논문 00-02-10

The Header Compression Scheme for Real-Time Multimedia Service Data in All IP Network
All IP 네트워크에서 실시간 멀티미디어 서비스 데이터를 위한 헤더 압축 기술

Sang-Ho Choi*, Kwang-Chun Ho**, Yung-Kwon Kim***
(崔 尙 鎬*, 扈 光 春**, 金 暎 權***)

Abstract

This paper remarks IETF based requirements for IP/UDP/RTP header compression issued in 3GPP2 All IP Ad Hoc Meeting and protocol stacks of the next generation mobile station, All IP Network, for real time application such as Voice over IP (VoIP) multimedia services based on 3GPP2 3G cdma2000. Frames for various protocols expected in the All IP network Mobile Station (MS) are explained with several figures including the bit-for-bit notation of header format based on IETF draft of Robust Header Compression Working Group (ROHC). Especially, this paper includes problems of IS-707 Radio Link Protocol (RLP) for header compression which will be expected to modify in All IP network MS's medium access layer to accommodate real time packet data service[1]. And also, since PPP has also many problems in header compression and mobility aspects in MS protocol stacks for 3G cdma2000 packet data network based on Mobile IP (PN-4286)[2], we introduce the problem of solution for header compression of PPP. Finally, we suggest the guidelines for All IP network MS header compression about expected protocol stacks, radio resource efficiency and performance

요 약

본 논문은 3GPP2 All IP Ad Hoc 회의에서 언급된 IP/UDP/RTP 헤더 압축을 위한 IETF의 요구사항에 대해 언급하고, 3GPP2 3G cdma2000에 기반을 두고 있는 VoIP 멀티미디어 서비스와 같은 실시간 응용을 위한 차세대 이동통신 단말기의 프로토콜 스택을 연구하였다. All IP 네트워크 단말기의 다양한 프로토콜에 대한 프레임은 IETF ROHC Working Group의 draft에 기반을 둔 헤더 형태의 비트별 설명을 포함하여 그림으로 설명하였다. 특히, 본 논문은 실시간 패킷 데이터를 수용하기 위해 All IP 단말기의 Medium Access Layer 계층에서 변형될 헤더 압축관련 IS-707 RLP 프로토콜의 문제점을 포함하고 있다. PPP프로토콜은 현재 이슈가 되고 있는 Mobile IP(PN-4286)에 기반을 둔 3세대 cdma2000 패킷 데이터 네트워크을 위한 단말기의 프로토콜 스택에서 헤더 압축과 이동성 측면에서 많은 문제를 가지고 있기에 PPP프로토콜의 헤더 압축을 위한 해결책의 문제점을 제시하였다. 마지막으로 우리는 예측 프로토콜 스택, 자원 효율성과 성능에 대한 All IP 네트워크 단말기 헤더 압축에 대한 guideline를 제안하였다.

Key Word : Header Compression, All IP Network, PPP, RTP, UDP

* 韓國電子通信硏究院
(Electronics Telecommunications Research Institute)
** 漢城大學校 情報通信工學科
(Dept. of Information and Communication Eng., Hansung University)
*** 建國大學校 電子工學科
(Department of Electronics Eng., Kon-Kuk University)

接受日: 2000年 8月25日, 修正完了日: 2001년 6月26日

I. Introduction

In the field of mobile communication, many engineers study and standardize the new concept about the next generation mobile network in order to improve the transport efficiency based on packet switch and support current and future IP-based user traffic and services in a globally common manner. This new concept, called by All IP Network [3], takes maximum advantage of existing and planned radio networks and provides a flexible and open API structure for the creation and support of new user services. However, there are many problems needed to solve in the technical view of All IP network. Due to the packet based migration of mobile network, mobility management and header compression are the first work item to evolve to the new generation mobile network for real time application services. Especially, Internet Engineering Task Force (IETF) has prepared for robust header compression over cellular link built using technologies such as WCDMA, EDGE, and cdma2000 as making a working group (WG). This WG has milestones such as the establishment of requirements for IP/UDP/RTP header compression and the introduction of IP/UDP/RTP header compression schemes submitted to Internet Engineering Study Group for publication by late 2000. In spite of bandwidth limitation, the goal of this working group is the development of header compression schemes that perform well over links with high error rates and long roundtrip times. By now, there doesn't exist the satisfied scheme over cellular links with high loss and long roundtrip latency. We suggest the protocol stack of All IP multimedia MS, the suitable header compression scheme and the guideline for All IP network MS header compression according to the expected protocol stacks

Therefore, through this paper, I will consider requirements for header compression over the cellular link in section II based on IETF. And also IP/UDP/RTP frame format is explained in section III including PPP of the protocol stack in MS for 3GPP2 3G cdma2000 packet data services. Especially, the issue in the aspect of the header compression and mobility management is described in this section. The remainder of this paper is organized with All IP network MS guidelines and conclusion.

