IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 14, 2020, accepted October 27, 2020, date of publication November 2, 2020, date of current version November 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035186

Evolution of Power Saving Technologies for 5G New Radio

TAEHYOUNG KIM[1], YOUNSUN KIM[1], (Senior Member, IEEE), QIONGJIE LIN[2], FEIFEI SUN[3], JINGXING FU[3], YOUNGBUM KIM[1], ARIS PAPASAKELLARIOU[2], HYOUNGJU JI[1], (Senior Member, IEEE), AND JUHO LEE[1], (Fellow, IEEE)
[1]Samsung Research, Samsung Electronics Company Ltd., Seoul 06765, South Korea
[2]Samsung Research America, Inc., Mountain View, CA 94043, USA
[3]Samsung Research and Development Institute China-Beijing (SRC-B), Beijing 10028, China

Corresponding author: Hyoungju Ji (hyoungju.ji@samsung.com)

ABSTRACT Battery life of mobile terminals has been a key performance indicator (KPI) for user experience since the early days of wireless communications. This KPI has become even more important for fifth generation (5G) cellular systems due to higher requirements such as the support of carrier frequency up to 100 GHz, carrier bandwidth up to 400 MHz, peak data rate up to 20 Gbps, latency as small as 1 ms, and enhanced receiver performance. In response, various power saving techniques have been studied and implemented so that devices can be used over a longer period without noticeable decrease in performance. In this article, we provide a comprehensive overview of the latest terminal power saving techniques which were specified in 5G New Radio (NR) over the last couple of years. The key feature and benefit of each power saving technique is analyzed with simulation results under various traffic scenarios. In addition, some of the candidate techniques for the next phase of NR on terminal power saving enhancement are proposed targeting not only conventional data services but also various vertical services.

INDEX TERMS Battery life, power saving, 5G, new radio.

I. INTRODUCTION

Mobile wireless communications have profoundly changed human life over the last 20 years. Terminals in various forms enable humans and machines to connect wirelessly without any restrictions on location or time. Taking advantage of such wireless connectivity, new services such as real time broadcasting, high definition video streaming, mobile cloud, wireless backhaul, and remote machine control are expanding at an exponential pace [1]. These new applications take advantage of the high data rate and reduced latency offered by the latest wireless technology such as fifth generation (5G) New Radio (NR). However, the high performance comes with a cost in the form of rapid battery power consumption. Over the last 10 years, wireless technology has evolved to achieve an improvement in data rate by a factor of more than 10. However, during the same period, battery capacity has only improved by a factor of 2 or 3 [2]. As a result of such disproportionate improvement in battery capacity compared to processing requirement, the battery life of the latest terminals is no better than what it was a decade ago and consumers still consider battery life as one of the most important criteria when choosing a phone [3]. Such demands have led to ITU defining energy efficiency improvement by a factor of 100 as a key requirement for 5G [4].

Compared to fourth generation (4G) Long Term Evolution (LTE), NR is designed to operate in a much higher frequency band (up to 100 GHz) with bandwidth that is 20 times larger (up to 400 MHz) using twice the number of receiver antennas (up to 4 mandatory antennas) at a terminal [5]. Each of these features is critical in realizing the key use cases of 5G such as enhanced mobile broadband (eMBB), ultra-reliable low latency communications (URLLC), and vehicular-to-everything communications (V2X) [6]. However, the downside of these features is that they all require significant battery power consumption. As a result, much effort has been spent in the 3rd Generation Partnership Project (3GPP) when specifying 5G so that minimal battery power is consumed when realizing the high performance use cases [7].

A. RELATED WORK

Over the last decade, there have been extensive efforts for enabling power-efficient mobile communication [8], [9]. Discontinuous reception (DRX) which is the most widely

The associate editor coordinating the review of this manuscript and approving it for publication was Pietro Savazzi.

used power saving mechanism in commercial networks has been extensively studied in [10]–[15]. DRX allows the terminal to reduce power consumption by requiring its receiver radio modules to be activated only over short time intervals [10]. In [11], the DRX mechanism is formulated as a multi-objective optimization problem satisfying the latency requirements of active traffic flows and the corresponding preferences for power saving. In [12], DRX is modeled as a Markov chain with long and short DRX cycles with analysis on the power-saving maximization satisfying a wake-up delay requirement. In [13], a new DRX mechanism called directional DRX (DDRX) is proposed for supporting the directional air interface expected in mmWave enabled 5G communications under non-standalone deployment scenario. The work in [13] has been expended to more generalized scenarios including standalone 5G system and sidelink in [14]. In [15], flexible mechanisms for DRX have been studied by exploiting channel capacity prediction to jointly optimize resource allocation and DRX configuration for video streaming applications.

To further minimize the power consumption at terminals operating with DRX, wake-up based approaches have been studied in [16]–[19]. [16] provides a comprehensive overview of various wake-up schemes for wireless sensor networks considering the benefits and the challenges. In [17], an energy-efficient wake-up receiver design that takes into account actual analog front-end and digital base-band architectures is provided along with analysis on wake-up beacon detection performance and false alarm probabilities. In [18], stochastic geometry is used to analyze the performance of an energy-efficient joint downlink and uplink wake-up solution for Internet of Things (IoT) devices over cellular networks. In [19], the average power consumption of a wake-up radio enabled mobile device is analyzed by using a semi-Markov process. In [20], a new concept of pre-grant message is introduced to improve the power saving gain of the terminal by minimizing the control channel monitoring.

On other aspects, several studies that handle the power consumption issue of terminals operating with carrier aggregation (CA) across multiple component carriers (CCs) [21], [22]. The terminal transmitting/receiving a signal through the multiple CCs consumes more power due to large signal processing complexity [23], [24]. In [21], a new dynamic CA scheduling scheme is proposed to improve the energy efficiency of uplink communications. This study has been expended in [22] for both downlink and uplink scenarios by jointly optimizing bandwidth and power allocation.

B. CONTRIBUTIONS & ORGANIZATION

In this article, we present the latest standardization features and potential future directions for 5G NR in 3GPP. The core of these efforts is to enable power saving by adaptively optimizing radio operation based on the user's data usage characteristics. Cellular data is often observed to have a random inter-arrival time that is often modeled with a heavy-tailed distribution [25]. Also, the time spent in queue and the time spent for actual transmission over the air make up only a very small portion of the entire time the terminal is connected. In other words, most users consume power not for processing actual data but from waiting for the data to be received. For example, in the downlink where a terminal receives data from a base station, it was observed in field tests that the percentage of occupied time resources per terminal and the probability for a terminal being scheduled data on two consecutive time resources were both less than 20% [7]. However, since a terminal does not have prior knowledge on when a base station schedules data, it has to constantly monitor the control messages, resulting in battery power wastage.

In order to realize 5G high performance use cases and also minimize power consumption during the standby time, NR supports specification features that target power saving without compromising the quality of services such as latency or system throughput. As a first step, in Rel-15 NR, the interval during which a terminal checks for downlink data reception was reduced [10], [26]. In addition, when a terminal wakes up to check, the observation bandwidth and the number of search trials on the control channel were reduced such that further power consumption would be possible [26], [27]. In Rel-16 NR, enhanced power saving techniques such as more efficient control if terminal's monitoring states, relaxation of required processing latency, and adaptive activation of antennas at the terminal [7] were supported. However, despite these efforts, the goal of meeting the 5G requirement of a power savings by a factor of 100 was still out of reach. This paper discusses the latest ongoing researches and standardization efforts in the wireless industry to achieve this goal.

Our main contributions are summarized as follows.

- This paper provides an overall picture of the efforts on terminal power consumption reduction that have been discussed so far and the features supported in current 5G NR as a result of this study. To this end, progressive evolution of power saving techniques in 5G NR over the recent releases is presented. Comprehensive examples are provided so as to illustrate how various power saving techniques can be combined and operated systematically.
- Key system models for terminal power state, power consumption, and activity which were used in the development of power saving technologies are provided. The differences between the models for each power saving technique are explained. In addition, the effectiveness of power saving technologies are compared and analyzed through simulations using these models under various traffic types.
- Potential future directions on the evolution of power saving techniques are provided. Such power saving techniques are being considered for application for not only conventional data services such as eMBB but also for new vertical services such as sidelink and low-cost terminals as well as data services in unlicensed spectrum.

