Network Working Group
INTERNET-DRAFT
Expires: April 21, 2000

Lars Westberg, Ericsson
Morgan Lindqvist, Ericsson
Sweden

October 21, 1999

Realtime Traffic over Cellular Access Networks
<draft-westberg-realtime-cellular-01.ps>

Status of this memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or cite them other than as _work in progress_.

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/lid-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This document is an individual submission to the IETF. Comments should be directed to the authors.

Abstract

The draft discusses problems with transport of realtime traffic over cellular access channels and their implications for protocol enhancements.

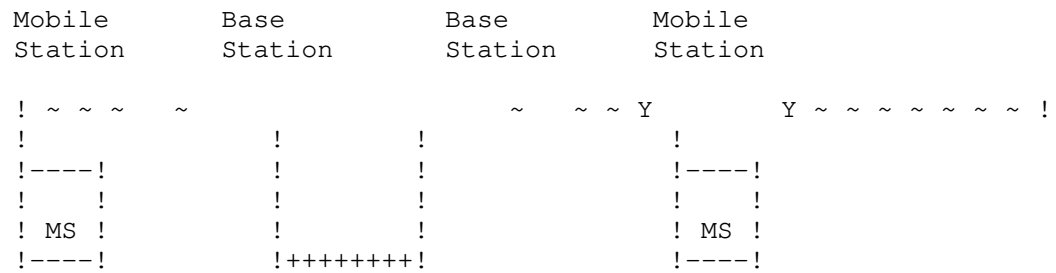
1. Realtime services over cellular access channels - background and motivation

Emerging realtime services in the Internet, such as VoIP (Voice over IP), impose new requirements on cellular access networks. Support for these new services in cellular access networks may be provided in a number of ways, ranging from interworking (e.g., terminating the IP protocols in the fixed network and using other optimized protocols over the cellular link) to transferring the IP packets end-to-end over the cellular links. Transferring the IP packets end-to-end allows the use of standard applications in the cellular terminal and is therefore an important alternative.

Most of the work so far has been focused on transmission of best effort traffic over wireless and not on the time critical applications. End-to-end VoIP applications are possible to use in the new generation of cellular networks, but the efficiency of radio spectrum usage must be improved for such applications. Combining spectrum efficiency, high quality speech and short delay calls for new solutions.

The usual way to transport the IP packets in a radio network is to use retransmissions over the radio link in order to obtain similar characteristics as in the fixed network. This, however, will cause long delays for speech, which in turn entails poor conversational quality. Instead we need to solve the problems arising in the radio network by enhance some parts of the protocol suite.

The scenario we are considering is one where two mobile stations (MSs) are connected to a common fixed network through cellular links.



The mobile stations contain a Voice-Over-IP application and a full IP stack. The application generates audio, video and application-specific session signaling, e.g., SIP/H.323. The audio/video is transported over RTP/UDP/IP, while the application-specific signaling

uses TCP and/or UDP. The cellular access is treated as a layer 2 (L2) network with functionality for optimizing the performance on the cellular link.

In this document, we summarize some of the requirements on the layers that must be met in order to achieve good speech quality and spectrum efficiency when transferring IP packets transparently over the cellular access. It should be noted that due to the spectrum cost of the transparent solution, alternative solutions such as interworking also deserve to be considered. Such solutions, however, are not discussed in this draft.

It should also be pointed out that some of the problems with realtime packets over cellular access might only be solvable with "wireless aware" terminals, meaning that not only the link layers, but also the IP stack must be "wireless aware". However, all terminals and applications will not be "wireless aware". Interworking between the two classes of terminals/applications can be solved by gateways in the fixed network.

2. Cellular access performance and system cost

The cellular radio access puts tough requirements on end-to-end packet transmission. Packet transmission over the cellular access is typically constrained by two factors:

- The high cost of cellular access links. Cellular bandwidth with high quality imposes high system cost.

- The lossy link behavior. The radio network generates a high BER. If retransmission over the radio link is not used, the BER may be in the order of 10e-3.

