Analyzing Traffic Characteristics and Performance for LTE Uplink Resource Allocation

Fang-Chang Kuo(✉)

National Ilan University, Yilan City, Taiwan
kfc@niu.edu.tw

Abstract. Many resource allocation schemes have been proposed based on different criteria such as system throughput, fairness, transmission power, user priority and others. According to the 3GPP specifications, however, as long as a GBR bearer is admitted, the eNB has to ensure the Guaranteed Bit Rate (GBR) of the bearer before it is disconnected or dropped. However, the 3GPP hasn't defined how to measure GBR. In previous papers, we present how to use Exponentially Weighted Moving Average (EWMA) algorithm to define the measurement of data rates, so as to facilitate resource allocation works. In this paper, we discuss the characteristics of traffic patterns under the constraint of EWMA, as well as the impact on the performance of eNB that serving such traffics. According the simulation results, we also suggest that the more bursty traffic patterns should be charged with higher rates.

Keywords: Uplink scheduling · Resource allocation · EWMA

1 Introduction

Dynamic Resource Allocation (DRA) in an LTE system concerns the process for an eNB (Evolved Node B) to allocate radio resource blocks (RBs) to user equipment (UE), such that the UE can transmit data to eNB. Many DRA schemes for have been proposed. Initially, some DRA schemes, such as Recursive Maximum Expansion (RME), emphasize on maximizing the system throughput of an eNB [1, 2]. Then, some researchers raised that fairness is also a very important criterion, and proposed some schemes, such as frequency domain Round Robin (RR) [3] and Proportional Fair (PF) [4, 5]. However, one of the important tasks of eNB is to admit or reject the establishment requests for new radio bearers, at the same time it also needs to ensure high radio resource utilization and to ensure proper QoS for the in-progress sessions [6]. As a result, some schemes emphasizing bearer QoS were proposed [7–9].

The throughput of a Guaranteed Bit Rate (GBR) bearer is expressed in terms of the GBR, which is declared when the UE requests to establish the bearer. GBR is the bit rate that can be expected when there is data pending for transmission [10]. However, 3GPP hasn't clearly defined how to measure GBR. As a result, it is hard to verify whether the QoS, in term of GBR, is guaranteed. Besides, it is quite general that a UE generates traffic with silences of different durations. As a result, it is also hard for a UE to declare the data rate and also hard for the eNB to allocate RBs for the UE.

Y.-B. Lin et al. (Eds.): IoTaaS 2017, LNICST 246, pp. 86–93, 2018.
https://doi.org/10.1007/978-3-030-00410-1_11

In order to facilitate DRA work, we employed Exponentially Weighted Moving Average (EWMA) [11] to define the GBR of a bearer, and designed a DRA scheme, named AAG (Allocate As Granted). We showed that AAG has outstanding performance in terms of UE satisfaction, packet delay, and system throughput, then we also extended it to new versions AAG-2 and AAG-LCG [12, 13]. The new versions aim at guaranteeing the QoS of all GBR bearers, while efficiently distributing RBs to non-GBR bearers, so as to improve resource utilization.

In the next section we describe our previous works about how EWMA is employed by AAG scheduling scheme to facilitate resource allocation work. The characteristics of EWMA are analyzed in Sect. 3, and the corresponding simulation is depicted in Sect. 4. We then make conclusion and describe future works in Sect. 5.

2 Previous Works

In order to analysis the characteristics of EWMA, here we first briefly describe its principle and how it is employed by AAG to obtain outstanding performance.

