Performance Comparison of Control-less Scheduling Policies for VoIP in LTE UL

Haiming Wang
Nokia Device R&D/Wireless System Research
Nokia (China) Investment Corporation Limited, 100176
Beijing, China
haiming.wang@nokia.com

Dajie Jiang
School of Telecommunication Engineering
Beijing University of Posts and Telecommunications,100876
Beijing, China
jiangdajie@gmail.com

Abstract— **Performance comparison of control-less scheduling policies for VoIP in LTE UL is conducted in this work. In LTE system, the main challenges of effectively supporting VoIP service are the control channel restriction from the large number of simultaneous users and the frequent arrival of small packets. Semi-persistent scheduling and group scheduling are the main methods to solve above problem and achieve QoS requirements. By simulation, we prove that semi-persistent scheduling outperforms group scheduling in the case of no group interaction for group scheduling.**

Keywords- VoIP, LTE, Semi-persistent Scheduling, Group Scheduling

I. Introduction

Voice over Internet Protocol (VoIP) in wireless cellular networks has drawn more and more interest recently because of its integrated IP infrastructure and larger user capacity than traditional voice of circuit-switch mode [1-5]. The Evolved Universal Terrestrial Radio Access Networks (E-UTRAN), targeting for a Long Term Evolution (LTE) of 3G UMTS, is currently being standardized by the 3rd Generation Partnership Project (3GPP). As packet-optimized radio access technology, the air interface of LTE system is using OFDMA (orthogonal frequency division multiple access) and SC-FDMA (single carrier-FDMA) in downlink (DL) and uplink (UL) respectively [6-7]. Then VoIP is not just an attractive feature in LTE system for effective speech conveying but is also the only way to provide voice service in E-UTRA network. However, the low user data rate and tight delay requirements imply that to enable high VoIP capacity several users need to be allocated to the same transmission time interval (TTI) by using different frequency resources. Signaling the allocations to each user within one TTI would increase the control channel overhead. Unfortunately E-UTRAN is designed to avoid large control channel overhead and includes restrictions to the number of control channels per TTI. These restrictions limit the number of allocation signaling per TTI. It was noticed in 3GPP that the control channel limits the VoIP capacity heavily [8-9]. Thus solutions to reduce signaling from new scheduling mechanism point of view were developed in uplink, which include semi-persistent scheduling [9-10] and group scheduling [11]. Semi-persistent scheduling has been agreed in 3GPP already.

The main structure of this paper is organized as follow: Section II expresses the main VoIP problems related to LTE system and Section III explains principle of semi-persistent scheduling and group scheduling for VoIP traffic. Numerical results are given and analyzed in Section IV before conclusions are drawn in Section V.

II. Main VoIP Problems In LTE System

The VoIP traffic is mainly characterized by small packets in regular intervals with tight delay constraints.

In order to fulfill the requirements of high data rate and low transmission latency, the simplified radio architecture for LTE was adopted where all user plane functionalities for the radio access were grouped in one network node only: the evolved Node B (eNB). eNB hosts all the Radio Resource Management (RRM) related work besides other radio functions. With all radio protocol layers and scheduling located in the eNB, efficient allocation of resources to UEs with minimum latency and protocol overhead becomes possible. This is especially important for real time services like VoIP, for which, high capacity requirements were set: at least 200 users per cell should be supported for spectrum allocations up to 5 MHz, and at least 400 users for higher spectrum allocation. Without signaling restriction, to achieve high VoIP capacity, full dynamic scheduling is actually the best choice because it can make full use the resources and enable good link adaptation by dynamically scheduling each packet (both initial transmissions and retransmissions). However, the major drawback with dynamic scheduling applied to VoIP is its large amount of signaling that is required to allocate resources for every transmission and the possible retransmission of each speech frame. Thus, it is valuable to investigate some control-less scheduling schemes to further enhance the VoIP performance when control channel is limited in full load case. As a result, semi-persistent scheduling and group scheduling are invented.

III. Control-less Scheduling Policy For VoIP

To support high VoIP capacity with reasonable control signaling, semi-persistent and group scheduling are proposed and discussed in 3GPP meetings.

