Performance Analysis of Buffer Aware Scheduling for Video Services in LTE Network

Meng-Hsien Lin[1, 2] and Yen-Wen Chen[1]
[1] Department of Communication Engineering, National Central University, Taoyuan, Taiwan 32001
[2] Chunghwa Telecom Laboratories Taoyuan, Taiwan 32661
*Corresponding author: Yen-Wen Chen
[e-mail : ywchen@ce.ncu.edu.tw]

Received April 10, 2015; revised July 6, 2015; accepted July 25, 2015; published September 30, 2015

Abstract

Recent advancements in broadband wireless communication technologies enable mobile users to receive video streaming services with various smart devices. The long term evolution (LTE) network provides high bandwidth and low latency for several emerging mobile applications. This paper proposes the buffer aware scheduling (BAS) approach to schedule the downlink video traffic in LTE network. The proposed BAS scheme applies the weighting function to heuristically adjust the scheduling priority by considering the buffer status and channel condition of UE so as to reduce the time that UE stays in the connected state without receiving data. Both of 1080P and 2160P resolution video streaming sources were applied for exhaustive simulations to examine the performance of the proposed scheme by comparing to that of the fair bandwidth (FB) and the best channel quality indicator (CQI) schemes. The simulation results indicate that the proposed BAS scheme not only achieves better performance in power saving, streaming delivery time, and throughput than the FB scheme while maintaining the similar performance as the best CQI scheme in light traffic load. Specifically, the proposed scheme reduces streaming delivery time and generates less signaling overhead than the best CQI scheme when the traffic load is heavy.

Keywords: LTE, resource allocation, scheduling

A preliminary version of this paper appeared in IEEE WPMC 2014, September 7-10, Sydney, Australia. This version includes a concrete analysis and more results especially for the heavy traffic load with detail discussions. The enhancement part is more than 30% of the previous conference paper. This research work was supported in part by the grants from the National Science Council (NSC) (NSC 98-2221-E-008-063, NSC 99-2218-E-159-001, NSC 100-2221-E-008-097, NSC 101-2221-E-159-026, and NSC 102-2221-E-008 -015), Taiwan, ROC.

http://dx.doi.org/10.3837/tiis.2015.09.017 **ISSN : 1976-7277**

1. Introduction

Broadband wireless communication has created a convenient environment in which people can ubiquitously access to the Internet and has facilitated the deployment of several novel mobile services. New applications, such as social network services, high definition video, and Internet of Things (IoT), are developed to satisfy emerging demands of daily life. In [1] and [2], Ericsson and Cisco estimated the video service is the fastest growing mobile data traffic and it is expected that over 50% of mobile data traffic will come from video service by 2019. The video service requires high bandwidth and severe quality demand; therefore, the radio resource allocation plays an important role for the service deployment in crowded environment [3]. In addition to the effective resource allocation, due to the limited power of mobile device, efficient power management is critical toward the successful service.

To facilitate high-speed and convenient wireless communication, the long term evolution (LTE) technology, which is evolved from Universal Mobile Telecommunications System (UMTS), has been proposed to support high bandwidth for mobile novel services [4, 5].In the LTE network, the Adaptive Modulation and Coding (AMC) scheme is applied to effectively use the spectrum of diverse channel conditions of user equipment (UE). Thus, a high level of the Modulation-and-Coding Scheme (MCS) is adopted to achieve high bandwidth if the channel condition is acceptable. Due to the distance increases, MCS index and throughput will also decrease. Moreover, the Orthogonal Frequency Division Multiple Access (OFDMA) technique is used for downlink transmission and thus improves system capacity. The evolved Node B (eNB), i.e., the base station, is responsible to deliver data to UEs and to schedule the radio resource for the uplink data transmission of UEs.

UEs consume more transmission power to gain higher data rates in the LTE network than in Wi-Fi and 3G [6] under the same channel condition. To save power, the specification allows UE to release the data path and transit to idle state if it does not need to transmit or receive data. Although the Discontinuous Reception (DRX) mechanism [5] can also provide power saving in the connected state, it only turns off the radio receiver and the other components still consume power. Thus entering idle mode is much more efficient to save power when compared to the DRX mechanism. As a result, eNB had better let UE transition to idle state if there is no data for the UE or eNB has no sufficient radio resource to allocate. And from the video streaming service satisfaction perspective, UE does not need receive video frames from eNB in hurry if the stock of video data in its local buffer is sufficient to play. Thus UE may also enter idle state to save power while playing video streaming and transition to connected state if it is going to exhaust its local backlog. However, frequent switching between the connected and idle states causes substantial radio resource control (RRC) related signaling overhead on the control plane. Such a high amount of RRC related signaling can affect eNB and core network directly in real operator environment [7, 8]. Additionally, if without proper arrangement, it may introduce wasting energy, and high latency when UE is resumed connected state from the idle state frequently. Thus the tradeoff among power saving and signaling overhead is a quite complex issue to achieve effective scheduling.

The available bandwidth for individual users in LTE networks is limited compared to the throughput demand of high-quality video streaming [19-21]. Due to limitation of spectrum resources, network component capability, and battery lifetime, how to balance resource scheduling, power consumption and user experience will become more challenging. The ideal scheduler is designed to carefully arrange limited radio resource to maximize the spectrum efficiency so that the signaling overhead and power consumption of UE can be minimized while the quality of user experience can be maintained. In this paper, the buffer aware scheduling (BAS) scheme is proposed based on considering buffer and channel condition of UE. This study was conducted to investigate the performance of the scheduler when the DRX mechanism is disabled. Both 1080P and 2160P video streaming were adopted for experimentation and comparison. The remainder of this paper is organized as follows. Section 2 presents the background and related works. In Section 3, the BAS scheme is illustrated. In Section 4, experimental results, analysis, and discussions are provided. The conclusions are drawn and suggestions for future research are provided in the last section.

2. Background and Related Work

With recent developments in wireless communications technology and popularity of smart devices, people require high speed and low latency wireless communication systems to connect to the Internet any time and location. To achieve this objective, the LTE network was proposed and specified by the 3GPP. The LTE is backwards-compatible with the UMTS and uses an architecture based on using pure packet switching to handle all types of applications. The packet switch based architecture provides the flexibility in arranging radio resource when compared to the circuit switch based approach. LTE comprises Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Evolved Packet Core (EPC), and relevant system interfaces with other communication networks. E-UTRAN is the radio-accessing part of the LTE system, and UE accesses the LTE core network through the air interface, named as LTE-Uu. The EPC network includes the Mobility Management Entity (MME), the Home Subscriber Server, the Serving Gateway (S-GW), and a Packet Data Network Gateway. E-UTRAN accesses the MME and S-GW through the S1-MME and S1-U interfaces, respectively. The OFDMA is applied for downlink transmission in LTE and provides the advantages of flexible subcarrier allocation and adaptive modulation with respect to changing channel conditions. In LTE network, ten 1-ms subframes compose one 10 ms frame to convey information. Although the Resource Element (RE) is the smallest transmission unit in LTE, the basic unit of resource allocation is either the Resource Block (RB) or the Resource Block Group (RBG), which are group of subcarriers. The LTE Downlink resource grid and LTE system features three types of resource allocation [9, 10, 11], named as group resource blocks based allocation, physical resource block based allocation, and virtual resource block based allocation. As the radio resource can be flexibly allocated, eNB could schedule the radio resource to the UEs that starve for the resource to meet their demand first and postpone the requirements of other UEs that are not badly in need of resource. In addition to the desired quality of services for each UE, eNB usually needs to consider the changeable channel condition during resource allocation so as to maximize system throughput and achieve fair treatment among UEs.

