

VoIP End-to-End Performance in HSPA with Packet Age Aided HSDPA Scheduling

¹Tao Chen, ²Gilles Charbit, ³Karri Ranta-aho, ⁴Oscar Fresan, ⁵Tapani Ristaniemi
^{1,2}Nokia Devices, ³Nokia Siemens Networks, ⁴University of Tampere, ⁵University of Jyväskylä
tao.t.chen@nokia.com

Abstract—In this paper, we present an enhanced VoIP scheduling for the high speed downlink packet access (HSDPA) in UMTS, which takes the age of the VoIP packet into account. The downlink capacity can be significantly improved by this way, especially for shorter uplink transmission delay. In order to quantify the achievable performance improvement, we present results obtained from extensive system-level uplink and downlink simulations. Inter alia, it is shown that using the proposed scheme can lead to an increase in the downlink cell capacity of up to 16%. By applying the proposed method, the downlink performance can be improved considerably while the uplink performance remains the same, which ensures the good end-to-end performance.

Index Terms—Scheduling, VoIP, HSUPA, HSDPA

I. INTRODUCTION

High-speed Uplink Packet Access (HSUPA), which is included in 3GPP Release 6 specification (see, e.g., [1],[2]), has been designed to increase the uplink packet data throughput by means of Hybrid Automatic Repeat request (HARQ), shorter frame size and fast scheduling controlled by the base station. High Speed Downlink Packet Access (HSDPA) is the downlink counterpart of HSUPA in 3GPP Release 5 (see, e.g., [3]). Furthermore, both HSUPA and HSDPA have been improved in Release 7, which is so called High-speed Packet Access (HSPA). Recently, there has been an increasing interest to use HSDPA/HSUPA for real-time (RT) packet switched (PS) services such as Voice over IP (VoIP). The reasoning behind the increased interest towards VoIP is to use VoIP in All-IP network instead of using circuit switched (CS) speech, which will imply cost savings for operators as CS related part of the core network would not be needed anymore. The challenge for the VoIP on HSPA will be the delay requirement for the voice traffic.

Some enhanced HSDPA VoIP schedulers considering the packet delay requirement have been developed to offer VoIP services more efficiently. However, this kind of static delay budget used for scheduling is mainly based on a rough estimation, which is typically not accurate enough due to the retransmissions and varying channel conditions. Furthermore, an overestimated delay budget may lead to poor end-user

experience, whereas an underestimated delay budget may sacrifice the radio efficiency. Therefore, a packet age aided scheduling scheme is proposed in this paper by exploiting the packet delay information in order to improve VoIP performance over HSPA.

The rest of this paper is organized as follows. In Section II, the principle of VoIP over HSPA is described, and an enhanced scheduling scheme in HSDPA aided by the packet age is proposed. Then, the system modeling is described in Section III whereas both uplink and downlink simulation results are presented and analyzed in section IV. And finally, the conclusions are drawn in Section VI.

II. PRINCIPLE OF VOIP OVER HSPA

A. VoIP Packet Delay requirement

VoIP as a real-time service requires the transmission and reception within the reasonable time. According to [4], the maximum acceptable mouth-to-ear delay for voice is on the order of 250 ms. If the delay for core network, RNC processing and Iub transport is approximately 100 ms, the one-way tolerable delay for MAC buffering/scheduling, transmission and reception would be roughly 150 ms assuming VoIP call is between two mobiles. Furthermore, this maximum allowed delay is typically used as the delay budget for the evaluation of VoIP service over HSPA.

B. VoIP over HSPA

HSPA was originally designed to carry high bit rate delay tolerant data. A number of features have been introduced to 3GPP Release 6 and 7 to improve the efficiency of low bit rate delay critical applications like VoIP. In [5], the VoIP enabling features such as packet bundling, uplink gating for HSUPA, non-scheduled HSUPA transmission and advanced HSDPA receivers are summarized. Uplink gating, which will disable DPCC transmission when there is no data to be sent on E-DCH and no L1 feedback signaling on High-speed Dedicated Physical Control Channel (HS-DPCC), can minimize the L1 control overhead, reduce uplink interference and save UE power consumption. Packet bundling by sending multiple VoIP packets together can minimize the control overhead improving the performance in both uplink and downlink.

