Pervasive and Mobile Computing 52 (2019) 29–45

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Energy-efficient scheduling scheme with spatial and temporal aggregation for small and massive transmissions in LTE-M networks[☆]

Jia-Ming Liang [a,b], Po-Yen Chang [c,d], Jen-Jee Chen [e,*]

[a] Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan 33302, Taiwan
[b] Department of General Medicine, Chang Gung Memorial Hospital, Taoyuan 33378, Taiwan
[c] Department of Computer Science, College of Computer Science, National Chiao Tung University, Hsin-Chu 30010, Taiwan
[d] Information and Communications Research Laboratories, Industrial Technology Research Institute, Hsinchu 31057, Taiwan
[e] Department of Electrical Engineering, National University of Tainan, Tainan 70005, Taiwan

ARTICLE INFO

Article history:
Received 3 January 2018
Received in revised form 2 November 2018
Accepted 14 November 2018
Available online 26 November 2018

Keywords:
Internet of Things (IoT)
Discontinuous reception/discontinuous transmission (DRX/DTX)
LTE-M
Machine-to-machine communication (M2M)
Massive connectivity
Power saving
Resource allocation
Small data

ABSTRACT

Machine-to-machine (M2M) communication is one of the key technologies to realize *Internet of Things (IoT)*. Since IoT applications are mainly for smart sensing, such as metering, home surveillance, disaster detection, and e-health, their special sensing/uploading behaviors will result in periodic and/or event-driven *small data* transmissions, which may potentially decrease the radio resource efficiency. On the other hand, due to the frequent communication nature of sensing data for IoT applications, the power consumption is increasing dramatically. To reduce the power consumption of IoT devices, the *3rd Generation Partnership Project (3GPP)* has defined the *discontinuous reception/discontinuous transmission (DRX/DTX)* mechanism to allow devices to turn off their radio interfaces and go to sleep in various patterns. However, how to optimize the DRX/DTX scheduling while improving the resource efficiency is still an open issue. In this paper, we investigate an uplink resource allocation problem over *long-term evolution machine-to-machine (LTE-M)* networks, which is standardized by 3GPP to improve performance on IoT. In this network, we consider the periodic, event-driven, and query-based IoT traffic while minimizing the devices' power consumption. We prove this problem to be NP-complete and propose an *aggregation-efficient DRX/DTX scheduling (AEDS)* scheme. This scheme takes advantage of both *spatial* and *temporal* data aggregation while applying DRX/DTX for energy saving. Specifically, the scheme consists of three phases. The first phase exploits long-term static scheduling for periodic data to ensure the latency and data rate while minimizing the devices' wake-up time. The second phase tries to decrease devices' power consumption through precisely determining their DRX/DTX configurations. Finally, the third phase employs short-term dynamic scheduling for event-driven and query-based data to improve transmission efficiency. Therefore, both small data and power consumption problems are relieved. Extensive simulation results show that the proposed scheme can improve resource efficiency, enlarge network capacity while reducing power consumption compared to the existing schemes.

☆ A preliminary version of this paper appeared in the *Proceeding of IEEE Wireless Communications and Networking Conference*, 2017 (Chang et al., 2017 [1]).

* Corresponding author.
E-mail address: jjchen@mail.nutn.edu.tw (J.-J. Chen).

https://doi.org/10.1016/j.pmcj.2018.11.002
1574-1192/

1. Introduction

The *Internet of Things (IoT)* is emerging as the next mega-trend in the development of the Internet. It can provide plenty of applications and powerful services such as home surveillance, disaster detection, and e-health care that make our lives more convenient and efficient. To realize the IoT, a promising technology, called *machine-to-machine (M2M) communication* is developed [1]. It has been standardized as the *machine type communications (MTC)* in LTE-M to enable networked devices exchanging information and performing actions without the assistance of humans. Specifically, the LTE-M specifies two important components for MTC: *device* and *gateway*. Each device in LTE-M networks can connect to a base station (i.e., *evolved NodeB, eNB*) directly or relay data via a gateway over two hops [2]. In LTE-M, the radio resource unit for the data transmission is a *resource block (RB)*, which is capable of carrying hundreds of data bits for each transmission. Since the IoT applications are commonly used for environment monitoring, event detection and/or information reporting, these incur a huge number of '*small data*' on uplink transmission. When the high-capacity radio resource (i.e., RBs) is allocated to such small data transmission individually, it will significantly decrease the resource efficiency and network capacity [3]. In addition, those small data are with diverse traffic features, such as periodic, event-driven, and query-based, how to guarantee their *quality of service (QoS)* is also a challenge.

On the other hand, continuous low-rate small data may be reported from IoT devices over a long period of time, implying frequent communication nature of sensing data, and thus the power consumption of devices becomes a critical issue. To reduce the power consumption of IoT devices, the *3rd Generation Partnership Project (3GPP)* has defined the *discontinuous reception/discontinuous transmission (DRX/DTX) mechanism* to allow devices to turn off their radio interfaces and go to sleep in various patterns. However, how to effectively integrate uplink data aggregation with optimal configuring DRX/DTX[1] parameters for IoT devices is still left as an open issue.

To address the above issues, we propose an *aggregation-efficient DRX scheduling (AEDS)* scheme, which takes advantage of data aggregation in both spatial and temporal manners while leveraging optimal DRX scheduling to conserve devices' energy. Specifically, for the spatial data aggregation, this scheme tries to classify the traffic based on the devices' QoS features and make the devices nearby with each other to connect with the special gateways. Thus, the transmission data from different devices can be aggregated efficiently. In addition, by utilizing spatial reuse technique, the data transmission from different devices to different gateways can be transmitted concurrently. Therefore, the resource efficiency can be further increased. For the temporal data aggregation, this scheme adopts a mix of static and dynamic resource allocation strategy, which leverages a long-term scheduling to aggregate the periodic small data under the constraints of limited delay and exploits a short-term scheduling for the event-driven/query-based small data by piggybacking them into the vacant space of long-term scheduling results. Thus, the latency of both the event-driven and query-based data can be guaranteed and the unnecessary allocated resource can be reduced. On the other hand, to maximize the sleep efficiency of devices, the DRX cycle of devices are set by referring to the delay constraint. To prevent extra wake-up time of devices caused by resource collision, the setting of the DRX cycles follow the exponential increment rule. Based on the above results, the RB resource and the DRX parameters of each device are scheduled to increase the capacity of the network, guarantee the QoS of services and optimize the saving power of devices.

Major contributions of this paper are threefold. First, this is the first work joint DRX scheduling optimization with small data aggregation issues for IoT applications in the LTE-M network. We formulate this problem and prove it to be NP-complete by reducing it to the *multiple choice knapsack* problem. Second, we develop an energy-efficient scheduling scheme to optimize the DRX mechanism with the long-term and short-term scheduling for small data to ensure the latency and data rate while minimizing the devices' wake-up time. Third, through the simulation results, we give a constructive summary that the joint problem of data aggregation and DRX scheduling for IoT communications is very important and we validate the proposed scheme can improve resource efficiency, enlarge network capacity, decrease devices wake-up periods while reducing power consumption compared to the existing schemes.

The rest of this paper is organized as follows. Section 2 surveys the related work. Section 3 introduces the background and formally defines the problem. Section 4 presents the proposed scheme. Numerical results are shown in Section 5. Finally, concluding remarks are given in Section 6.

2. Related work

In the literature, the study [2] has surveyed the existing M2M service platforms and discussed the research challenges when enabling M2M services. In Refs. [4–7], they proposed several path selection strategies to improve network performance. Specifically, in Ref. [4], the proposed scheme selected the paths based on the best channel condition to improve the throughput. In work [5], the approach determined the uplink paths according to the specific received channel quality to balance the tradeoff between transmission efficiency and resource utilization. In Ref. [6], it enforced the devices to connect to their nearest gateways to get better transmission quality. In Ref. [7], the devices chose their gateways by the shortest end-to-end transmission time. In this way, the transmission efficiency can be improved. However, the above studies [4–7] all preferred each device to own an independent uplink route which produced lots of '*individual paths*' in the network. This would decrease the radio resource efficiency due to consuming too much bandwidth. Thus, the study [8] proposed to

[1] The operation of DTX is similar to DRX. Thus, we state DRX in the following.

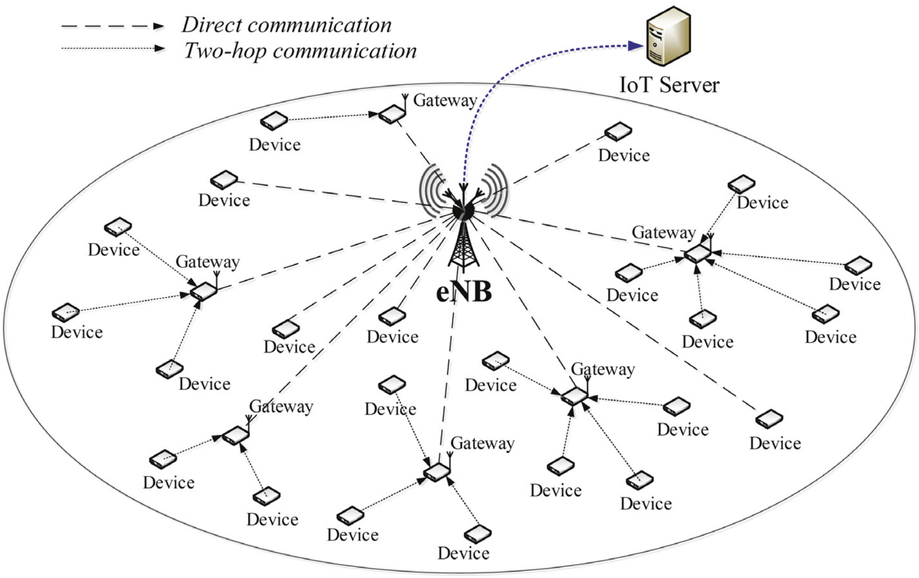


Fig. 1. The LTE-M network architecture.

share the physical resource blocks among different devices and thus can improve the data aggregation efficiency. However, these studies neglect the traffic characteristics such as the periodic, event-driven, and query-based small data. The power consumption of devices is also ignored.

