On Device Grouping for Efficient Multicast Communications in Narrowband-IoT

Galini Tsoukaneri
The University of Edinburgh
Email: G.Tsoukaneri@sms.ed.ac.uk

Mahesh K. Marina
The University of Edinburgh
Email: mahesh@ed.ac.uk

Abstract—Narrowband IoT (NB-IoT) is a new cellular network technology that has been designed for low capability, low power consumption devices that are expected to operate for more than 10 years on a single battery. These types of devices will be inexpensive (less than $5) and deployed on massive scales. This long life expectancy will lead to the need for occasional software updates, to very large groups of devices. While a new multicast mechanism has recently been proposed for the efficient multicast transmission of such updates, it assumes that devices can be grouped together and synchronized in order to receive the multicast data. In this paper, we explore three different approaches to achieve device grouping, with different trade-offs between bandwidth usage, energy consumption and compliance with the NB-IoT standard. To assess the performance of each approach, we conducted a thorough experimental evaluation under realistic operating conditions.

Index Terms—NB-IoT, Group Communications, Device grouping, DRX synchronization

I. INTRODUCTION

NarrowBand Internet-of-Things (NB-IoT) [1] is a new cellular network technology for 5G, that focuses on large numbers of low-cost, low-capability devices with minimal energy consumption (sensors, meters, etc.) These devices are expected to operate for long periods (more than 10 years), so it important to be able to distribute firmware updates efficiently to keep them secure and up-to-date. While such firmware updates are typically handled through group communications (multicast), the Single Cell-Point to Multipoint (SC-PTM) scheme [2] used in 5G to enable such communications was not designed with the resource constraints of NB-IoT in mind and is thus quite inefficient. Recently, [3] proposed an on-demand scheme as an alternative to SC-PTM, that is more efficient both in terms of bandwidth usage and device energy consumption. In that scheme, whenever there is multicast content to be transmitted, the evolved NodeB (eNB) notifies the devices of the upcoming multicast transmission. That way, the devices do not need to periodically wake up and check for potential multicast transmissions, as in SC-PTM, and thus their scheme reduces both the energy consumption of the devices as well as the signaling overhead of the eNB.

However, while [3] describes an efficient mechanism for informing the devices about upcoming multicast transmissions, it does not address how these devices could be synchronized so that they wake up simultaneously and receive the data with a single transmission. This synchronization is essential in the context of NB-IoT, as the available bandwidth is extremely limited, and multiple transmissions for the same multicast data would severely degrade the system’s performance. Therefore, it is crucial to have an efficient mechanism that can group devices with different Discontinuous Reception (DRX) cycles and synchronize them, to minimize the number of multicast transmissions (optimally to a single transmission).

In this paper we explore three different mechanisms to achieve such grouping and synchronization, and we experimentally evaluate their performance under realistic operating conditions. Each mechanism makes different trade-offs between three important aspects: the device energy consumption, bandwidth usage and compliance with the NB-IoT standards. Our first mechanism respects the DRX cycle of the devices completely and aims to transmit the multicast data with the fewest transmissions required to cover all devices. This approach is standards compliant and has the lowest energy consumption, at the expense of increased bandwidth due to multiple transmissions. The second mechanism modifies the DRX cycles of the devices in order to synchronize them at the time of multicast transmission so that only a single multicast transmission is needed. This approach minimizes the bandwidth usage and is also standards compliant, but increases the energy consumption of the devices. The third mechanism uses a small modification to the paging protocol to notify the devices well in advance of the time of the multicast transmission, so that the devices can wake up to receive it without the need for further signaling. This minimizes both energy consumption and bandwidth usage, but is not compliant with the NB-IoT standard.

The rest of the paper is organized as follows. In Sec. II we summarize the background of our work. In Sec. III we present our three different grouping mechanisms. In Sec. IV we present our experimental procedure and our results. In Sec. V we discuss related work and we conclude in Sec. VI.

II. BACKGROUND

A. Group Communications in NB-IoT

Recently Single Cell - Point to Multipoint (SC-PTM) [2] was standardized as the protocol to be used for multicast transmissions in NB-IoT. SC-PTM is based on the Multimedia Broadcast Multicast Service (MBMS) standard [4] which was first introduced in release 6, and is used to facilitate multicast transmissions in cellular networks through a number of
different procedures (i.e. subscription, service announcement, joining, notification, session start/stop). In MBMS the eNB announces the available and upcoming services and each device subscribes to the services it wants to receive. SC-PTM differs from MBMS on the fact that MBMS can be applied on multiple cells at the same time by synchronizing the eNBs, while SC-PTM can only be applied on a single cell.

While this type of multicast mechanism has been successfully applied in LTE/LTE-A networks, the limited resources available in NB-IoT make this subscription-based approach inefficient in terms of bandwidth usage and device energy consumption, due to its large signaling overhead requirements [3].