2.1. System Cost - Selection of BER for the radio link

In wireless systems there is a close relationship between the BER and the SNR (signal-to-noise ratio) of the radio channel. Furthermore, the required SNR (corresponding to a selected BER requirement) can be directly related to the system capacity, i.e. the number of users per cell. Fewer users result in lower income, which in the end result in higher system cost.

The cellular access network has wide area coverage and if the service requires more spectrum, the whole network capacity will be affected due to the mobility of the users. A poor spectrum usage for Voice over IP will cause much more cost for hardware in base stations than other more optimized services e.g. circuit switched voice.

Another important factor is the cost of the spectrum itself. The spectrum is a limited resource and is licensed by the cellular operator. The cost for licensed RF-spectrum is estimated to be 30% [18] of the total operating cost and therefore represents a significant cost for the cellular operator.

A typical circuit switched voice service requires a target BER of 10e-3. If the BER requirement is changed from 10e-3 to 10e-6, this might result in a decreased spectrum efficiency of 25-50%, depending on the type of cellular system and the underlying radio conditions. This will increase system costs significantly. An efficient use of the spectrum resource is crucial.

2.2. Lossy links - Design of link layer protocol based on radio requirements

If the radio link characteristics are not considered in the link layer design, the services will be more costly, or the performance in terms of speech quality will be poor. The design of current protocols is based on the transmission characteristics of fixed networks, so these are not well suited to the radio requirements. A good trade-off between the requirements of transmission (BER and packet loss) and the design of the protocols is crucial.

The quality in a radio system (expressed for instance as the BER) typically changes from one 20 ms radio frame to another due to fading in the radio channel. For example, in a cellular system where the target BER has been set to 10e-3, the BER might vary, roughly speaking, from 10e-4 to 10e-2. However, the average BER will equal the target BER. Thus, for the system capacity to remain unaffected, the link layer protocols and speech coders must tolerate that the BER exceeds the BER target for limited periods of time.

Another important aspect is the long round-trip time (RTT). Even if we use channels without retransmission over the radio link, the unidirectional delay might be important to consider for some link layer functions, such as header compression.

Delay in cellular systems is caused by several reasons. Forward Error Correction (FEC) and interleaving are used to increase the radio channel performance. A typical cellular radio channel gives, without any improvements, approximately 10% BER. With interleaving and FEC, the radio channel can provide the proper channel performance (as discussed above) for the service. But this introduces a fixed delay for the channel. In some cellular systems, the algorithmic channel RTT-delay is in the order of 80ms.

Other sources of delay are DSP-processing, node internal delay and transmission. In all cellular system, the overall delay budget is

optimized to achieve optimal performance for service quality and spectrum efficiency.

It is difficult to state a generally valid value for the RTT. Some RTT figures (e.g. 200 ms) are mentioned in [16], but the delay might be shorter for the case of radio channels optimized for real-time, 100-200 ms if the speech coding is excluded. One may compare the RTT (Mobile->PSTN-GW->Mobile) for circuit switched speech, with the "long-term objective value" which is stated to be less than 180 ms for GSM-FR (GSM full-rate speech codec) in [17].

3. Transport of realtime IP flows over cellular

We summarize the problems of transporting realtime packets over cellular links and the implications of these problems for protocol enhancements in wireless transmission.

3.1. Layer 2 enhancements for realtime traffic

To efficiently transfer audio and video streams over the radio channels, these flows should be identified and de-multiplexed. Identification of realtime flows could be carried out by heuristic rules, as proposed in [12]. One of the problems is that the radio channels still need to be adapted to the characteristics of the compressed information. The BER assignment might be different for audio and video. We might also differentiate the redundancy coding of the compressed payload, something that requires a detailed knowledge of the payload.

Therefore, for RTP flows that have dynamically assigned payload type indicator (PTI) values [13], the identification of codec type is important in order to allow simple layer 2 identification of the compressed payload type.

Apart from the problems of transporting the payload, we also have to perform optimization of the protocol headers [14] for real-time traffic. One of the major problems here is that the compression algorithm must work well in the radio environment with its link delay, and also be resistant to bit errors. The current header compression scheme [14] is sensitive to bit errors, as is shown in [15].