Let $R_{m,GBR}^{grant}$ be the GBR that is granted by the eNB for the m-th UE (UE_m). We may convert $R_{m,GBR}^{grant}$ to another form $B_{m,GBR}^{grant}$ which stands for "the number of bits that is permitted to be transmitted per TTI (Transmission Time Interval)." For UE_m, let $B_m(n)$ denotes the number of bits transmitted at the n-th TTI. Define $\bar{B}_m(n)$ as the average number of bits per TTI that has transmitted after the n-th TTI. Then, based on the definition of EWMA, we obtain

$$\bar{B}_m(n) = (1-\alpha)\bar{B}_m(n-1) + \alpha B_m(n), \quad (1)$$

where $\alpha \in (0,1)$ is the weighting factor. The $\bar{B}_m(n)$ should not be less than $B_{m,GBR}^{grant}$ if eNB wishes to provide UE_m with average data rate that is no less than $R_{m,GBR}^{grant}$. As a result, the eNB should plan to allocate some RBs to transmit the following number of bits

$$B_m^{plan}(n) = \min\left\{ \max\left[\frac{B_{m,GBR}^{grant} - (1-\alpha)\bar{B}_m(n-1)}{\alpha}, 0 \right], L_m(n-1) \right\}, \quad (2)$$

where the new term $L_m(n-1)$ is the total queue length of the UE_m. This term is obtained through the Buffer Status Report (BSR) sent by UE_m to indicate how many bits are waiting for transmission. With this term, eNB can prevent wasting RBs if there is not so many bits pending in the buffer.

In order to ensure the throughput of every UE, AAG allocates RBs to UEs according to the descending order of the priority values defined as follows

$$P_m(n) = \frac{B_{m,GBR}^{grant} - \bar{B}_m(n-1)}{B_{m,GBR}^{grant}}. \quad (3)$$

This term is also an indication of how a UE is satisfied with the current average data rate as compared with the declared one.

3 Characteristic Analysis of EWMA

As mentioned above, there could be lots of silences for the duration of a session. Based on EWMA, the number of bits that a UE authorized to transmit GBR traffic has been expressed in (2) and the corresponding priority metric is expressed in (3). Let's count out the term $L_m(n-1)$ representing the pending number of bits, and concentrate on discussing the characteristics of EWMA. As a result, the longer duration the UE keeps silent, the more bits it can transmit with higher priority metric. To the extreme condition, if the UE_m keeps silent for a very long period (duration), the term $\bar{B}_m(n-1)$ would approach zero. As a result, based on (2), the number of bits that the UE authorized to transmit as a burst would approach

$$\frac{B^{grant}_{m,GBR}}{\alpha}. \tag{4}$$

We define this value as the Maximum Burst Size (MBS) corresponding to a bearer specified with the set of parameters $(B^{grant}_{m,GBR}, \alpha)$. For example, when α is 0.01 the MBS is 100 times of $B^{grant}_{m,GBR}$. By the way, the corresponding priority metric for sending such a burst would approach to one, which is the maximum value. However, if the silent duration is not long enough, the UE is only allowed to transmit a smaller burst size, which can be derived as follows.

Starting from the n-th TTI, assume UE_m keeps silent for τ TTIs, then

$$\bar{B}_m(n+\tau) = (1-a)^{\tau} \cdot \bar{B}_m(n) \tag{5}$$

Based on (2), we have

$$B^{plan}_{in}(n+\tau) = \frac{B^{grant}_{m,GBR} - (1-\alpha)^{\tau} \cdot \bar{B}_m(n)}{\alpha}. \tag{6}$$

If $\bar{B}_m(n) = B^{grant}_{m,GBR}$ before it get into silent mode, we may express the burst size that the UE can transmit after a silence of τ TTIs as following

$$\mathrm{MBS}[1-(1-\alpha)^{\tau}], \tag{7}$$

where the maximum burst size, MBS, is defined in (4).

