Wireless Personal Communications (2005) 32: 23–41
DOI: 10.1007/s11277-005-7733-2

RObust Header Compression (ROHC) Performance for Multimedia Transmission over 3G/4G Wireless Networks

FRANK H.P. FITZEK[1], STEPHAN REIN[2], PATRICK SEELING[3] and MARTIN REISSLEIN[3]

[1]*Aalborg University, Denmark*
E-mail: ff@kom.aau.dk
[2]*Electronic Measurement and Diagnostics group, Technical University Berlin, Germany*
E-mail: stephan.rein@tu-berlin.de
[3]*Department of Electrical Engineering, Arizona State University, Goldwater Center MC 5706, Tempe AZ 85287-5706*
E-mails: patrick.seeling@asu.edu, reisslein@asu.edu, Web: http://www.fulton.asu.edu/∼mre

Abstract. RObust Header Compression (ROHC) has recently been proposed to reduce the large protocol header overhead when transmitting voice and other continuous media over IP based protocol stacks in wireless networks. In this paper we evaluate the real-time transmission of GSM encoded voice and H.26L encoded video with ROHC over a wireless link. For the voice transmission we examine the impact of ROHC on the consumed bandwidth, the voice quality, and the delay jitter in the voice signal. We find that for a wide range of error probabilities on the wireless link, ROHC roughly cuts the bandwidth required for the transmission of GSM encoded voice in half. In addition, ROHC improves the voice quality compared to transmissions without ROHC, especially for large bit error probabilities on the wireless link. The improvement reaches 0.26 on the 5-point Mean Opinion Score for a bit error probability of 10^{-3}. For the video transmission we examine the impact of ROHC on the consumed bandwidth. We find that the bandwidth savings with ROHC depend on the quantization scale used for the video encoding and the video content and ranges between 5–40% for typical scenarios.

Keywords: bandwidth savings, Internet Protocol, robust header compression, subjective voice quality, video transmission, wireless links

1. Introduction

Third and fourth generation wireless systems are being designed to support a wide range of services, including audio and video applications. This service flexibility is achieved by relying on the Internet Protocol (IP), typically in conjunction with the User Datagram Protocol (UDP) and the Real Time Protocol (RTP). One major problem with the IP based protocol architectures is the large overhead, which affects the limited bandwidth of wireless channels. A low bit rate speech application can result in IP packets with a ratio of 30 bytes of payload to 60 bytes of overhead. Recently, RObust Header Compression (ROHC) [1] has been proposed to compress the protocol headers for packet transmission over a wireless link.

In this paper we present extensive performance results for the wireless transmission of multimedia with ROHC. In particular, we consider the real-time wireless transmission of GSM encoded voice and H.26L encoded video with ROHC. For the voice transmission with ROHC our evaluation metrics are the bandwidth compression gain (reduction in header and total packet size), the voice quality, and the delay jitter. Importantly, our employed voice quality evaluation methodology considers a wide array of objective voice quality metrics, including both the traditional and the segmental signal to noise ratio (SNR), spectral distance metrics, and parametric distance metrics. The considered parametric distance metrics include the cepstral distance metric, which can be transformed into the mean opinion score (MOS), thus enabling us to quantify the effect of ROHC on the voice quality in terms of the MOS.

We find in our evaluation that for a wide range of bit error probabilities on the wireless link, ROHC reduces the protocol overhead for voice transmission with IPv4 by approximately 85%, which reduces the bandwidth required for a GSM coded voice transmission by about 47%. On top of these bandwidth savings, ROHC improves the voice quality. On the 5-point MOS scale the improvement increases roughly exponentially with the bit error probability. The improvement is about 0.028 for an error probability of $10^{-4.5}$ and reaches 0.134 and 0.264 as the error probability increases to $10^{-3.6}$ and 10^{-3}. We also find that ROHC slightly increases the jitter for small error probabilities and slightly reduces the jitter for large error probabilities. For wireless video transmission with ROHC we find that the bandwidth savings generally vary with the video content and increase with increasing quantization scale in the video encoding. The total bandwidth savings are typically in the range from 5–40%.

This paper is organized as follows. In the following subsection we review related work. In Section 2 we describe the principles and integration of ROHC in the IP protocol stack. In Section 3 we describe our evaluation methodology. In Section 4 we present the ROHC performance results for wireless voice transmission, i.e., the bandwidth reduction, objective voice quality, and delay jitter results. In Section 5 we present the bandwidth reduction results for wireless video transmission with ROHC. In Section 6 we summarize our contributions.

1.1. Related Work

There exists a large body of literature on the development of header compression schemes for wireless networks and on the evaluation of these schemes in terms of the network metrics of throughput, packet delay, and packet jitter. We now briefly review this literature and refer to [2] for a more detailed survey. Following the seminal work [3] which proposed a header compression scheme for wireline links, a number of studies have examined the problem of header compression for error prone links, see for instance [4, 5]. In [6] a header compression scheme that provided robustness at the expense of reduced compression was introduced and further optimized in [7, 8]. On the other hand, in [9] a number of extensions to the seminal scheme [3] were proposed; these extensions provide support for UDP, IPv6, and additional TCP features and incorporate further refined compression mechanisms [10]; a performance evaluation for packet voice over the wired internet is reported in [11]. We also note that a header compression scheme specifically for IPv6 based communication with mobile wireless clients has been developed [12].

Building on the insights from the preceding research efforts, the compressed real time protocol (CRTP) [13] was developed, primarily to compresses the RTP/UDP/IP headers for real-time multimedia transmission, and evaluated in [14] for wireless links. Robust checksum–based compression (ROCCO) [15] is a refinement of CRTP, which improves the header compression performance for highly error–prone links and long round trip times [16]. Similarly, Enhanced Compressed RTP (ECRTP) [17] is a refinement of CRTP.

ROHC is envisioned as an extensible framework for robust and efficient header compression over highly error–prone links with long round–trip times. This design is motivated by the large bit error rates (typically on the order of 10^{-4}–10^{-2}) and long round trip times (typically 100–200 msec) of cellular networks. The design of ROHC is based on the lessons learned from the header compression schemes reviewed above. In particular, ROHC incorporates elements from ROCCO and Adaptive Header Compression (ACE) [18], which may be viewed as preliminary forms of ROHC. The ROHC impact on packet loss and delay has been studied in [19–24].

The existing evaluations of header compression schemes have primarily focused on "network metrics", such as packet loss, delay and jitter, whereas the impact of the header compression on the quality of the transmitted medium (e.g., the received voice quality) has received very little attention so far. The only study in this direction that we are aware of is [16]. In [16] the objective speech quality degradation is studied (using the traditional SNR which has only a weak correlation with user perception) for ROCCO and CRTP. In contrast, in this paper we consider the state-of-the-art ROHC compression scheme and evaluate the voice quality using a thorough evaluation methodology that was developed in [25]. This methodology employs an array of objective metrics that allow accurate predictions of the subjective voice quality of hearing tests. We note that [25] provides a few voice quality results for ROHC, which are mainly intended to illustrate the interpretation of the voice quality results obtained with the methodology explained in [25]. (We also note that a preliminary form of the voice quality evaluation methodology was presented in [26].) In contrast, in this paper we provide a comprehensive set of performance results for the voice transmission with ROHC including not only detailed voice quality results, but also bandwidth compression and delay jitter results. In addition, we provide bandwidth compression results for the transmission of H.26L compressed video (of which preliminary results were presented in [27]) and which we contrast with the results for voice transmission in this paper.

