QoS-provisioning Downlink Resource Management in 4G Cellular Systems

You-Chiun Wang
Department of Computer Science
and Engineering,
National Sun Yat-sen University,
Kaohsiung 80424, Taiwan
Email: ycwang@cse.nsysu.edu.tw

Song-Yun Hsieh
Department of Computer Science
and Engineering,
National Sun Yat-sen University,
Kaohsiung 80424, Taiwan
Email: m023040061@student.nsysu.edu.tw

Abstract—**Long term evolution (LTE) has become one of the dominating communication standards for 4G cellular systems. It employs the orthogonal frequency division multiple access (OFDMA) technique to enhance the efficiency of downlink transmissions, which divides the communication spectrum into multiple *resource blocks*. How to efficiently distribute these resource blocks among user equipments (UEs) is a challenging issue, which also significantly affects the system performance of LTE. Furthermore, in order to provide different quality of service (QoS) for various types of traffic flows, LTE defines two categories of traffic flows: *guaranteed bit rate (GBR)* and *non-guaranteed bit rate (non-GBR)*. In the paper, we develop an efficient resource management mechanism to allocate downlink resource blocks to UEs so as to support QoS for GBR traffic flows while avoid starving non-GBR traffic flows. Our mechanism takes into account various network conditions such as channel quality, head-of-line (HOL) packet delay, QoS class identifier (QCI) value, and buffer length. Through the LTE-Sim network simulator, we demonstrate that the proposed mechanism can reduce both packet dropping and delay of GBR traffic flows. In addition, non-GBR traffic flows can still keep relatively high throughput to improve system performance.**

Keywords—*4G network, cellular system, downlink traffic, long term evolution (LTE), quality of service (QoS).*

I. INTRODUCTION

Nowadays, we have entered the epoch of 4G cellular communication and look forward to the coming of 5G. Numerous countries have launched their service of *long term evolution (LTE)* to provide wireless, pervasive Internet access. In order to support broad-band, high-speed data communication, LTE introduces several emerging techniques. For instance, it adopts the communication technology of orthogonal frequency division multiple access (OFDMA), which allows simultaneous transmissions of multiple downlink traffics [1], [2]. LTE also supports heterogeneous networks, in the sense that various types of base stations (called E-UTRAN Node B, and usually abbreviated as *eNodeB*) are able to coexist and cooperate in the same network, which provides flexible system deployment and configuration [3]. In addition, the technique of carrier aggregation (defined in the advanced version, LTE-A) helps the eNodeB to integrate multiple carriers located in different frequency bands, so as to increase the overall transmission bandwidth [4].

LTE adopts the concept of *traffic flows* to provide quality of service (QoS) to different applications, and it classifies all traffic flows into *guaranteed bit rate (GBR) traffic flows* and *non-guaranteed bit rate (non-GBR) traffic flows*. As their name would suggest, GBR traffic flows (for example, conversational voice and video, or real-time gaming) usually have strict delay and bandwidth demands. Therefore, they have to be prioritized over non-GBR traffic flows so as to meet their demands [5]. On the other hand, LTE divides the downlink spectral resource into multiple *resource blocks*, where each user equipment (UE) needs to obtain some resource blocks before receiving the data of a traffic flow sent from the eNodeB. Apparently, how to allocate resource blocks to UEs (and their traffic flows) plays an important role in LTE downlink performance. However, the LTE standard leaves the implementation of resource-block allocation to manufacturers and researchers [6].

Therefore, this paper focuses on investigating how to efficiently distribute resource blocks among UEs with two primary objectives. First, the QoS requirements of GBR traffic flows should be satisfied in order to reduce their packet latency (and also the packet dropping ratio). Second, non-GBR traffic flows can still obtain the necessary amount of downlink resource to prevent them from starvation. To achieve these two objectives, this paper proposes a *QoS-provisioning downlink resource management (QDRM) mechanism* for LTE systems. The idea is to first give the preliminary allocation of resource blocks to UEs based on their current channel conditions. Then, the eNodeB can calculate a portion of resource to be 'reallocated' later (in order to improve the system performance) by referring to the QoS class identifier (QCI) value, the buffer length, and the number of data bits carried by one resource block. Finally, traffic flows will compete such resource according to their head-of-line (HOL) packet delay and QCI values. By conducting simulations on LTE-Sim, experimental results demonstrate that our QDRM mechanism can outperform common LTE resource scheduling approaches, in terms of the above two objectives.

