Two-Phase Multicast DRX Scheduling for 3GPP LTE-Advanced Networks

Jia-Ming Liang, Jen-Jee Chen, *Member, IEEE*, Po-Chun Hsieh, and Yu-Chee Tseng, *Fellow, IEEE*

Abstract—For next-generation wireless communications, the *3GPP Long Term Evolution-Advanced (LTE-A)* is the most promising technology which provides transmission rate up to 1 Gbps and supports various broadband multimedia services, such as IPTV and Voice/Video-over-IP services. To reduce the energy consumption of *user equipments (UEs)*, the LTE-A standard defines the *Discontinuous Reception Mechanism (DRX)* to allow UEs to turn off their radio interfaces and go to sleep when no data needs to be received. However, how to optimally configure DRX for UEs is still left as an open issue. In this paper, we address the DRX optimization problem for multicast services. This problem asks how to guarantee the *quality of service (QoS)* of the multicast streams under the *Evolved Node B (eNB)* while minimizing the UEs' wake-up time. We prove this problem to be NP-complete and propose an energy-efficient heuristic. This heuristic consists of two phases. The first phase tries to aggregate the required bandwidth of the multicast streams for UEs to reduce their wake-up periods. The second phase further minimizes UEs' unnecessary wake-up periods by optimizing their DRX configurations. Extensive simulation results show that our scheduling is close to the optimum in most cases.

Index Terms—Discontinuous reception mechanism (DRX), energy saving, long term evolution-advanced (LTE-A), multimedia broadcast multicast service (MBMS), quality of service (QoS)

1 Introduction

The *3rd Generation Partnership Project (3GPP)* develops *Long Term Evolution-Advanced (LTE-A)* [2] for next-generation wireless communications, which provides transmission rates up to 1 Gbps for low-mobility *UEs (user equipments)* and 100 Mbps for high-mobility UEs. This enables a lot of broadband multicast/broadcast services such as *Live Mobile IPTV*, *Multiparty Video/Voice Conferencing*, *Interactive Group Gaming*, *Telmatics Services*, *Stock Quote Streaming* [3], [4], [5]. Because such services are with one-to-many or many-to-many communication features [6], LTE-A exploits *Multimedia Broadcast Multicast Service (MBMS)* to enhance their transmission efficiency. Specifically, when multiple UEs request the same video stream, the base station (called *Evolved Node B*, *eNB*) only needs to transmit one copy of the video data to the multicast group to save bandwidth.

On the other hand, as mobile devices become more and more powerful, the *multi-tasking* concept [7] is realized such that multiple media streams can be subscribed and displayed simultaneously on UEs. For example, the *Multi-Window* mode on Samsung Galaxy Note 4 [8], *QSlide feature* on LG Optimus G [9], *Quasar* for Apple iPad [10], and *Multi-View* feature of *Stick it!* (e.g., Pop-up play) app on Android devices [11]. With these features, users can run multiple apps at the same time to improve the life quality, such as watching Live IPTV (e.g., the CNN news, MLB baseball, NFL football, and NBA basketball) when playing interactive group games or joining multiparty conference calls [12]. However, streaming services on mobile devices are usually constrained by the limited battery capacity [13]. To address the issue, LTE-A defines the *Discontinuous Reception Mechanism (DRX)* [14], [15] to allow UEs to turn off their radio interfaces and go to sleep when no data needs to be delivered to them. That means once DRX is enabled, UEs will wake up and sleep in a periodical manner to save their energy. However, how to optimize the DRX configurations from a system perspective is still left as an open issue.

In this paper, we address the DRX optimization problem by asking how to satisfy the *quality of service (QoS)* of multicast streams while minimizing UEs' wake-up periods following the DRX specification. We formulate this problem and show it to be NP-complete. We then design an energy-efficient heuristic, which includes two phases. In the first phase, a *minimal cost first* strategy is adopted to reduce the necessary wake-up periods for UEs. In the second phase, we optimize the DRX parameters to further decrease UEs' wake-up periods. Extensive simulation results show that the proposed scheme can significantly decrease the wake-up period of UEs.

The rest of this paper is organized as follows. Related work is discussed in Section 2. Preliminaries are given in Section 3. Section 4 presents our scheme. Simulation results are shown in Section 5. Section 6 concludes the paper and gives some future directions.

- *J.-M. Liang is with the Department of Computer Science and Information Engineering, Chang Gung University, Kweishan, Taoyuan 33302, Taiwan. E-mail: jmliang@mail.cgu.edu.tw.*
- *P.-C. Hsieh and Y.-C. Tseng are with the Department of Computer Science, National Chiao-Tung University, Hsin-Chu 30010, Taiwan. E-mail: {pchsieh, yctseng}@cs.nctu.edu.tw.*
- *J.-J. Chen is with the Department of Electrical Engineering, National University of Tainan, Tainan 70005, Taiwan. E-mail: jjchen@mail.nutn.edu.tw.*

Manuscript received 16 June 2014; revised 15 Mar. 2015; accepted 3 Sept. 2015. Date of publication 18 Sept. 2015; date of current version 1 June 2016. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TMC.2015.2480090

2 Related Work

In the literature, [16], [17], [18] focus on maximizing the system utility under the consideration of multicast services. However, they neglect the power consumption of

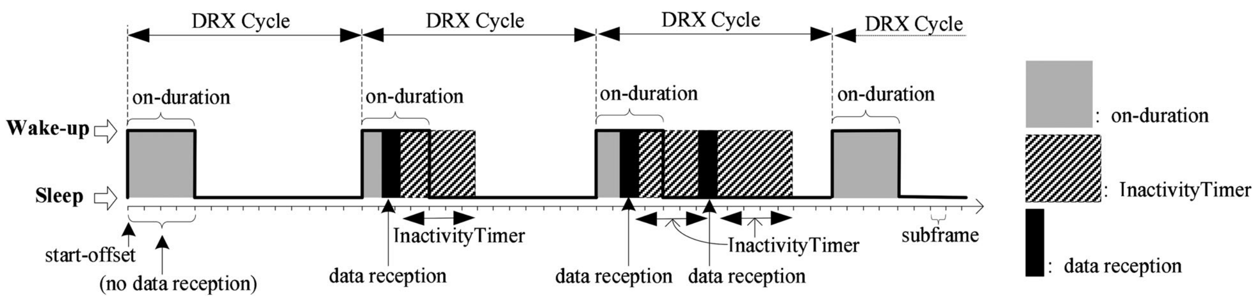

Fig. 1. An example of the DRX mechanism.

UEs. Reference [19] considers the allocation of *OFDM* symbols and tries to minimize the symbols allocated to UEs to save their power. The study [20] investigates the energy consumption incurred by UEs' state transitions and tries to minimize the number of state transitions to reduce unnecessary power consumption. Recently, several studies [21], [22], [23], [24], [25], [26] have considered the DRX mechanism. Although they concluded that enabling DRX can significantly reduce UEs' power consumption, they do not show how to optimize the DRX configurations. Reference [27] first proposes to adjust UEs' DRX configurations, but this scheme requires considerable control signaling for the adjustment. In [28], a dynamic DRX scheme is proposed to determine UEs' sleep cycle based on their traffic loads. In [29], an adaptive DRX is proposed to adjust UEs' extended wake-up periods based on their channel qualities to improve system utility. These studies all consider unicast. Up to now, the DRX optimization problem for multicast with QoS consideration is still an open issue to be studied.

3 Preliminaries

In this section, we first introduce the DRX mechanism. Then, we will define our problem and prove the problem to be NP-complete.

3.1 Discontinuous Reception Mechanism (DRX)

In LTE-A, the DRX mechanism is defined in the *Radio Resource Control (RRC)* [30]. An example is shown in Fig. 1. The basic unit of wake-up/sleep periods is subframe (with the length of 1 ms). The DRX mechanism supports short and long cycles for real-time and non-real-time streams, respectively. Since multicast streams are real-time streams, we focus on short cycles only. In the rest of this paper, "cycle" means "short cycle" unless otherwise stated. When DRX enables, four parameters need to be defined for each UE: 1) *DRX-cycle-length*, 2) *DRX-start-offset*, 3) *on-duration*, and 4) *InactivityTimer*. The *DRX-cycle-length* defines the basic period of the UE to receive the arriving stream data from the eNB. It should be shorter than the tolerable delay of the stream for QoS purposes. The *DRX-start-offset* defines the beginning subframe of a cycle. At the beginning of each cycle, the UE has to wake up during *on-duration* subframes. In each wake-up subframe, the UE will monitor its incoming data from the eNB. If any of its data is received, the UE will start an *InactivityTimer* and stay awake before the timer expires. Before *InactivityTimer* expires, if the UE detects any of its data arriving, it resets InactivityTimer to its original value. This is repeated until the UE's InactivityTimer expires. Then, the UE goes to sleep and turns off its radio interface to save energy. During the UE's sleep period, all the data for the UE will be buffered at the eNB until the next on-duration arrives.