More recently, [3] proposed a novel scheme for multicast transmission in NB-IoT, based on the idea of on-demand paging of the devices whenever a multicast transmission is imminent. In contrast to SC-PTM, the entity that provides the multicast data (e.g. device manufacturer, telecommunication company) now also supplies the network with the list of the devices that need to receive it. The mobile network operator then distributes both the list and the data to all the eNBs that the devices are attached to. Finally the eNBs page the devices to connect to the network and receive the transmission.

As there are no subscriptions that need to be managed, the MBMS model can be greatly simplified, to reduce its complexity and signaling overhead. More specifically, the subscription and service announcement stages [4] have been removed, as the co-ordination entity is the one that decides which devices need to receive the data. Additionally, the notification stage has also been removed as the devices are now paged individually. Finally, the joining procedure is performed at the network side to set up a generic multicast bearer based on the capabilities of the devices that will use it.

B. DRX/eDRX

Although the aforementioned scheme is simple and efficient in terms of resource utilization and energy consumption, it makes the assumption that the paged devices are synchronized and a single multicast transmission is needed. However, in practice this is not the case, as devices may have different DRX/eDRX cycles and different starting points to their cycles.

DRX [5] is an important feature in cellular networks as it allows a device that is not sending or receiving data to turn off its reception (RF) and transmission (TX) modules and enter a sleep mode, with minimal energy consumption. During the active period of the DRX cycle, the device switches its RF module on and checks its paging occasion (PO) [6] on the paging channel to determine if there is a message for it. If yes, the device connects to the network to receive the downlink data, otherwise it falls to sleep. During which time its modules are turned off to reduce the energy consumption. After data reception, the device starts an inactivity timer (usually 10 – 30 sec. in commercial networks), waiting for the arrival of additional data. Upon expiration of the inactivity timer, the device goes back to sleep and starts a new DRX cycle (Fig. 1).

The length of a DRX cycle is usually negotiated between the eNB and the device at connection time. However, the eNB can unilaterally decide on the DRX cycle, which is something that can be used to forcibly synchronize the devices (sec. III-B). In LTE/LTE-A, the DRX cycle ranges from 0.32 to 2.56 seconds [2], while in NB-IoT, extended DRX (eDRX) cycles may also be used, that span from 20.48 seconds to 175 minutes [7], to prolong the battery life even more. Furthermore, DRX values are always twice as long as the immediately shorter DRX value (e.g. 20.48 sec, then 40.96 sec, then 81.92 sec and so forth until 10485.76 sec). The shorter the DRX cycle the more often the device will go into the active period, resulting in increased uptime [8]. This is an important consideration for NB-IoT devices, for which the battery life is expected to be more than 10 years.

III. GROUPING MECHANISMS

If multiple devices could be grouped together and synchronized so that they have a PO within the inactivity timer ($T_i$) before a multicast transmission, they will all be able to receive the multicast data with a single transmission. This is because none of the device will go back to sleep before the others have been paged, and they will all be active at multicast transmission time. Otherwise, the inactivity timer of some devices will expire before the other devices can be paged. An example is depicted in Fig. 2. In Fig. 2(a), devices 2 and 3 are within the range of the $T_i$ from device 1 so one multicast transmission will cover all of them. However, in Fig. 2(b), only...
device 2 is within range from device 1 so a second multicast transmission is required for device 3.

Here, we explore three grouping mechanisms to transmit multicast data in NB-IoT. Each mechanism focuses on different but equally important aspects: device energy consumption, bandwidth usage and compliance with the NB-IoT standard.

A. DRX Respecting, Standards Compliant (DR-SC)

In this mechanism, the DRX cycle of the devices is respected and devices share a multicast transmission only if their POs happen to be closer in time than $T_I$. As such, the devices do not use any more energy than what they normally would have under normal operation, aside from the multicast data transmission itself. However, as there is no guarantee that multiple devices will coincide at a multicast transmission, numerous transmissions will most likely be needed to cover all devices (Sec. IV), leading to higher bandwidth usage.

In this scenario, we would ideally like to find the minimum number of transmissions needed to cover all devices, so that the bandwidth usage is minimized. We can formulate the POs in time as a bipartite graph of devices and frames where each edge indicates that the device has a PO on that frame (Fig. 3). Finding the minimum set of frames that would cover all devices corresponds to the set cover problem, which is known to be NP-hard [9]. Here, we follow the approximate solution would be frames 4 and 5.

Here we formulate the POs in time as a bipartite graph. Each edge from a device to a frame indicates that the device has a PO on that frame. For simplicity, and without loss of generality, we consider the inactivity timer $T_I$ to be a single frame here. Finding the minimum set of frames that would cover all devices corresponds to the set cover problem, which is known to be NP-hard [9]. Here, the optimal solution would be frames 4 and 5.

