Reception Cycle-aware Resource Block Allocation for Real Time Video Traffic in 3GPP LTE System

Chin Li, Jen-Yi Pan, and Ming-Hung Hsu
Department of Communications Engineering
National Chung Cheng University
Chiayi, Taiwan
xxu0627@gmail.com , jypan@comm.ccu.edu.tw , abc19891010@hotmail.com

***Abstract*— Discontinuous reception (DRX) in Long-Term Evolution reduces power consumption of users' equipment (UE) by periodically turning on and off the radio reception module. As the module falling asleep, packets from base stations are deferred and even useless due to exceeding delay bounds in real time applications such as video streaming. Current resource allocators only skip transmission opportunities of those sleeping UEs, rather than consider their reception cycle and traffic delay bounds.**

This study proposes a mechanism of resource block allocation called Reception Cycle-aware resource block Allocation (RCA), which defers most transmissions until the last moment, and reserves a minimal resource block for waking up each UEs that has traffic going expired, to reduce packet loss/miss rate (PLR). Simulation results show that RCA can more effectively reduce power consumption and cope with packets missing deadline than traditional methods do. Users may enjoy real-time applications and also a longer battery life just with realization of our mechanism only at the base station side.

Keywords—Discontinuous reception (DRX), Transmission Scheduler, Long Term Evolution (LTE), Power Saving, Energy Efficiency, Real Time Video.

I. INTRODUCTION

Long Term Evolution (LTE) is one of the fourth-generation (4G) wireless technologies. It offers higher transmit data rate comparing with 3G technologies with novel techniques such as multi-input multi-output (MIMO), orthogonal frequency-division multiplexing (OFDM), and so on. However, the receiver requires computationally intensive circuit to decode 4G signal. The circuit drains users' equipment (UE) battery power quickly. Discontinuous Reception (DRX) is a power saving mechanism in LTE, which extends battery life by turning on and off radio reception module in cycles.

The most ideal power saving case is that the UE only wakes up when the packets just arrive at the base station (BS), and then the BS transmits those packets to the UE. In other words, DRX's performance strongly depends on the traffic pattern. The BS shall determine the most appropriate DRX parameters according to different traffic pattern. For example, real time video traffic is characterized by a fixed video frame rate, and within a frame there are several packets of various sizes. For receiving this traffic, a long period of UE's sleeping may cause long delay, even resulting in packets missing deadline of playing out that deteriorate user's experience.

In addition to traffic pattern, DRX's performance also relates to scheduling and resource block allocation. Orthogonal frequency-division multiple access (OFDMA) is a method of encoding data on multiple sub-carriers with multiple access using time and frequency separation of multiple users, which is adopted by 3GPP Long Term Evolution (LTE) in downlink radio resource allocation. OFDMA splits wireless bandwidth into many subcarriers. The size of a subcarrier is 15 kHz. A resource block (RB) is consistent of twelve subcarriers and is the minimal allocation unit in LTE. We assume the BS allocates RBs to N UEs every transmission time interval (TTI), which is 1 millisecond long. In DRX mode, even though packets for the UE arrive at the BS and the UE is awake, the packets still can't be delivered if no RB is allocated for the UE. As the UE doesn't get any RBs, the transmission is delayed to the next TTI at which the UE wakes up.

Existing resource allocations have a problem. When the UE is in DRX, it might wake up several times within delay bound, causing redundant receptions. The proposed method defers most transmissions until the last moment (reception) so that reduces active time and saves power. The proposed method also reserves a minimal resource block for waking up each UEs that has traffic going expired, to reduce packet loss/miss rate (PLR).

Reference [1] proposed a QoS-Aware downlink scheduling algorithm for VoIP. If a UE is achieving lower than average data rate required by the Guaranteed Bit Rate (GBR), then the scheduler will increase priority of that UE to fulfill the GBR requirement and vice versa. However there still has the redundant receptions. Reference [2] proposed an efficient scheme and a DRX-aware packet scheduling method. The scheme reduces UEs' unnecessary wake-up periods incurred by resource competition by interleaving DRX cycles. For those UEs going into DRX cycles at the next TTI, the packet scheduler will allocate them one RB if their virtual queues contain stringent packets which might be dropped in the next TTI.