Recently, several studies [9–13] have evaluated the performance of DRX and shown that enabling DRX can significantly save device' energy. The work in [14] proposes to adjust devices' DRX configurations, but this scheme requires considerable control signaling for the adjustment. In [15], a dynamic DRX scheme is proposed to determine devices' sleep cycle based on their traffic loads. In [16], an adaptive DRX is proposed to adjust devices' extended wake-up periods based on their channel qualities to improve system utility. However, these studies neglect both the low-rate small data and QoS requirements for the IoT devices. The work in [17] investigates an energy-efficient paging mechanism to integrate with DRX to reduce signaling overhead and alleviate network load. In [18], a novel energy harvesting DRX mechanism is proposed to minimize the device wakeup delay while keeping the energy consumption in cellular networks for IoT devices. The work in [19] proposes three different mechanisms to achieve device grouping in IoT networks. Each mechanism makes different trade-offs between bandwidth usage, device energy consumption for multicast transmissions. The study [20] analyzes the impact of DRX parameters based on Markov chain model and proposes an IoT-aware algorithm to balance power consumption, delay, and signal load for IoT networks. The research [21] proposes a grouping-based energy efficient DRX strategy for massive IoT to conserve the energy of devices. However, the above studies neglect to fully aggregate data through gateways and leverage spatial reuse, which significantly limits the network performance in IoT networks.

Based on the above observations, it motivates us to address the problem of small data transmissions with special traffic requirements while optimizing the DRX configurations to save energy. To the best of our knowledge, this is the first paper to address both the issues of resource efficiency and energy conservation in LTE-M networks.

3. Preliminary

In this section, we first introduce the network architecture in LTE-M. Then, we describe the traffic features and QoS requirements of the MTC devices. Finally, we formally define this uplink resource allocation problem.

3.1. LTE-M network architecture

In LTE-M network, there are two special machines designed for communication: one is *MTC device* and the other is *MTC gateway*. The *MTC device* is an end device that can communicate with an eNB directly or connect to a *MTC gateways* over two hops. Thus, the MTC gateway can help devices to relay and transfer their data to the eNB to increase the transmission efficiency. The LTE-M network architecture is shown in Fig. 1.

3.2. Traffic features and QoS requirements

For the IoT applications, the traffics can be divided into three types: *periodic traffic*, *event-driven traffic*, and *query-based traffic* [22]. Periodic traffic (such as smart metering, traffic sensing, and e-health monitoring, etc.) needs to periodically

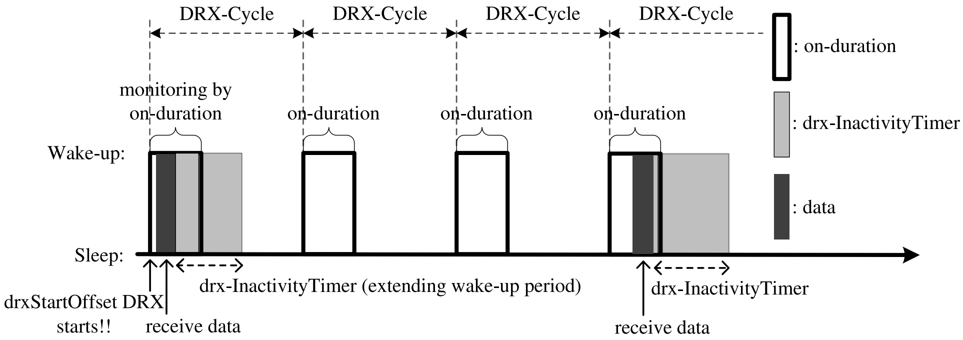


Fig. 2. Overview of the DRX operation.

upload the sensing data. Thus, it needs to guarantee a minimum uplink data rate with limited transmission delay for QoS purposes. Event-driven traffic (such as disaster detection and home surveillance, etc.) has to ensure a specific request size with the strict delay to guarantee its delivery quality. Query-based traffic (such as requesting the temperature reading, noise decibel, or other environment status remotely) has to ensure a specific request data size and strict delay, which is similar to the event-driven traffic.

3.3. Discontinuous reception (DRX) mechanism

The 3GPP standard defines *Discontinuous Reception (DRX)* and *Discontinuous Transmission (DTX)* to support sleep mode, which is controlled by the *Radio Resource Control (RRC)*. The central base station (*eNB*) can initialize DRX mechanism by sending MAC control signal [23]. Each device can operate according to its DRX sleep parameters and all these parameters are configured by eNB. An overview of the DRX operation is shown in Fig. 2. The length of DRX wake-up time and sleep time are in unit of subframes (1 subframe = 1 ms). DRX supports two types of sleep cycles. One is the short cycle; the other is long cycle. Since the behavior of long cycle is similar to that of short cycle, this paper focuses on the short cycle. In DRX, there are four sleep parameters which are used in this work. They are (1) DRX-Cycle, (2) on-duration, (3) drxStartOffset and (4) drx-InactivityTimer. DRX-Cycle is the basic cycle for devices to perform wake-up and sleep operations. On-duration is the least number of wake-up subframes that the device must keep active at the beginning of a DRX cycle in which the device monitors the incoming data. The drxStartOffset is the subframe where the DRX starts. Once the device receives data packets during on-duration, the device will start the drx-InactivityTimer. If the device receives any data packets before the timer expires, the drx-InactivityTimer will be reset; otherwise, the device will go back to sleep once the timer expires.

3.4. Problem definition

In this paper, we consider the uplink transmission for an eNB with N devices and M gateways, where each device D_i, $i = 1..N$ can communicate via a gateway $G_j, j = 1..M$ over two hops or upload its data directly to the eNB (for ease of presentation, we use G_0 to represent the eNB in the following). Note that we assume the related distance between devices, gateways, and the eNB can be estimated according to the *received signal strength (RSS)*. In addition, each device D_i may have *periodic*, *event-driven*, or *query-based* traffic, where periodic traffic has to ensure its admitted data rate R_i (bits/ms) and both the event-driven and query-based traffic have to ensure a specific data request Q_i (bits). Without loss of generality, we assume that there are n devices with periodic traffic (i.e., D_i, $i = 1..n$) and $(N - n)$ devices with the event-driven or query-based traffic (i.e., D_i, $i = (n + 1)..N$), respectively. For QoS guarantee, all periodic, event-driven, and query-based traffic have its transmission delay constraint L_i (ms) to ensure the transmission latency between the devices to the eNB. We consider the LTE-M under *FDD (frequency division duplex)* mode. For the uplink transmission, the radio resource is divided into *subframes* which is with the length of 1 *ms*. The basic resource allocation unit in the uplink subframe is a *resource block (RB)* and each subframe is with fixed number of RBs, denoted by Ω. In each RB, the device with a higher channel quality can carry more data bits to achieve higher channel rate. The channel rate of device D_i to its destination $G_j, j = 0..M$ is denoted as $C_{i,j}$ (bits/RB), which may not change frequently due to the static feature of IoT. Our problem is to ask how to determine the best uplink transmission path for each device $D_i, i = 1..N$ with different types of traffic, and schedule their data efficiently by taking advantage of *data aggregation* and *concurrent transmission* techniques such that the QoS of devices can be guaranteed (including the traffic delay L_i, data rate R_i, and request size Q_i) and more devices can be served while the power consumption of devices can be saved by well determining the sleep parameters such as on-duration (T_i^{on}), DRX-Cycle (T_i), drxStartOffset (O_i) and drx-InactivityTimer (A_i), respectively.

We can summarize the problem as follows.

Definition 1. Given a network $G = (V, E)$, the admitted data rate or request size of each device, traffic delay constraint, channel rates of each device to different gateways, and the available number of RBs per subframe, the *energy-conserved uplink resource allocation (EURA)* problem is to find an interference-free DRX transmission schedule such that the total wake-up periods of devices is minimized.

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

Proof. To simplify the proof, we consider the case where there is no spatial reuse among the device-to-gateway links and the delay constraints of N devices are identical, i.e., $L_i = L_1, i = 2..N$. Thus, the best DRX-Cycle length T_i of device $D_i, i = 1..N$ will be the same as their delay constraint L_i, i.e., $T_i = L_i, i = 1..N$ and the best drx-InactivityTimer A_i of each device D_i will be 0 due to static feature of IoT. In addition, the required number of RBs per cycle for each path is unique and thus the wake-up period of a device on each path is also uniquely determined. Then, the EURA problem can be revised as a decision problem: *Energy-conserved uplink resource allocation decision (EURAD)* problem: Given the network topology $\mathcal{G}$ and the admitted request data of each device, we ask whether or not there exists a path set $\mathcal{R}_p$ on $\mathcal{G}$ such that all devices can reduce the total wake-up periods of $\mathcal{Q}$ to satisfy their requests. Then, we show EURAD problem is NP-complete.