A. Semi-persistent Scheduling

The principle of semi-persistent scheduling includes two parts: persistent scheduling for initial transmissions and dynamic scheduling for retransmissions. In this case UE is preconfigured (using RRC or L1/L2 signaling) with a limited set of time/frequency resources where initial transmissions can be sent. All the retransmissions are scheduled dynamically using the L1/L2 control channel. Since the retransmissions are always scheduled, the retransmissions can be freely allocated on any free resources, e.g., on those remaining unused by silent users. From system and signaling overhead point of view, this method is promising. The main points of UL semi-persistent scheduling are as follows, depicted in Figure 1:

- For voice packets, persistent allocation is used for new transmissions (a sequence of chunks located at every 20ms are allocated for one user 'persistently' until this user steps into DTX periods) and dynamic allocation for retransmissions.
- All SID packets (initial transmission and retransmission) are dynamically scheduled.
- Scheduling priority: [*new transmissions of voice packets > retransmissions > SID packets*].
- Talk-spurt-based: eNB reallocates or releases resources when UE steps from active periods into DTX periods.

Semi-persistent scheduling can decrease the signaling burden by persistently allocating resources to the initial transmissions. If VAF (Voice Activity Factor, v) is 50%, then the average number of control channels ($No.CCHs$) needed per TTI can be estimated by the below equation.

$$No.CCHs = \lambda[nv / I_1 + n(1-v) / I_2] \quad (1)$$

n is the number of total VoIP users and λ means the average transmission numbers (here we assume λ is1.2), I_1 and I_2 denote the inter-arrival time of voice packets (20ms) and SID packets (160ms), respectively. In high load situation, the signaling saving reaches to 75% compared to the full dynamic scheduling. The detail signaling overhead analysis is shown in reference [9].

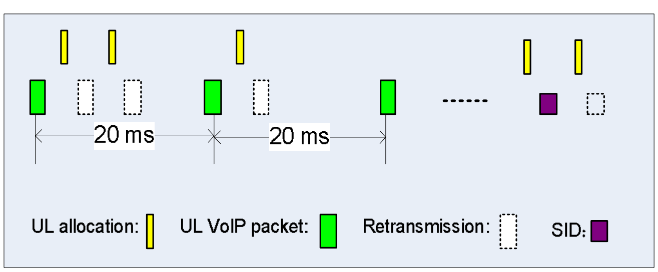

Figure1 Semi-persistent scheduling
(SID is scheduled dynamically)

B. Group Scheduling

Group Scheduling is firstly proposed by Motorola in [11]. The basic idea is to divide users into several user groups, and then a group based signaling is send to a class of users and does scheduling instead of sending separate signaling to each user. The purpose of group scheduling is to decease the signaling overhead from resource allocation signaling. However at meanwhile complex group management signaling may be needed to do dynamic resource sharing among groups to get a better performance. The main points of UL group scheduling are depicted in Figure 2 (take 5 groups which are distinguished by different colurs as an example):

- UEs are assigned randomly to several different VoIP groups.
- The number of groups is several times of the number of HARQ processes. i.e. 6, 12 or 18 groups (given Sync. HARQ for UL and 6 HARQ processes), that's to say, each VoIP group can only be activated in one HARQ process.
- Delay dependant scheduling with each group.
  - ✓ Retransmissions always get priority over first transmissions.
- No inter-group management within one call and no related signalling overhead.
- In the TTIs when some group is activated, one scheduling signalling will be sent out by eNB in advance and the users in this group can hear this signalling to find their related resource allocations.
- Compared to a normal signalling where only one user is involved, a group signalling relates to several users. Due to the restriction of signalling size, the group signalling can only include parts of information, such as allocated resources to each user. The other information like MCS (modulation and coding scheme) is not included in group scheduling and it must be preconfigured or blindly judged by UE.