The next section introduces DRX and conventional resource allocation in detail. Section III explains the proposed framework of Reception Cycle-aware resource block Allocation (RCA). Section IV describes the simulation and results. Finally, the section V is the conclusion.

II. BACKGROUND

A. Discontinuous Reception

Discontinuous Reception (DRX) [3] saves UE's power consumption by periodically turning on and off the radio reception module according to parameters as shown in Fig. 1.

1) Inactivity Timer: This timer specifies the number of consecutive Transmission Time Intervals (TTI) since the latest reception during which the UE shall monitor if any downlink data transmission for this UE. The UE goes into Short DRX cycle when no packet arrives in monitored consecutive TTIs; otherwise, the UE stays in active cycle and always turns on the reception module.

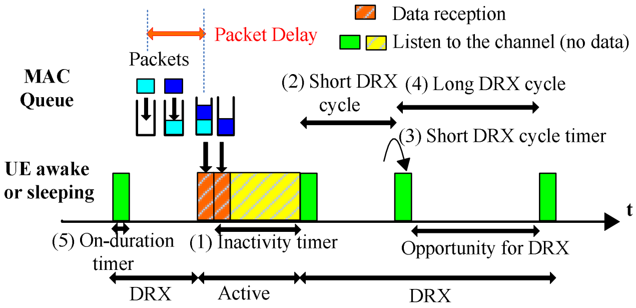

Fig. 1. DRX operation and its parameters

2) Short DRX Cycle: This parameter specifies the number of consecutive TTIs meaning a shorter operational cycle, consisted of on-duration (in this paper also called transmission opportunity (TxOP)) and opportunity for DRX [4]. The UE periodically turns on the reception module for an on-duration and then turns off for an opportunity for DRX, during which may save power consumption.

3) Short DRX cycle Timer (Ns): Short DRX cycle Timer specifies the number of times of short DRX cycles. When no packet arrives for Ns cycle, the UE goes into long DRX cycle.

4) Long DRX cycle: Long DRX cycle specifies the number of consecutive TTIs, meaning a longer operational cycle than short DRX cycle. The period of opportunity for DRX in long DRX cycle is usually longer than short DRX cycle's, but on-duration period is the same.

5) On-duration timer: Specifies the number of consecutive TTIs during which the UE shall monitor the downlink transmission for possible allocations. If the transmission for the UE is scheduled, the UE will go active to receive packets. Otherwise, the UE keeps sleeping.

6) DRX offset: DRX offset stands for when the on-duration timer starts if the UE turns into DRX cycle. The DRX offset in the case of Fig. 1 is set to zero and not shown for simplicity of drawing.

An UE is awake when it is in active mode or in on-duration of DRX mode. As the UE falling asleep, packets are deferred in the base station and are received until the next time of the UE wakeing up. So the longer sleeping time, the longer packet delay. If the packet delay exceeds delay bounds in real time applications, the packet will be dropped .

B. Resource allocation

Dynamic RB scheduler runs every TTI and consists of two parts, time-domain and frequency-domain schedulers [5]. In the first part, time-domain scheduler selects those UEs who are awake and have pending downlink packets. The selection result is called Scheduling Candidate Set (SCS).

In the second part, frequency-domain scheduler allocates N_{RB} resource blocks (RB) for UEs in SCS. The value of N_{RB} is defined in [6] according to different band-width operated in the BS. There are many different resource allocation policies introduced in the survey [7]. In the literature [8], Proportional Fair (PF) and Modified Largest Weighted Delay First (M-LWDF) are the most popular policies, which allocate resource block according to their defined metric function for each UE.