In LTE, the UE has only two RRC states (i.e., RRC_Idle and RRC_Connected), which is simpler than the UMTS system has five RRC states [12]. To save power, UE can generally switch between RRC_Connected and RRC_Idle states. And RRC_Connected state can be further subdivided into continuous reception and discontinuous reception/transmission (DRX) modes as shown in **Fig. 1**. There are two inactivity timers, named as T_i and T_{tail}, to enable UE to transition to DRX mode and RRC_Idle state, respectively, when it does not received any data during the timer period. In DRX mode, the long cycle is designed for UE to lengthen its sleep period if it has not received data in the short cycle for a specific time period, i.e., the timer T_{is}. Basically, when UE is in DRX mode, the established data path is kept and therefore, it spends less latency to recover from DRX mode to transmit or receive data. However, it still consumes more power when compared to the UE that is in RRC_Idle state. In addition, the duration of DRX is usually much shorter than the RRC_Idle state and the efficiency of saving power is not significant due to higher frequency of switching between continuous reception mode and DRX mode [6]. Hence this paper focuses on the power management of UE in switching between RRC_Connected and RRC_Idle states.

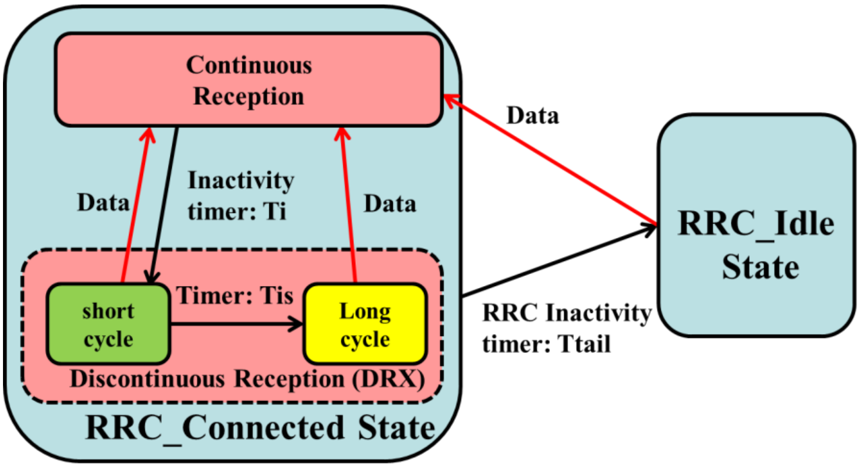


Fig. 1. Operations of RRC_Connected and RRC_Idle States

When the UE is in RRC_Idle state, it needs to spend time to execute Random Access Procedure and adopt RRC related signaling to seizure eNB via LTE-Uu interface to establish data path. The eNB needs to respond with S1 Application Protocol (S1AP) related signaling to MME via S1-MME interface to allow UE to transition the RRC_Connected state [12, 13]. **Fig. 2** illustrates the numbers of control messages required for RRC connection establishment and release between RRC_Connected state and RRC_Idle state. The RRC related signaling messages include RRC connection and security mode command. The RRC_Connected state offers continuous data reception capacity to reduce latency at the expense of more power consumption, while the RRC_Idle state is the other way round. Although it is a contradiction issue, eNB shall properly schedule the radio resource to achieve better tradeoff under the QoS consideration.

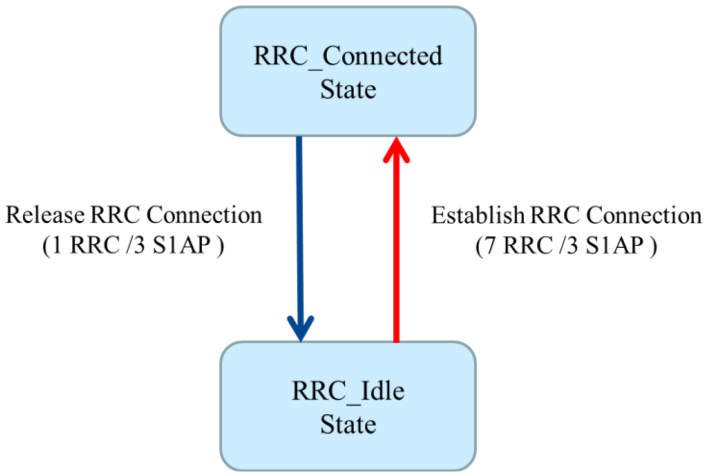


Fig. 2. Operations of RRC_Connected and RRC_Idle States

Numerous methods have been proposed to improve either power saving or scheduling efficiency. In [14], the researchers described a novel cross layer optimized mechanism, which enhanced eNB architecture to determine DRX operation during video streaming. The proportional fair scheduler at the upper and lower levels for downlink traffic was proposed by adopting the discrete-time linear control theory [15]. The new resource allocation scheme for RT video streaming traffic in downlink LTE by using downlink SNR values, average throughput, packet delay and buffer information was proposed. The resource allocation scheme for downlink video streaming traffic by using Signal-to-Noise Ratio (SNR) values was proposed in [16]. And the authors mentioned that, according to the simulation results, their scheme outperforms opportunistic technique. The authors in [17], the investigators presented a Quality of Service (QoS) aware scheduling algorithm, which jointly considered not only the system throughput, application QoS, but also fairness, for wireless real-time video delivery in multiple users' environment. In [18], a Quality of Experience aware radio resource management framework was proposed to allow the operator to enhance the video capacity. However, most of the previous studies have focused on either power saving or QoS scheduling management, few studies have investigated the effectiveness of the scheduler in power saving, signaling overhead, streaming delivery time and throughput. The main objective of the proposed BAS scheme is to improve the performance of the video streaming service with different resolutions in heavy traffic loading.

3. The Proposed Buffer Aware Scheduling Scheme

To investigate the power cost caused by switching between RRC_Connected and RRC_Idle states in LTE network, the power consumption model was provided in **Fig. 3** for analysis. The power consumptions of each state as well as the state transitions are also provided accordingly [6]. UE can only be allocated with radio resource when it is in RRC_Connected state. The figure shows that UE in RRC_Connected state, even without data reception (i.e., DRX mode), consumes much more power than in RRC_Idle state. Moreover, the power consumptions of each state and the state transition are specified.