Furthermore, within one group, statistical multiplexing of users through voice activity factor and HARQ retransmission activity is allowed (see figure 2)

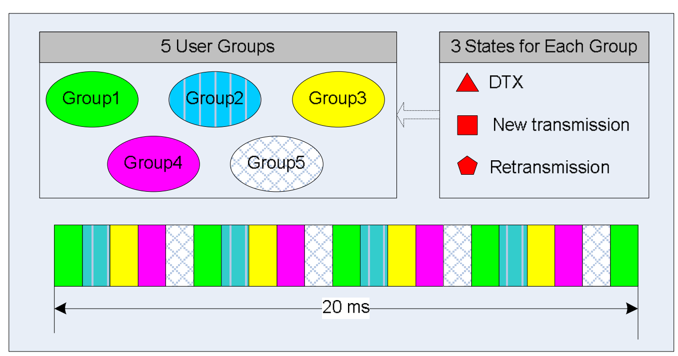

Figure 2 Group scheduling

IV. Simulation and Performance Analysis

In this section, we introduce system level simulation principles in investigating the performance for VoIP on LTE UL with the above two scheduling methods. To make a fair comparison between semi-persistent scheduling and group scheduling, most parameters/terms are common to both schemes. For example, both schemes are talk-spurt based, which means eNodeB reallocates or releases resources when UE steps from active periods into DTX periods. In both methods, slow link adaptation is enabled. It means different number of RUs (resource units, 180 kHz per RU) is allocated to each user in terms of its path-loss information. The band allocation is done from one direction to another (i.e., from RU 0 to RU 20) to avoid possible frequency fragmentation. The SNR based PC parameters are also the same. No control channel error is modeled.

A. System Model

We use a quasi-static system level simulator based on the 3GPP LTE system simulation scenario [7]. Our simulation platform models interference from 57 sectors (19 3-sector cells) in a wrap-around model and users are dropped uniformly in entire sector. Inter-site distance is 0.5 km and the shadowing standard deviation is 8 dB. 2 long blocks per TTI are reserved for pilots and 4 out of 25 RUs for control channels, so there are 21 RUs left for data. Synchronous adaptive HARQ is used with 6 HARQ processes and chase combining is assumed. The maximum number of retransmissions is determined by the VoIP delay budget (DB) and number of HARQ processes. Given 50ms DB and 6 HARQ processes, then the maximum retransmission number is 8. An SNR target based power control is used to make up for the overall path loss and to control the inter-cell interference. Link to system interface used is AVI assuming practical FDE receiver and realistic channel estimation. The other parameters are listed in Table 1.

TABLE I. Simulation Parameters

Parameter	Value
Propagation Model (dB)	128.1 + 37.6 LogR, R in kilometers
Sector antenna pattern	70 deg (-3dB) with 20dB front-to-back ratio
Shadowing correlation between cells / sectors	0.5 / 1.0
Penetration loss	20dB
Channel model	6-ray Typical Urban (TU)
UE speed	3km/h
UE maximum power	24dBm
Carrier frequency	2000MHz
Bandwidth	5MHz
TTI length	1ms
Channel update	per slot (0.5 ms)
eNB receiver	2 antennas with MRC (maximum ratio combining)
UE transmitter antenna	1 antennas
eNB/UE antenna gain	14dBi / 0dBi
Thermal noise density	-174dBm/Hz
Frequency re-use	1

B. Traffic Model

Voice traffic is modeled based on a 2-state voice activity model [12]. 2-state voice activity model is shown in figure3:

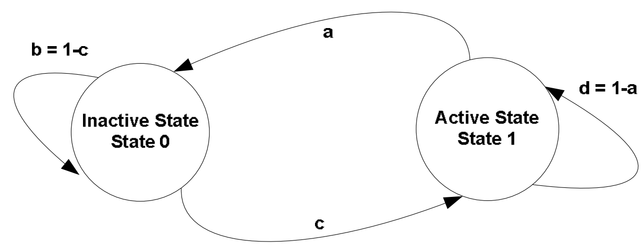

Figure 3 2-state voice activity models

In the model, the probability of transitioning from state 1 (the active state) to state 0 (the inactive or silent state) while in state 1 is equal to a, while the probability of transitioning from state 0 to state 1 while in state 0 is c. The model is assumed updated at the speech encoder frame rate $R = 1/T$, where T is the encoder frame duration (typically, 20ms). The steady-state equilibrium of the model requires that

$$P_0 = \frac{a}{a+c}, \qquad P_1 = \frac{c}{a+c} \tag{2}$$

Where P_0 and P_1 are respectively the probability of being in state 0 and state 1.