1) Proportional Fair (PF): The PF algorithm allocates resource block based on (1), where $R_i(t)$ is average data rates of the UE_i. Every RB, said RB_j, is allocated to the UE who has the maximum metric among N UEs in (1). $r_{i,j}(t)$ is the instantaneous data rate of UE_i in resource block j.

$$\max\ metric = \max_{1\le i\le N} \frac{r_{i,j}(t)}{R_i(t)} \tag{1}$$

The $R_i(t)$ is defined in (2), where the value of β is between 0 and 1, and set to 1/1000 in [9]. $r_i(t)$ is instantaneous data rate of UE_i, which is also the rate summation of its allocated resource blocks at time interval t.

$$R_i(t+1) = (1-\beta)\cdot R_i(t) + \beta \cdot r_i(t) \tag{2}$$

2) M-LWDF: The selective metric in M-LWDF algorithm is defined in (3), where $r_i(t)$ and $R_i(t)$ are the same definition as in PF algorithm. $w_i(t)$ is the head-of-line delay of the UE_i in the MAC queue. T_i and δ_i are the tolerable delay of packets and packet loss rate.

$$\max metric = \max_{1\le i\le N} \frac{-\log \delta_i \cdot r_{i,j}(t)\cdot w_i(t)}{T_i \cdot R_i(t)} \tag{3}$$

3) DRX-Aware scheduling: In addition to conventional mechanisms, reference [10] proposes the DRX-Aware scheduling (DAS) mechanism, whose metric definition is shown in (4). An UE_i with a shorter remaining active time has a higher priority by the term $f_i(t)$.

$$\max metric = \max_{1\le i\le N} \frac{-\log\delta_i \cdot r_{i,j}(t)\cdot w_i(t)}{T_i \cdot R_i(t)} \cdot f_i(t) \qquad (4)$$

Conventional resource block allocation, for example PF and M-LWDF, might be appropriate for the UE without DRX, but not for the UE with DRX. As shown in Fig. 2, the UE_x has only two transmission opportunities before the delay bound expiring. In these two TxOPs, The BS possibly allocates the RBs for other UEs with higher values of metric function, and the packets of UE_x will be dropped.

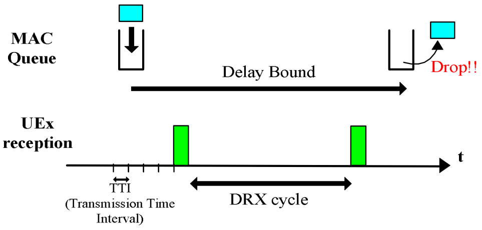

Fig. 2. Resource allocation for UE with DRX

Instead, reference [10] considers the DRX parameters with the modified metric function (4) in resource allocation. The closer the UE is to the sleeping moment, the easier the UE gets allocated RBs.

III. PROPOSED FRAMEWORK

Discontinuous Reception (DRX) mechanism can conserve the battery of the UE by periodically turning on and off radio modules. Base stations (BS) have DRX parameters of each UE, and dispatch its traffic when the UE is awake. Each UE goes to sleep after a predefined period of receiving no traffic, and periodically wakes up to receive BS's dispatching. If its traffic arrives, then the UE stops DRX and goes to active mode. Meanwhile, the real time video traffic has burst data arriving at regular intervals, which should be played out within a delay bound. For a short DRX cycle, this traffic often wakes up the UE, causing little battery conservation. For a long DRX cycle, these burst data may be queued up to miss an opportunity of delivering to periodically active UE, even dropped until the deadline expires.

Hence, we propose a framework with two added algorithms in the conventional resource block allocation and traffic scheduling for DRX, as shown in Fig. 3. The proposed algorithm 1 picks up UEs with critical traffic from the candidate set of time-domain scheduler to the next stage and hence conserves battery life of non-critical UEs. The algorithm 2 reserves resource blocks for UEs with the last chance and hence avoids packets' expiring when the downlink resource is congested.