2. Overview of ROBust Header Compression

A multimedia stream packet composed for an IP network transmission typically consists of a 20 byte IP header, an 8 byte UDP header, and a 12 byte RTP header, as shown in Figure 1. The

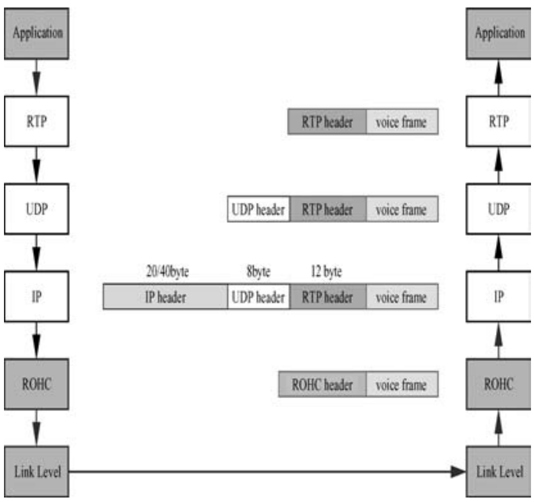


Figure 1. The 40 byte RTP/UDP/IPv4 header (or 60 byte RTP/UDP/IPv6 header) is compressed to a smaller ROHC header.

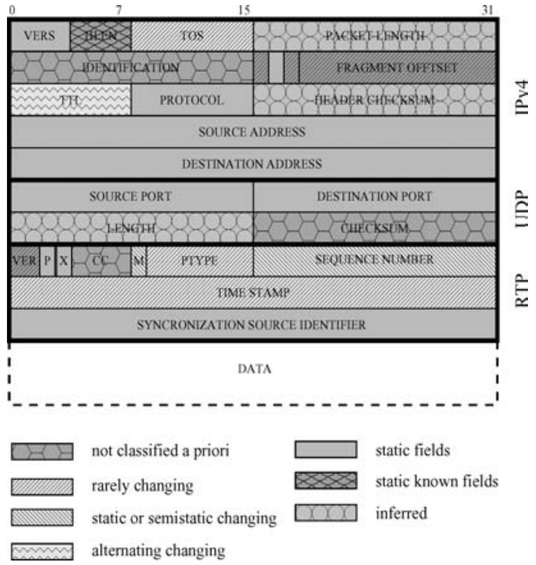


Figure 2. Header fields for RTP/UDP/IP packets (Version 4) with the appropriate dynamics.

IPv6 version requires a 40 byte IP header, so the total RTP/UDP/IP header size can sum up to 60 bytes. A speech application generates compressed data at a low bit rate of around 13 kbit/s. Considering a typical payload smaller than 40 bytes, the ratio of header size to payload results in an significant waste of link bandwidth. The ROHC compressor replaces the RTP/UDP/IP headers by its own, much smaller header. On the receiver side the decompressor transforms the ROHC header into the original protocol layer headers.

Many IP header fields of a given data flow are static, i.e., do never change, see Figure 2. ROHC stores the values of these static header fields as *static context* at the decompressor. More challenging for the compression scheme is the treatment of the changing (dynamic) fields in the IP header. ROHC uses linear functions based on the packets' sequence numbers to derive the values of the dynamic header fields, see [1] for details. The parameters characterizing these linear functions are stored and updated as so-called *full context* at the decompressor. We also note that some IP header fields can be derived (inferred) from other header fields at the decompressor and are therefore not transmitted once the static context is established.

The fundamental challenge in header compression for transmission over wireless links is to maintain the correct context at the decompressor in the face of quite frequent bit errors in the received packets. ROHC supports three different modes for maintaining the context in different wireless systems. The *unidirectional mode* is designed for systems without a feedback channel from the decompressor to the compressor, i.e., where the decompressor can not acknowledge the correct receipt of context information. To overcome this limitation, the compressor periodically retransmits the context information. The *bidirectional optimistic mode* and the *bidirectional reliable mode* are designed for systems with a feedback channel from the decompressor to the compressor, i.e., where the decompressor can acknowledge the correct receipt of context information and/or send negative acknowledgements to request the retransmission of context information. With the bidirectional optimistic mode, bit errors in the compressed header are detected with a 3-bit cyclic redundancy check (CRC) code. When the CRC check fails the decompressor generally discards the affected packet and attempts to repair its context either locally or by requesting a context update from the compressor. The reliable mode extends the optimistic mode by a more complex error detection and correction which uses a larger number of coding bits. Our evaluation concentrates on the optimistic mode, since it gives generally the best compression efficiency. Also, with the results of the optimistic mode, it is possible to predict the results for the reliable mode.

In the commonly used RTP/UDP/IP protocol suite, ROHC is installed between the network and the link layer, see Figure 1. In the third generation Universal Mobile Telecommunication System (UMTS), ROHC compressor and decompressor are part of the UMTS mobile phone and the corresponding UMTS radio network controller (RNC). (Other solutions are possible, but ROHC always resides above the link layer.) The other Internet components do not notice the usage of a compression scheme, but the wireless service provider can take advantage of a significant reduction of the required bandwidth, as demonstrated by our results in Sections 4 and 5. ROHC requires from the link layer that the packets are sent in a strictly sequential order. Also the packets are not allowed to contain routing information (single hop restriction).

3. Evaluation Methodology

The ROHC measurements were conducted on a testbed consisting of two Linux PCs. The Linux kernels had been enhanced by an ROHC implementation (provided by the acticom GmbH, www.acticom.de). For the evaluation of the voice transmission we used three different voice files (track 49, track 53, and track 54) of the sound quality assessment material from the European Broadcasting Union [28], as shown in Table 1. The files, given in the wave file mono format, are first down sampled to 8 kHz and then transferred from one PC over a simulated wireless link to the other PC. On the sender's side the wave file is GSM encoded using the full-rate GSM encoder. The GSM coder operates on speech frames of 20 msec, i.e., 160 samples (each 8 bits worth of uncompressed voice data), which are encoded into 33 byte GSM frames. For the real-time transmission, each individual GSM frame is packaged into an RTP packet and passed down the protocol stack through the UDP and IP layers to the ROHC and link layers, as illustrated in Figure 1. The two Linux machines are connected by an unloaded Ethernet network. Recent channel characterization studies [29] and ROHC evaluation studies [23] have revealed that uncorrelated bit errors give a good approximation of the error process experienced by the ROHC layer in 3G networks. Consequently, we simulate uncorrelated bit errors on the link layer. We use nine different bit error probabilities ranging from 10^{-6} to 10^{-3}. On the receiver side, the packet, which may or may not have suffered bit errors in transit, passes up through the protocol stack to the GSM decoder (we do not employ any packet retransmissions in the measurement set-up). If the compressed header suffered bit errors and the ROHC CRC check fails, then the packet is discarded. If bit errors in the compressed header are not detected by the CRC check and/or if only the payload suffered bit errors, then the header is decompressed and the IP packet is handed up further the protocol stack. (If there

Table 1. Test material for voice quality evaluation

File	Text	Language	Speaker	Duration [sec]	Total size (kB)	GSM size (kB)
49.wav	A	English	Female	19.15	306.45	31.6
53.wav	B	German	Female	16.64	266.21	27.46
54.wav	B	German	Male	16.79	268.72	27.72

were undetected bit errors in the header the decompression may give invalid header values, resulting in the packet being discarded.) Finally, the GSM decoder decompresses the received GSM frame into 160 audio samples. Note that bit errors in the received GSM frame may result in decoded samples that differ from the original voice samples, i.e., distortion of the speech.