II. Header Compression Requirements

The following requirements shown as very simple description in this paper have, more or less arbitrarily, been divided into three groups. The first group deals with requirements concerning the impact of a header compression scheme on the rest of the Internet infrastructure. The second group concerns what kind of headers that must be compressed efficiently. The final group concerns efficiency requirements and requirements which stem from the properties of the anticipated link technologies.

- Impact on Internet infrastructure

- Transparency : Through process of the compression / decompression, the resulting header must be identical to the original header.
- Ubiquity : Must not require modifications to existing IP (v4 or v6), UDP, or RTP implementations.

- Supported headers and kinds of RTP streams

- Ipv4 and Ipv6 : Must support both Ipv4 and Ipv6.
- Mobile IP : The kinds of headers used by Mobile IP{v4,v6} should be compressed efficiently.

- Efficiency

- Performance/spectral efficiency : Very good spectral efficiency same as legacy mobile network
- Error propagation : Should be kept at an absolute minimum
- Hand-off loss events : seamless hand-off
- Hand-off context recreation
- Link delay : Minimum
- Processing delay : Minimum
- Multiple links : The scheme must perform well when there are two or more cellular links in end-to-end path.
- Packet misordering - The scheme should be able to compress when there are misordered packets in the

RTP stream reaching the compressor.

- Configurable frame size fluctuation - It should be possible to restrict the number of different frame sizes used by the scheme.
- Residual errors - At most 10^{-05} bit error rate

III. Header Compression Issues

3.1 Header Format[4]

For the current All IP network service, protocol stacks of All IP MS is studied as RTP/UDP/IP/PPP/MAC/PHY shown in figure 1 in our research, and also can be changed arbitrarily by international standard organization. However, I would like to describe about frame formats of above protocol stacks.

Real time application
RTP
UDP
IP
PPP
MAC
Physical

Figure 1. Protocol stack of All IP network MS
그림 1. All IP 네트워크 단말기의 프로토콜 스택

a) RTP(Real-time Transport Protocol)

RTP provides end-to-end delivery services for data with real-time characteristics, such as interactive audio and video. Those services include payload type identification, sequence numbering, timestamping and delivery monitoring. Frame format of RTP is shown in Figure 2.

0 31

V	P	X	CC	M	PT	Sequence Number
Timestamp						
Synchronization Source(SSRC) ID						
Contributing Source(CSRC) ID						
Application Data						

V : Version=2 (2bit), P : Padding (1bit), X : Extension(1bit), CC : Contributor Count(4bit)
M : Marker(1bit), PT : Payload Type(7bit)

Figure 2. RTP Fixed header fields format
그림 2. RTP 헤더 필드 형태

b) UDP(User Datagram Protocol)

UDP is a simple, datagram-oriented, transport layer protocol: each output operation by a process produces exactly one UDP datagram, which causes one IP datagram to be sent. This is different from a stream-oriented protocol such as TCP where the amount of data written by an application may have little relationship to what actually gets sent in an IP datagram. Figure 3 shows the frame format of UDP.

0 15 16 31

source port number	destination port number
UDP-length	UDP checksum
Data	

Figure 3. Frame format of UDP
그림 3. UDP 프레임 형태

c) IP(Internet Protocol)

IP provides an unreliable, connectionless datagram delivery service. By unreliable we mean there are no guarantees that an IP datagram successfully gets to its destination. IP provides a best effort service. When something goes wrong, such as a router temporarily running out of buffer, IP has a simple error handling algorithm: throw away the datagram and try to send an

ICMP message back to the source. Any required reliability must be provided by the upper layers(e.g., TCP).

0 31

Version(4)	header length(4)	TOS(8)	total length(16)
Identification(16)	flags(3)	fragment offset(13)	
TTL(8)	protocol(8)	header checksum(16)	
source IP address			
destination IP address			
option(if any)			
data			

TOS : Type of Service, TTL : Time to Live

Figure 4. Frame format of IP
그림 4. IP의 프레임 형태

d) PPP

The Point-to-Point Protocol consists of the three components.

- A way to encapsulate IP datagram on a serial link.
- A Link Control Protocol (LCP) to establish, configure, and test the data link connection.
- A family of network control protocols(NCP) specific to different network layer protocols.

