Online Energy Management in IoT Applications

Ali Sehati and Majid Ghaderi
Department of Computer Science, University of Calgary
{asehati, mghaderi}@ucalgary.ca

***Abstract*—This paper considers energy management on LTE-enabled Internet of Things (IoT) devices. A characteristic feature of IoT applications is the periodic generation of small messages, whose transmission over LTE is highly energy inefficient. In this paper, we consider application message bundling to alleviate the effect of short message transmissions on energy consumption. Specifically, we model the interplay between energy consumption and the extended DRX mechanism introduced in LTE to deal with IoT traffic. We formulate bundling as a cost minimization problem and develop an online algorithm to solve the problem. Detailed analysis shows that, depending on DRX and application parameters, our algorithm is 1, 2, or 4-competitive with respect to the optimal offline algorithm that knows the entire sequence of application messages a priori. We evaluate the performance of the proposed algorithm and the accuracy of our analysis in a range of realistic scenarios using both model-driven simulations and real experiments on an IoT testbed. Our results show that, i) depending on application requirements, energy savings ranging from zero to about 100% can be achieved using our algorithm, and ii) ignoring DRX could significantly overestimate or underestimate energy consumption.**

I. Introduction

A. Background and Motivation

In recent years, an ever increasing number of smart devices with sensing and communication capabilities has given rise to the Internet of Things (IoT). IoT envisions a world where a variety of smart objects are connected to the Internet and communicate with each other. Currently, 4G cellular networks based on LTE are widely deployed around the world, which makes LTE a natural candidate for IoT connectivity [1]. LTE, however, was designed for high data rate applications to respond to the growing traffic demand of smartphones. It is not optimized for low data rate and low power applications that are envisioned for IoT. A major drawback of LTE is that it consumes a lot of power. Many IoT devices run on battery, and often the cost of replacing batteries is a major operational expenditure.

A characteristic feature of IoT applications is the periodic generation of small messages [2]. It is well-known that periodic transmission of small messages over LTE is highly energy inefficient [3]. Specifically, to transmit a message, the LTE radio has to switch to a high power active mode from the idle mode. To avoid frequent switching, and consequently reduce the network signalling load, once the LTE radio switches to the active mode, it lingers in that mode for some *tail time* even after the transmission of the message is completed. As a result, every time a small message is transmitted, an additional tail energy is consumed by the LTE radio. In contrast to large messages on smartphones, when transmitting small messages on IoT devices, the tail energy is significant compared to the energy consumed for transmitting the message itself.

Recently, 3GPP has proposed a number of LTE enhancements for low-power wide area communications including Machine Type Communication in LTE-M specifications [4]. To reduce power consumption, LTE-M includes several power saving mechanisms such as the extended Discontinuous Reception (DRX) [3]. Specifically, in the radio resource control (RRC) protocol for LTE networks [5], RRC_IDLE (or idle state) represents the lowest energy state and sending or receiving a message in this state will cause a promotion to the RRC_CONNECTED state. Once promoted, the user equipment (UE) enters the Continuous Reception mode (hereafter called *active* state) and consumes high power as it continuously monitors the physical downlink control channel (PDCCH) for scheduling information. The UE also starts an *inactivity timer*, T_i, in this mode which gets restarted every time a message transfer request is scheduled before the timer is expired. Otherwise, the expiry of the timer moves the UE to DRX mode. In DRX mode, the UE periodically wakes up to monitor PDCCH only for short intervals (referred to as ON durations) and then goes to sleep at other times. As a result, power consumption in DRX mode is higher than the idle state but is lower than the active state. Along with T_i, the UE also starts a timer called RRC *tail timer*, T_t, every time it enters the RRC_CONNECTED state. When there is no network activity for the duration of tail timer, the UE moves to the idle state.

Although the proposed power saving mechanisms are effective in reducing LTE radio energy consumption, they are oblivious to the specific requirements of different IoT applications in terms of delay and energy. In particular, while different IoT applications have different delay requirements, most are generally delay-tolerant [2]. As such, it makes sense to *bundle* multiple message transfer requests together and grant them later at once instead of granting individual requests immediately upon their arrivals [6]–[10], specially in scenarios when an IoT device aggregates sensor readings from multiple sensors.

Clearly, bundling reduces radio energy consumption as it consolidates multiple tail energies into one tail energy. The side effect of bundling is the increased delay experienced by applications. The challenge is to design a flexible bundling algorithm to allow applications to trade increased delay for reduced energy consumption. Recently, there have been several works on bundling for smartphones [6], [8], [10]. However, these works focus on traditional 3G networks and consider an *On/Off radio model* to characterize the radio energy consumption of a smartphone. While such an On/Off model is suitable

in 3G cellular networks, it does not accurately represent the radio behavior in LTE networks. In particular, the On/Off model does not capture the effect of DRX, which would lead to significant overestimation or underestimation of the radio energy consumption (this will be shown in Section V), and hence, the inefficient management of the IoT device energy.

B. Our Work

To mitigate the energy inefficiency resulting from small IoT messages, we design a message bundling algorithm with provable performance that is tailored to the specific operation of LTE radios and the DRX mechanism. The difficulty in designing a bundling algorithm is that the bundling decisions have to be made online without knowing the timing of future message transfer requests. For instance, many IoT applications generate messages in an event-based manner with no predictable schedule [11]. To this end, we consider the problem of balancing energy and delay and formulate it as an *online* optimization problem. The objective of the problem is to minimize the bundling cost defined as a weighted summation of energy and delay costs. Energy cost is modeled based on the behavior of the LTE radio as described earlier. Preference over delay versus energy is controlled by including a weight factor in the objective function. Such a weight factor can be used to balance the energy-delay tradeoff based on the IoT traffic type (delay tolerant or delay sensitive) and also power constraints of the IoT device.

To solve the optimization problem, we develop an *online* deterministic bundling algorithm called Energy Optimizer (EO). EO's design is motivated by the classic Ski-rental problem [12]. Specifically, EO does not automatically grant each message transfer request upon its arrival. Rather, it bundles them together and makes a single grant when the energy cost and weighted delay cost associated with that grant become equal. As a benchmark for evaluation of EO, we also design an offline algorithm that relies on the *unrealistic* assumption of knowing the entire arrival times of message transfer requests in advance. We present a detailed analysis of the EO's performance using the well-known notion of *competitive ratio* (CR). We prove that, depending on DRX and application parameters, EO achieves a competitive ratio of 1, 2, or 4 compared to the optimal offline algorithm. It is worth noting that while energy management in IoT applications was the main motivation for our work, *the presented online algorithm and its analysis are applicable to any LTE device*.

To assess the performance of EO, we conducted an extensive set of model-driven simulations under a wide variety of realistic conditions. We also collected real traces from an experimental IoT testbed and used them on an Android-based LTE smartphone to empirically evaluate our online algorithm. Our results show that in most realistic scenarios, EO achieves an empirical competitive ratio less than 2. Also, depending on application requirements, energy savings ranging from zero to about 100% can be achieved using our algorithm. We also observe that DRX has a significant effect on energy consumption, which is neglected by the existing On/Off models.