We first show that the EURAD problem belongs to NP. Given a problem instance and a solution containing the path set, it can be verified whether or not the solution is valid in polynomial time. Thus, this part is proved.

We then reduce the *multiple-choice knapsack (MCK)* problem [24], which is known to be NP-complete, to the EURAD problem. Consider that there are N disjointed classes of objects, where each class i contains N_i objects. In each class i, every object $x_{i,j}$ has a profit $q_{i,j}$ and a weight $u_{i,j}$. Besides, there is a knapsack with capacity of $\mathcal{U}$. The MCK problem asks whether or not we can select exact one object from each class such that the total object weight is no larger than $\mathcal{U}$ and the total object profit is $\mathcal{Q}$.

We then construct an instance of the EURAD problem as follows. Let N be the number of devices. Each device D_i has N_i paths to the eNB. When device D_i selects a path $x_{i,j}$, it will conserve wake-up period of $q_{i,j}$[2] and the system should allocate the number of RBs $u_{i,j}$ to transmit device$_i$'s data to the eNB. The total frame space is $T_1 \cdot \Omega = \mathcal{U}$. Our goal is to let all devices conserve the wake-up periods of $\mathcal{Q}$ and satisfy their demands. We show that the MCK problem has a solution if and only if the EURAD problem has a solution.

Suppose that we have a solution to the EURAD problem, which is a path set $\mathcal{R}_p$ with devices' conserved wake-up period and RB allocations. Each device can choose exact one path and we need to assign paths to all devices to satisfy their requests. The total number of RBs cannot exceed the frame space $\mathcal{U}$ and the reduced wake-up period of all devices is $\mathcal{Q}$. By viewing the paths of a device as a class of objects and the frame as the knapsack, the paths in $\mathcal{R}_p$ all constitute a solution to the MCK problem. This proves the *if* part.

Conversely, let $\{x_{1,\alpha_1}, x_{2,\alpha_2}, \ldots\}$ be a solution to the MCK problem. Then, for each device $D_i, i = 1..N$, we select a path such that D_i conserves wake-up period of q_{i,α_i} and the size of allocated RB(s) to transmit D_i's data to the eNB is u_{i,α_i}. In this way, the conserved wake-up periods of all devices will be $\mathcal{Q}$ and the overall RBs size is no larger than $\mathcal{U}$. This constitutes a solution to the EURAD problem, thus proving the *only if* part.

4. The proposed scheme

Since the EURAD problem is NP-complete, finding the optimal solution is impractical due to the time complexity. Thus, we propose the *aggregation-efficient DRX scheduling (AEDS)* scheme. The flowchart of the scheme is shown in Fig. 3. The main idea of our scheme is to leverage the concept of *spatial* and *temporal* data aggregation with DRX scheduling. For the spatial data aggregation, our scheme will classify traffic based on the QoS features and enforce the devices close to each other to connect with the gateways with higher aggregation efficiency. Then, it intends to maximize the *spatial reuse* for the device-to-gateway links so that more data can be transmitted concurrently. On the other hand, for the temporal data aggregation, our scheme adopts a *mix of static and dynamic resource allocation* strategy, which exploits a long-term scheduling approach to aggregate the periodic small data under the constraints of limited delay and exploits a short-term scheduling approach for the event-driven/query-based small data by piggybacking to the long-term scheduling results. Finally, the DRX parameters including the drxStartOffset, on-duration, and DRX-Cycle are assigned together according to the resource allocation.

4.1. Phase1: grouping devices and initializing sleep cycles

Step 1. For each device with periodic traffic, i.e., $D_i, i = 1..n$, we determine the DRX-Cycle T_i for *long-term scheduling* by

$$T_i = min\{L_i | i = 1..n\}. \quad (1)$$

Eq. (1) implies that the sleep cycle $T_i \le L_i$ for each device $D_i, i = 1..n$, which guarantees to meet their delay constraints.

2 Note that the reduced wake-up period of a device's path is compared to the same device's path with the longest wake-up period.

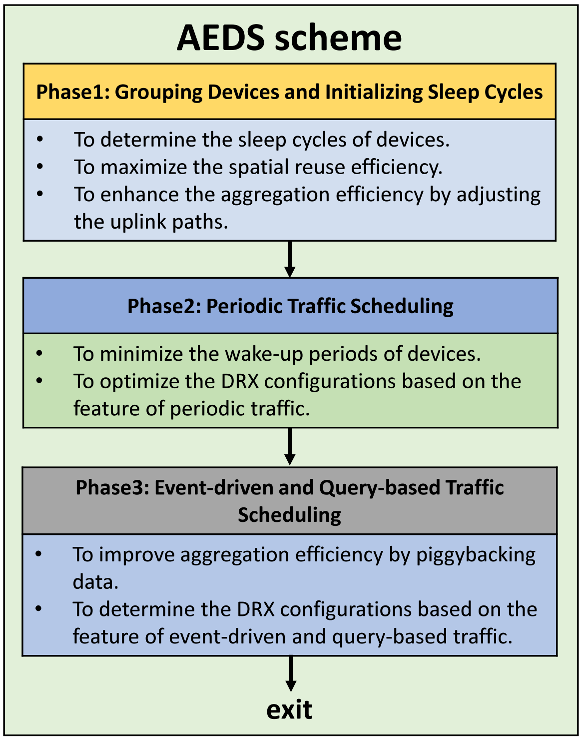


Fig. 3. Flowchart of our proposed scheme.

Step 2. Then, let R be the number of available RBs in the cycle, where $R = T_1 \times \Omega$. Then, we use G^{set} and A^{set} as the candidate set of gateways and the allocated device set; initially, $G^{set} = \phi$ and $A^{set} = \phi$. In the following, we let (f^{avail}, r^{avail}) be the *index of subframe* and *index of RB* in the cycle to represent the starting point of the *available* resource to serve devices; initially, $(f^{avail}, r^{avail}) = (1, 1)$. In addition, several spatial reuse groups would be constructed, denoted as $SP = \{g_m, m = 1, 2, ..\}$, where each group g_m uses (f_m, r_m, ω_m) as its uplink data schedule, where f_m is the subframe index, r_m is the RB index, and ω_m is the maximum required amount of RBs, which is calculated by the amount of data bits over the channel rate (this will be clear later on); initially, $SP = \phi$. Now, we determine the devices'uplink transmission paths and data scheduling results in the following steps.

Step 3. Let $D^{set} = \{D_i, i = 1..n\}$ be the unallocated device set. Then, for each device $D_i \in D^{set}$, we calculate its expected data size Q_i that would arrive during its DRX-Cycle T_i by

$$Q_i = R_i \times T_i.$$

Step 4. Then, for each device $D_i, i = 1..n$, we first assign a *temporary* uplink path P_i as its default transmission path by $P_i = G_{j^*}$, where

$$G_{j^*} = \arg\min_{j=0..M} \{RB^D_{i,j} + RB^G_{i,j}\}. \tag{2}$$

Note that $RB^D_{i,j} = \lceil \frac{Q_i}{C_{i,j}} \rceil$ is the number of required RBs of the first hop (i.e., the link of device-to-gateway, denoted as $link(D_i, P_i)$) and $RB^G_{i,j} = \lceil \frac{Q_i}{C_{j,0}} \rceil$ is the number of required RBs of the second hop (i.e., the link of gateway-to-eNB, denoted as $link(P_i, G_0)$), respectively. Also note that $RB^G_{i,P_i} = 0$ if P_i is the eNB (i.e., $P_i = G_0$).

Step 5. Now, for those devices in D^{set} with $P_i \neq G_0$, we choose the device D_{i^*} with the maximal $RB^D_{i^*,P_{i^*}}$ which can be served within the cycle, i.e.,

$$D_{i^*} = \arg\max_{i \in D^{set}} \{RB^D_{i,P_i} \mid RB^{Total}(D_i, SP, A^{set}, G^{set}) \leq R\}, \tag{3}$$

where

$$RB^{Total}(D_i, SP, A^{set}, G^{set})$$

$$= (RB^D_{i,P_i} + \sum_{g_m \in SP} \omega_m) + GB^{A^{set} \cup D_i}_{G^{set} \cup P_i}$$

is the total number of required RBs to serve $link(D_i, P_i)$ and $link(P_i, G_0)$ for the devices in $A^{set} \cup D_i$ and the gateways in $G^{set} \cup P_i$, respectively. Note that R is the total remaining RBs in the cycle and

$$GB^{A^{set} \cup D_i}_{G^{set} \cup P_i} = \sum_{G_j \in G^{set} \cup P_i} \left\lceil \frac{\sum_{\forall D_{i'} : P_{i'} = G_j, D_{i'} \in A^{set} \cup D_i} Q_{i'}}{C_{P_{i'},0}} \right\rceil \tag{4}$$

is the total required number of RBs of the gateways in $G^{set} \cup P_i$ which helps the devices in $A^{set} \cup D_i$ to relay their data by data aggregation.