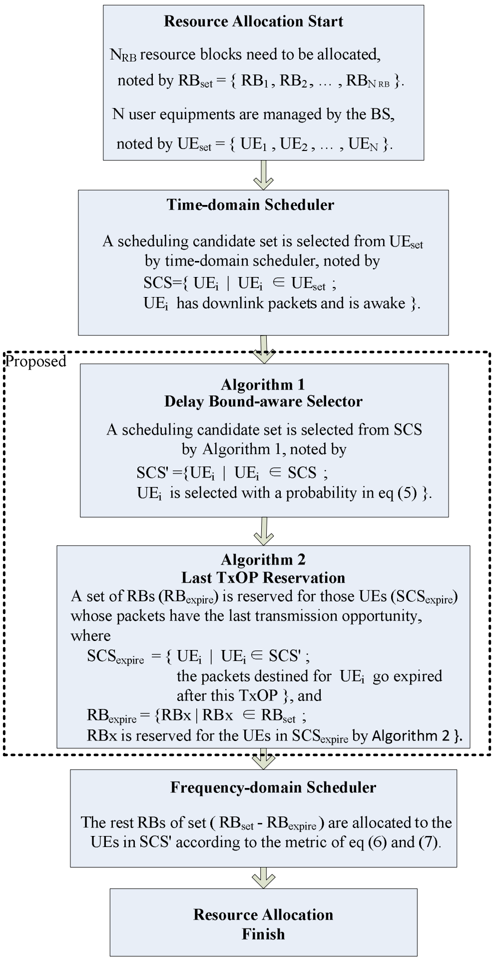

Fig. 3. Framework of Reception Cycle-aware Resource Block Allocation

A. Algorithm 1: Delay bound-aware selector

The original scheduler allocates resource blocks for each awake UEs until all the blocks are occupied or all the awake UEs are served. Fig. 4 gives an example of real-time periodic traffic. The UE wakes up to receive every packet just arrived, even though the delay bound of packets does not run out.

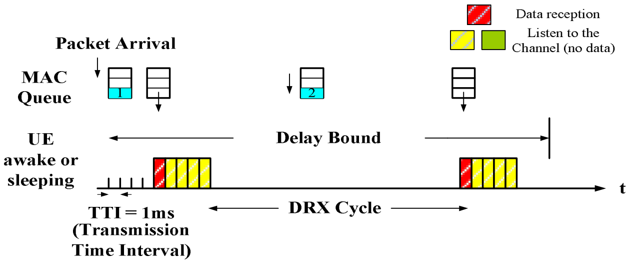

Fig 4. An example of original time-domain scheduler

Fig. 5 shows the basic idea of the algorithm 1. If the UE can tolerate the delay of packets less than the delay bound, then these packets can be queued before expiration to conserve battery energy. For example, the UE in Fig. 5 turns on radio module for 6 TTIs, which are less than 10 TTIs of the case in Fig. 4.

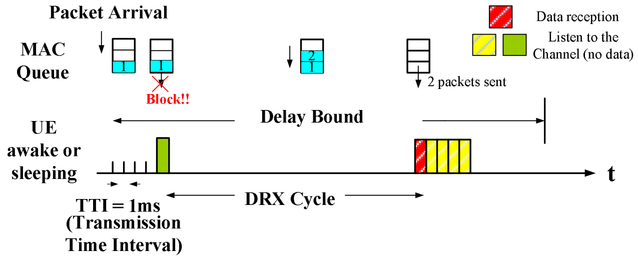

Fig 5. An example of time-domain scheduler with algorithm1

Fig. 6 shows the selection algorithm considering delay bound for each UE's traffic. To prevent packet congestion in the last moment, we define a probability to pick up those UEs to be scheduled. The more critical delay bound of the UE is, the easier the UE is picked up. The probability is defined as

$$P\{UE_i \text{ is chosen}\} = 1 - \frac{remainingTxOP_i}{numberTxOP_i}, \qquad (5)$$

where remainingTxOP and numberTxOP are the left number of transmission opportunity before expiration and the total number of DRX transmission opportunity of the head-of-line packet for the UE, respectively.

```
Algorithm1 ( ){
  SCS' = $\phi$
  For each UE$_i$ in SCS
  {
      PacketArrivalTime$_i$ = arrival time of the head-of-line packet
                              in the queue destined to UE$_i$
      RemainingTxOP$_i$ =floor( (DelayBound – (Now –
                             PacketArrivalTime$_i$)) / DRXcycle$_i$ )
      NumberTxOP$_i$ = floor( DelayBound / DRXcycle$_i$)
      Select UE$_i$ with probability (1- RemainingTxOP$_i$ / NumberTxOP$_i$ )
      If (UE$_i$ is selected) {
          SCS' = SCS' $\bigcup$ { UE$_i$ }
      }
  }
  return SCS' ;
}
```