TABLE 1. **Summary of notations.**

Symbol	Description
P_{ds}	Relative power for deep sleep state
P_{ls}	Relative power for light sleep state
P_{ms}	Relative power for micro sleep state
P_c	Relative power for PDCCH only state
P_d	Relative power for PDCCH+PDSCH state
P_p	Relative power for paging monitoring
P^{total}	Total power consumption
P^{total}_{ref}	Total power consumption for baseline scheme
P^{total}_{prop}	Total power consumption for power saving scheme
L_{ref}	Latency for baseline scheme
L_{prop}	Latency for power saving scheme
t_{arrival}	Time for packet arrival
t_{schedule}	Time for packet scheduling
$L_{\max}$	Number of maximum MIMO layers
p_1	Default periodicity for control channel monitoring
p_2	Indicated periodicity for control channel monitoring
t_{inact}	Inactivity timer
t_{sleep}	Sleep duration
γ_g	Group paging rate
γ_{ue}	Individual paging rate
M	Number of terminals in a paging group
N_{ss}	Number of monitored synchronization signal

For convenience, notations used throughout this article are summarized in Table 1.

The rest of this article is organized as follows. Section II provides some fundamental models used in evaluation and design of power saving mechanisms. Section III introduces the conventional techniques specified in current NR specifications. Section IV describes candidates of power techniques that are under study for the continuing evolution of 5G NR. Simulation results are provided to verify the benefits of these new approaches. The final section provides concluding remarks.

II. KEY MODELS ON TERMINAL POWER SAVING

Specification support for terminal power saving in a cellular network refers to a set of features that enables the cellular network to apply a proper configuration to balance the trade-off between the terminal's receiver modem power consumption and network performance. To better understand the mechanics of terminal power saving, some understanding of the key models used for terminal state, power consumption including scaling, and traffic, is necessary. A brief summary of such models is described in this section.

A. TERMINAL'S STATE IN THE NETWORK

The energy efficiency of wireless communications can be determined by the efficiency of data transmission in a loaded network and by energy consumption when there is no data at the terminal [4]. This aspect can be simplified to two modes of a terminal in the network: network access mode (NAM) and power saving mode (PSM) (see Figure 1). NAM corresponds to the active state of the terminal where it can transmit or receive data. For NR system, in NAM, the terminal monitors the control channel called physical downlink control channel (PDCCH) and receives a downlink data channel named physical downlink shared channel (PDSCH) or transmits uplink data channel called a physical uplink shared channel (PUSCH). PSM corresponds to the sleep state in which a terminal operates in a very low power by switching off radio frequency (RF) and/or baseband components.

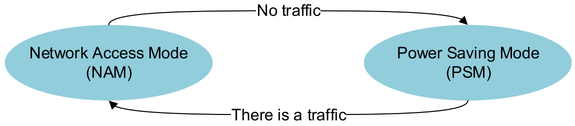


FIGURE 1. **Transitioning model between network access mode and network idle mode according to data traffic.**

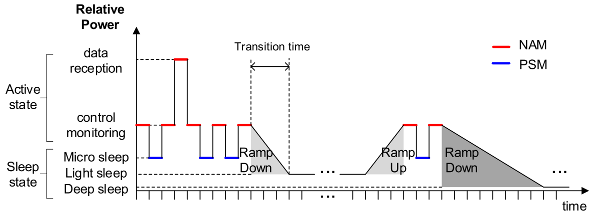


FIGURE 2. **Terminal power consumption model according to functionalities.**

Power saving technique was designed in the following two objectives. The first objective is to maximize the energy efficiency of data transmission during NAM and the second objective is to develop access technology that maximizes the period of PSM. To achieve the first objective, the power saving techniques focus on reducing and/or adapting various parameters, including bandwidth size for receptions or transmissions, the number of active antennas, and the terminal processing time. To achieve the latter objective, techniques focus on minimizing the switching latency from NAM to PSM and vice versa based on dynamic signaling.

B. POWER CONSUMPTION AND SCALING MODEL

Figure 2 depicts an exemplary view of a terminal's power consumption according to key functionalities. In general, the state in which a terminal receives a control channel or data channel in NAM is referred to as the active state. On the other hand, the state in which a terminal is in standby in PSM is referred to as a sleep state. In the active state, a terminal switches on its RF and baseband components to transmit/receive control and data information so that it consumes a lot of power. In the sleep state, the terminal can deactivate and bypass its RF and baseband components except for essential functionalities such as synchronization and tracking.

Depending on the degree of deactivation, the sleep state can be categories into deep sleep, light sleep, and micro sleep [29]. The deep sleep allows the lowest power consumption with limited activities in baseband such as minimum timing activity and no activity in RF circuits. In light sleep, a more stringent timing is maintained by an internal clock and the terminal's activity level which allows a transition to an active state with a reasonably small delay. A terminal in micro sleep cannot transmit or receive data but is able to do a quick transition to an active state. The more deactivated components, the longer it takes by transition to the

TABLE 2. **Power Consumption Model.**

Power state	Relative power	Transition energy	Transition time
Deep sleep	1(ref)	450	20 ms
Light sleep	20	100	6 ms
Micro sleep	45	0 (ref)	0 ms
PDCCH-only	100	-	-
PDCCH+PDSCH	300	-	-

TABLE 3. **Power Scaling Model.**

Scaling parameters	Scaling model $\gamma(X)$
Bandwidth	$\gamma(X) = 0.4 + 0.6 \times (X - 20)/80$ where for $X = 10, 20, 40, 80,$ and 100 MHz
Carrier	$\gamma(X) = 1.7 \times X/2$ (e.g., power for 2 carriers = power for 1.7×1 serving cell)
Antenna	$\gamma(X) = 0.7 \times 2X$ (e.g., power for 2 RX = power for 0.7×4 RX)
Scheduling reservation	$\gamma = 0.7$ where it is applicable for P_c (e.g., P_c for same-slot scheduling = power for 0.7 $\times$ P_c for cross-slot scheduling)

active state. For active state, the two states, PDCCH-only and PDCCH+PDSCH state, are considered in Rel-16 in NR specification. Relative power consumption for each state described above is summarized in Table 2 [7].

The power consumption for each state follows the scaling model defined in Table 3 [7]. This table is modeled based on the actual measurements from terminal manufacturers. A scaling model was derived for the following four items (bandwidth, number of carriers, number of antennas, and scheduling reservation) among various RF-related system parameters. Note that the power consumption increases as the bandwidth or the number of antennas used increases but decreases as the activity decreases.

C. TRAFFIC MODEL

Today's commercial wireless networks offer a variety of data applications such as voice, video streaming, web browsing, gaming, and instant messaging. These applications commonly have burst traffic characteristics with each data packet having a short duration, followed by an arrival interval of a few milliseconds to a few seconds before the next data packet is generated [25].

To reflect such traffic characteristics, three representative traffic models are considered: file transfer protocol (FTP), instant messaging, and voice-over internal protocol (VoIP) as in Table 4 [7]. FTP model 3 is used to model transmission control protocol (TCP) based file transfer including instant messaging. In FTP model 3, the packet arrival is modelled with the Poisson process which results in an exponential distribution of inter-arrival time [30]. For the VoIP model, a two-state activity model is used to generate an encoder of 40 bytes of packet size every 20 msec [31].

In such bursty traffic environments, the terminal is scheduled data only in a very small portion of the available time resources. However, even if the user has no scheduled data, the terminal still needs to monitor the control channel which ends up using the majority of its battery power. Based on the power consumption model defined in Table 2, it is observed that 99.75%, 99.97% and 98.6% of the active time is consumed for control channel monitoring without any scheduled data for FTP, instant messaging, and VoIP, respectively when no power saving technique is used [7]. With this observation, maintaining the sleep state by minimizing the time duration for wake-up as much as possible when there is no data is the most important thing for terminal's power saving.

TABLE 4. **Traffic Model.**

Traffic Type	Model	Packet size	Inter-arrival time
FTP traffic	Model 3 [30]	0.5 MB	200 ms
Instant messaging	Model 3 [30]	0.1 MB	2 sec
VoIP	Real-time Transport Protocol (RTP) for Adaptive Multi-Rate (AMR) as defined in [31]		

III. CONVENTIONAL POWER SAVING TECHNIQUES

This section briefly discusses conventional power saving techniques including discontinuous reception, bandwidth adaptation with flexible control channel which are specified in the first release of NR.