C. Related Work

There is a large body of work on DRX-aware radio energy management schemes. The most relevant categories related to our work include:

1) DRX optimization: There have been several studies on improving the energy efficiency of IoT devices by optimal configuration of DRX parameters (*e.g.*, see [13], [14], and references therein). The common approach in these works is to model the effect of DRX parameters on energy and delay at the UE side, and then determine the optimal DRX parameters that achieve a desired tradeoff between energy and delay. To model energy and delay, typically modeling assumptions are made about the traffic arrival process and other aspects of the system. For example, [13] developed a Markov model to characterize DRX effect on energy and delay assuming a Poisson traffic arrival. In contrast, our analysis is independent of any specific assumption about the incoming traffic. Also, the problem considered in this paper is orthogonal (and complementary) to the existing work on DRX. In particular, we design our algorithm assuming that the DRX parameters are already configured and set using one of the above optimization models.

2) DRX enhancement: There have also been proposals for modified versions of the DRX. The authors of [15] proposed to enhance the DRX mechanism with a quick sleep indication, where the base station (called eNB) can inform the device to go to sleep when there is no incoming traffic. [16] proposed a packet coalescing mechanism, where the eNB delays transmitting packets to the UEs in DRX mode until their downstream queues reach a tunable threshold. However, these schemes require changes to the operation of current eNBs which could hinder their deployment and adoption.

3) DRX-aware scheduling: Several recent works have proposed scheduling strategies that take DRX operations into account. For example, Liang *et al.* [17] suggested using a DRX selection algorithm along with a cooperating DRX-aware scheduling algorithm at eNB in order to satisfy QoS requirements of IoT applications. However, most IoT traffic is uplink, and hence their algorithm is not sufficient for most IoT applications. Wang *et al.* [11] proposed an IoT device-based uplink scheduler that balances device power consumption and the network signaling load. Their design choice requires end devices cooperating in signal load reduction of the network.

4) Smartphone request bundling: The seminal work addressing energy-delay tradeoff for bundling was TailEnder [6], which is a threshold-based bundling algorithm to reduce the mobile device energy consumption, while satisfying prespecified request deadlines. To avoid relying on a priori knowledge of request deadlines, [8] and [10] proposed online bundling algorithms to jointly minimize energy and delay costs. As mentioned earlier, all these works consider an On/Off radio model which does not capture the operation of the DRX mechanism in LTE networks.

D. Paper Organization

The rest of the paper is organized as follows. In Section II we present a formal specification of the problem. We intro-

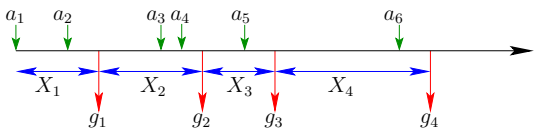


Fig. 1: Relation between arrivals, grants and intervals.

duce our online algorithm in Section III. Then we perform competitive analysis of the online algorithm in Section IV. Performance evaluation results are discussed in Section V. Finally, Section VI concludes the paper.

II. Problem Statement

Consider a sequence of message transfer request arrivals $\mathcal{A} = \langle a_1, \ldots, a_n \rangle$, where a_i denotes the arrival time of request i. *The sequence $\mathcal{A}$ is not known in advance*. Without loss of generality, we assume that the radio is in idle state when the first request arrives. The goal is to design an online algorithm to bundle multiple requests together and grant them at once as opposed to individually granting each request. A message transfer request may involve uploading environmental readings received at an LTE-enabled IoT device from multiple IoT sensors. The LTE-enabled IoT device may communicate with its sensors using short-range low-power wireless technologies such as Bluetooth Low Energy (BLE) [11]. Let $\mathcal{G}_A = \langle g_1, \ldots, g_k \rangle$ denote the sequence of grants made by some algorithm A, for the arrival sequence $\mathcal{A}$, where g_i denotes the time of grant i. Let $\mathcal{X}_A = \{X_1, \ldots, X_k\}$ denote the set of all grant intervals of algorithm A, where $X_1 = [a_1, g_1]$ and $X_i = (g_{i-1}, g_i]$, for $i \geq 2$. All requests that arrive during the interval X_i are bundled together and granted at time g_i. Throughout the paper, we use the notation X_i to refer to the i-th grant interval as well as the length of that interval, when there is no ambiguity.

Fig. 1 shows the relation between arrivals and grants. The objective of the bundling algorithm is to determine the grant times g_i that minimize the cost $C_A = E_A + \alpha D_A$, where E_A and D_A denote the *energy cost* and *delay cost* of algorithm A, respectively. The coefficient α is a control parameter that can be used to specify the relative importance of delay cost over energy cost depending on the IoT application requirements.

A. Energy Cost

The energy cost E_A is the tail energy consumed because of inactivity periods between grants of algorithm A. Let P_C and P_D denote the base powers consumed by the radio during the active and DRX substates of the RRC_Connected state, respectively, where $P_C > P_D$. Also, let T_i denote the length of the inactivity timer in active state and T_t denote the overall RRC tail time, where $T_t > T_i$. Similar to [18], [19], we use the following function to characterize the tail energy:

$$\varepsilon(\tau) = \begin{cases} P_C \tau & 0 \leq \tau \leq T_i, \\ P_C T_i + P_D(\tau - T_i) & T_i < \tau \leq T_t, \\ P_C T_i + P_D(T_t - T_i) & T_t < \tau, \end{cases} \quad (1)$$

where, τ is the time passed since the last grant of the algorithm. Then, the energy cost of grant interval X_i is given by $E_A(X_i) = \varepsilon(X_i)$. Consequently, the energy cost of the algorithm A is given by,

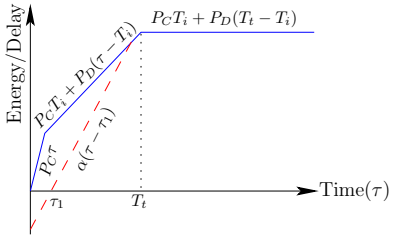


Fig. 2: Intersection of energy and delay functions.

$$E_A = \sum_{X_i \in \mathcal{X}_A} E_A(X_i) + \varepsilon(T_t), \quad (2)$$

where the additional term $\varepsilon(T_t)$ is added to account for a tail time after the last grant of the algorithm. To simplify the analysis, similar to [8], [20], we ignore the transfer time of bundles as this time is the same for every bundling algorithm.

B. Delay Cost

The delay cost of the algorithm is defined as the sum of delay costs of all the bundles. We use the notation $D_A(X_i)$ to denote the delay cost of bundle i, which includes all requests that arrive during interval X_i. Consider a request $a_j \in X_i$. The delay cost of request a_j is given by $(g_i - a_j)$. The delay cost of bundle i is then expressed as $D_A(X_i) = \max_{a_j \in X_i}(g_i - a_j)$. In other words, the delay cost of a bundle is the maximum of all the delays of the requests in the bundle [20]. Equivalently, the delay cost of bundle i is given by $g_i - a_{first,i}$, where $a_{first,i}$ is the arrival time of the first request in bundle i. It then follows that,

$$D_A = \sum_{X_i \in \mathcal{X}_A} D_A(X_i) = \sum_{X_i \in \mathcal{X}_A} \max_{a_j \in X_i} (g_i - a_j). \quad (3)$$

C. Optimal Offline Algorithm

As a point of comparison for our online algorithm, we design an optimal offline algorithm, called OPT. OPT is not a realistic algorithm, since *it knows the entire request arrival sequence in advance*. An important observation used in designing OPT is the fact that the optimal algorithm always makes grants right at the time of some request arrivals and never makes a grant in-between two arrivals. Based on this observation, we design OPT using a dynamic programming algorithm similar to the one in [20]. OPT has the runtime of $O(n^2)$, where n is the length of the arrival sequence. A detailed discussion of the optimal offline algorithm and its analysis can be found in our technical report [21].

