一种在蜂窝链路上高效传输实时 IP 业务的机制*

吴亦川+, 黄 奎, 孙利民, 吴志美

(中国科学院 软件研究所,北京 100080)

A Mechanism to Efficiently Transport Wireless Real-Time IP Services over Cellular Links

WU Yi-Chuan+, HUANG Kui, SUN Li-Min, WU Zhi-Mei

(Institute of Software, The Chinese Academy of Sciences, Beijing 100080, China)

+ Corresponding author: Phn: +86-10-62645407 ext 801, E-mail: wyc@iscas.ac.cn, http://www.iscas.ac.cn

Received 2003-04-17; Accepted 2003-05-27

Wu YC, Huang K, Sun LM, Wu ZM. A mechanism to efficiently transport wireless real-time IP services over cellular links. ***Journal of Software*****, 2004,15(1):112~119.**

http://www.jos.org.cn/1000-9825/15/112.hml

Abstract: When transporting real-time IP services such as VoIP in a cellular communication system with limited channel bandwidth and high bit error rate, there exist some problems such as the relatively large overhead imposed by protocol headers and poor data discarding policies. In this paper, a new scheme, ROHC/UDP Lite, is proposed to solve these problems, which combines ROHC with UDP Lite protocol effectively. With the ROHC/UDP Lite, it is possible to efficiently transport IP-based packet-switched voice services in the future IP cellular systems.

Key words: cellular link; real-time IP service; ROHC; UDP Lite

摘 要: 针对在传输信道带宽有限和传输误码率高的蜂窝移动通信系统中,传输基于 IP 包交换的话音等实时业务时所面临的协议报头开销大、有用包被丢弃等问题,提出一种 ROHC/UDP Lite 机制,将 ROHC 报头压缩算法和 UDP Lite 传输协议有机结合,以解决上述问题.给出了实验模型,分析了实验结果,表明未来的全 IP 蜂窝移动通信系统采用此机制,能够高效传输基于 IP 包交换的话音业务.

关键词: 蜂窝链路;实时 IP 业务;ROHC;UDP Lite

中图法分类号: TP393 **文献标识码**: A

Cellular system is currently starting to embrace the IP (Internet protocol) as the choice of a communication

---

* Supported by the National Natural Science Foundation of China under Grant No.60272098 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2001AA112051 (国家高技术研究发展计划(863))

WU Yi-Chuan was born in 1977. He is a Ph.D. candidate at the Institute of Software, the Chinese Academy of Sciences. His current research interests include wireless IP and multimedia communication. **HUANG Kui** was born in 1977. He is a Ph.D. candidate at the Institute of Software, the Chinese Academy of Sciences. His current research interests include multicast and wireless network. **SUN Li-Min** was born in 1966. He is a professor at the Institute of Software, the Chinese Academy of Sciences. His current research interests include wireless network and broad band access network. **WU Zhi-Mei** was born in 1942. He is a professor at the Institute of Software, the Chinese Academy of Sciences. His current research interests include multimedia communication.

framework for the future. To ensure its application independence and decrease the transporting and switching costs, it is attractive to go IP all the way over the air interface to the end user equipment, i.e., not to terminate Internet protocols before the air interface. IP all the way enables service flexibility, i.e., there are no dependencies between applications and the radio access network[1]. A base is created where many players can participate and develop new applications.

The Internet Protocol offers an unreliable connectionless network-layer service. When used in cellular system to transport real-time services, IP can only offer best effort service[2]. To provide a guaranteed quality of service for real-time audio and video using public net-works, UDP/RTP[3,4] is typically deployed over IP in the IP-based protocol stack. However, the main drawback of using IP/UDP/RTP is the relatively large overhead imposed by these headers, which leads to a significant waste of wireless bandwidth and spectrum inefficiency, particularly for low-bitrate voice over IP (VoIP) services. Generally, an IPv4 packet with speech data will have an IP header, a UDP header, and an RTP header, making a total of 20+8+12=40 octets. With IPv6[5], the IP header is 40 octets for a total of 60 octets. The size of the speech data depends on the codec, but can be 15~30 octets. These figures present a major reason for not using IP over the air interface.