Based on (7), we illustrate burst size as a function of silence duration and parameter α in Fig. 1, where the burst size is expressed as a value that is normalized by the corresponding MBS expressed in (4). Take a UE with α being 0.01 as an example, if it keeps silent for 0.1 or 0.01 s, it has the credit to respectively transmit a burst which is 0.63 or 0.09 times of the MBS. As we can see, when a UE goes into silence, it begins to accumulate credit for transmitting more bits. At the same time, however, it begins to

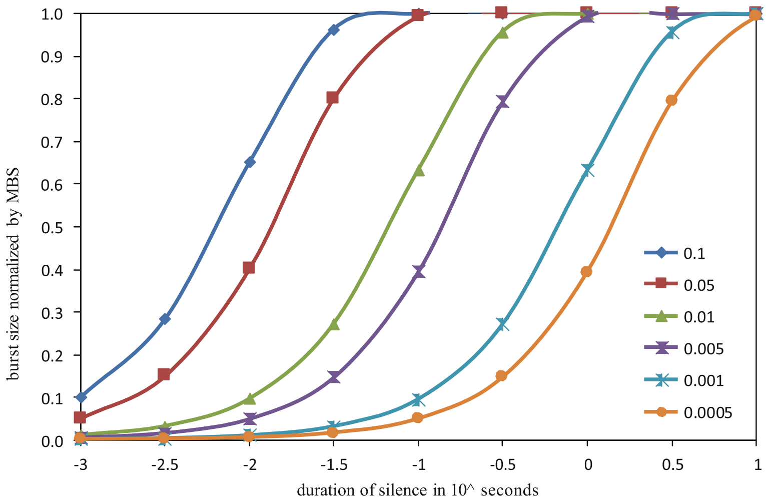

Fig. 1. Burst sizes as a function of silence duration and parameter α

lose some of its transmission credit. Let's consider another scenario where the UE_m keeps transmitting for τ TTIs at a data rate $B^{grant}_{m,GBR}$, rather than keeps silent. Then the number of bits that it transmits is $B^{grant}_{m,GBR} \cdot \tau$, which is larger than the legal burst size expressed in (7).

In the position of a UE which has nothing to transmit for a while, it would like to accumulate the transmission credit and then transmit a bigger burst at a later moment. As we can see, the EWMA can help preserve the transmission credit to a certain degree during the silence period. However, from the aspect of an eNB, if lots of UEs transmit bursts at a moment, the eNB is prone to be overloaded at this moment, and some of the packets will suffer from longer delay and even overdue. That means for some of the UEs, their GBRs aren't guaranteed within limited delay. This is what an eNB needs to avoid. Fortunately, by employing EWMA, the legal burst size is bounded by (7). If a UE generates a bigger burst, only the QoS of the legal part should be ensured. The eNB is irresponsible to immediately allocate RBs for the illegal part because it also needs to ensure the QoS of other UEs. As a result, we may regard EWMA as a suitable algorithm to compromise the expectation of both UEs and eNBs.

4 Performance Evaluation

In order to observe the impact of traffic parameters on the performance of eNB, we simulate the environment where some UEs declare the same GBR and generate bursty traffic patterns, while the eNB dynamically allocates RBs based on the AAG scheme presented in [12].

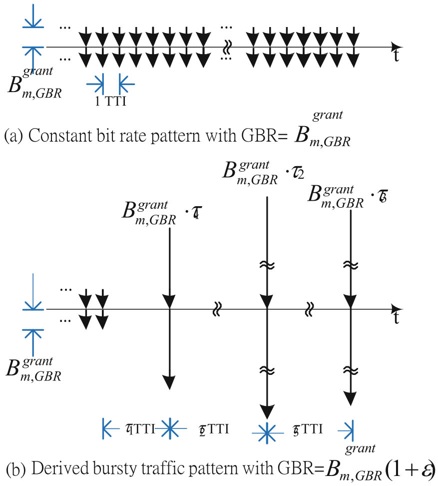

Fig. 2. Derived bursty traffic pattern that complied with a specific traffic parameters.

