PROVIDING VOIP SERVICE IN UMTS-HSDPA WITH FRAME AGGREGATION

Gee Rittenhouse

Wireless Research Laboratory
Bell-Labs, Lucent Technologies
Murray Hill, NJ, U.S.A
Email: gee@lucent.com

Haitao Zheng

Microsoft Research Asia
49 Zhichun Road, Haidian District
Beijing, P.R. China
Email: htzheng@microsoft.com

ABSTRACT

In this work, we evaluate the performance of Voice over Internet Protocol (VOIP) services over high speed downlink shared channel (i.e. UMTS-HSDPA), that is used in third generation cellular systems to provide high speed packet data service. We focus on examining system outage based on delay budget in radio access network. We conclude that VoIP system performance is highly sensitive to delay budget. In order to support VoIP services with end-to-end delay comparable to that of circuit switched voice, HSDPA system needs to be carefully engineered, and design alternatives that relax delay budget in radio access network could be considered.

1. INTRODUCTION

Future generation radio-access technology will extend to an IP based transport and service platform that offers mobile users a multitude of real-time and interactive services. Third generation wireless networks are evolving towards a packet-switched radio access network, marked by recent deployments of 1X-EVDO/DV and future plan of high-speed downlink packet access (HSDPA)[1, 2]. One candidate service would be VoIP, which can relieve capacity from the circuit switched network, and save network operating and maintenance cost. In order for VoIP to be a viable alternative to traditional circuit switched voice, it needs to retain or at least approach the performance exhibited by the corresponding circuit voice. We have previously studied the performance aspects of providing VoIP of HSDPA [8]. In particular, we examine the network capacity of VoIP service under very tight delay requirements (20-80ms in radio access network), and conclude that the network capacity is less than that in circuit switched systems. One fundamental problem is the mismatch between the size of VoIP packet (40byte) and that of HSDPA MAC frame (usually larger than 100bytes). This often results in uncessary bit-stuffing and thus under-utilization of radio resources.

Recently, there have been new proposals to aggregate multiple voice packets (frame aggregation) to reduce the occurrence of bit-stuffing and improve resource efficiency. The penalty is extra packet delay due to buffering multiple frames. As a result, the delay budgets should be relaxed. In this paper, we continue previous effort and focus on the study of VoIP performance with relaxed delay budgets. Particulary, we investigate the impact of frame aggregation scheme, admission control policy, multiplexing capability and multiple antenna on overall system performance.

We assume a fixed-to-mobile scenario, where voice calls are originated from land phones to mobiles. VoIP packets traverse wired IP network, wireless core network, until they reach radio access network (i.e. base station), which forwards the packets to mobile users. We assume that the transmission path from wired IP network to base station has QoS support so that packets encounter fixed delay. This allows us to focus our study on packet delay within radio access network, particular in downlink direction, from base station to mobile users[1].

2. END TO END VOIP COMMUNICATIONS

We begin with a brief introduction on some fundamental elements of VoIP service, including codec and protocol, performance metrics and delay components.

2.1. Codecs and Protocols

Voice packets usually are generated at a constant rate during talk spurt, *e.g.*every 20ms. There are a number of voice coding standards, with different data rate, quality and complexity. Each voice packet consists of a payload representing voice signal and a header that contains protocol information. The header fields include overhead required for RTP/UDP/IP, PDCP and RLC. The combined RTP/UDP/IP header contribute a overhead of 40bytes for IPv4 and 60 bytes for IPv6, respectively. While voice payload typically ranges between 20 and 40 bytes for typical 8-16kbps AMR codec, the RTP/UDP/IP overhead is substantial and would consume bandwidth unnecessarily when transmitted over the air interface. Efficient and robust header compression techniques, *i.e.* Robust Header Compression(ROHC)[3] can reduce RTP/UDP/IP headers to 1 (IPv4) and 3 bytes (IPv6).