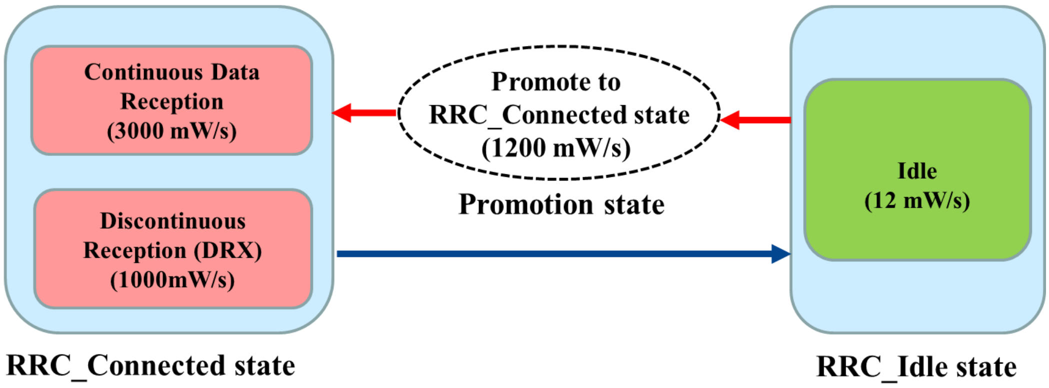


Fig. 3. The State Transitions of UE with Power Consumptions

From the throughput perspective, scheduler needs to maximize the spectrum efficiency. In terms of spectrum efficiency, OFDMA can be characterized as a type of elastic bandwidth. The total bandwidth of OFDMA system is determined by the MCS level, which is, in turn, determined by the channel condition between eNB and the UE. Hence, the higher levels of the MCS are allocated to the connected UEs, who have better channel condition, to achieve higher throughput. From the power saving perspective, the proper approach is to let UE receive as much data as possible to fill its buffer as soon as possible when UE is in RRC_Connected state so that UE can transition to RRC_Idle state with enough backlog to play. As mentioned in Section 2, the RRC inactivity timer is designed to serve as the entrance that allows UE to transition from connected state to RRC_Idle state. The UE wastes power during the inactivity timer countdown period if no radio resource is allocated. Consequently, shorting UE in RRC_Connected state without receiving data can achieve high power-saving efficiency. From the user experience perspective, the video stream shall continuously be played without interruption. Thus, UE needs to play the backlog video frames in the UE buffer periodically regardless of which state. As UE is going to exhaust its backlog in RRC_Connected state, the scheduler allocates radio resource to the UE. By contrast, while UE is going to exhaust its backlog in RRC_Idle state, the UE needs to wait for the paging cycle to request for the connection to eNB. Then UE can transition to RRC_Connected state to be allocated with radio resource for receiving video data. However, the radio resource is limited and it is meaningless to let too many UEs to stay in the RRC_Connected state. For the extreme case, the number of RRC_Connected UEs had better not exceed the number of RBG in one frame because RBG is the basic radio resource allocation unit. Therefore, the proposed BAS scheme considers the number of UEs in RRC_Connected state to determine whether it is suitable to allow the transition of UE from RRC_Idle state to RRC_Connected state. From the signaling perspective, frequent state transitions introduce massive RRC and S1-AP related signaling messages, to reduce the state transitions can also minimize the control plan overhead.

Normally, the resource is scheduled according to the priorities of UEs. In this paper, in addition to the channel quality of each UE, the proposed BAS scheme determine the priority by referring to the estimated backlog for playing in the UE buffer. The purpose behind the proposed BAS scheme is to schedule data in burst manner so as to reduce the frequency of state transition in streaming video traffic. The following notions are considered for scheduling priority in the proposed BAS scheme:

1. For RRC_Connected state UE, which has better channel condition (i.e., higher MCS levels), then has higher priority to be scheduled with data to achieve high system throughput.
2. For RRC_Connected state UE, whose backlog is going to be exhausted, has higher priority to be scheduled with data to reduce streaming delivery time and enhance user experience.
3. For RRC_Connected state UE, whose buffer is going to be full, is preferred to be scheduled with data so as to transition to RRC_Idle state as soon as possible for power saving;
4. For RRC_Idle state UE (in either state), who is going to exhaust its backlog, has higher priority to transit to RRC_Connected state for scheduling and will be constrained by the number of allowable UE in connected state to avoid too many UEs contending the limited radio resource, cause power wasting and signaling overhead.

In the proposed BAS scheme, eNB allocates RBG to UE according to the reported channel quality and estimated backlog video data in each UE. And, according to the above notions, the connected UE, whose buffer is going to be full or to be exhausted, will be given higher priority. The weighting function $W_{i,j}$ is heuristically designed in equation (1) to represent the preference of assigning RBG j to UE$_i$. A higher weighting value means the higher scheduling priority.

$$W_{i,j} = \begin{cases} (M_{i,j})^n X_{l(i)} & if\ 0 \le B_i \le B_l \\ (M_{i,j})^n & if\ B_l < B_i < B_h \\ (M_{i,j})^n X_{h(i)} & if\ B_h \le B_i \le B_{Max} \end{cases} \quad (1)$$

where B_i, B_l, B_h and B_{Max} are the estimated buffer length (or the backlog) of UE_i, the low buffer length threshold, high buffer length threshold and maximum buffer length of UE, respectively. The $M_{i,j}$ is the MCS level of UE_i on the *j-th* RBG. The parameters *n* denotes the impact degrees of channel condition. The parameter $X_{l(i)}$ in the equation (1) is applied to increase the UE's weight when the buffer length of UE is close to empty as stated in above notion 2. And the parameter $X_{h(i)}$ can increase the UE's weighting when the buffer of UE is going to be full and let UE transition to the RRC_Idle state as stated in above notion 3. These additional factors were proposed in equations (2) and (3), respectively.

$$X_{l(i)} = \alpha_l + \lambda_l \frac{(B_l - B_i)}{B_l} \tag{2}$$

$$X_{h(i)} = \alpha_h + \lambda_h (e^{(B_i - B_h)/(B_{Max} - B_h)} - 1) \tag{3}$$

where α_l and α_h are the initial values to determine the weighting value; λ_l and λ_h are the proportionality factors. The weighting function is illustrated in **Fig. 4**.