Fig 6. Pseudo code of Algorithm1 Delay bound-aware selector

B. Algorithm 2: Last transmission opportunity reservation

The original frequency-domain scheduler allocates resource blocks just according to the value of the metric equation. However, the scheduler does not consider whether the UEs in scheduling candidate set are available in the next transmission time interval. If some UEs cannot get resource blocks in this TTI, they might go back to sleep in the next TTI, probably resulting in their packets missing the delay bound, especially for those UEs in the last transmission opportunity. This situation may be worse when a lot of packets are pending in the queue of base station. The left of Fig. 7 shows an example of this case.

The UE 5, 7, and 8 are in DRX mode and selected by algorithm 1 to join the scheduling in this interval. However, the UE 5 has the highest metric value and needs five resource blocks to convey its downlink packets. The frequency scheduler allocates five resource blocks to the UE 5, and one to the UE 7. The UE 8 does not get any resource block in this round, and goes to sleep in the next round. Its packets then miss the deadline and are dropped by the video player. Although the algorithm 1 has a probability option to mitigate this effect, we still design the algorithm 2 to reserve resource blocks for such case.

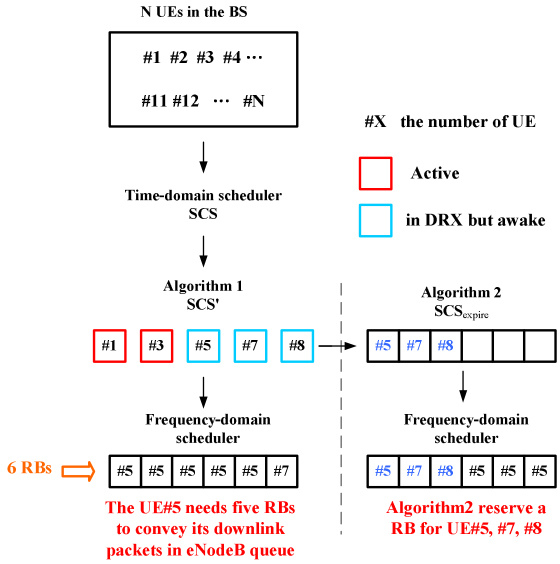

Fig 7. An example of resource block allocation of original and algorithm2

The right of Fig. 7 shows the proposed solution to reduce the packet loss in missing delay bound. The algorithm 2 picks up those UEs with the last chance, e.g. UE 5, 7, and 8 in this example, and then reserves a resource block for each of these UEs before frequency-domain scheduler. After getting the resource blocks, these UEs go to active mode and continuously receive their queued packets without loss. The algorithm 2 is shown in Fig. 8.

```
Algorithm2 ( ) {
    SCS$_{expire}$ = $\phi$
    RB$_{expire}$ = $\phi$
    For each UE$_i$ in SCS' {
      if (Now – PacketArrivalTime$_i$ + DRXcycle$_i$ $\geq$ DelayBound) {
         SCS$_{expire}$ = SCS$_{expire}$ $\bigcup$ { UE$_i$ }
      }
    }
    For each UE$_i$ in SCS$_{expire}$ {
      Randomly allocate an RB$_x$ in RB$_{set}$ to UE$_i$
      RB$_{expire}$ = RB$_{expire}$ $\bigcup$ { RB$_x$ }
    }
}
```