If D_{i^*} is found, we create a new spatial reuse group g_m and add D_{i^*} as its member. Then, update $\omega_m = RB^D_{i^*,P_{i^*}}$, $A^{set} = A^{set} \cup D_{i^*}$, and $G^{set} = G^{set} \cup P_{i^*}$. Finally, we remove D_{i^*} from D^{set} accordingly. Then, go to step 6 to find more members for the current spatial reuse group. Otherwise, if no spatial reuse group meets the requirement, for those devices D_i in D^{set} with $P_i = G_0$, we sequentially check if it can be served within the cycle, i.e.,

$$RB^{Total}(D_i, SP, A^{set}, G^{set}) \leq R.$$

If yes, we create a new *singleton* special group $g_m = \{D_i\}$ and then update $\omega_m = RB^D_{i,P_i}$, $A^{set} = A^{set} \cup D_i$ while removing D_i from D^{set} accordingly. After that, go to Phase2 for determining the precise resource index and DRX parameters.

Step 6. Now, we find the possible $link(D_{i^*}, P_{i^*})$ which can concurrently transmit with the new group g_m and can be served within the cycle if $D_{i^*} \in D^{set}$, $P_{i^*} \in G^{set}$, i.e.,

$$(D_{i^*}, P_{i^*}) = \arg \min_{D_i \in D^{set}, D_{i'} \in g_m} \{dst(D_i, D_{i'}) + dst(P_i, P_{i'}) \mid IntfFree(g_m, D_i, P_i)$$
$$\text{and } RB^{Total}(D_i, SP, A^{set}, G^{set}) \leq R\}, \tag{5}$$

where $dst(D_i, D_{i'})$ is the related distance between D_i and $D_{i'}$, and $IntfFree(g_m, D_i, P_i)$ returns 'TRUE' if no interference is incurred between g_m and $link(D_i, P_i)$.

If (D_{i^*}, P_{i^*}) is found, we update $g_m = g_m \cup D_{i^*}$, $G^{set} = G^{set} \cup P_{i^*}$ and add D_{i^*} to A^{set} while removing D_{i^*} from D^{set} accordingly. This step is repeated until no other $link(D_{i^*}, P_{i^*})$ is found. Then, go to step 7.

Step 7. Here, we adjust the temporary paths of the remaining devices in D^{set} if they can transmit their requests via the current candidate gateway $G_{j'} \in G^{set}$ without additional RBs, i.e., updating $P_i = G_{j'}$ if $D_i \in D^{set}$ and

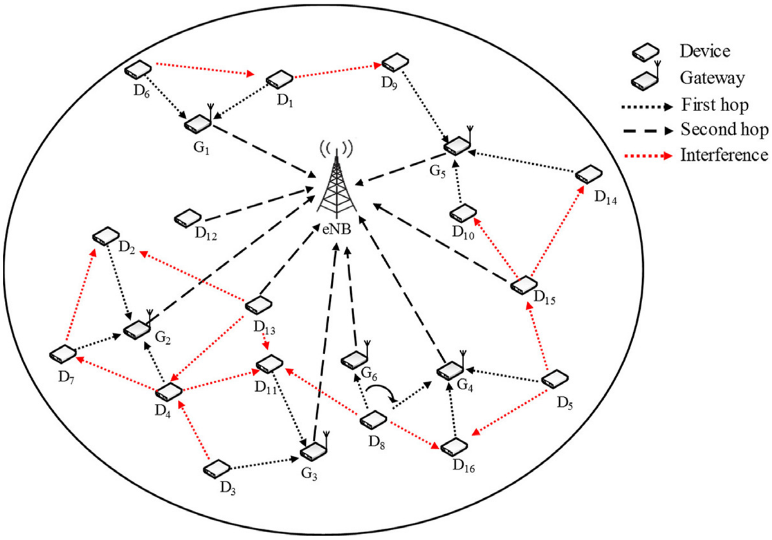
$$(RB^D_{i,P_i} - RB^D_{i,j'}) + (GB^{A^{set} \cup D^{set}}_{\{P_i\}} - GB^{A^{set} \cup D^{set}}_{\{G_{j'}\}}) \leq 0. \tag{6}$$

Note that the left and right parentheses in Eq. (6) are the extra number of RBs of $link(D_i, P_i)$ and $link(P_i, G_0)$ after replacing $P_i = G_{j'}$, respectively. In this way, more devices can leverage the gateways with higher aggregation efficiency to relay data and thus can improve the resource efficiency. Then, go back to step 5 for next group construction.

Below, we give an example for the operation of Phase1. We consider the network with one eNB, $N = 16$ devices and $M = 6$ gateways, where $D_1 \sim D_{14}$ are with periodic traffic and $D_{15} \sim D_{16}$ are with event-driven and query-based traffic, respectively, as shown in Fig. 4. Assume each subframe has $\Omega = 10$ RBs. In Step 1, for the devices with periodic traffic, it determines DRX-Cycle based on the minimum delay constraint by $T_i = \min\{50, 90, 75, .., 90\} = 50$, $i = 1..14$. In Step 2, the number of available RBs in the cycle is set by $R = T_1 \times \Omega = 50 \times 10 = 500$ initially. In Step 3, it uses $D^{set} = \{D_1..D_{14}\}$ as the unallocated device set and calculates the expected data size for each device by $Q_1 = R_1 \times T_1 = 0.8 \times 50 = 40$, $Q_2 = R_2 \times T_2 = 80, .., Q_{14} = R_{14} \times T_{14} = 80$. In Step 4, it assigns a temporary uplink path with the minimum number of required RBs. Thus, we have $P_1 = P_6 = G_1, P_2 = P_4 = P_7 = G_2, P_3 = P_{11} = G_3, P_5 = G_4, P_9 = P_{10} = P_{14} = G_5$, and $P_8 = G_6$, respectively. In Step 5, for the two-hop devices, it choose D_1 which has the maximal number of required RBs at the first hop if the available number of RBs is sufficient. Then, a new spatial reuse group $g_1 = \{D_1\}$ is created; then, it removes D_1 from D^{set}. In Step 6, it finds D_2 which can transmit concurrently with the group g_1. Then, it updates $g_1 = g_1 \cup D_2$ accordingly. In Step 7, no path is adjusted. Similar to D_2, the devices D_3 and D_{10} are added to g_1. After that, we have $g_1 = \{D_1, D_2, D_3, D_{10}\}$, $SP = \{g_1\}$, and the candidate gateway set $G^{set} = \{G_1, G_2, G_3, G_5\}$. For the next round, it selects D_6 as the first member for a new spatial reuse group $g_2 = \{D_6\}$ and finds the devices to achieve parallel transmission similarly. After that, we have $g_2 = \{D_6, D_4, D_5, D_{14}\}$ and $G^{set} = \{G_1, G_2, G_3, G_4, G_5\}$. Note that the path of D_8 is adjusted from G_6 to G_4 because it can transmit request via current candidate gateways without additional RBs. For the third round, the new group $g_3 = \{D_7, D_{11}, D_9, D_8\}$ is created accordingly. Finally, for those devices directly connecting to the eNB, i.e., D_{12} and D_{13}, the corresponding singleton special groups are created by $g_4 = \{D_{12}\}$ and $g_5 = \{D_{13}\}$, respectively.

4.2. Phase2: periodic traffic scheduling

Based on the results of Phase1, Phase2 will allocate resource to each group and determine the drxStartOffset, on-duration, and drx-InactivityTimer for each D_i in the groups. The details is depicted as follows.

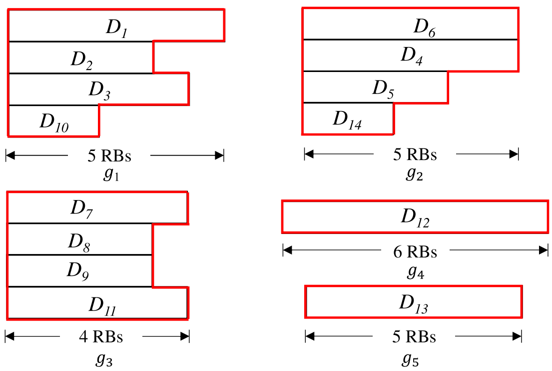


(a) Network topology

D_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
R_i	0.8	0.8	0.8	1.6	2.4	0.8	0.8	1.6	0.8	0.8	1.6	0.8	2.4	0.8	-	-
*Traffic**	P	P	P	P	P	P	P	P	P	P	P	P	P	P	E	Q
L_i	50	90	75	80	60	75	80	100	65	85	55	75	85	90	30	80
T_i	50	50	50	50	50	50	50	50	50	50	50	50	50	50	-	-
Q_i	40	80	60	160	120	60	40	160	60	60	80	60	180	80	60	80
P_i	G_1	G_2	G_3	G_2	G_4	G_1	G_2	$G_6 \rightarrow G_4$	G_5	G_5	G_3	G_0	G_0	G_5	-	-
$RB^D_{i,Pi}$	5	3	4	5	4	5	4	3	3	2	4	6	5	2	4	5