We assess the voice quality using the voice quality evaluation methodology developed in [25]. This methodology employs an array of *objective* voice quality metrics, i.e., metrics that can be mathematically calculated, to predict the *subjective* voice quality, i.e., the voice quality that would be the outcome of extensive listening experiments with human subjects. More specifically, the methodology employs three families of objective metrics, namely signal-to-noise ratio (SNR) metrics (in particular the traditional SNR and the segmental SNR), spectral distance metrics (in particular the inverse linear unweighted distance, the unweighted delta form, and the logarithmic root mean square), and parametric distance metrics (in particular the log area ratio, the energy ratio, the logarithmic likelihood, and the cepstral distance). The SNR metrics are positively correlated with the voice quality, i.e., larger SNR values indicate higher voice quality, whereas the other metrics are negatively correlated with the voice quality, i.e., smaller metric values indicate higher voice quality. Importantly, the spectral distance and parametric distance metrics are generally more highly correlated with the voice quality and give therefore more reliable estimates of the subjective voice quality. The voice evaluation methodology first synchronizes the transmitted and the received voice signals using a cross correlation approach and evaluates the absolute values for the individual quality metrics. Next, the improvement (gain) in voice quality from the communication system without ROHC to the communication system with ROHC is evaluated by evaluating the differences between the absolute metric values and transforming the difference into a decibels (dB) scale. The gain obtained with the cepstral distance metric is then transformed into the gain terms of the 5-point mean opinion score, given in Table 2. In a final step, the consistency of the voice quality results is validated with a scatter plot. The scatter plot is obtained by expressing the cepstral distance as linear functions of the other metrics and plotting the cepstral distance values estimated with the linear functions as a function of the actual cepstral distance value.

To assess the interplay between the ROHC and audio application independent of the IP/UDP/RTP layers in-between, we disabled the UDP checksum in our measurements and did not employ the RTP timestamps and sequence numbers for the signal synchronization. Instead we employed the cross correlation synchronization conducted as part of the voice quality evaluation.

For the evaluation of the video transmission we encoded the widely used video reference sequences, such as *foreman*, *container*, *akiyo*, and *silent* as well as our own productions *bridge close* and *highway* [30], see Table 3, in the QCIF format (176 $\times$ 144 pixel) with the H.26L encoder [31, 32] using the IBBPBBPBBPBB Group of Pictures (GoP) structure. The

Table 2. Mean opinion score

MOS	
5	Imperceptible
4	Just perceptible but not annoying
3	Perceptible and slightly annoying
2	Annoying but not objectionable
1	Very annoying and objectionable

quantization scale is set to 30 for all streams, which corresponds to a relatively high quality level as per our visual assessment. Furthermore we encoded the *container* sequence with different quantization scales.

We repeat each transmission experiment numerous times with independent bit errors to obtain 95% confidence intervals that are smaller than 10% of the corresponding sample mean for all performance metrics.

4. Performance Results for Voice Transmission

In this section we present the results of our evaluations of voice transmission with ROHC. We first evaluate the compression performance (bandwidth reduction) achieved by ROHC. Next, we employ the outlined voice quality evaluation methodology to assess the impact of ROHC on the voice quality. Finally, we evaluate the impact of ROHC on the jitter in the voice signal.

4.1. Bandwidth Compression Gain

To assess the bandwidth compression achieved by ROHC, we consider both the compression gain for the packet header and for the total packet. Let H denote the size of the packet header (in byte) without compression and let L denote the size of the packet payload (in byte). Furthermore, we let H_{ROHC} denote the average size of the compressed header (in byte). We define the *header compression gain* G^H as the relative reduction of the header size, i.e., as

$$G^H = \frac{H - H_{\mathrm{ROHC}}}{H} = 1 - \frac{H_{\mathrm{ROHC}}}{H}. \tag{1}$$

We define the *packet compression gain* G^P as the relative reduction of the total packet size, i.e., as

$$G^P = \frac{(H + L) - (H_{\mathrm{ROHC}} + L)}{H + L} = \frac{H - H_{\mathrm{ROHC}}}{H + L}. \tag{2}$$

To assess the maximum compression gain (packet size reduction) with header compression we consider an ideal compression scheme that reduces the header size to $H_{\mathrm{ROHC}} = 0$ bytes. Clearly, such an ideal compression scheme has a packet compression gain, i.e., reduces the bandwidth required for the voice transmission by

$$G^P_{\max} = \frac{H}{H + L}. \tag{3}$$

With a GSM codec generating $L = 33$ byte voice frames, the maximum saving potential is 55% when using IPv4 (which has $H = 40$ bytes of overhead) and grows to 65% when using IPv6 (which has $H = 60$ bytes of overhead). As the overhead for a given version of IP is constant, the maximum saving with compression increases as the payload size decreases. Therefore ROHC is well suited for low bit rate voice streams, where the header size is typically larger than the payload.

In Figure 3 we plot the actually achieved header compression gain G^H for IPv4 with ROHC as a function of the wireless bit error probability for the three different voice tracks used in our performance evaluation. For track 49 we plot the 95% confidence intervals, but not for the

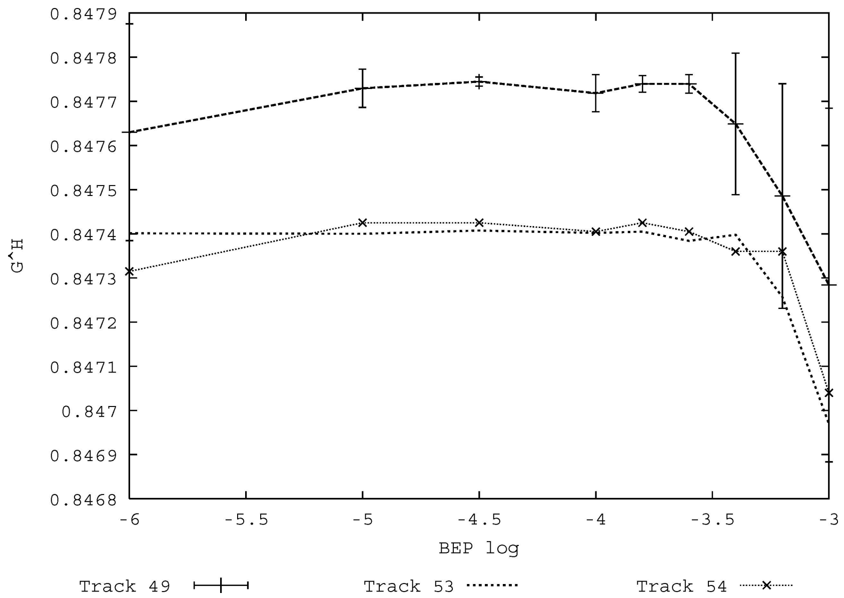


Figure 3. Header compression gain G^H as a function of bit error probability on wireless link for the three voice tracks.