Fig. 3. Set cover problem: Here we formulate the POs in time as a bipartite graph. Each edge from a device to a frame indicates that the device has a PO on that frame. For simplicity, and without loss of generality, we consider the inactivity timer $T_I$ to be a single frame here. Finding the minimum set of frames that would cover all devices corresponds to the set cover problem, which is known to be NP-hard [9]. Here, the optimal solution would be frames 4 and 5.

In more detail, the eNB chooses a time $t$ to transmit the multicast data. The time $t$ should be at least $2 \times \text{max}_{DRX}$ where $\text{max}_{DRX}$ is the longest DRX cycle of the devices to receive the multicast data, so that there will be at least one PO of every device before $t$. The eNB then finds each device that does not have a future PO within $[t - T_I, t)$, and decreases its DRX cycle to the maximum that creates a PO within that time period. Since the shortest DRX cycle (2.5s) is typically much shorter than $T_I$, such a DRX cycle is guaranteed to exist. To keep the energy consumption introduced by the adapted DRX cycle as low as possible, the adaptation happens in the last PO before $t - T_I$ (Fig. 5).

To enforce the new DRX cycle, the eNB pages the device, which then proceeds to connect to the network through the typical Random Access process [11], and receives the new DRX value in the RRC Connection Reconfiguration message. The eNB then instructs the device to switch back to sleep immediately (without waiting the inactivity timer to expire), using the RRC Connection Release procedure, to reduce the device uptime and resulting energy consumption. After the multicast transmission, the original DRX cycles are restored with an additional RRC Connection Reconfiguration message.

B. DRX Adjusting, Standard Compliant (DA-SC)

The second mechanism seeks to proactively change the DRX cycle of some devices (for a limited time), so that all devices will have a PO within $T_I$, and receive the multicast transmission simultaneously. This has the benefit of minimizing the number of multicast transmissions needed to just a single one, making good use of the limited bandwidth available in NB-IoT. However, devices will need to use more energy, as they will be using a smaller DRX cycle. Since a DRX value is exactly twice as long as the previous one (Sec. II), increasing the DRX cycle respects the original periodicity of the device.

In more detail, the eNB chooses a time $t$ to transmit the multicast data. The time $t$ should be at least $2 \times \text{max}_{DRX}$ where $\text{max}_{DRX}$ is the longest DRX cycle of the devices to receive the multicast data, so that there will be at least one PO of every device before $t$. The eNB then finds each device that does not have a future PO within $[t - T_I, t)$, and decreases its DRX cycle to the maximum that creates a PO within that time period. Since the shortest DRX cycle (2.5s) is typically much shorter than $T_I$, such a DRX cycle is guaranteed to exist. To keep the energy consumption introduced by the adapted DRX cycle as low as possible, the adaptation happens in the last PO before $t - T_I$ (Fig. 5).

To enforce the new DRX cycle, the eNB pages the device, which then proceeds to connect to the network through the typical Random Access process [11], and receives the new DRX value in the RRC Connection Reconfiguration message. The eNB then instructs the device to switch back to sleep immediately (without waiting the inactivity timer to expire), using the RRC Connection Release procedure, to reduce the device uptime and resulting energy consumption. After the multicast transmission, the original DRX cycles are restored with an additional RRC Connection Reconfiguration message.

C. DRX Respecting, Standard Incompliant (DR-SI)

The third mechanism uses a new, non-critical extension (named mltc_transmission) to the existing paging message in order to notify the devices in advance about an imminent multicast transmission. This allows the devices to retain their preferred DRX cycles as in DR-SC, maintaining the normal energy usage, with a single multicast transmission, as in DA-
The figure depicts the adaptation of DRX cycles. Subfigure (a) shows the devices with their original DRX cycles. Device c does not need adapting as it has a PO within \([t - T_1, t]\). In steps (b) and (c) the DRX cycles of the devices a and b are adapted, as they don't have any POs in \([t - T_1, t]\). Please note that the adaptation happens on the last PO before \(t - T_1\) so that the energy consumption introduced is kept to a minimum.

SC. However, although the extension to the paging message is simple, this solution is no longer standards compliant.

In more detail, whenever the eNB has multicast data to transmit it sends an extended paging message to the devices that do not have a PO within \([t - T_1, t]\). The paging message contains the new non-critical extension which comprises of the device identity and the time remaining until the multicast transmission. The device ID is only present in the non-critical extension and not in the PagingRecordList field of the paging message, so devices can distinguish between a paging to receive downlink data and multicast transmissions. As the device is not paged to receive downlink data, it does not need to wake up and connect to the network, keeping the energy consumption similar to that in normal operation.