A. DISCONTINUOUS RECEPTION

DRX is the most typical power saving method in which terminals receive signals from the network discontinuously where the wake up time and sleeping time are pre-determined according to the network configuration [10]. In addition, when the terminal wakes up, the network can instruct the terminal whether to stay awake or fall asleep again to save power. DRX was firstly introduced in the 2G Global System for Mobile Communications (GSM) and it continues to play an important role in 3G, 4G, and 5G networks.

Three system parameters are basically used to operate DRX. The first is the *drx-Cycle*, the second is *drx-onDurationTimer*, and the last is *drx-InactivityTimer* [32]. *drx-Cycle* refers to the periodicity that the terminal wakes up from the sleep state and monitor the control channel. *drx-onDurationTimer* is a timer for checking whether a control channel exists or not after the terminal wakes up. If the control channel is not detected until the timer ends, the terminal goes back to sleep mode. If a control channel is detected, *drx-InactivityTimer* is incremented so that the DRX active time increases. When all incremented timers are expired, the terminal enters sleep mode (refer to upper sub-figure in Figure 3).

B. BANDWIDTH ADAPTATION

Compared to the maximum bandwidth of 20 MHz supported by LTE, NR supports up to 400 MHz of bandwidth for ultra-fast transmission [5]. However, in order to receive a wideband signal, the terminal needs a wide filter, a high-dimensional fast Fourier transform (FFT), and a high-performance analog-to-digital converter (ADC), and these factors consume significantly more power than LTE. To solve this problem, 5G introduced the concept of bandwidth part (BWP) [27]. BWP refers to the portion of the system bandwidth (often

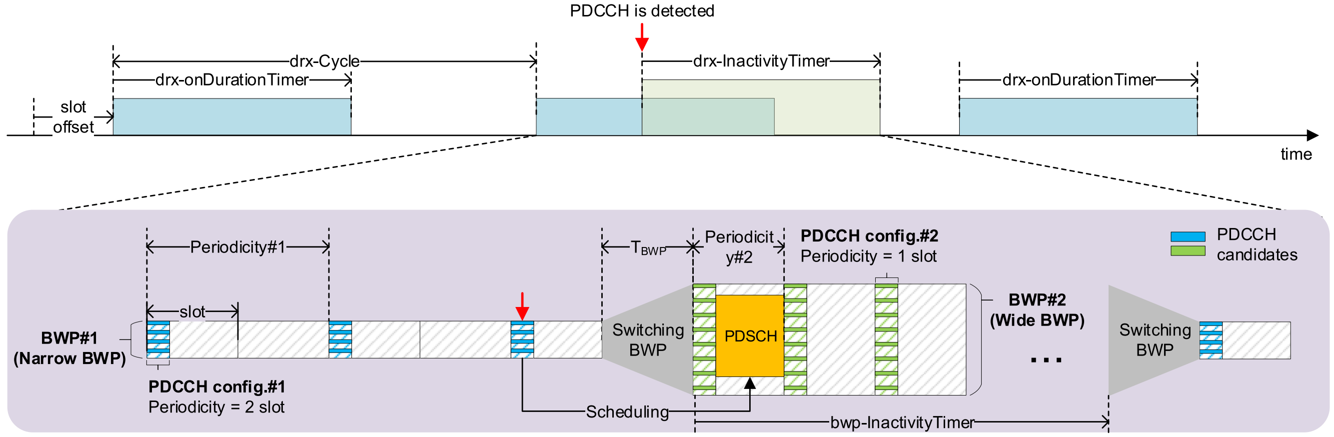


FIGURE 3. **Illustration of power saving techniques in Rel-15 NR; DRX (upper) and BWP adaptation (lower).**

called channel bandwidth) that the terminal is configured to receive or transmit signals from. In other words, even if the system bandwidth is large, if the BWP is configured to be smaller than the system bandwidth, the terminal can only tune the RF bandwidth to the BWP and thereby avoid unnecessary power consumption. BWP plays an important role in power saving, especially when the data rate served by the terminal is small compared to the system's peak data rate. In 5G, the system can allocate up to 4 different sizes of BWP, and only one of them is activated at a given time. The network can switch active BWP based on an indication via a control channel or based on a timer called *bwp-InactivityTimer*. If the timer is expired due to lack of traffic, the terminal switches its active BWP to the pre-configured default BWP which typically has a narrow bandwidth for reasons of power saving. If BWP switching is triggered, the terminal can activate the new BWP until the time no later than BWP switching delay T_{BWP} [28].

Figure 3 illustrates how DRX and BWP operate together. The terminal periodically repeats wake-up and go-to-sleep based on the DRX operation. Once the terminal wakes up at a certain DRX occasion, the terminal monitors its control channel at a narrow BWP (BWP#1) until traffic actually arrives. At this time, if data is generated and the control channel is detected as shown in the red arrow in Figure 3, the system can instruct the terminal to use a wider BWP (BWP#2). When all the data is transmitted, the terminal is changed back to the narrow BWP (BWP#1) based on the *bwp-InactivitiyTimer* expiration to minimize power consumption. To further adjust the trade-off between power saving and data rate, the control channel monitoring can be differently configured per BWP. As depicted in Figure 3, a control channel monitoring configuration in BWP#1 can target sparse or latency-tolerant traffic by having a smaller number of candidates and a longer monitoring periodicity to save terminal power. On the other hand, a control channel monitoring configuration in BWP#2 can target frequent or low latency traffic through a larger number of candidates and a shorter monitoring periodicity to enable higher data rate or lower latency [26].

IV. ENHANCED POWER SAVING TECHNIQUES

The power saving technique in the time and frequency domains that was initially introduced has been consequently expanded to antenna domain, carrier domain, and other areas in the second release of NR [7]. In the following paragraphs, four such techniques and one assistance information are introduced: 1) Dynamic indication of control channel monitoring, 2) Dynamic scheduling reservation, 3) Dynamic antenna adaptation, 4) Enabling dormancy behavior of secondary cells (SCells), and terminal assistance information.

A. TECHNIQUE#1: DYNAMIC INDICATION OF CONTROL CHANNEL MONITORING

The disadvantage of the existing DRX method is that the terminal needs to wake up periodically even though there is no data arrival. Due to this fact, the first method, called dynamic indication of control channel monitoring (DIC), was introduced to avoid unnecessary wake up in the upcoming DRX cycle. This indication is provided by a new downlink control information (DCI) format, DCI format 2_6, which is carried on a control channel. The network configures the terminal to monitor a control channel for DCI format 2_6 at occasion(s) prior to slots corresponding to each DRX occasion. In consequence, the terminal monitors DCI format 2_6 outside DRX active time. Based on the information in the detected control information in DCI format 2_6, the terminal can determine whether to start *drx-onDurationTimer* or not (see upper subfigure in Figure 4). Through this method, the network can reduce the terminal's power consumption by adjusting the DRX cycle on/off according to traffic conditions.

B. TECHNIQUE#2: DYNAMIC SCHEDULING RESERVATION

A data channel can be scheduled by either same-slot scheduling or cross-slot scheduling. Same-slot scheduling refers to the case where the data channel is transmitted in the same slot as the corresponding control channel, while cross-slot scheduling refers to the case where the data channel is transmitted in a slot that comes after the slot on which the corresponding control channel is transmitted. Typically, same-slot