The Voice Activity Factor (VAF) λ is given by

$$\lambda = P_1 = \frac{c}{a+c} \tag{3}$$

50% VAF is assumed to model half probability of active and inactive periods ($a = c = 0.01$). The duration of both active and inactive periods is negative exponentially distributed with an average of 2 seconds. Including the compressed RTP/UDP/IP headers and other overhead, there are totally 40 bytes per voice packet for 12.2kbps AMR codec and 15 bytes for an SID packet. Adaptive Transmission Bandwidth (ATB) is available by allocating 1 or 2 RUs within one TTI to transmit one packet according to users' path loss (no frequency selective scheduling is used for UL VoIP considering the huge number of sounding pilots it need).

Two main QoS requirements of VoIP are the packet delay and packet loss rate. The packet delay in UL we modeled consists of the queuing delay in the scheduler at eNB and the packet transmission delay which mainly depends on number of retransmission. Note that the packet delay here doesn't include delay of RR (Resource Request) from UE to eNB and resource allocation from eNB to UE. According to [12], the radio interface delay budget is 50 ms. A VoIP user is in outage (unsatisfied) if more than 2% of its packets in a 60 seconds call cannot be correctly received within 50ms DB. The system capacity is defined as the maximum load in which more than 95% of the users are satisfied.

C. Simulation Results

The VoIP capacity for 12.2 kbps AMR codec on 5 MHz carrier bandwidth is shown in Table 2. Herein, parameter p means "proportion of users with 2RU allocation". It can be seen that with number of groups increasing (6 groups to 12 groups), VoIP capacity becomes lower. This is because that more groups lead to less statistical multiplexing gains among groups. In fact, dynamic scheduling is an extreme case of group scheduling where all users belong to one group. Dynamic scheduling without control channel restriction should get the highest capacity because it can get a full statistical multiplexing among users. On the other hand, if each group includes only one user, then there is no statistical multiplexing gain among users and this is the worst case. In general, more users in one group, more statistical multiplexing gain can be got which will result into higher capacity. This, however, would require larger control channel and thus lead to more overhead. On the other hand, larger control channel is not possible if the group scheduling control is required to have the same size as normal UL allocation. Semi-persistent scheduling outperforms group scheduling. Figure 4 explain this problem clearly:

- The red curve denotes total number of active users in all groups divided by number of groups (12). Curves in other colors mean number of active users in each group.
- Imbalance of active users among groups in different time leads to ineffective resource utilization—lack of statistical multiplexing among groups!
- So group management will be necessary. However, complexity and extra signaling by group management will be introduced as well.

TABLE II COMPARISON OF VoIP CAPACITY

	p=100%	p=90%	p=70%
Semi-persistent	222	234	244
Group scheduling (total 6 groups)	208	224	235
Group scheduling (total 12 groups)	185	193	201