*Traffic type: P: Periodic, E: Event-driven, Q: Query-based

(b) Device table



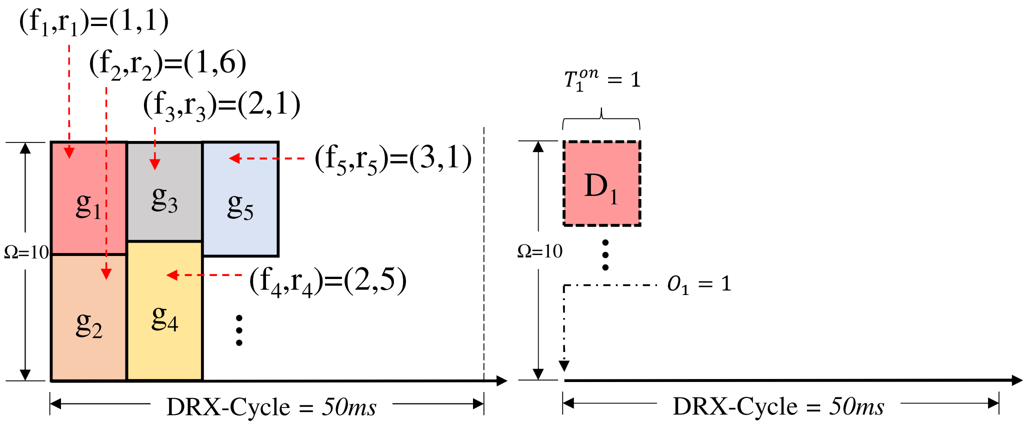
(c) Spatial reuse groups

Initially: $D^{set} = \{D_1..D_{14}\}$, $G^{set} = \emptyset$, $SP=\emptyset$

- Round 1: $SP=\{g_1\}$, $g_1=\{D_1, D_2, D_3, D_{10}\}$, $G^{set} = \{G_1, G_2, G_3, G_5\}$
- Round 2: $SP=\{g_1, g_2\}$, $g_2=\{D_6, D_4, D_5, D_{14}\}$, $G^{set} = \{G_1..G_5\}$
- Round 3: $SP=\{g_1, g_2, g_3\}$, $g_3=\{D_7, D_8, D_9, D_{11}\}$, $G^{set} = \{G_1..G_5\}$
- Round 4: $SP=\{g_1, g_2, g_3, g_4, g_5\}$, $g_4=\{D_{12}\}$, $g_5=\{D_{13}\}$, $G^{set} = \{G_1..G_5\}$

(d) Assignment results

Fig. 4. An example of the operation of Phase 1.



(a) resource allocation for groups $g_1 \sim g_5$ (b) on-duration and drxStartOffset of D_1

Fig. 5. An example of the operation of Phase2.

For each of the group $g_m \in SP$, we allocate the resource to g_m at the resource index $(f_m, r_m) = (f^{avail}, r^{avail})$. Then, set the drxStartOffset O_i for each D_i in g_m by $O_i = f_m$. Next, set the on-duration $T_i^{on} = \lceil (r_m + RB_{i,P_i}^D)/\Omega \rceil$ and set drx-InactivityTimer $A_i = 0$ since the channel rate of devices will not change frequently due to the static feature of IoT. Then, update the corresponding results by $R = R - \omega_m$ and

$$f^{avail} = f^{avail} + \left\lfloor \frac{r^{avail} + \omega_m - 1}{\Omega} \right\rfloor,$$

and $r^{avail} = (r^{avail} + \omega_m)\%\Omega$. This is repeated until all g_m in SP are examined. Similarly, for the $link(P_i, G_0)$ of each g_m, we allocate the resource to G_j at the resource index (f^{avail}, r^{avail}) and then update R, f^{avail} and r^{avail}, respectively. This is repeated until all $G_j \in G^{set}$ are examined. Then, go to next phase for short-term scheduling.

After that, we can make devices obtain precise on-duration (T_i^{on}) and drxStartOffset (O_i), and drx-InactivityTimer (A_i) to achieve better energy efficiency. In the following, we give an example for the operation of Phase2 in Fig. 5, which is based on the previous example results of Fig. 4 in Phase1. For each of the group in $SP = \{g_1, g_2, g_3, g_4, g_5\}$, it allocates the resource to $g_1 \sim g_5$ accordingly and starts from the resource index $(f^{avail}, r^{avail}) = (1, 1)$. Thus, the resource indices of the first three groups (f_m, r_m), $m = 1..3$ are updated by $(f_1, r_1) = (1, 1)$, $(f_2, r_2) = (1, 6)$, and $(f_3, r_3) = (2, 1)$, separately, as shown in Fig. 2(a). Then, it sets the drxStartOffset for each device in the same spatial reuse groups by $O_1 = O_2 = O_3 = O_{10} = 1$ for g_1, $O_6 = O_4 = O_5 = O_{14} = 1$ for g_2, and $O_7 = O_{11} = O_9 = O_8 = 2$ for g_3, separately. Next, it sets the on-duration by $T_1^{on} = T_2^{on} = T_3^{on} = .. = T_{16}^{on} = 1$ accordingly. Then, the corresponding results R and f^{avail} are updated. Similarly, for the devices connecting directly with eNB, i.e., D_{12} and D_{13}, it allocates the resource indices by $(f_4, r_4) = (2, 5)$ and $(f_5, r_5) = (3, 1)$ and sets DRX parameters by $O_{12} = 2$, $T_{12}^{on} = 1$, and $O_{13} = 3$, $T_{13}^{on} = 1$, respectively. Finally, the drx-InactivityTimer for all devices are set by $A_i = 0$, $i = 1..14$.

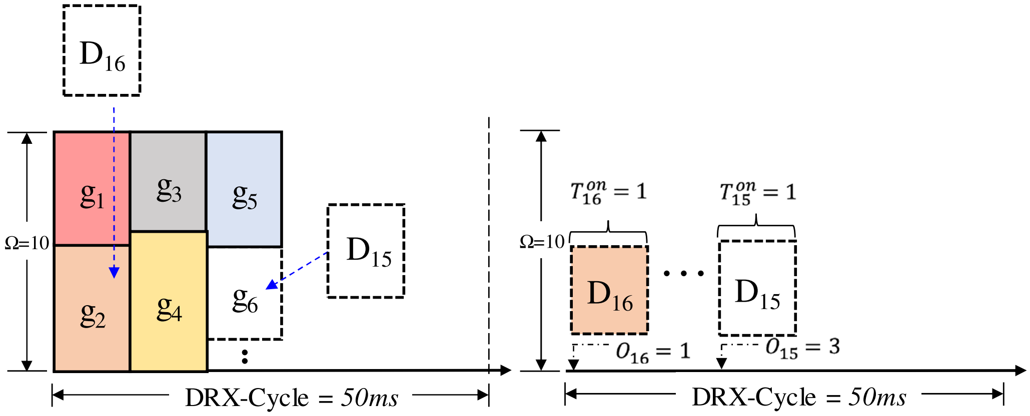
4.3. Phase3: event-driven and query-based traffic scheduling

Consider the device D_i, $i = (n + 1)..N$ with event-driven/query-based traffic, we try to find a potential gateway G_{j^*} in the candidate gateway set G^{set} for D_i which could leverage both the data aggregation and spatial reuse with g_{m^*}, i.e.,

$$(m^*, j^*) = \arg \min_{g_m \in SP, G_j \in G^{set}} \{(m, j) \mid IntfFree(g_m, D_i, G_j) = 0, RB_{i,j}^D \leq \omega_m,$$
$$GB_{\{G_j\}}^{A^{set}} - GB_{\{G_j\}}^{A^{set} \cup D_i} = 0\}. \quad (7)$$

If found, we add D_i to group g_{m^*} and update $P_i = G_{j^*}$. Otherwise, we assign an uplink path $P_i = G_{j'}$ according to Eq. (2), then constructing a new spatial reuse group $g_{new} = \{D_i\}$ and adding it into the SP. Finally, we allocate g_{new} at (f^{avail}, r^{avail}) and update $A^{set} = A^{set} \cup D_i$ and R, f^{avail}, r^{avail}, O_i, T_i^{on}, and A_i accordingly. This phase is repeated until all D_i, $i = (n + 1)..N$ are examined.

Here, we give an example in Fig. 6 to show the operation of Phase3. This example is based on the previous example results in Fig. 4 (Phase1) and Fig. 5 (Phase2), respectively. In this example, it considers the devices with event-driven and query-based traffic, i.e., D_{15} and D_{16}, and tries to find the potential gateways from the candidate gateway set to leverage the data aggregation and spatial reuse. Since D_{16} can transmit through gateway G_4 concurrently with group g_2, it is added to g_2 and set path by $P_{16} = G_4$; the DRX parameters are also set accordingly. Note that the last device D_{15} will cause interference



(a) D_{16} is added to g_2 and g_6 is created for D_{15} (b) on-duration and drxStartOffset of D_{16} and D_{15}

Fig. 6. An example of the operation of Phase2.

Table 1
System parameters.

Simulation parameter	Value
Channel bandwidth	5 MHz
Number of resource blocks (Ω)	25 RBs/subframe
Number of M2M devices (N)	1600 ∼ 7200
Number of M2M gateways (M)	320 ∼ 1440
Path loss model (gateway-to-eNB) [28]	128.1+37.6log(d), d in km.
Path loss model (device-to-gateway) [28]	48.9+40log(d)

to the existing spatial reuse groups, i.e., g_1, g_2, and g_3; thus, it creates a new spatial reuse group $g_6 = \{D_{15}\}$ and then sets $O_{15} = 3$, $T_{15}^{on} = 1$, and $A_{15} = 0$ accordingly.