3.2 DRX and Multicast

In this paper, we investigate how to adopt DRX for multicast streams. Considering the DRX configuration, we define two metrics: *internal* cost and *external* cost, to evaluate the wake-up costs incurred by the allocation order of multicast groups and the DRX configuration. An internal cost occurs when a UE has to stay awake to wait for the next service (i.e., the next multicast group) that it subscribes to. We give an example in Fig. 2. In this example, there are five UEs ($UE_1 \sim UE_5$) and six multicast groups ($G_1 \sim G_6$) in the network. The groups $G_1 \sim G_3$ will be allocated data every 20 subframes and the groups $G_4 \sim G_6$ will be allocated data every 40 subframes. That means that 20 and 40 subframes are the *group-cycle-length* of $G_1 \sim G_3$ and $G_4 \sim G_6$, respectively. In addition, the eNB will allocate 2, 3, 4, 2, 2, 4 subframes for G_1, G_2, G_3, G_4, G_5, G_6 every cycle, to satisfy their data rates. Each UE can join more than one multicast groups (e.g., UE_3 joins G_1 and G_3). We illustrate two allocation orders of these groups in the following. Let the *DRX-cycle-length* of UE_3 be 20 subframes. Fig. 2a shows a poor allocation order for UE_3 because it has to stay awake for three extra subframes per cycle. These extra subframes are called internal cost. Contrarily, Fig. 2b shows a better allocation order because there is no internal cost for UE_3. This example indicates that the allocation order of multicast groups matters.

On the other hand, the external cost is the idle waiting time even after a UE has received the data of its last multicast group during a wake-up period. With the same allocation order from Fig. 2b, Fig. 3 shows three different DRX configurations for UE_2 who subscribes to G_1, G_5, and G_6. In the first case, UE_2 sets on-duration $= 2$ and InactivityTimer $= 4$. There is always an external cost of four subframes at the end of each wake-up period. In the second case, UE_2 sets on-duration $= 5$ and InactivityTimer $= 1$. The external cost is 4 in the first wake-up period and 1 in the second period. In the third case, UE_2 sets on-duration $= 11$ and InactivityTimer $= 0$. The external cost is 9 in the first period and 0 in the second period. These examples indicate that the allocation order of multicast groups and the DRX configurations significantly affect UEs' wake-up costs. This strongly motivates us to study the DRX problem.

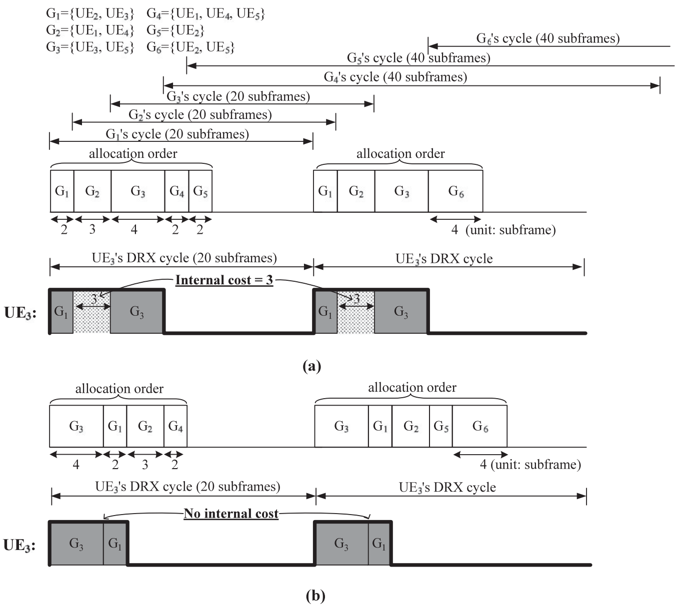

Fig. 2. Effects of multicast allocation orders on the internal cost.

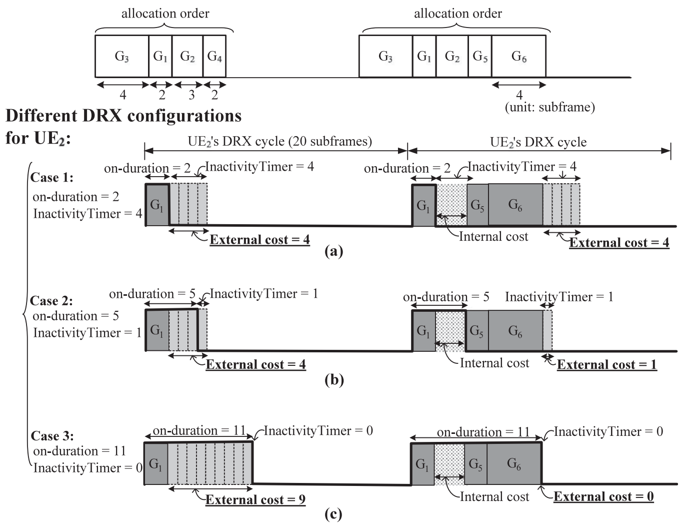

Fig. 3. Effects of on-duration and InactivityTimer on the external cost.

3.3 The DRX Optimization Problem

We consider an LTE-A network with an eNB serving N UEs. Assume that M multicast[1] streams are available for UEs to request in the network, which means that each $\mathrm{UE}_i, i = 1..N$, can request at most M multicast streams. For simplicity, we use $\mathrm{G}_j, j = 1..M$, to represent the multicast group which collects the UEs that subscribe to multicast stream j. To maintain the QoS, each multicast stream $j, j = 1..M$, is with a delay constraint D_j (subframes) and a data arrival rate R_j (bits/subframe). For each $\mathrm{UE}_i, i = 1..N$, the transmission rate from the eNB to UE_i is C_i (bits/subframe), which depends on the *modulation and coding scheme (MCS)*. The *multicast DRX scheduling (MDS) problem* asks how to schedule M multicast groups (by determining the group allocation order in terms of the group-cycle-length $\hat{L}_j$ and group-start-offset $\hat{S}_j$[2] and the required number of subframe $T_j, j = 1..M$) and well configure UEs' DRX parameters (including each UE_i's DRX-cycle-length L_i, DRX-start-offset S_i, on-duration O_i, and InactivityTimer I_i) such that the delay constraint D_j of each stream j is not violated, the perceived data arrival rate by UE_i under its DRX configuration is R_j, and the total number of wake-up subframes of all UEs is minimal.

Theorem 1. *The addressed problem is NP-complete.*

Proof. To simplify the proof, we consider a special case where the delay constraints of M multicast groups are identical, i.e., $D_j = D_1, j = 2..M$. Thus, the best group-cycle-length $\hat{L}_j$ of stream $j, j = 1..M$ will be the same as their delay constraint D_j, i.e., $\hat{L}_j = D_j, j = 1..M$ and the best InactivityTimers I_i of each UE_i will be 0. In addition, consider that each multicast group $\mathrm{G}_j, j = 1..M$, requests $T_j = 1$ subframe to serve its arrival data in a cycle to meet its QoS. We formulate the addressed problem as a decision problem, called *Multicast DRX scheduling decision (MDSD) problem*, which is stated as follows. Given N UEs and M multicast groups, where each $\mathrm{UE}_i, i = 1..N$ can subscribe to one or more than one multicast groups from $\mathrm{G}_1..\mathrm{G}_M$. Let $X_i = \{G_j | UE_i \in G_j\}$ be the multicast groups that UE_i subscribes to. We ask whether or not there exists a multicast group allocation order $\hat{A} = [A_1, A_2, .., A_M]$, which $A_j, j = 1..M$ is G_j's index in the allocation order, such that the total amount of wake-up periods of all UEs subscribing to them is less than or equal to λ, i.e., $\sum_{i=1..N} O_i \leq \lambda$, where $O_i = \max_{\forall G_j, G_{j'} \in X_i}(|A_j - A_{j'}| + 1)$. Next, we show that the MDSD problem is NP-complete.

We first show that MDSD problem belongs to NP. Given a problem instance and a solution containing the allocation order of M multicast groups $\mathrm{G}_1..\mathrm{G}_M$, this solution can be verified its validity in polynomial time. Thus, MDSD problem belongs to NP. We then reduce the *Hyper-graph Optimal Linear Arrangement (HOLA)* problem [32], [33], which is known to be NP-complete, to the MDSD problem. Consider a *hyper-graph* $\hat{H} = (V, E)$, where $V = \{v_i, i = 1..N\}$ and $E = \{e_j, j = 1..M\}$ are the vertex set and the hyper-edge set, respectively, where a hyper-edge e_i is a subset of V that can form a partial maximal complete graph. That means that each vertex v_j in e_i connects with each other. The HOLA problem asks whether or not we can find the function $\sigma(\cdot)$ to label each vertex with a distinct integer from 1 to M, i.e., $\sigma(\cdot) : V \to \{1, .., M\}$, such that the total lengths of all hyper-edges, denoted by $C_{\hat{H},\sigma}$, is less than or equal to $\lambda - N$, i.e., $C_{\hat{H},\sigma} \leq \lambda - N$, where $C_{\hat{H},\sigma}$ is defined as

$$C_{\hat{H},\sigma} = \sum_{\forall e_i \in E} \max_{\forall v_j, v_{j'} \in e_i} |\sigma(v_j) - \sigma(v_{j'})|.$$

Note that $\max_{\forall v_j, v_{j'} \in e_i} |\sigma(v_j) - \sigma(v_{j'})|$ is the length of the hyper-edge e_i. We show that the HOLA problem has a solution *if and only if* the MDSD problem has a solution.

1. In LTE-A, the multicast employs UDP and transmits data in a feedback-less manner [31].
2. The group-start-offset is defined as the beginning subframe when allocating the data to stream j.

Suppose that we have a solution to the MDSD problem, which is an allocation order of the multicast groups, i.e., $\hat{A} = [A_1, A_2, A_3, .., A_M]$, where $A_j, j = 1..M$, is group G_j's index in the allocation order. Next, we regard the multicast group G_j and UE_i's multicast groups X_i as the vertex v_j and hyper-edge e_i in $\hat{H}$, respectively. Then, the allocation order $[A_1, A_2, A_3, .., A_M]$ for the multicast groups can constitute a one-to-one labeling function $\sigma(\cdot)$ such that the total lengths of hyper-edges is $C_{\hat{H},\sigma} \leq \lambda - N$. This proves the *if* part.