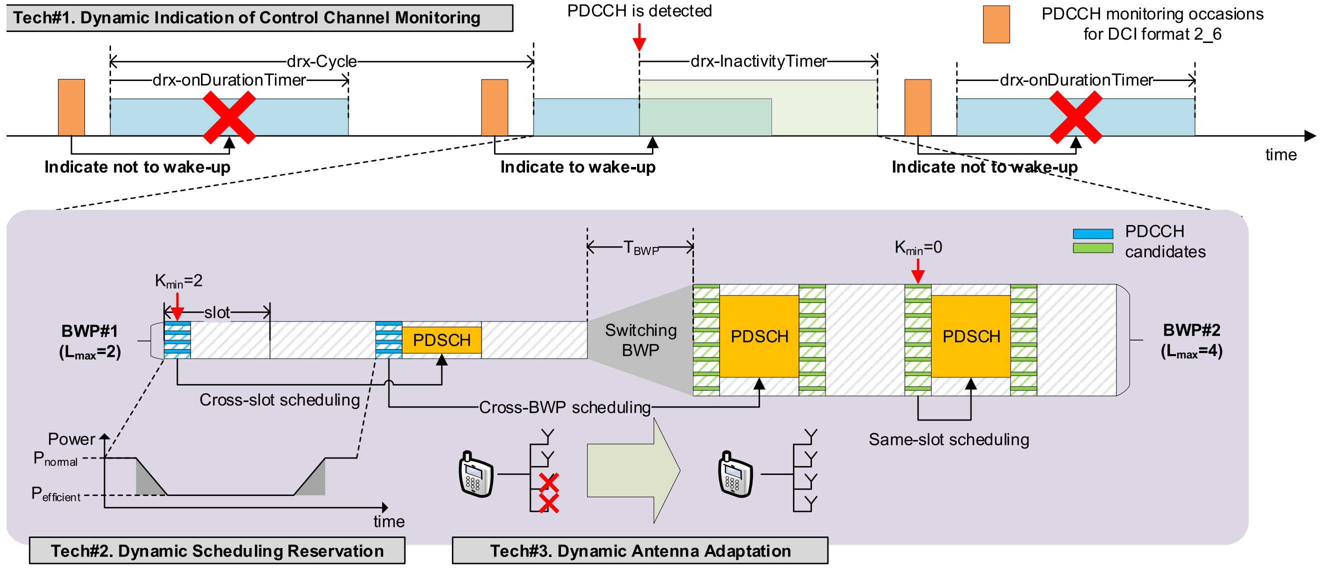


FIGURE 4. Illustration of enhanced power saving techniques in Rel-16 NR; dynamic indication of control channel monitoring, dynamic scheduling reservation, and dynamic antenna adaptation.

scheduling requires a tighter processing timeline compared to cross-slot scheduling. Since the terminal does not know the scheduling type prior to decoding the control channel, the terminal needs to assume the worst case scenario in terms of processing timeline, i.e., same-slot scheduling, and be ready for immediate data channel reception and processing. Consequently, this increases terminal's power consumption.

Dynamic scheduling reservation (DSR) is a method where the terminal is informed in advance that the data channel would be in a slot that comes after the slot of a control channel. The network can indicate to the terminal that no data channel is to be scheduled during a number of slots (denoted by K_{min} in Figure 4) after the control channel reception. If the terminal knows in advance that a scheduled data would be in a later slot than the slot of a control channel reception, the terminal power consumption could be reduced due to at least:

- A relaxed control channel processing time
- Avoidance of buffering for potential data channel reception that could start prior to the completion of the control channel processing
- Ability for the terminal to operate in micro sleep mode during the time where data channel will not be scheduled

As illustrated in Figure 4, if a terminal can assume cross-slot scheduling for data, the terminal can operate in low power ($P_{\text{efficient}}$) mode during a time gap between control and data transmission. Operation in low power mode shows up to 30% power reduction compared to operation in normal power (P_{normal}) mode as described in Table 3 [7].

C. TECHNIQUE#3: DYNAMIC ANTENNA ADAPTATION

The increase in the number of antenna is one of the key factors that consume power. Although the benefit of increasing channel capacity obtained by using MIMO systems is important, for low traffic loads and/or latency-tolerant traffic, MIMO transmission may not be necessary. Thus, it is desirable to minimize the number of active antennas in order to reduce power consumption while maintaining coverage. To this end, dynamic antenna adaptation (DAA) is introduced that enables changing a number of maximum MIMO layers (L_{max}). The terminal can activate only L_{max} antennas in order to conserve power since a data with a number of layers larger than L_{max} will not be scheduled. DAA is achieved under the BWP framework. Different L_{max} values can be assigned for each BWP. As illustrated in Figure 4, if BWP#1 is an active BWP the terminal can activate $L_{\text{max}} = 2$, while if BWP#2 is activated, the terminal can adapt its antennas to $L_{\text{max}} = 4$. By reducing the maximum number of active antennas, the terminal can reduce power consumption by skipping channel estimation for unused antennas as well as turn off RF components connected to each antenna. The DAA technique can be applied for both downlink data reception and uplink data transmission by adjusting the maximum number of MIMO layers for downlink and uplink, respectively. For downlink, the maximum number of MIMO layers is explicitly configured to the terminal, while for uplink it is implicitly informed to the terminal in terms of uplink reference signal configurations [33], [34].

D. TECHNIQUE#4: DORMANCY BEHAVIOR OF SECONDARY CELLS

In order to support larger bandwidth, carrier aggregation (CA) has been supported since LTE. Even in NR, CA is an important function but has a disadvantage in terms of power consumption. Basically, in CA operation, there are a primary cell (Pcell) that manages network access and SCell(s) for additional bandwidth expansion. In order to increase the data rate, it is natural to operate all the cells at the same time, but it is a great loss from a power saving point of view to continuously receive data on SCell(s) even when there is no data transmission. However, deactivating and reactivating the SCell(s) requires a lot of power consumption and a long latency since a series of core procedures such as automatic gain control (AGC), channel state information (CSI) measurement, and beam management (BM) are required in the process.

To solve this problem, a new technique, referred to as dormancy secondary cell (D-SCell), has been introduced to put the SCell to sleep temporarily. For a SCell under dormancy, a terminal does not monitor control channel but continues the core processes. Thereby, the terminal can save power by avoiding unnecessary control channel monitoring when there is no traffic on the SCell. Furthermore, when traffic occurs on a SCell, the SCell can be switched to non-dormancy state with lower latency compared to the conventional latency for activating the SCell since the terminal still performs the core processes. Dormancy indication for a SCell is based on the BWP framework. A SCell can be configured with non-dormant BWP and dormant BWP as shown in Figure 4. If dormant BWP of a SCell is activated, it is assumed that the SCell is on dormancy state by the terminal. Since the terminal can see up to 16 cells, a cell group-based indication was introduced to switch to the dormancy mode more effectively. The terminal can be configured with multiple SCells and set SCell groups (SCGs) such as SCG#1, SCG#2 as illustrated in Figure 4. The bitmap information for dormancy indication is applicable to each SCG. The network can control the power consumption generated by the CA by switching the SCG from the active state to the dormancy state according to the traffic status.

E. TERMINAL ASSISTANCE INFORMATION

The method presented so far is a method where the network triggers the terminal's power saving operation by adjusting the state of RF or baseband of the terminal. However, since the network cannot grasp all the terminal's status (mobility, battery status, and changes in link quality) in real time, there is limitation to use the power saving techniques only relying on the instructions of the network. In order to make it more efficiently, terminal can deliver the following information to the network:

- Mobility log and history
- Preferred DRX, BWP and SCell configurations
- Minimum scheduling offset for cross-slot scheduling
- Preferred parameters for power saving
- Number of maximum MIMO layers

After receiving assistance information, the network can reconfigure system parameters. A more detailed reporting procedure can be found in [33].

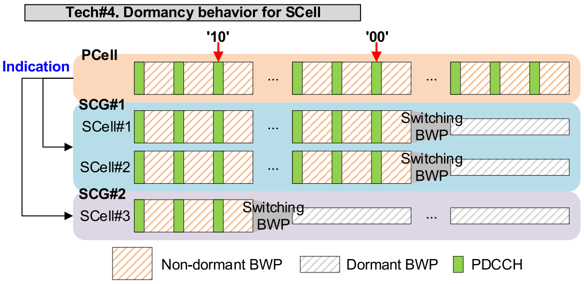


FIGURE 5. Illustration of enhanced power saving techniques in Rel-16 NR; dormancy behavior of SCell.

TABLE 5. Evaluation Assumptions for DRX Configuration.

Traffic Type	FTP traffic	Instant messaging	VoIP
drx-Cycle	160 ms	320 ms	40 ms
drx-InactivityTimer	100 ms	80 ms	10 ms
drx-onDurationTimer	8 ms	10 ms	4 ms

F. PERFORMANCE EVALUATION

In this sub-section, we provide evaluation results for each of power saving technique. Typically, power saving technique can reduce power consumption, on the other hand, there can be a loss in data rate due to the increase in latency. Given that, power saving gain (PSG) and latency loss (LL) are considered as performance metrics which are respectively defined as

$$PSG = \frac{P_{ref}^{total} - P_{prop}^{total}}{P_{ref}^{total}}, \tag{1}$$

$$LL = \frac{L_{prop} - L_{ref}}{L_{ref}}, \tag{2}$$

where P_{ref}^{total}, P_{prop}^{total}, L_{ref}, and L_{prop} are total power consumption of baseline scheme, total power consumption of power saving technique, latency of baseline scheme, and latency of power saving technique, respectively [8]. Total power consumption is calculated as

$$P^{total} = \sum_{i \in \mathcal{S}} P_i \times T_i, \tag{3}$$

where $\mathcal{S}$ is a set of power states given in Table 2, i.e., $\mathcal{S} = \{ds, ls, ms, c, d\}$, and P_i, T_i, denote relative power for state i and time occupied for state i where $i \in \mathcal{S}$, respectively. Latency is calculated as

$$L = t_{\text{arrival}} - t_{\text{schedule}}, \tag{4}$$

where t_{arrival} is time where the packet is arrival and t_{schedule} is time where the arrived packet is scheduled and transmitted.