4.4. Time complexity analysis

We analyze the time complexity of the proposed scheme as follows. In phase one, the step 1 costs $O(N)$ to determine the DRX-Cycle due to N devices. In step 2, it costs $O(1)$ to determine the number of available RBs in the cycle. In step 3, it costs $O(N)$ to calculate the expected data size for each device. In step 4, it costs $O(N \cdot M)$ to determine the temporary uplink path for each device transmitting to all possible gateways. In step 5, it costs $O(N) + O(1)$ to determine the maximal number of required RBs for the first hop and create a new spatial reuse group. In step 6, it costs $O(N)$ to find the possible link for each device with the concurrent transmission group. This operation repeats $O(N)$ times until all links have been determined. Since we have at most N devices to adjust temporary paths to M possible gateways, the time complexity of step 7 is $O(N \cdot M)$. Besides, it will go back to step 5 $O(N)$ times because of N devices. Thus, Phase1 totally costs $O(N)+O(1)+O(N)+O(N\cdot M)+((O(N)+O(1))+O(N^2)+O(N\cdot M))\cdot O(N) = O(N\cdot M+N^3+N^2M)$. In the phase two, it costs $O(N)$ to determine resource index for each reuse group and determine drxStartOffset, on-duration and drx-InactivityTimer for each device. In the phase three, it costs $O(N \cdot M)$ to determine resource index for each event-driven/query-based device with at most M potential gateways.

In brief, the time complexity of the proposed scheme incurred by Phase1, Phase2 and Phase3 is $O(N \cdot M + N^3 + N^2M) + O(N) + O(N \cdot M) = O(N^3)$ if $N \gg M$.

5. Performance evaluation

In the section, we present the simulation results to verify the effectiveness of the proposed scheme. We develop a simulator in C language. The system parameters of the simulator are listed in Table 1. Fifteen channel qualities and several types of IoT applications are considered in the simulation [25–27]. The detailed QoS parameters of the applications are shown in Table 2.

We compare our scheme against three schemes, including ***direct-communication (DC)*** scheme, ***shortest-distance (SD)*** [6], ***shortest-time (ST)*** scheme [7], and ***IoT-aware (IoT-A)*** scheme [20]. The ***DC*** scheme prefers the devices to upload their data directly to the eNB without the help of gateways. The ***SD*** scheme selects the closest gateways of the devices in order to improve the channel quality. The ***ST*** scheme determines the paths with the minimum end-to-end transmission time to improve the transmission efficiency. The ***IoT-A*** uses a mathematical model to analyze DRX parameters and adaptively

Table 2
QoS features of M2M applications [25–27].

Application	Date rate (bps)	Average message size (bits)	Traffic delay (ms)
Traffic light/sensors	1.33	8	150, 300
Roadway signs	0.26	8	100, 150
Home security system	0.26	8	50, 100
Other IoT services	1.0 ∼ 10 k	8 ∼ 16	100, 150, 300

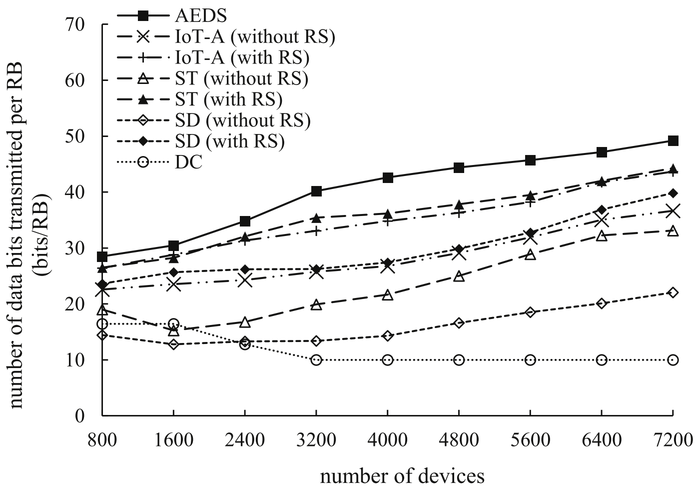


Fig. 7. Comparisons on the number of data bits transmitted per RB (bits/RB) of all schemes.

controls DRX to balance power consumption and delay. Note that we also compare to these schemes with/without resource sharing (RS) [8] to investigate the importance of spatial reuse on performance.

We consider three types of performance metrics: *resource efficiency*, *network capacity* and *power saving*. The resource efficiency includes (1) *the average number of data bits transmitted per RB* and (2) *the average number of RBs allocated per second*. The networks capacity includes (1) *the number of served connections* and (2) *the eNB throughput*. The power saving includes (1) *the average sleep ratio* and (2) *the average power consumption per device*. Note that each experiment is averaged by at least 2000 simulation results.

5.1. Resource efficiency

First, we investigate the impact of the number of IoT devices on the number of data bits transmitted per RB, as shown in Fig. 7. Apparently, except for the DC scheme, the performance of all schemes increases as the number of devices increases. This is because, for the DC scheme, more and more devices are farther away from the eNB and thus their link qualities significantly decrease when they communicate with the eNB directly. Contrarily, the other four schemes have higher probability to transmit data through the gateways when more devices have requests; thus, the resource efficiency increases. Note that our scheme outperforms the other schemes because our scheme can always leverage the temporal and spatial aggregation manners while exploiting the long-term and short-term scheduling for diverse traffic. In this way, the small data can be aggregated more efficiently.

Next, we investigate the impact of the number of IoT devices on the total number of RBs allocated per second. As shown in Fig. 8, the results of all schemes increase as the number of devices increases. This is because more data requests need to be served and thus more RBs are allocated. The DC scheme consumes the most resource because it neglects to utilize the gateways to relay data. The ST scheme performs better than the SD scheme because the former chooses the gateways with higher transmission efficiency. Although the IoT-A scheme uses mathematical models to schedule data, it does not fully leverage the gateways and spatial reuse, which may decrease the performance. Our scheme needs the least resource even when the number of devices is up to 4800. This is because our scheme has higher data aggregation efficiency and thus the allocated resource can be fully utilized.

5.2. Network capacity

Now, we investigate the impact of the number of IoT devices on the eNB throughput, as shown in Fig. 9. We can see that the eNB throughput of all schemes increases as the number of devices increases. The DC scheme has the lowest throughput because the link quality between devices to the eNB is decreasing when more devices are farther away from the eNB.

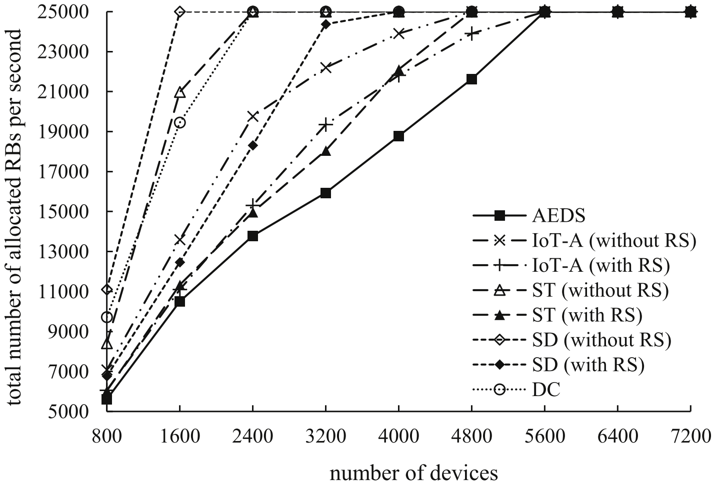


Fig. 8. Comparisons on the total number of RBs allocated per second of all schemes.

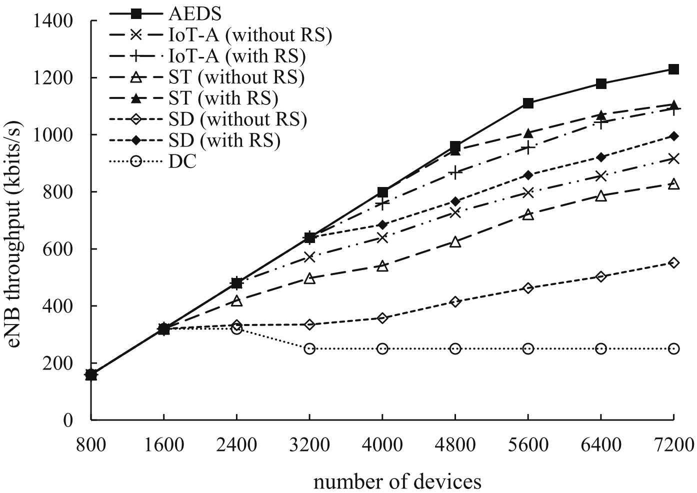


Fig. 9. Comparisons on the eNB throughput (kbps) of all schemes.

Similarly, the ST scheme is better than the SD scheme because the later one determines the path without considering the transmission efficiency. The IoT-A scheme schedules based on mathematical analysis, but it neglects to fully make use of gateways to aggregate data and well leverage spatial reuse. Note that our scheme has the highest eNB throughput because our scheme can well utilize resource by effectively aggregating small data so that more requests can be transmitted. The trend is very similar when we evaluate the impact of the number of IoT devices on the number of served connections.

5.3. Power saving

Next, we investigate the impact of the number of IoT devices on the average sleep ratio. In Fig. 10, we can see that when the number of devices increases, the average sleep ratio decreases because more devices need to wake up for longer time to compete for resource. Note that the value of the proposed scheme is close to 1 when the number of devices is 800. This is because most of devices can wake up with the minimal transmission time without resource competition when the spectrum resource is sufficient. The DC, ST, and SD schemes have the worse average sleep ratio because they do not optimize DRX and make data reception with periodic features, so devices need to keep waking up for possible data delivery. For our proposed scheme, when the number of devices increases, the average sleep ratio decreases. The reason is that our scheme has to satisfy the QoS requirements of devices and thus potentially decreases devices' wake-up periods.