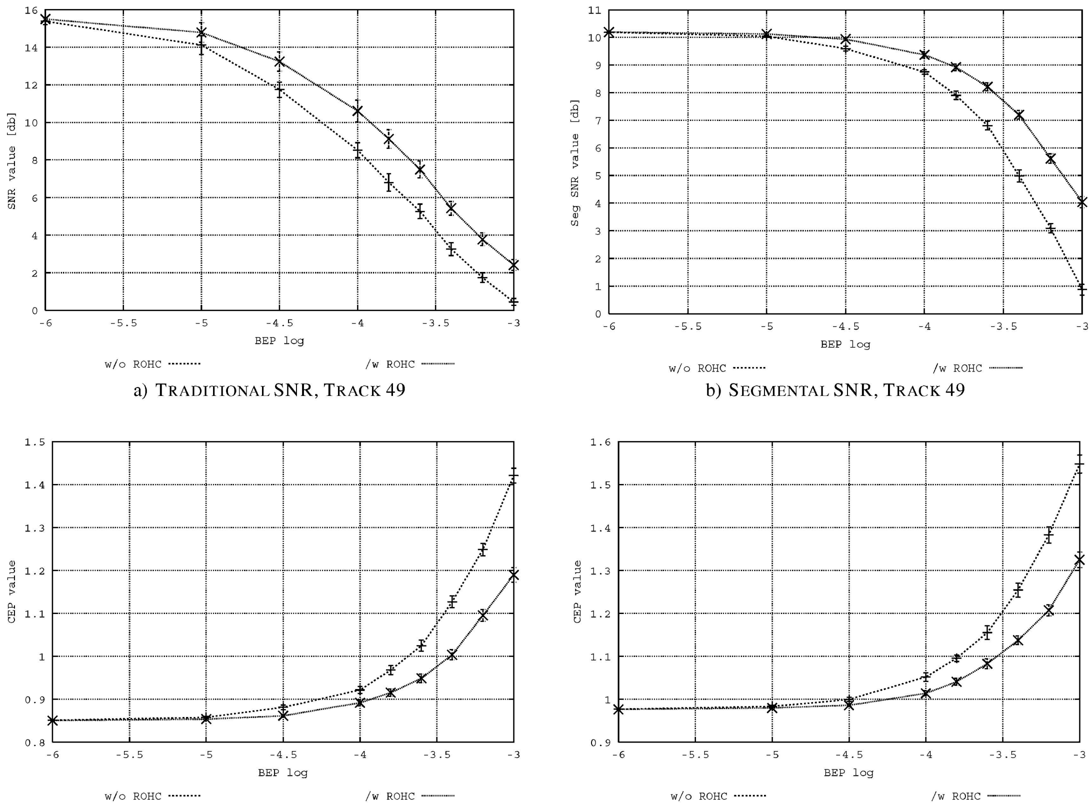
other tracks to avoid visual clutter. We observe that the header compression gain is slightly above 84.7 % for all tracks for the entire range of considered error probabilities from 10^{-6} to 10^{-3}. Upon closer inspection we observe that the header compression gain is essentially constant in the bit error probability range from 10^{-6} to $10^{-3.5}$ and then tends to drop off as the bit error probability increases to 10^{-3}. This is because the more frequent bit errors tend to disrupt the context at the decompressor, forcing ROHC to re-establish the context for the compression. Overall, this result indicates that the header compression is relatively robust in the sense of being able to maintain a constant header compression gain in the face of different bit error probabilities that make it more or less difficult to maintain the correct context at the decompressor. This observed header compression gain implies that the IPv4 header size is in the long run average reduced from 40 to approximately 6 bytes. The compression gain for the total RTP/UDP/IP packet with a payload of 33 bytes is calculated as approximately

$$G^P = \frac{H - H_{\text{ROHC}}}{H + L} = \frac{(40-6)\text{ bytes}}{(40+33)\text{ bytes}} = 0.47.$$

This actual compression gain of 47% for the total IP packet is close to the maximum gain of $G^P_{\max} = 55\%$. Next we address the question whether this significant reduction in consumed bandwidth affects the voice quality.

4.2. VOICE QUALITY

In Figure 4 we plot the voice quality as a function of the logarithm with base 10 of the wireless bit error probability for different tracks. We plot the 95% of the confidence intervals of different objective voice quality metrics. We observe that the SNR based metrics generally decrease with increasing bit error probability, while the cepstral distance metric increases with increasing bit error probability. Keeping in mind that the SNR based metrics are positively



a) TRADITIONAL SNR, TRACK 49

b) SEGMENTAL SNR, TRACK 49

c) CEPSTRAL DISTANCE, TRACK 49

d) CEPSTRAL DISTANCE, TRACK 53

Figure 4. 95% confidence intervals for voice quality without and with ROHC as a function of wireless bit error probability for different objective voice quality metrics and tracks.

correlated with the voice quality, while the cepstral distance metric is negatively correlated with the voice quality, these results indicate that the voice quality generally decreases with increasing bit error probability. Importantly, we observe that the SNR qualities with ROHC are above the SNR qualities without ROHC and the cepstral distance values with ROHC are below the values without ROHC. This indicates that the voice quality with ROHC is higher than the voice quality without ROHC. The differences are especially pronounced in the range from 10^{-4} to 10^{-3} and in this range the differences between the results without ROHC and with ROHC are significantly larger than the confidence intervals, indicating that the voice quality improvement with ROHC is statistically significant.

For additional statistically reliability we average the results for the three tracks in all subsequent investigations.

Following the methodology developed in [25] we proceed to evaluate the impact of ROHC on the voice quality by considering the gain in dB in the voice quality for the set of objective quality metrics. In Figures 5(a)–(c), we plot the gain (in dB) as a function of the logarithm with base 10 of the bit error probability on the wireless link. We observe that all gain metrics indicate an increasing positive gain (i.e., improvement in voice quality) with increasing error probabilities. As an exception, the gain for the traditional SNR decreases for bit error probabilities above $10^{-3.8}$. This is due to the shortcomings of the traditional SNR metric which applies unequal weighing of soft and loud voice frames. Overall, the SNR measures indicate a gain