Theorem 1. *When $\alpha \geq P_C$,* OPT *makes a grant for every request arrival.*

Proof. See our technical report [21]. ∎

III. Online Energy Management Algorithm

With Theorem 1 characterizing the behavior of the optimal algorithm for $\alpha \geq P_C$, our online algorithm called Energy Optimizer (EO) will imitate OPT in that regime. In other

regimes, EO works as follows. Assume that the most recent grant was at time g_i and the algorithm has to decide when to make its next grant g_{i+1}. Let τ denote the time duration since the last grant g_i. Also, let $E_{\text{EO}}(\tau)$ and $D_{\text{EO}}(\tau)$ denote the energy and delay cost incurred by EO during τ. Then, EO makes a grant at time $g_{i+1} = g_i + \tau$, when $E_{\text{EO}}(\tau) = \alpha D_{\text{EO}}(\tau)$ holds. Fig. 2 portrays a plot illustrating the behavior of EO. In this figure, τ_1 is the time between the last grant of the algorithm (*i.e.*, g_i) and the first arrival after that. This way, the weighted delay cost associated with the grant at $g_i + \tau$ will be $\alpha(\tau - \tau_1)$, which is presented by the dashed line (called D-line). The solid polyline (called E-line) is the graphical presentation of the energy cost function defined in (1). The intersection of these two lines determines the desired τ value.

In the following sections, we will focus on the value of $\tau - \tau_1$ to analyse the performance of EO. If $\alpha \leq P_D$, D-line will intersect horizontal part of the E-line, and hence,

$$\tau - \tau_1 = \frac{P_C T_i + P_D(T_t - T_i)}{\alpha} = \frac{\varepsilon(T_t)}{\alpha}, \qquad \text{if } \alpha \leq P_D\,. \quad (4)$$

If $P_D < \alpha < P_C$, D-line can intersect the middle part or the horizontal part of the E-line. Specifically, we have,

$$\tau - \tau_1 = \begin{cases} f(\tau_1), & \text{if } (P_C - P_D)T_i + \alpha\tau_1 \leq (\alpha - P_D)T_t, \\ \frac{\varepsilon(T_t)}{\alpha}, & \text{otherwise,} \end{cases} \quad (5)$$

where, $f(\tau_1) = \frac{(P_C - P_D)T_i + P_D\tau_1}{\alpha - P_D}$. Notice that the first grant is treated differently, *i.e.*, when there is no g_i. The algorithm makes its first grant at some time τ that satisfies the equation $\varepsilon(T_t) = \alpha D_{\text{EO}}(\tau)$.

The algorithm EO can be implemented using timers. Specifically, for the first arrival after each grant g_i, EO sets a timer to time out after w time units, where the value of w can be computed from (4) or (5) depending on the values of α, τ_1, and their relations to the power model parameters. Upon expiry of the timer, a grant will be made and all the pending requests will be granted.

IV. Analysis of the Energy Optimizer Algorithm

As mentioned earlier, we will study the behaviour of EO in three different regimes. In the regime of $\alpha \geq P_C$, EO imitates the behavior of OPT as determined by Theorem 1. As such, EO is 1-competitive when $\alpha \geq P_C$. Therefore, it suffices to analyze EO for the remaining two regimes, namely when $\alpha \leq P_D$ and $P_D < \alpha < P_C$. In the sequel, we focus on proving the following theorems.

Theorem 2. *When $\alpha \leq P_D$, the Energy Optimizer algorithm is 2-competitive.*

Theorem 3. *When $P_D < \alpha < P_C$, the Energy Optimizer algorithm is 4-competitive.*

A. Preliminaries

The following observations will be used in our analysis.

Observation 1. *At least one request arrives during an interval $X_i = (g_{i-1}, g_i]$. If there is no arrival, then the algorithm does not make any grant at g_i as there is no request to grant.*

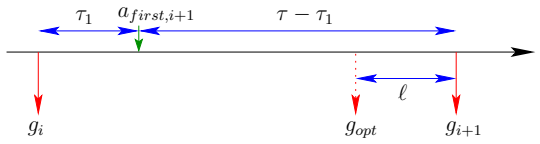


Fig. 3: $(g_i, g_{i+1}]$ is a sample interval created by EO.

Observation 2. *Energy function $\varepsilon(\tau)$ is a concave piecewise-linear function, where the following relations always hold,*

$$\varepsilon(\tau) \leq P_C\tau, \quad (6)$$
$$\varepsilon(\tau) \leq P_C T_i + P_D(\tau - T_i), \quad (7)$$
$$\varepsilon(\tau) \leq P_C T_i + P_D(T_t - T_i) = \varepsilon(T_t)\,. \quad (8)$$

B. Analysis of a Single Interval

We focus on individual grant intervals created by EO in isolation. Fig. 3 shows one such interval, which we refer to as X. Considering the distance of $(\tau - \tau_1)$ between the first arrival in X and its associated grant, the weighted delay cost incurred by EO will be $\alpha(\tau - \tau_1)$. Since EO makes a grant when $E_{\text{EO}}(X) = \alpha D_{\text{EO}}(X)$, the cost of interval X is,

$$C_{\text{EO}}(X) = 2\alpha D_{\text{EO}}(X) = 2\alpha(\tau - \tau_1) = 2E_{\text{EO}}(X)\,. \quad (9)$$

If OPT does not make any grant in X, it will incur at least a delay cost equal to $D_{\text{EO}}(X)$, and hence $C_{\text{OPT}}(X) \geq \alpha D_{\text{EO}}(Y)$, which establishes $C_{\text{EO}}(X) \leq 2C_{\text{OPT}}(X)$. In case OPT has at least one grant in X, we will use g_{OPT} to refer to its first grant in this interval. Let ℓ denote the time duration from g_{OPT} up to the end of interval X (Fig. 3). We will charge g_{OPT} with $\varepsilon(\ell)$ for its contribution to the energy cost and $\alpha(\tau - \tau_1 - \ell)$ for its contribution to the weighted delay cost. If $\ell > T_t$, then $\varepsilon(\ell) = \varepsilon(T_t) \geq E_{\text{EO}}(X)$, indicating $2C_{\text{OPT}}(X) \geq C_{\text{EO}}(X)$. Otherwise (*i.e.*, if $\ell \leq T_t$), we have,

$$\begin{aligned} C_{\text{OPT}}(X) &\geq \alpha(\tau - \tau_1 - \ell) + \varepsilon(\ell) \\ &= \begin{cases} \alpha(\tau - \tau_1) + (P_C - \alpha)\ell & \text{if } 0 \leq \ell \leq T_i, \\ \alpha(\tau - \tau_1) + (P_C - P_D)T_i & \\ \quad + (P_D - \alpha)\ell & \text{if } T_i < \ell \leq \tau - \tau_1\,. \end{cases} \end{aligned} \quad (10)$$

$C_{\text{OPT}}(X)$ is lower bounded in (10) because it is possible for OPT to have more than one grant in interval X.