4.1 Method for Generating User Traffic Patterns

We need to emulate some bursty traffic patterns generated by UEs. There could be numerous traffic patterns complying to a specific parameter set ($B^{grant}_{m,GBR}$, α_m). A constant bit rate (CBR) pattern shown in Fig. 2(a) is a special one. We may use such a CBR pattern to derive bursty patterns as shown in Fig. 2(b). The characteristic of such a bursty pattern is that after a silence of deration *i* TTIs, a burst with proportional number of bits $B^{grant}_{m,GBR}\cdot\tau_i$ is generated. Such a burst is regarded as a packet in this paper. A longer silence duration is followed by a larger burst size. However, as we have explained in Sect. 3, although keeping silent can accumulate transmission credit, some of that is lost at the same time. As a result, it is not allowed for a UE to send out so many bits at a TTI, unless the granted GBR is enlarged. Thus, we enlarge the granted GBR by an extra ratio ε. That is to say, we enlarge the granted GBR to $B^{grant}_{m,GBR}(1+\varepsilon)$. Then based on (7) we can calculate τ_{max} (ε, α), the maximum value of τ, by the following equation

$$\alpha\tau \leq [1-(1-\alpha)^{\tau}](1+\varepsilon). \tag{8}$$

Given α and ε, then τ_{max} (ε, α) is the maximum silence duration that the following burst, which is a packet with $B^{grant}_{m,GBR}\cdot\tau_i$ bits, is legal (authorized) to be sent out immediately. For example, if ε = 0.15, the τ_{max} equals 29 and 15 for α being 0.01 and 0.02, respectively. For the sake of diversity when emulating the generated traffic patterns, the chosen silence durations τ_i are uniformly distributed between 80% and 100% of τ_{max}.

4.2 Simulation Environment and Results

The simulation parameters employed are listed in Table 1. Here we consider a simplified environment where each UE generates only one traffic pattern. And we choose α being 0.01 and 0.02 for the traffic patterns used in scenario 1 and 2, respectively. However, the long term averages of these patterns are the same because they have the same $(1+\varepsilon)B_{m,GBR}^{grant}$. In average, the patterns with smaller α would result in longer silence durations and larger burst sizes. We regard such patterns as more bursty in this paper.

Table 1. Simulation parameters

Parameters	Values (scenario 1/2)
$B_{m,GBR}^{grant}$	1.5 Mbps
Extra ratio ε for generating input patterns	0.15
EWMA weight α for UEs and the eNB	0.01/0.02
τ_{max}	29/15
Extra ratio ε for allocating RBs at eNB	0.1, 0.15
No. of UEs (assume only one bearer per UE)	10–40

As for the eNB, different values of extra ratio ε are employed to observe the impact of load on packet delay. The packet delay of a packet is defined as the duration measured from its arrival at the UE buffer until the transmission is completed.

Figure 3 shows the 90 percentile packet delay corresponding to the number of UEs served by the eNB before being overloaded. The delay increases with the increase of load because of the increased possibility that more burst packets are generated simultaneously. The delay decreases with the increase of the extra ratio at the eNB because higher mean rate is allowed.

Note that in order to have the same 90 percentile delay, less number of UEs can be accommodated for smaller value of smaller α, which means more bursty traffic pattern. For example, with ε being 0.15, the delay for 15 UEs with α being 0.01 is almost the same as that for 30 UEs with ε being 0.02. That means the more bursty patterns should be declared with smaller values of α, while they would induce higher impact on the eNB. The result is that less UEs can be accommodated if the same delay is expected. From the billing point of view, the more bursty traffic patterns should be charged with higher rates if the same QoS is expected.