DRX operation is considered as the baseline scheme with parameters given in Table 5. The evaluations are performed based on the parameters for the power consumption model in Table 2, the power scaling model in Table 3, and the traffic model in Table 4. Especially for DAA, we assume that a terminal is configured with two BWPs with different L_{max} as shown in Figure 4. For D-SCell, SCell is assumed to be configured with 20 MHz dormancy BWP and 100 MHz non-dormancy BWP. For dormancy behavior switching criterion, the network switches dormancy BWP to non-dormancy BWP

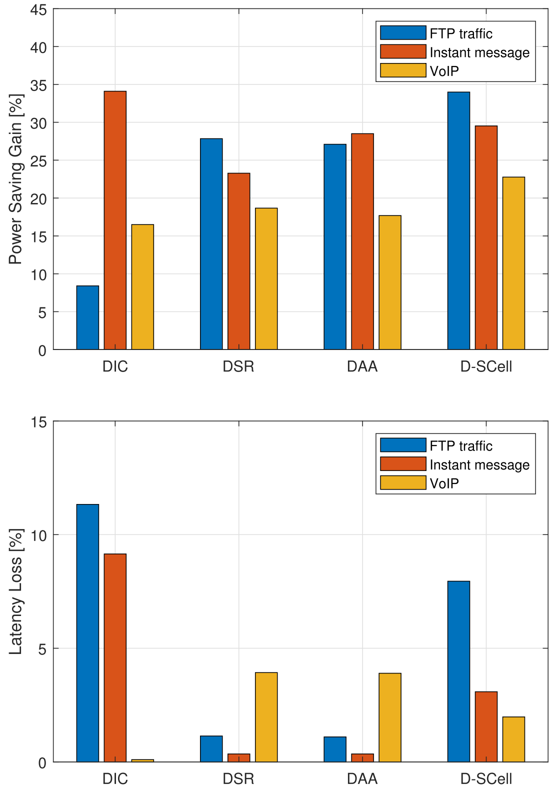


FIGURE 6. **Evaluation results for enhanced power saving techniques; power saving gain (upper) and latency loss (lower).**

when the data buffer has more than 2 packets. The performance of PSG and LL for each power saving scheme is shown in Figure 6.

First, in the performance of each power saving scheme in FTP traffic, using the D-SCell method has the best gain with 34% PSG. When using DSR and DAA, PSG is lower than D-SCell, but it can be seen that LL performance is better. On the other hand, when DIC is used, PSG is the smallest (8.4%) and LL is very large (11.3%). What these results imply is that when the traffic is large but sparse, it is advantageous to process data as quickly as possible without considering power consumption and minimize the power consumption for control channel monitoring from the rest of the time.

Second, observing the results of instant message traffic, the performances of DSR, DAA, and D-SCell are generally similar to FTP model (about 25%∼30% PSG with very low LL). However, in the case of DIC, we can observe that its performance is very different from FTP. LL is slightly reduced but PSG is very high, which is superior to the other three techniques. The difference between FTP and instant message is that the instant message has a file size of 1/5 less and the interval between traffic occurrences has increased 10 times. That is, when the file size is small, the terminal can reduce power consumption by adjusting the control channel monitoring more dynamically. This tendency is more evident when considering with VoIP where file size is very small.

Finally, looking at the performance in VoIP, all four techniques can have PSG benefit (approximately 20% PSG).

TABLE 6. **Summary of power saving techniques supported in 5G NR.**

Release	Power saving techniques	Key features
Rel-15 NR	DRX	Discontinuous reception based on periodic wake-up and sleep
	BWP	Allocating and adapting bandwidth size of terminal
Rel-16 NR	DIC	Dynamic indication of wake-up for the next DRX cycle
	DSR	Dynamic indication of minimum scheduling offset for data
	DAA	Dynamic adaptation for maximum MIMO layers
	D-SCell	Dynamic adapting control channel monitoring for SCell

Interestingly, in the DSR and DAA techniques, LL is increased from other methods (e.g., 4%). This is because, in the case of DSR, even when scheduling in advance considering traffic pattern of VoIP, this benefit disappears when packet generation is dynamically determined from AMR operation [31]. In the case of DAA, reducing the number of antennas decreases the channel capacity, so it can take more time to transmit the subsequent packets. For DIC, it is observed that almost no penalty in latency.

Since power saving techniques are not mutually exclusive, the network can mix and use various power saving techniques together according to the traffic situation of the terminal.

V. ADVANCED POWER SAVING TECHNIQUES IN FURTHER EVOLUTION OF 5G

The power saving techniques supported in current NR specification are briefly summarized in Table 6. The various power saving technologies introduced so far have focused on the power consumption reduction of the terminal in the connected mode where the terminal is connected to a network and receives data services from the network. However, the users often do not use their cell phones for a long time, but they still see that the battery is exhausted. This problem occurs when the terminal is in the idle mode. When there is no more data to send or receive, the terminal's connection can be released and it falls into the idle state. From the user's point of view, it would be natural that a battery is consumed by using a mobile phone, but it could be strange that the battery is still consumed quite a lot even though it is not used. In order to further reduce power consumption even for the idle state, further advanced power saving techniques are presented in this section. In addition, a more powerful power saving method is introduced for the connected mode and the various vertical services for the future release of NR.

A. ADVANCED POWER SAVING TECHNIQUE FOR IDLE MODE

In the idle mode, the terminal periodically monitors the control channel to check for the presence of paging messages from the network (of course, much less than the active mode) [35]. When a control information for paging message is detected, the terminal continues decoding the subsequent

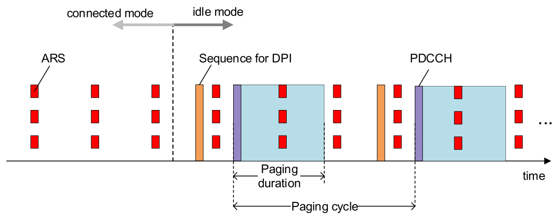


FIGURE 7. **Illustration of advanced power saving techniques for idle mode; dynamic paging indication.**

data channel to determine whether the paging message is actually targeting itself or not. Since the terminal has not been connected to the network for a long time, the user naturally needs an additional operation to acquire synchronization before receiving the paging message. Therefore, the paging process consumes power by repeating the following three processes: synchronization acquisition, control channel monitoring and data channel receptions for the paging message.

At first glance, the paging process doesn't seem to have any problems, but two points need to be revisited. One point is that one paging message is delivered for a large number of users rather than one, and the other is that maintaining synchronization is essential to receiving paging messages. Basically, the paging mechanism divides users into groups and delivers paging messages on a group basis to reduce the overhead of paging messages [36]. That is, any users in a group monitor the same paging occasion(s) assigned to that group. For example, if the network sends a paging message to one user in the group, all users in the group try to decode the same paging message at the same time. Therefore, the rest of the users continue to consume power even though the paging message does not intend for those users.

In addition, the terminal should maintain time and frequency synchronization at the modem to receive the paging message, and thus keep monitoring the synchronization signal. The period of the synchronization signal is much shorter than the time the user wakes up in idle mode (e.g., typically 20 ms period for synchronization signal [37]) and may not aligned with the period of paging cycle. Consequently, the terminal needs to wake-up more frequently to monitor the synchronization signal during the paging procedure. To solve these two problems, a dynamic paging indication (DPI) technique is proposed. The DPI shares the similar design fundamentals to DIC.

Figure 7 shows the DPI operating principle for power saving in idle mode. To reduce unnecessary decoding attempts during the paging procedure, the network informs the terminal whether or not to attempt paging monitoring in the next paging cycle. For DPI purpose, a sequence can be transmitted that does not require a complicated decoding process of the terminal, unlike the DIC, to minimize power consumption to receive this additional signal. For this sequence-based DPI, the terminal can determine the presence or absence of the sequence transmission through a simple correlation operation, so that power consumption is very low. In addition, since time and frequency synchronization can be obtained from the sequence, there is an advantage of not having to use a conventional synchronization signal.