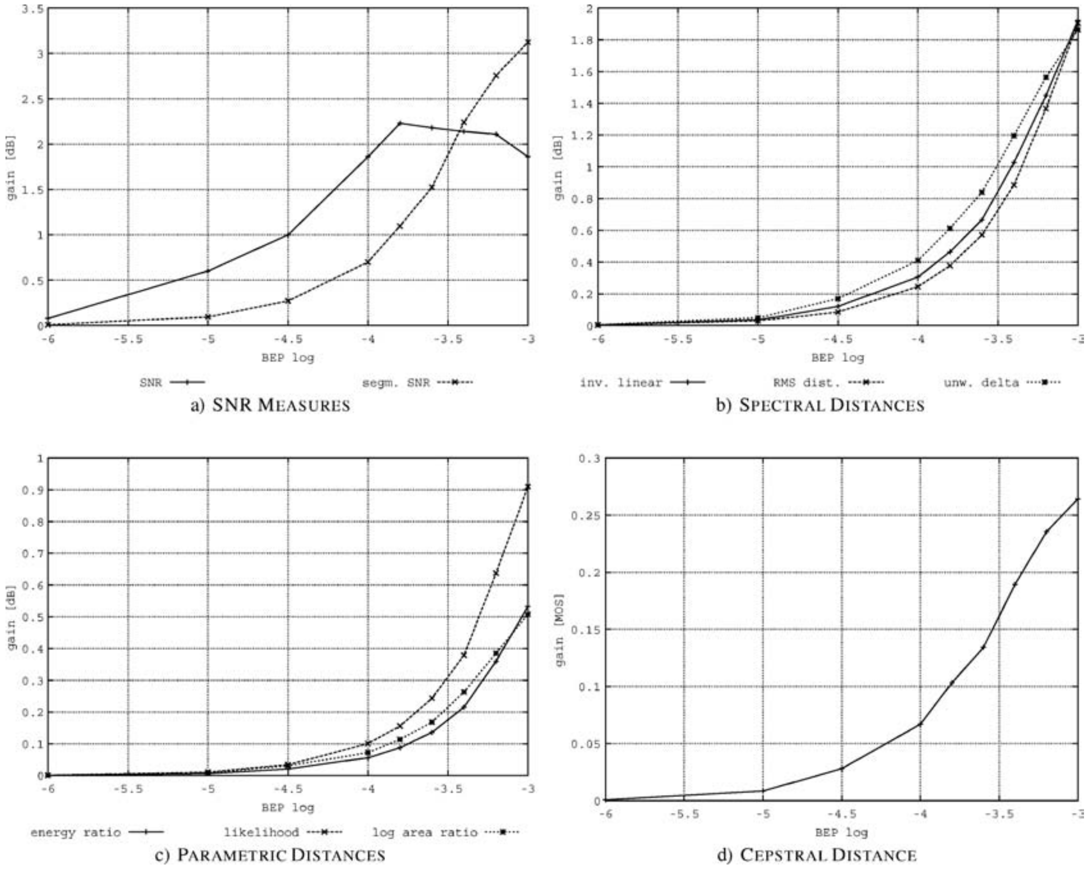


Figure 5. Gain in voice quality with ROHC as a function of bit error probability.

between two and three decibels for link error probabilities in the $10^{-3.4}$ to 10^{-3} range. Similarly, the spectral distances indicate gains between 0.02 and 2 dB for link error probabilities of 10^{-3} and the parametric distances give gains between 0.5 and 1 dB. Overall, these results indicate that the voice quality does not suffer from header compression. On the contrary, it is improved, especially for large bit error probabilities on the wireless link. Note that these gain values in dB represent the improvement in terms of objective voice quality and not in terms of user perception.

In order to asses the impact on the user perception we investigate the voice quality improvements on the subjective 5-point MOS scale by transforming the values of the cepstral distance to the predicted MOS. We plot the gain in terms of the MOS as a function of the bit error probability in Figure 5(d) and observe that the predicted gain for ROHC increases roughly exponentially with increasing error probability and reaches 0.26 for error probabilities of 10^{-3}.

As noted in [25] in objective voice quality evaluation it is generally advisable to consider a variety of metrics since each individual metric (including the cepstral distance which is used to obtain the MOS) has been evaluated for a limited set of distortions. Following the employed voice quality evaluation methodology we examine the consistency of the obtained objective voice quality results with the so-called scatter plot. The scatter plot is obtained by expressing the cepstral distance as linear functions (mappings) of each of the other spectral

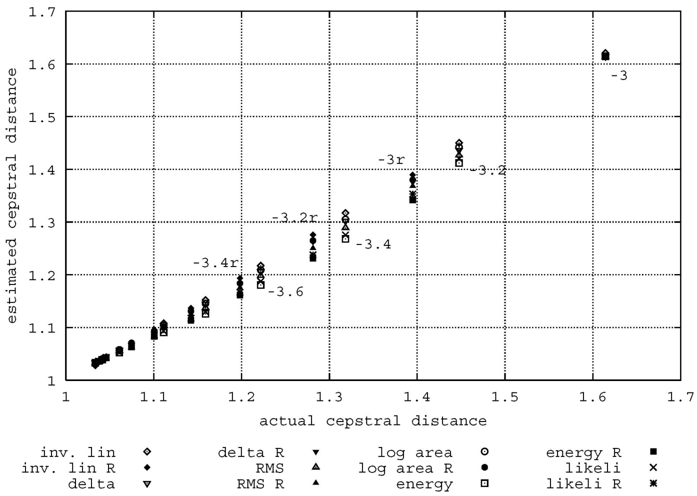


Figure 6. Scatter plot of cepstral distance estimated from linear mappings of the other voice quality metrics as a function of actual cepstral distance.

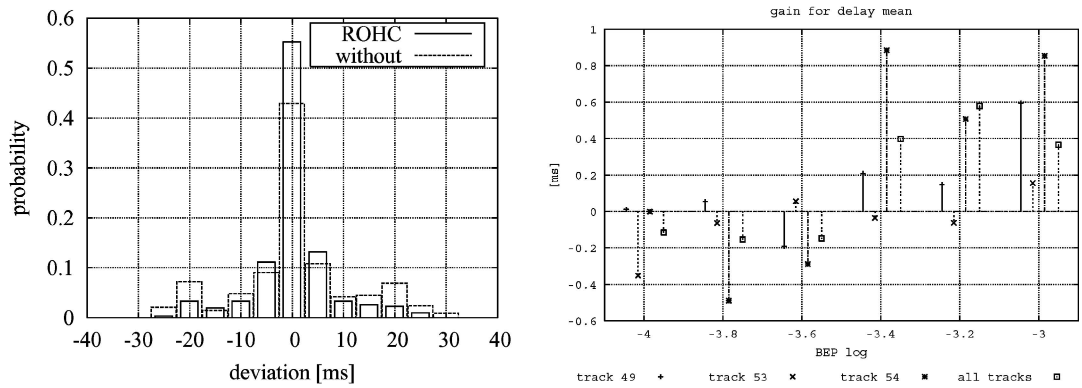
distance metrics (inverse linear unweighted distance, unweighted delta form, and logarithmic root mean square) and parametric distance metrics (log area ratio, energy ratio, and logarithmic likelihood) of the voice quality, and then plotting the cepstral distances estimated from the linear mappings as functions of the actual cepstral distance. The thus obtained scatter plots for voice transmission without ROHC and with ROHC are given in Figure 6. In the figure the filled (shaded) symbols also identified by R in the legend of the figure correspond to the qualities with ROHC. The unfilled symbols correspond to the qualities without ROHC. We observe that the points are fairly closely scattered around a straight line with slope one. This indicates that there is a consistently high correlation between the voice qualities obtained with the other metrics and the cepstral distance, which validates the gains in the voice quality reported above.

Overall the consistent improvements in the voice quality indicate that the effect of smaller packets with ROHC and thus the smaller probability of suffering bit errors dominates over the effect of packet headers that are incorrectly uncompressed due to incorrect context at the decompressor.

4.3. Delay Jitter Results

The voice quality metrics considered in the preceding section do not capture the signal delays. Therefore, we investigate the delay, or more precisely, the delay variation (jitter) separately in this section. Recall that the employed voice quality evaluation methodology performs delay corrections to the received (distorted) voice signal before evaluating the voice quality metrics. The amount of these delay corrections gives the delay jitter in the voice signal.