Conversely, suppose that we have a solution to the HOLA problem, which is a one-to-one labeling function $\sigma(\cdot) : V \rightarrow \{1, .., M\}$ such that the total length of all hyper-edges in E is $C_{\hat{H},\sigma} \leq \lambda - N$. Then, we assign the multicast group G_j with an allocation order A_j equal to $\sigma(v_j)$. In this way, the overall wake-up periods of all UEs in a cycle is less than or equal to λ. This constitutes a solution to the MDSD problem. This proves the *only if* part. □

4 The Proposed Scheme

Since the MDSD problem is NP-complete, finding the optimal solution is impractical due to the time complexity. Thus, we propose an energy-efficient heuristic to tackle the problem. The heuristic consists of two phases. The first phase exploits the "*minimal cost first*" strategy to reduce the overall internal costs of UEs incurred by the allocation order of multicast groups. The second phase is to further optimize DRX configurations of UEs including InactivityTimers and on-durations to reduce their overall external costs. The detail of the scheme is described as follows.

4.1 Phase 1 - Minimal Cost First

The goal of this phase is to determine the allocation order of multicast groups in terms of the group-cycle-length and group-start-offset such that the overall internal costs of UEs can be minimized while the data rate and delay constraint of each multicast stream can be guaranteed. In the following, we first select the group-cycle-length and group-start-offset for each multicast group and then determine each UE's DRX-cycle-length and DRX-start-offset based on these results. The detailed steps are depicted as follows.

Step 1: Sorting all multicast groups $G_j, j = 1..M$, by their delay constraints in an increasing order. Without loss generality, let $D_1 \leq D_2 \leq ... \leq D_M$.

Step 2: Let $\hat{L}_j$ be the group-cycle-length of $G_j, j = 1..M$. Assume that $1 \leq \hat{L}_1 \leq D_1$ is known, we determine $\hat{L}_j$ for multicast group $G_j, j = 2..M$ as follows.

$$\hat{L}_j = \left\lfloor \frac{D_j}{\hat{L}_{j-1}} \right\rfloor \times \hat{L}_{j-1}. \tag{1}$$

By Eq. (1), the delay constraint of stream j can be guaranteed because $\hat{L}_j \leq D_j, j = 1..M$. Also, Eq. (1) makes the group-cycle-length of G_j be integer multiple of that of $G_{j'}$ when $j' < j$ (i.e., $\frac{\hat{L}_j}{\hat{L}_{j'}} \in \mathbb{Z}$ if $j' < j$). This design can avoid two multicast groups colliding with each other. Finally, Eq. (1) implies that $\hat{L}_1$ is the basic group-cycle-length in the system and the allocation results will repeat every $\hat{L}_M/\hat{L}_1$ basic cycles. Note that $\hat{L}_1$ will be specified later on.

Step 3: Let $T_j, j = 1..M$, be the number of subframes required to serve stream j's arrival data during group-cycle-length $\hat{L}_j$, such that

$$T_j = \left\lceil \frac{R_j \times \hat{L}_j}{\min\{C_i | UE_i \in G_j\}} \right\rceil. \tag{2}$$

In Eq. (2), the numerator part is the amount of arrival data (in bits) of stream j during cycle $\hat{L}_j$ and the denominator part is the available transmission rate (bits/subframe) for group G_j, which is restricted by the lowest transmission rate of the UEs in G_j. This ensures the UEs in G_j to successfully receive the arriving data of stream j.

Then, we set $\hat{L}_1 = \hat{L}_1^*$, where $\hat{L}_1^*$ is the best basic group-cycle-length which minimizes the overall necessary wake-up ratio of all UEs in corresponding groups, i.e.,

$$\hat{L}_1^* = \arg \min_{1 \leq \hat{L}_1 \leq D_1} \sum_{j=1..M} \frac{T_j}{\hat{L}_j} \times |G_j| \tag{3}$$

subject to Eq. (1). Here, $T_j/\hat{L}_j$ is the necessary wake-up ratio of stream j. Note that a smaller ratio $T_j/\hat{L}_j$ will make the UEs subscribing G_j with less necessary wake-up cost. By finding the best basic group-cycle-length $\hat{L}_1^*$ from Eq. (3), the total necessary wake-up periods for the UEs subscribing to $G_j, j = 1..M$ can be minimized. Thus, by Eq. (1), Eq. (2), and Eq. (3), we can guarantee those UEs' QoS requirements on the aspects of data arrival rate and delay constraint.

Step 4: Based on the results of previous step, we collect those groups with the same group-cycle-length into the same "class" and sort these classes by their group-cycle-lengths in an increasing order, i.e., for any group G_j in class_k, its group-cycle-length is less than that of any group $G_{j'}$ in class_ℓ if $k < \ell$. Later, we will follow the order to schedule multicast groups class by class.

Next, we divide the time line into basic cycles and assume S_∂ be the first available subframe in cycle ∂. To schedule multicast groups, only the first $\frac{\hat{L}_M}{\hat{L}_1}$ cycles need to be considered because the allocation behavior will repeat every $\frac{\hat{L}_M}{\hat{L}_1}$ cycles. Initially, we set $S_\partial = 0, \partial = 1..\frac{\hat{L}_M}{\hat{L}_1}$. Then, let $\hat{G}$ be the set of unallocated groups and $\hat{A}_\partial$ be the set of allocated groups in cycle ∂; initially, $\hat{G} = \{G_1, G_2, .., G_M\}$ and $\hat{A}_\partial = \phi$. In addition, let $\hat{S}_j$ be the group G_j's group-start-offset; initially, $\hat{S}_j = 0$.

Then, we let $\varphi_\partial = \{UE_i | UE_i \in \bigcup_{G_\ell \in \hat{A}_\partial} G_\ell\}$ represent the subset of UEs who subscribe to the groups in the allocation list $\hat{A}_\partial$ and let $\hat{\varphi} = \{UE_i | UE_i \in \bigcup_{G_\ell \in \hat{G}} G_\ell\}$ represent the subset of UEs who subscribe the groups that have not been allocated. Based on these, we define the *potential-internal-cost function* $E_{j,\partial}$ for each group G_j in cycle $\partial, \partial = 1..\frac{\hat{L}_j}{\hat{L}_1}$, to evaluate the overall potential increased internal cost when allocating G_j in cycle ∂, i.e.,

$$E_{j,\partial} = |K_{j,\partial}| \times T_j, \tag{4}$$

where

$$K_{j,\partial} = \{UE_i | UE_i \notin G_j, UE_i \in \varphi_\partial, UE_i \in \hat{\varphi}\} \quad (5)$$

which is the set of UEs that do not subscribe to G_j but have subscribed at least one multicast group in $\hat{A}_\partial$ and at least one group still in $\hat{G}$. We will give an example later.

Based on above cost function, we design two procedures (*procedure 1 and procedure 2*) to select groups into the allocation list. Procedure 1 (the first group selection) selects the first group into the empty allocation list (i.e., $\hat{A}_\partial = \phi$) and Procedure 2 (the middle group selection) selects the groups when the allocation lists are not empty. These group selections are executed class by class and terminated until $\hat{G} = \phi$. The detailed descriptions of the two procedures are described as follows.

- *Procedure 1: the first group selection.* In procedure 1, we select the first multicast group into the allocation list $\hat{A}_\partial$ from class_1. Recall that the group-cycle-length of the groups in class_1 are equal to the length of basic cycles. We select the first group G_{j*} as follows:

$$G_{j*} = \arg \max_{G_j \in class_1} \left\{ |K'_{j,1}| \times T_j \right\}, \quad (6)$$

where $K'_{j,1} = \{UE_i | UE_i \notin G_j, UE_i \in \hat{\varphi}\}$. This means that the group G_{j*} in class_1 is selected as the first group if G_{j*} would incur the maximal internal cost when it is allocated into the middle of the allocation list. Then, we add the group G_{j*} as the first group of all allocation lists $\hat{A}_\partial$ by $\hat{A}_\partial = \{G_{j*}\}, \partial = 1..\frac{\hat{L}_M}{\hat{L}_1}$, and updating $\text{class}_1 = \text{class}_1 - G_{j*}$. Next, set the offset of G_{j*} by $\hat{S}_{j*} = 0$, and update the next available offset of cycle ∂ by $S_\partial = T_{j*}, \partial = 1..\frac{\hat{L}_M}{\hat{L}_1}$. Finally, update $\hat{G}, \varphi_\partial$, and $\hat{\varphi}$ by $\hat{G} = \hat{G} - G_{j*}, \varphi_\partial = \{UE_i | UE_i \in G_{j*}\}$, for $\partial = 1..\frac{\hat{L}_M}{\hat{L}_1}$, and $\hat{\varphi} = \{UE_i | UE_i \in \bigcup_{G_\ell \in \hat{G}} G_\ell\}$.

With procedure 1, we can avoid group G_{j*} being put into the middle of the allocation list that would incur the maximal internal cost to the system.

