

# Optimized Wake-up Scheme with Bounded Delay for Energy-Efficient MTC

Soheil Rostami<sup>§</sup>, Sandra Lagen<sup>†</sup>, Mário Costa<sup>§</sup>, Paolo Dini<sup>†</sup>, and Mikko Valkama<sup>\*</sup>

<sup>§</sup>Huawei Technologies Oy (Finland) Co. Ltd, Helsinki R&D Center

<sup>†</sup>Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Barcelona, Spain

<sup>\*</sup>Department of Electronics and Communications Engineering, Tampere University of Technology, Finland

Emails:<sup>§</sup>{soheil.rostami1, mariocosta}@huawei.com, <sup>†</sup>{sandra.lagen, paolo.dini}@cttc.es, <sup>\*</sup>mikko.e.valkama@tut.fi

**Abstract**—The limitations of state-of-the-art cellular modems prevent achieving low-power and low-latency Machine Type Communications (MTC) based on current power saving mechanisms alone. Recently, the concept of wake-up scheme has been proposed to enhance battery lifetime of 5G devices, while reducing the buffering delay. The existing wake-up algorithms use static operational parameters that are determined by the radio access network at the start of the user’s session. In this paper, the average power consumption of the wake-up enabled MTC UE is modeled by using a semi-Markov process and then optimized through a delay-constrained optimization problem, by which the optimal wake-up cycle is obtained in closed form. Numerical results show that the proposed solution improves the power consumption of an optimized DRX scheme by up to 40% for a given delay requirement. The numerical results show that the proposed solution, as compared to an optimized discontinuous reception (DRX)-based reference scheme, can improve power consumption by up to 40% for a given delay requirement.

**Index Terms**—Energy efficiency, MTC, DRX, wake-up schemes, 5G.

## I. INTRODUCTION

The 3rd generation partnership project (3GPP) has employed a set of computationally-intensive physical layer techniques in order to satisfy high quality-of-service (QoS) requirements in fifth generation (5G) networks [1], which inevitably lead to an increased energy consumption of mobile terminals. At the same time, 5G mobile networks are expected to support and drive a large variety of existing, emerging and unforeseen internet of things (IoT) use cases, such as low-power and low-latency Machine Type Communications (MTC). One of the key challenges in MTC is to extend the lifetime of the user equipment (UE) without the need for replacing its battery. However, due to the slow improvement in battery efficiency, the cellular modem’s energy consumption needs to be reduced [2].

In general, the MTC UE consists of a transmitter, a receiver, and interfaces to other logical and physical units of the terminal. Due to the fact that the transmitter may easily

have considerably higher instantaneous power consumption as compared to the receiver, the UE turns on its transmitter only for data transmission whenever is needed, thereby reducing the energy consumption of the UE. However, the UE in receive mode is required to monitor a control channel from the network. Depending on the use case, the receiver side energy consumption, when accumulated over a long period of time, can eventually surpass the transmitter energy consumption [3]. Therefore, the power saving mechanisms for cellular modem in receive mode are of great importance.

The current *de facto* power saving mechanism for Long-Term Evolution (LTE), LTE-Advanced, and 5G New Radio (NR) networks is discontinuous reception (DRX) [1], [4], [5]. DRX specifies a sleep/active pattern that is followed by both UE and network sides. During DRX, the energy consumption of the UE is reduced by switching off some radio modules for long periods of time, and the UE wakes up for short intervals to check for potential control signals from the network. However, it has been shown that the existence of unscheduled DRX cycles has adverse impact on the UE energy consumption [6]; see also Fig. 1.

In particular, the energy consumption of DRX can be reduced even further, by removing the unscheduled DRX cycles. To end this, the wake-up scheme (WuS) has been recently proposed and discussed in [2], [7]. In the WuS, the wake-up receiver (WRx) monitors periodically a narrowband wake-up signal that triggers it to wake-up. The wake-up signal detection is prone to false alarm and misdetection due to presence of noise and fading channel. Because of the strict delay requirements and scarcity of radio resources, WRx and the corresponding wake-up signal can be designed in such a way to have very low probabilities of misdetection and false alarm [7]. However, the existing wake-up concepts and algorithms, such as those described in [2], [7], [8], utilize static operational parameters, which are determined by the base station at the beginning of the user’s session. Therefore, the optimization of such parameters to tailor the WuS to the specific user’s session profile is needed. This is the main objective of the present work.

