Azari, A., Miao, G. (2016)
Battery Lifetime-Aware Base Station Sleeping Control with M2M/H2H Coexistence.
In: IEEE conference proceedings

Battery Lifetime-Aware Base Station Sleeping Control with M2M/H2H Coexistence

Amin Azari and Guowang Miao
KTH Royal Institute of Technology
Email: {aazari, guowang}@kth.se

***Abstract*—Fundamental tradeoffs in green cellular networks with coexistence of machine-oriented and human-oriented traffics are investigated. First, we present a queuing system to model the uplink transmission of a green base station which serves two types of distinct traffics with strict requirements on delay and battery lifetime. Then, the energy-lifetime and energy-delay tradeoffs are introduced, and closed-form expressions for energy consumption of the base station, average experienced delay in data transmission, and expected battery lifetime of machine devices are derived. Furthermore, we extend the derived results to the multi-cell scenario, and investigate the impacts of system and traffic parameters on the energy-lifetime and energy-delay tradeoffs using analytical and numerical results. Numerical results show the impact of energy saving for the access network on the introduced tradeoffs, and figure out the ways in which energy could be saved by compromising on the level of performance.**

***Index Terms*—M2M communications, Energy efficiency, Green cellular network, Lifetime, Coexistence, Tradeoff.**

I. INTRODUCTION

During the last decade, energy efficiency is becoming increasingly important for cellular networks to decrease operational costs, and reduce negative environmental footprints [1]. Motivated by the facts that in cellular networks, (i) base stations (BSs) consume almost 60% of the total energy [2], (ii) BS deployment is based on the maximum traffic demand, and (iii) network traffic varies over time, BS sleeping has been investigated in literature to realize substantial reduction of energy consumption [3, 4]. To optimize overall network performance, the impact of sleep mode on other network performance measures has been investigated in several research works [5–7]. For example, the delay and outage constrained energy minimization problems have been investigated in [5, 6] in order to achieve good energy-delay and energy-coverage tradeoffs. For a cellular network with decoupled control and data planes, characteristics of the energy-delay tradeoff have been studied in [7]. The existing research works on energy saving strategies for wireless access networks have mainly focused on the downlink transmission [8]. Besides the existing voice and data services, the next generation of cellular networks is expected to support new communication paradigms such as internet of things (IoT) and machine-type communications (MTC) [9]. MTC, which is also known as cellular-based machine to machine (M2M) communications, means the communications between smart devices without human intervention, and is generally characterized by the massive number of concurrent active devices, small packet payload, and vastly diverse quality-of-service (QoS) requirements [10]. Most smart devices are battery driven and once deployed, their batteries will not be replaced. Then, long battery lifetime is of paramount importance for battery-limited devices, especially for devices in remote areas [11]. Scalable yet energy efficient system design for serving the coexistence of M2M and human-to-human (H2H) communications in cellular networks constitutes an interesting research problem which has attracted lots of interests in recent years [9, 12]. To the best of our knowledge, the energy/lifetime/delay analysis of a system which serves mixed machine- and human-oriented traffics is absent in literature.

In this work, we investigate the tradeoffs between energy saving for the access network, experienced delay in the uplink transmission, and battery lifetime for machine-type devices in order to figure out when and how to trade the tolerable delay/lifetime for energy saving in the access network. The main contributions of this paper include:

- Develop a tractable framework to model the operation of a green BS which serves mixed M2M and H2H traffics. Derive closed-form expressions for energy, delay, and battery lifetime. Introduce the fundamental tradeoffs. Explore the impact of system and traffic parameters on the introduced tradeoffs.
- Extend the proposed framework to the multi-cell scenario, derive closed-form expressions, and investigate the performance impact of the control parameters.

The rest of this paper is organized as follows. System model is presented in the next section. Performance modelings for the single- and multi-cell scenarios are presented in section III and IV respectively. Performance tradeoffs are investigated in section V. Numerical results are presented in section VI. The concluding remarks are given in section VII.

II. SYSTEM MODEL

Consider a single cell with one base station at the center, and a massive number of user equipments (UEs) which are uniformly distributed in the cell. The links between the UEs and the BS experience Rayleigh fading. Unlike the previous works on green cellular network design [4]-[8], which have been focused on the downlink data transmission, we focus on the coexistence of human- and machine-oriented traffic in the uplink direction. Motivations for considering only the uplink direction can be the proposed uplink/downlink decoupling architecture for the 5G [13], and the proposed multi radio access technology (RAT) architecture [14], in which cellular users can be served with LTE and WiFi for uplink and downlink transmissions respectively.