- *Procedure 2: the middle group selection.* In procedure 2, from the first non-empty class (denoted by class_z), we find the best pair of group G_{j*} and cycle ∂^* which conduct the minimal potential cost E_{j^*,∂^*} as the next allocated group, i.e.,

$$E_{j^*,\partial^*} = \min_{G_j \in class_z, \partial = 1..\frac{\hat{L}_j}{\hat{L}_1}} \{E_{j,\partial}\}. \quad (7)$$

Here, if more than one group G_{j*} or cycle ∂^* match, choose the one with the smallest index. Then, after adding G_{j*} to the allocation lists $\hat{A}_{\partial^*+k\times\frac{\hat{L}_{j*}}{\hat{L}_1}}, k = 0..\frac{\hat{L}_M}{\hat{L}_{j*}} - 1$, we update the offset of G_{j*} by $\hat{S}_{j^*} = (\partial^* - 1) \times \hat{L}_1 + S_{\partial^*}$ and update the next available group-start-offset of cycles by $S_{\partial^*+k\times\frac{\hat{L}_{j*}}{\hat{L}_1}} = S_{\partial^*+k\times\frac{\hat{L}_{j*}}{\hat{L}_1}} + T_{j*}$, for $k = 0..\frac{\hat{L}_M}{\hat{L}_{j*}} - 1$, and update $\varphi_{\partial^*+k\times\frac{\hat{L}_{j*}}{\hat{L}_1}} = \varphi_{\partial^*+k\times\frac{\hat{L}_{j*}}{\hat{L}_1}} \cup \{UE_i | UE_i \in G_{j^*}\}$, for $k = 0..\frac{\hat{L}_M}{\hat{L}_{j*}} - 1$. Next, update $\hat{G}$ and $\hat{\varphi}$ by $\hat{G} = \hat{G} - G_{j^*}$ and $\hat{\varphi} = \{UE_i | UE_i \in \bigcup_{G_\ell \in \hat{G}} G_\ell\}$, and then remove G_{j*} from class_z.

With above procedures, we can effectively reduce the potential internal costs of UEs. Now, we give a simple example in Fig. 4 to show the evaluation of the potential internal cost, where $N = 5, M = 4, \partial = 1$ and all the groups are with the same group-cycle-length (i.e., in the same class). In this example, we assume that {G_1, G_2} have been allocated in the allocation list $\hat{A}_1$ but {G_3, G_4} have not been allocated. Thus, if adding G_3 as the next group of the allocation list $\hat{A}_1$, it will increase some idle periods of the UEs (i.e., $K_{3,1} = \{UE_2, UE_5\}$) because these UEs do not subscribe to stream 3 but subscribes to the streams ordered in front of and in back of stream 3 in the allocation list (here, UE_2 subscribes to $\{G_2, G_4\}$ and UE_5 subscribes to $\{G_1, G_4\}$). Thus, if allocating G_3 into allocation list $\hat{A}_1$, it takes increased internal cost $E_{3,1} = |K_{3,1}| \times T_3 = 2 \times 4 = 8$ subframes.

Step 5: Finally, calculate the DRX-start-offset for each $UE_i, i = 1..N$, as follows:

$$S_i = \min_{j=1..M} \{\hat{S}_j | UE_i \in G_j\}. \quad (8)$$

Here, the DRX-cycle-length of UE_i follows the shortest wake-up cycle-interval that UE_i has to awake to receive their data of its subscribed multicast groups.

With phase one, we can determine the DRX-cycle-length L_i, DRX-start-offset S_i for each $UE_i, i = 1..N$ while ensuring to meet the data rates and delay constraints of multicast streams. Most important of all, it can reduce unnecessary internal wake-up costs of UEs incurred by the allocation order of multicast groups.

Below, we give a complete example in Fig. 5 to show the operation of phase 1, where $N = 8, M = 6$, and $\partial = 2$. In the example, assume $G_1 = \{UE_1, UE_2, UE_3, UE_4, UE_7, UE_8\}$, $G_2 = \{UE_2, UE_5\}$, $G_3 = \{UE_3, UE_4\}$, $G_4 = \{UE_3, UE_4, UE_6, UE_7, UE_8\}$, $G_5 = \{UE_6, UE_7\}$, and $G_6 = \{UE_6, UE_8\}$, where the streams' group-cycle-lengths and the required subframes to meet their data rates determined by steps $1 \sim 3$ of phase one are $\hat{L}_1 = \hat{L}_2 = 20$, $\hat{L}_3 = \hat{L}_4 = \hat{L}_5 = \hat{L}_6 = 40$ subframes and $T_1 = 2, T_2 = 4, T_3 = 3, T_4 = 2, T_5 = 1, T_6 = 4$ subframes, respectively. Consider the step 4 of phase 1, the class_1 is $\{G_1, G_2\}$ because their group-cycle-lengths are equal to the basic cycle-length. The scheme calculates the potential internal costs for such groups by $|K'_{1,1}| \times T_1 = 3 \times 2 = 6$, $|K'_{2,1}| \times T_2 = 6 \times 4 = 24$. Then, it selects G_2 as the first group of the allocation lists $\hat{A}_1$ and $\hat{A}_2$ because G_2 has the maximal potential cost of 24. In the second round, it chooses the middle group through procedure 2 because all the allocation lists are not empty. After calculating the potential internal cost by $E_{1,1} = |K_{1,1}| \times T_1 = 0 \times 2 = 0$, group G_1 is chosen as the middle group for both the allocation lists $\hat{A}_1$ and $\hat{A}_2$ because G_1 is the remaining group in the current class and it incurs the minimal potential cost. Consider the third round, the current class is $\{G_3..G_6\}$. Because the remaining groups $G_3 \sim G_6$ are with the double group-cycle-length

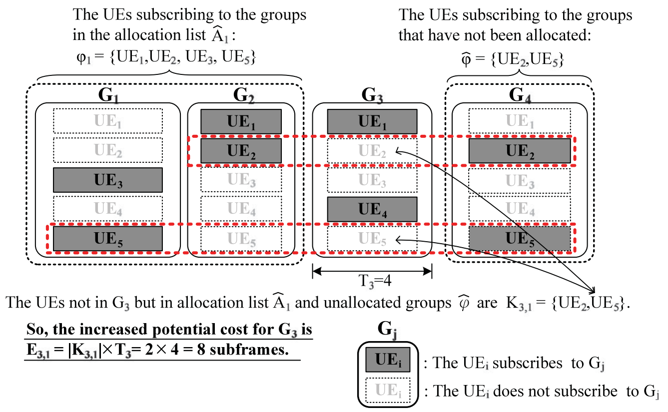

Fig. 4. An example of calculating the potential internal cost.

compared to that of the allocated groups $\{G_1, G_2\}$, these groups have double choices to be added to the allocation list in cycle 1 or cycle 2. So, after calculating the potential costs by $\{E_{3,1} = 2 \times 3 = 6, E_{4,1} = 1 \times 2 = 2, E_{5,1} = 3 \times 1 = 3, E_{6,1} = 3 \times 4 = 12\}$ for cycle 1 and $\{E_{3,2} = 2 \times 3 = 6, E_{4,2} = 1 \times 2 = 2, E_{5,2} = 3 \times 1 = 3, E_{6,2} = 3 \times 4 = 12\}$ for cycle 2, the pair of group G_4 and cycle 1 is chosen because G_4 incurs the minimal cost in cycles 1 and 2, and thus we simply choose the cycle with the smaller index for representation (i.e., cycle 1) to add. Next, in the fourth round, it chooses G_5 as the middle group for allocation list $\hat{A}_1$ because $\{E_{3,1} = 3 \times 3 = 9, E_{5,1} = 3 \times 1 = 3, E_{6,1} = 3 \times 4 = 12\}$ by cycle 1 and $\{E_{3,2} = 2 \times 3 = 6, E_{5,2} = 3 \times 1 = 3, E_{6,2} = 3 \times 4 = 12\}$ by cycle 2. In the fifth round, group G_3 is selected in $\hat{A}_2$ due to $\{E_{3,1} = 2 \times 3 = 6, E_{6,1} = 2 \times 4 = 8\}$ and $\{E_{3,2} = 1 \times 3 = 3, E_{6,2} = 2 \times 4 = 8\}$. In the last round, group G_6 is added similarly. The final allocation results are $\hat{A}_1 = [G_2, G_1, G_4, G_5, G_6]$ and $\hat{A}_2 = [G_2, G_1, G_3]$, respectively. The detailed iterations are shown in Fig. 5.

4.2 Phase 2 - Optimizing DRX Parameters

The goal of phase two is to determine the best DRX parameters, including InactivityTimer I_i and on-duration O_i for each UE$_i$, $i = 1..N$, to reduce its unnecessary *external* wake-up cost. Recall that $\hat{S}_j$ is the group-start-offset of G_j determined in phase one. Below, we calculate each UE$_i$'s I_i and O_i based on the results of $\hat{S}_j, j = 1..M$. Specifically, we first define the *external cost function* to represent the external wake-up period of UE$_i$ with InactivityTimer I_i and on-duration O_i. Then, we choose the best pair of InactivityTimer and on-duration which incur the minimum external cost for each UE. The detailed operations are depicted as follows.