The rest of this paper is organized as follows. In the next section, we first briefly introduce the LTE network model, and then discuss the related work. Section III proposes our QDRM mechanism for LTE downlink communication. Simulation study is presneted in Section IV. We finally make a conclusion in Section V.

TABLE I. LTE TRAFFIC FLOWS AND THEIR QOS REQUIREMENTS, WHERE DELAY IS MEASURED IN MILLISECONDS.

Category	QCI	Delay	Loss rate	Applications
GBR flows	1	100	10^{-2}	VoIP
	2	150	10^{-3}	Live-streaming video
	3	50	10^{-3}	Real-time gaming
	4	300	10^{-6}	Buffered-streaming video
Non-GBR flows	5	100	10^{-6}	IMS signaling
	6	300	10^{-6}	Video, TCP-based services
	7	100	10^{-3}	Interactive gaming
	8	300	10^{-6}	Video, TCP-based services
	9	300	10^{-6}	Video, TCP-based services

TABLE II. TRANSMISSION BANDWIDTHS DEFINED IN LTE AND THE CORRESPONDING NUMBER OF RESOURCE BLOCKS.

Bandwidth (MHz)	1.4	3	5	10	15	20
Resource blocks	6	15	25	50	75	100

II. PRELIMINARY

A. Overview of LTE

LTE defines GBR and non-GBR traffic flows for message transmission of various applications, as presented in Table I. Broadly speaking, GBR traffic flows support real-time applications, for example, voice over IP (VoIP), video (live streaming or buffered streaming), and real-time gaming. On the other hand, non-GBR traffic flows usually support non-real-time applications such as IP multimedia subsystem (IMS) signaling and TCP-based services (including www, email, ftp, chatting, etc.). A UE can possess multiple traffic flows, each corresponding to one application. The QoS requirement of each traffic flow can be described by both packet delay and loss rate. For ease of management, LTE associates every traffic flow with a scalar identifier called *QoS class identifier (QCI)* to indicate its demand for packet delay and loss rate. LTE totally defines nine QCI values for different types of applications (listed in Table I) [7].

In LTE, the OFDMA technique is applied to transmit downlink traffics, which means that the eNodeB is able to simultaneously employ multiple sub-carriers to transmit data to one single UE. Specifically, the eNodeB partitions its downlink spectral resource into a number of *resource blocks*, which are the basic unit for resource allocation. Each resource block is a two-dimensional array with twelve successive sub-carriers and one time slot (occupying 0.5 milliseconds, which contains six or seven OFDMA symbols). The transmission bandwidth of one sub-carrier is 15 kHz, which means that a resource block has the overall bandwidth of 180 kHz. There are totally 44 frequency bands reserved for LTE operation in the world, and different LTE operating bands are employed in different countries or regions [8]. Each LTE band may have different amount of bandwidth and therefore possess different number of resource blocks. Table II presents available transmission bandwidths and their corresponding number of resource blocks supported by LTE.