In particular, the contributions of this paper are described next. Firstly, the average power consumption and buffering

This work has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 675891 (SCAVENGE), and Tekes TAKE-5 project. Also, work at CTTC has been partially funded by Spanish MINECO grant TEC2017-88373-R (5G-REFINE) and Generalitat de Catalunya grant 2017 SGR 1195.

delay of WuS are quantified under Poisson traffic model for a given set of wake-up related parameters. Secondly, by utilizing such mathematical model, a closed-form solution for the optimal wake-up cycle (w-cycle) that minimizes the UE's power consumption is obtained. Finally, simulation-based numerical results are provided in order to compare the power consumption of our proposed solution with respect to the optimized DRX-based reference mechanism that is proposed in [9].

The rest of this paper is organized as follows. Section II models the wake-up enabled UE using a semi-Markov process. Section III derives the power consumption and buffering delay of the wake-up based access scheme. Then, building on these mathematical models, Section IV formulates the optimization problem and derives the optimal solution for minimum power consumption subject to a delay constraint in closed form. Numerical results and conclusions are presented in Sections V and VI, respectively.

## II. SYSTEM MODEL

### A. Basic Wake-Up Radio Concept

To avoid periodical monitoring of unscheduled control signaling that can reduce the battery lifetime quickly, a new WuS has been recently developed. The WuS is characterized by a companion low-complex single-purpose receiver, known as WRx, and a specifically designed narrowband wake-up signaling [7].

Generally, the WRx monitors the wake-up signaling periodically, every w-cycle ( $t_w$ ), to check whether there is any scheduled data, i.e. wake-up indicator (WI) equals 1 or not (WI=0). We assume that WRx is needed to be ON for a duration of  $t_{on}$  milliseconds to receive such a signal. If WI=1, the UE switches ON the main cellular modem and proceeds to decode both physical downlink control channel (PDCCH) and physical downlink shared channel (PDSCH), otherwise (WI=0) the UE remains in sleep state. As soon as processing of PDCCH/PDSCH is completed, the cellular modem immediately re-initiates its inactivity timer with duration of  $t_i$  to receive the upcoming short-dependent packets. When this timers expires, the cellular modem switches OFF and enters the sleep state. We denote the times for start-up and power-down of the baseband unit as  $t_{su}$  and  $t_{pd}$ , respectively.

Fig. 1 (a) and (b) depict the basic operation and representative power consumption behaviors of the conventional DRX-enabled cellular module and the cellular module with WRx, respectively, at a conceptual level. As illustrated, the WuS eliminates the unnecessarily wasted energy in the first and second DRX cycles, while also reducing the buffering delay compared to DRX. Due to the specifically-designed narrowband signal structure of WuS, the WRx power consumption ( $PW_2$ ) is much lower than that of the baseband unit at active state ( $PW_1$ ) [7].

In this study, different time intervals of wake-up related procedures are defined as integer multiples of a transmission time interval (TTI) with duration of 1 ms, in order to provide consistent and accurate timing definitions. Terminology-wise,

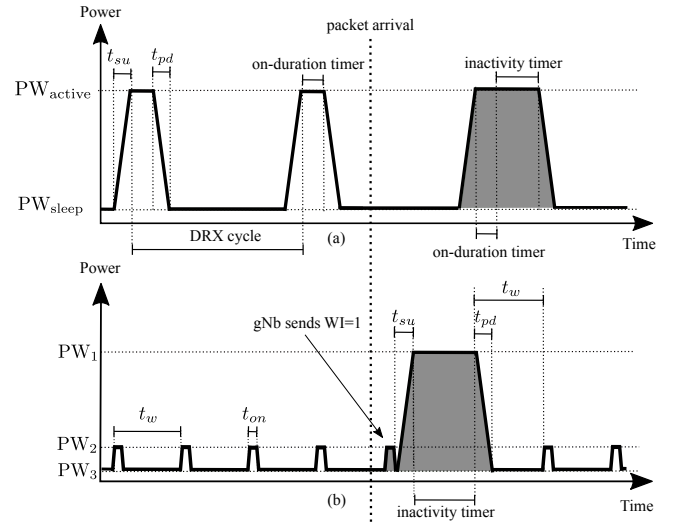


Figure 1: Power consumption profiles of (a) a typical DRX mechanism, and (b) a WRx-enabled cellular module. Gray and white colors illustrate scheduled and unscheduled cycles, respectively.

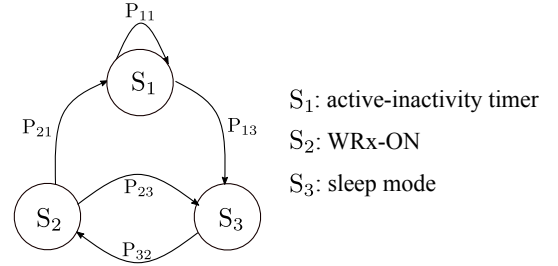


Figure 2: Semi-Markov process for the state transitions of the wake-up scheme, with states  $S_1$ ,  $S_2$  and  $S_3$ .

we use gNB and UE to refer to the base-station unit and the MTC device, respectively, according to 3GPP NR specifications [1].