The UEs are heterogeneous in nature such that one can mainly categorize them into two different categories. The first category consists of high-priority devices, called $\mathcal{P}_1$. The second category consists of low-priority devices, called $\mathcal{P}_2$. The $\mathcal{P}_2$ devices are usually battery-limited machine-type devices with strict constraints on transmit power, and battery lifetime; however, they can tolerate delay in data transmission. We assume that $\{\mathcal{P}_n\}_{n=1,2}$ devices arrive according to a Poisson process with arrival rate of λ_n to the system, and $\mathcal{P}_1$ devices have non-preemptive priority over $\mathcal{P}_2$ devices. After serving UEs which are queued to be served, the BS waits for new requests. The listening window is exponentially distributed with rate μ. If no user arrives in the listening window, BS goes to the sleep mode, where the sleep window follows a general distribution with cumulative distribution function (CDF), probability distribution function (PDF), Laplace-Stieltjes transform, and first and second moments as $V(x)$, $v(x)$, $v^*(s)$, $\bar{v}$, and $\hat{v}$, respectively. When the sleeping window expires, the BS wakes up and starts listening for serving potential arriving UEs.

The number of deployed $\mathcal{P}_2$ devices in future cellular networks is expected to become much higher than the existing user equipments [10]. To prevent network congestion due to the surge in the number of connected devices, we utilize an access class barring (ACB) scheme [15]. Based on our ACB scheme, when the BS is asleep or busy, $\mathcal{P}_2$ devices retry after a random backoff time which is exponentially distributed with rate α. The $\mathcal{P}_1$ devices are queued to be served based on a processor sharing service discipline when the BS is busy, and keep listening to find the BS available when the BS is asleep. In the following, the energy consumption of the access network, the experienced delay, and the battery lifetime of UEs are formulated as the key performance indicators (KPIs) of interest.

III. Formulating The KPIs in a Single-Cell Scenario

Here, we focus on the cases where the sleep mode operation doesn't not affect UEs' arrival processes [7]. The cell breathing [1], in which UEs in the switched-off cells are served by the neighboring BSs will be investigated in section IV.

A. Energy Consumption Model for the BS

Denote by P_s, P_{sl}, and P_l the average power consumption in the service, sleep, and idle listening modes respectively. Then, one can compute the average energy consumption of the BS per unit time as:

$$E^b_{cons} = T_sP_s + T_lP_l + \big[1 - T_s - T_l\big]P_{sl} + F_{sw}E_{sw}, \quad (1)$$

where T_s and T_l are the percentages of time the BS spends in the service and listen modes, E_{sw} is the energy consumption for switching BS on/off, and F_{sw} is the mode transition frequency, i.e. the mode transition times between listen and sleep modes per unit time. In order to derive a closed-form expression for E^b_{cons}, we utilize a vacation-based queuing system with mixed-priority users in which, server is subject to traffic-aware interruptions [16]. Towards this end, in the following we investigate the service time of the UEs.

1) Service time: Denote noise power spectral density at the receiver as N_0, and the channel gain between node i and the BS as $g_i = hr_i^{-\sigma}$ where $h \sim \exp(1)$, r_i is the communication distance, and σ is the pathloss exponent. Also, to guarantee a constant received signal to interference and noise ratio (SINR), i.e. γ_n, one can use the Shannon capacity formula and derive the required transmit power of node i as:

$$P_{t_{(i,n)}} = \gamma_n \Phi \gamma_0 r_i^{\sigma}/[G_{tr}h], \quad (2)$$

where G_{tr} is the multiplication of transmit and receive antenna gains, γ_0 the SINR gap between channel capacity and a practical coding and modulation scheme, $\Phi = I + N_0B$, B the bandwidth, I the upperbound on the out-of-cell interference, and $n \in \{1,\, 2\}$ shows the priority category of the device. Regarding the uniform distribution of UEs in the cell, the PDF of the distance between a UE and the BS is $f(r) = \frac{2r}{R_c^2}$, where R_c is the cell radius. Then, one can find the long-term average required transmit power as:

$$\bar{P}_{t_n} = \int_0^{R_c} \frac{\gamma_n \Phi \gamma_0 r^{\delta}}{G_{tr}} \frac{2r}{R_c^2} dr = \frac{2(\delta+1)R_c^{\delta-2}\gamma_n\Phi\gamma_0}{G_{tr}}. \quad (3)$$

Also, one can derive the unsuccessful transmission probability for node i as:

$$q_{(i,n)} = pr(P_{t_{(i,n)}} > P_{max}) \overset{(a)}{=} 1 - \exp\big(-\frac{r_i^{\delta}\gamma_0\Phi\gamma_n}{G_{tr}P_{max}}\big), \quad (4)$$

where P_{max} is the maximum allowed transmit power, and (a) is due to the fact that h is exponentially distributed. Then, regarding the uniform distribution of UEs in the cell one can drive the average unsuccessful transmission probability as:

$$\bar{q}_n = \int_0^{R_c} \Big[1 - \exp(-\frac{r^{\delta}\gamma_0\Phi\gamma_n}{G_{tr}P_{max}})\Big] \frac{2r}{R_c^2} dr. \quad (5)$$