Header compression can significantly reduce the size of protocol headers so as to reducing the bandwidth required for headers. Since header compression allows smaller packet sizes, the end-to-end delay can be significantly lower and the QoS (quality of service) of services can be improved. In addition, as the overall spectrum efficiency of the cellular link is increased, other applications will also benefit from header compression. CRTP[6] (compressed RTP) can compress IP/UDP/RTP headers from 40 or 60 octets down to 2~4 octets. However, cellular system cannot provide enough QoS for real-time IP services using CRTP due to the error-proneness and long round trip time of cellular links[7]. ROHC[8] (RObust Header Compression) solves this problem by making the IP-/UDP/RTP header compression more tolerant of loss of compressed packets. Headers can be compressed down to 1~3 octets. Comparing ROHC over CRTP over cellular links shows that CRTP losses 8 times more packets than ROHC, resulting in decreased quality of services.

On the other hand, in order to get a better QoS in cellular system, the applications of real-time IP services are often designed to be able to handle damaged data to some extent. But UDP protocol provides a service that can throw away the whole packets when any bit error is detected by its checksum which covers the whole datagram. Thus, over cellular links of high bit error rate, the checksum policy of UDP may cause discarding a lot of packets useful to the applications whose payloads include some errors, resulting in decreasing the throughput of channel and QoS of services. With this mind, the UDP Lite[9] protocol is designed. It provides the option to define a part of each datagram as insensitive to errors. Errors in the insensitive part will be ignored by the checksum mechanisms, and the decision whether a damaged datagram is usable or not is moved from the transport protocol to the application itself. Even with this extra flexibility, the UDP Lite protocol is still compatible with the UDP protocol.

Based on above, by combining ROHC with UDP Lite effectively, it becomes possible to efficiently transport real-time IP services over cellular links. In this paper, ROHC/UDP Lite scheme is proposed and its performance is detailedly evaluated by simulation. Section 1 describes this scheme in detail. Simulation environment is presented in Section 2, together with discussions about the results. Finally, the conclusions are presented in Section 3.

1 ROHC/UDP Lite Scheme

1.1 Header compression

1.1.1 The fundamental of header compression

Header compression is possible thanks to a high degree of redundancy between header fields within each packet and especially between consecutive packets. Header compression is applied on a per-hop basis. The process of header compression is shown in Fig.1.

Header compression algorithms maintain, at both sides of a link, a shared state called context including essentially the uncompressed version of the last header sent over the link. Compression and decompression are done relative to the context, and only unpredictable changes are transferred. While static header information needs to be transmitted only once, dynamic information can be transmitted as changes from the previous packet using only a few bits. When packets are lost over the link, the decompressor context will be brought out of sync with compressor context, and decompressing of subsequent headers will fail. Thus, in addition to methods for initializing and updating the context, header compression algorithms must also have mechanisms for the detection of incorrect decompression and ways to repair an invalidated decompressor context.

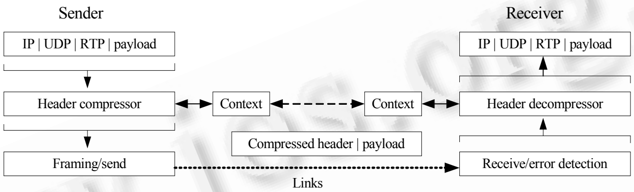


Fig.1 IP/UDP/RTP header compression

1.1.2 CRTP algorithm

CRTP is suitable for compressing IP/UDP/RTP headers over low speed serial links which can compress the 40 or 60 octets IP/UDP/RTP headers to 2~4 octets. To repair an invalidated context, CRTP makes use of explicit signaling messages called context_state by which the decompressor requests a context update from the compressor. The link round trip time will thus affect the time it takes to update the decompressor context. While the context is invalid, all packets received by the decompressor will be discarded, meaning that a long link round trip time will generate a large number of packets lost.

On lossy links with long round trip times, such as most cellular links, CRTP does not perform well. This has been studied in Ref.[7], where it was found that the compression ratio of CRTP is satisfactory but the packer loss rate after decompression at the receiver is too high over such links.