B. ADVANCED POWER SAVING TECHNIQUE FOR CONNECTED MODE

Among the results of various techniques tested in Section IV.F, DIC shows a relatively high LL compared to other power saving schemes. This means that if one DRX cycle is skipped by the DIC, the terminal should wait until the next DRX cycle even if data occurs before the next DRX cycle. This behavior is not suitable for use in delay-sensitive services such as video streaming and gaming. To overcome this drawback of DIC, enhanced DIC (E-DIC) which corresponds to a method of applying DIC with finer time granularity is proposed. For E-DIC, control information can be transmitted within the DRX active time and dynamically indicates how to monitor the subsequent control channel monitoring occasions during the DRX active time. Two types of E-DIC signaling can be considered which are E-DIC for adaptation (E-DIC-A) and E-DIC for skipping (E-DIC-S) as depicted in Figure 10. E-DIC-A dynamically changes the control channel configurations such as monitoring periodicity and E-DIC-S indicates to skip the a number of control channel monitoring occasions within a certain time period.

For E-DIC-A, two control channel monitoring periodicities p_1 and p_2 where p_1 corresponds to a default periodicity and an inactivity timer t_{inact} are considered. The terminal is dynamically indicated to switch the monitoring periodicity of control channel from p_1 to p_2 when there is frequent traffics. If there is no traffic so that the inactivity timer t_{inact} is expired, the terminal switches the monitoring periodicity from p_2 to p_1 to save power consumption.

For E-DIC-S, there are a default monitoring periodicity p_1, sleep duration t_{sleep} and a inactivity timer t_{inact}. The terminal keeps monitoring the control channel with periodicity p_1 and the inactivity timer t_{inact} is running. If the terminal receives scheduling information, the inactivity timer t_{inact} is initialized. If either the inactivity timer t_{inact} is expired or the terminal dynamically indicated to stop to monitor the control channel, the terminal goes to sleep for t_{sleep} duration.

C. PERFORMANCE EVALUATION

In this sub-section, we provide evaluation results for the advanced power saving techniques introduced in Section V-A and V-B.

To evaluate the benefits of using the DPI, we take the conventional paging process relying on the synchronization signal as the baseline scheme. In addition, for the comparison, it is further considered to use an additional reference signal (ARS) where the occasions of the ARS are aligned with the paging cycle. By using ARS, the number of wake-ups for the synchronization process of the terminal can be minimized. Therefore, it is expected that ARS gives power saving gain over the baseline scheme using a synchronization signal.

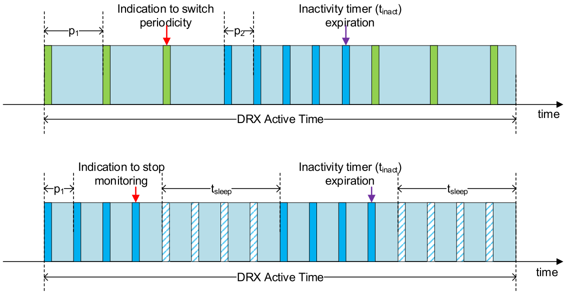


FIGURE 8. Illustration of advanced power saving techniques for connected mode; E-DIC-A (upper) and E-DIC-S (lower).

The power consumption model for the paging process in idle mode (P_p) can be modeled as below:

$$P_p = (1 - \gamma_g(M)) \times P_c + \gamma_g(M) \times P_d, \qquad (5)$$

where P_c and P_d is relative power consumptions for receiving control channel and data channel, respectively as shown in Table 2. $\gamma_g(M)$ is the group paging rate defined by

$$\gamma_g(M) = 1 - (1 - \gamma_{ue})^M, \qquad (6)$$

where γ_{ue} is the paging rate for a terminal and M is the number of terminals in a paging group [38]. We consider that the relative power consumption for sequence detection and subsequent synchronization process is 60. It is assumed that the periodicity of synchronization signal, paging duration, paging rate γ_{ue} are 20 ms, 4 ms and 1%, respectively. Sequence for DPI is transmitted 2 ms before the start of paging cycle. For the baseline scheme, it is assumed that N_{ss} synchronization signals are monitored by the terminal before receiving paging message.

Figure 8 shows the PSG performance of DPI and ARS according to N_{ss} and M. It is observed that the PSG increases as N_{ss} increases for both DPI and ARS. When $N_{ss} = 3$ with $M = 10$, the DPI and ARS achieved 29% and 41% of PSGs over the baseline scheme. Through this result, it can be seen that the DPI and ARS methods provide a great benefit especially for terminals with poor signal quality (e.g., the terminals in cell boundary), which typically require more synchronization signal observations. It can be seen that DPI surpasses ARS since the terminal can avoid unnecessary paging decoding by using DPI. For DPI, the PSG decreases as M increases. This is because the PSG of DPI is affected by the group paging rate, and it is desirable to reduce the number of users per paging group as much as possible.

Figure 9 describes the PSG results of DPI and ARS according to the length of the paging cycle are observed. The PSGs of the DPI and ARS decrease as the paging cycle increases. Under the extremely long paging cycle such as 10.25 sec, the PSG of DPI is limited to about 10%. Therefore, the network can determine how to utilize the power saving techniques for idle mode according to the paging configuration that would be depending on the various network situations.

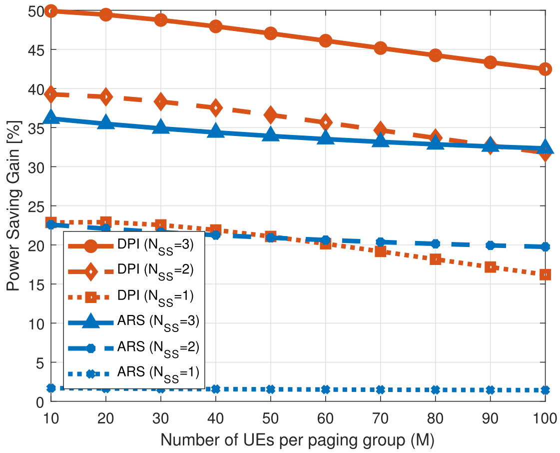


FIGURE 9. Power saving gain for DPI and ARS according to the number of terminals per paging group where N_{ss} = 1, 2, and 3.

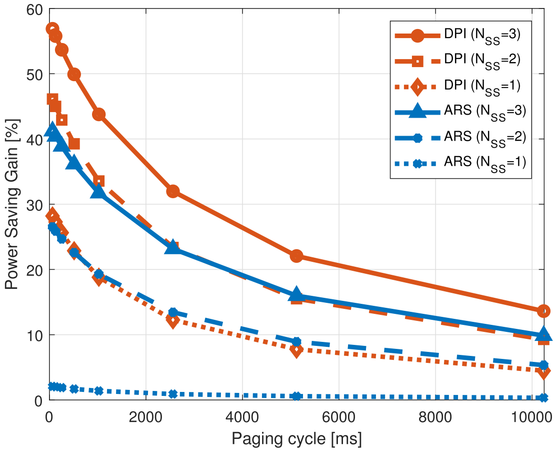


FIGURE 10. Power saving gain for DPI and ARS according to the paging cycle where N_{ss} = 1, 2, and 3.

To evaluate the benefit of using E-DIC, FTP model 3 with 1 MB packet size with various inter-arrival time is evaluated. For DRX configuration, we set 40 ms DRX cycle, 100 ms inactivity timer, and 10 ms on-duration timer. For E-DIC-A, $p_1 = 2$ ms, $p_2 = 0.5$ ms, and $t_{\text{inact}} = 2$ ms are assumed and for E-DIC-S, $p_1 = 0.5$ ms, $t_{\text{sleep}} = 2$ ms, and $t_{\text{inact}} = 2$ ms are assumed. DRX without any additional power saving technique is considered as the baseline scheme.