3.2. Transport of audio over cellular

3.2.1. Properties of speech codecs designed for cellular networks

A cellular access link is very lossy and expensive compared to fixed lines. Furthermore telephony is a real-time service where retransmissions should be avoided. Thus the speech decoder should receive not only error-free speech frames but also frames with errors [1]. If all erroneous speech frames are dropped, the frame error rate (FER) will be high. With a high FER, it is not possible to produce speech with an acceptable quality. In cellular telephony systems, this problem is overcome by delivering all frames to the speech decoder, regardless of bit errors.

The sensitivity to errors varies widely between different bits in a frame of encoded speech. High error sensitivity means that an error in that bit results in a severe degradation in speech quality. In most cellular speech codecs for 2nd generation mobile systems (GSM, TDMA or PDC), the bits are divided into three classes: 1a, 1b and 2. Class 1a (the most sensitive bits) and 1b (medium sensitive bits) are protected by convolutional coding. Class 1a bits are in addition covered by a CRC. Class 2 bits (the least sensitive bits) are not protected at all. This scheme results in a reduced FER, since a frame is considered erroneous only if there are errors in the class 1a bits (which on average amount to one third of the total number of bits). On the other hand, the scheme also leaves undetected residual errors in class 1b and class 2. However, it is better, from a speech quality point of view, to allow some errors in these bits than to discard the whole frame as soon as bit errors occur, and let the ECU (see 3.2.2 below) reconstruct the frame [2][3][4].

3.2.2. Error Concealment Unit (ECU)

If the CRC for the class 1a bits is corrupt, there are severe errors in the speech frame, which probably would give rise to annoying distortions. The frame is therefore discarded, and instead the speech decoder generates artificial speech that closely resembles that of the previous frames. In this way the decoder attempts to mask the distortion. The component carrying out this task is called the error concealment unit (ECU). The ECU reconstructs the frame based on the corrupt version of it that was received as well as the last good frame. Using the bad frame itself increases the speech quality if the number of damaged bits in the frame is not too high.

In most speech coders with a 20 ms frame size, the ECU is a state machine with 6 states. When several consecutive lost/bad frames are encountered, the ECU proceeds to the next state for each new frame (until it reaches state 6). The amplitude of the generated speech is

gradually reduced in the states, and in state 6, the speech is completely muted.

If the decoder receives a good frame when the ECU is in state 1-5, it immediately switches back to normal decoding mode. If the ECU is in state 6 when a good frame is received, the decoder awaits the next frame. If this frame is also good, the decoder switches back to normal decoding mode, otherwise it remains in state 6. This feature is added to eliminate sound spikes when recovering from a long sequence of lost/bad frames. It also mitigates the effects of residual bit errors in the class 1a bits.

3.2.3. Adaptive buffer management

In order to minimize the end-to-end delay, an adaptive buffer manager (ABM) is useful, another term is adaptive play-out buffer. The function of the adaptive buffer manager is to change the buffer size in order to allow as many packets as possible to reach the speech decoder in time, while keeping the buffering delay to a minimum.

To achieve good performance, the ABM should treat the packets with bit errors in the payload as normal packets, not as late packets. Otherwise, the buffer size might be larger than necessary.

3.3. Transport of Video

The transport of compressed video, intended for conversational services (i.e., videophone) over cellular links, entails some unique requirements. One is that, delay must be kept under strict control. Cellular links also have other error characteristics than fixed networks, something that may cause problems.

One way to fulfil the requirements (real-time and limited delay) is to allow bit errors in the payload in the same way as for speech. Errors in the payload are of course not permissible for all type of packets, and not even for all video streams. However, a number of existing video compression standards (e.g. H.263 and MPEG4) have error-resilient modes that facilitate resynchronization of the decoder and improves performance on parts with errors in them.

3.3.1. Conversational video in a wireless environment

Wireless channels have high bit error rates. These high bit error rates will result in requirements on retransmission, if all packets delivered to the application must be error-free. This is not a good solution for conversational applications, as was explained above.