Fig 8. Pseudo code of Algorithm2 Last transmission opportunity reservation

C. Modification of metric function

After going through the algorithm 2, the frequency-domain scheduler allocates the remaining resource blocks according the metric function. We proposed a modified metric function based on the one in M-LWDF for UE_i and RB_j at time interval t, as follows

$$metric(i,j,t)=\frac{-\log\delta_i\cdot r_{i,j}(t)\cdot w_i(t)}{T_i\cdot R_i(t)}\cdot f_i(t) \qquad (6),$$

where at time interval t, $r_{i,j}(t)$ is the instantaneous data rate of UE_i in RB_j and $R_i(t)$ is for the average data rate of UE_i. δ_i and T_i are the maximum tolerable PLR and delay of packets, and $w_i(t)$ is the waiting time of head-of-line packet in the buffer. The $f_i(t)$ is the major different term from M-LWDF, and defined as follows.

$$f_i(t)=\frac{Inactivity\ \ counter_i}{Inactivity\ \ timer_i}\quad \text{; if UE}_\text{i}\text{ is in active mode}$$

$$f_i(t)=\frac{On-duration\ \ counter_i}{On-duration\ \ timer_i}\quad \text{; if UE}_\text{i}\text{ is in DRX mode}$$

Inactivity counter reset to zero when packets arrive in active mode. On-duration reset to zero when the UE goes to DRX mode. These two counters count up every TTI until they reach the value of inactivity timer and on-duration timer, respectively.

At time interval t, the frequency-domain scheduler allocates each resource block j to UE k according to the following equation,

$$k=\arg\max_{1\le i\le N}\ metric(i,j,t) \qquad (7)$$

In a short summary, the proposed framework filters out those non-critical UEs with a probability, and reserves resource blocks for critical UEs. The remaining resource blocks are then scheduled according to the modified metric function by prioritizing those UEs with a shorter remaining active time.

Comparing to the conventional scheduling and a method in reference [10], those schemes still allocate RBs to those UEs whose metric values are small, for example that packets just arrive at the BS, resulting in frequent waking up and more power consumption. In addition, those schemes may cause starvation of few critical UEs in the case of congested downlink resource. On the other hand, our proposed framework can alleviate these two cases in a simple way, and conserve battery life under a controllable delay bound.

IV. NUMERICAL RESULT

Our simulator deals with events of packet generation, RB allocation, RLC layer processing and UE DRX state transition at the time scale of TTIs. Table I lists the system parameters. We assume the UEs are around BS, so the channel quality of RBs is always the best. In the start of simulation, 105 users watch the video simultaneously. We use the traffic model called near real time video traffic [11]. There are 8 packets within a frame whose length is 0.1 second. The packet size and inter-arrival time between packets both are truncated Pareto distribution. When the all 128kbps video streams are downloaded, the simulation is finished. Inactivity timer and On-duration timer are setting to 10 TTIs and 2 TTIs respectively for each UE. The setting of DRX offset refers to the reference [13], and we staggered the DRX offset among UEs. Delay bound and tolerated packet loss rate (PLR) are 0.1s and 1%. We don't adopt the short DRX cycle.

Table I
System Parameters

The number of BS	1
The number of UE	105
The number of RB per TTI	25
Delay Bound	100ms
UE Location	around BS
Video Size	10 MBytes
Promised PLR	1%
Modulation and coding	64 QAM (I_{MCS} = 28)
DRX parameters	
Inactivity Timer	10ms
On-duration Timer	2ms
DRX Offset	Staggered [13]

We use three measures to assess the performance: power consumption, packet loss rate and packet delay. The UE power consumption model refers to the reference [12]. The consumptions for awake and sleeping modes are 500 mW and 20 mW, respectively. Equation (8) is the equation of average power consumption. Packet delay and loss rate are defined in (9) and (10).

$$\text{Average power (mW) of UE}_\text{i}=\frac{500\cdot awake\ time_i+20\cdot sleeping\ time_i}{awake\ time_i+sleeping\ time_i} \qquad (8)$$

$$\text{Packet Delay} = the\ time\ \text{of}\ packets\ being\ received\ in\ the\ \text{UE} - packet's\ arrvial\ time\ in\ the\ BS \qquad (9)$$

$$PLR=\frac{the\ number\ \text{of}\ packets\ exceeding\ delay\ bound}{the\ number\ of\ all\ packets} \qquad (10)$$