Proof of Theorem 2:
When $\alpha \leq P_D$ in (10), ℓ's coefficient in both cases becomes positive (also $P_C - P_D > 0$). As a result, it always holds that $\alpha(\tau - \tau_1) \leq C_{\text{OPT}}(X)$, which implies that $C_{\text{EO}}(X) \leq 2C_{\text{OPT}}(X)$. ∎

Notice that when $P_D < \alpha < P_C$, the coefficient of ℓ in (10) is positive only in the first case. As a result, we have $\alpha(\tau - \tau_1) \leq C_{\text{OPT}}(X)$ only in the first case, and hence $C_{\text{EO}}(X) \leq 2C_{\text{OPT}}(X)$. On the other hand, ℓ's coefficient is negative in the second case of (10). Therefore, the minimum value of the lower bound in (10) is obtained by the maximum possible value for ℓ, which is given by $\min\{\tau - \tau_1, T_t\}$. If $T_t < \tau - \tau_1$, we obtain that

$$C_{\text{OPT}}(X) \geq \alpha(\tau - \tau_1 - T_t) + \varepsilon(T_t) \geq E_{\text{EO}}(X), \quad (11)$$

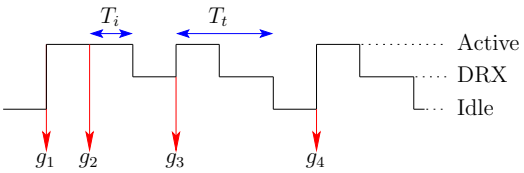


Fig. 4: Grants and radio state transitions of OPT.

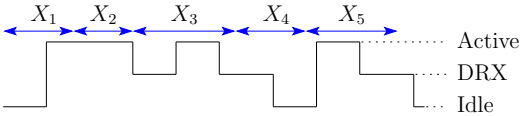


Fig. 5: Grant intervals of EO overlaid on radio states of OPT.

which results in $C_{\text{EO}}(X) \leq 2C_{\text{OPT}}(X)$. However, if $T_t \geq \tau - \tau_1$, we obtain that,

$$C_{\text{OPT}}(X) \geq P_D(\tau - \tau_1) + (P_C - P_D)T_i\,. \tag{12}$$

C. Cost of Grant Intervals

In the previous subsection, we studied EO intervals in isolation. Specifically, in each EO interval we charged EO with the energy cost of the entire interval. In contrast, we charged potential OPT grants only with the energy cost of a portion of the interval (*i.e.*, $\tau - \tau_1$), ignoring energy consumption during τ_1 period. This can result in a pessimistic bound for the competitive ratio. Thus, in this section, we study the entire set of intervals altogether.

We consider $\mathcal{G}_{OPT} = \langle g_1, \ldots, g_l \rangle$ to be the grants made by OPT. Because of these grants, the radio will transition between different states (idle, DRX and active state) several times. Fig. 4 depicts an example consisting of a few OPT grants and radio state transitions resulting from them. In general, the radio is initially in the idle state, then goes through several transitions and finally goes to the idle state after T_t time from the last grant. The main technique used in this section is to overlay grant intervals of EO over the radio states under OPT (see Fig. 5). This way, we can identify two classes of intervals:

1) Intervals with no radio state transition,
2) Intervals with one or more radio state transitions.

The next subsections, analyze these two classes of intervals.

D. Intervals with No Radio Transition

Let X represent one such interval. Considering the radio state during X, the following cases can be identified.

1) *Radio is in active state during interval X:* In this case, we only focus on the energy cost incurred by OPT as its delay cost could be as low as zero. Since the OPT radio is in the active state during interval X, we have $E_{\text{OPT}}(X) = P_C X$. Based on Observation 2, it is obtained that,

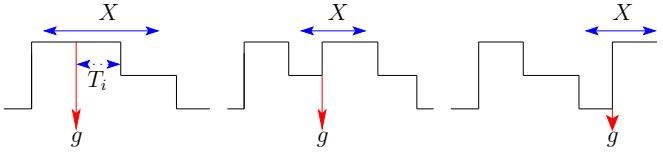
$$C_{\text{OPT}}(X) \geq P_C X \geq E_{\text{EO}}(X)\,. \tag{13}$$

Recall that $C_{\text{EO}}(X) = 2E_{\text{EO}}(X)$, which yields $2C_{\text{OPT}}(X) \geq C_{\text{EO}}(X)$.

2) *Radio is in DRX or idle states during interval X:* Based on Observation 1, the fact that EO makes a grant at the end of interval X implies the arrival of at least one request during X. When the radio for OPT spends the entire interval X in the DRX or idle states, it means that OPT did not make a grant during X, because otherwise it would have transitioned to the active state. Therefore, OPT incurs at least a delay cost equal to $D_{\text{EO}}(X)$. Thus, we have,

$$C_{\text{OPT}}(X) \geq \alpha D_{\text{EO}}(X)\,. \tag{14}$$

Combining (14) with $C_{\text{EO}}(X) = 2\alpha D_{\text{EO}}(X)$ results in $2C_{\text{OPT}}(X) \geq C_{\text{EO}}(X)$.



(a) First grant of OPT happens when the radio is in active state.

(b) First grant of OPT happens when the radio is in DRX state.

(c) First grant of OPT happens when the radio is in idle state.

Fig. 6: Intervals with one or more radio transitions.

E. Intervals with One or More Radio Transitions

Let X refer to one such interval. If OPT does not have a grant in X then $C_{\text{EO}}(X) \leq 2C_{\text{OPT}}(X)$ will hold because OPT will suffer from at least the delay cost $D_{\text{EO}}(X)$. Therefore, in the following, we assume OPT makes at least one grant during interval X.

When the first grant of OPT in X happens, the OPT radio can be in any of the following three possible power states.

1) *The first grant happens when the radio is in active state:* Fig. 6(a) depicts this case. Since OPT makes grants only at request arrival times (and not in-between them), we know that OPT can not have a grant during the period τ_1. Also notice that only a grant can cause the radio to transition to the active state (in contrast to the expiry of an RRC timer and state demotion). As such, the radio of OPT should be in the active state from the beginning of X until the first OPT grant in X (and this time period covers τ_1).
As shown in subsection IV-B, $C_{\text{OPT}}(X)$ and $C_{\text{EO}}(X)$ always satisfy the relation $2C_{\text{OPT}}(X) \geq C_{\text{EO}}(X)$, except in one case. The only case where this relation does not hold leads to the lower bound in (12). Given that this lower bound is obtained without accounting for the radio energy consumption during τ_1, we can adjust the lower bound by considering the fact that the OPT radio will be in the active state during the period τ_1. Therefore, given that $P_C > P_D$, we have,

$$\begin{aligned} C_{\text{OPT}}(X) &\geq P_D(\tau - \tau_1) + (P_C - P_D)T_i + P_D\tau_1, \\ &= P_C T_i + P_D(\tau - T_i), \\ &\geq E_{\text{EO}}(X)\,. \end{aligned} \tag{15}$$

The last inequality comes form (7) in Observation 2 by considering that τ represents the length of interval X. Finally, the relation $C_{\text{EO}}(X) = 2E_{\text{EO}}(X)$ yields $2C_{\text{OPT}}(X) \geq C_{\text{EO}}(X)$.