The integral in (5) can be evaluated using the integral tables in [17] in order to derive the exact $\bar{q}_n$. For example, when $\delta = 4$ we have:

$$\bar{q}_n = 1 - \sqrt{\pi}\mathrm{Erfc}(AR_c^2)/2AR_c^2,$$

where $A = \sqrt{\frac{\gamma_0\Phi\gamma_n}{G_{tr}P_{max}}}$, and $\mathrm{Erfc}(\cdot)$ is the error function [17]. Assume each UE requires a random amount of uplink service which is exponentially distributed with average length of $\bar{d}_n$ bits. The average service time for a successful transmission is derived as:

$$u_n = \bar{d}_n[B\log_2(1+\gamma_n)]^{-1}, \quad n \in \{1{,}2\}. \quad (6)$$

By considering the retransmission policy, the average service time is written as:

$$\bar{u}_n = \sum_{k=1}^{\infty} \bar{q}_n^{[k-1]}[1-\bar{q}_n]k\; u_n = \frac{1}{1-\bar{q}_n}\frac{\bar{d}_n}{B\log_2(1+\gamma_n)}. \quad (7)$$

Denote the CDF of the service time by U_n. Without considering the retransmissions, the service time is exponentially distributed with mean of u_n. By considering the retransmission policy, the service time can be well approximated by an exponential distribution with mean of $\bar{u}$[illegible]. In Fig. 1a of section VI, we will show the tightness of the derived expressions based on this assumption.

2) Steady state analysis: Denote by ξ the state of the BS, where $\xi = 0, 1, 2,$ and 3 refer to the sleep, busy with a $\mathcal{P}_1$ device, busy with a $\mathcal{P}_2$ device, and listen states, respectively. Using theorem 5.2 in [18], the stability condition for this queuing system is $\rho = \rho_1 + \rho_2 < 1$, where $\rho_n = \bar{u}_n \lambda_n$. In order to investigate the steady state performance of the system, the generation functions P_0, P_1, P_2, and P_3 are defined in Appendix A. From these functions, the steady state probabilities are derived as:

$$pr(\xi = 0) = P_0(1,1) = \mu P_3(1)\bar{v}, \tag{8}$$

$$pr(\xi = n) = P_n(1,1) = \bar{u}_n \lambda_n, \quad n \in \{1, 2\}, \tag{9}$$

$$pr(\xi = 3) = P_3(1) = \frac{1-\rho}{1+\mu\bar{v}}. \tag{10}$$

Regarding the steady state probabilities, the mode transition frequency is found as the ratio between the average time in the sleep mode and the average length of each sleeping window as

$$F_{sw} = 2pr(\xi = 0)/\bar{v} = 2\mu P_3(1).$$

Finally, (1) can be rewritten as:

$$E^b_{cons} = \rho P_s + \frac{1-\rho}{1+\mu\bar{v}}\big[P_l + \mu\bar{v}P_{sl} + 2\mu E_{sw}\big]. \tag{11}$$

B. Battery Lifetime Model for Battery-Limited UEs

The average energy consumption from data generation to successful data transmission can be modeled as:

$$E^{ac}_n + \big[P_c + \eta\bar{P}_{t_n}\big]\bar{u}_n, \tag{12}$$

where P_c is the average circuit power consumption, and η is the inverse of power amplifier efficiency. Also, E^{ac}_n is the average energy consumption from data generation until successful transmission, and is formulated as $P_c W_n \alpha\tau$, where τ is the average spent time in each trial, and W_n is the average delay from data generation until successful transmission, to be discussed in the following subsection. For most MTC applications, the packet generation at each device can be modeled as a Poisson process [19], and hence, one can define a duty cycle for operation of each device. Let us define the expected battery lifetime as the ratio between battery capacity, i.e. E_0, and the average energy consumption in each duty cycle [20]. Then the expected battery lifetime for a $\mathcal{P}_n$ device is found as:

$$\mathcal{L}_n = \frac{E_0}{E_{st} + E^{ac}_n + \big[P_c + \eta\bar{P}_{t_n}\big]\bar{u}_n} T, \tag{13}$$

where T is the duty cycle, and E_{st} is the average static energy consumption in each duty cycle for data gathering, synchronization, and etc.

C. Transmission Delay for the UEs

In order to evaluate the delay performance of the system, in the following we take the derivative of the introduced generation functions in Appendix A at z_1=z_2=1, as follows:

$$\mathbb{E}(N_1^{(0)}) = 0.5\mu\lambda_1 P_3(1)\hat{v}, \tag{14}$$

$$\mathbb{E}(N_1^{(1)}) = \lambda_1^2 u_1[u_1 + \mu P_3(1)\hat{v}/2 + u_2^2\lambda_2]/[1-\lambda_1 u_1], \tag{15}$$

$$\mathbb{E}(N_1^{(2)}) = \lambda_1\bar{u}_2^2\lambda_2, \;\; \mathbb{E}(N_1^{(3)}) = 0, \tag{16}$$

$$\mathbb{E}(N_2^{(0)}) = \mu\big[\mathbb{E}(N_2^{(3)})\bar{v} + 0.5\lambda_2 P_3(1)\hat{v}\big], \tag{17}$$