Besides, some VoIP delay sensitive scheduling algorithms in HSDPA (see [6], [7]) were proposed and evaluated

previously, in which the power allocation, codes allocation and user multiplexing are scheduled considering the VoIP packet delay requirement as well as the channel conditions. The channel conditions are obtained from the Channel Quality Indicator (CQI) report transmitted on uplink HS-DPCCH channel. The packet delay requirement used for the scheduling is just a static delay budget, e.g., 80ms, which is based on a rough estimation by dividing 150ms end-to-end air interface delay budget into uplink and downlink components. It has been proved that taking the packet delay requirement into account in the downlink scheduling can improve the VoIP performance while retaining the service quality (see [7]).

III. VOIP PACKET AGE AIDED HSDPA SCHEDULING

Even though early scheduling schemes provides some improvement in the downlink capacity by considering the packet delay requirement, the estimated static delay budget applying for all VoIP packets may not be reasonable, because it is supposing that all VoIP packets would suffer similar delay before arriving at NodeB scheduling queue. Hence, there is no difference on the priority in the terms of the packet delay for VoIP packets. Furthermore, this kind of static delay budget leads to the independent downlink scheduling without considering any end-to-end user experience. For example, the packet suffering shorter uplink delay can be transmitted in the downlink earlier than the packet suffering longer uplink delay due to the better channel condition and the same delay budget. Further, it may result in the outage of the latter because the overall delay may exceed 150ms even though the downlink transmission is done successfully within 80ms static delay budget. Besides, each packet arriving at NodeB scheduling queue may suffer different delay in practice. The suffered delay would vary with the uplink sources. For example, the uplink delay from the mobile would be larger than from the fixed-line phone because of the retransmission in the complicated channel conditions. Even for mobiles there can still be some difference on the uplink transmission delay due to the varying channel conditions. Therefore, it does make sense to exploit the specific delay budget for each packet in order to fully utilize the radio resource efficiently for higher VoIP capacity while maintaining end-user good experience.

Here, an enhanced HSDPA scheduling aided by the packet age is proposed. The downlink delay budget used for scheduling would be derived from the one way end-to-end delay budget and the “age” of the VoIP packet as expressed in equation (1).

$$PDB_{ui}^{pj} = E2E_{ui}^{pj} - PAge_{ui}^{pj} = 250\text{ ms} - PAge_{ui}^{pj} \quad (1)$$

where PDB_{ui}^{pj} stands for the delay budget used in the downlink scheduling for packet j of user i , $E2E_{ui}^{pj}$ the end-to-end delay budget for packet j of user i and $PAge_{ui}^{pj}$ the age of packet j of user i . Here, the end-to-end delay budget is the maximum acceptable mouth-to-ear delay. The packet age is defined as the accumulative delay from the generation of the

VoIP packet to the instant of packet arriving at the end-user serving NodeB in the downlink.

Due to the trend of migrating to the ALL-IP network, the packet age information can be derived from the time-stamp in the IP header. The birth of the packet can be marked by UE in the uplink through the IP layer or by the calling node in the IP network. In IP header (see [8]), there is an option for the internet timestamp in milliseconds as shown in Table I.

TABLE I.
TIMESTAMP IN IP HEADER

1000100	Length	Pointer	OfIw	flg
Internet address				
Timestamp				

Thus, for example, the UE packet timestamp information in 3G-GGSN (see Fig. 1) could firstly be obtained by calculating the accumulative delay according to the time difference between current time and the packet timestamp. Then, the timestamp information can subsequently be appended for each hop. Finally, the NodeB may derive the age of the each packet according to IP header in the case of direct IP connection to NodeB, assuming that most RNC functions have been moved to Node-B and IP transmission is supported.