First, let $\mathbb{X}_i$ be the *external cost function* for UE$_i$, which is defined by

$$\mathbb{X}_i = \sum_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} (\max\{(EPO_i - F_{i,\partial}), 0\} + I_i), \tag{9}$$

where EPO_i is the end point (in subframe) of UE$_i$'s on-duration in each of its wake-up cycle (which will be specified

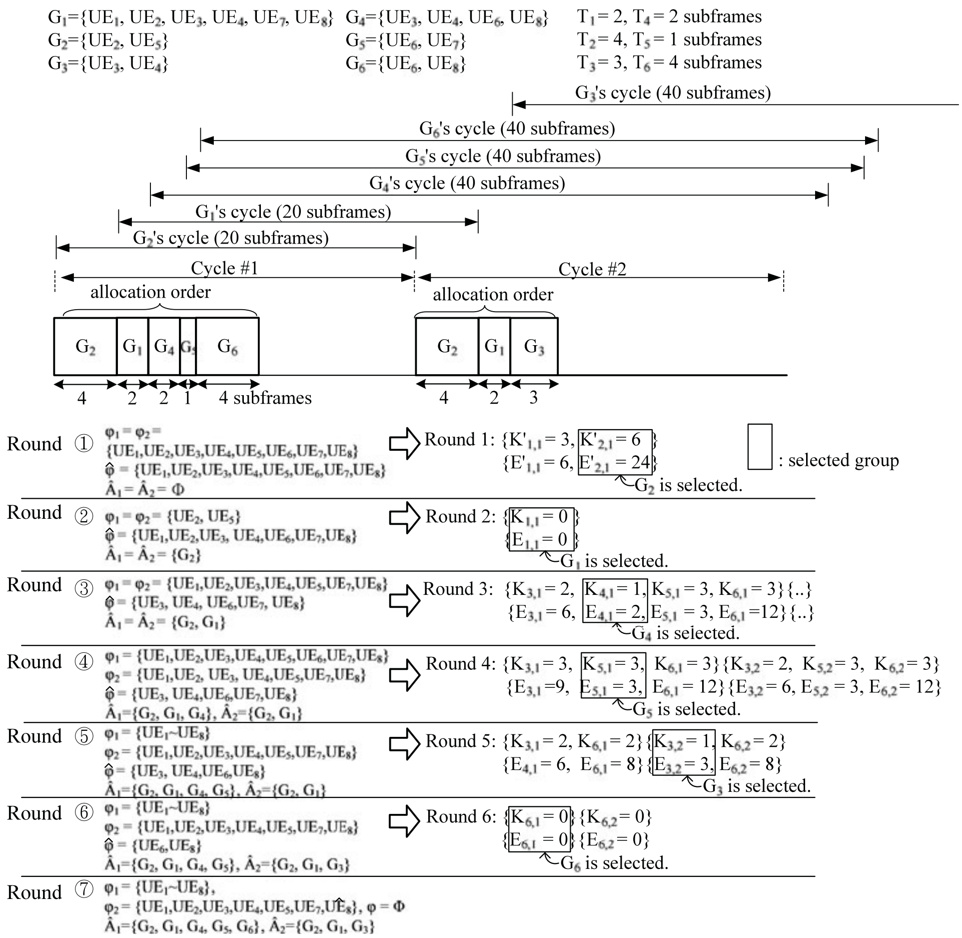

Fig. 5. An example of the operation of phase one.

later) and $F_{i,\partial}$ is the end point (in subframe) of the last group, that UE$_i$ subscribes to, in cycle ∂, i.e.,

$$F_{i,\partial} = \max_{j=1..M}\Big\{(\hat{S}_j \bmod \hat{L}_1) + T_j | UE_i \in G_j \text{ and } G_j \in \hat{A}_\partial\Big\}. \quad (10)$$

Eq. (9) shows that the external cost of UE$_i$ depends on the *necessary wake-up cost* and the *extended wake-up cost*. The necessary wake-up cost is incurred by the idle on-duration after the end point of the last group that UE$_i$ subscribes to in each wake-up basic cycle. The extended wake-up cost is incurred by the InactivityTimer I_i. Now, we use two rules to find the feasible I_i and O_i paris for each UE$_i$. Then, we choose the best pair from these solutions for each UE$_i$ which conducts the minimal external cost. The details of the two rules are depicted as follows.

Rule 1: Set $I_i = 0$ and set O_i by $O_i = EPO_i - SPO_i$. Here, EPO_i is the farthest end point of the last group, that UE$_i$ subscribes to in all basic cycles, and SPO_i is the closest start point (in subframe) of the first group, that UE$_i$ subscribes to in all basic cycles, i.e.,

$$EPO_i = \max_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} \{F_{i,\partial}\}, \quad (11)$$

$$SPO_i = \min_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} \Big\{ \min_{j=1..M} \Big\{(\hat{S}_j \bmod \hat{L}_1) | UE_i \in G_j \text{ and } G_j \in \hat{A}_\partial\Big\}\Big\}. \quad (12)$$

If I_i is set to 0, Eq. (12) will set O_i be large enough to cover all UE$_i$'s subscribed multicast groups in each cycle (we will give an example later).

Rule 2: Let $H_i = \{h_{i,1}, h_{i,2}, .., h_{i,k}\}$ be the set of "hole" where a hole means an idle period of UE$_i$ when it wakes up but no data to receive. Thus, a hole $h_{i,k} \geq 0$ occurs when two adjacent groups G$_j$ and G$_{j'}$ subscribed by UE$_i$ are separated by other groups unsubscribed by it in a cycle. Note that the case of UE$_i$ subscribing single group in a cycle is regarded as the hole with zero size. We let $\ddot{G}_{i,\partial}$ be the set of group pair $(G_j, G_{j'})$ in basic cycle ∂, where both groups G$_j$ and G$_{j'}$ are subscribed by UE$_i$, and G$_{j'}$ is the next group after G$_j$ in $\hat{A}_\partial$. Note that to make all UE$_i$'s subscribed groups have a partner in $\ddot{G}_{i,\partial}$, we add two virtual groups in each cycle ∂: one is G_0 with $T_0 = 0$ at the beginning of each cycle ∂ and the other is G_{M+1} with $T_{M+1} = 0$ behind UE$_i$'s last group in cycle. Note that this will not affect the wake-up behaviors of all UEs because such added groups do not require any resource subframes due to $T_0 = T_{M+1} = 0$. Thus, H_i can be derived as $H_i = \bigcup_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} \{h_{i,k}\} = \bigcup_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} \{(\hat{S}_{j'} \bmod \hat{L}_1) - [(\hat{S}_j \bmod \hat{L}_1) + T_j] | \forall (G_j, G_{j'}) \in \ddot{G}_{i,\partial}\}$. Then, we try to decrease the length of O_i conducted from Rule 1 to reduce the external cost of UE$_i$. To avoid missing multicast data caused by a smaller O_i, a suitable value of I_i has to be set. The worst case is to try $O_i = 1..(EPO_i - SPO_i)$ and calculate each of their corresponding I_i. Actually, only $|H_i|$ possible on-durations have to be calculated (we will explain this later). For each $h_{i,k}$ in H_i, a pair of feasible on-duration and InactivityTimer $(O_{i,k}, I_{i,k})$ for UE$_i$ can be derived as follows:

$$\begin{cases} O_{i,k} = EPO_{i,k} - SPO_i, \\ I_{i,k} = h_{i,k} + 1, \end{cases} \quad (13)$$

where $EPO_{i,k}$ is the end point of UE$_i$'s on-duration $O_{i,k}$ against $h_{i,k}$. $EPO_{i,k}$ can be derived below.

$$EPO_{i,k} = \max\{(SPO_i + 1), Y_{i,\partial}\}, \quad (14)$$

where $Y_{i,\partial} = \max_{\partial=1..\frac{\hat{L}_M}{\hat{L}_1}} \{(\hat{S}_{j'} \bmod \hat{L}_1) - (I_{i,k} - 1) | \forall (G_j, G_{j'}) \in \ddot{G}_{i,\partial}, (\hat{S}_{j'} \bmod \hat{L}_1) - [(\hat{S}_j \bmod \hat{L}_1) + T_j] \geq I_{i,k}\}$. Eq. (14) guarantees all holes (i.e., idle periods) after on-duration $O_{i,k}$ in each cycle ∂ to be covered by the InactivityTimer $I_{i,k}$. This can avoid UE$_i$ missing the multicast data because the InactivityTimer $I_{i,k}$ will expire during an idle period which is greater than or equal to $I_{i,k}$. Moreover, $EPO_{i,k}$ must be greater than SPO_i by 1 in order to trigger $I_{i,k}$.

Theorem 2. *Given any group allocation order, above two rules can find the optimal pair of on-duration O_i and InactivityTimer I_i for UE$_i$ which incurs the minimum external cost.*

Proof. For the rule 1 (by setting $I_i = 0$ and $O_i = EPO_i - SPO_i$), it is clear that it incurs the minimum external cost for the case of all groups with the same group-cycle-lengths. Consider the rule 2 (by setting $I_i = I_{i,k} = h_{i,k} + 1$ and $O_i = O_{i,k} = EPO_{i,k} - SPO_i$), where we assume $H_i = \{h_{i,1}, h_{i,2}, .., h_{i,k}\}$. Let $h_{i,k}, h_{i,k'} \in H_i$ where $h_{i,k} < h_{i,k'}$ and $\nexists h$ in H_i such that $h_{i,k} < h < h_{i,k'}$. Then, we let $\mathbb{X}_{i,\partial}(h_{i,k})$ be the external cost derived from $h_{i,k}$ for UE$_i$ in cycle ∂. Now, we show that the external cost derived from $h_{i,k}$ must be less than or equal to that of any other non-hole size $(h_{i,k} + t)$ where $0 < t < h_{i,k'} - h_{i,k}$, i.e., $\mathbb{X}_{i,\partial}(h_{i,k}) \leq \mathbb{X}_{i,\partial}(h_{i,k} + t)$ if $h_{i,k} < (h_{i,k} + t) < h_{i,k'}$. Without loss of generality, let $\mathbb{X}_{i,\partial}(h_{i,k}) = (\max\{EPO_{i,k}, F_{i,\partial}\} + I_{i,k})$. Recall that $EPO_{i,k} = \max\{(SPO_i + 1), Y_{i,\partial}\}$, it is true when $EPO_{i,k} = SPO_i + 1$ because $\mathbb{X}_{i,\partial}(h_{i,k} + t) = (\max\{(SPO_i + 1), F_{i,\partial}\} + (I_{i,k} + t)) > (\max\{(SPO_i + 1), F_{i,\partial}\} + I_{i,k}) = \mathbb{X}_{i,\partial}(h_{i,k})$ for each cycle ∂. Consider $EPO_{i,k} = Y_{i,\partial}$ where $Y_{i,\partial}$ is determined by a hole larger than $h_{i,k}$ and $(h_{i,k} + t)$. Then, we have $\mathbb{X}_{i,\partial}(h_{i,k} + t) = (\max\{Y_{i,\partial} - t, F_{i,\partial}\} + (I_{i,k} + t)) \geq (\max\{Y_{i,\partial}, F_{i,\partial}\} + I_{i,k}) = \mathbb{X}_{i,\partial}(h_{i,k})$ for each cycle ∂. Otherwise, $EPO_{i,k} = Y_{i,\partial}$ where $Y_{i,\partial}$ is determined by a hole larger than $h_{i,k}$ but smaller than or equal to $(h_{i,k} + t)$, which contradicts the previous assumption of $h_{i,k} < (h_{i,k} + t) < h_{i,k'}$. Thus, it can find the optimal pair of on-duration O_i and InactivityTimer I_i for UE$_i$ which incurs the minimum external cost by above two rules. □