B. Related Work

Maximum Throughput (MT) and *Proportional Fair (PF)* are two traditional resource management schemes in LTE [9]. As its name would suggest, MT always selects the UE with the best channel quality to allocate resource, so as to maximize the overall throughput. To provide certain degree of fairness,
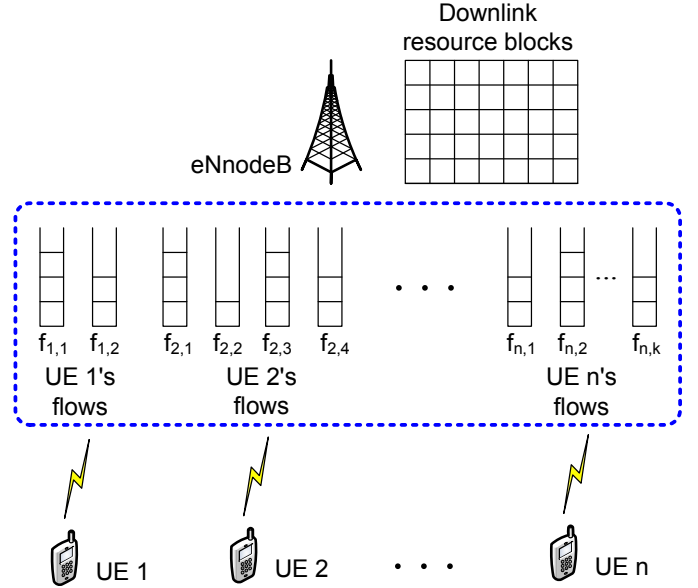

Fig. 1. System architecture of the QDRM mechanism.

PF compares the current data rate of every UE with its average (past) throughput. Then, it takes the UE with the largest value for scheduling. However, none of these two methods consider the delay requirement of real-time applications. On the other hand, both *Exponential Proportional Fair (EXP/PF)* and *Modified Largest Weighted Delay First (M-LWDF)* [10] introduce the concept of average HOL packet latency and, they prefer scheduling those UEs which encounter large packet latency. In [11], Liu et al. modify the PF method by taking into account both the HOL packet delay and the packet remaining time in the buffer before being discarded. In this way, the above three methods favor real-time applications over non-real-time ones. However, they do not take advantage of the property of GBR/non-GBR traffic flows and their QCI values defined in LTE (that is, Table I).

A large number of research efforts [12], [13], [14], [15] investigate how to manage LTE downlink resource with the help of *multi-input multi-output (MIMO)*. Specifically, MIMO is an advanced transmission technology used to improve system performance by exploiting the spatial domain of wireless fading channels [16]. There are two common types of MIMO in LTE systems. In a *single-user MIMO* system, every resource block should be given to no more than one UE. On the other hand, in a *multi-user MIMO* system, different spatial streams can be assigned to multiple UEs, so that they could share the same resource block. Therefore, multi-user MIMO usually results in better system performance comparing with single-user MIMO. However, MIMO requires the support of specialized hardware (for example, multiple antennas) and may incur higher designing complexity. In this paper, we thus consider that resource blocks are not sharable, and develop an efficient downlink resource management mechanism with the goals of providing QoS to GBR traffic flows, while keeping the throughput of non-GBR traffic flows higher to avoid starving them.

III. QoS-provisioning Downlink Resource Management (QDRM) Mechanism

Figure 1 presents the system architecture of our QDRM mechanism, where the eNodeB is serving n UEs and attempts to distribute its downlink resource blocks among these UEs. Each UE can have more than one traffic flow for communication. The eNodeB will maintain a *buffer* for every traffic flow to store its outstanding packets (to be received by the corresponding UE). Because LTE defines nine types of traffic flows (referring to Table I), the eNodeB needs to maintain at most nine buffers for each UE to keep track of its traffic flows.

According to the LTE specification, each UE has to notify the eNodeB of its current channel quality for resource-allocation reference. This operation involves multiple signalling and message exchanges, and we leave the details in [8]. Without loss of generality, we assume that the eNodeB will automatically discard those out-of-date packets in each buffer. Therefore, UEs will not receive stale data which have already passed their deadlines.

Based on the system architecture in Figure 1, the QDRM mechanism is composed of the following three stages:

- **Stage 1 – Preliminary resource allocation:**
  By referring to the channel quality of each UE, the eNodeB first computes how many resource blocks can be allocated to the UE. Then, each UE deals out the received resource blocks to its traffic flows.

- **Stage 2 – Calculation of reallocating resource:**
  The eNodeB then searches for 'disposable' resource blocks, which can be the residual resource blocks in the system or a subset of resource blocks allocated to traffic flows in the previous stage.