$$\mathbb{E}(N_2^{(1)}) = \big[R_1 + R_2\big]\bar{u}_1 + \lambda_1\lambda_2\bar{u}_1^2, \tag{18}$$

$$\mathbb{E}(N_2^{(2)}) = \big[H + \lambda_2\mathbb{E}(N_2^{(3)})\big]\bar{u}_2 + \lambda_2^2\bar{u}_2^2, \tag{19}$$

$$\mathbb{E}(N_2^{(3)}) = \lambda_2[1 - P_3(1)]/\alpha, \tag{20}$$

where $\mathbb{E}(N_i^{(j)})$ is the derivative of P_j with respect to z_i, and

$$R_1 = \frac{H + \lambda_2\mathbb{E}(N_2^{(3)})\rho_1\rho_2}{[1-\rho_1]\lambda_2\bar{u}_1} - \frac{\lambda_2\rho_1\rho_2}{1-\rho_1} + \frac{\bar{u}_1[R_3 - 2\lambda_2^2\bar{u}_2^2]}{2\bar{u}_1^2},$$

$$R_2 = \frac{\lambda_2\rho_1[1-\rho][1-P_3(1)]}{\alpha\bar{u}_1[1-\rho_1]P_3(1)} + \frac{P_3(1)}{2\lambda_2^2\bar{u}_1^2}\Bigg\{-\frac{2\lambda_2^3\bar{u}_1^2\rho_1[1-\rho]}{P_3(1)[1-\rho_1]} + \lambda_2\bar{u}_1\Bigg[\frac{2\lambda_1\lambda_2^2\bar{u}_1^2}{[1-\rho_1]^3}(1+\mu\bar{v}) + \mu\frac{\lambda_2^2\hat{v}\rho_1[2-\rho_1]}{[1-\rho_1]^2})\Bigg]\Bigg\},$$

$$R_3 = [\frac{\lambda_2}{1-\rho_1}]^2\Bigg[2\bar{u}_2^2 + \frac{2\lambda_1\bar{u}_2\bar{u}_1^2}{1-\rho_1}\Bigg],$$

$$H = \frac{(1-\rho_1)P_3(1)}{2\big[1-\rho\big]}\Bigg\{\frac{\lambda_1^2\lambda_2 2\bar{u}_1^2}{\big[1-\rho_1\big]^3} + \frac{\lambda_2 R_3}{P_3(1)} + \Bigg[\hat{v} + \frac{2\lambda_1\bar{v}\bar{u}_1^2}{1-\rho_1}\Bigg] \times \mu[\frac{\lambda_2}{1-\rho_1}]^2\Bigg\} + \alpha\big[\mathbb{E}(N_2^{(3)})\big]^2/P_3(1).$$

Then, one can use the Little's lemma to obtain the average delay from data generation until successful transmission, as follows:

$$W_n = \sum\nolimits_{m\in\{0,1,2,3\}}\mathbb{E}(N_n^{(m)})/\lambda_n. \tag{21}$$

By adding the service time to W_n, the average experienced delay in data transmission is derived as $D_n = W_n + \bar{u}_n$.

IV. THE MULTI-CELL SCENARIO

In this section, we present the characteristics of the energy-delay-lifetime tradeoffs in a multi-cell scenario. We consider a homogeneous network of BSs which are deployed as independent homogeneous PPPs of intensities λ_b. To be adaptive to traffic variations, we turn off $(1-\beta)\lambda_b$ BSs, and hence, the density of active BSs is $\beta\lambda_b$, where β is the traffic-dependent control parameter. The arriving UEs in the switched-off cells are served by the active BSs. We further assume that each UE is associated to the closest BS, and the out-of-cell is upper-bounded by I.

A. Energy Consumption Model for the Access Network

The average energy consumption of the access network per unit area can be formulated as:

$$E^{net}_{cons} = \beta\lambda_b E^{ab}_{cons} + (1-\beta)\lambda_b P_{sl}, \tag{22}$$

where the average energy consumption for an active BS is found from (1) as:

$$E^{ab}_{cons} = T_s P_s + [1 - T_s] P_l. \tag{23}$$

To model the system performance, we reuse the presented vacation-based queuing system with mixed-priority users in section III-C where $\mu \to 0$, i.e. the BS is either in the service or listen mode. Denote the channel gain between node i and the corresponding BS as $g_i = h[r_i]^{-\sigma}$, where $h \sim \exp(1)$, and the PDF of the communication distance r_i is [21]:

$$f_{r_i}(x) = 2\pi\beta\lambda_b x \exp(-\pi\beta\lambda_b x^2). \tag{24}$$

Using the derived transmit power expression in (2), and the PDF of the communications distance in (24), one can derive the long-term average uplink transmit power of a type-n device as:

$$\bar{P}_{t_n} = \int_0^\infty \frac{\gamma_n \Phi \gamma_0 x^\delta 2\pi\beta\lambda_b x}{G_{tr}\exp\left(\pi\beta\lambda_b x^2\right)} dx = \frac{\gamma_n\Phi\gamma_0}{G_{tr}} \frac{\Gamma(1+\frac{\delta}{2})}{\left[\pi\beta\lambda_b\right]^{\frac{\delta}{2}}}, \tag{25}$$

where Γ is the gamma function. One sees that the average uplink transmit power is a decreasing function of β. Also, the probability of unsuccessful transmission for node i of type n is found as:

$$q_{(i,n)} = pr(P_{t_{(i,n)}} > P_{max}) = 1 - \exp\big(-\frac{r_i^\delta \gamma_n \Phi \gamma_0}{G_{tr} P_{max}}\big).$$