The data rate required for conveying VoIP packets is in general less than 20kbps, much smaller compared to typical peak data rate (0.5-2Mbps) in HSDPA. Transmitting one VoIP packet during each MAC frame (2ms) would under-utilize bandwidth. It is thus natural to aggregate several successively generated voice frames into one VoIP packet in order to improve bandwidth utilization. However, frame aggregation introduces buffering delay, *i.e.* aggregating N frames into one packet leads to an extra delay of $(N-1)$ frame interval. In addition, a single packet loss results in N frame burst loss which is in general difficult to conceal compared to single frame loss.

[1] Although VoIP performance depends on both downlink (i.e. base station to mobile users) and uplink (reverse direction) performance. However, in this work, we only consider the impact of downlink transmissions. A comprehensive study that considers both directions will be included in a future study.

Table 1. Delay Component

Source	Delay Component	Simulated value (ms)
Calling party	Voice encoding	20
	Aggregation	0 (1 Frame AGG)
		60 (4 Frame AGG)
	Protocol(RTP/UDP/IP)	N/A
IP backbone network	Processing	N/A
	Transport	N/A
Base station	Header Compression	N/A
	RLP+MAC processing	4
	Scheduling + HARQ	Variable
	Propagation	N/A
Terminal	UE processing	4 + (queuing delay)
	Playback buffering	N/A
	Voice decoding	N/A

2.2. Performance Metric

The VoIP performance objectives are derived from those observed for circuit switched voice.

Delay: In telephony terms, delay is the measure of time that it takes the talker's voice to reach the listener's ear. High end-to-end delay leads to echo and talker overlap.

Jitter: Since voice packets are generated at a constant rate, each packet would be expected to arrive at the destination exactly every 20ms. However, this is not always the case. since network devices cause an unpredictable amount of delay to occur between these packets. Receiver play-out buffers can be used to smoothen play-out of packets, but yield extra delay and packet loss.

Packet Loss: As packet switched networks do not guarantee that packets will be delivered at a timely manner, packets will be dropped under heavy network traffic loads and during periods of congestion. Although packet loss of any kind is undesirable, some loss can be tolerated. Packet-loss-concealment techniques are effective when packet loss is limited to a few packets. In general, the voice quality is not adversely affected as long as the amount of packet loss is less than 2-5% for the total number of calls[4, 5].

2.3. End-to-End Delay Components

End-to-end delay is the measure of time that it takes the talker's voice to reach the listener's ear. In this study, we consider a fixed-to-mobile scenario where delays can be attributed to caller, IP backbone network, cellular access network and mobile terminal. Among them, delays introduced by codec, buffer, protocol processing are predefined by standards and implementations, while others particularly those within cellular access network depend on the network topology and design choices. Table 1 summarizes the delay components. The delay marked as N/A is in general bounded by a fixed value, (e.g. one way delay in IP backbone network is in general less than 32ms from coast to coast) and hence not considered in this study. Such simplification allows us to focus on the delay within radio access network.

3. HIGH SPEED SHARED PACKET CHANNEL

In HSDPA, mobile users share downlink transmission in a time and code multiplexing manner. In each transmission frame, MAC scheduler determines the users to be served. While traditional CDMA systems dedicate a Walsh code to each user, HSDPA assigns Walsh codes dynamically over time. In each MAC frame, only a subset of users who are experiencing good channel conditions will be served. HSDPA also employs link adaptation and hybrid ARQ(HARQ) modules for throughput enhancement. Next, we briefly describe these modules.

3.1. Link Adaptation and Hybrid ARQ

While traditional CDMA systems use power control to mitigate channel fading, HSDPA employs rate control based link adaptation where base station transmits at full power and adjusts modulation and coding scheme (MCS) according to channel variations, maximizing instantaneous usage of the wireless channel. An important design choice of link adaptation, is the mapping between MCS and the measured signal-to-interference-noise-radio (SINR). In general, the mapping is set to keep transmission failure below a certain value (*i.e.* 1%.

The performance of link adaptation largely depends on the accuracy of SINR measurement, which is difficult to maintain as mobility increases. HARQ scheme is thus introduced to recover transmission failures. When mobile detects a transmission failure (through cyclic redundancy check(CRC)), it sends a request to base station for retransmission. Mobile combines soft signals of both original and subsequent retransmissions, and the combined signal has higher probability of being successfully decoded.