- **Stage 3 – Reallocation of resource blocks:**
  The aforementioned disposable resource blocks are finally reallocated to traffic flows according to their HOL packet delay and QCI values.

Below, we discuss the details of each stage.

A. Preliminary resource allocation

In the first stage, our major objective is to increase the overall throughput. Therefore, we employ the strategy of MT (discussed in Section II-B), where the eNodeB first selects the UE U_i which has the best channel quality to serve. In particular, we have

$$U_i = \arg \max_{j=1..n} (u_j(t)), \qquad (1)$$

where $u_j(t)$ is the channel rate of a UE U_j at the current time t. Then, the eNodeB allocates the necessary number of resource blocks to U_i to satisfy its traffic demand. The aforementioned operation is repeated until either 1) all UEs have been served or 2) all resource blocks have been dealt out. The first condition means that the eNodeB has sufficient downlink resource to meet the traffic demand of every UE. In this case, the QDRM mechanism need not execute the remaining two stages and thus it works the same as the MT scheme. On the other hand, the second condition indicates that some UEs may not receive enough resource blocks and

therefore the eNodeB has to consider adjusting the allocation of resource blocks to improve the system performance.

When the second condition occurs, we deal out the resource blocks of each UE to all of its traffic flows. Specifically, let r_i be the number of allocated resource blocks to a UE U_i, and $F_i = \{f_{i,1}, f_{i,2}, \cdots, f_{i,k}\}$ be the set of traffic flows of U_i, where $k \leq 9$. Then, we consider two possible cases. In the case of $r_i \leq |F_i|$, which implicitly indicates that some traffic flows of U_i cannot get any resource block, we then assign resource blocks according to the QCI values of traffic flows. In particular, staring from the traffic flow with the smallest QCI value (that is, $f_{i,1}$), we then allocate one resource block to every traffic flow of U_i (until there is no resource block). On the other hand, in the case of $r_i > |F_i|$, which means that some traffic flows of U_i can receive more than one resource block, we then assign resource blocks based on the buffer lengths of traffic flows. In particular, the number of resource block assigned to a traffic flow $f_{i,j}$ should be proportional to its buffer length. For example, suppose that UE U_i has received 8 resource blocks and F_i contains three traffic flows with buffer lengths of 6, 8, and 2 packets. Then, we assign traffic flows $f_{i,1}$, $f_{i,2}$, and $f_{i,3}$ with 3, 4, and 1 resource blocks, respectively. When there still remain some resource blocks, we assign them to traffic flows according to the rules in the previous case.

B. Calculation of reallocating resource

After assigning resource blocks to UEs (and their traffic flows), the eNodeB then seeks to find out disposable resource blocks to be reallocated later. There are two possible sources for such disposable resource blocks. First, the eNodeB can check whether the system has residual resource. In particular, we roughly distribute resource blocks to traffic flows in the first stage. However, each resource block in fact can carry different number of data bits, depending on the current channel quality of the corresponding UE. More specifically, when the UE experiences a better channel condition, the eNodeB can employ a more complex *modulation and coding scheme* (MCS, for example, 64QAM) to encode data bits in its resource blocks. In this case, the UE can actually use fewer resource blocks for communication. Therefore, for each traffic flow $f_{i,j}$ of a UE U_i, the eNodeB can calculate the necessary number of resource blocks required to transmit its packets:

$$r_{i,j} = \left\lceil \frac{b_{i,j} \times p_{i,j}}{RB_i} \right\rceil, \qquad (2)$$

where $b_{i,j}$ is the number of packets in traffic flow $f_{i,j}$'s buffer (that is, the buffer length), $p_{i,j}$ is the size of a packet in traffic flow $f_{i,j}$, and RB_i is the number of data bits which can be carried by one resource block allocated to UE U_i (under its current channel quality). Apparently, when traffic flow $f_{i,j}$ is given more than $r_{i,j}$ resource blocks from the first stage, the eNodeB can get back the 'extra' resource blocks by marking them as *disposable* (otherwise, they will be simply wasted because traffic flow $f_{i,j}$ does not have sufficient packets to consume these resource blocks).