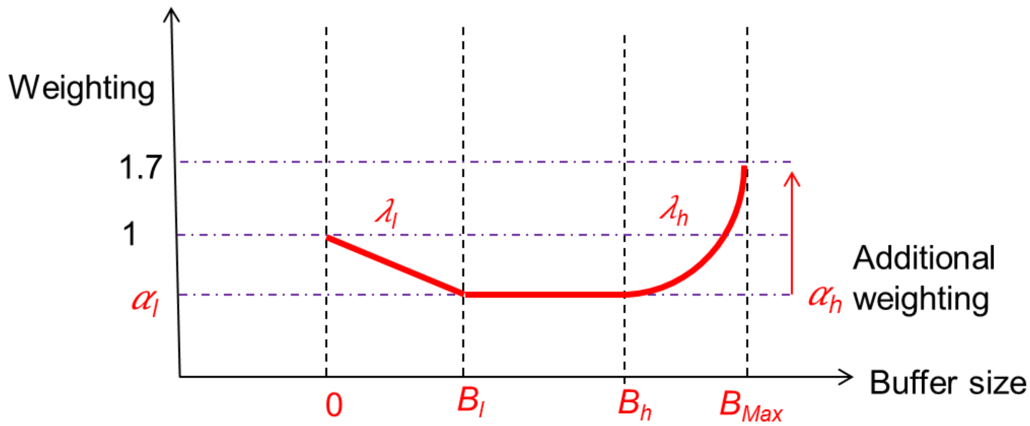


Fig. 4. Representation of Additional Weighting and Buffer Size

As shown in **Fig. 3**, UE will waste power if it is in RRC_Connected state, however, without data reception and promoting to RRC_Connected state. In the forward strategy, an attempt is made to increase the scheduling priority of a data packet that has a low weighting value, according to equation (1), by aggregating the packet with the current transmission packet, so that UE transitions to RRC_Idle state as soon as possible. **Fig. 5** presents an example in which Packets *P1* and *P2* will be scheduled to be transmitted to a specific UE. If the UE has been in the RRC_Connected state and its buffer is almost full, then, according to equation (3), this UE may get higher scheduling priority for *P2* due to the higher value of B_i and then UE can transition to RRC_Idle state for power saving after receiving *P1* and *P2*. This approach can decrease the period with which a UE stays the RRC_Connected state and thereby lengthen the RRC_Idle state period relative to that attained using the original approach. It is noted that the number of state transitions is three for the original scheduling; however, it is only one when applying the proposed approach.

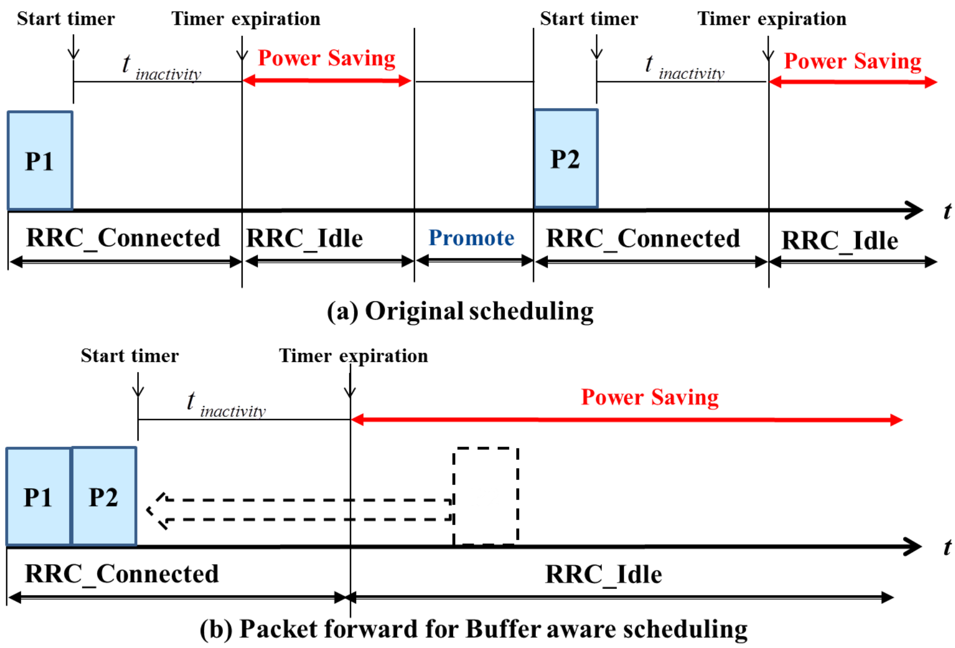


Fig. 5. Buffer Aware Scheduling – Forward Packet Concept

The backward strategy is based on the idea that eNB may postpone the scheduling of a data packet that is going to be scheduled according to original assigned priority; so that the UE can stay the RRC_Idle state if the backlog video of UE is enough to play. **Fig. 6** compares the original scheduling and the backward strategy applied by the proposed BAS scheme. The example in **Fig. 6** shows that the scheduling of Packet *P2* is postponed because the local buffer of the UE is enough to play; eNB allows the UE to stay the RRC_Idle state and aggregates *P2* and *P3* to transmit them in burst. Thus, using the backward approach, the power saving period is longer than that attained using the original approach.

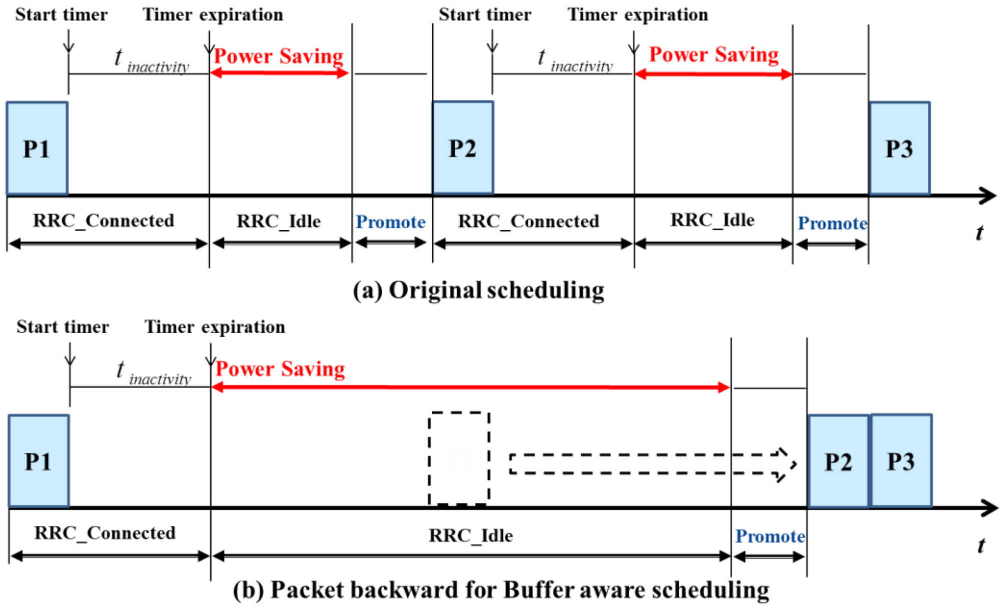


Fig. 6. Buffer Aware Scheduling –Packet Backward Concept

It is inappropriate to let too many UEs in the RRC_Connected state to contend limited radio resource and cause signaling overhead. For example, the number of UE in connected state shall not exceed the number of RBG. To avoid too many UEs contending the limited radio resources, the proposed BAS scheme estimates the allowable maximum number of UE in RRC_Connected state, denoted as δ, by considering the maximum capacity of each frame f_{bw} ,the maximum allocated bandwidth for one UE in previous LTE frame UE_{bw} , the consumed video data (i.e., video playing) of the UE in one LTE frame C_{bw} and the number of RBG in one frame R_n, as shown in equation (4). The proposed BAS scheme estimates the value of δ as follows.