Finally, we investigate the impact of the number of devices on the average power consumption per device. The calculation of the power consumption for the devices is based on Fig. 11. As shown in Fig. 12, we can see that when the number of devices increases, the average power consumption increases. Similarly, the DC, ST, and SD schemes incur the most power

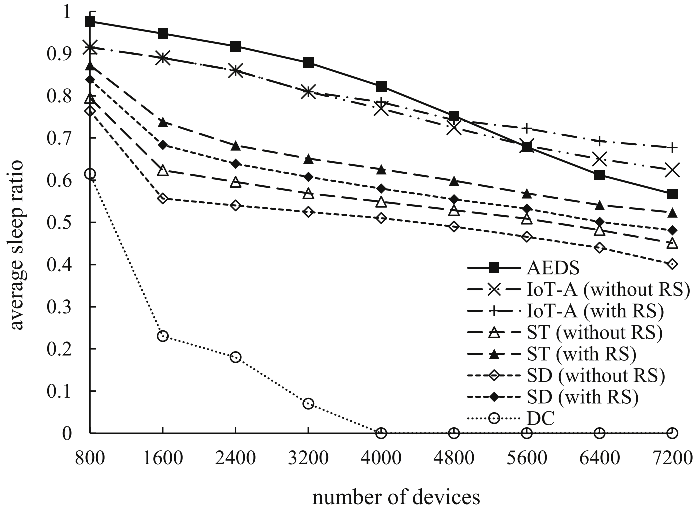


Fig. 10. Comparisons on the average sleep ratio of all schemes.

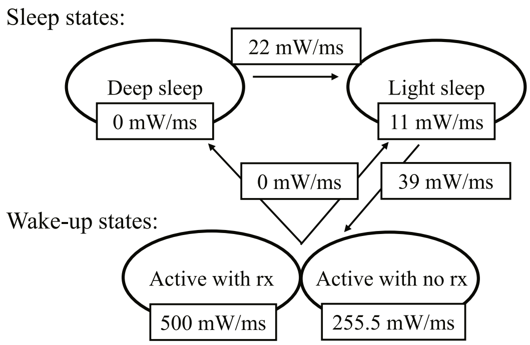


Fig. 11. The power consumption model [29].

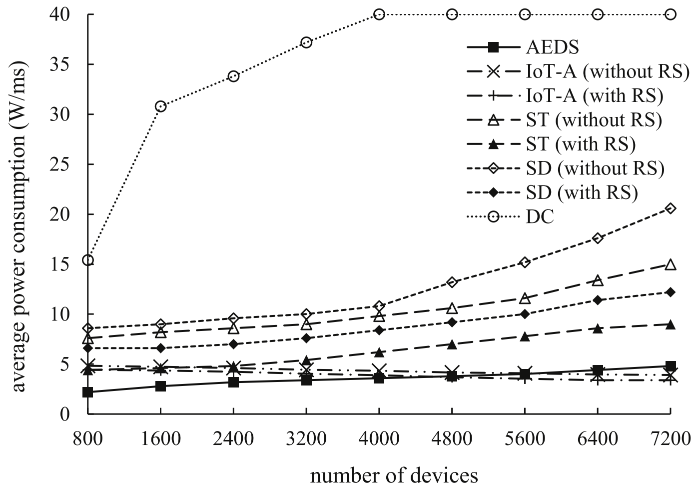


Fig. 12. Comparisons on the average power consumption of all schemes.

consumption because these devices need to keep waking up to wait for possible data delivery. Our scheme has lower power consumption because it well integrates with DRX and thus the data delivery and resource scheduling follow the regular wake-up and sleep patterns. This effectively reduces the wake-up time and power consumption of devices. Note that when

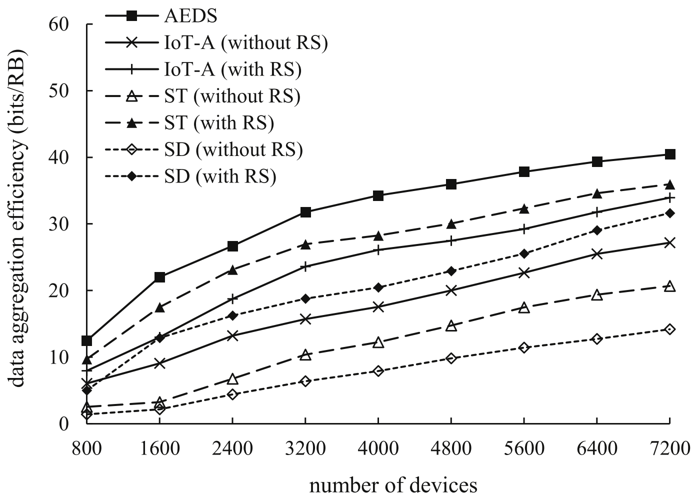


Fig. 13. Comparisons on data aggregation efficiency of all schemes.

the number of devices is greater than 4800, the IoT-A scheme has a slightly lower power consumption than our scheme (15% ∼ 30%) but the proposed scheme can serve more requests and ensure QoS (up to 40%), which are more essential in the system.

5.4. Observations

5.4.1. Data aggregation efficiency

Now, we investigate the impact of the number of IoT devices on data aggregation efficiency, which is evaluated by the number of data bits transmitted per RB on gateways, as shown in Fig. 13. We can see that the increase trend is similar to that in Fig. 7, which represents the transmission efficiency. This implies that the transmission efficiency strongly depends on the data aggregation efficiency, which is one of the important factor for IoT networks.

5.4.2. Traffic latency distribution

Here, we observe the traffic delay distribution of all schemes, where the delay ratio is the traffic delay time over the delay constraint when the number of devices is 4000, as shown in Fig. 14. We can see that all schemes increases as the delay ratio increases. The DC scheme performs the worst because lots of devices are far away from the eNB leading to low link quality. Thus, more resource is consumed and fewer requests can be transmitted within delay constraint. Contrarily, the other four schemes have higher probability to transmit data through the gateways; thus, more requests can be served before delay deadline. Although the IoT-A scheme schedules based on analysis, it neglects to fully leverage gateways which may decrease transmission efficiency and increase requests' delay. Note that the curve of our scheme increases the fastest in all schemes because it can make use of both the temporal and spatial aggregation while exploiting the long-term and short-term scheduling for diverse traffic. Therefore, the small data can be aggregated more efficiently and the delay of most requests can be satisfied.

5.4.3. Satisfaction ratio

Next, we investigate the impact of number of devices on satisfaction ratio, which is defined by the ratio of the amount of satisfied requests to the total amount of requests. When the satisfaction ratio is 1, it means that the scheme can satisfy all devices' requests. In Fig. 15, we can see that the results of all schemes decrease as the number of devices increases due to network saturation. The DC scheme performs the worst because most requests are transmitted by using low channel rate (most devices are far away from the eNB). Contrarily, the other four schemes have higher probability to transmit data through the gateways; thus, the resource is utilized well and more requests can be satisfied. Note that our scheme has the highest satisfaction ratio because it can aggregate small data based on temporal and spatial manners with the long-term and short-term traffic scheduling. Therefore, more small requests can be aggregated efficiently and transmitted successfully.

5.4.4. Path assignment ratio

Finally, we observe the path assignment ratio of all schemes in Fig. 16, where the number of devices is 4000. As can be seen, except for the DC scheme, most schemes assign two-hop paths for devices to transmit data through the gateways. Specifically, our scheme assigns up to 95% of devices with two-hop path; thus, it can fully elaborate the temporal and spatial aggregation manners to achieve higher performance.

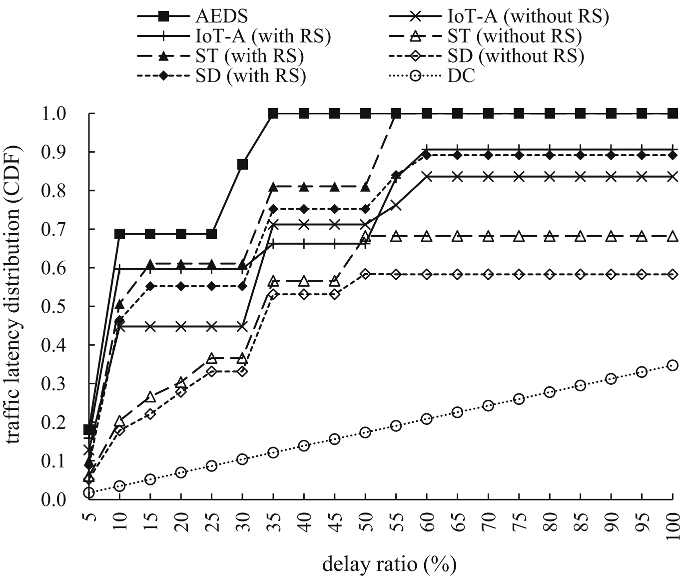


Fig. 14. Comparisons on traffic latency distribution of all schemes.

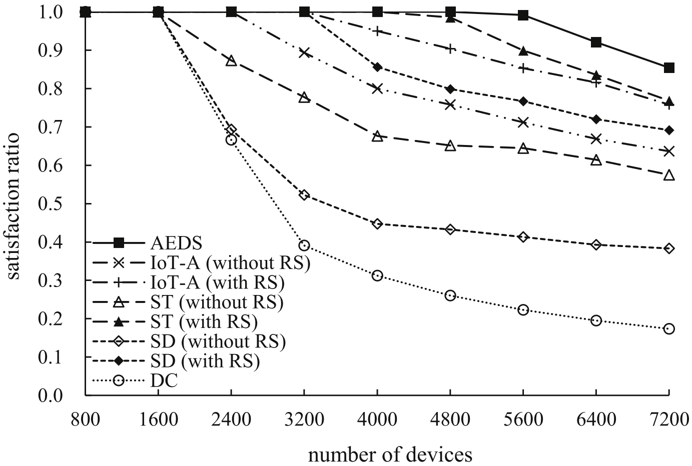


Fig. 15. Comparisons on satisfaction ratio of all schemes.