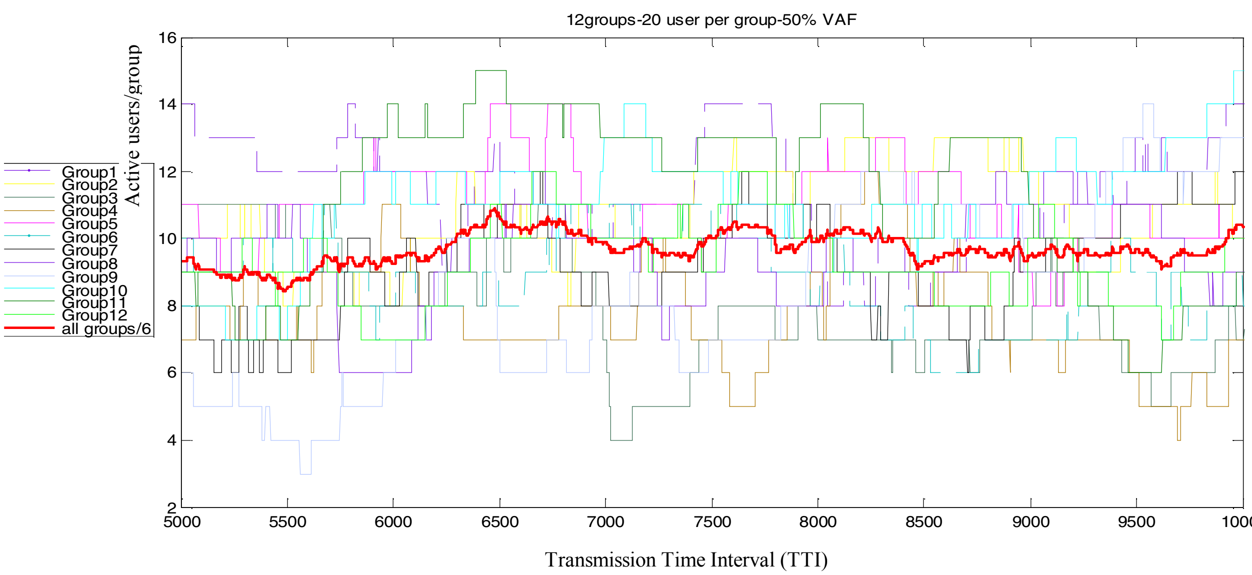

Figure 4 Analysis to group scheduling in 5MHz (12 groups)

TABLE III AVERAGE IOT (DECIBEL), 2RU ALLOCATION, 12.2 KBPS AMR CODEC, 5 MHZ

No. of user per sector:	180	200	220	240
IoT of Semi-persistent:	-	7.79	7.96	8.26
IoT of Group scheduling (total 6 groups):	7.64	7.89	7.95	-
IoT of Group scheduling (total 12 groups):	7.5	7.67	7.94	-

The average Interference over Thermal (IoT) value can denote the stability of wireless system. The higher IoT is, the much unstable the system is. The average IoT Value with 2RU allocation is shown in table3. It can be seen that IoT is increased with load increasing and the scheduling methods have little impact on IoT value.

V. CONCLUSIONS

We have presented the principle and performance of semi-persistent scheduling and group scheduling in LTE UL system. Simulation results show the semi-persistent has the larger capacity than group scheduling (without group interaction). The capacity difference is due to lack of statistic multiplexing between groups in group scheduling. Without group management, resources in different groups can not be exchanged and the overall resources can not be effectively utilized. If group management is enabled, the extra complexity and signaling load will be introduced.

ACKNOWLEDGMENT

The authors would like to acknowledge a number of colleagues from Nokia who have been involved in the development of LTE VoIP simulations and the enhanced LTE features in 3GPP.

REFERENCES

[1] Harri Holma, Markku kuusela, Esa Malkamaki, Karri Ranta-aho, Chen Tao, "VoIP over HSPA with 3GPP Release 7", *IEEE PIMRC'06*, September 2006,pp.1-5

[2] Tao Chen, Markku Kuusela, Esa Malkamaki, "Uplink Capacity of VoIP on HSUPA", *IEEE VTC 2006 spring*, Vol.1, May 2006. pp. 451-455

[3] Petteri Lunden, Markku Kuusela, "Enhancing performance of VoIP over HSDPA", *IEEE VTC 2007 spring*, April 2007, pp. 825-829

[4] W. Bang, K.I.pederen, T.E. Kolding, P.E. Mogensen, "Performance of VoIP on HSDPA", *IEEE VTC 2005 spring*, Vol.4, May 2005, pp. 2335-2339

[5] P.J.Black et al., "cdma2000 1xEV-DO Revision A: A Physical layer and MAC Layer overview", *IEEE communication magazine*, Feb.2006, pp. 75-87

[6] 3GPP TR 25.913 V7.3.0, March 2006

[7] 3GPP TR 25.814 V7.1.0, Sep 2009

[8] 3GPP TSG-RAN WG2 Meeting #57bis, R2-071227, "Number of Control Symbols", St. Julian's, Malta, March 26-30, 2007

[9] Dajie Jiang, Haiming Wang, Esa Malkamaki and Esa Tuomaala, "Principle and Performance of Semi-persistent Scheduling for VoIP in LTE System". *IEEE Wi'Com'07*, September 2007. pp. 2861-2864

[10] Haiming Wang, Dajie Jiang and Esa Tuomaala, "UL Capacity of VoIP on LTE UL System", *IEEE APCC'07*, October 2007, pp. 397-400

[11] 3GPP TSG RAN WG1#44, R1-060398, "VoIP Group Scheduling", Denver, USA. Feb13-17, 2006

[12] 3GPP TSG-RAN WG1 Meeting #48, R1-070674, "LTE physical layer framework for performance verification", St. Louis, USA, Feb 12-16, 2007