2) *The first grant happens when the radio is in the DRX state:* Fig. 6(b) depicts this case. This case is similar to the previous case. Notice that the only way the radio

can transition to the DRX state is through the expiry of the RRC inactivity timer and demotion from the active state. In other words, making a grant will always bring the radio to the active state and not the DRX state. Thus, the only difference compared to the previous case is that the OPT radio can be in the active and DRX states during τ_1 but not in the idle state. As a result, the lower bound in (12) can be adjusted by adding $P_D\tau_1$, which means that the relation in (15) still holds. Therefore, we have $2C_{\mathrm{OPT}}(X) \geq C_{\mathrm{EO}}(X)$.

3) *The first grant happens when the radio is in the idle state:* Fig. 6(c) depicts this case. In this case, OPT's first grant (called g_{OPT}) could be right at the arrival time of the first request in the interval, which implies zero delay cost. On the other hand, if g_{OPT} is close to the end of interval X, the energy cost incurred by OPT during interval X could also be zero. As a result, the cost of $C_{\mathrm{OPT}}(X)$ could be as low as zero. Based on Observation 2, the following upper bound is obtained for $C_{\mathrm{EO}}(X)$,

$$C_{\mathrm{EO}}(X) = 2E_{\mathrm{EO}}(X) \leq 2\varepsilon(T_t)\,. \tag{16}$$

Let $\mathcal{X}^0_{\mathrm{EO}} \subseteq \mathcal{X}_{\mathrm{EO}}$ denote the set of all such intervals of EO. Instead of establishing a lower bound for C_{OPT} over individual intervals, we will bound the cost of OPT over the entire set of such intervals (*i.e.*, over $\mathcal{X}^0_{\mathrm{EO}}$). This will be discussed in the proof of Theorem 3 presented next.

Proof of Theorem 3:

For computing C_{OPT} and C_{EO}, we will use the following:

$$C_{\mathrm{OPT}} = \sum_{X_i \in \mathcal{X}_{\mathrm{EO}}} C_{\mathrm{OPT}}(X_i) + \varepsilon(T_t), \tag{17}$$

$$C_{\mathrm{EO}} = \sum_{X_i \in \mathcal{X}_{\mathrm{EO}}} C_{\mathrm{EO}}(X_i) + \varepsilon(T_t)\,. \tag{18}$$

Based on the analysis in previous subsections, every interval X_i in $\mathcal{X}_{\mathrm{EO}}/\mathcal{X}^0_{\mathrm{EO}}$ satisfies $C_{\mathrm{EO}}(X_i) \leq 2C_{\mathrm{OPT}}(X_i)$. Also based on (16), every interval X_i in $\mathcal{X}^0_{\mathrm{EO}}$ satisfies $C_{\mathrm{EO}}(X_i) \leq 2\varepsilon(T_t)$. Therefore, we have,

$$\begin{aligned} C_{\mathrm{EO}} &= \sum_{X_i \in \mathcal{X}_{\mathrm{EO}}} C_{\mathrm{EO}}(X_i) + \varepsilon(T_t) \\ &= \sum_{X_i \in \mathcal{X}_{\mathrm{EO}} \setminus \mathcal{X}^0_{\mathrm{EO}}} C_{\mathrm{EO}}(X_i) + \sum_{X_i \in \mathcal{X}^0_{\mathrm{EO}}} C_{\mathrm{EO}}(X_i) + \varepsilon(T_t) \\ &\leq 2 \sum_{X_i \in \mathcal{X}_{\mathrm{EO}} \setminus \mathcal{X}^0_{\mathrm{EO}}} C_{\mathrm{OPT}}(X_i) + \sum_{X_i \in \mathcal{X}^0_{\mathrm{EO}}} 2\varepsilon(T_t) + \varepsilon(T_t) \\ &\leq 2C_{\mathrm{OPT}} + 2\varepsilon(T_t)\,|\mathcal{X}^0_{\mathrm{EO}}|, \end{aligned} \tag{19}$$

where, $|\mathcal{X}^0_{\mathrm{EO}}|$ denotes the cardinality of set $\mathcal{X}^0_{\mathrm{EO}}$. Thus, our analysis is reduced to establishing an upper bound on $|\mathcal{X}^0_{\mathrm{EO}}|$. To this end, we focus on the behavior of OPT and observe that for the entire arrival sequence, the OPT radio will transition between different states several times. Assume that for $\mathcal{K}$ times, there is a transition from idle to active state. Accordingly, we have $C_{\mathrm{OPT}} \geq \mathcal{K}\varepsilon(T_t)$, because every time the radio goes to the active state, it incurs at least the energy cost of $\varepsilon(T_t)$ before going back to the idle state. Also notice that after overlaying EO intervals over the radio states under OPT, we cannot have more than $\mathcal{K}$ intervals belonging to $\mathcal{X}^0_{\mathrm{EO}}$. Thus, it is obtained that,

$$|\mathcal{X}^0_{\mathrm{EO}}| \leq \mathcal{K} \leq \frac{C_{\mathrm{OPT}}}{\varepsilon(T_t)}\,. \tag{20}$$

By replacing $|\mathcal{X}^0_{\mathrm{EO}}|$ in (19) with its upper bound from (20), the following relation is obtained,

$$\begin{aligned} C_{\mathrm{EO}} &\leq 2C_{\mathrm{OPT}} + 2\varepsilon(T_t)|\mathcal{X}^0_{\mathrm{EO}}|, \\ &\leq 2C_{\mathrm{OPT}} + 2C_{\mathrm{OPT}} = 4C_{\mathrm{OPT}}\,. \end{aligned} \tag{21}$$

■

F. Remarks on the competitive ratio of Theorem 3

We note that the competitive ratio of 4 proved in Theorem 3 is not tight. This can be observed by the discussions in subsection IV-B, where grant intervals were considered in isolation. Specifically, using (9) and (12), the following upper bound for $\frac{C_{\mathrm{EO}}}{C_{\mathrm{OPT}}}$ can be established,

$$\begin{aligned} \frac{C_{\mathrm{EO}}}{C_{\mathrm{OPT}}} &\leq \frac{2\alpha(\tau - \tau_1)}{P_D(\tau - \tau_1) + (P_C - P_D)T_i} \\ &\leq \frac{2\alpha(\tau - \tau_1)}{P_D(\tau - \tau_1)} = \frac{2\alpha}{P_D}\,. \end{aligned} \tag{22}$$

This implies that, for example, when α/P_D is 1.5, the competitive ratio of EO is bounded by 3.