The other source of disposable resource blocks comes from those traffic flows which obtain 'too many' resource blocks from the first stage. This situation would occur when some traffic flows have smaller QCI values, request a large amount of traffic transmission, and experience better channel conditions.

One representative is a *video* traffic flow (with QCI value = 2). In this case, the preliminary resource allocation in the first stage could let the video traffic flow get a large number of resource blocks, which would starve other traffic flows. To deal with this problem, we set a threshold number of $\alpha \in \mathbb{N}$ resource blocks for each traffic flow. Here, the threshold α should be larger than one, because we have to reserve at least one resource block for the traffic flow to transmit its packets. In particular, suppose that one traffic flow $f_{i,j}$ has obtained a number of $r_{i,j}$ resource blocks from the first stage, where $r_{i,j}$ is larger than α. Then, the eNodeB marks a number of $\lceil \beta \times (r_{i,j} - \alpha) \rceil$ resource blocks of traffic flow $f_{i,j}$ as *disposable*, where $0 < \beta < 1$, which means that these resource blocks can be reallocated to other traffic flows in the next stage.

C. Reallocation of resource blocks

In the last stage, the eNodeB reallocates the disposable resource blocks to those traffic flows whose demands have not been satisfied yet. Because the number of disposable resource blocks may not be sufficient to meet the demands of all traffic flows, the eNodeB first allocates the disposable resource blocks to the traffic flows which will encounter packet dropping soon. In particular, the eNodeB can check the HOL packet delay of each traffic flow. When the HOL packet of a traffic flow $f_{i,j}$ will pass its deadline (by referring to the delay requirement mentioned in Table I) if the HOL packet cannot be sent out at the current scheduling period, then traffic flow $f_{i,j}$ has a higher priority to obtain the disposable resource block(s). When there are two or more such traffic flows, the eNodeB allocates the disposable resource blocks according to their QCI values (from smaller to larger). After the aforementioned allocation, if there still remain disposable resource blocks, they are reallocated to traffic flows whose demands are not satisfied yet (also based on the QCI sequence).

IV. SIMULATION STUDY

We adopt LTE-Sim, which is an open-source framework developed to simulate LTE systems [17], to evaluate the performance of our QDRM mechanism. In our simulations, we consider an LTE macro-cell coordinated by one eNodeB. The transmission range of the eNodeB is two kilometers, and it covers a number of UEs. UEs will roam inside the macro-cell following the random waypoint mobility model [18], with the average moving speed of three kilometers per hour. Each UE may have multiple traffic flows. For GBR traffics, we consider both VoIP (with QCI value = 1) and H.264-video (with QCI value = 2) traffic flows. For non-GBR traffics, we employ constant-bit-rate (CBR, with QCI value = 6) traffic flows.

The transmission bandwidth of the eNodeB is 20 MHz, so it can support up to 100 resource blocks (referring to Table II). To simulate the LTE communication channel, we apply the following equation to model the effect of path loss (PL):

$$PL = 128.1 + 37.6 \log_{10} d, \tag{3}$$

where d is the distance between the eNodeB and one UE (measured in kilometers). To model the shadowing fading effect, we employ the *log-normal distribution* with zero mean and eight decibels of standard deviation. We compare the proposed QDRM mechanism with three popular LTE downlink resource management schemes, including the MT, M-LWDF,

(a) VoIP traffic flows (GBR)