### B. Semi-Markov Model for Wake-up Scheme

It is assumed that the packet generation follows a homogeneous Poisson point process with an average packet arrival rate of  $\lambda$  packets per TTI. The state machine of WuS is then modeled as a semi-Markov process with three different states,  $S_1$ ,  $S_2$  and  $S_3$ , as shown in Fig. 2. At  $S_1$ , inactivity timer is running, and if the device is scheduled before the expiry of timer, it restarts itself and remains at  $S_1$ , otherwise the device transfers to  $S_3$ . At  $S_2$ , WRx monitors the wake-up signal, and if WRx receives WI=0, the UE transfers to  $S_3$ , otherwise (WI=1) it transfers to  $S_1$ . At  $S_3$ , the device is in sleep state, and cannot receive any data, opposed to being fully-functioning at  $S_1$ . Moreover, at the end of its w-cycle the UE moves to  $S_2$ .

Let us denote  $P_{kl}$  the transition probability from state  $S_k$  to  $S_l$  for those  $k, l \in \{1, 2, 3\}$  that correspond to feasible state transitions, as shown in Fig. 2. When the UE is at  $S_1$ , it restarts its inactivity timer if the next packet is received before the end of  $t_i$ , otherwise it moves to  $S_3$ . Therefore,  $P_{11}$  and

$P_{13}$  can be expressed as  $P_{11} = \Pr[t_p \leq t_i] = 1 - e^{-\lambda t_i}$ , and  $P_{13} = 1 - P_{11}$ , where  $t_p$  denotes the exponentially-distributed inter-packet arrival time. When the UE is at  $S_2$ , it moves to  $S_1$  either because of false alarm or correct detection, otherwise it moves to  $S_3$ . Accordingly,  $P_{21}$  and  $P_{23}$  can be expressed as

$$P_{21} = \Pr[t_p > t_w]P_{fa} + \Pr[t_p \leq t_w](1 - P_{md}) \quad (1)$$

$$= e^{-\lambda t_w}P_{fa} + (1 - e^{-\lambda t_w})(1 - P_{md}),$$

and  $P_{23} = 1 - P_{21}$ , where  $P_{md}$  is the probability of mis-detection and  $P_{fa}$  the probability of false alarm. Finally, at the end of every sleep cycle, the UE decodes wake-up signal, and therefore,  $P_{32} = 1$ .

The stationary probability of each state  $S_k$ , denoted by  $P_k$ , can be calculated by utilizing the set of balance equations ( $P_k = \sum_{l=1}^3 P_l P_{lk}$ ) and the sum of probabilities ( $\sum_{k=1}^3 P_k = 1$ ). It follows that,

$$P_1 = \frac{P_{21}}{2P_{13} + P_{21}}, \quad (2)$$

$$P_2 = P_3 = \frac{P_{13}}{2P_{13} + P_{21}}. \quad (3)$$

Let us refer to the holding time for state  $S_k$  as  $\omega_k$ . Then, with the assumption that the per-packet service time is negligible, thanks to 5G NR new latency-optimized frame structure and high data rates, it follows that  $\mathbb{E}[\omega_1] = \mathbb{E}[T_I]$  [10], where  $\mathbb{E}[T_I]$  is the average value of the inactivity timer after waking-up the baseband unit, known as average inactivity period. If a packet arrives before  $t_i$ , the inactivity period is equal to the inter-packet call time, otherwise the inactivity period equals to  $t_i$ . Therefore, the inactivity period ( $T_I$ ) can be calculated based on  $t_p$  as

$$T_I(t_p) = \begin{cases} t_p, & \text{for } t_p \leq t_i \\ t_i, & \text{for } t_p > t_i. \end{cases} \quad (4)$$

Hence,  $\mathbb{E}[\omega_1]$  can be expressed as

$$\mathbb{E}[\omega_1] = \mathbb{E}[T_I] = \int_0^\infty T_I(t) f_p(t) dt = \frac{1 - e^{-\lambda t_i}}{\lambda}, \quad (5)$$

where  $f_p(t) = \lambda e^{-\lambda t}$  is the probability density function of the exponentially distributed packet arrival time. Additionally, the holding times for  $S_1$  and  $S_2$  are given by  $\mathbb{E}[\omega_2] = t_{on}$  and  $\mathbb{E}[\omega_3] = t_w - t_{on}$ .