The PSGs and LLs of DIC, E-DIC-A, and E-DIC-S over the baseline scheme according to inter-arrival time of data traffic are described in Figure 11. It is observed that both E-DIC-A and E-DIC-S outperform DIC in terms of PSG. This implies that although DRX and DRX with DIC are used, there are still a lot of power consumptions at the terminal due to unnecessary control channel monitoring within DRX active time. The proposed E-DIC-A and E-DIC-S can minimize such unnecessary decoding attempts for the control channel of the terminal as much as possible. The E-DIC-A and E-DIC-S show smaller LLs compared to the DIC. This is because of that E-DIC-A and E-DIC-S adapt the terminal's control channel monitoring behavior with much finer time granularity that corresponds to slot-level adaptation. E-DIC-A is better than E-DIC-S from PSG perspective and vice versa from LL perspective. Considering the trade-off,

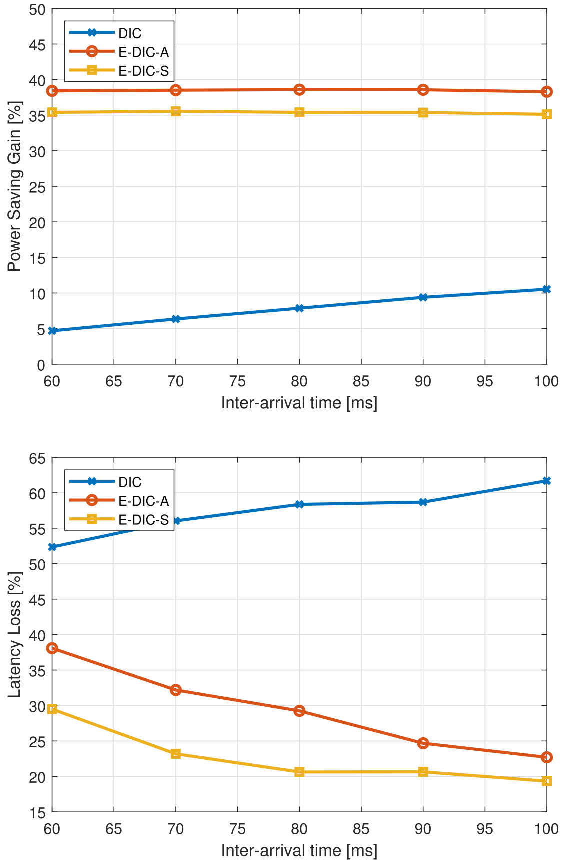


FIGURE 11. **Evaluation results for advanced power saving techniques for connected mode; power saving gain (upper) and latency loss (lower).**

the network can decide to use a more suitable power saving technique between E-DIC-A and E-DIC-S.

D. POWER SAVING TECHNIQUES FOR VARIOUS VERTICAL SERVICES

In this subsection, we briefly introduce power saving techniques that can be applied for 5G special functions such as direct terminal communication as known as sidelink, unlicensed band communication, and services for low-complexity (reduced complexity) terminals.

- *Sidelink:* Sidelink is for inter-vehicle communication, vehicle-to-user communication, and the ability to transfer data between terminals to ensure coverage in disaster situations [39]. It is a very useful function in a variety of situations, such as lifesaving and accident situation delivery, but it acts as a relay for data transmission and checks whether it is available before data transmission, causing high power consumption. To prevent this, partial resource sensing or random resource selection scheme could be the candidates of solutions to minimize the sensing function of the terminal before data transmission.
- *Unlicensed spectrum:* For unlicensed access, since transmissions and receptions typically include an additional energy sensing operation called listen-before-talk (LBT), higher power consumption is expected compared to operation in the licensed spectrum [40]. Moreover, the situation is even worse since the terminal needs to do LBT frequently to minimize access latency. To reduce the power consumption in unlicensed access, dynamic adaptation of control channel monitoring according to LBT success or failure is specified in Rel-16 NR [40]. It can be further enhanced in future release by considering partial LBT which enables that the terminal monitors only a part of beam directions for LBT operation to minimize the sensing attempts.
- *Low-complexity terminal:* the low-complexity terminal is for low-performance terminals such as smart watches, smart glasses, sensors, surveillance cameras, etc [41]. Since the low-complexity terminal does not have high performance in many parts, it is desirable to take into account much lower requirements for processing time, control channel monitoring budgets, number of active antennas, etc. In addition, the services for low-complexity terminal may have different traffic characteristics compared to normal eMBB such as a materially longer idle period and smaller maximum data packet size. The power saving techniques should be developed by considering these aspects. For example, processing time relaxation, extremely long DRX operation, and control-less data transmissions can be the candidates of potential solutions for the low-complexity terminals.

VI. CONCLUSION

The power saving technology is becoming more and more important as customers expect higher and higher performance on their terminals despite the disproportionate to improve battery capacity. Developing terminal power saving techniques applicable even for high performance use cases are essential for the success of 5G. In this paper, we provided a comprehensive overview of the on-going research direction for terminal power saving techniques in 5G NR standardization. First step taken in 5G NR was to specify DRX operation which reduces terminal power consumption by managing the terminal's time domain behavior. DRX operation was further supplemented with other techniques that addressed the power consumption issue in the frequency domain and antenna domain. To further improve power savings, techniques to effectively adjust various sleep states through more flexible and faster control were introduced. The design principles and benefits of the power saving techniques were comparatively analyzed with various simulation results considering key system models for evaluating terminal's power consumption. We further proposed some of potential candidates of power saving techniques that will be handled in future releases of 5G NR. Such power saving techniques are going to be considered for conventional data services as well as various vertical services.

REFERENCES

[1] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A survey on 5G usage scenarios and traffic models," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 905–929, 2nd Quart., 2020.

[2] M. Lauridsen, "Studies on mobile terminal energy consumption for LTE and future 5G," Ph.D. dissertation, Dept. Electron. Syst., Aalborg Univ., Aalborg, Denmark, Jan. 2015.

[3] M. Lauridsen, P. Mogensen, and T. B. Sorensen, "Estimation of a 10 Gb/s 5G Receiver's performance and power evolution towards 2030," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC-Fall)*, Boston, MA, USA, Sep. 2015, pp. 1–5.

[4] *Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s)*, document M.2410-0, ITU-R, Nov. 2017.

[5] Y. Kim, F. Sun, Y. Wang, Y. Qi, J. Lee, Y. Kim, J. Oh, H. Ji, J. Yeo, S. Choi, H. Ryu, H. Noh, and T. Kim, "New radio (NR) and its evolution toward 5G-advanced," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 2–7, Jun. 2019.

[6] J. Yeo, T. Kim, J. Oh, S. Park, Y. Kim, and J. Lee, "Advanced data transmission framework for 5G wireless communications in the 3GPP new radio standard," *IEEE Commun. Standards Mag.*, vol. 3, no. 3, pp. 38–43, Sep. 2019.

[7] *Study on User Equipment (UE) Power Saving in NR (Release 16), V16.0.0*, document TR 38.840, 3GPP, Jun. 2019.

[8] M. Ismail, W. Zhuang, E. Serpedin, and K. Qaraqe, "A survey on green mobile networking: From the perspectives of network operators and mobile users," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1535–1556, 3rd Quart., 2014.

[9] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, "Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 33–56, 1st Quart., 2017.

[10] S.-R. Yang and Y.-B. Lin, "Modeling UMTS discontinuous reception mechanism," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 312–319, Jan. 2005.

[11] A. T. Koc, S. C. Jha, R. Vannithamby, and M. Torlak, "Device power saving and latency optimization in LTE–A networks through DRX configuration," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2614–2625, May 2014.

[12] H. Ramazanali and A. Vinel, "Performance evaluation of LTE/LTE–A DRX: A Markovian approach," *IEEE Internet Things J.*, vol. 3, no. 3, pp. 386–397, Jun. 2016.

[13] M. K. Maheshwari, M. Agiwal, N. Saxena, and A. Roy, "Hybrid directional discontinuous reception (HD-DRX) for 5G communication," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1421–1424, Jun. 2017.

[14] M. K. Maheshwari, M. Agiwal, N. Saxena, and A. Roy, "Directional discontinuous reception (DDRX) for mmWave enabled 5G communications," *IEEE Trans. Mobile Comput.*, vol. 18, no. 10, pp. 2330–2343, Oct. 2019.

[15] F. Moradi, E. Fitzgerald, M. Pioro, and B. Landfeldt, "Flexible DRX optimization for LTE and 5G," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 607–621, Jan. 2020.

[16] I. Demirkol, C. Ersoy, and E. Onur, "Wake-up receivers for wireless sensor networks: Benefits and challenges," *IEEE Wireless Commun.*, vol. 16, no. 4, pp. 88–96, Aug. 2009.

[17] N. S. Mazloum and O. Edfors, "Performance analysis and energy optimization of wake-up receiver schemes for wireless low-power applications," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7050–7061, Dec. 2014.

[18] N. Kouzayha, Z. Dawy, J. G. Andrews, and H. ElSawy, "Joint Downlink/Uplink RF wake-up solution for IoT over cellular networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1574–1588, Mar. 2018.