Now, the average unsuccessful transmission probability is found as

$$\bar{q}_n = \int_0^\infty \frac{q_{(i,n)} 2\pi\beta\lambda_b x}{\exp(-\pi\beta\lambda_b x^2)} dx \overset{\text{(b)}}{\simeq} 1 - \exp(-\frac{\bar{P}_{t_n}}{P_{max}}), \tag{26}$$

where Jensen's inequality [22] has been used in (b). Also, one sees that $\bar{q}_n$ is a decreasing function of β. Inserting the above derived $\bar{q}_n$ and $\bar{P}_{t_n}$ in (7), one can find u_n and $\bar{u}_n$ for the multi-cell scenario. Following the same steps as in section III-A2, one can derive the steady state probabilities as:

$$pr(\xi = 0) = 0, \tag{27}$$

$$pr(\xi = n) = P_n(1,1) = \bar{u}_n \lambda^u_n / [\beta\lambda_b], \quad n \in \{1,2\}, \tag{28}$$

$$pr(\xi = 3) = P_3(1) = 1 - pr(\xi = 1) - pr(\xi = 2), \tag{29}$$

where λ^u_n is the arrival rate of type-n UEs per unit area. Now, one can rewrite (22) as:

$$\begin{aligned} E^{net}_{cons} &= \beta\lambda_b[P_l + \rho(P_s - P_l)] + (1-\beta)\lambda_b P_{sl}, \\ &= \beta\lambda_b[P_l - P_{sl}] + \beta\lambda_b\rho(P_s - P_l) + \lambda_b P_{sl}, \end{aligned} \tag{30}$$

where

$$\rho = \bar{u}_1 \frac{\lambda^u_1}{\beta\lambda_b} + \bar{u}_2 \frac{\lambda^u_2}{\beta\lambda_b} \simeq \sum_{n\in\{1,2\}} u_n \frac{\lambda^u_n}{\beta\lambda_b} \exp\left(G_n [\beta\lambda_b]^{-\frac{\delta}{2}}\right),$$

$$G_n = \gamma_n \Phi \gamma_0 \Gamma(1 + 0.5\delta) / [G_{tr} \pi^{\frac{\delta}{2}}]. \tag{31}$$

By inserting the above derived expressions for $\bar{u}_n$, ρ, $\bar{P}_{t_n}$, and $\mu = 0$ in (13)-(21), one can derive the respective lifetime and delay expressions for the UEs in the multi-cell scenario, which are skipped here.

V. ENERGY-LIFETIME-DELAY TRADEOFFS

Single-Cell Scenario: From (14) and (17), one sees that the experienced delay decreases in mean listening time, i.e. $1/\mu$. Then, the access probability for UEs increases in the listening time, which in turn improves UEs' energy efficiency and battery lifetimes. The increase in the lifetime of machine devices comes at the cost of increase in the energy consumption of the BS because the latter increases in the idle listening time, as found in (8) and (11). Another tradeoff can be seen between battery lifetime and delay performance of machine devices. From (18) and (20), one sees that the experienced delay for $\mathcal{P}_2$ devices decreases in α. On the other hand, (12) and (13) show that the battery lifetime decreases in α. In the following, we present a delay/lifetime constrained energy minimizing problem which casts light to the aforementioned tradeoffs.

Assume the sleeping window is exponentially distributed with mean $\bar{v}$, and D^{th}_1 and L^{th}_2 are given as the threshold delay and lifetime requirements for $\mathcal{P}_1$ and $\mathcal{P}_2$ devices respectively. Then, one can formulate the constrained energy minimizing problem as:

$$\min_{\mu,\bar{v}} \; E^b_{cons} \tag{32}$$

$$\text{subject to: } D_1 \le D^{th}_1, L^{th}_2 \le L_2.$$

Using the lifetime and delay expressions in (13) and (21) respectively, the constraints in (32) can be rewritten as:

$$\frac{\bar{u}_1 + \mu P_3(1)\hat{v}/2 + \lambda_2 \bar{u}_2^2}{1 - \bar{u}_1\lambda_1} \le D^{th}_1, \tag{33}$$

$$\sum_m \mathbb{E}(N_2^{(m)}) \le \frac{E_0 T - L^{th}_2\big[[P_c + \eta\bar{P}_{t_2}] + E_{st}\big]}{L^{th}_2 P_c \alpha\tau / \lambda_2}. \tag{34}$$