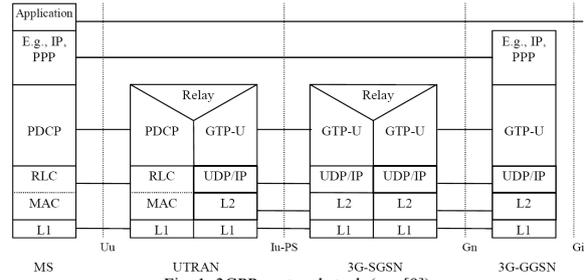


Fig. 1. 3GPP protocol stack (see [9])

If RNC and NodeB functions are split and Iub IP transmission is not supported, we may attach the age information into the VoIP packet via the extended fields of frame protocol.

Each VoIP UE can have its own specific delay budget for HSDPA scheduling. The serving NodeB can prioritize packets to the UEs based on the specific delay budget of each packet. Otherwise, the scheduler can even discard the packet directly when the packet age is exceeding the 150ms end-to-end delay budget. So it is a better use of resources to avoid transmitting a VoIP packet that is known to have higher end-to-end delay than what is required for acceptable end-user experience.

The scheduling for VoIP users aided by the packet age is such that, it selects at most M users, which is corresponding to the maximum allowed multiplexing user number, from a scheduling candidate set (SCS) by using Proportional Fair (PF) packet scheduling algorithm which exploits the instantaneous channel state information for each user, and therefore provides a multi-user diversity gain. The selected VoIP users are to be transmitted in the following 2ms Transmission Time Interval (TTI).

The SCS includes VoIP users that fulfill the following

criteria:

(i) Users whose packet delay is equal to or larger than a threshold corresponding to the packet specific delay budget.

(ii) Users with pending retransmissions in their hybrid automatic repeat request (HARQ) manager.

If M users are not fulfilled by SCS, other users can be selected according to the priority of PF scheduling.

The novelty of the proposed scheme is two-fold.

1. The packet specific delay budget for each VoIP packet will be used instead of the static delay budget by considering the “age” of the packet.

2. The age information can be obtained from the optional fields of the IP header assuming IP connection to NodeB in the transmission network. Alternatively without IP connection to the NodeB, the RNC could inform the NodeB, for example, by appending similar information to each VoIP packet via some extended fields of the frame protocol.

With this proposed delay sensitive scheduling the other end HSDPA scheduler can benefit from the knowledge that some packets arrived faster to the buffer than some other packets, and it can thus have longer scheduler delay for those arriving faster. Hence, it is more viable than the conventional method with the static radio delay budget as described in prior art in which the scheduling may not consider the specific end-user experience. It can also de-jitter the buffer delay by scheduling users more efficiently according to the packet specific end-to-end delay budget. Consequently, it can enhance downlink radio transmission efficiency. Otherwise, the method could be widely used in other IP-based system such as Long Term Evolution (LTE) system and Wimax for the downlink scheduling of real-time traffic.

IV. MODELING OF VOIP OVER HSPA

A. VoIP traffic

The duration of each VoIP call is sampled from a negative exponential distribution function with an average call length of 60 seconds. Discontinuous transmission (DTX) is simulated by 50% probability of transmission and silent periods, i.e., on and off periods. The duration of both on and off periods is negative exponentially distributed with an average of 3 seconds. During on periods, a voice encoder generates a payload per VoIP packet of 31 bytes every 20 ms, corresponding to a source rate of approximately 12.2 kbps. Another 9 bytes from the headers of RTP/UDP/IP/PDCP/RLC with header compression will be included in each VoIP packet.

B. VoIP QoS

The outage is introduced to evaluate the performance of VoIP on HSPA. Here, an outage is counted if 5% of the VoIP packets for one user are not received within the delay budget when monitored over a 10 seconds on period.

Each newly arrived VoIP packet in the UE is associated with a discard timer. Whenever a buffered packet has been transmitted, it is moved to the Hybrid-ARQ manager and its

discard timer is deactivated. Hence, whenever a packet has been transmitted, it can be dropped only if it has not been successfully received after the maximum number of allowed transmissions.

Furthermore, the capacity is defined as the number of VoIP users per sector without exceeding 5% cell outage level. The typical criterion for the uplink, noise rise of the sector, is also applied for the benchmarking of uplink results.