We compare the power consumption with PF, M-LWDF and DAS1 [10] in Fig. 9. A larger DRX cycle length saves more power. Algorithm1 of RCA buffers the packets until later TxOPs, so the waking-up frequency is less than other methods. Notably, the packets almost have only one chance to send when the DRX cycle of the UE about equals to delay bound (0.1s). In this situation, the proposed method doesn't defer the transmission, performing the similar action as in DAS1 because they have similar metric functions in frequency domain schedulers. Therefore, the longer DRX cycle the UE has, the closer the line of RCA to the DAS1's.

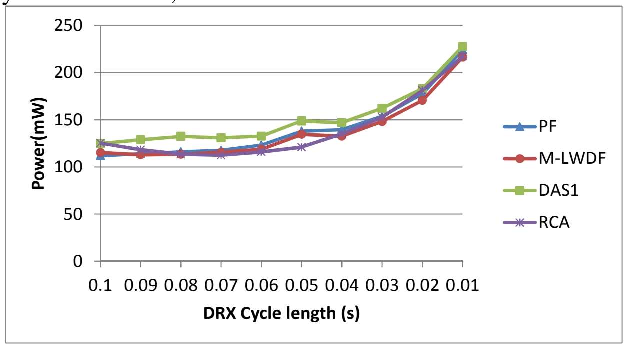

Fig 9. Power consumption of four methods with varying DRX cycle length

Fig. 10 shows when the DRX cycle length is longer, the PLR is higher. The PF scheduler has the worst performance in real-time traffic because it doesn't consider the head-of-line delay. DAS1 method considers that the closer the UE to sleeping moment, the higher weight the UE gets, so DAS1 performs better than M-LWDF. In RCA, UEs in DRX mode gets at least an RB to receive packets if they wake up in the last TxOP. Therefore, RCA performs better in the PLR comparison.

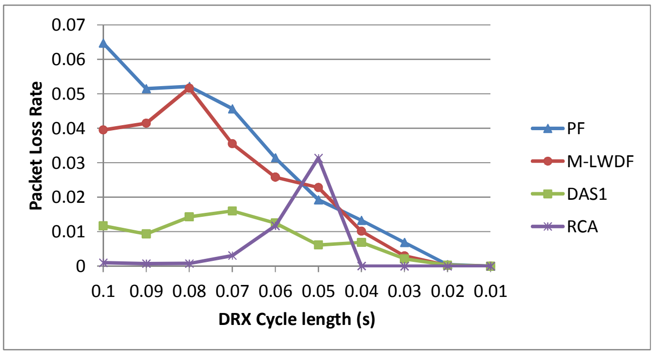

Fig. 10. Packet loss rate of four methods with varying DRX cycle length

Fig. 11 shows that RCA temporary blocking the packets in the BS raises the packet delay. The PF, M-LWDF and DAS1 have similar performance. In RCA, the longer DRX cycle length, the longer packet delay. When the cycle length is longer than 0.05s, the TxOPs are almost the last chance such that algorithm 1 doesn't filter out these UEs, hence reducing the packet delay.

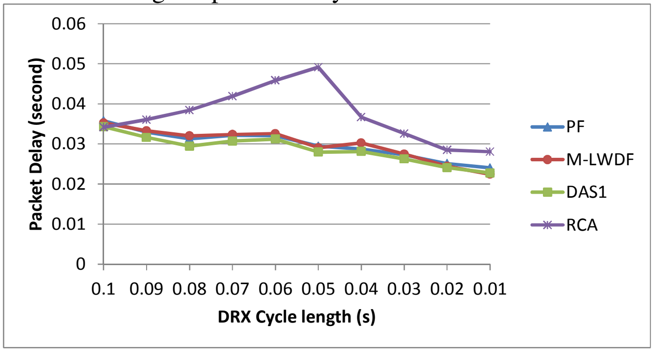

Fig. 11. Packet delay of four methods with varying DRX cycle length

Table II

The minimum power with guaranteed PLR of 1%

Scheduling Type	**PF**	**M-LWDF**	**DAS1**	**RCA**
DRX Cycle length (s)	0.037	0.040	0.089	0.075
Average Power Consumption (mW)	142.370170	132.474487	122.005336	112.056937
PLR	0.982%	0.8894%	0.914%	0.1628%