6. Conclusion

In this paper, we have addressed the DRX scheduling and small data aggregation problems under the consideration of resource efficiency and power saving for IoT communications in LTE-M networks. We have shown this problem to be NP-complete and proposed the AEDS scheme, which consists of three phases. The first phase exploits long-term static scheduling for periodic data to ensure the latency and data rate while minimizing the devices' wake-up time. The second phase exploits DRX technology to further decrease devices' power consumption through optimizing their DRX configurations. The third phase employs short-term dynamic scheduling for event-driven, query-based data to improve transmission efficiency. Extensive simulation results show that the proposed scheme can improve resource efficiency, enlarge network capacity while reducing power consumption compared to the existing schemes.

Acknowledgments

This research is co-sponsored by MOST, Taiwan 106-2221-E-024-004-, 107-2221-E-024-001-MY3, 105-2221-E-182-051, 106-2221-E-182-015-MY3, 105-2745-8-182-001, 106-2221-E-024-004, 105-2221-E-009-100-MY3, 105-2218-E-009-029, 105-2923-E-009-001-MY2, 104-2221-E-009-113-MY3, MoE ATU Plan, Taiwan, Delta Electronics, Taiwan, ITRI, Taiwan, Institute for Information Industry, Taiwan, Academia Sinica, Taiwan AS-105-TP-A07, and Chang Gung Memorial Hospital, Taoyuan, Taiwan.

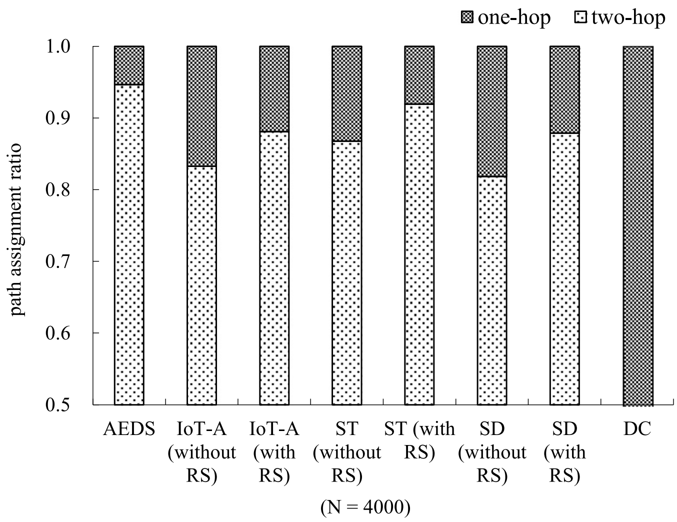


Fig. 16. Comparisons on path assignment ratio of all schemes (N = 4000).

References

[1] P.Y. Chang, J.M. Liang, J.J. Chen, K.R. Wu, Y.C. Tseng, Spatial and temporal aggregation for small and massive transmissions in LTE-M Networks, IEEE Wirel. Commun. Netw. Conf. (WCNC) (2017) 1–5.

[2] K. Zheng, F. Hu, W. Wang, W. Xiang, M. Dohler, Radio resource allocation in LTE-advanced cellular networks with M2M communications, IEEE Commun. Mag. 50 (7) (2012) 184–192.

[3] S.N.K. Marwat, Y. Zaki, et al., A novel machine-to-machine traffic multiplexing in LTE-A system using wireless in-band relaying, Mob. Netw. Manag. Lect. Notes 125 (2013) 149–158.

[4] L.1. Greenstein J. Luo, A.M. Haimovich, Link Failure Probabilities for Practical Cooperative Relay Networks, IEEE Veh. Technol. Conf. (VTC) (2005) 1489–1493.

[5] A.S. Lioumpas G.A. Elkheir, A. Alexiou, Energy Efficient AF Relaying under Error Performance Constraints with Application to M2M Networks, Int. Symp. Pers. Indoor and Mob. Radio Commun. (PIMRC) (2011) 56–60.

[6] D. Zhang, G. Li, K. Zheng, X. Ming, Z.H. Pan, An energy-balanced routing method based on forward-aware factor for wireless sensor networks, IEEE Trans. Ind. Inf. 10 (1) (2014) 766–773.

[7] C. Zhou H. Li, W. Zhuang, Routing metrics for minimizing end-to-end delay in multiradio multichannel wireless networks, IEEE Trans. Parallel Distrib. Syst. 24 (11) (2013) 2293–2303.

[8] Y. Mehmood, S.N.K. Marwat, C. Gorg, Y. Zaki, A. Timm-Giel, Evaluation of M2M data traffic aggregation in LTE-A Uplink, ITG Mob. Commun. Conf. (2015).

[9] C.W. Chang, J.C. Chen, UM Paging: Unified M2M Paging with Optimal DRX Cycle, IEEE Trans. Mob. Comput. 16 (3) (2017) 886–900.

[10] N.M. Balasubramanya, G. Vos L. Lampe, S. Bennett, On timing reacquisition and enhanced primary synchronization signal (ePSS) design for energy efficient 3GPP LTE MTC, IEEE Trans. Mob. Comput. 16 (8) (2017) 2292–2305.

[11] D.P. Van, B.P. Rimal, S. Andreev, T. Tirronen, M. Maier, Machine-to-machine communications over fiwi enhanced LTE networks: A power-saving framework and end-to-end performance, J. Lightwave Technol. 34 (4) (2016) 1062–1071.

[12] N. Mysore Balasubramanya, L. Lampe, G. Vos, S. Bennett, DRX With Quick Sleeping: A Novel Mechanism for Energy-Efficient IoT Using LTE/LTE-A, IEEE Int. Things J. 3 (3) (2016) 398–407.

[13] S.C. Jha, A.T. Koc, R. Vannithamby, Device power saving mechanisms for low cost MTC over LTE networks, IEEE International Conference on Communications (ICC) (2014) 412–417.

[14] H. Xu, H. Tian, B. Huang, P. Zhang, An improved dynamic user equipment power saving mechanism for lte system and performance analysis, Sci. China Inf. Sci. 53 (10) (2010) 2075–2086.

[15] K.T. Feng, W.C. Su, Y.P. Yu, Design and analysis of traffic-based discontinuous reception operations for LTE systems, IEEE Transactions on Wireless Communications 16 (12) (2017) 8235–8249.

[16] L.P. Tung, Y.D. Lin, Y.H. Kuo, Y.C. Lai, K.M. Sivalingam, Reducing power consumption in LTE data scheduling with the constraints of channel condition and QoS, Comput. Netw. 75 (Part A) (2014) 149–159.

[17] C.-W. Chang, J.-C. Chen, UM Paging: unified M2M paging with optimal DRX cycle, IEEE Trans. Mob. Comput. 15 (3) (2017) 886–900.

[18] M. Lee, T. Lee, Energy harvesting discontinuous reception (DRX) mechanism in wireless powered cellular networks, IET Commun. 11 (14) (2017) 2206–2213.

[19] G. Tsoukaneri, M.K. Marina, On device grouping for efficient multicast communications in narrowband-IoT, Int. Conf. Distrib. Comput. Syst. (ICDCS) (2018) 1442–1447.

[20] X. Wang, M. Sheng, Y. Lou, Y. Shih, M. Chiang, Internet of things session management over LTE-balancing signal load, power, and delay, IEEE Int. Things J. 3 (3) (2016) 339–353.

[21] S. Xu, Y. Liu, W. Zhang, Grouping-based discontinuous reception for massive narrowband internet of things systems, IEEE Int. Things J. 5 (3) (2018) 1561–1571.

[22] S.K. Lee, J.G. Koh, C.R. Jung, An energy-efficient qos-aware routing algorithm for wireless multimedia sensor networks, Int. J. Multimedia Ubiquitous Eng. 9 (2) (2014) 245–252.

[23] 3GPP TS 36.321, E-UTRA; Medium Access Control (MAC) Protocol Specification.

[24] H. Kellerer, U. Pferschy, D. Pisinger, Knapsack Problems, Springer-Verlag (2004).

[25] J.M. Liang, J.J. Chen, H.H. Cheng, Y.C. Tseng, An energy-efficient sleep scheduling with qos consideration in 3GPP LTE-advanced networks for internet of things, IEEE J. Emerg. Sel. Topics Circuits Syst. 3 (1) (2013) 13–22.
[26] L.S. Kapov, M. Matijasevic, Analysis of QoS requirements for E-health services and mapping to evolved packet system QoS classes, Int. J. Telemed. Appl. 2 (2010) 1–18.
[27] R. Ratasuk, J. Tan, A. Ghosh, Coverage and Capacity Analysis for Machine Type Communications in LTE, IEEE Veh. Technol. Conf. (VTC) (2012) 1–5.
[28] A. Aijaz, M. Tshangini, M.R. Nakhai, X. Chu, A.-H. Aghvami, Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical QoS guarantees, IEEE Trans. Commun. 62 (7) (2014) 2353–2365.
[29] 3GPP TSG RAN WG2 LTE contribution, R2-071285DRX Parameters in LTE.