Based on the external cost function, we can find the best pair of on-duration and InactivityTimer for each UE and can effectively reduce their external wake-up cost. Below, we give an example in Fig. 6 (based on the group allocation order resulted from Fig. 5) to show the determination of the best InactivityTimer and on-duration for UEs. We use UE$_8$ as the example, where UE$_8$ subscribes to G$_1$, G$_4$, and G$_6$. First, the rule one finds a feasible pair of on-duration and InactivityTimer by $O_8 = EPO_8 - SPO_8 = 13 - 4 = 9$ and $I_8 = 0$ due

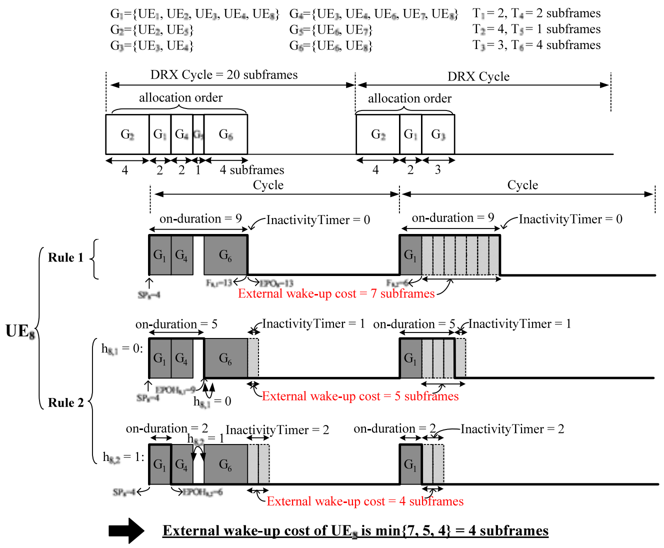

Fig. 6. An example to determine the best pair of on-duration and InactivityTimer for the UE.

to $F_{8,1} = 13$ and $F_{8,2} = 6$. We can see if InactivityTimer is set by $I_i = 0$, the on-duration O_i has to start from the closest point of the first group, i.e., $SPO_8 = 4$, and end at the farthest end point of the latest group i.e., $F_{8,1} = 13$, that the UE$_i$ subscribes to, to cover all the data receiving periods in each cycle. Thus, it costs $\mathbb{X}_8 = (\max\{13-13, 0\}+0)+ (\max\{13-6, 0\}+0) = 7$ subframes for the external cost. Considering the rule two, the hole set is $H_8 = \{h_{8,1} = 0, h_{8,2} = 1\}$ because the holes (i.e., idle periods) between (G$_1$, G$_4$) and between (G$_4$, G$_6$) are 0 and 1, respectively. For the hole length $h_{8,1} = 0$, the UE$_8$'s on-duration O_8 is set by $O_8 = O_{8,1} = EPO_{8,1} - SPO_8 = 9 - 4 = 5$ and InactivityTimer I_8 is set by $I_8 = I_{8,1} = h_{8,1} + 1 = 0 + 1$. Thus, it costs $\mathbb{X}_8 = (\max\{9-13, 0\}+1)+ (\max\{9-6, 0\}+1) = 5$ subframes for the external cost. For the hole length $h_{8,2} = 1$, the UE$_8$'s on-duration is set by $O_8 = O_{8,2} = EPO_{8,2} - SPO_8 = 5 - 4 = 1$ and InactivityTimer is set by $I_8 = I_{8,2} = h_{8,2} + 1 = 1 + 1 = 2$. Thus, it costs $\mathbb{X}_8 = (\max\{5-13, 0\}+2) + (\max\{5-6, 0\}+2) = 4$ subframes. Thus, the optimal pair of on-duration and InactivityTimer for UE$_8$ is $O_8^* = 1$ and $I_8^* = 2$ which takes the minimal external cost of 4 subframes.

4.3 Time Complexity Analysis

We analyze the time complexity of the proposed heuristic as follows. In phase one, the step 1 costs $O(M \cdot \log(M))$ to sort the delay constraints of M multicast streams. In steps 2 and 3, it costs $O(M)$ to determine the group-cycle-lengths and corresponding required subframes for M multicast groups. Thus, it costs $D_1 \cdot O(M)$ to find the best basic group-cycle-length to incur the minimal necessary wake-up cost, where D$_1$ is the minimal delay constraint among all UEs'. In step 4, the procedure 1 costs $O(M \cdot N)$ to evaluate the potential internal costs for M multicast groups where each multicast group checks whether or not each UE disappears in this group but appears in $\hat{\varphi}$. For procedure 2, slightly deferent from procedure 1, each time it costs $O(M \cdot N \cdot \frac{\hat{L}_M}{\hat{L}_1})$ to evaluate the potential internal costs for at most M multicast groups where each multicast group checks whether or not the UE disappears in this group but appears in $\hat{\varphi}$ and $\varphi_\partial, \partial = 1..\frac{\hat{L}_M}{\hat{L}_1}$. This operation repeats $(M-1)$ times until all multicast groups have been allocated. Finally, the step 5 costs $O(N \cdot M \cdot \frac{\hat{L}_M}{\hat{L}_1})$ to determine the DRX-start-offsets and DRX-cycle-length for N UEs from $\frac{\hat{L}_M}{\hat{L}_1}$ allocation lists where each allocation list may have M groups at most. Thus, phase one totally costs $O(M \cdot \log(M)) + D_1 \cdot O(M) + [O(M \cdot N) + (M-1) \cdot O(M \cdot N \cdot \frac{\hat{L}_M}{\hat{L}_1})] + O(N \cdot M \cdot \frac{\hat{L}_M}{\hat{L}_1}) = O(M^2 \cdot N)$.

In the phase two, for each UE, the rule 1 costs $O(1)$ to determine the InactivityTimer and costs $O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M)$ to determine the on-duration according to the group-start-offset and the end point of the groups (at most M groups), that the UE subscribes to, from $\frac{\hat{L}_M}{\hat{L}_1}$ allocation lists. So, the rule one costs $N \cdot (O(1) + O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M)) = O(N \cdot M)$. For the rule 2, each UE costs $O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M)$ to find the hole set from $\frac{\hat{L}_M}{\hat{L}_1}$ allocation lists because each UE has at most $(M+1)$ holes if the UE subscribes to at most M groups in a cycle. Then, each UE costs $O(1)$ and $O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M)$ to determine its InactivityTimer and on-duration, respectively, from each hole. Thus, the rule two totally costs $N \cdot O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M) \cdot (O(1) + O(\frac{\hat{L}_M}{\hat{L}_1} \cdot M)) = O(M^2 \cdot N)$. Therefore, the phase two totally costs $O(N \cdot M) + O(M^2 \cdot N) = O(M^2 \cdot N)$.

In brief, the time complexity of the proposed heuristic incurred by phases one and two is $O(M^2 \cdot N) + O(M^2 \cdot N) = O(M^2 \cdot N)$. We also note that the number of UEs is much larger than the number of multicast streams in a real network, i.e., $M \ll N$.

5 Simulation Results

In this section, we develop a simulator in C++ language to verify the effectiveness of the proposed scheme. The system parameters of the simulator are listed below. The frequency bandwidth is 10 MHz. The number of UEs is $N = 150$. The total number of multicast streams available for the UEs to subscribe is $M = 1, 2, 3, 4, 5, 6$, and 7. Each UE$_i$ will randomly decide its number of subscribed groups $|X_i|$ $(1 \leq |X_i| \leq M)$ and then select $|X_i|$ out of M groups randomly in each experiment. Each multicast stream j has the delay constraint $D_j = 50 \sim 300$ subframes [34] and an admitted data rate $R_j = 600 \sim 1{,}900$ bits/ms, which are admitted by an admission control. The transmission rate of each UE is generated according to Table 1. In the simulation, we compare our *two-phase* scheme (2-Phase) to our previous *minimal-cost* scheme (MC) [1], the *random selection* scheme *(Random)*, and the *optimal* scheme *(OPT)*. The *MC* scheme can determine the best group allocation order to reduce internal costs but it neglects the optimization of other DRX parameters, such as InactivityTimer and on-duration, for each UE. The *Random* scheme always adopts the default DRX parameters and determines the group allocation order randomly for simplicity. The OPT scheme can find the optimal group allocation order, InactivityTimer, and on-duration for UEs by a brute force manner. Thus, it incurs high computational cost.