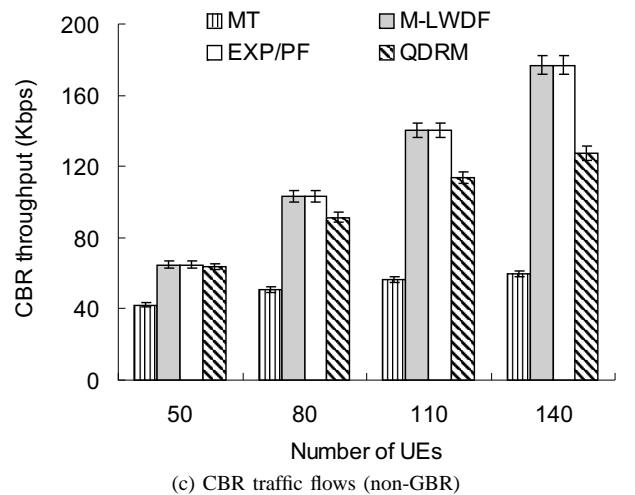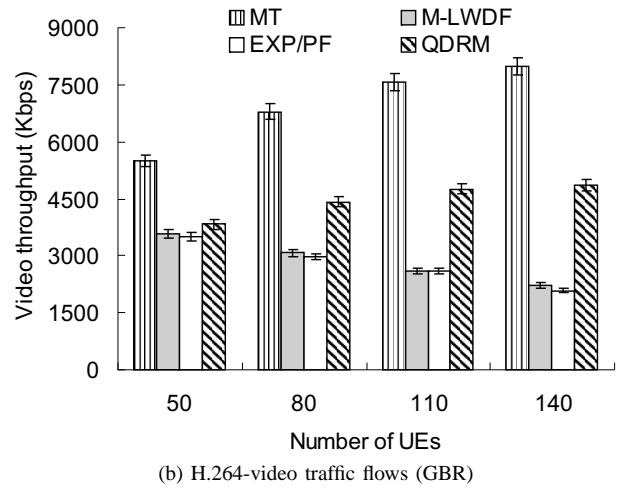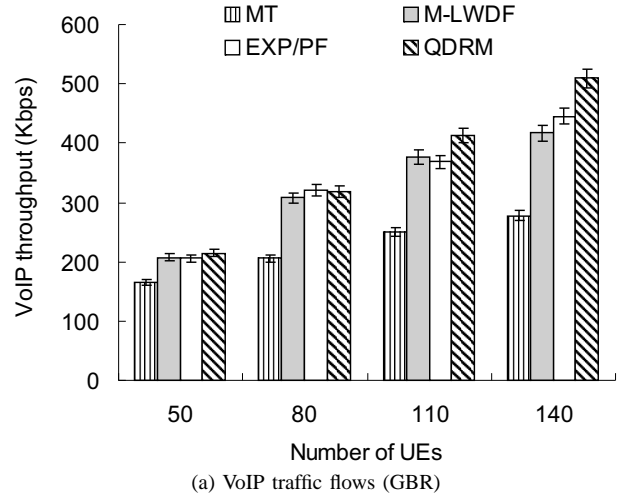(b) H.264-video traffic flows (GBR)

(c) CBR traffic flows (non-GBR)

Fig. 2. Comparison on the average data throughput of different traffic flows.

and EXP/PF schemes discussed in Section II-B. (Here, we skip the PF scheme because the EXP/PF scheme is an improved version of the PF scheme by considering the HOL packet latency.) For our QDRM mechanism, we set $\alpha = 3$ and $\beta = 0.5$. The total simulation time is 90 seconds.

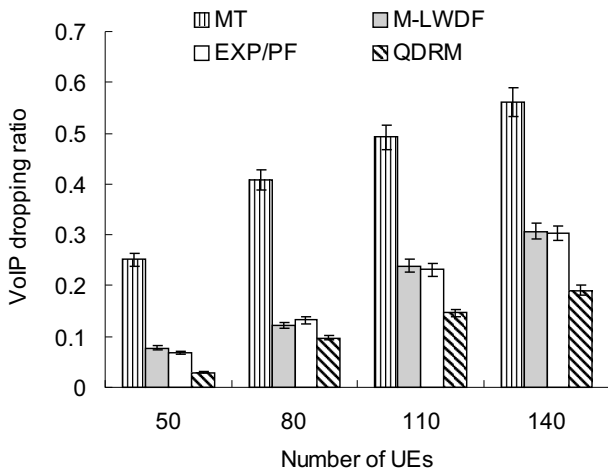Figure 2 presents the average data throughput of VoIP,