[19] S. Rostami, S. Lagen, M. Costa, M. Valkama, and P. Dini, "Wake-up radio based access in 5G under delay constraints: Modeling and optimization," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1044–1057, Feb. 2020.

[20] S. Rostami, K. Heiska, O. Puchko, K. Leppanen, and M. Valkama, "Pre-grant signaling for energy-efficient 5G and beyond mobile devices: Method and analysis," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 2, pp. 418–432, Jun. 2019.

[21] F. Liu, K. Zheng, W. Xiang, and H. Zhao, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 197–207, Feb. 2014.

[22] G. Yu, Q. Chen, R. Yin, H. Zhang, and G. Y. Li, "Joint downlink and uplink resource allocation for energy-efficient carrier aggregation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3207–3218, Jun. 2015.

[23] K. Pedersen, F. Frederiksen, C. Rosa, H. Nguyen, L. G. Garcia, and Y. Wang, "Carrier aggregation for LTE-advanced: Functionality and performance aspects," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 89–95, Jun. 2011.

[24] H. Lee, S. Vahid, and K. Moessner, "A survey of radio resource management for spectrum aggregation in LTE-advanced," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 745–760, 2nd Quart., 2014.

[25] *LTE Radio Access Networks (RAN) Enhancements for Diverse Data Applications, V11.0.0*, document TR 36.822, 3GPP, Sep. 2012.

[26] K. Takeda, H. Xu, T. Kim, K. Schober, and X. Lin, "Understanding the heart of the 5G air interface: An overview of physical downlink control channel for 5G new radio," *IEEE Commun. Standards Mag.*, vol. 4, no. 3, pp. 22–29, Sep. 2020.

[27] J. Jeon, "NR wide bandwidth operations," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 42–46, Mar. 2018.

[28] F. Abinader, "Impact of bandwidth part (BWP) switching on 5G NR system performance," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Dresden, Germany, 2019, pp. 161–166.

[29] *Radio Frequency (RF) and Co-Existence Aspects (Release 14), V14.2.0*, document TR 38.803, 3GPP, Sep. 2017.

[30] *Study on Licensed-Assisted Access to Unlicensed Spectrum, V13.0.0*, document TR 36.889, 3GPP, Jun. 2015.

[31] *LTE Physical Layer Framework for Performance Verification*, document R1-070674, 3GPP, Feb. 2018.

[32] *Medium Access Control (MAC) Protocol Specification (Release 16), V16.1.0*, document TS 38.321, 3GPP, Jul. 2020.

[33] *Radio Resource Control (RRC) Protocol Specification (Release 16), V16.1.0*, document TS 38.331, 3GPP, Jul. 2020.

[34] *Physical Layer Procedures for Data (Release 16)*, document TS 38.214, 3GPP, Jun. 2020.

[35] M. Bagaa, T. Taleb, and A. Ksentini, "Efficient Tracking Area Management Framework for 5G Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, p. 4117–4131, Jun. 2016.

[36] *User Equipment (UE) Procedures in Idle Mode and RRC Inactive State (Release 16)*, document TS 38.304, 3GPP, Jul. 2020.

[37] J. Liu, K. Au, A. Maaref, J. Luo, H. Baligh, H. Tong, A. Chassaigne, and J. Lorca, "Initial access, mobility, and user-centric multi-beam operation in 5G new radio," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 35–41, Mar. 2018.

[38] Q. Pan, X. Wen, Z. Lu, W. Jing, and L. Li, "Cluster-based group paging for massive machine type communications under 5G networks," *IEEE Access*, vol. 6, pp. 64891–64904, 2018.

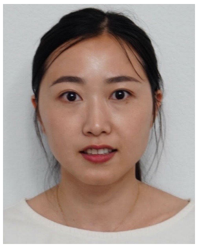
[39] S. A. Ashraf, R. Blasco, H. Do, G. Fodor, C. Zhang, and W. Sun, "Supporting Vehicle-to-Everything services by 5G new radio release-16 systems," *IEEE Commun. Standards Mag.*, vol. 4, no. 1, pp. 26–32, Mar. 2020.

[40] *Study on NR-based Access to Unlicensed Spectrum (Release 16), V16.0.0*, document TR 38.889, 3GPP, Dec. 2018.

[41] *New SID Support Reduced Capability NR Devices*, document RP-193238, 3GPP, Dec. 2019.

TAEHYOUNG KIM received the B.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2010 and 2016, respectively. He joined Samsung Electronics Company Ltd., in 2016, and has been working on the standardization of 5G wireless communication systems in 3GPP. His research interests include signal processing for communications, massive MIMO systems, and frame structure design.

YOUNSUN KIM (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Yonsei University, in 1996 and 1999, respectively, and the Ph.D. degree in electrical engineering from the University of Washington, in 2009. He joined Samsung Electronics Company Ltd., in 1999, and has since worked on the physical layer standardization of cdma2000, HRPD, LTE, and NR. He has been serving as the Vice-Chairman of the 3GPP RAN1 (physical layer) working group since 2017. He is currently a Master (Technical VP) with Samsung Electronics Company Ltd., and leads the research on physical layer standardization of wireless communications.



QIONGJIE LIN received the B.S. degree in communications engineering from the Nanjing University of Posts and Telecommunication, China, in 2012, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Tech, in 2015 and 2017, respectively. During her Ph.D. degree, she was working on physical layer system design and implementation of distributed arrays on software-defined radios for a variety of packet-based wireless systems. She has been joined Samsung Research America, Inc., as a Senior Standards Engineer, since 2018. She is currently working on standardizing innovative technologies for 5G New Radio (NR) system.

FEIFEI SUN received the B.S. degree in telecommunication engineering and the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunication, Beijing, China, in 2006 and 2011, respectively. She was a Visiting Student with the University of California, Santa Barbara, from 2009 to 2010. From 2011 to 2017, she worked as a Technical Manager with MediaTek (Beijing) Inc. Since 2011, she has been working with 4G/3G physical layer standardization in 3GPP TSG RAN Group 1. She is currently working as a Staff Engineer with the Samsung Research and Development Institute China-Beijing (SRC-B), Samsung Electronics Service Center, Beijing. She focused her research on machine-type communications and the Internet of Things for nine years. She was a co-rapporteur of 3GPP TSG RAN working Item on LTE for machine-type communication (MTC) in Release 15.

JINGXING FU received the B.S. degree in system safety engineering from Northeast University, Shenyang, China, in 1989, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, China, in 2004. Since 2004, he has been working as an Engineer with the Beijing Samsung Research and Development Center on 3GPP standard work.

YOUNGBUM KIM received the B.S. degree in electrical and electronic engineering from Chung-Ang University, in 2000, and the M.S. degree in electrical and electronic engineering from Yonsei University, South Korea, in 2002. He has been with Samsung Electronics Company Ltd., since 2002, and has worked on wide range of wireless communications. He has contributed to standardization activity in 3GPP for HSPA, LTE, LTE-Advance, LTE-Advance Pro, and NR. His current research interests include coexistence of multi-radio systems, communication coverage, and the cellular IoT communications.

ARIS PAPASAKELLARIOU received the Ph.D. degree in electrical and computer engineering from Rice University, in 1996. He is currently the Vice President of technology standardization with Samsung Electronics Company Ltd. For the last 15 years, he worked on 3GPP standardization and contributed to the specification of much fundamental functionality for LTE/4G and NR/5G. He received the 3GPP Excellence Award in 2017. He is also the Editor of TS 38.213 (NR; physical layer procedures for control).

HYOUNGJU JI (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Sogang University, and the Ph.D. degree in electrical engineering from Seoul National University. He joined Samsung Electronics Company Ltd., in 2007, where he has been involved in 3GPP RAN1 LTE, LTE-Advanced, LTE-Advanced Pro, and 5G NR technology developments and standardizations. He is currently a Principal Engineer with Samsung Electronics Company Ltd. His current research interests include multi-antenna techniques, FD-MIMO, massive connectivity, machine type communications, and IoT communications.

JUHO LEE (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 2000. He is currently a Fellow (Technical SVP) with Samsung Electronics Company Ltd., where he is leading research and standardization for mobile communications. He joined Samsung Electronics Company Ltd., in 2000, and has worked on 3G, 4G, and 5G technologies and is leading research for future technologies, such as beyond 5G and 6G. He was a Vice Chairman of 3GPP RAN WG1 from February 2003 to August 2009 and chaired LTE/LTE-Advanced MIMO sessions.

• • •