The frame format of PPP is shown in figure 5

flag	addr	control	protocol	information	CRC	Falg
1	1	1	2	up to 1500 byte	2	1

flag : 7E, addr : FF, control : 03,
protocol : 0021(IP datagram), C021(link control data), 8021(network control data)

Figure 5. Format of PPP frame
그림 5. PPP 프레임의 형태

3.2 RTP/UDP/IP Header Compression Issues

As the voice codec, G.729, an ITU standard of 8kbps packetized voice, is quite popular in landline IP telephony, such as that used for a 28.8 kbps modem connection. G.729 uses 10 ms per frame (80bits per frame for full rate transmission during speech and 15 bits per frame during silence) and 20 ms per packet(160 bits per packet for full-rate transmission and 30bits per packet for silence). Thus, 20bytes of packetized voice are transmitted every 20ms during speech. The short interval of 20 ms helps reduce latency, lessen the impact of packet loss, and improve voice quality. For wireless voice-over-IP to be economical, we must understand and analyze the additional overhead incurred by the headers of an IS-95 traffic frame, radio link protocol (RLP), point-to-point protocol (PPP), IP, and UDP. Since the header overhead of UDP, IP, and PPP is well known, we will focus on Layer 1 and RLP.[5] The information on minimum header sizes of UDP, IP, and PPP is from the IETF ROHC WG. Voice signals in IS-95 CDMA are digitized to produce voice traffic frames, and each frame is transmitted across the air interface at 20ms intervals. This short interval produces a voice quality comparable to that of wireline. CDMA maximizes the efficiency of RF resource usage by employing digital compression technology and exploiting some well-known voice activity patterns during voice conversations.[4] In digital compression, for example, only about an 8 kb/s transmission rate over the air is needed to achieve voice quality equivalent to landlines, where 64kb/s voice encoding is typically used. While full-duplex connections are necessary for a good experience in two-way conversation, most of the time only one end speaks and the other end listens. CDMA takes advantage of this by dynamically switching from the so-called full-rate frame (9,600b/s) during speech to a 1/8rate frame during silence (1,200b/s) which further improves the usage efficiency of RF power. As a result, the average traffic frame is only 3,936b/s every 20ms, even though the maximum is 9,600b/s. Similarly, the average RF resource usage is 3,325b/s every 20ms for information or voice bits, even though the maximum is 8,550b/s.

Four traffic rates are used at different levels of voice activity; their statistics, based on the Markov service option, may vary slightly from actual field data. The RF power resources consumed by 1/8rate frames during the 60% silence, or about 18%of the total RF power resources consumed by an average circuit-mode voice call (assuming that frame transmit power is proportional to frame rate), are not completely wasted. In fact, packet-mode voice may not be able to provide advantages of RF power resource efficiency compared to circuit-mode voice for three reasons.

First, the 1/8rate frames are needed to support traffic channel supervision when voice and signaling are carried over the same radio channel. The supervision procedures detect any loss of the signaling channel and remove calls with degraded RF conditions from the system.

Second, it is possible for 3Gsystems to carry packetized voice and signaling on separate radio channels. This scheme could enable a base station to selectively transmit nothing on a voice channel during silence, while maintaining traffic channel supervision and power control over a separate signaling channel. The new dedicated signaling channel, however, would still consume RF power resources during silent periods.

Third, if traffic channel supervision were somehow changed to allow a 3G system in a base station to turn off its voice/signaling channel during silence without experiencing the adverse effects described above, any advantages that might benefit packet-mode voice could also benefit circuit-mode voice through new circuit voice options defined to fully exploit this change. To compete with circuit voice on both quality and price, any last-hop voice-over-IP scheme using CDMA must meet the following two challenges:

- A voice packet must be transmitted every 20 ms during speech, and
- The maximum RF bandwidth consumption must not exceed 9,600 b/s, with an average of 3,936 b/s.

This is not possible, given that the overhead of uncompressed UPD/IP/PPP headers is already 33 (8+20+5) bytes, or 264 bits, per packet, far exceeding the192 bits of an existing CDMA full-rate traffic frame. This calculation does not even consider the RLP overhead and the packetized voice payload of 160 bits using the G.729 format. To close up the gap, the UDP/IP overhead must be significantly reduced, leading to 3G All IP Network Design.