$$\delta = \min\left\{ {}^{f_{bw}}\!/_{UE_{bw}}, {}^{f_{bw}}\!/_{C_{bw}}, R_n \right\} \tag{4}$$

The proposed BAS algorithm is illustrated in **Fig. 7**.

```
Algorithm: Buffer Aware Scheduling (BAS)
For all UEs are in RRC_Idle state
    If the estimated buffer length of UE <( C_bw * promotion delay)
        If the number of UEs are in RRC_Connected and promotion state< equation (4)
        the eNB pages the UE and promote UE to RRC_Connected state
        End If
    End If
End for
For all UEs are in promotion state
    If the promote is complete
        UE is in RRC_Connected state
    End If
End for
For each UE i in RRC_Connected state
    For each RBG j
    Calculate W_i,j according to equation (1),(2) and (3)
    Allocate the RBG j to UE i that has the highest W_i,j
    End for
End for
Allocate the radio resource and send the packet to UEs
```

Fig. 7. The Proposed BAS Algorithm

4. Experimental Results and Analysis

To investigate the performance of the proposed approach, exhaustive simulations were performed, with the focus being on analyzing the power consumption, signaling overhead, streaming delivery time and throughput per UE. The simulation environment included one eNB, which provided a channel bandwidth of 10 MHz and several UEs. The International Telecommunication Union (ITU) Veh-A channel model was applied for each UE and six MCS levels were adopted according to the condition. **Table 1** specifies the parameters used in the simulation.

Table 1. Parameters of Simulation Environment

Parameters	Content
Channel model	ITU Veh-A
Channel bandwidth	LTE FDD 10MHz
RB Number	50
RBG size	3
RBG Number	17
CQI report interval	10ms
Resource allocation type	Type 0
Subcarrier per RB	12
Subcarrier	15 KHz
RRC Inactivity Timer	10ms
Promotion delay	260ms
Paging cycle	320ms
MCS, Kbits/ RB per LTE frame	QPSK 1/2, 1 QPSK 3/4, 2.5 16QAM 1/2, 4 16QAM 3/4, 5 64QAM 1/2, 6 64QAM 3/4, 7.2

1080P and 2160P resolutions video traffic models with H.265 were applied in the simulations. The average data rates of 1080P and 2160P video traffics were 1.65 Mbps and 2.65 Mbps, respectively. And the local buffer size of each UE, max buffer length, low buffer length threshold and high buffer length threshold were assumed to be 3600Kbits, 3600Kbits, 500Kbits and 3000Kbits, respectively. In addition to power consumption, the volume of RRC related signaling messages was measured and investigated in the simulations. The streaming delivery time indicates the user need spend and wait watching video streaming. The throughput per UE represents the average bandwidths that are allocated to UE.

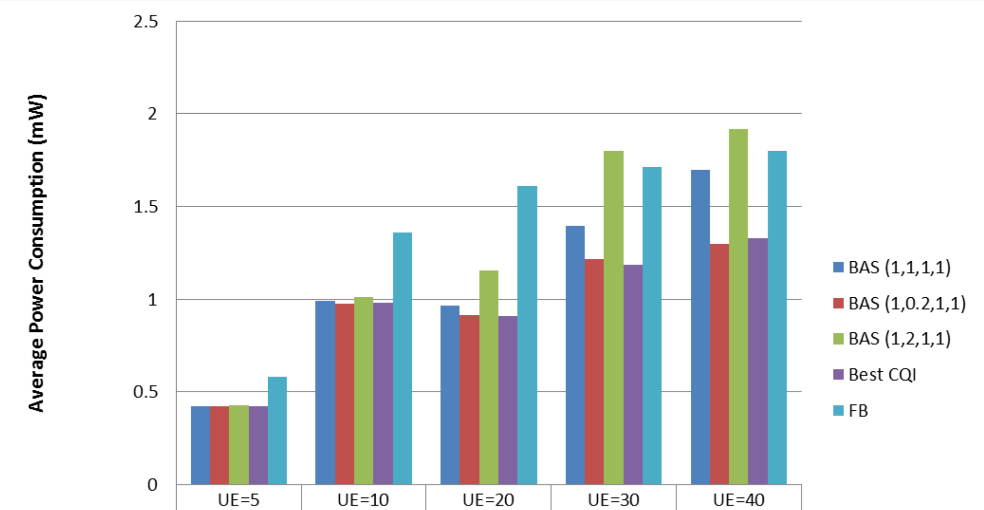
First, the 1080P resolutions video traffic simulation case was designed to examine the performances by varying the number of UE. The changes in the parameters of the proposed weighting functions were examined and the proposed BAS scheme was compared with two comparison schemes, named best CQI scheme and fair bandwidth (FB) scheme. The BAS scheme with three weighting-function parameters and the two comparison schemes are provided as follows:

- The best CQI scheme: the UE with the good channel condition always receives the highest scheduling priority
- FB scheme: eNB treat all UE in fair manner and maintain bandwidth balance between UEs
- BAS (1,1,1,1): the proposed weighting function with n = 1, α_l =1, λ_l =1, α_h =1, and λ_h =1
- BAS (1,0.2,1,1): the proposed weighting function with n = 1, α_l = 1, λ_l = 0.2, α_h =1 , and λ_h =1
- BAS (1,2,1,1): the proposed weighting function with n = 1, α_l =1, λ_l = 2, α_h =1 , and λ_h =1

Figs. 8, 9, 10, and 11 illustrate the average power consumed when 1 KB of data is transmitted, average signaling messages generated per UE in one second, average streaming delivery time and average system throughput per UE, respectively. In **Fig. 8**, the average power consumptions are compared by setting distinct the number of UE. To figure out the power utilization, the average power consumption, P_{avg}, is calculated by equation (5).