We consider four performance metrics: (i) *wake-up ratio*: the ratio of wake-up subframes over the total execution

TABLE 1
The Modulation and Coding Schemes Supported by the LTE-A Standard [35]

channel quality identifier (CQI)	modulation	code rate × 1,024	rate (bits/subframe)
1	QPSK	78	1,279.2
2	QPSK	120	1,969.0
3	QPSK	193	3,166.8
4	QPSK	308	5,153.4
5	QPSK	449	7,366.8
6	QPSK	602	9,876.6
7	16QAM	378	12,403.4
8	16QAM	490	16,078.4
9	16QAM	616	20,212.8
10	64QAM	466	22,936.2
11	64QAM	567	27,907.4
12	64QAM	666	32,779.4
13	64QAM	772	37,997.0
14	64QAM	873	42,968.6
15	64QAM	948	46,659.4

subframes; (ii) *conserved energy*: the average saved energy per subframe for the DRX enabled UEs, where the power consumption model for downlink transmission is $P_d = \alpha_d \times t_d + \beta$ [36], where $\alpha_d = 51.97$ (mW/Mbps) is the power coefficient for downlink transmission, t_d (Mbps) is the downlink throughput, and $\beta = 1{,}288.04$ (mW) is the base power when throughput is 0 (note that we use the tail base power $P_{tail} = 1{,}060$ (mW) as the power consumption of a UE during the sleep period); (iii) *failed-to-sleep probability*: the probability that at least one UE fails to sleep; (iv) *computational complexity*: the average computational time to successfully determine DRX configurations for all UEs in each round. Note that each simulation result is averaged by at least 5,000 experiments.

5.1 Wake-up Ratio

We first investigate the effects of number of multicast streams on wake-up ratio of all schemes. As shown in Fig. 7, when the number of streams increases, the wake-up ratio of all schemes increases. The reason is that more streams are able to be subscribed by UEs, and thus UEs have to stay awake for a longer period to receive the corresponding data and meet the required data rates. The *Random* scheme incurs the largest wake-up ratio because it neglects the optimization of the group allocation order and DRX parameters. The *MC* scheme is better than *Random* because it uses potential cost metric to reduce the internal cost; however, the external cost still remains. Note that the proposed *2-Phase* scheme tries to reduce both the internal and external costs by optimizing multicast group allocation order and DRX parameters through the phases 1 and 2. Thus, the unnecessary wake-up periods of UEs are significantly reduced and thus its performance is close to that of the OPT scheme.

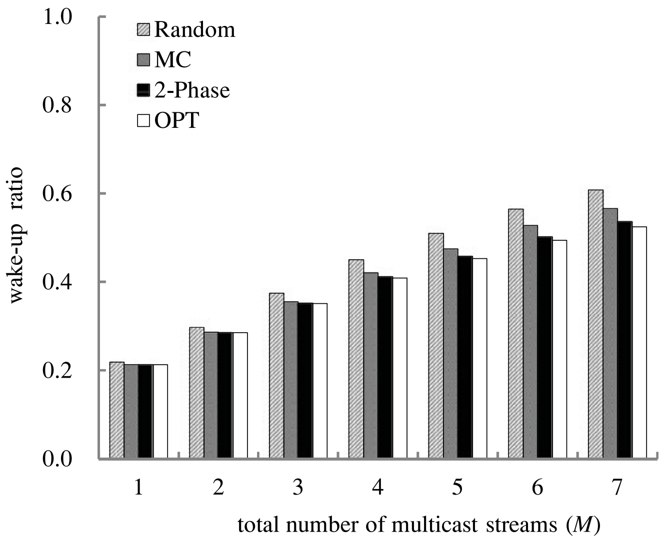

Fig. 7. Comparisons on the wake-up ratio of all schemes under $M = 1 \sim 7$ multicast streams.

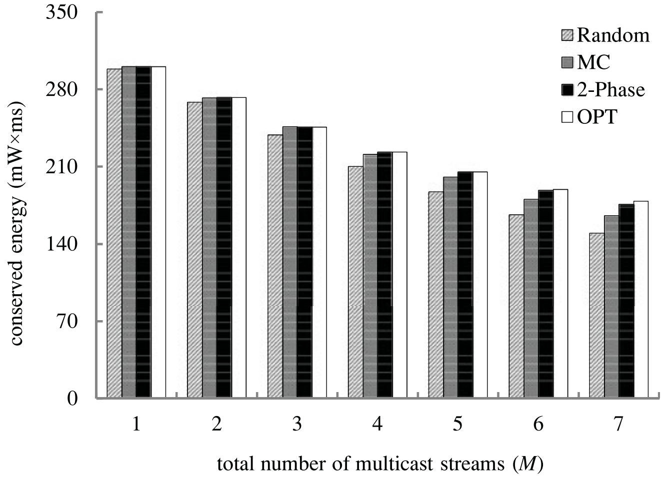

Fig. 8. Comparisons on conserved energy of all schemes under $M = 1 \sim 7$ multicast streams.

5.2 Conserved Energy

We then investigate the effects of number of multicast streams on the average conserved energy of all schemes. As shown in Fig. 8, when the number of streams increases, the conserved energy of all schemes decreases. This is because the conserved energy of UEs is inversely proportional to their wake-up periods. The *Random* scheme saves the least energy and the *MC* scheme is slightly better than the *Random* scheme because *Random* neglects to determine the best group allocation order compared to the *MC* scheme. Note that the performance of our *2-Phase* scheme is close to that of the *OPT* scheme, because our scheme can significantly reduce the wake-up costs incurred by the internal and external costs.

5.3 Failed-to-Sleep Probability

Next, we investigate the effects of number of multicast streams on failed-to-sleep probability of all schemes. As shown in Fig. 9, when the number of streams increases, the probability of all schemes increases. The reason is that the network resource becomes insufficient when more multicast streams are requested by UEs. We can see that the failed-to-sleep probability of the *Random* scheme and the *MC* scheme is the worst and second worst one, respectively, but our *2-Phase* scheme is keeping close to the *OPT* scheme. This is because our scheme can well utilize the resource by determining appropriate group allocation orders based on UEs' wake-up/sleep behaviors when they subscribe to multicast streams.

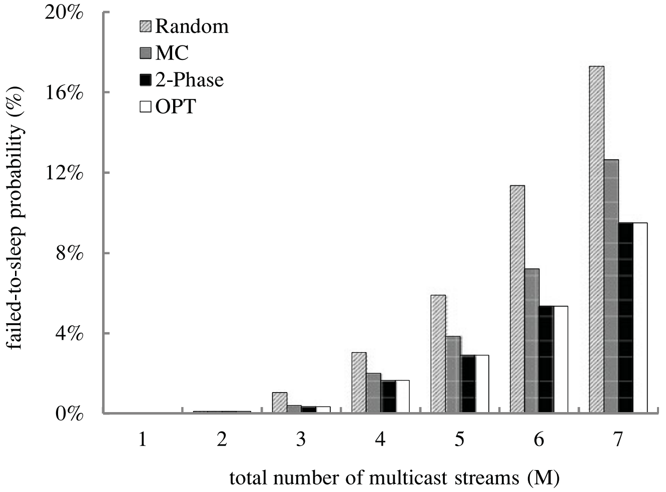

Fig. 9. Comparisons on failed-to-sleep probability of all schemes under $M = 1 \sim 7$ multicast streams.

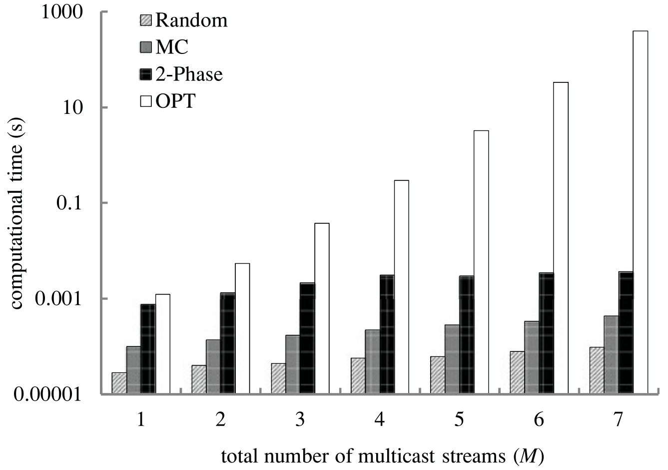

Fig. 10. Comparisons on computational complexity of all schemes under $M = 1 \sim 7$ multicast streams.

5.4 Computational Complexity

Finally, we investigate the effects of number of multicast streams on computational time of all schemes. Here, the computation time is measured by the platform of DELL Optiplex 990 with Intel i7-2600 3.4 GH and 4 GB DDR3 SDRAM. As shown in Fig. 10, the computing time of the *OPT* scheme increases exponentially as the number of streams grows (note that the y-axis is drawn with exponential scales). This is because it takes more combinations to find the optimal allocation orders and corresponding DRX parameters for UEs when the number of streams increases. Contrarily, our *2-Phase* scheme takes much less computational time and it is lower than 0.01 second per round even when the amount of multicast streams is up to 7. The reason is that our *2-Phase* scheme sequentially determines the group allocation orders in phase 1 and optimizes the best DRX configurations by the distinct rules in phase 2. Thus, the time complexity of the *2-Phase* scheme is reduced significantly. Note that the *Random* scheme takes the least computing time but it incurs the worst wake-up ratio, the lowest conserved energy, and the highest failed-to-sleep probability. On the other hand, the *MC* scheme takes less computing time than our *2-Phase* scheme because ours needs extra effort to optimize DRX parameters and thus has higher energy efficiency.