There are, at least, two possible ways to implement a conversational service over channels with high bit error rate. One is to use strong forward error correction; another is to have an error-robust method of video compression (usually called error-resilient source coding). Many experiments (ITU, MPEG, ARIB, 3GPP) have shown that for a wireless channel, error-resilient source coding [5] outperforms methods using forward error correction codes.

4. Conclusions

Spectrum efficient transmission of audio and video packets is extremely important in cellular access. Spectrum constitutes a significant cost for the operator and must be considered in the development of end-user services. To facilitate effective transport of voice and video in a cellular system, some improvements of the IP protocol suite are needed. Some of the changes are related to the link layer and some to the behavior of RTP (real-time protocol).

The following improvements are identified:

- Simple identification of codec type in the link layer. The knowledge of codec type enables enhancement of the performance over the wireless link.

- BER-resistant header compression algorithm for RTP/UDP/IP. The header compression algorithm also has to work well in an environment with long round-trip delays.

- No dropped packets due to bit errors in the payload. The speech decoder and the buffer manager perform better if they can access all packets, also those that contain bit errors.

- Use of CRC for the most sensitive bits in the payload in order to detect bit errors. This improves the performance of the speech decoder.

5. References

[1] European digital cellular telecommunication system (Phase 2): Radio transmission and reception (GSM 05.05), European Telecommunications Standards Institute, October 1993.

[2] European digital cellular telecommunication system (Phase 2): Channel coding (GSM 05.03), European Telecommunications Standards Institute, October 1993.

[3] TIA/EIA/IS-641, Interim Standard, TDMA Cellular/PCS radio interface - Enhanced Full-Rate Speech Codec, May 1996.

[4] Association of Radio Industry and Business, RCR STD-27F, 1997.

[5] Signal Processing, Image Communication, Special Issue on Error Resilience, Volume 14, Nos. 6-8, May 1999, page 443-676, ISSN 0923-5965, Guest Editors: J.C Brailean, T. Sikora and T. Miki.

[6] Video coding for low bit rate communication, Recommendation H.263 (02/98), International Telecommunication Union.

[7] Multiplexing protocol for low bit rate multimedia communication, Recommendation H.223 (03/96) with later annexes (A,B,C 02/98), International Telecommunication Union.

[8] Information Technology -- Very low bitrate audio-visual coding - Part 2: Visual, ISO/IEC 14496-2 ("MPEG4").

[9] Association of Radio Industry and Business, Test results of video multimedia codec simulation, ICWG 16-4, July 17, 1998.

[10] Association of Radio Industry and Business, Report of ARIB IMT-2000 Video Multimedia Codec Simulation Test, ICW-VMG35-, March 18, 1999.

[11] 3rd Generation Partnership Project (3GPP), TSG-SA Coding Working Group, "QoS for Speech and Multimedia Codec Quantitative performance evaluation of H.324 Annex C over 3G", TR 26.116 (working document).

[12] Heuristics for utilizing ISSL Mechanisms for A/V Streams over Low Bandwidth Links in the absence of Announcement Protocols, IETF, draft-putzolu-heuristic-00.txt (work in progress).

[13] RTP Profile for Audio and Video Conferences with Minimal Control, IETF, ietf-avt-profile-new-05.txt (work in progress).

[14] Compressing IP/UDP/RTP Headers for Low-Speed Serial Links, IETF, RFC 2508.

[15] CRTP over cellular radio links, IETF, draft-degermark-crtp-cellular-00.txt (work in progress).

[16] Long Thin Networks, IETF, draft-montenegro-pilc-ltn-02.txt (work in progress).

[17] European digital cellular telecommunication system (phase 1): Technical Performance Objectives, GSM 03.05, version 3.2.0.

[18] Wireless Voice-over-IP and Implications for Third-Generation Network Design, Bell Labs Technical Journal, September 1998

6. Authors' addresses

Lars Westberg
Ericsson Research
E-mail: rtiow@era-t.ericsson.se

Morgan Lindquist
Ericsson Research
E-mail: morgan.lindqvist@era.ericsson.se

This Internet-Draft expires April 21, 2000.