V. Performance Evaluation

In this section, we evaluate EO using both model-driven simulations and real experiments on an IoT testbed. We compare EO with two algorithms: 1) OPT, and 2) *Default*, which grants requests as soon as they arrive. Notice that the delay cost of the Default is always zero.

A. Model-Driven Evaluation

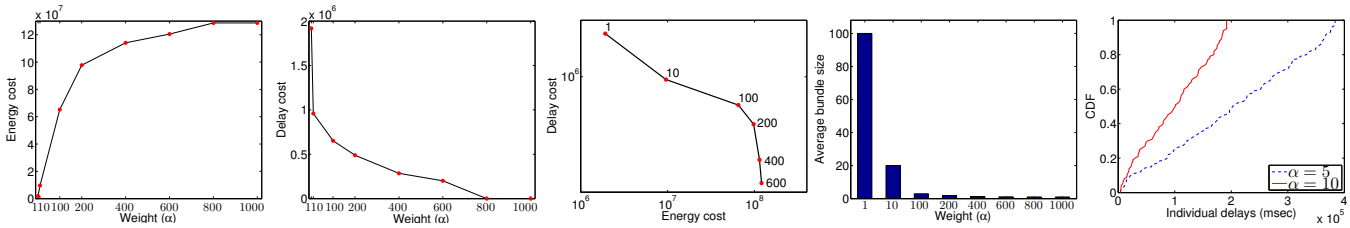
In this part, we study the performance of different algorithms using a custom-developed discrete-event simulator. The simulator takes as input the weight factor α, parameters of the power model (T_i, T_t, P_C, P_D) and the transfer request arrival sequence. Unless otherwise stated, parameters of the power model are set to the values reported in Table I. These values are reported in [22] based on measurements in an LTE network. The DRX base power (P_D) is computed by taking the weighted average of LTE tail base power and power consumption of ON durations in each DRX cycle.

TABLE I: Power model parameters.

State	Power (mW)	Duration (ms)
Active	$P_C = 788$	$T_i = 200$
DRX	$P_D = 163$	$T_t = 11000$

1) Exploring Energy-Delay Tradeoff: We used a sequence of size 100 requests with normal inter-arrival times (μ = 7000 ms, σ = 6000 ms) to perform this experiment. About 2% and 68% of the inter-arrival times in the sequence are less than T_i and T_t, respectively.

Figs. 7(a) and 7(b) show the energy and delay costs for different values of α, respectively. These two plots are combined in Fig. 7(c) which shows the pairwise energy and delay values next to their corresponding weight factors. It is observed that



(a) Energy cost (in micro-Joules). (b) Delay cost (in ms). (c) Energy-delay tradeoff. (d) Average bundle size. (e) CDF of individual delays.

Fig. 7: Performance of EO: By controlling α, different energy-delay tradeoffs can be achieved.

the energy consumed by EO decreases with lower values of the weight α. For example, $\alpha = 1$ results in 98.5% energy saving compared to $\alpha = 800$. Fig. 7(d) which shows the average size of the bundles, illustrates EO's ability to adapt its behavior depending on the weight assigned to the delay cost. Fig. 7(e) presents the cumulative distribution function (CDF) of the delay experienced by the individual requests in the sequence. Notice that in our considered cost function, the delay cost (D_{EO}) is defined as the summation of the maximum delays experienced in each bundle. However, as we can see in Fig. 7(e), D_{EO} is directly related to the delays experienced by individual requests.

The empirical competitive ratios for different values of α are listed in Table II. These results conform the properties claimed in Theorems 1, 2, and 3. Also in the settings characterized by $P_D < \alpha < P_C$, EO exhibits a performance significantly better than the one predicted by the competitive ratio of 4.

TABLE II: Empirical competitive ratio of EO.

α	$C_{\text{EO}}/C_{\text{OPT}}$	α	$C_{\text{EO}}/C_{\text{OPT}}$
1	1.41	400	1.78
10	1.93	600	1.87
100	1.59	800	1
200	1.52	1000	1

2) Performance under Different Arrival Patterns: Similar to [23], here we change the fluctuation level of the inter-arrival times to generate different patterns of request arrivals. In particular, based on the coefficient of variation (CV) of inter-arrival times, we consider arrival sequences of *low* (CV $= 0.5$), *medium* (CV $= 1.5$) and *high* (CV $= 5$) fluctuations. The inter-arrival times are normally distributed with mean 7000 ms.

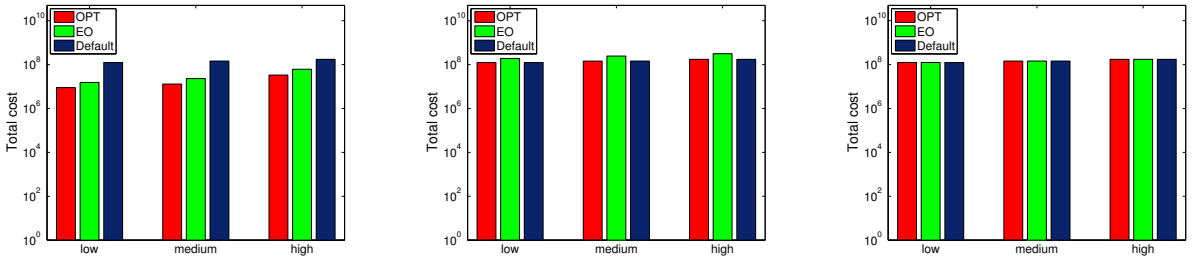
Fig. 8 shows the total cost achieved with the three algorithms under varying fluctuation levels. We consider three weight values corresponding to three regimes identified by Theorems 1, 2, and 3. In Fig. 8(b), the cost of EO is 1.53 and 1.81 times the cost of OPT for sequences with low and high fluctuation, respectively. This implies that for a specific delay weight, the performance of EO changes depending on the characteristics of the arrival sequence. The total costs presented in Fig. 8 also verify our analysis, since the maximum value of $C_{\text{EO}}/C_{\text{OPT}}$ is 1.84 among all the considered scenarios. In scenarios where energy is more important ($\alpha \leq P_D$), EO outperforms the Default algorithm. For example, in a setting with $\alpha = 10$ and high fluctuation, EO results in 64.6% reduction in the total cost compared to Default.

Across all α values, EO's worst performance is achieved when $P_D < \alpha < P_C$. While in this regime EO results in lower energy consumption compared to the Default algorithm, the higher weight assigned to the delay causes the total cost of EO to be higher. In this regime, Default performs better in sequences with long inter-arrival times, where most of the gaps are longer than EO's timer value. In that case, not only EO misses chances of bundling, but also incurs higher cost due to unnecessary waiting. In contrast, EO outperforms Default in sequences with shorter inter-arrival times. For example, in an experiment characterized by $\alpha = 200$ and normal inter-arrival times of mean 400 ms and standard deviation 200 ms, the Default's total cost is 14% higher than EO's cost. As Fig. 8(c) illustrates, in scenarios with high delay importance ($\alpha \geq P_C$), all three algorithms have an identical performance as they grant requests as soon as they arrive.