We examine both the delay jitter histogram and the standard deviation of the delay jitter. Figure 7 a) shows a typical histogram of delay jitter for the bit error probability 10^{-3}. Each bar represents a delay jitter range of 5 msec. (The bars of ROHC are slightly thinner to avoid overlap in the plots.) We observe from the histogram that the jitter – if at all present – is typically quite small. A jitter on the order of 20 msec, which would occur as a result of a packet discard



a) TYPICAL DELAY JITTER HISTOGRAM. THE PROBABILITY OF A DELAY BETWEEN -2.5 AND $+2.5$ MSEC IS HIGHER FOR ROHC TRANSMISSIONS.

b) JITTER GAIN AS A FUNCTION OF BIT ERROR PROBABILITIES: NEGATIVE GAIN (I.E., LARGER JITTER) FOR ERROR PROBABILITIES $10^{-4} \ldots 10^{-3.6}$, POSITIVE GAIN (I.E., SMALLER JITTER) FOR $10^{-3.4} \ldots 10^{-3}$.

Figure 7. Delay jitter results.

is quite rare. The small jitter values are mainly due to varying processing delays in the end systems in our experimental set-up (in particular due to the ROHC and the GSM codec) and signal alignments to correct for severe distortions in audio samples at the edges of frames. Figure 7 b) depicts the ROHC gain for jitter (i.e., reduction in delay standard variation). For the small bit error probabilities up to $10^{-3.6}$ there is a loss of around 5 msec. Track 54 is mainly responsible for the loss, for all other tracks ROHC mostly causes a gain. The observed larger jitter for Track 54 is mainly due to the delay variations caused by ROHC, which dominate over the variations due to the GSM codec for these small error probabilities. For the larger bit error probabilities in the range from $10^{-3.4}$ to 10^{-3}, there is a gain between 0 and 10 msec for the average of all tracks. This is primarily due to delay variations caused by the GSM codec, which dominate over the ROHC delay variations for these larger error probabilities. Overall, our results indicate that ROHC does not significantly deteriorate the delay jitter.

5. Performance Results for Video Transmission

In this section we present the bandwidth compression results of ROHC for wireless video transmission. A fundamental difference between the voice transmission and the video transmission is that the voice frames have a fixed payload size L whereas the video frames have variable size and thus result in variable size payloads. In particular, the frame sizes vary over time depending on the video content as well as the used encoder and its settings (quantization scale) [33]. Formally we let n, $1, \ldots, N$, denote the frame index and assume that each frame forms the payload of one packet (which is motivated by the real-time transmission requirement to immediately send each generated frame). We define the packet compression gain for packet n as the relative reduction of the size of packet n due to the header compression, i.e.,

$$G^P(n) = \frac{(H + L(n)) - (H_{\text{ROHC}} + L(n))}{H + L(n)} = \frac{H - H_{\text{ROHC}}}{H + L(n)}. \tag{4}$$

We define the average packet compression gain for a sequence of packets (video frames) $n, n = 1, \ldots, N$, as the average of the packet compression gains for the individual packets, i.e., as

$$\bar{G}^P = \frac{1}{N}\sum_{n=1}^{N} G^P(n). \tag{5}$$

We define the maximum packet compression gain for packet n as the relative reduction in packet size achieved by an ideal header compression scheme that compresses the original header down to $H_{\text{ROHC}} = 0$ bytes, i.e.,

$$G^P_{\max}(n) = \frac{H}{H + L(n)}. \tag{6}$$

We define the average maximum packet compression gain for a sequence of packets (video frames) $n, n = 1, \ldots, N$, as

$$\bar{G}^P_{\max} = \frac{1}{N}\sum_{n=1}^{N} G^P_{\max}(n). \tag{7}$$

The average size of the compressed header H_{ROHC} and the header compression gain G^H are defined as for the voice transmission. Before we proceed it is in order to clarify a subtlety relating to the averaging in the defined performance metrics. The defined average compressed header size R_{ROHC} is obtained by averaging the sizes of the compressed headers of the packets carrying a given voice stream (with fixed payload size) or video stream (with variable payload size). We did not explicitly write out this averaging procedure in mathematical equations to keep the notation simple. On the other hand, to highlight the impact of the variable payload sizes in video transmission we explicitly wrote out the averaging over the video frame sizes as detailed above.

In Table 3 we report the average maximum packet compression gains $\bar{G}^P_{\max}$ for the considered video sequences. A number of observations are in order. First, the video sequences differ in the average maximum compression gains (saving potentials). In contrast to voice services, the savings for video services depend on the content. As noted above the potential savings for IPv6 are larger than for IPv4 because of the IP header length (IPv4 with 40 byte and IPv6 with 60 byte). We also observe that the savings potentials are larger for the lower bite rate video streams. This is because the smaller the IP packets, the higher the impact of the IP overhead and therefore the larger the saving potential $\bar{G}^P_{\max}$. The largest savings potential is obtained for video encoded with large quantization scales and transported with IPv6 with a savings potential of up to 61%. For the video encodings with a fixed quantization scale of 30, the savings potential varies from 6.6% to 24.5%.

Next, we conducted measurement experiments of the actual video transmission using the same set-up as for the voice transmission experiments. We fixed the wireless bit error rate at 10^{-5} for these experiments. In Table 3 we report the achieved header compression gain G^H and the achieved average packet compression gain $\bar{G}^P$. We observe that the achieved header compression gain is approximately 84% for all video sequences and quantization scales. This corresponds to a compression of the 40 byte IPv4 headers to a ROHC header with an average size of approximately 6 byte, which is very similar to the results for the

Table 3. Average maximum packet compression gains (avg. max. comp. gain), average packet compression gains (avg. comp. gain), and header compression gains (head. comp. gain) for H.26L encoded video

Video	Qu. scale	Avg. bit rate (kbps)	Avg. max comp. gain $\bar{G}^P_{\max}$ IPv4	Avg. max comp. gain $\bar{G}^P_{\max}$ IPv6	Head. comp. gain G^H IPv4	Avg. comp. gain $\bar{G}^P$ IPv4	Comp. head. size H_{ROHC} (byte)
Container	10	855.0	1.1	1.6	84.2	0.93	6.37
Container	20	213.0	4.3	6.3	83.7	3.58	6.52
Container	30	65.8	12.7	17.9	84.3	10.5	6.29
Container	40	24.1	28.5	37.5	84.3	23.0	6.27
Container	51	9.1	51.0	61.0	83.7	40.0	6.51
Bridge close	30	69.9	10.3	14.6	84.5	8.6	6.19
Carphone	30	135.4	6.6	9.6	83.9	5.4	6.44
Claire	30	44.3	17.8	24.5	84.6	14.7	6.19
Foreman	30	121.9	7.28	10.5	83.9	6.0	6.44
Grandma	30	56.4	14.5	20.3	84.5	12.0	6.19
Highway	30	57.2	12.3	17.3	84.3	10.2	6.29
Mother	30	66.4	12.6	17.8	84.6	10.5	6.16
News	30	100.9	8.7	12.5	83.6	7.1	6.55
Salesman	30	81.5	10.5	15.0	84.5	8.8	6.21
Silent	30	101.7	8.6	12.4	84.3	7.2	6.28

voice transmission. While the achieved header compression gain is nearly constant when comparing the different videos and quantization scales, the average achieved packet compression gain varies significantly for the different video sequences and quantization scales. We observe that the actually achieved compression gains are only slightly smaller than the average maximum compression gains. The average achieved total compression gain for the investigated video sequences of quality level 30 ranges from 5.4% to 14.7%. For a lower video quality the compression gain reaches 40% for the *container* sequence. Overall, we observe that the achievable compression gain increases as the average bit rate of the compressed video decreases, i.e., for larger quantization scales and for video content that is more amenable to compression (e.g., sequences with a low level of motion such as *claire* and *grandma*).