Finally, it is important to note that from a system-level viewpoint, the tunable parameters of the WuS are the w-cycle ( $t_w$ ) and the inactivity timer ( $t_i$ ). However, for the sake of clarity, we mainly focus on scenarios with low packet arrival rates ( $\lambda \leq 0.2$  p/ms), e.g., MTC, for which, without loss of generality, the inactivity timer can be configured to a single TTI ( $t_i = 1$ ). The remaining parameters of the WuS ( $t_{on}$ ,  $t_{pd}$ ,  $t_{su}$ ) depend on physical constraints and signal design, and accordingly we assume them to be fixed. Therefore, in the rest of this paper, we focus on modeling the WuS performance and then optimizing the w-cycle ( $t_w$ ) to address the energy-latency trade-off for given values of  $t_{on}$ ,  $t_{pd}$ ,  $t_{su}$ , and  $t_i = 1$ .

### III. POWER AND DELAY ANALYSIS

#### A. Average Power Consumption

The average power consumption of the UE, denoted by  $\bar{P}_c$ , can be calculated as the ratio of the average energy consumption and the corresponding overall observation period, expressed as

$$\bar{P}_c = \frac{e_t + \sum_{n=1}^3 P_n \mathbb{E}[\omega_n] PW_n}{t_t + \sum_{n=1}^3 P_n \mathbb{E}[\omega_n]}, \quad (6)$$

where  $e_t$  and  $t_t$  are the average energy consumption of transition states and the mean overall time period that the UE spends on transition periods, which respectively read

$$t_t = P_2 P_{21} t_{su} + P_1 P_{13} t_{pd}, \quad (7)$$

and

$$e_t = t_t \frac{PW_1 - PW_3}{2}. \quad (8)$$

For modeling simplicity, and based on the actual settings and measurements of WRx [7], it is reasonable to assume that  $t_{on} \approx 0$ ,  $PW_3 \approx 0$ ,  $P_{fa} \approx 0$  and  $P_{md} \approx 0$ . Therefore, Eq. (6) can be expanded as a function of  $t_w$  as follows

$$\bar{P}_c(t_w) = PW_1 \frac{\frac{e^\lambda - 1}{\lambda} + \frac{1}{2}(t_{su} + t_{pd})}{\frac{e^\lambda - 1}{\lambda} + \frac{t_w}{1 - e^{-\lambda t_w}} + t_{su} + t_{pd}}. \quad (9)$$

In order to provide better insight into  $\bar{P}_c$ , the derivative of the power consumption with respect to w-cycle is next calculated, assuming continuous variable. The result is included in the following Lemma 1.

**Lemma 1.** *The average power consumption  $\bar{P}_c(t_w)$  is a strictly decreasing function with respect to  $t_w$  at  $t_w \geq 0$ .*

*Proof.* For continuous variable, the derivative of power consumption  $\bar{P}_c(t_w)$  with respect to  $t_w$  is given by

$$\frac{d\bar{P}_c}{dt_w} = PW_1 \frac{\left(\frac{e^\lambda - 1}{\lambda} + \frac{1}{2}(t_{su} + t_{pd})\right) \left((1 + \lambda t_w)e^{-\lambda t_w} - 1\right)}{\left((1 - e^{-\lambda t_w}) \left(\frac{e^\lambda - 1}{\lambda} + t_{su} + t_{pd}\right) + t_w\right)^2}. \quad (10)$$

It can be seen from Eq. (10),  $\frac{d\bar{P}_c}{dt_w} < 0$ , due to fact that  $(1 + \lambda t_w)e^{-\lambda t_w} < 1$ .  $\square$

As expected, increasing the w-cycle reduces the power consumption.

#### B. Average Buffering Delay

We next assume that the packets arriving during  $S_3$  are buffered at gNB until the UE enters  $S_1$ , thus causing buffering delay. We assume that the radio access network experiences unsaturated traffic conditions. Therefore, all packets that arrive during  $S_1$  are served promptly without any further scheduling or other processing delays. Furthermore, to simplify the delay modeling, we omit the buffering delay caused by packets arriving on  $S_2$  or at the start-up state of the modem. This is because the additional buffering delay of such packet arrivals is anyway very small ( $t_{on} + t_{su}$ ).

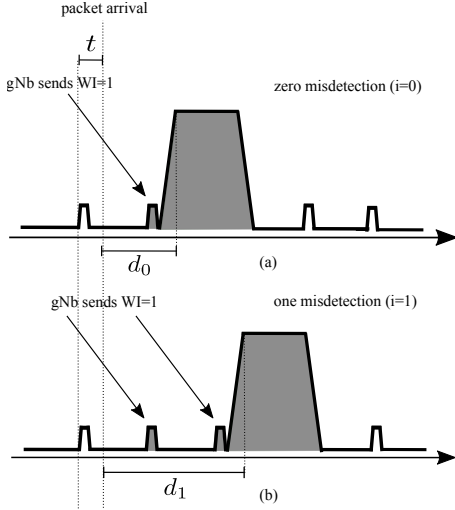


Figure 3: Buffering delay caused by wake-up scheme when (a) there is no misdetection, and (b) there is a single misdetection.

