decoder can bias the transition probabilities of the Markov model for the mode indications towards the requested mode.

Table 1 gives examples for the performance achieved by a priori knowledge based decoding of codec mode information some simulation results for the GSM AMR Half-Rate speech traffic channel. The table provides error rates of codec mode information encoded with different inband coding methods. Compared are rate 1/4 convolutional coding, rate 1/2 block coding with a priori knowledge based decoding without knowledge of the requested codec mode (A priori 1), and rate 1/2 block coding with a priori knowledge based decoding with knowledge of the requested codec mode (A priori 2). For further comparison, FER figures for the low rate speech codec mode operating at 5.15 kbps are also provided. It is found that the a priori knowledge based block coding method clearly outperforms the convolutional coding scheme even though the latter has half the code rate. Using the knowledge of the requested codec mode gives an additional advantage. However, this knowledge can only be used for the transmission of codec mode indications and provided that the transmitting side grants the mode requests. Comparing the figures with the FER figures of the speech codec mode, it is seen that the codec mode transmission performance is about one order of magnitude superior. It can be concluded that the overall performance will not significantly affect by codec mode transmission errors.

<table>
<thead>
<tr>
<th>C/I [dB]</th>
<th>Conv. 1/4</th>
<th>Block 1/2</th>
<th>A priori 1</th>
<th>Block 1/2</th>
<th>A priori 2</th>
<th>FER MR515</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/I = 1 dB</td>
<td>11%</td>
<td>5.5%</td>
<td>1.3%</td>
<td>1.3%</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>C/I = 4 dB</td>
<td>4.2%</td>
<td>1.8%</td>
<td>0.19%</td>
<td>0.19%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>C/I = 7 dB</td>
<td>0.82%</td>
<td>0.55%</td>
<td>0.06%</td>
<td>0.06%</td>
<td>2%</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Word error performance of codec mode transmission methods

V. SUMMARY

This paper has presented various aspects of the AMR concept, which has been standardized for the GSM system. The different AMR adaptation dimensions were presented and special focus was put on AMR codec mode adaptation. Various concepts and solutions for codec mode adaptation were discussed and the standardized solutions were motivated. Besides the link adaptation itself, the necessary inband transmission techniques were presented which are required for it. In particular a robust channel coding method was described and compared with other conceivable methods. It was shown that the inband transmission technique presented allows for proper AMR operation even under adverse channel conditions.

REFERENCES


[3] GSM 05.09: “Digital cellular telecommunications system (Phase 2+); Link Adaptation”; version 7.0.0, Release 1998