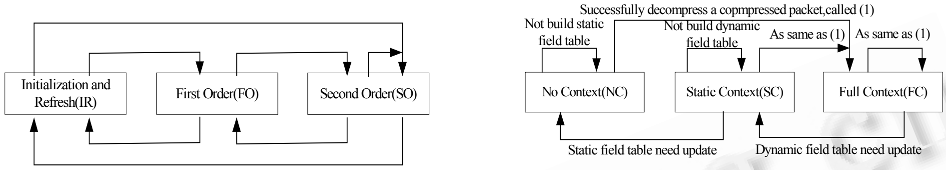
1.1.3 ROHC algorithm

ROHC is developed to be adaptable to the characteristics of wireless links which can compress the 40 or 60 octets IP/UDP/RTP headers to 1~3 octets. Thanks to its adaptability, ROHC exhibits an excellent performance both in term of compression efficiency and robustness against packet loss on cellular links.

ROHC is heavily geared towards a local decompressor repair of the context. A CRC covering the original uncompressed header is included in the compressed header, providing a reliable way to verify the correctness of the reconstructed headers. ROHC defines three types of packet field according to their change patterns during the lifetime of packet streams: STATIC, DYNAMIC and INFERRABLE fields. STATIC fields do not change; DYNAMIC fields may change by variable quantities; INFERRABLE fields can be determined through the knowledge of STATIC and DYNAMIC fields.

Header compression with ROHC can be characterized as an interaction between two state machines, one compressor machine and one decompressor machine, each instantiated once per context. The compressor bases on the state machine in Fig.2(a), with each state representing a different level of compression. Transition conditions between the three states are defined by modes of operation.

The compressor starts in the IR state with the transmission of all static and dynamic header fields, allowing the decompressor to acquire the correct context. In the FO state, the compressor considers only the dynamic fields that have changed respect to the header of the previous packet and transmits only encoded variations of these dynamic fields. In the SO state, where we have the highest level of compression, the compressor enters only when the stream becomes regular, i.e., when dynamic field changes are predictable by the knowledge of a “leader” header field (for example, RTP sequence number for IP/UDP/RTP streams). In the SO state, it is sufficient to send only the encoded “leader” field variations, so that, compressed header size can be 1~2 bytes when UDP checksum is not used.



(a) state transition of compressor (b) state transition of decompressor

Fig.2 State transition of ROHC

The decompressor bases on the state machine in Fig.2(b), with each state representing a different level of decompression according to context state. Initially, the decompressor works in the NC state. Once a packet has been decompressed correctly, the decompressor transmits all the way to the FC state. When the dynamic fields of context needs to update, the decompressor transits back to the SC state. Only when the static fields of context needs to update will the decompressor go all the way back to the NC state. By updating context step-by-step, ROHC can further reduce the size of compressed headers and the possibility of context out of sync between the compressor and decompressor.

ROHC defines three modes of operation: Unidirectional mode, Bidirectional Optimistic and Bidirectional Reliable modes. The optimal mode to operate in depends on the characteristics of the links over which ROHC is used, such as feedback abilities, error probabilities and distributions etc. The transition conditions between these three modes are decided by the decompressor. In addition, ROHC can get further compression gain due to the synergy between the sophisticated encoding methods and the wide availability of formats needed for sending compressed header. ROHC elasticity derives from the concept of the compression profiles. Actually, ROHC specifies four compression profiles:

- Profile 1 for IP/UDP/RTP streams;
- Profile 2 for UDP/IP streams;
- Profile 3 for ESP/IP streams;
- Profile 0 for uncompressed streams.

With the above mentioned mechanisms, ROHC is suitable for compressing IP/UDP/RTP headers on wireless links due to the significant reduction of the negative impact imposed by the high BER of a wireless channel and high round trip time that usually characterizes cellular links.