Fig. 3. Comparison on the average packet dropping ratio of VoIP traffic flows.
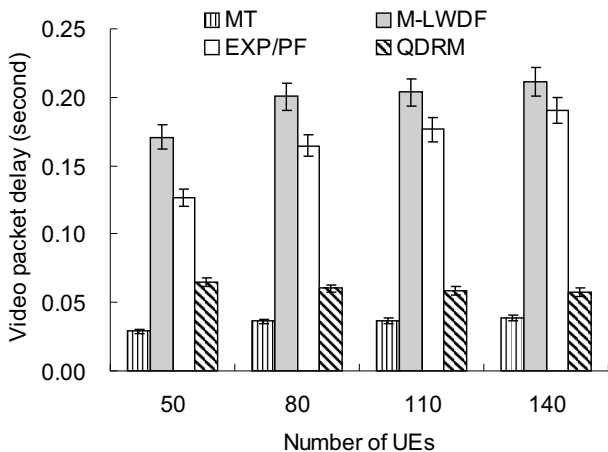

Fig. 4. Comparison on the average packet delay of H.264-video traffic flows.

H.264-video, and CBR traffic flows (measured in kilo-bits per second, Kbps), where the number of UEs is ranged from 50 to 140. For the MT scheme, H.264-video traffic flows require quite large bandwidth, and the MT scheme always selects the UE with the best channel quality for resource allocation. Thus, the MT scheme will let H.264-video traffic flows have the highest data throughput, at the cost of other traffic flows. On the other hand, because of the PF property, both the M-LWDF and EXP/PF schemes would prefer allocating more resource to VoIP and CBR traffic flows. In this case, because the amount of downlink resource is constant, the average data throughput of H.264-video traffic flows will decrease when the number of UEs grows. Our QDRM mechanism takes into account not only the HOL packet delay but also the QCI values of traffic flows. Because VoIP and H.264-video traffic flows have smaller QCI values, they will be allocated more resource blocks for transmission. On the other hand, the QDRM mechanism will not allocate the most portion of downlink resource to GBR traffic flows in order to avoid starving non-GBR traffic flows. Therefore, it results in much higher data throughput of CBR traffic flows compared with the MT scheme.

We then observe the effect of different resource man-

agement schemes on GBR traffic flows. Figure 3 shows the average packet dropping ratio of VoIP traffic flows. Because most of downlink resource is occupied by H.264-video traffic flows in the MT scheme, VoIP traffic flows will suffer from serious packet dropping. In this case, the voice quality will significantly degrade, especially when there are more UEs. For example, when there are 140 UEs in the LTE macro-cell, the eNodeB will drop more than 50% packets of a VoIP traffic flow in the MT scheme. On the contrary, by allocating resource blocks according to the QCI value, our QDRM mechanism can always have the lowest packet dropping ratio for VoIP traffic flows.

On the other hand, Figure 4 illustrates the average packet delay of H.264-video traffic flows (measured in seconds). As mentioned earlier, the MT scheme allocates the most downlink resource to H.264-video traffic flows. Therefore, it will have the lowest packet delay. By giving a higher priority to GBR traffic flows, the QDRM mechanism can significantly reduce the average packet delay of H.264-video traffic flows, as compared with both the M-LWDF and EXP/PF schemes.

To sum up, the MT scheme helps reduce the packet delay of H.264-video traffic flows, but it drops a large number of VoIP packets and causes non-GBR CBR traffic flows to starve. On the other hand, the M-LWDF and EXP/PF schemes attempt to fairly distribute downlink resource to all traffic flows. However, they do not consider the large bandwidth requirement of H.264-video traffic flows, which results in pretty higher video packet delay. Finally, our QDRM mechanism not only meets the QoS demand of GBR traffic flows (where VoIP traffic flows have a smaller packet dropping ratio and H.264-video traffic flows have lower packet delay), but also allows non-GBR traffic flows to receive the necessary amount of downlink resource for transmission (which results in much higher CBR throughput than the MT scheme).