3) Comparison with On/Off Radio Models: To study the difference between the LTE radio model and an On/Off model, we compare the performance of EO under 3 different power profiles. Specifically, **Profile-1** and **Profile-2** are On/Off radio models, where the radio consumes P_D and P_C for the entire tail period (T_t), respectively. **Profile-3** represents the LTE radio model characterized by parameters P_C, P_D, T_i, T_t.

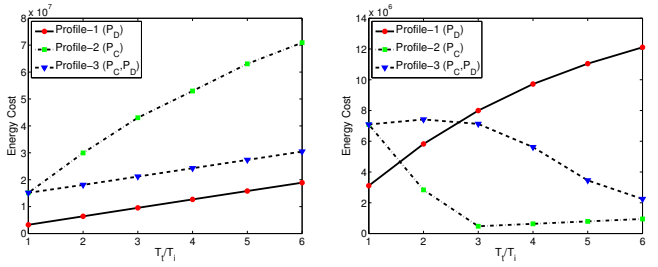
We performed experiments using a sequence of size 100 with normal inter-arrival times ($\mu = 7000$ ms, $\sigma = 6000$ ms). For P_C, P_D and T_i we used values reported in Table I, but we changed T_t between 200 and 1200 ms. For $\alpha = 200$, Fig. 9(a) shows the energy cost of EO under the three power profiles as a function of tail time ratio T_t/T_i (called **TTR**). As seen, for all the TTR values, **Profile-2** and **Profile-1** result in the highest and the lowest energy consumption, respectively.

We also observed a behavior similar to the one in Fig. 9(a) when using input sequences with short inter-arrival times. However, depending on the characteristics of the input sequence, OPT can exhibit a different behaviour. For example, Fig. 9(b) compares the performance of OPT under the three power profiles using an input sequence with short inter-arrival times (normal with $\mu = 700$ ms and $\sigma = 600$ ms). We can see a different ordering between power profiles as **Profile-2** results in the lowest energy consumption. This stems from the fact that OPT can make bundling decisions based on its knowledge about the entire sequence. In particular, when arrival times are close to each other, the delay cost of bundling would be low compared to the reduction in its energy cost. Thus, with an increase in power dissipation rate (**Profile-2**), OPT more aggressively reduces the number of grants resulting in low inter-grant time gaps. In contrast, when inter-arrival times are longer, OPT cannot adopt such an aggressive policy as it



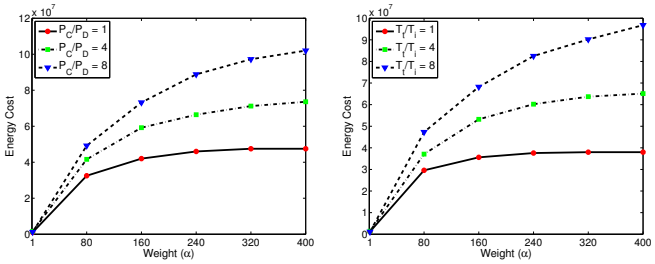
(a) $\alpha = 10$ (b) $\alpha = 200$ (c) $\alpha = 800$

Fig. 8: Comparing the performance of EO with OPT and Default under different fluctuation levels of request inter-arrival times.



(a) EO's energy cost under long inter-arrival times (b) OPT's energy cost under short inter-arrival times

Fig. 9: Energy cost under different power models.



(a) Energy cost for different ratios of $\frac{P_C}{P_D}$. $(P_D, T_i, T_t) = (500, 200, 1000)$. (b) Energy cost for different ratios of $\frac{T_t}{T_i}$. $(P_D, P_C, T_i) = (500, 2000, 200)$.

Fig. 10: Energy cost under different ratios of parameters.

would result in high delay costs. That is why OPT's energy cost follows a similar trend to the one in Fig. 9(a) under sequences with longer inter-arrival times.

4) Effect of Power Model Parameters: In this experiment, we study the effect of power model parameters (P_C and P_D) on the performance of EO. Specifically, we examine the performance of EO under different ratios of P_C/P_D (called power ratio) by using a fixed value for P_D and changing values of P_C. To better capture the effect of power ratio on EO's energy cost, we use the power model parameters characterized by $P_D = 500$ mW, $T_i = 200$ ms, and $T_t = 1000$ ms, which are different than the ones in Table I. We performed experiments using a sequence of 100 requests with normal inter-arrival times ($\mu = 7000$ ms, $\sigma = 6000$ ms). Given the high importance of energy in IoT scenarios, we consider the regime of $\alpha \leq P_D$, where energy is more important than delay.

Fig. 10(a) plots the energy cost of EO under different power ratios as a function of the weight factor α. We can see that increasing the power ratio leads to higher energy consumption. Also notice that the increase in the energy cost becomes more pronounced in settings with higher values of α. This is because with an increase in delay importance, EO tends to avoid bundling and grants requests as soon as they arrive. This will create more inter-grant idle gaps which in turn will highlight the role played by a higher power dissipation rate.

5) Effect of Tail Times: Here we study the effect of timers T_i and T_t on the performance of EO. Specifically, we perform experiments under three different ratios of **TTR** by using a fixed value for T_i and changing values of T_t. As in the previous subsection, we use parameter values of $(P_D, P_C, T_i) =$ (500 mW, 2000 mW, 200 ms) to better represent the impact of timers on the performance of EO. Also, we perform these experiments in the regime of $\alpha \leq P_D$ using the same sequence described in the previous subsection.

Fig. 10(b) depicts the energy cost for different values of TTR as a function of the delay weight (α). As observed, EO's energy cost is influenced by the values of the timers. Specifically, increasing TTR contributes to a higher energy consumption in all cases (α values). For example, in the case of $\alpha = 80$, raising TTR from 1 to 8 leads to 59.8% increase in the energy cost. This is due to the fact that with larger values of T_t, the radio stays longer in the DRX state instead of switching to the idle state. Similar to the previous section, in Fig. 10(b), an increase in the delay weight intensifies the impact of long tail time, which is the result of higher number of grants and longer inter-grant idle gaps.

B. Experiments on IoT Testbed

To assess the performance of EO in real-life conditions, we also performed experiments on Grenoble platform of the FIT IoT-LAB testbed [24]. IoT-LAB is a large scale open testbed for IoT research which provides access to IoT devices with IEEE 802.15.4-based radio transmitters. We created a topology consisting of 30 M3 Open nodes configured with Contiki operating system. Fig. 11 shows the topology used for this experiment. One of the nodes (node 231 in Grenoble platform) was configured to act as a gateway and the rest of the nodes were configured with a program that periodically (every 60 seconds) reads the value of atmospheric pressure from node's sensor and sends it to the gateway over a UDP connection. As mentioned in [25], upon experiment initialization, each node goes through a slightly different setup phase required for establishing the routing-tree structures in the network. This creates a random delay before each node starts generating traffic which is one of the reasons for desynchronization among nodes. For a duration of 20 minutes, we captured radio

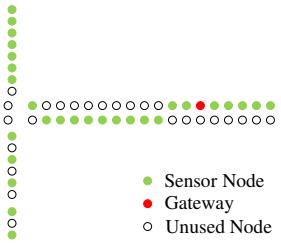
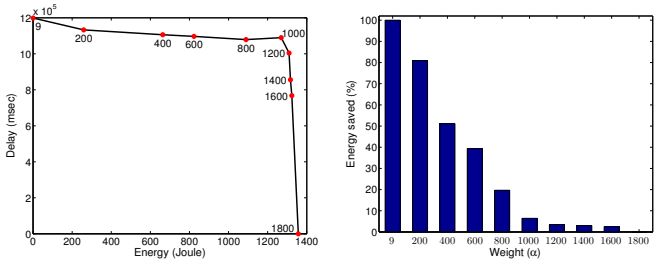


Fig. 11: Topology of the experiment run on IoT-LAB testbed.