1.2 UDP Lite protocol

In future IP cellular systems, ROHC addresses the problem with an expensive band-width, but IP packets across a cellular link might be discarded by UDP due to a failing checksum. If damaged data packets are usable by the application, it is not desirable that they should be discarded by the transport layer. In IPv4, disabling the UDP checksum can enable delivery of damaged packets to the application, but doing so will remove the original verification for UDP and RTP headers as well. Moreover, the UDP checksum is mandatory and cannot be discarded in IPv6. Therefore, it is usually undesirable to disable the UDP checksum to ensure that usable packets are not discarded by the transport layer.

The UDP Lite protocol can solve this problem. It replaces the redundant Length field in the UDP headers with a coverage field shown in Fig.3. This modification will not disturb intermediate nodes along a path since the length field in the UDP header is used only at the endpoints. Only the first coverage octets of the UDP Lite datagram are verified by the checksum. By setting the coverage value to less than the datagram length, the datagram is divided into one sensitive and one insensitive part. Errors in the sensitive part will cause datagram discarded, while errors in the insensitive part are ignored.

0	15	16 31
Source address		
Destination address		
Zero	Protocol	UDP Length
Source port		Destination port
Length		Checksum

0	15	16 31
Source address		
Destination address		
Zero	Protocol	UDP Length
Source port		Destination port
Coverage		Checksum

Fig.3 UDP and UDP Lite header, respectively

UDP Lite provides the flexibility for cellular systems to move the decision whether a damaged packet is usable from the transport layer to the application itself. By setting the coverage field reasonably, UDP Lite can nearly ensure not discarding usable data to applications so as to improve the throughput of cellular links and the QoS. Moreover, UDP Lite is still compatible with UDP. If the coverage field value is equal to the datagram length, the UDP Lite datagram is identical to a UDP datagram. The simple design of the UDP Lite protocol makes it easy to implement.

1.3 The combination of ROHC and UDP lite

ROHC/UDP Lite scheme is proposed by combining ROHC and UDP Lite effectively in the following way:

• In UDP Lite, a coverage field value is set to the length of datagram headers, which means that checksum covers only headers and not the payload data.

• As to the protocol field in IP header, UDP Lite uses a different identification number from UDP. Hence, a new profile 4 is added to ROHC to compress IP/UDP Lite/RTP headers. Except the compression methods for the coverage and checksum fields in UDP Lite header, profile 4 is fully identical to profile 1.

• Profile 4 compresses the coverage field of UDP Lite header using the method similar to profile 1 for compressing the length field of UDP header.

• Since the CRC of ROHC provides a stronger error protection for the same original headers data than the checksum of UDP Lite, the UDP Lite checksum does not have to be transmitted over the link. At the decompressor, the checksum is reconstructed if the ROHC CRC succeeds. Thus, the average compressed header size is further reduced 2 octets.

With the ROHC/UDP Lite scheme, it becomes possible to efficiently transport real-time IP services over cellular links.

2 Simulation and Analysis

2.1 Simulatin model

In order to more closely evaluate the performance differences among CRTP/UDP, CRTP/UDP Lite, ROHC/UDP and ROHC/UDP Lite, detailed simulations have been made according to the model shown in Fig.4. UDP uses checksum policy. The checksum of UDP Lite covers only headers.

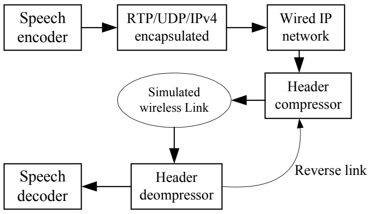


Fig.4 Simulation model

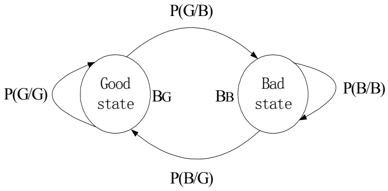


Fig.5 2-State Markov model

2.1.1 Speech encoder

The speech source generates a packet with a fixed size, 264 bits, every 20 ms (13.2kbps), corresponding to the GSM speech codec. The length of the talk spurts and the silence intervals between them are both exponentially distributed with an expected length of 1 second. Silence suppression is used.