V. Conclusion

Conventional LTE resource management schemes such as MT, PF, M-LWDF, and EXP/PF do not differentiate GBR traffic flows from non-GBR ones, so they may not meet the QoS requirement of GBR traffic flows. Therefore, this paper proposes the QDRM mechanism to manage the downlink spectral resource in an LTE system. Our QDRM mechanism allows the eNodeB to search for disposable resource blocks after a preliminary resource allocation to UEs. Then, these disposable resource blocks are reallocated to the traffic flows according to their QCI values and HOL packet latency. In this manner, we can satisfy the delay and bandwidth demands of GBR traffic flows while prevent non-GBR traffic flows from starvation. Through simulations, we show that the proposed QDRM mechanism can alleviate packet dropping of VoIP traffic flows, reduce average packet delay of H.264-video traffic flows, and keep relatively higher throughput of non-GBR CBR traffic flows, which demonstrates its effectiveness.

REFERENCES

[1] J. M. Liang, J. J. Chen, Y. C. Wang, and Y. C. Tseng, "A cross-layer framework for overhead reduction, traffic scheduling, and burst allocation in IEEE 802.16 OFDMA networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1740–1755, 2011.

[2] S. Srikanth, P. A. M. Pandian, and X. Fernando, "Orthogonal frequency division multiple access in WiMAX and LTE: a comparison," *IEEE Communications Magazine*, vol. 50, no. 9, pp. 153–161, 2012.

[3] Y. C. Wang and C. A. Chuang, "Efficient eNB deployment strategy for heterogeneous cells in 4G LTE systems," *Computer Networks*, vol. 79, pp. 297–312, 2015.

[4] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, "Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 122–130, 2012.

[5] Y. C. Wang, S. R. Ye, and Y. C. Tseng, "A fair scheduling algorithm with traffic classification in wireless networks," *Computer Communications*, vol. 28, no. 10, pp. 1225–1239, 2005.

[6] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[7] 3GPP TS 23.203, "Technical specification group services and system aspects; policy and charging control architecture (release 8)," V8.9.0, 2010.

[8] 3GPP TS 36.101, "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception (release 10)," V10.3.0, 2011.

[9] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in UTRA long term evolution downlink," in *IEEE International Symposium on Wireless Pervasive Computing*, 2008, pp. 308–313.

[10] R. Basukala, H. A. M. Ramli, and K. Sandrasegaran, "Performance analysis of EXP/PF and M-LWDF in downlink 3GPP LTE system," in *Asian Himalayas International Conference on Internet*, 2009, pp. 1–5.

[11] B. Liu, H. Tian, and L. Xu, "An efficient downlink packet scheduling algorithm for real time traffics in LTE systems," in *IEEE Consumer Communications and Networking Conference*, 2013, pp. 364–369.

[12] H. Zhang, N. Prasad, and S. Rangarajan, "MIMO downlink scheduling in LTE systems," in *IEEE INFOCOM*, 2012, pp. 2936–2940.

[13] J. Niu, D. Lee, X. Ren, G. Y. Li, and T. Su, "Scheduling exploiting frequency and multi-user diversity in lte downlink systems," in *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, 2012, pp. 1412–1417.

[14] Y. Xu, H. Yang, F. Ren, C. Lin, and X. Shen, "Frequency domain packet scheduling with MIMO for 3GPP LTE downlink," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1752–1761, 2013.

[15] J. Fan, G. Y. Li, and X. Zhu, "Multiuser MIMO scheduling for LTE-A downlink cellular networks," in *IEEE Vehicular Technology Conference*, 2014, pp. 1–5.

[16] C. X. Wang, X. Hong, X. Ge, X. Cheng, G. Zhang, and J. Thompson, "Cooperative MIMO channel models: A survey," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 80–87, 2010.

[17] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE cellular systems: an open-source framework," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 498–513, 2011.

[18] W. H. Yang, Y. C. Wang, Y. C. Tseng, and B. S. P. Lin, "Energy-efficient network selection with mobility pattern awareness in an integrated WiMAX and WiFi network," *International Journal on Communication Systems*, vol. 23, no. 2, pp. 213–230, 2010.