3.3 RLP Traffic Frame Issues

This section analyzes the RLP, which sits on top of IS-95. Figure 6 depicts the IS-95 RLP full-rate traffic frame and two of its formats, A and B. The latter supports the so-called transparent mode. In this mode the layer above RLP will be responsible for retransmission if a transmission error occurs over the air. Each format is 171bits in length which, as a payload, fits into the IS-95 CDMA Rate Set 1 traffic frame shown in Figure6. During speech (40% of the time), Format Bwhose fixed payload is 160bitscan be used to carry the voice-over-IP information. However, packetized voice using G.729 is 160bits every 20ms, and the overhead of uncompressed UPD/IP/PPP/RLP is 275(33 * 8+8+3) bits, making it impossible to squeeze the 435(160+275) bits into the 160data bits of Format B without using more RF resources. If a voice-over-IP designer were allowed to allocate more RF resources and still keep the 20ms per packet requirement, then the 9,600b/s full-rate traffic frame would be replaced by a new 22,800((192+33 * 8) * 50)b/s full-rate traffic frame, at least doubling the RF resource usage when compared with the circuit voice counterpart. This 22,800b/s full-rate traffic frame is purely hypothetical; no such frame exists in current IS-95 CDMA standards. In other words, CDMA voice-over-IP for licensed cellular/PCS radio spectrum is either uneconomical at 20ms per packet, or the use of longer packetization intervals will result in poor voice quality because of increased latency and the impact of packet loss. However, the situation could be improved significantly by combining some deep UPD/IP header compression and other techniques, as we discuss below.

Format A

8	1	1	6	Up to 152 bits	Variable
SEQ	CTL	RFI	LEN	Data	Padding

Format B

8	160bits	3
SEQ	Data	Type

CTL : RLP Frame Type, RFI : Retransmitted frame indicator, SEQ : Sequence Number, LEN : Length

Figure 6. RLP full-rate primary traffic and its two possible formats
그림 6. RLP 프레임의 두 가지 형태

a) An Analysis with Optimistic Assumptions

To assess the theoretical bounds and thus the ultimate possibilities for voice-over-IP using CDMA, we may take an optimistic view with some hypothetical assumptions. Suppose:

- Deep UDP/IP header compression were able to reduce the UDP/IP overhead from 28(=8+20)bytes to only 1byte; and
- A new, packet-centric variation of CDMA (non-IS-95) were invented that could squeeze out an extra 16bits (for example, the SEQ field of Format B) from the 192bits of the 9,600b/s full-rate traffic frame. (Some of the 192bits must be reserved for CDMA and RLP use.)

If we were to make these assumptions, the extra 16bits could be used to carry a portion of UDP/IP/PPP headers that now would be only 6bytes (1byte of compressed UDP/IP header plus 5 of PPP). If a 60 ms packetization interval were used, the 6 bytes of overhead could be amortized over three full-rate traffic frames, each transmitted every 20ms at 9,600b/s. While the voice quality using a 60 ms interval would not be as good as that of a 20 ms interval, the RF resource usage would be the same. If, on the other hand, the 20 ms interval were kept, the 9,600b/s frame would be replaced by a new 11,200((192+4 * 8) * 50)b/s full-rate traffic frame, representing a 16% increase in RF usage. If only UDP/IP compression is used without the new, packet-centric scheme, a 12,000((192+6 * 8) * 50)b/s full-rate traffic frame would be needed instead of the 9,600b/s, representing a 25% increase in RF usage.

b) Header Compression in the 3G LAC Layer

One alternative to having PPP perform header compression is to compress the headers of UDP, IP, and PPP in the 3G LAC layer. To reduce the 33bytes of the UDP/IP/PPP headers down to, say, 6bytes, the LAC software needs to peek a few levels into the PPP payload. It also needs to make sense out of PPP states, IP addressing, and UDP states, as well as many other things that PPP is designed to do in the first place. (For example, the PPP performs Van Jacobsons TCP/IP header compression, which reduces the number of TCP/IP headers from 40bytes to 5).[7] To achieve these results, certain software functionality would have to be duplicated in PPP and LAC, which could be an issue for the mobile device. Even if this LAC approach does not present a big problem in the voice-over-IP case, its applicability is likely to be limited for multimedia-over-IP, which requires a deep understanding of all the involved signaling protocols (such as voice, video, and data). After all, the LACs major role is to provide reliable transmission over the air, even if its use makes it necessary to reserve some additional bits for its header.

IV. Guidelines for All IP Network MS Header Compression

a) General Guidelines

- If applications like VoIP is important, All IP systems should be designed to support the types of applications that require low end-to-end latency.
- If multimedia-over-IP in general and voice-over-IP in particular are to provide the transparent quality services that mobile users have come to expect, 3G

systems should be designed to support packet data mobility in a manner as seamless as that of its voice counterpart in various handoff scenarios.