3.2. Packet Scheduling

Packet scheduling takes advantage of independent temporal channel variation at each mobile, by giving priority to users with transitorily better channel condition. Design choice of scheduling impacts throughput and latency performance, as well as fairness. We choose to use Proportional Fair scheduler [6] that selects the mobile with the maximum ratio between instantaneous service rate and the average service rate in the past. Statistically, for a symmetric system where mobiles experience the same average channel condition, PF scheduler distributes transmission time uniformly among mobiles. Each mobile is served when its channel condition is above average. By favoring the poor mobiles (who in bad channel condition) less emphatically, it provides an attractive trade-off between cell capacity and user fairness.

3.3. Code Division Multiplexing

Unlike 1X-EVDO system who transmits to only one mobile at any given time, *i.e.* time division multiplexing(TDM), HSDPA allows simultaneous transmissions to multiple mobiles. This is done through a hybrid code division multiplexing (CDM) and TDM. Base station serve a subset of mobiles in a single MAC frame using different set of Walsh codes for each mobile. TDM can maximize system throughput by exploring temporal channel variations; while CDM provides a constant rate that is important for real-time, constant bit rate services like VoIP and streaming video. The major concern of CDM is implementation complexity and protocol overhead due to maintaining multiple ”connections”.

4. PERFORMANCE EVALUATION

In this section, we evaluate VoIP performance in a simulated HSDPA network. We have developed an Opnet based simulation platform that captures dynamic processes in a radio access network. We reuse the standard RTP/UDP/IP modules in Opnet and develop MAC and PHY modules based on HSDPA specifications [7]. To reduce simulation scale, we reduce HSDPA bandwidth from 5MHz to 1.25MHz. As stated before, we focus our study on downlink performance and assume that uplink (mobile to base station) operates at a constant data rate of 64kbps without any transmission error. We choose 8bps AMR vocoders with 20ms voice frame, 50% voice activity. ROHC header compression reduces RTP/UDP/IP header to 2 bytes.

We consider a fixed-to-mobile scenario. There are multiple mobiles in the network, each running a VoIP session connecting to a VoIP source (Internet user). For each VoIP session, the performance is the packet loss rate for a given delay budget. A packet loss occurs when the packet is not delivered to voice decoder by the end of the packet's usefulness, i.e. the packet generation time plus the delay budget. We focus on packet delay within radio access network and exclude the delay due to IP backbone and play-out buffer. Values of other delay components are summarized in Table 1. Given the number of VoIP sessions and delay budget, the system performance is the percentage of sessions experiencing more than 2% of packet loss, hereby referred to as outage.

In each simulation, we randomly place mobiles in a cell[2]. Transmissions to each mobile experience frequency flat rayleigh fading at 3km/h. Mobiles estimate channel SINR based on pilot signal, and feedback the estimation to base station. Hence, base station performs link adaptation based on SINR estimated at previous MAC frame, which inevitably leads to errors. HARQ is used to recover channel errors, with maximum 3 retransmissions for each packet. Admission control is based on the distance to the base station, represented by geometry value (*i.e.* the average SINR that the mobile experiences). Only mobiles with geometry value larger than G will be admitted to the system. For each deployment, we record the system outage, *i.e.* the percentage of sessions experiencing more than 2% packet loss under different delay budgets. The overall outage is averaged over 100 mobile deployments.

Next, we examine the impact of frame aggregation, G, CDM capability (the maximum number of mobiles to be served in each MAC frame) and multiple antennas on VoIP performance. By default, we assume there are 50 mobiles, G=-5dB, CDM capability of 2 and single antenna on both mobile and base station.