6. Conclusions

We have examined the performance of RObust Header Compression (ROHC), an IP header compression scheme, for voice and video applications in 3rd and 4th generation mobile networks.

Our evaluations of RObust Header Compression (ROHC) for wireless voice transmission indicate that with ROHC the RTP/UDP/IPv4 header size is reduced by approximately 85%, which for the considered GSM encoded voice with 33 byte GSM frames cuts the total bandwidth required for the voice transmission almost in half. This reduction of the total bandwidth is expected to be even larger for lower bit rate encoders with smaller voice frames. Our extensive voice quality evaluations, which employ a thorough objective voice quality evaluation methodology indicate that this enormous reduction in used bandwidth does not deteriorate the voice quality. On the contrary, the voice quality is improved by ROHC. All of the considered parametric and spectral distances indicate improvements in the objective voice quality. In addition, the cepstral distance predicts a subjective quality improvement of 0.26 on the 5-point Mean Opinion Score (MOS) for a wireless bit error probability of 10^{-3}. Our phase timing measurements indicate that ROHC does not significantly deteriorate the delay jitter in the voice signal. Overall, we note that even if the voice quality improvements with ROHC are moderate and barely perceivable in many practical settings (with ambient noise), the compression gain of ROHC promises remarkable benefit for wireless service providers. The number of 3rd generation mobile cell phone users could nearly be doubled by employing ROHC without allocating more link bandwidth.

For the wireless video transmission with ROHC we found that the RTP/UDP/IPv4 packet header size is also reduced by approximately 85%. Importantly, we found that the packet compression gain (total bandwidth reduction) for the video transmission depends on the video content and the quantization scale of the video encoding. We found that for typical scenarios the bandwidth reduction for video encoded with a fixed quantization scale of 30 varies between 5–15% depending on the video content.

There are a number of interesting and important avenues for future work. One direction is to thoroughly examine the impact of ROHC in wireless video transmission on the visual quality of the received video. We expect that such an investigation will reveal that ROHC has a relatively small impact on the quality of the received video. We expect this result due to the generally larger sizes and smaller compression gains of the video packets compared to the significantly more compressed and smaller voice packets. Another direction is to jointly examine the received voice and video quality to comprehensively evaluate the perceived multimedia quality. It is also important to examine future, yet to be developed robust header compression schemes for wireless networks in terms of the media quality.

Acknowledgment

We are grateful for interactions with Professor Thomas Sikora of the Technical University Berlin, Germany, throughout this work. We gratefully acknowledge the use of the ROHC source code from acticom GmbH, Berlin, Germany. This work was supported in part by the National Science Foundation under Grant No. Career ANI-0133252 and Grant No. ANI-0136774 and the state of Arizona through the IT301 initiative.

References

1. C. Bormann, C. Burmeister, M. Degermark, H. Fukushima, H. Hannu, L.-E. Jonsson, R. Hakenberg, T. Koren, K. Le, Z. Liu, A. Martensson, A. Miyazaki, K. Svanbro, T. Wiebke, T. Yoshimura, and H. Zheng, "RObust Header Compression (ROHC): Framework and Four Profiles: RTP, UDP, ESP, and uncompressed", July 2001.
2. F.H.P. Fitzek, P. Seeling, and M. Reisslein, "Header Compression Schemes for Wireless Internet Access, Chapter 10", in A. Salkintzis, and A. Poularikas (eds.), *Wireless Internet: Technologies and Applications*, CRC Press, 2004, pp. 10.1–10.24.

3. V. Jacobson, "Compressing TCP/IP Headers for Low-Speed Serial Links, Request for Comments 1144", February 1990.
4. A. Calveras, M. Arnau, and J. Paradells, "A Controlled Overhead for TCP/IP Header Compression Algorithm over Wireless Links", in *Proceedings of The 11th International Conference on Wireless Communications (Wireless'99)*, Calgary, Canada, 1999.
5. A. Calveras, M. Arnau, and J. Paradells, "An Improvement of TCP/IP Header Compression Algorithm for Wireless Links", in *Proceedings of Third World Multiconference on Systemics, Cybernetics and Informatics (SCI'99) and the Fifth International Conference on Information Systems Analysis and Synthesis (ISAS'99)*, Vol. 4, Orlando, FL, July/August 1999, pp. 39–46.
6. S.J. Perkins and M.W. Mutka, "Dependency Removal for Transport Protocol Header Compression over Noisy Channels", in *Proceedings of IEEE International Conference on Communications (ICC)*, vol. 2, Montreal, Canada, June 1997, pp. 1025–1029.
7. M. Rossi, A. Giovanardi, M. Zorzi, and G. Mazzini, "Improved Header Compression for TCP/IP Over Wireless Links", *Electronics Letters*, Vol. 36, No. 23, pp. 1958–1960, November 2000.
8. M. Rossi, A. Giovanardi, M. Zorzi, and G. Mazzini, "TCP/IP Header Compression: Proposal and Performance Investigation on a WCDMA Air Interface", in *Proceedings of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 1, Sept. 2001, pp. A-78–A-82.
9. M. Degermark, B. Nordgren, and S. Pink, "IP Header Compression, Request for Comments 2507", February 1999.
10. M. Degermak, M. Engan, B. Nordgren, and S. Pink, "Low–Loss TCP/IP Header Compression for Wireless Networks", in *Proceedings of ACM MobiCom'96*, vol. 3, New York, New York, Oct. 1997, pp. 375–387.
11. C. Perkins and J. Crowcroft, "Effects of Interleaving on RTP Header Compression", in *Proceedings of IEEE Infocom 2000*, Tel Aviv, Israel, 2000, pp. 111–117.
12. J. Lim and H. Stern, "IPv6 Header Compression Algorithm Supporting Mobility in Wireless Networks", in *Proceedings of the Southeastcon 2000*, 2000, pp. 535–540.
13. S. Casner and V. Jacobson, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links, Request for Comments 2508", Feb. 1999.
14. M. Degermark, H. Hannu, L. Jonsson, and K. Svanbro, "Evaluation of CRTP Performance Over Cellular Radio Links", *IEEE Personal Communications*, Vol. 7, No. 4, pp. 20–25, 2000.
15. K. Svanbro, H. Hannu, L.-E. Jonsson, and M. Degermark, "Wireless Real–time IP Services Enabled by Header Compression", in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Vol. 2, Tokyo, Japan, 2000, pp. 1150–1154.
16. A. Cellatoglu, S. Fabri, S. Worral, A. Sadka, and A. Kondoz, "Robust Header Compression for Real–Time Services in Cellular Networks", in *Proceedings of the IEE 3G 2001*, London, GB, Mar. 2001, pp. 124–128.
17. W.-T. Chen, D.-W. Chuang, and H.-C.Hsiao, "Enhancing CRTP by Retransmission for Wireless Networks ", in *Proceedings of the Tenth International Conference on Computer Communications and Networks*, 2001, pp. 426–431.
18. L. Khiem, C. Clanton, L. Zhigang, and Z. Haihong, "Efficient and Robust Header Compression for Real-Time Services", in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Vol. 2, Chicago, IL, 2000, pp. 924–928.
19. G. Boggia, P. Camarda, and V.G. Squeo, "ROHC+: A New Header Compression Scheme for TCP Streams in 3G Wireless Systems", in *Proceedings of the IEEE International Conference on Communications (ICC)*, Vol. 5, pp. 3271–3278, 2002.
20. R. Cuny and A. Lakaniemi, "VoIP in 3G networks: An End-to-End Quality of Service Analysis", in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, Jeju, Korea, Apr. 2003, pp. 930–934.
21. A. Minaburo, L. Nuaymi, K. D. Singh, and L. Toutain, "Configuration and Analysis of Robust Header Compression in UMTS", in *Proceedings of IEEE Personal, Indoor and Mobile Radio Communications*, Sept. 2003, pp. 2402–2406.
22. Z. Yin and V. C. M. Leung, "A Proxy Architecture to Enhance the Performance of WAP 2.0 by Data Compression", in *Proceedings of IEEE Wireless Communications and Networking Conference*, Mar. 2003, pp. 1322–1327.
23. B. Wang, H. Schwefel, K. Chua, R. Kutka, and C. Schmidt, "On Implementation and Improvement of Robust Header Compression in UMTS", in *Proceedings of the 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, 2002, pp. 1151–1155.