From (11), one sees that the objective function in (32) is not jointly convex over μ and $\bar{v}$. Hence, in order to find the optimal operating point, we can optimize μ as a function of $\bar{v}$, and then search over $\bar{v}$ to minimize the objective function in (32). Given $\bar{v}_x$ as the length of the sleeping window, the optimization problem in (32) is reduced to:

$$\min_{0\le\mu\le\mu_{max}} \quad \rho P_s + \frac{1-\rho}{1+\mu\bar{v}_x}(P_l + \mu\bar{v}_x P_{sl} + 2\mu E_{sw}), \tag{35}$$

where μ_{max} is the upper bound on μ, and is found by joint satisfying (33) and (34). Taking the first derivative of E^b_{cons} with respect to μ shows that it is a strictly decreasing function of μ when the following condition is held:

$$2E_{sw} < [P_l - P_{sl}]\bar{v}_x. \tag{36}$$

One sees that the right hand side of (36) represents the average amount of saved energy in each sleeping window, and the left hand side of (36) is the energy consumption in switching the BS off and on. Then, the optimal choice of μ is μ_{max} whenever $\bar{v}_x$ satisfies (36), and 0 otherwise.

Multi-Cell Scenario: Taking the second derivative of E^{net}_{cons} in (30) with respect to β shows that E^{net}_{cons} is a strictly convex function of β, and chooses its minimum value at β^*, where

TABLE I: Parameters for numerical analysis

Parameters	*Value*
Cell inner and outer radius	50 m, 500 m
Pathloss	$128.1 + 37.6\log(\frac{d}{1000})$
Thermal noise power	-204 dBW/Hz
Bandwidth	1.8 MHz
λ_1, λ_2	0.2, 2
d_1, d_2	5 Mbit, 50 Kbit
γ_1, γ_2	6, 6
$\bar{v}$, α, ξ	5, 0.1, 2
E_0, E_{sw}, E_{st}	2000 J, 25 J, 250 μJ
P_c, P_s, P_{sl}, P_l	0.05, 130, 30, 100 W

β^* is found by solving:

$$\Delta E(\beta) \triangleq P_l - P_{sl} - 0.5\lambda_b[P_s - P_l]\times$$
$$\sum\nolimits_{n\in\{1,2\}} G_n\delta[\beta\lambda_b]^{-\frac{\delta}{2}-1}u_n\lambda_n^u \exp\left(G_n[\beta\lambda_b]^{-\frac{\delta}{2}}\right) = 0.$$

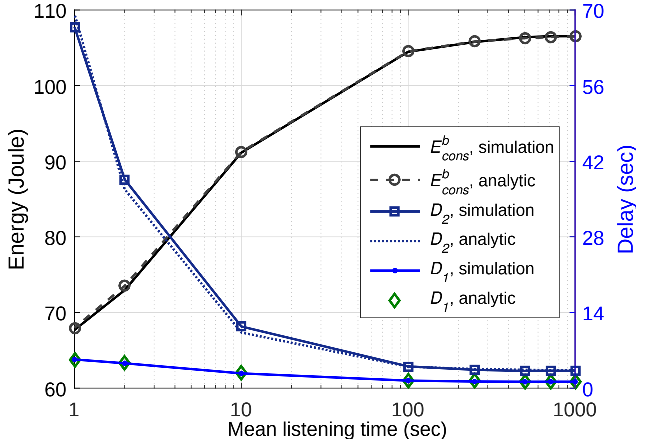
where $\Delta E(\beta)$ is the derivative of E_{cons}^{net} with respect to β. Then, the network energy consumption decreases by turning off $(1-\beta)\lambda_b$ BSs when $\Delta E(\beta) > 0$. In the special case that UEs are homogeneous, e.g. only $\mathcal{P}_1$ devices are present, and $\delta = 2$, we have:

$$\beta^* = 0.5G_1/\text{LW}\left(0.5\sqrt{\frac{G_1[P_l - P_{sl}]}{2[P_s - P_l]u_1\lambda_1^u}}\right) \qquad (37)$$

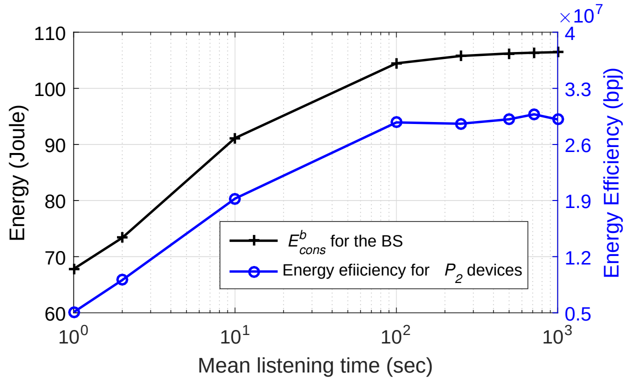
where LW(x) function is the inverse of the function $f(x) = x\exp(x)$ [23]. Valuable insights to the green network design problem can be drawn from (37). For example, one sees how β^* decreases in the difference between listening and sleeping power consumption levels. Also, we see that for $\beta < \beta^*$, $\Delta E(\beta)$ is a decreasing function of β, and hence, here one can not trade delay and lifetime of the UEs to save energy for the access network.