$$P_{avg} = \sum_{i=1}^{N}(P_{i\_Idle} + P_{i\_Connected})\,/N \qquad (5)$$

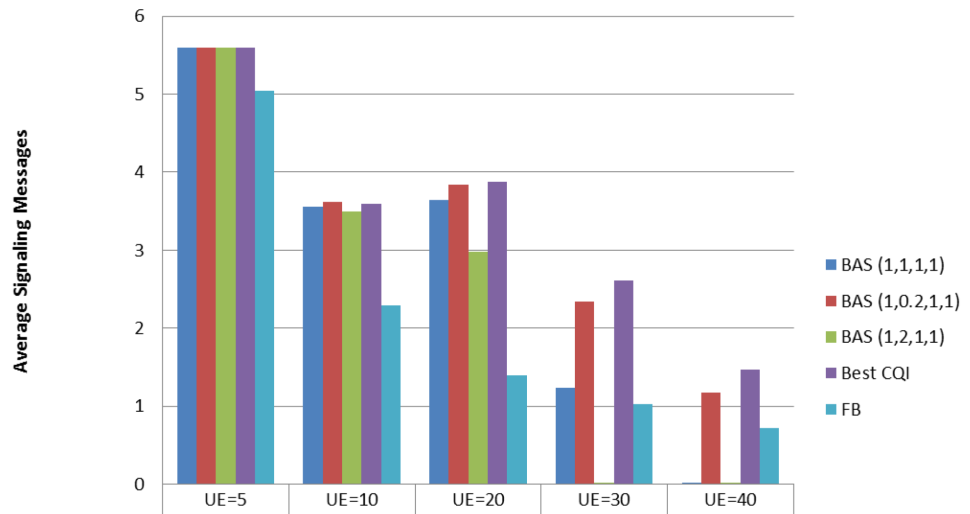
where P_{i_idle}, $P_{i\_connected}$ and N are the power consumption of UE in RRC_Idle state, power consumption of UE in RRC_Connected state, according to **Fig. 3**, and the number of UE, respectively. As the proposed BAS scheme tends to increase the scheduling priority to deliver packet to UE when the buffer of UE is going to full by using the factor $X_{h(i)}$, the power can be utilized more effectively. When the $X_{h(i)} > X_{l(i)}$ (i.e., $\lambda_h > \lambda_l$), the BAS scheme prefer UE to gain higher priority to be assigned with RBG for scheduling. Therefore, UE can transition to RRC_Idle state as soon as possible when its buffer is full and then its staying time in the idle state can also be lengthened and achieve power saving. The results indicate that when the number of UE is getting larger, the BAS (1,1,1,1) and BAS (1,0.2,1,1) can save more power than the FB schemes. By contrast, the BAS (1,2,1,1) tends to increase the scheduling priority to deliver packet to UE when the buffer of UE is going to be empty, therefore, power is used inefficiently, especially when the number of UE is more than 30. As mentioned above in equation 5, setting the higher value of λ_h leads UE to easily transition to RRC_Idle state for power saving when its buffer is almost full. And setting the higher value of λ_l makes UE to increase the scheduling priority when its buffer is almost empty, though UE is still in RRC_Connected state. The simulation results illustrate that the BAS (1,0.2,1,1) scheme, which has higher values of λ_h and lower value of λ_l, consumes the least power among the proposed BAS schemes. In addition, as the number of connected UE is restricted, the probability of UE being in connected state can get more bandwidth in the RRC_Connected state so that the possibility of UE to stay in the RRC_Connected state without data reception can be minimized.



	UE=5	UE=10	UE=20	UE=30	UE=40
BAS (1,1,1,1)	0.424311739	0.990499188	0.966918686	1.395822611	1.698072783
BAS (1,0.2,1,1)	0.422344212	0.976469409	0.911200603	1.213435154	1.297013523
BAS (1,2,1,1)	0.426128958	1.0108272	1.15407447	1.797954868	1.91857422
Best CQI	0.422716091	0.978802336	0.909599595	1.186324024	1.330354295
FB	0.579550655	1.358122527	1.608646227	1.712784208	1.798717289

Fig. 8. Comparisons of average power consumption

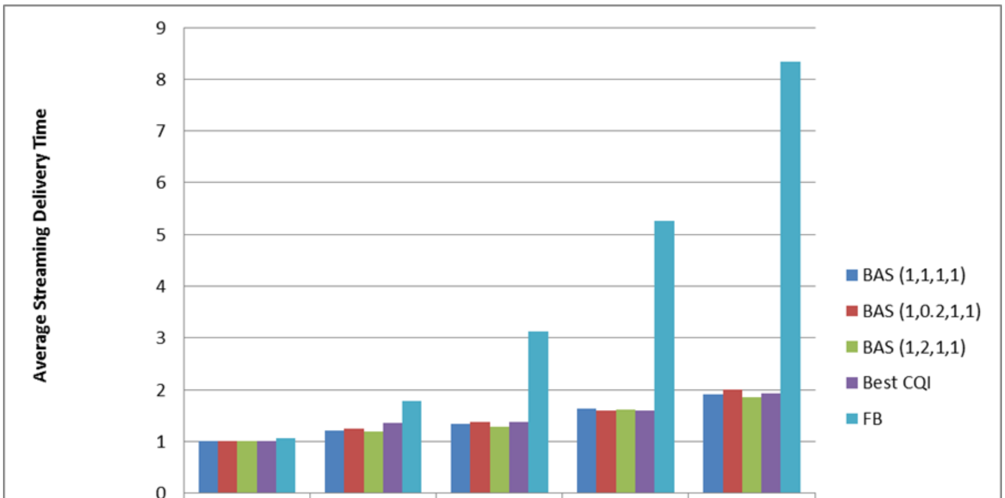
As shown in **Fig. 9**, the volumes of average signaling messages are compared. The results show that the BAS (1,0.2,1,1) has almost the same signaling messages as best CQI scheme and higher average signaling messages than that of the other proposed BAS schemes and FB scheme when the number of UEs exceeds 20. The main reason is that the BAS (1,0.2,1,1) encourages eNB to fill the buffer of UE as soon as possible and easier and then let UE transition to RRC_Idle state. After a period of time, the UE needs to transition from RRC_Idle state to RRC_Connected state to download video packet. As the result of UEs transition RRC state frequently, it will generate more signaling messages, however, save more power. The other proposed BAS schemes and FB scheme tend to allow connected UEs to contend limited radio resource in the RRC_Connected for video download, and, therefore, generate less signaling messages.



	UE=5	UE=10	UE=20	UE=30	UE=40
BAS (1,1,1,1)	5.596159616	3.559355936	3.641864186	1.233790046	0.01080108
BAS (1,0.2,1,1)	5.596159616	3.612561256	3.842684268	2.341900857	1.172817282
BAS (1,2,1,1)	5.598159816	3.490749075	2.98129813	0.010467713	0.012051205
Best CQI	5.600160016	3.593959396	3.871487149	2.613928059	1.461846185
FB	5.04370437	2.286428643	1.389238924	1.028036137	0.720672067

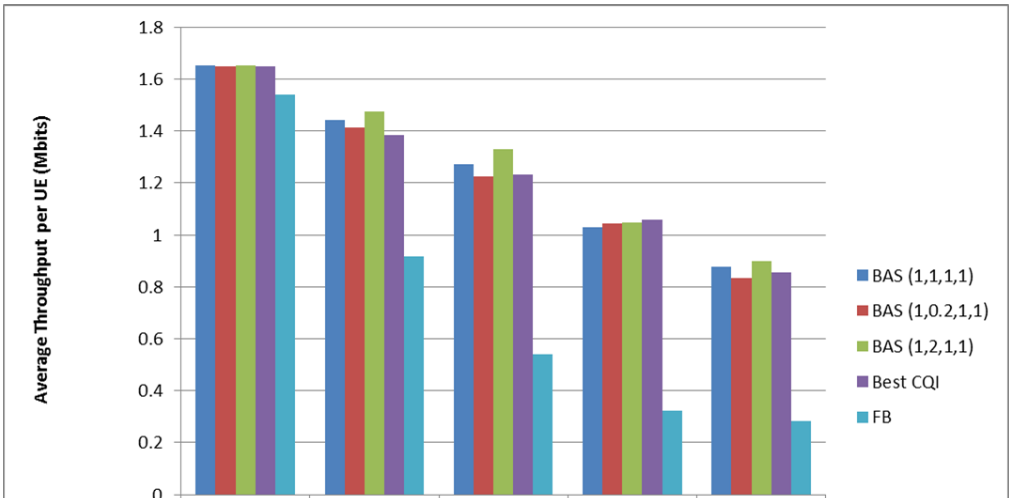
Fig. 9. Comparisons of average signaling message