4.1. Impact of Frame Aggregation

We first examine the impact of frame aggregation on system performance. Fig. 1 compares outage performance when aggregating 1, 2 and 4 voice frames into one packet. Frame aggregation increases packet arrival interval and packet size to utilize peak channel rate efficiently. Since service rate (*i.e.* peak channel rate) is in general much larger than VoIP data rate, a VoIP packet can be transmitted using one MAC frame. Hence, aggregating voice frames is equivalent to reducing packet arrival rate. As a result, transmission delay is smaller. However, results show that reduction of transmission delay can not fully compensate the aggregation delay. We also observe that mobiles with low geometry (those who are far away from base station) benefit from frame aggregation due to significant improvement of transmission delay, while others still suffer from the penalty of aggregation delay. This can

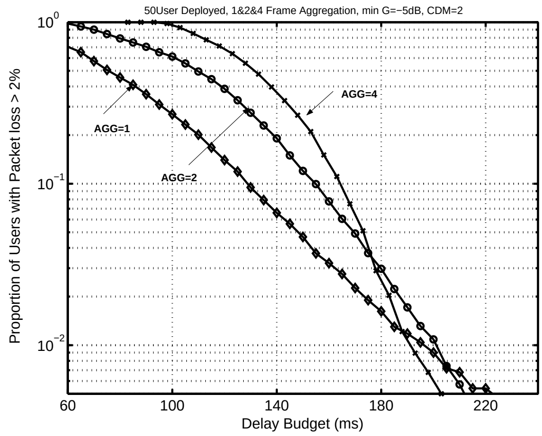

Fig. 1. Outage vs. delay performance with different frame aggregation.

be observed from the improved performance of 2 and 4 frame aggregations at large delay budgets (around 200ms) and low outage rate (1%).

4.2. Impact of Admission Control and Service Coverage

From simulations, we observe that system outage is critically limited by the performance of mobiles with low geometry. Traditional CDMA systems use soft-handoff technique to enhance their channel conditions. As HSDPA supports fast cell switching instead of soft-handoff, we can use distance/location based admission control to address such limitation[3]. Fig. 2 compares the outage performance, when minimum geometry G is set to C15dB, -5dB and C1dB. Clearly, the outage performance is quite sensitive to the choice of G. It should be noted that increasing G is equivalent to reducing service coverage. In this example, the service coverage are 100%, 97%, and 70% for those G choices, respectively.

4.3. Impact of Code Division Multiplexing Capability

CDM capability is critical to VoIP service. Theoretically, CDM capability should be as large as possible. However, the caveat of serving multiple users simultaneously is increased signaling and hardware complexity. As signalling and data share radio resources like power and Walsh codes, signalling consumes additional bandwidth and would degrade the performance of data transmission. Fig. 3 compares the outage performance when CDM capacity is set to 1 and 2. We observe that setting CDM capability to 2 significantly reduces packet delay. Using 10% outage as a reference, the delay is reduced from 320ms to 130ms.

4.4. Impact of Antenna Diversity

During simulations, we found that most VoIP packets were delivered within a reasonable time. Delay spike only happens when a mobile experiences several channel fading. One solution is to use diversity techniques to reduce the occurrence of channel fading. In this work, we choose to use one additional antenna at the mobile

[2]In this work, we only investigate single cell performance without considering any handoff. Multi-cell VoIP performance has been addressed in another study.

[3]Such assumption is apparently only suitable for immobile system. Investigation of handoff and fast cell switching will be included in another study.

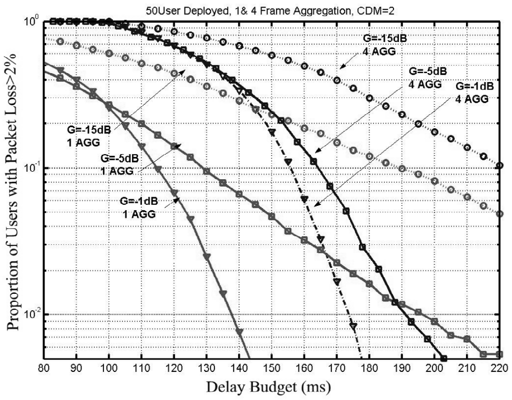

Fig. 2. Outage vs. delay with different minimum geometry G.

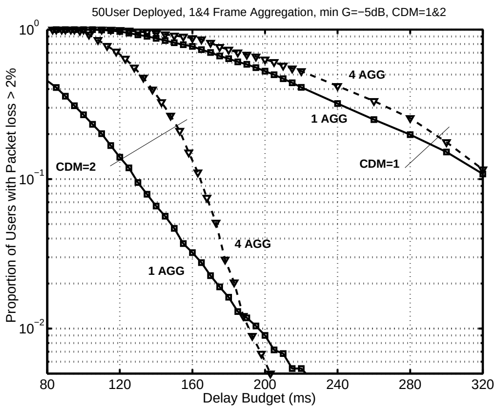

Fig. 3. Outage vs. delay with different CDM capabilities.