24. H. Wang and K.G. Seah, "An Analytical Model for the ROHC RTP Profile", in *Proceedings of IEEE Wireless Communications and Networking Conference*, Mar. 2004, pp. 126–131.
25. S. Rein, F.H.P. Fitzek, and M. Reisslein, "Voice Quality Evaluation in Wireless Packet Communication Systems: A Tutorial and Performance Results for ROHC", *IEEE Wireless Communications, to appear*, 2004, preprint available at http://www.fulton.asu.edu/∼mre/voice_qu_method.pdf.
26. S. Rein, F.H.P. Fitzek, and M. Reisslein, , "Voice Quality Evaluation for Wireless Transmission with ROHC", in *Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Applications*, Honolulu, HI, Aug. 2003, pp. 461–466.
27. F.H.P. Fitzek, S. Hendrata, P. Seeling, and M. Reisslein, "Evaluation of Wireless Video Transmission with Robust Header Compression", in *Proceedings of the International Conference on Information, Communications & Signal Processing and IEEE Pacific-Rim Conference On Multimedia (ICICS-PCM 03),*, Singapore, Dec. 2003, pp. 1346–1350.
28. G. Waters, "Sound quality Assessment Material – Recordings for Subjective Tests: User's handbook for the ebu–sqam compact disk", European Broadcasting Union (EBU), Technical Report, 1988, available at http://www.ebu.ch/tech_32/tech_t3253.pdf.
29. M. Rossi, A. Philippini, and M. Zorzi, "Link Error Characteristics of Dedicated (DCH) and Common (CCH) UMTS Channels", Universita di Ferrara, FUTURE Group, Italy, Tech. Rep., July 2003.
30. "YUV 4:2:0 video sequences", 2002, available at http://trace.eas.asu.edu.
31. T. Wiegand, "H.26L Test Model Long–Term Number 9 (TML-9) draft0", ITU-T Study Group 16, Dec. 2001.
32. JVT, "JM/TML Software Encoder/decoder", available at http://bs.hhi.de/∼suehring/tml/, 2002, h.26L Software Coordination by Carsten Suehring.
33. F. Fitzek and M. Reisslein, "MPEG-4 and H.263 Video Traces for Network Performance Evaluation", *IEEE Network*, Vol. 15, No. 6, pp. 40–54, November/December 2001.

Frank H.P. Fitzek is an Associate Professor in the Department of Communication Technology, University of Aalborg, Denmark heading the Future Vision group. He received his diploma (Dipl.-Ing.) degree in electrical engineering from the University of Technology – Rheinisch-Westflische Technische Hochschule (RWTH) – Aachen, Germany, in 1997 and his Ph.D. (Dr.-Ing.) in Electrical Engineering from the Technical University Berlin, Germany in 2002. As a visiting student at the Arizona State University he conducted research in the field of video services over wireless networks. He co-founded the start-up company acticom GmbH in Berlin in 1999. In 2002 he was Adjunct Professor at the University of Ferrara, Italy giving lectures on wireless communications and conducting research on multi-hop networks. His current research interests are in the areas of 4G wireless communication, QoS support for multimedia services, access techniques, security for wireless communication, and the integration of multi hop networks in cellular systems. Dr. Fitzek serves on the Editorial Board of the IEEE Communications Surveys and Tutorials. He is the program chair for the International Conference on Advances in Computer Entertainment Technology (ACE2004) and serves in the program committee for VTC2003, VTC2004, ACE2004, and IEEE MWN2004.

Stephan Rein studied Electrical Engineering at the Technical University of Aachen, Germany, and the Technical University of Berlin (TUB), Germany. He received the Dipl.-Ing. degree in electrical engineering from the TUB in 2003. From March 2003 to October 2003 he visited the multimedia networking group in the Department of Electrical Engineering at Arizona State University, Tempe. He is currently pursuing the Ph.D. degree at the Institute for Energy and Automation Technology, Technical University of Berlin. His current research interests include data compression and digital signal processing with emphasis on wavelet theory.

Patrick Seeling received the Dipl.-Ing. degree in industrial engineering and management (specializing in electrical engineering) from the Technical University of Berlin (TUB), Germany, in 2002. Since 2003 he has been a Ph.D. student in the Department of Electrical Engineering at Arizona State University. His research interests are in the area of video communications in wired and wireless networks. He is a student member of the IEEE and the ACM.

Martin Reisslein is an Assistant Professor in the Department of Electrical Engineering at Arizona State University, Tempe. He received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Germany, in 1994, and the M.S.E. degree from the University of

Pennsylvania, Philadelphia, in 1996. Both in electrical engineering. He received his Ph.D. in systems engineering from the University of Pennsylvania in 1998. During the academic year 1994–1995 he visited the University of Pennsylvania as a Fulbright scholar. From July 1998 through October 2000 he was a scientist with the German National Research Center for Information Technology (GMD FOKUS), Berlin. While in Berlin he was teaching courses on performance evaluation and computer networking at the Technical University Berlin. He is editor-in-chief of the *IEEE Communications Surveys and Tutorials* and has served on the Technical Program Committees of *IEEE Infocom*, *IEEE Globecom*, and the *IEEE International Symposium on Computer and Communications*. He has organized sessions at the *IEEE Computer Communications Workshop (CCW)*. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG-4 and H.263 encoded video, at http://trace.eas.asu.edu. He is co-recipient of the Best Paper Award of the *SPIE Photonics East 2000 – Terabit Optical Networking* conference. His research interests are in the areas of Internet Quality of Service, video traffic characterization, wireless networking, and optical networking.