The proposed BAS schemes illustrate much shorter average streaming delivery time than that of the FB scheme in **Fig. 10**. The main reason is that the proposed BAS scheme cares the buffer length of UE and applies the factor $X_{l(i)}$ to increase the scheduling possibility for UE when its local buffer is going to be exhausted. Consequently, the UE can be allocated with radio resource to download video data and achieve less streaming delivery time. **Fig. 11** shows the comparison of average throughput per UE. The simulation results indicate that the proposed BAS schemes achieve higher system throughputs than that of the FB scheme and have almost the same throughputs as the best CQI scheme. Because the FB scheme only considers the fairness to achieve balance throughput at every scheduling moment, however, the UE, which needs to be allocated with radio resource for fairness at a specific scheduling instant, may be in poor channel condition, and, therefore, sacrifice the spectrum efficiency. In the proposed BAS scheme, the designed weighting function considers the spectrum efficiency; therefore, it can schedule the radio resource to UEs at proper time moment.



	UE=5	UE=10	UE=20	UE=30	UE=40
BAS (1,1,1,1)	1	1.216545012	1.335113485	1.639344262	1.908396947
BAS (1,0.2,1,1)	1	1.25	1.370863014	1.594501587	2
BAS (1,2,1,1)	1	1.19047619	1.282051282	1.612903226	1.851851852
Best CQI	1	1.351351351	1.369863014	1.587301587	1.923076923
FB	1.07	1.785714286	3.125	5.263157895	8.333333333

Fig. 10. Comparisons of streaming delivery time



	UE=5	UE=10	UE=20	UE=30	UE=40
BAS (1,1,1,1)	1.6503177	1.44101262	1.271138085	1.02947997	0.876128258
BAS (1,0.2,1,1)	1.6493739	1.41282366	1.226189595	1.04357695	0.83385765
BAS (1,2,1,1)	1.65032496	1.47342621	1.328435385	1.04596936	0.89916942
Best CQI	1.64772414	1.38450645	1.233599025	1.05738353	0.855699585
FB	1.53814914	0.91643337	0.538923795	0.32111173	0.283692127

Fig. 11. Comparisons of system throughput per UE

Table 2 provides the normalized performance. The results show that the BAS (1,0.2,1,1) scheme and the best CQI scheme have almost the same performance. In order to further examine the performances of both schemes under heavy loading, the video source with 2160P resolution were applied for the same number of UEs and their comparisons are shown in **Table 3**.

Table 2 Normalization performance compare for 1080P video with 40UEs

1080P Video / UE=40	**best CQI scheme**	**BAS (1,0.2,1,1) scheme**	**Normalized to the best CQI scheme**
			BAS (1,0.2,1,1)
Avg. Power Consumption	1.330354	1.297014	0.975
Avg. Signaling message	1.4617	1.1727	0.802
Avg. Streaming delivery time	1.923076923	2	1.04
Avg. Throughput per UE	0.8557	0.8334	0.974

It is hard to develop a perfect schedule algorithm to satisfy the above performance indexes, different indexes shall have different priorities for satisfaction during tradeoff. For example, the scheduler shall not save power at the expense of sacrificing the service quality. The results show that the proposed BAS scheme demonstrates much better performance than the best CQI scheme in signaling overhead while still achieve almost the same power consumption, streaming delivery time, and throughput as the best CQI scheme. The main reason is that UE consumes much backlog to play 2160P video frames than that of 1080P and every UE needs more bandwidth to download video data from eNB. The proposed BAS scheme not only promotes UE, whose buffer is going to exhaust and has better channel condition, to higher scheduling priority to download data, but also restricts the number of UE to transition to RRC_Connected state when there is no sufficient radio resource to decrease the signaling overhead. The above restriction policy can effectively avoid UE contending limited radio resource. Therefore, UE needs not to frequently switch between states and the buffer underflow condition can be minimized as well.

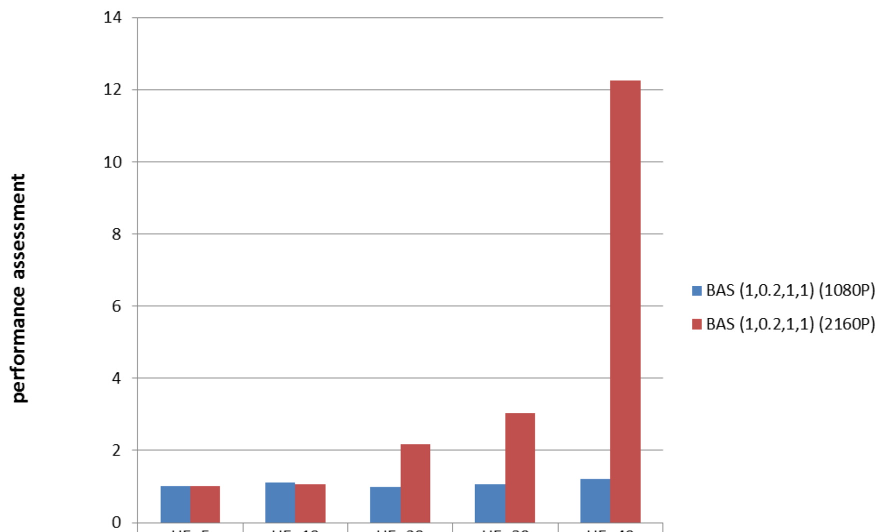
Table 3. Normalization performance compare for 2160P video with 40UEs

2160P Video / UE=40	**Best CQI scheme**	**BAS (1,0.2,1,1) scheme**	**Normalized to the best CQI scheme**
			BAS (1,0.2,1,1)
Avg. Power Consumption	1.242506	1.180451	0.95
Avg. Signaling message	0.0298	0.0083	0.279
Avg. Streaming delivery time	15.55936	8.414453	0.541
Avg. Throughput per UE	0.883172	0.916337	1.038

As power usage, signaling overhead, streaming delivery time, and throughput are tradeoff and correlated to each other, a fair performance assessment P_A is designed in equation (6) to compare both schemes.

$$P_A = T_U/(P_C S_M S_T) \tag{6}$$

where T_U is average throughput per UE, P_C is average power consuming, S_M is average signaling message, and S_T is average streaming delivery time. It is noted that the results of the proposed BAS scheme are normalized to the best CQI scheme; P_A of the best CQI scheme is 1. The performances of the proposed BAS scheme for both 1080P and 2160P resolution sources are illustrated in **Fig. 12**. The results indicate that when the UE number is larger than 20, the proposed scheme achieves much better performance than the best CQI scheme in applying the 2160P resolution video traffic.