C. HSPA system level simulators

The quasi-static system level simulators respectively for HSUPA and HSDPA, where all necessary RRM algorithms as well as their interactions are modeled, are used to investigate the performance of VoIP on HSPA with the proposed method. These tools include a detailed simulation of the users within multiple cells. The fast fading is explicitly modeled for each user according to the ITU Vehicular-A profile. Regarding to the methodology this kind of quasi-static simulator is based on descriptions in [10]. A wrap-around multi-cell layout modeling several layers of interference is utilized in this study.

However, different with the conventional VoIP simulations with independent uplink and downlink simulators, both simulators would be coupled by the delay information, i.e., the uplink packet delay from the uplink simulation result would be used as the input of the downlink simulation in order to calculate the packet specific delay budget for the downlink scheduling.

On the other hand, CQI reporting on 10ms period is modeled, which is supposed to provide good balance between the uplink interference from CQI reporting and efficient downlink scheduling based on CQI information. Therefore, uplink interference from HS-DPCCH carrying 10ms CQI report and CQI information available on 10ms period for downlink scheduling are modeled in HSUPA and HSDPA simulators separately to study the end-to-end performance in a practical way.

V. SIMULATION SCENARIO AND RESULTS ANALYSIS

Main parameters used in the system simulation are summarized in Table II.

TABLE II.
SIMULATION PARAMETERS SETTINGS

Parameter	Value
System configuration	
Inter site distance	2.8 km
Cell configuration	ITU Veh-A, Macrocell
Voice call mean length	60 seconds
UE speed	3 kmph
Outage observation window length	10 seconds
Cell outage threshold	5%
End-to-End air interface delay budget	150 ms
Residual FER	1%
HSUPA specific configuration	
Frame Size	2ms / 10ms
Channels	E-DCH/DPCCH
Number of HARQ channels	8 (2ms TTI) / 4 (10ms TTI)

Max number of L1 transmissions	4 (2ms TTI) / 2 (10ms TTI)
Uplink delay budget	80ms
Scheduling algorithm	Non-scheduled
HSDPA specific configuration	
Frame Size	2ms / 10ms
Channels	HS-DSCH/CPICH/HS-DPCCH
Number of HARQ channels	8 (2ms TTI) / 4 (10ms TTI)
Max number of L1 transmissions	4 (2ms TTI) / 2 (10ms TTI)
Max CDM Users	4
Number of code	10
HSDPA power / Max Power	12.5w / 20w
Downlink delay budget	(150ms - UplinkPacketDelay)

A. Performance of VoIP over HSUPA

In a first step, we investigated the VoIP performance that can be achieved by applying HSUPA features for different TTI lengths. In Table III, the uplink VoIP capacity results with two criteria for both 2ms and 10ms TTI lengths are summarized. It is observed that uplink gating can give 40-50% further gain compared to the case of continuous DPCCH transmission, whatever the TTI length and QoS criteria are.

As can be seen, for uplink gating with average 1.25 HARQ transmissions in 10ms TTI length, the performance gain in terms of the cell capacity is higher compared to the case of average 1.67 HARQ transmissions. That is because fewer transmissions may spare more time for DPCCH gating, which can further lead to less uplink interference. Consequently, it provides more potential for significant performance improvement with DPCCH gating.

On the other hand, fewer transmissions would require higher power leading to the power shortage of the UE in the cell edge. This can result in capacity loss as shown in the case of 10ms TTI with continuous DPCCH transmission, where the capacity is 65 UEs per sector for 1.25 transmissions whereas 80UEs for 1.67 transmissions in terms of the cell outage criteria.

Thus, FER target setting after 1st transmission used in outer loop power control can be adjusted to optimize the average HARQ transmission in the case of DPCCH gating for exploring the potential capacity gain.