Now, as already shortly mentioned in Section II-A, misdetections can in general increase the buffering delay. For this purpose, Fig. 3 illustrates the buffering delay experienced by the UE with no misdetections, and a single misdetection. The number of consecutive misdetections and the corresponding buffering delay are referred to as  $i$  and  $d_i$ , respectively, and their dependency as a function of  $t$  can be written as  $d_i(t) = (i+1)t_w + t_{su} + t_{on} - t$ , for  $i \in \{0, 1, \dots\}$ . Furthermore, due to the small value of misdetection probability ( $P_{md} \approx 0$ ), the contribution of multiple consecutive misdetections to average delay is small. Therefore, the average buffering delay, denoted by  $\bar{D}$ , can be expressed as

$$\begin{aligned} \bar{D} &= P_2 \sum_{i=0}^{\infty} (1 - P_{md})(P_{md})^i \int_0^{t_w} f_p(t) d_i(t) dt \\ &\approx P_2 \int_0^{t_w} f_p(t) d_0(t) dt, \end{aligned} \quad (11)$$

which can be further expanded and solved as a function of  $t_w$  as follows

$$\bar{D}(t_w) = \frac{t_w + (t_{su} - \frac{1}{\lambda})(1 - e^{-\lambda t_w})}{2 + (1 - e^{-\lambda t_w})e^{\lambda}}. \quad (12)$$

Similarly to  $\bar{P}_c(t_w)$ , the derivative of  $\bar{D}(t_w)$  with respect to the continuous variable  $t_w$  can be calculated, which leads to the result in the following Lemma 2.

**Lemma 2.** *The average buffering delay  $\bar{D}(t_w)$  is a strictly increasing function with respect to  $t_w$  at  $t_w \geq 0$ .*

*Proof.* For continuous variable  $t_w$ , the derivative of  $\bar{D}(t_w)$  with respect to  $t_w$  is given by

$$\frac{d\bar{D}}{dt_w} = \frac{e^{\lambda t_i} (1 - (1 + \lambda t_w)e^{-\lambda t_w}) + 2(1 - e^{-\lambda t_w}) + 2\lambda t_{su}e^{-\lambda t_w}}{(2 + (1 - e^{-\lambda t_w})e^{\lambda t_i})^2}. \quad (13)$$

From Eq. (13), it can be easily concluded that  $\frac{d\bar{D}}{dt_w} > 0$ , due

to fact that  $(1 - (1 + \lambda t_w)e^{-\lambda t_w}) > 0$ .  $\square$

As expected, contrary to the behavior of  $\bar{P}_c$ , increasing the w-cycle increases the buffering delay. Therefore, a clear energy-delay trade-off appears in the selection of  $t_w$  for the WuS.

#### IV. OPTIMIZATION PROBLEM AND SOLUTION

In this section, the optimal w-cycle ( $t_w$ ) in terms of power consumption for a given buffering delay is found as the solution of a constrained optimization problem. In particular, the average buffering delay is constrained to be less than or equal to a pre-specified maximum tolerable delay or delay bound, denoted by  $\bar{D}_m$ , whose value is set based on the service type. To this end, building on the modeling results from the previous section, the optimization problem is expressed as

$$\begin{aligned} \underset{t_w}{\text{minimize}} \quad & \bar{P}_c(t_w) \end{aligned} \quad (14)$$

$$\text{subject to} \quad \bar{D}(t_w) \leq \bar{D}_m, \quad (15)$$

$$t_w \in \{1, 2, \dots\}, \quad (16)$$

where  $\bar{P}_c(t_w)$  and  $\bar{D}(t_w)$  are defined in Eq. (9) and Eq. (12), respectively.

In general, the optimization problem in (14)-(16) belongs to a class of intractable non-linear integer programming problems [11]. However, thanks to the properties that we have developed in Lemma 1 and Lemma 2, we are able to find the optimal solution to the problem in closed form.

**Theorem 1.** *The optimal parameter value of the optimization problem in (14)-(16) is  $t_w^* = \lfloor t_{wm} \rfloor$ , being  $t_{wm}$  the boundary point of the delay constraint:*

$$t_{wm} = \frac{1}{\lambda} \left( F + \mathcal{W}_0 \left( \frac{H}{e^F} \right) \right) \geq 1, \quad (17)$$

where  $F = ((e^\lambda + 2)\bar{D}_m - t_{su})\lambda + 1$ ,  $H = \lambda t_{su} - e^\lambda \bar{D}_m \lambda - 1$ , and  $\mathcal{W}(x)$  is the Lambert W function [12].