Supposing the promised PLR is set to 1%, Table II shows the minimum power of these four methods. The RCA method has the minimum power consumption, when DRX cycle equals to 0.075s. Notably, the PLR of RCA is much smaller than the promised PLR because the algorithm 2 always efficiently controls the packet delay.

V. CONCLUSION

We propose the RCA method. At first we find that the UE has redundant receptions. The proposed RCA could reduce power consumption by buffering packets and then sending in the later TxOp. However, the more buffered packets exceed the delay bound, the more the packets get lost. RCA allocates RBs to those UEs in the last TxOP and wake them up to receive packets. Our simulation shows that RCA effectively saves power under the promised PLR in real time video traffic and the BS could choose the appropriate DRX parameters when users watch videos. Users may have longer usage time without service quality deterioration. In the future, we will discuss DRX mechanism for other traffic types.

ACKNOWLEDGEMENT

The authors thank the National Science Council, Taiwan, for partially supporting this research under contract NO. NSC 101-2221-E-194-005 and NSC 102-2221-E-194-036.

REFERENCES

[1] J.M. Liang, J.J. Chen, H.H. Cheng and Y.C. Tseng, "An Energy-Efficient Sleep Scheduling With QoS Consideration in 3GPP LTE-Advanced Networks for Internet of Things," IEEE Emerging and Selected Topics in Circuits and Systems, vol.3 no.1, Mar. 2013

[2] M.S. Mushtaq, A. Shahid and S. Fowler, "QoS-Aware LTE Downlink Scheduler for VoIP with Power Saving," IEEE 15th Computational Science and Engineering, Dec. 2012

[3] Chandra S. Bontu and Ed Illidge, "DRX mechanism for power saving in LTE," IEEE Communications Magazine, vol.47 no.6, p.48-55, June 2009

[4] 3GPP TS 36.321 V10.5.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification," Mar. 2012.

[5] D. Vinella and M. Polignano, "Discontinuous reception and transmission (DRX/DTX) strategies in Long Term Evolution(LTE) for voice-over-IP traffic under both full-dynamic and semi-persistent packet scheduling policies," master dissertation, Aalborg University, Denmark, November 20th, 2009.

[6] 3GPP Std. TS 36.101 V8. 1.0, "User Equipment (UE) radio transmission and reception", Mar. 2008

[7] 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 Commun. Surveys & Tutorials, vol. 15,no. 2, Jun., 2012.

[8] S.C. Nguyen, "Optimised proportional fair algorithm for longterm evolution-advanced system with multiple component carriers," IET Communications , vol. 6, no. 11, pp. 1579 - 1586 Jul. 2012

[9] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R.Vijayakumar, "Providing quality of service over a shared wireless link,"IEEE Commun. Mag., vol. 39, no. 2, pp. 150-154, Feb,2001.

[10] H. Bo, T. Hui, C. Lan, and Z. Jianchi, "Drx-aware scheduling method for delay-sensitive traffic," IEEE Communications Letters, vol. 14, no. 12, pp. Dec. 1113-1115, 2010.

[11] IEEE 802.16m-08/004r5, IEEE 802.16m evaluation methodology document (EMD), 2009 [online] Available: http:/www.ieee802.org/16/tgm/core.html

[12] 3GPP TSG RAN WG2 meeting 57bis, R2-071285, DRX parameters in LTE, Mar. 2007.

[13] 3GPP TSG RAN WG2 meeting 63, R2- 084003, DRX Configuration of DRX, Aug. 2008.