VI. NUMERICAL RESULTS

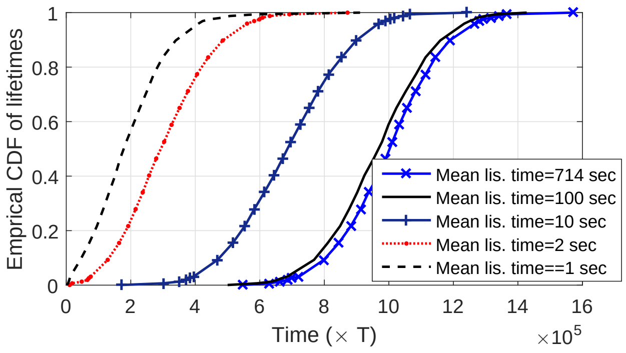
The implemented system model in this section is based on the uplink of a single cell with coexistence of M2M and H2H traffics. The UEs are randomly distributed according to a spatial Poisson point process in the cell with a minimum distance of 50 meters from the BS. The sleeping window is exponentially distributed with rate $1/\bar{v}$. The other parameters can be found in Table I.



(a) Tradeoffs between the average energy consumption of the BS per unit time (left y-axis), and the experienced delay (right y-axis)



(b) Tradeoff between average energy consumption of the BS per unit time, and energy efficiency for $\mathcal{P}_2$ devices (in bit-per-joule)



(c) Empirical CDF of lifetimes for different mean listening time values

Fig. 1: Performance tradeoffs

Fig. 1a illustrates the tradeoff between energy saving for the BS and the experienced delay in data transmission. The solid lines are the numerical results, while the dashed lines have been derived from the analytical expressions in section III. This figure confirms that the presented analytical results in section III match well with the numerical results. One sees while the operational costs of the BS increase in the listening time, the delay decreases in the listening time. Fig. 1b depicts the tradeoff between energy saving for the base station and energy efficiency in data transmission. One sees that both the average energy consumption of the BS and the energy efficiency of machine devices in data transmission increase in the idle listening time. The increase in energy efficiency is due to the fact that the access probability for UEs increases in the listening time, and hence, UEs need to wait for a shorter time to be served. Fig. 1c presents the empirical CDF of individual lifetimes versus mean listening time. Comparing Fig. 1b with

Fig. 1c, one sees a higher level of energy efficiency in data transmission provides a higher level of battery lifetime. For example, one sees that the shortest and the mean individual lifetime of $\mathcal{P}_2$ devices when the mean listening time is 100 seconds, are approximately 200% and 40% higher than the case that the mean listening time is 10 seconds.

VII. CONCLUSION

In this work, we have investigated the energy-lifetime-delay tradeoffs in green cellular network design with M2M/H2H coexistence. For a single cell in which a green BS serves two types of distinct traffics, we have derived analytical expressions for the total energy consumption of the BS, the experienced delay in data transmission, and the expected battery lifetime of connected devices, as a function of system and traffic parameters. Then, the results have been extended to the multi-cell scenario, and the performance impacts of the control parameters on the energy-lifetime and energy-delay tradeoffs have been studied. Significant impact of the BSs' energy saving strategies on the UEs' lifetimes has been presented using numerical and analytical results. These results promote revisiting traditional energy saving strategies to cope with the ever increasing number of connected machine-type devices in cellular networks.

APPENDIX A

Define $\bar{X}$ to be the elapsed duration of the random variable X, and

$$\begin{aligned} p_{ij}^{(0)}(x)dx &= pr\big(N_1^{(0)}=i, N_2^{(0)}=j, x<\bar{V}\leq x+dx\big),\\ p_{ij}^{(m)}(x)dx &= pr\big(N_1^{(m)}=i, N_2^{(m)}=j, x<\bar{U}_n\leq x+dx\big),\\ p_j^{(3)} &= pr(N_1^{(3)}=0, N_2^{(3)}=j), \end{aligned}$$

where $N_n^{(m)}$ shows the number of queued $\mathcal{P}_n$ devices is state m. The corresponding generating functions are given by:

$$\begin{aligned} P_m(z_1,z_2,x) &= \sum_{i\geq 0}\sum_{j\geq 0} p_{ij}^{(m)}(x) z_1^i z_2^j dx, \quad m\in\{0,1,2\},\\ P_3(z_2) &= \sum_{j\geq 0} p_j^{(3)}(x) z_2^j. \end{aligned}$$