In Fig. 2, some uplink packet delay information is collected for both cases of DPCCH gating and no gating (i.e., continuous DPCCH transmission). It is noted that uplink packet transmission can always be done within the uplink delay budget (i.e., 80ms). The most reason is that synchronous HARQ transmission in HSUPA would always force the packet

transmission in the period, which can ensure the packet delay not to exceed some threshold in accordance with the maximum allowed HARQ transmission number. It is proved in the case of 2ms TTI that there are 4 slopes corresponding to 4 HARQ transmissions. Therefore, it also implies that the packet delay is not the bottleneck for the uplink capacity.

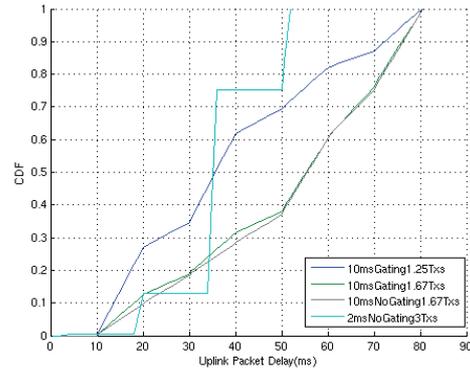


Fig. 2. Uplink packet delay distribution.

Besides, more HARQ transmissions, more delay VoIP packet would suffer. However, there is not much difference on the packet delay between gating DPCCH and continuous DPCCH with the same average transmission times.

B. Performance of VoIP over HSDPA with the packet age aided scheduling

Based on aforementioned uplink packet delay information, the downlink delay budget for each packet can be calculated. Then, this packet specific delay budget is used for the downlink scheduling. Fig. 3 depicts the cell outage level as a function of downlink user number per sector. It is noted that the static delay budget used in early studies, 80ms and 150ms, can just provide the lower bound and upper bound of the downlink capacity. And 150ms delay budget for only downlink air interface transmission is even not reasonable when the call is between mobiles. So the noticeable gain of the scheme with the packet specific delay budget can be observed compared to the case with the 80ms fixed delay budget.

Shorter delay in the uplink can offer more time for the downlink scheduling which can exploit the relative better channel condition for the UE transmission. This will make the downlink transmission more efficiently, and then offer more

TABLE III.
SUMMARY OF THE UPLINK VOIP CAPACITY RESULTS

TTI	Average number of transmissions	Capacity criteria: 5% FER over 10 s			Capacity criteria: Noise rise 6 dB		
		Continuous DPCCH	Gated DPCCH	Gating Gain	Continuous DPCCH	Gated DPCCH	Gating Gain
2 ms	~3	82 users	123 users	50%	75 users	106 users	41%
10 ms*	1.25	65 users	115 users	77%	61 users	93 users	52%
	1.67	80 users	120 users	50%	73 users	103 users	41%

* 10 ms TTI results using bundling of 2 VoIP frames in a single TTI

UE leading an increase in the downlink capacity of up to 16% in the case of 2ms TTI length compared to the static delay budget. Consequently, the simulation results are also quite in line with uplink packet delay results, i.e., the downlink capacity with the uplink packet transmission in 2ms TTI is higher than the case with 10ms TTI due to the shorter delay. It is also verified that DPCCCH gating can improve the uplink performance significantly without sacrificing any downlink performance assuming the same transmission times for DPCCCH gating and continuous DPCCCH. Uplink HARQ transmission times would also affect the downlink performance because of its effect on the packet delay with this proposed scheduling.

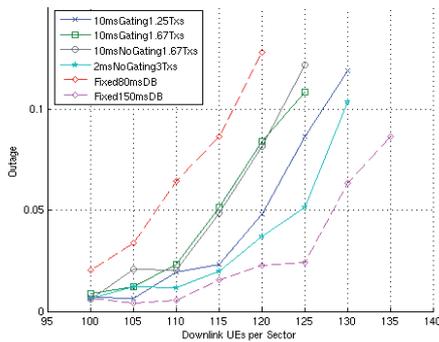


Fig. 3. Downlink VoIP capacity with different uplink transmission.

Hence, the packet specific delay budget can provide significant capacity gain in the downlink compared to the static delay budget. It can ensure the end-to-end performance but not only uplink or downlink performance. Assuming the uplink is from a land-line user, it is expected to provide more gain as shown in the case of 150ms, where the uplink transmission delay can be ignored.