	UE=5	UE=10	UE=20	UE=30	UE=40
BAS (1,0.2,1,1) (1080P)	1.00250047	1.099832249	0.998270343	1.062675915	1.211032314
BAS (1,0.2,1,1) (2160P)	1.004122912	1.063730948	2.172178712	3.030748799	12.24245796

Fig. 12. Comparisons of performance assessment for 1080P and 2160P video

5. Conclusion

The power saving and throughput are critical but complex issues when the spectrum efficiency and QoS are considered for service deployment. In this paper, the BAS scheme is proposed for video streaming service to examine the power-saving efficiency, signaling overhead, streaming delivery time and throughput. The proposed BAS scheme determines the scheduling priority based on the channel condition and the buffer of UE to adapt the scheduling preference by the designed weighting function. The designed weighting function considers the state transitions of power consumption and applies the burst scheduling concept to aggregate packets in either forward or backward manner so as to reduce the state transition frequency and then to minimize power consumption. Furthermore, the proposed BAS scheme restricts the number of UE in RRC_Connected state to avoid the UE wasting power when there is no sufficient radio resource. Thus the proposed BAS scheme can flexibly schedule radio resource to meet the critical performance, such as the streaming delivery time, through the heuristic weighting function in crucial resource condition. The simulation results show that the proposed scheme has better performance than FB scheme and achieves almost the similar performance as the best CQI scheme in light traffic load. When in heavy traffic load, the proposed BAS scheme can effectively reduce the signaling overhead and streaming delivery time when compared to the best CQI scheme. Due to delay-constrained video applications, the tradeoff between the power consumption and latency is a critical issue for service deployment [22]. The bandwidth requirements, user's experience, and signaling overhead are different from service to service. The scheduler shall be designed to be more service aware to satisfy the demands of emerging mobile services and this is one of our ongoing research issues.

References

[1] Ericsson Mobility Report , Nov.2014
http://www.ericsson.com/res/docs/2014/ericsson-mobility-report-november-2014.pdf

[2] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018 http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html

[3] Video Boom Forces Verizon to Upgrade Network, The Wall Street Journal http://online.wsj.com/news/articles/SB10001424052702304173704579260354018157042

[4] 3GPP TS 36.300 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Overall description Stage 2

[5] 3GPP TS 36.321 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA) Medium Access Control (MAC) protocol specification.

[6] J. Huang, F. Qian, and A. Gerber, "A Close Examination of Performance and Power Characteristics of 4G LTE Networks," in MobiSys, 2012.

[7] The Evolution of the Signaling Challenge in 3G & 4G http://downloads.lightreading.com/wplib/cisco/HR_Cisco_LTE_Signaling_WP_Final.pdf, June 2012

[8] Gupta, M. ; Jha, S.C., Koc, A.T., Vannithamby, R , " Energy impact of emerging mobile internet applications on LTE networks: issues and solutions," *IEEE Communications Magazine*, Vol. 51, Issue 2, pp. 90-97, February 2013. Article (CrossRef Link)

[9] 3GPP TS 36.211 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical Channels and Modulation

[10] 3GPP TS 36.213 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical layer procedures

[11] 3GPP TS 36.101, V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA), User Equipment (UE) radio transmission and reception

[12] 3GPP TS 36.331 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC) Protocol specification

[13] 3GPP TS 36.413 V12.0.0, Evolved Universal Terrestrial Radio Access (E-UTRA), S1 Application Protocol (S1AP)

[14] Kalle, R.K., Nandan, A.K., and Das, D., "La VoLTE: novel cross Layer optimized mechanism of Video transmission over LTE for DRX," in *Proc. of Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, pp 1-5, May. 2012. Article (CrossRef Link)

[15] Piro, G., Grieco, L.A., Boggia, G., Fortuna, R. and Camarda, P., "Two-Level Downlink Scheduling for Real-Time Multimedia Services in LTE Networks," *IEEE Transactions on Multimedia*, Vol. 13, Issue 5, pp. 1052 – 1065, Oct. 2011. Article (CrossRef Link)

[16] Ramli, H.A.M., Sandrasegaran, K., Basukala, R., Patachaianand, R., Minjie Xue and Cheng-Chung Lin, "Resource Allocation Technique for Video Streaming Applications in the LTE System," in *Proc. of Wireless and Optical Communications Conference (WOCC), 2010 19th Annual*, pp 1-5, May, 2010. Article (CrossRef Link)

[17] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, " Quality-Driven Cross-Layer Optimized Video Delivery over LTE," IEEE Communications Magazine, Vol. 48 , Issue 2, pp. 102-109, Feb. 2010. Article (CrossRef Link)

[18] Singh, S., Oyman, O., Papathanassiou, A., Chatterjee, D. and Andrews, J.G., "Video Capacity and QoE Enhancements over LTE", , 2012 IEEE International Conference on Communications (ICC), pp. 7071 - 7076,Jun. 2012. Article (CrossRef Link)

[19] J. Wu, B. Cheng, C. Yuen, Y. Shang, J. Chen, "Distortion-Aware Concurrent Multipath Transfer for Mobile Video Streaming in Heterogeneous Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 4, pp. 688-701, 2015. Article (CrossRef Link)

[20] J. Wu, C. Yuen, N.-M. Cheung, J. Chen, "Delay-Constrained High Definition Video Transmission in Heterogeneous Wireless Networks with Multi-homed Terminals," *IEEE Transactions on Mobile Computing*, Apr. 2015 Article (CrossRef Link)

[21] J. Wu, C. Yuen, M. Wang, J. Chen, "Content-Aware Concurrent Multipath Transfer for High-Definition Video Streaming over Heterogeneous Wireless Networks," *IEEE Transactions on Parallel and Distributed Systems*, March, 2015.Article (CrossRef Link)

[22] J. Wu, B. Cheng, C. Yuen, N.-M. Cheung, J. Chen, "Trading Delay for Distortion in One-Way Video Communication over the Internet," *IEEE Transactions on Circuits and Systems for Video Technology*, March, 2015. Article (CrossRef Link)

Meng-Hsien Lin received the Ph.D. degree in Communication Engineering from National Central University in 2015. His research interests include broadband mobile networks, social network services and cloud computing architecture. He is currently with Chunghwa Telecommunication Laboratories, Taiwan.

Yen-Wen Chen received the Ph.D. degree in Electronic Engineering from National Taiwan University of Science and Technology (NTUST) in 1997. During 1983 to 1998, he worked at Chunghua Telecommunication Laboratories, Taiwan and was a project manager of the broadband switching systems. From August 1998 to July 2000, he joined the Department of Information Management, Central Police University. Since August 2000, Dr. Chen has joined the Department of Communication Engineering of National Central University. Currently, he is a professor. His research interests include broadband mobile networks, QoS management, sensor networks, network applications, and multimedia networks. Dr. Chen is a member of the IEEE communication society.