*Proof.* Let  $t_{wm}$  refer to the boundary point of the delay constraint, i.e., the continuous  $t_w$  that satisfies  $\bar{D}(t_w) = \bar{D}_m$ . Based on (12), and by induction on  $\bar{D}(t_{wm}) = \bar{D}_m$ , we can derive it as follows

$$\begin{aligned} e^\lambda &= \frac{t_{wm} + (t_{su} - \frac{1}{\lambda})(1 - e^{-\lambda t_{wm}}) - 2\bar{D}_m}{\bar{D}_m(1 - e^{-\lambda t_{wm}})} \Rightarrow \\ e^{-\lambda t_{wm}} &= -\frac{-\lambda t_{wm} + ((e^\lambda + 2)\bar{D}_m - t_{su})\lambda + 1}{\lambda t_{su} - e^\lambda \bar{D}_m \lambda - 1} \Rightarrow \\ e^{-\lambda t_{wm} + F} &= -e^F \frac{-\lambda t_{wm} + F}{H} \Rightarrow \\ -\lambda t_{wm} + F &= -\mathcal{W} \left( \frac{H}{e^F} \right) \Rightarrow \\ t_{wm} &= \frac{1}{\lambda} \left( F + \mathcal{W} \left( \frac{H}{e^F} \right) \right). \end{aligned} \quad (18)$$

For typical  $\bar{D}_m$  and  $t_{su}$  values,  $1 \ll \frac{F}{\lambda}$  and  $H < 0$ . Therefore, the main branch of the Lambert W function ( $\mathcal{W}_0$ ) can be considered as a solution for (18) that has a value greater than  $-1$ . Accordingly,  $t_{wm} \geq 1$  and Eq. (17) is derived.

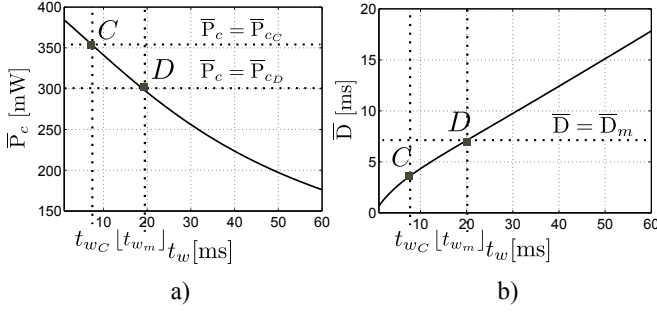


Figure 4: Schematic proof of Theorem 1.

Thanks to Lemma 1 and Lemma 2, we can easily show that  $t_w = \lfloor t_{w_m} \rfloor$  is the optimal solution to the optimization problem in (14)-(16), as detailed next. Figures 4.(a) and 4.(b) show the decreasing trend of the power consumption and increasing behaviour of the delay constraint as a function of  $t_w$ , respectively, which satisfies  $\frac{dP_c}{dt_w} < 0$  and  $\frac{dD}{dt_w} > 0$  (as proved in Lemmas 1 and 2). Consider an arbitrary point  $C$  in the interior of the feasible region for  $t_w$  ( $t_{w_C} < \lfloor t_{w_m} \rfloor$  where  $\bar{D}(t_{w_m}) = \bar{D}_m$ ). As it can be seen from Fig. 4, there is always a point close to the boundary of the delay constraint, denoted by  $D$  ( $t_{w_D} = \lfloor t_{w_m} \rfloor$ ), where its power consumption  $\bar{P}_{c_D}$  is lower than that of  $C$  ( $\bar{P}_{c_D} < \bar{P}_{c_C}$ ). Then, we can conclude that under a given delay constraint, the point  $\lfloor t_{w_m} \rfloor$  always exists and attains the lowest power consumption within the feasible region, and hence it is the optimal solution to the problem in (14)-(16).  $\square$

## V. NUMERICAL RESULTS

In this section, extensive numerical results are provided in order to validate WuS and analytical results, as well as to show and compare the average power consumption of the optimized WuS to that of DRX for low packet arrival rates. Power consumption of the UE in different operating states is highly dependent on the implementation, and also its operational configurations. Therefore, for the numerical results, the power consumption model used in [2], [7], [13] is employed. Its parameters for DRX and WuS are shown in Table I and Table II, respectively. LTE-based power consumption values, shown in Table I, are considered as a practical example while those of the emerging NR modems are not yet publicly available.