2.1.2 Used link layer

CRTP needs to have the packet type identification provided by the link layer, whereas ROHC has the packet type identification integrated. Hence one extra octet of link layer overhead is added for the CRTP case. This octet is not included in the presented result.

A packet is considered as lost if it is not passed up to the speech codec. There are four possible reasons for packet loss in these simulations:

- A bit error has occurred in the compressed header;
- A bit error has occurred in the link layer packet type;
- The header compression algorithm has a faulty context and cannot decompress any received compressed header (context damage). This can happen even if the compressed header is error free.
- When using UDP, a packet is also lost if it has any bit error in the payload. But in the cases with UDP Lite a packet with errors in the payload is not regarded as lost as long as it is considered ‘ok’ by a header compression algorithm.

2.1.3 The simulation of error model of cellular link

The error model of wireless link used in the simulation is modeled as a two-state Markov[10,11] process with two states: good and bad which is illustrated in Fig.5. The good state represents the error free behavior of the link, while the bad state represents a fade. In each state, bit errors are considered to be uniformly distributed, and their mean bit error rates are B_B and B_G respectively. The durations of each state are also uniformly distributed with average durations T_B and T_G. Then the link conditions can be modeled by the following transition probability matrix

$$P=\begin{pmatrix} P(B/B) & P(B/G) \\ P(G/B) & P(G/G) \end{pmatrix} \tag{1}$$

where

$$\begin{aligned} P(B/B)&=1-P(B/G) \\ P(G/G)&=1-P(G/B) \end{aligned} \tag{2}$$

Suppose

$$\begin{aligned} &T_G>1 \\ &T_B>1 \\ &P(B/G)=1/T_B \\ &P(G/B)=1/T_G \end{aligned} \tag{3}$$

The steady-state probability that the link is in the bad state is given by

$$\varepsilon=P(G/B)/(P(B/G)+P(G/B)) \tag{4}$$

The average bit error rate of the link is

$$BER = \varepsilon B_B + (1-\varepsilon) B_G \tag{5}$$

Putting (3) and (4) into (5), we get

$$BER = (B_B T_B + B_G T_G)/(T_B + T_G) \tag{6}$$

The error characteristics of various real cellular links can be simulated by the link average bit error rate given by (6) through adjusting the values of the four parameters B_B, B_G, T_B and T_G. In addition, the reverse link used in simulation is supposed not to damage the FEED-BACK messages. The RTT is set to 120 ms which is similar to the RTT of real cellular links.

2.2 Simulation results and analyses

2.2.1 Packet loss pattern

Figure 6 shows the packet loss patterns for different schemes at BER 10^{-3}. It is evident from this figure that the majority of CRTP losses are such that 9 and 10 consecutive packets are lost. This comes from its context repair mechanism, which is heavily dependent on the round trip time of the link. For ROHC, all loss events are nearly 1 lost packets, which means that it does not suffer from context damage. Figure 6 also shows that the choice of transport protocol does not significantly affect the distribution of consecutive packet losses. This is mainly due to the fact that the consecutive packet loss is mainly caused by header compression.

2.2.2 Packet loss rate

Figure 7 shows the packet loss rate versus channel BER. It shows the enhanced robustness of the ROHC compared to the CRTP. A packet loss between CRTP compressor and decompressor results in a burst of additional packet loss due to CRTP's repair mechanism. At a BER of 10^{-3}, the loss rate for CRTP is approximately 8 times higher than for ROHC. There is also a clear difference in packet loss rate when using UDP and UDP Lite. The packet loss rate of ROHC/UDP is about triple as high with ROHC/UDP Lite. The packet loss rate of CRTP/UDP is about one time higher than for CRTP/UDP Lite.

2.2.3 Average header size

Figure 8 shows the average header size plotted against channel BER. CRTP needs a context update packet (17 octets) for each loss, so the average compressed header size increases as the link BER increases. Since the context is rarely brought out of sync for ROHC, the average header size is slightly larger than the minimum header size. Furthermore, since ROHC/UDP Lite does not need to transmit UDP Lite checksum over the link, its average header size is only bigger than the minimal header size of ROHC (1 octet). While in other three schemes, the checksum must be included in each compressed packet.