By applying the supplementary variable method, one can form the generating functions as [18]:

$$\begin{aligned} &P_0(z_1,z_2) = \int_0^\infty \mu P_3(z_2)[1-V(x)]e^{-f_1 x}dx,\\ &P_1(z_1,z_2) = \frac{\bar{u}_1}{z_1(1+\bar{u}_1 f_1)-1}\Big\{\frac{f_3}{1+\bar{u}_2 f_1} - \frac{f_3}{1+\bar{u}_2 f_4} - f_5\\ &\qquad\qquad + P_3(z_2)\lambda_1 z_1 + P_3(z_2)\mu[v^*(f_1)-v^*(f_2)]\Big\},\\ &P_2(z_1,z_2) = \bar{u}_2 f_3/\big[1+\bar{u}_2 f_1\big],\\ &P_3(z_2) = \frac{\alpha z_2 - \frac{\alpha}{1+f_4\bar{u}_2}}{\mu(v^*(f_1)-1)+\lambda_1 x(z_2)+\frac{\lambda_2}{1+f_2\bar{u}_2}-\lambda}\frac{\partial P_3(z_2)}{\partial z_2}, \end{aligned}$$

where

$$\begin{aligned} &f_1 = \lambda-\lambda_1 z_1-\lambda_2 z_2, &&f_2=\lambda-\lambda_2 z_2,\\ &f_3 = \lambda_2 P_3(z_2)+\alpha\partial P_3(z_2)/\partial z_2, &&f_4 = \lambda-\lambda_1 y(z_2)-\lambda_2 z_2, \end{aligned}$$

$$f_5 = P_3(z_2)\lambda_1 y(z_2) + P_3(z_2)\mu\big[v^*(f_4)-v^*(f_2)\big],$$

in which $\lambda=\lambda_1+\lambda_2$, and $y(z_2)$ is the smallest positive real root of $z_1=\frac{1}{1+\bar{u}_1 f_1}$.

REFERENCES

[1] G. Miao and G. Song, *Energy and Spectrum Efficient Wireless Network Design*. Cambridge University Press, 2014.

[2] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *IEEE International Conference on Communications Workshops*, 2009, pp. 1–5.

[3] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1525–1536, September 2011.

[4] J. Wu, S. Zhou, and Z. Niu, "Traffic-aware base station sleeping control and power matching for energy-delay tradeoffs in green cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4196–4209, August 2013.

[5] J. Wu, Y. Bao, G. Miao, S. Zhou, and Z. Niu, "Base station sleeping control and power matching for energy-delay tradeoffs with bursty traffic," *IEEE Trans. Veh. Technol.*, no. 99, pp. 1–15, May 2015.

[6] H. Tabassum, U. Siddique, E. Hossain, and M. J. Hossain, "Downlink performance of cellular systems with base station sleeping, user association, and scheduling," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5752–5767, 2014.

[7] Z. Niu, X. Guo, S. Zhou, and P. R. Kumar, "Characterizing energy–delay tradeoff in hyper-cellular networks with base station sleeping control," *IEEE Journal on Selected Areas in Communications,*, vol. 33, no. 4, pp. 641–650, 2015.

[8] J. Hämäläinen *et al.*, "A novel multi-objective cell switch-off method with low complexity for realistic cellular deployments," *arXiv preprint arXiv:1506.05595*, 2015.

[9] S. Andreev *et al.*, "Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 32–40, September 2015.

[10] 3GPP TS 22.368 V13.1.0, "Service requirements for machine-type communications," Tech. Rep., 2014. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/22_series/22.368/22368-d10.zip

[11] Nokia Networks, "Looking ahead to 5G: Building a virtual zero latency gigabit experience," Tech. Rep., 2014.

[12] A. Azari and G. Miao, "Lifetime-aware scheduling and power control for cellular-based M2M communications," in *IEEE Vehicular Technology Conference (VTC Spring)*, 2015.

[13] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer, "Downlink and uplink decoupling: A disruptive architectural design for 5G networks," in *IEEE Global Communications Conference*, Dec 2014, pp. 1798–1803.

[14] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G Multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1224–1240, June 2015.

[15] S. Duan, V. Shah-Mansouri, and V. W. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *IEEE Global Communications Conference*, 2013, pp. 4747–4752.

[16] I. Dimitriou, "A mixed priority retrial queue with negative arrivals, unreliable server and multiple vacations," *Applied Mathematical Modelling*, vol. 37, no. 3, pp. 1295–1309, 2013.

[17] D. Zwillinger, *Table of integrals, series, and products*. Elsevier, 2014.

[18] I. Dimitriou, "Analysis of a priority retrial queue with dependent vacation scheme and application to power saving in wireless communication systems," *The Computer Journal*, pp. 1363–1380, 2012.

[19] 3GPP, "USF capacity evaluation for MTC," Tech. Rep., 2010, TSG GERAN 46 GP-100894.

[20] A. Azari and G. Miao, "Lifetime-aware scheduling and power control for cellular-based M2M communications," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2015.

[21] J. Peng, P. Hong, and K. Xue, "Stochastic analysis of optimal base station energy saving in cellular networks with sleep mode," *IEEE Communications Letters*, vol. 18, no. 4, pp. 612–615, 2014.

[22] W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, vol. 30, no. 1, pp. 175–193, 1906.

[23] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the LambertW function," vol. 5, no. 1, pp. 329–359, 1996.