To conclude, the performances of our *2-Phase* scheme in terms of wake-up ratio, energy conservation, and failed-to-sleep probability, are much close to that of the *OPT* scheme but ours only takes $\frac{1}{2} \sim \frac{1}{100{,}000}$ computational time of the OPT scheme.

6 Conclusions

In this paper, we have addressed the DRX optimization problem under the consideration of QoS guarantee for multicast streams and energy conservation for UEs. We have proved the problem to be NP-complete and developed an energy-efficient scheme. Specifically, this scheme consists of two phases. The first phase uses the minimal cost strategy to reduce the internal cost of UEs and the second phase adopts two special rules to further reduces the external cost of UEs. Extensive simulation results have verified the effectiveness of our scheme and shown that our scheme can significantly save UEs' energy by decreasing their wake-up ratio in an effective way. We also note that the performance of our scheme approximates that of the optimal scheme in most cases. For future directions, we will study how to optimally deploy our solution in system aspects of LTE-A and conduct the experimental field trial for future optimization. In addition, we expect to cooperate DRX for the UEs in both RRC_CONNETED and RRC_IDLE states to further optimize their energy efficiency.

Acknowledgments

This research is co-sponsored by MOST-102-2218-E-009-014-MY3, MOST 104-2221-E-024-005-, MOST 104-2221-E-009-113-MY3, MOST 104-2218-E-009-009, CGURP, MoE ATU Plan, Academia Sinica AS-102-TP-A06, ITRI, Delta, and D-Link. This work was also supported by MOST, NTU, and Intel under Grants MOST 103-2911-I-002-001, NTU-ICRP-104R7501, and NTU-ICRP-104R7501-1. A preliminary version of this paper appeared in the *Proceeding of IEEE Wireless Communications and Networking Conference*, 2013[1]. J.-J. Chen is the corresponding author.

References

[1] J.-M. Liang, P.-C. Hsieh, J.-J. Chen, and Y.-C. Tseng, "Energy-efficient DRX scheduling for multicast transmissions in 3GPP LTE-Advanced wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2013, pp. 551–556.

[2] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, S. Parkvall, and E. Research, "LTE: The evolution of mobile broadband," *IEEE Commun. Mag.*, vol. 47, no. 4, pp. 44–51, Apr. 2009.

[3] O. Oyman, J. Foerster, Y. J. Tcha, and S.-C. Lee, "Toward enhanced mobile video services over WiMAX and LTE," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 68–76, Aug. 2010.

[4] H.-S. Liu, C.-H. Wang, and R.-I. Chang, "The design and implementation of a future internet live TV system over 4G networks," *Telecommun. Syst.*, vol. 54, no. 1, pp. 203–214, 2013.

[5] G. K. Walker, J. Wang, C. Lo, X. Zhang, and G. Bao, "Relationship between LTE broadcast/eMBMS and next generation broadcast television," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 185–192, Jun. 2014.

[6] J. Liu, Q. Huang, B. Yang, and Y. Zhang, "Efficient multicast key distribution using HOWP-based dynamic group access structures," *IEEE Trans. Comput.*, vol. 62, no. 8, pp. 1652–1672, Aug. 2011.

[7] A. Ali, "A review of different comparative studies on mobile operating system," *Res. J. Appl. Sci., Eng. Technol.*, vol. 7, no. 12, pp. 2578–2582, 2014.

[8] (2014). Getting started with Multi-Window on the Samsung Galaxy Note 4. [Online]. Available: http://www.cnet.com/how-to/getting-started-with-multi-window-on-the-samsung- galaxy-note-4/

[9] (2012). Mobile news: QSlide on LG Optimus G. [Online]. Available: http://metroplexbbq.blogspot.tw/2012/10/qslide-on-lg-optimus-g.html

[10] (2012). Quasar cydia tweak lets you run multiple apps in separate windows simultaneously. [Online]. Available: http://www.ijailbreak.com/cydia/quasar-cydia-tweak-ios/

[11] (2013). Stick it! (Pop-up Player). [Online]. Available: https://play.google.com/store/apps/ details?id=com.myboyfriendisageek.stickit

[12] M.-H. Park, Y.-H. Park, H.-Y. Jeong, and S.-W. Seo, "Key management for multiple multicast groups in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1712–1723, Sep. 2013.

[13] A. Elnashar and M. El-Saidny, "Extending the battery life of smartphones and tablets: A practical approach to optimizing the LTE network," *IEEE Veh. Technol. Mag.*, vol. 9, no. 2, pp. 38–49, Jun. 2014.

[14] 3GPP TS 36.321, "E-UTRA; Medium Access Control (MAC) protocol specification," v12.4.0 (Release 12), January 2015.

[15] 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 J. Emerging Select. Topics Circuits Syst.*, vol. 3, no. 1, pp. 13–22, Mar. 2013.

[16] C.-C. Kao, S.-R. Yang, and H.-C. Chen, "A sleep-mode interleaving algorithm for layered-video multicast services in IEEE 802.16e networks," *Comput. Netw.*, vol. 56, pp. 3639–3654, 2012.

[17] M. Han and J. P. Hong, "Balancing QoS and fairness of layered video multicast in LTE networks," in *Proc. Int. Conf. Adv. Commun. Technol.*, 2012, pp. 387–392.

[18] W.-H. Kuo, W. Liao, and T. Liu, "Adaptive resource allocation for layer-encoded IPTV multicasting in IEEE 802.16 WiMAX wireless networks," *IEEE Trans. Multimedia*, vol. 13, no. 1, pp. 116–124, Feb. 2011.

[19] Y.-J. Yu, P.-C. Hsiu, A.-C. Pang, and C.-P. Lai, "Power-aware scalable video multicast in 4G wireless system," in *Proc. IEEE Global Telecommun. Conf.*, 2010, pp. 1–6.

[20] C. Wu, X. Sun, and T. Zhang, "A power-saving scheduling algorithm for mixed multicast and unicast traffic in MBSFN," in *Proc. Int. Comput., Commun. Appl. Conf.*, 2012, pp. 170–174.

[21] L. Zhou, H. Xu, H. Tian, Y. Gao, L. Du, and L. Chen, "Performance analysis of power saving mechanism with adjustable DRX cycles in 3GPP LTE," in *Proc. IEEE Veh. Technol. Conf.*, 2008, pp. 1–5.

[22] J. Wigard, T. Kolding, L. Dalsgaard, and C. Coletti, "On the user performance of LTE UE power savings schemes with discontinuous reception in LTE," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2009, pp. 1–5.

[23] S. Fowler, "Study on power saving based on radio frame in LTE wireless communication system using DRX," in *Proc. IEEE Global Telecommun. Conf. Workshops*, 2011, pp. 1062–1066.

[24] S. Jin and D. Qiao, "Numerical analysis of the power saving in 3GPP LTE advanced wireless networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1779–1785, May 2012.

[25] F. W. Li, Y. Q. Zhang, and L. W. Li, "Enhanced discontinuous reception mechanism for power saving in TD-LTE," in *Proc. IEEE Int. Conf. Comput. Sci. Inf. Technol.*, 2010, pp. 682–686.

[26] C. Zhong, T. Yang, L. Zhang, and J. Wang, "A new discontinuous reception DRX scheme for LTE-Advanced carrier aggregation systems with multiple services," in *Proc. IEEE Veh. Technol. Conf.*, 2011, pp. 1–5.

[27] H. Xu, H. Tian, B. Huang, and P. Zhang, "An improved dynamic user equipment power saving mechanism for LTE system and performance analysis," *Sci. China Inf. Sci.*, vol. 53, no. 10, pp. 2075–2086, 2010.

[28] E. Liu, J. Zhang, and W. Ren, "Adaptive DRX scheme for beyond 3G mobile handsets," in *Proc. IEEE Global Telecommun. Conf.*, 2011, pp. 1–5.

[29] S. Gao, H. Tian, J. Zhu, and L. Chen, "A more power-efficient adaptive discontinuous reception mechanism in LTE," in *Proc. IEEE Veh. Technol. Conf.*, 2011, pp. 1–5.

[30] 3GPP TS 36.331, "E-UTRA; Radio resource control (RRC) protocol specification," v12.4.0 (Release 12), Jan. 2015.

[31] C. Bouras, N. Kanakis, V. Kokkinos, and A. Papazois, "Application layer forward error correction for multicast streaming over LTE networks," *Int. J. Commun. Syst.*, vol. 26, no. 11, pp. 1459–1474, 2013.

[32] J. Bhasker and S. Sahni, "Optimal linear arrangement of circuit components," *J. VLSI Comput. Syst.*, vol. 2, no. 1, pp. 87–109, 1987.

[33] R. Jin, Y. Xiang, D. Fuhry, and F. F. Dragan, "Overlapping matrix pattern visualization: A hyper graph approach," in *Proc. IEEE Int. Conf. Data Mining*, 2008, pp. 313–322.

[34] 3GPP TS 23.203, "Policy and charging control architecture," v12.9.0 (Release 12), Jun. 2015.

[35] 3GPP TS 36.213, "E-UTRA; Physical layer procedures," v12.4.0 (Release 12), Jan. 2015.

[36] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "A close examination of performance and power characteristics of 4G LTE networks," in *Proc. Int. Conf. Mobile Syst., Appl. Serv.*, 2012, pp. 225–238.