(a) Energy-delay tradeoff. (b) Energy savings.

Fig. 12: LTE experiments using IoT trace.

communications at the gateway and created a trace from packet arrival times at the gateway.

Then for different values of α, we run EO and Default algorithms with the collected trace as their input sequence. For each value of α, we recorded the resulting grant times in a separate file. We then fed those grant files to our Android app installed on a Nexus smartphone. This app, which is developed for the purpose of radio energy measurement, performs message transfers at user-specified times. Each *run* of the app uses a grant sequence file as input. It then repeatedly sends message transfer requests at the times specified in the grant file. We also blocked all background traffic from OS services and other apps. To measure the energy consumption of the radio interface, we used the AT&T ARO tool [26], configured with AT&T LTE network parameters [5].

Fig. 12(a) presents the pairwise energy-delay values of EO next to their corresponding α values. Notice that in this figure, the energy cost is expressed in Joules. Here EO exhibits a behavior similar to the one in simulations as it is able to cover the broad spectrum of the energy-delay tradeoff. In this experiment, because of the specific power model parameters of AT&T's LTE network, the maximum energy saving and maximum delay reduction are achieved at $\alpha = 9$ and $\alpha = 1800$, respectively. Also, our experiment with the Default algorithm resulted in zero delay and an energy expenditure of 1356.63 Joules. Fig. 12(b) plots the energy savings of EO compared to the Default algorithm for different values of α. As the relative importance of delay decreases, higher energy savings are achieved. Across all the values of α, EO can achieve energy savings ranging from 0% to 100%.

VI. Conclusion

In this work, we studied the problem of IoT energy management in LTE networks. Based on the specific characteristics of the LTE radio, we proposed an online bundling algorithm that has the flexibility of achieving different energy-delay tradeoffs. We performed competitive analysis of the algorithm and evaluated it using an extensive set of simulations and real experiments. Our results indicate that in realistic scenarios, our algorithm exhibits a performance better than the one implied by the competitive ratio. Design and analysis of a randomized version of our algorithm are possible avenues for future research.

References

[1] Nokia, "LTE evolution for IoT connectivity." [Online]. Available: https://resources.ext.nokia.com/asset/200178

[2] M. Z. Shafiq *et al.*, "A first look at cellular machine-to-machine traffic: large scale measurement and characterization," in *Proc. ACM SIGMETRICS*, 2012.

[3] C. S. Bontu and E. Illidge, "DRX mechanism for power saving in LTE," *IEEE Commun. Mag.*, vol. 47, no. 6, 2009.

[4] R. Ratasuk *et al.*, "Overview of LTE enhancements for cellular IoT," in *Proc. IEEE PIMRC*, 2015.

[5] J. Huang *et al.*, "A close examination of performance and power characteristics of 4G LTE networks," in *Proc. ACM MobiSys*, 2012.

[6] N. Balasubramanian *et al.*, "Energy consumption in mobile phones: a measurement study and implications for network applications," in *Proc. ACM IMC*, 2009.

[7] S. Deng and H. Balakrishnan, "Traffic-aware techniques to reduce 3G/LTE wireless energy consumption," in *Proc. ACM CoNEXT*, 2012.

[8] L. Xiang *et al.*, "Ready, set, go: Coalesced offloading from mobile devices to the cloud," in *Proc. IEEE INFOCOM*, 2014.

[9] B. Zhao *et al.*, "Energy-aware web browsing on smartphones," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 3, 2015.

[10] A. Sehati and M. Ghaderi, "Energy-delay tradeoff for request bundling on smartphones," in *Proc. IEEE INFOCOM*, 2017.

[11] X. Wang *et al.*, "Internet of Things session management over LTE—balancing signal load, power, and delay," *IEEE Internet Things J.*, vol. 3, no. 3, 2016.

[12] A. R. Karlin *et al.*, "Competitive snoopy caching," *Algorithmica*, vol. 3, no. 1-4, 1988.

[13] K. Zhou *et al.*, "LTE/LTE-A discontinuous reception modeling for machine type communications," *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, 2013.

[14] H. Ramazanali and A. Vinel, "Performance evaluation of LTE/LTE-A DRX: A Markovian approach," *IEEE Internet Things J.*, vol. 3, no. 3, 2016.

[15] N. M. Balasubramanya *et al.*, "DRX with quick sleeping: A novel mechanism for energy-efficient IoT using LTE/LTE-A," *IEEE Internet Things J.*, vol. 3, no. 3, 2016.

[16] Herrería-Alonso *et al.*, "Adaptive DRX scheme to improve energy efficiency in LTE networks with bounded delay," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, 2015.

[17] J.-M. Liang *et al.*, "An energy-efficient sleep scheduling with QoS consideration in 3GPP LTE-advanced networks for Internet of Things," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 1, 2013.

[18] Y. Cui *et al.*, "Performance-aware energy optimization on mobile devices in cellular network," *IEEE Trans. Mobile Comput.*, vol. 16, no. 4, 2017.

[19] J. Song *et al.*, "EDASH: Energy-aware QoE optimization for adaptive video delivery over LTE networks," in *Proc. IEEE ICCCN*, 2016.

[20] D. R. Dooly *et al.*, "On-line analysis of the TCP acknowledgment delay problem," *Journal of the ACM*, vol. 48, no. 2, 2001.

[21] A. Sehati and M. Ghaderi, "Online energy management in IoT applications," *Tech. Rep.* [Online]. Available: "https://cpsc.ucalgary.ca/~asehati/docs/infocom18-IoT.pdf"

[22] X. Chen *et al.*, "Smartphone energy drain in the wild: Analysis and implications," in *Proc. ACM SIGMETRICS*, 2015.

[23] W. Wang *et al.*, "Dynamic cloud resource reservation via cloud brokerage," in *Proc. IEEE ICDCS*, 2013.

[24] C. Adjih *et al.*, "FIT IoT-LAB: A large scale open experimental IoT testbed," in *Proc. IEEE WF-IoT*, 2015.

[25] A. Betzler *et al.*, "Experimental evaluation of congestion control for CoAP communications without end-to-end reliability," *Ad Hoc Networks*, vol. 52, 2016.

[26] F. Qian *et al.*, "Profiling resource usage for mobile applications: A cross-layer approach," in *Proc. ACM Mobisys*, 2011.