C. Effect of uplink packet delay variation

In order to further investigate the uplink delay effect on the downlink capacity with the packet specific delay budget, the downlink delay budget is calculated based on a uplink packet delay of each VoIP call, which is sampled from a normal distribution function as shown in Fig. 4, where the average uplink packet delay is 80 ms and 90 ms with the standard deviation of 0, 10 and 20.

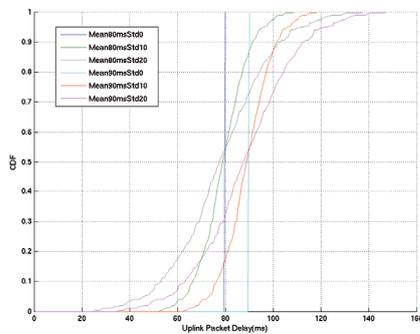


Fig. 4. Generated uplink packet delay distribution.

The downlink capacity results are shown in Fig. 5. It can be noted that larger standard deviation of uplink packet delay will lead to worse downlink performance at the same mean values. In some sense, the delay variation implies how burst the uplink VoIP packets arrive into the downlink scheduling buffer. The smaller variation means less bursty. This kind of bursty characteristic is mainly because the uplink air interface (re)transmission may result in uncertain packet delay and even disorder VoIP packets. However, the packet age aided scheduling can somehow relieve the pain from the bursty by prioritizing VoIP packets according to their ages.

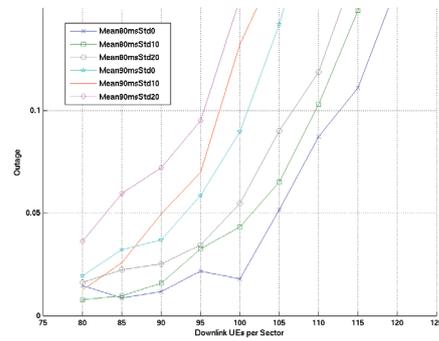


Fig. 5. Downlink VoIP capacity with generated uplink delay distribution

VI. CONCLUSION

A packet age aided scheduling in HSDPA is proposed to improve the downlink VoIP capacity by utilizing the packet specific delay budget instead of the static one. The end-to-end system performance can be significantly improved by this way, especially for shorter uplink transmission delay. The results obtained from extensive system-level uplink and downlink simulations have shown that using the proposed scheduling can lead to an increase in the downlink cell capacity of up to 16%. By applying the proposed method, the downlink performance can be improved considerably while the uplink performance remains the same, which can ensure the good end-to-end performance.

References

- [1] 3GPP TS 25.309, "FDD Enhanced Uplink; Overall description; Stage 2"
- [2] 3GPP TR 25.896, "Feasibility Study for Enhanced Uplink (FDD)"
- [3] 3GPP TS 25.308, "HSDPA; Overall description; Stage2"
- [4] ITU-T Recommendation G.114, "One way transmission time"
- [5] H. Holma, M. Kuusela, E. Malkamäki, K. Ranta-aho, Tao Chen, "VoIP over HSPA with 3GPP Release 7" PIMRC, Helsinki, September 2006.
- [6] W. Bang, K. I. Pedersen, T. E. Kolding, and P. E. Mogensen (2005), "Performance of VoIP on HSDPA", IEEE Proc. Vehicular Technology Conference, Stockholm, June.
- [7] Petteri Lunden, Markku Kuusela, "Enhancing Performance of VoIP over HSDPA". IEEE Proc. Vehicular Technology Conference, Spring, 2006.
- [8] W. Richard Stevens, "TCP/IP Illustrated, Volume 1: The Protocols", Addison-Wesley, 1994, ISBN 0-201-63346-9
- [9] R. Cuny, A. Lakaniemi, "VoIP in 3G Networks: An end-to-end quality of service analysis," IEEE, VTC spring, vol. 2, pp. 930-934, April 2003.
- [10] 3GPP2 Technical Specification Group C, 3GPP2/TSG-C.1002: "1X EV-DV Evaluation Methodology"