Additionally, the approach described in [9] is chosen as a benchmark, since it provides the optimal power consumption in DRX-based reference systems. Furthermore, to the best of our knowledge, virtually all the DRX-based reference work ignore the start-up and power-down energy consumption, and therefore we have modified the approach in [9] slightly in order to consider such additional energy consumption in the optimization.

Two different sets of power consumption results are presented based on the optimal configuration of the w-cycle  $t_w^* = \lfloor t_{w_m} \rfloor$ . Namely, a) with all simplifying assumptions of zero false alarm/mis-detection rates, and  $t_{on} \approx 0$  ms,

Table I: Average transition time and representative power consumption values of LTE-based cellular module during short and long DRX.

DRX Cycle	PW <sub>sleep</sub>	PW <sub>active</sub>	$t_{su}$	$t_{pd}$
short	395 mW	850 mW	1 ms	1 ms
long	$\approx 0$ mW	850 mW	15 ms	10 ms

Table II: Assumed power consumption parameters of the wake-up scheme.

PW <sub>1</sub>	PW <sub>2</sub>	PW <sub>3</sub>	$t_{su}$	$t_{pd}$	$t_{on}$
850 mW	$\approx 0$ mW	57 mW	15 ms	10 ms	1/14 ms

Table III: Optimal values of w-cycle under different delay requirements and packet arrival rates.

$\lambda$ [p/ms]	0.01		0.1		0.2	
$\bar{D}_m$ [ms]	100	600	100	600	100	600
$t_w^*$ [ms]	381	1891	305	1858	312	1922

equivalent to our analytical results (ana.), and b) with the realistic assumptions of  $P_{fa} = 10\%$ ,  $P_{md} = 1\%$ ,  $t_{on} = 1/14$  ms obtained from simulations and [7], termed simulation results (sim.).

Table III shows the optimal values of  $t_w^* = \lfloor t_{w_m} \rfloor$  for different values of  $\lambda$  (0.01, 0.1, 0.2 p/ms) and  $\bar{D}_m$  (100, 600 ms). As it can be observed, for tight delay requirements ( $\bar{D}_m = 100$  ms),  $t_w^*$  tends to be small, enabling the UE to reduce the duration of packet buffering. Interestingly, for mid range of packet arrival rates ( $\lambda = 0.1$  p/ms), optimal w-cycle for a given delay bound is shorter than for both lower and higher packet arrival rates. The justification is as follows. For higher packet arrival rates,  $t_w^*$  becomes larger, the reason being that the inactivity timer is ON most of the time. Therefore, the need for smaller w-cycles decreases and correspondingly higher energy overhead is induced. For lower packet arrival rates, in turn, the value of  $t_w^*$  is higher due to infrequent packet arrival rates, hence imposing smaller delay.

Fig. 5 illustrates the power consumption of the proposed WuS under ideal and realistic assumptions as a function of the packet arrival rate ( $\lambda < 0.2$  p/ms), for different maximum tolerable delays. As it can be observed, for both analytical and simulation results, and for all delay bounds, the average power consumption initially increases, while then remains almost constant as  $\lambda$  increases, due to the configuration of shorter w-cycles for mid packet arrival rates (see Table III). Moreover, the UE consumes higher power in order to satisfy tighter delay requirements, which, as shown in Table III, can be translated to shorter  $t_w^*$ . Furthermore, Fig. 5 shows that the simulation results follow very closely the analytical results, the non-zero gap being due to the non-zero false alarm and mis-detection rates. The relative gap between simulation-based and analytical results is somewhat larger for shorter delay bounds, which is due to the correspondingly higher number of wake-up instances.

Moreover, for comparison purposes, the relative power saving of WuS ( $\eta$ ) representing the amount of power that can be saved compared to the DRX-based reference system



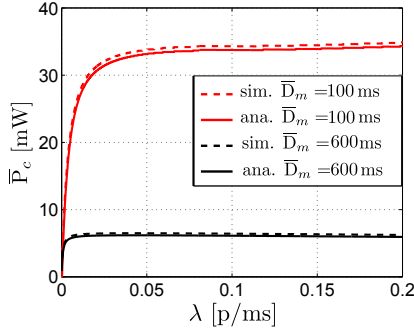


Figure 5: Average power consumption of the optimized wake-up scheme, under ideal and realistic assumptions as a function of delay bound and packet arrival rate.