(*i.e.* (1,2) mobile). We define system capacity as the average number of VoIP sessions supported with 5% overall outage (2% outage in VoIP session, 97% coverage). Fig. 4 compares system capacity as a function of delay budget for different design choices. We see that using two antennas at mobiles (1,2) results in more than 100ms delay reduction compared to using single antenna (1,1).

This figure also shows that VoIP capacity is highly sensitive to delay budget. Relaxing delay requirements in radio access network provides significant capacity improvement. For example, increasing delay budget from 80ms to 100ms could increase capacity from 20 to 32 under normal network setting (no aggregation, (1,1), 2% packet loss based voice quality). Hence, an alternative to improve VoIP performance could be seeking delay improvement in core network or IP backbone network, to compensate for delay relaxation in radio access network.

5. CONCLUSION

In this work, we conduct simulations to evaluate the performance of VoIP services over HSDPA. We focus on examining packet delay within radio access networks, and system outage under different delay budgets. Impact of critical design choices like frame aggregation, admission control, multiplexing capability and antenna diversity is discussed. We conclude that VoIP capacity is highly sensitive to delay budget, and a small relaxation (20ms)

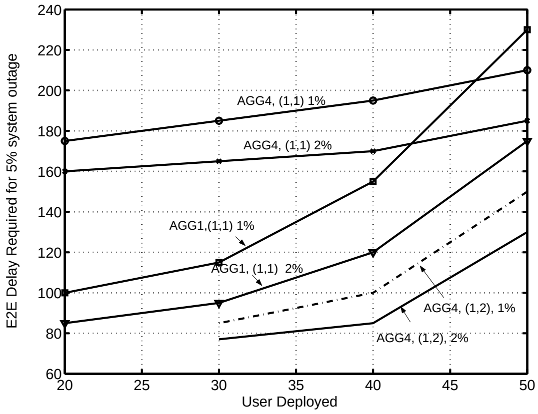

Fig. 4. System capacity vs. delay with different antenna configurations, frame aggregation schemes (1%, 2% represent the packet loss bound).

of delay budget can lead to more than 50% of capacity improvement. As system performance is critically limited by mobiles in bad channel condition, location based admission control, frame aggregation and antenna diversity can be utilized to mitigate channel fading. Overall, although HSDPA and other packet shared channel based system are promising choices for future data services, they need to be carefully engineered in order to support real-time delay-sensitive applications like VoIP and streaming multimedia. Alternatives (*e.g.* core network improvement) that relax the delay constraint within radio access network could be considered to improve VoIP end-to-end delay.

6. REFERENCES

[1] 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Physical Layer Aspects of UTRA High Speed Downlink Packet Access; (Release 2000), (3G Technical Report (TR) 25.848)

[2] 3GPP2 , CDMA2000 high rate packet data air-interface specification, Technical Specifications, Document No. C. S00024, 2002.

[3] "Robust Header Compression (ROHC)," IETF RFC 3095, July 2001.

[4] ETIS, "New Report on VOIP Quality," Report of ETSI Project Telecommunications and Internet Protocol Harmonization Over Networks (EP TIPHONTM), available at http://www.etsi.org/plugtests/04History/2002_voipsqa.htm.

[5] T. Enderes, S. C. Khoo, C.A. Somerville and K. Samaras, "Impact of statistical multiplexing on voice quality in cellular networks," *Mobile Networks and Applications*, Vol. 7, pp. 153-161, April 2002.

[6] P. Viswanath, D. Tse and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inform. Theory*, 2002.

[7] 3GPP, UTRAN High Speed Downlink Packet Access: UTRAN Overall Description, 3GPP TS 25. 308.

[8] H. Zheng et.al, "The Performance of Voice over IP over 3G Downlink Shared Packet Channels under Different Delay Budgets," in *Proceedings of IEEE VTC Fall*, 2003.