- RTP/UDP/IP header compression must be performed, requiring cooperation between the mobile and the network, to make wireless VoIP economical.
- RTP/UDP/IP header compression should be performed within the wireless network to make it completely transparent to external networks.
- All IP networks should be designed to efficiently support the second model of VoIP, which use circuit voice over the air and performs voice-to-IP conversion within the wireless access network.
- Must provide low relative overhead under expected operating conditions; compression efficiency should be better than for previous header compression scheme under equivalent operating conditions. The error rate should only marginally increase the overhead under expected operating conditions. In other word, spectrum efficiency is a primary goal.

b) Special Guidelines

- Deployment of header compression algorithm over the air must be performed by the phased evolution of All IP Network Architecture Model.
- Although ROHC WG define and support just transparent header compression in the context, we must consider the non-transparent header compression with spectral efficiency.
- The Interface between PPP and RLP will be developed; the modification of RLP and the link layer frame overhead will be minimized by the negotiation between 3GPP2 and IETF.
- The IP layer and link layer QoS and over-the-air QoS shall be coordinated to ensure the seamless mobility and the reliable performance of VoIP.
- The security broken due to header compression and decompression will be confirmed in MS and All IP networks

V. Conclusion

As the next generation wireless mobile infrastructure, All IP systems provide high speed data bit rate over the air and advanced service such as multimedia, Internet access, and seamless roaming and mobility across the ANSI-41 domain and IP multimedia domain networks [8]. In current 3GPP2 All IP network standardization, the architecture of All IP network must be evolved to protect IMT-2000 carrier and legacy network carrier. Evolution, rather than revolution, from 2G or 3G to All IP means that All IP multimedia services may be provided more cost effectively by packet based multimedia or wireless multimedia over IP, as opposed to the more expensive ATM revolutionary approach. In our study deep header compression within All IP network is a must for multimedia over IP to compete effectively with ATM based multimedia. The investment of header compression will be benefit not only multimedia over IP, but also voice over IP and IP over PPP. The compression should be carried out layer 2, in a manner completely transparent to the external networks. While such compression will also help packet mode voice, wireless voice over IP in general may not be economical for licensed various cellular radio spectrum because the cost of licensing the RF spectrum is high and the circuit mode air interface is already efficient. However, the effort to overcome technical problems must be endless and a more practical approach is to use circuit voice over the air and perform voice to IP conversion in the wireless network. The packet mode multimedia service of All IP network in the next generation wireless network, betting on packet data mobility in addition to voice mobility may be a key element for the wireless industry in fueling the explosive growth of mobile and Internet subscribers worldwide.

References

[1] Over-the-Air Voice over IP(VoIP) Issues and recommended Phases (ALLIP-20000323-006_QUA_VoIP_Issues). Hsu, Odenwalser, Chen, Tiedemann, 3GPP2 All IP, March 2000

[2] 3GPP2 " IP Network Architecture Model for cdma2000 Spread Spectrum Systems" Draft revision

0.8.0(NAM Focus Group) 3GPP2 SC.P000x Oct. 2000
[3] 3GPP " Architecture for All IP Network", 3G TR23.922, version1.0.0 Oct. 1999
[4] W. Richard Stevens, TCP/IP Illustrated, Vol. 1: The Protocols, Addison-Wesley, Reading, Mass., 1994.
[5] End-to-End QoS for VOIP and Multimedia in 3GPP2 ALL IP Networks (SC-ALLIP-20000426-16009_LUC_End-to_End), Tom Hiller, Pete McCann, 3GPP2 All IP, April 2000
[6] Enhancements to IP/UDP/RTP Header Compression (draft-koren-avt-crtp-enhance-01.txt), Koren, Casner, Ruddy, Thompson, Tweedly, Geevarghese, March, 2000 (IETF draft)
[7] Robust Checksum-based header Compression (ROCCO) (draft-jonsson-robust-hc-04.txt), Lars-Erik Jonsson, Mikael Degermark, Hans Hannu,Krister Svanbro, March 10, 2000 (IETF draft)
[8] Adaptive Header Compression (ACE) for Real-Time Multimedia (draft-ace-robust-hc-01.txt), Le, Clanton, Liu, Zheng, March 2000 (IETF draft)

저 자 소 개

崔 尙 鎬 (正會員)
제 5권제 1호 논문 00-02-09 참조
현재 한국전자통신연구원 선임연구원

扈 光 春 (正會員)
제 5권제 1호 논문 00-02-09 참조
현재 한성대학교 정보통신공학과 조교수

金 暎 權 (正會員)
제 5권제 1호 논문 00-02-09 참조
현재 건국대학교 전자공학과 교수