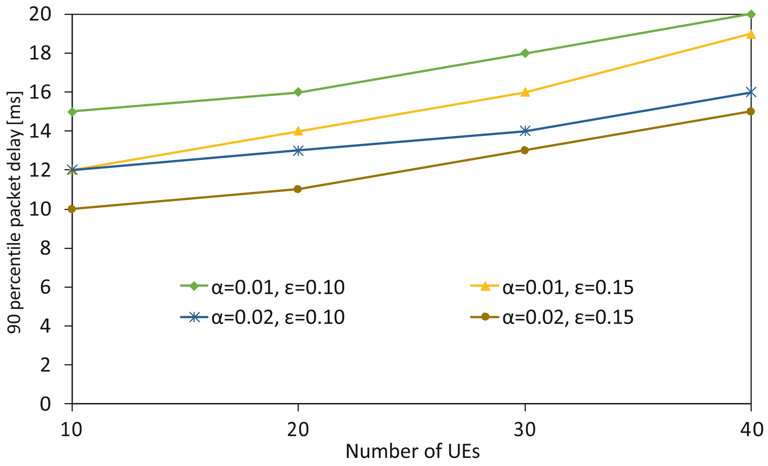

Fig. 3. 90 percentile packet delay vs. system load

5 Conclusion and Discussion

In this paper, we present how to use EWMA algorithm to define the measurement of GBR of a bearer and show how it facilitates uplink resource allocation. We discuss the characteristics of traffic patterns under the constraint of EWMA and explain that EWMA is a suitable algorithm to accommodate the expectation of both UEs and eNBs. Simulation results also suggest that the more bursty traffic patterns, which need to be declared by smaller EWMA weighting factors, should be charged with higher rates because they induce higher impact on the eNB.

It is for further study about how to select parameter values for different kinds of practical traffic. It would also be worthy to analyze the performance when an eNB accommodates UEs with traffics of different parameter values.

Acknowledgements. This paper is supported by Ministry of Science and Technology of Taiwan under grant No. 104-2221-E-197-016. The author would like to thank Jia-Hao Xu and Zhen-Hao Huang for helping the simulation work.

References

1. Ruiz de Temino, L., Berardinelli, G., Frattasi, S., Mogensen, P.: Channel- aware scheduling algorithms for SC-FDMA in LTE uplink. In: IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC (2008)
2. Liu, F., She, X., Chen, L., Otsuka, H.: Improved Recursive Maximum Expansion Scheduling Algorithms for Uplink Single Carrier FDMA System. In: IEEE 71st Vehicular Technology Conference (VTC 2010-Spring), pp. 1–5 (2010)
3. Iosif, O., Banica, I.: LTE uplink analysis using two packet scheduling models. In: 19th Telecommunications Forum (TELFOR), pp. 394–397 (2011)

4. Lee, S.-B., Pefkianakis, I., Meyerson, A., Xu, S., Lu, S.: Proportional fair frequency-domain packet scheduling for 3GPP LTE uplink. In: IEEE INFOCOM, pp. 2611–2615 (2009)
5. Calabrese, F.D., et al.: Search-tree based uplink channel aware packet scheduling for UTRAN LTE. In: IEEE Vehicular Technology Conference, VTC Spring 2008, pp. 1949–1953
6. 3GPP, TS 36.300 v. 12.4.0 (2015)
7. Hatoum, R., Hatoum, A., Ghaith, A., Pujolle, G.: Qos-based joint resource allocation with link adaptation for SC-FDMA uplink in heterogeneous networks. Presented at the 12th ACM International Symposium on Mobility Management and Wireless Access (2014)
8. Safa, H., El-Hajj, W., Tohme, K.: A QoS-aware uplink scheduling paradigm for LTE networks. In: IEEE 27th International Conference on Advanced Information Networking and Applications, AINA, pp. 1097–1104 (2013)
9. Marwat, S.N.K., Zaki, Y., Goerg, C., Weerawardane, T., Timm-Giel, A.: Design and performance analysis of bandwidth and QoS aware LTE uplink scheduler in heterogeneous traffic environment. In: 8th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 499–504 (2012)
10. 3GPP, TS 23.401 v. 13.2.0. General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access (2015)
11. Exponentially weighted moving average (EWMA). https://en.wikipedia.org/wiki/Moving_average
12. Kuo, F.-C., Ting, K.-C., Wang, H.-C., Tseng, C.-C., Chen, M.-W.: Differentiating and scheduling LTE uplink traffic based on exponentially weighted moving average of data rate. Mob. Netw. Appl. (2016). https://doi.org/10.1007/s11036-016-0693-9. (First Online: 15 February 2016)
13. Kuo, F.-C., Ting, K.-C., Wang, H.-C., Tseng, C.-C.: On demand resource allocation for LTE uplink transmission based on logical channel groups. Mob. Netw. Appl. (2017)