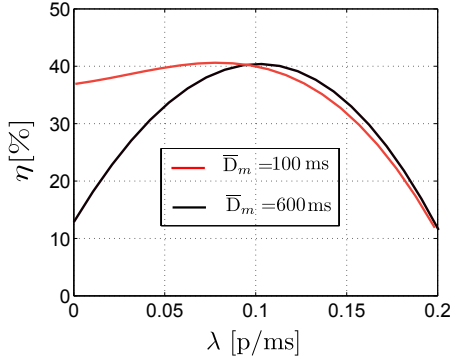


Figure 6: Achieved relative power saving values ( $\eta$ ) of the proposed optimized wake-up scheme as a function of delay bound and packet arrival rate.

is utilized, under the same delay constraints in both methods. The value of the relative power saving ranges from 0 to 100%, and a greater value indicates that the WuS saves energy better than the DRX. Formally, we express the relative power saving as

$$\eta = \frac{\bar{P}_{\text{DRX}} - \bar{P}_c}{\bar{P}_{\text{DRX}}} \times 100, \quad (19)$$

where  $\bar{P}_{\text{DRX}}$  refers to average power consumption of DRX. Furthermore, for a fair comparison, we consider an exhaustive search over a large parameter set of DRX configuration, developed in [9]. However, in order to take start-up and power-down power consumption into account, the solution is slightly modified to account for the transition states.

Fig. 6 shows the power saving results. It is observed that the proposed WuS, under realistic assumptions, outperforms DRX for low packet arrival rates, especially for lower packet arrival rates with tight delay requirements. The main reason is that in such scenarios, DRX-based devices need to decode the control channel very often, operating mainly in short DRX cycle, which causes additional power consumption. The WRx, in turn, needs to decode the wake-up signaling frequently, but with lower power overhead. Additionally, as expected, regardless of the delay requirements, for higher packet arrival rates, DRX has relatively similar power consumption as the WuS. The reason is that in such cases, the DRX parameters can be configured in such a way that there is a low amount of

unscheduled DRX cycles, either by utilizing short DRX cycles for very tight delay bounds or by employing long DRX cycles for less strict delay requirements. Overall, the results in Fig. 6 clearly demonstrate that WuS can provide substantial energy-efficiency improvements over DRX.

## VI. CONCLUSIONS

In this paper, the average energy consumption and delay of wake-up enabled MTC UE are analyzed and then optimized by means of a constrained optimization problem. It is shown that the performance of the wake-up scheme is mainly governed by the w-cycle. The optimal w-cycle configuration in terms of energy consumption is obtained in closed form as a function of the maximum delay requirement and fixed packet arrival rate. Analytical and simulation results show that the proposed optimized WuS outperforms the corresponding optimized DRX-based reference system in terms of energy consumption. Thus, it is an efficient approach that should be considered in future MTC devices to reduce their energy consumption, while ensuring a latency target as well as QoS.

## REFERENCES

- [1] "NR; overall description; stage-2," 3GPP TS 38.300 v15.5.0 Release 15, Tech. Rep., Apr. 2019.
- [2] M. Lauridsen, "Studies on mobile terminal energy consumption for LTE and future 5G," Ph.D. dissertation, Aalborg University, Jan. 2015.
- [3] S. Rostami, K. Heiska, O. Puchko, K. Leppänen, and M. Valkama, "Robust pre-grant signaling for energy-efficient 5G and beyond mobile devices," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.
- [4] "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," 3GPP TS 36.213 version 10.1.0 Release 10, Tech. Rep., APR. 2010. [Online]. Available: <http://www.3gpp.org>.
- [5] "NR; User Equipment (UE) procedures in idle mode and in RRC Inactive state," 3GPP TS 38.304 v15.2.0 Release 15, Tech. Rep., Jan. 2019.
- [6] "UE power consideration based on days-of-use," Qualcomm Incorporated, R1-166368, Tech. Rep., Aug. 2016.
- [7] S. Rostami, K. Heiska, O. Puchko, J. Talvitie, K. Leppänen, and M. Valkama, "Novel wake-up signaling for enhanced energy-efficiency of 5G and beyond mobile devices," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec 2018, pp. 1–7.
- [8] S. Rostami, K. Heiska, O. Puchko, K. Leppänen, and M. Valkama, "Wireless powered wake-up receiver for ultra-low-power devices," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018, pp. 1–5.
- [9] H. Ramazanali and A. Vinel, "Tuning of LTE/LTE-A DRX parameters," in *2016 IEEE 21st International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, Oct 2016, pp. 95–100.
- [10] L. Kleinrock, *Theory, Volume 1, Queueing Systems*. Wiley-Interscience, 1975.
- [11] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, p. 1–131, 2013.
- [12] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, Dec 1996. [Online]. Available: <https://doi.org/10.1007/BF02124750>
- [13] C. C. Tseng, H. C. Wang, F. C. Kuo, K. C. Ting, H. H. Chen, and G. Y. Chen, "Delay and power consumption in LTE/LTE-A DRX mechanism with mixed short and long cycles," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1721–1734, March 2016.