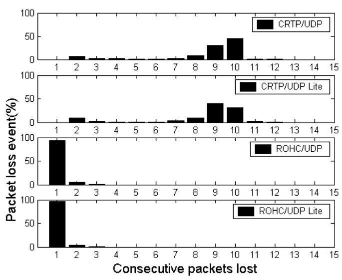


Fig.6 Packet loss pattern, channel BER: 10^{-3}

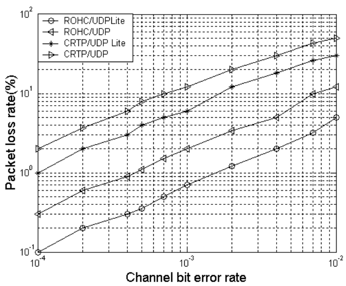


Fig.7 Packet loss rate versus channel BER

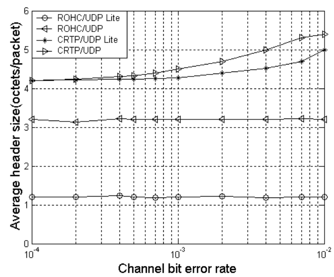


Fig.8 Average header size versus channel BER

3 Conclusions

This paper proposes a new scheme ROHC/UDP Lite which combines ROHC and UDP Lite effectively. With this scheme, it is able to not only significantly reduce the relatively large overhead imposed by protocol headers, but also deliver those packets whose payloads include errors to the up applications so as to efficiently transport real-time IP services over cellular networks.

Simulation results show that at a realistic bit error rate (10^{-3}), the packet loss rate of CRTP/UDP is about 8 times higher than for ROHC/UDP. CRTP/UDP Lite can deliver about twice as many packets to the applications as CRTP/UDP. Compared with CRTP/UDP, ROHC/UDP Lite can deliver at least an order of magnitude more packets to the applications. Simulation results also show that when using ROHC/UDP Lite scheme in wireless VoIP, the packet loss rate is less than 1% at a realistic bit error rate (10^{-3}) and voice quality is sufficiently better. With ROHC/UDP Lite, it is possible to use IP-based packet-switched voice services in future IP cellular systems with a sufficient speech quality and spectrum efficiency.

References:

[1] Svanbro K, Hannu H, Jonsson L.E, Degermark M. Wireless real-time IP services enabled by header compression. In: Proc. of the Vehicular Technology Conf. IEEE Press, 2000. 15~18.

[2] Postel J. Internet protocol. IETF RFC791, 1981.

[3] Postel J. User datagram protocol. IETF RFC768, 1980.

[4] Schulzrinne H, Casner S, Frederick R, Jacobson V. RTP: A transport protocol for real-time applications. IETF RFC1889, 1996.

[5] Deering S, Hinden R. Internet protocol, version 6(IPv6). IETF RFC2460, 1998.

[6] Casner S, Jacobson V. Compressing IP/UDP/RTP headers for low-speed serial links. IETF RFC2508, 1999.

[7] Degermark M, Hannu H, Jonsson L, Svanbro K. Evaluation of CRTP performance over cellular radio links. Personal Communications, 2000,7(4):20~25.

[8] Bormann C, Degermark M, Fukushima H, Hannu H, Jonsson L-E. Hakenberg R, Koren T, Le K, Liu Z, Martensson A, Miyazaki A, Svanbro K, Wiebke T, Yoshimura T, Zheng H. RObust header compression (ROHC): framework and four profiles: RTP, UDP, ESP, and uncompressed. IETF RFC3095, 2001.

[9] Larzon LA, Degermark M, Pink S. The UDP Lite Protocol. IETF Internet Draft, 2002.

[10] Fritchman B.D. A binary channel characterization using partitioned Markov chains. IEEE Trans. on Information Theory, 1967, 13:221~227.

[11] Zorzi M, Rao RR, Milstein LB. On the accuracy of a first order Markov Model for data transmission on fading channels. In: Proc. of the IEEE ICUPC'95. IEEE Press, 1995. 211~215.