Upon receiving the paging message, the device selects a random time value between \([t - T_1, t]\) and sets a new timer (T322) to expire at the selected time. When T322 expires, the device wakes up and connects to the network to receive the multicast data. Finally, to indicate that the connection is made for the multicast transmission and not for unicast downlink data, the device sets the EstablishementCause field of the RRC Connection Request message to the new value of multicastReception.

**IV. Evaluation**

**A. Experimental Setup**

To assess the impact of each grouping approach in terms of energy consumption and bandwidth usage, we conducted a thorough experimental evaluation, considering a single eNB scenario serving a large number of NB-IoT devices. The effect on the bandwidth usage is dependent on the number of required transmissions to cover all devices for DR-SC. While the DA-SC and DR-SI approaches only need a single transmission, the DR-SC approach requires a variable number transmissions (Sec. III-A) depending on how many devices happen to be synchronized. Therefore we use the number of multicast transmissions as a proxy for the bandwidth utilization. As the probability of devices being synchronized increases as the number of devices increases, we evaluated a varying number of devices (100 to 1000), and averaged the results over 100 runs.

Specific energy consumption values are hard to estimate, as they are device specific and may change as technology evolves. However, increased uptime will lead to increased energy consumption irrespectively of the type of the device. Therefore, as a proxy for the energy consumption, we measure the relative increase of uptime compared to what would be required for unicast transmission (i.e. each device receiving the multicast data based on its own DRX and without waiting for other devices). Since unicast transmission would not introduce any additional processes, it is the most efficient way to receive the data in terms of energy consumption from the device perspective. Furthermore, we consider the uptime spent in light sleep mode (during the PO) and the active mode (during connection) separately as the energy usage in the latter is significantly higher [12], [13].

Following [3], we simulate a single cell with realistic NB-IoT traffic patterns based on [14]. Furthermore, we show results for multicast transmission data of three different sizes (100KB, 1MB and 10MB), which we believe covers the spectrum of typical firmware updates. All experiments were implemented on a custom simulator in Matlab.

**B. Results**

**Device Uptime:** First, we assessed the increase in uptime that each approach incurs to a device. This increase can be due to additional paging (DA-SC), DRX adjustment(DA-SC) or setting additional timers (DR-SI).

Fig. 6 shows the relative increase of uptime compared to unicast transmission (Sec. IV-A). In particular, Fig. 6(a) shows the uptime in light sleep mode, which is spent on POs and receiving paging messages. As we can see, the DR-SC approach requires exactly the same uptime as the unicast approach, as no extra POs are needed. The DA-SC induces a minor increase as additional POs are used with the adapted DRX, while the DR-SI introduces a negligible increase as only the reception of the paging message is required. Therefore, the energy consumption in the DR-SI is kept similar to that of unicast but with a single multicast transmission.

Fig. 6(b) shows the uptime in the connected mode, which is the mode that the device is in during the Random Access process, while waiting for the multicast transmission to begin and when receiving the multicast data. This is a more important metric than the light sleep uptime, as the energy consumed in this state is an order of magnitude greater [12]. In this case both DR-SC and DR-SI have slightly higher uptime than unicast transmission, as they will wait for \(\frac{T_1}{2}\) on average for the multicast transmission to start. DA-SC has the longest uptime, as it also needs to go through the Random Access process in order to connect to the eNB and get the DRX cycle adjusted. Overall however, compared to the actual time spent on receiving the multicast data, the relative increase in uptime is very low (Fig. 6(b)). In practice, the overhead introduced
Fig. 6. Relative uptime increase: The figure depicts the relative uptime increase compared to unicast transmission. Subfigure (a) depicts the uptime during light sleep mode, while subfigure (b) depicts the uptime in connected mode for different sizes of multicast data.

by the signaling of DA-SC becomes practically negligible as the multicast data size gets above 1MB.

**Number of multicast transmissions:** Here, we assess the number of multicast transmissions required to update all devices. While DA-SC and DR-SI mechanisms only require a single multicast transmission by design, DR-SC will typically require multiple transmissions as there is no guarantee that the devices will have synchronized POs. Fig. 7 shows the average number of multicast transmissions needed as the number of devices increases. Larger numbers of devices generally lead to a higher probability that multiple devices will be synchronized, so the number of required transmissions increases slower than the number of devices. However, the number of transmissions required is significant, and even for 1000 devices it is only 40% more bandwidth efficient than using unicast transmissions. Given that NB-IoT already operates on limited resources, this approach can seriously affect the existing traffic and result in significant performance degradation, ana thus, it is not deemed a practical grouping mechanism for multicast transmissions in NB-IoT.

V. RELATED WORK

A. Grouping in 4G/5G

Presently, device grouping for multicast transmissions (MBMS/SC-PTM), is done implicitly through the service an-

Fig. 7. Number of multicast transmissions: The figure depicts the average number of multicast transmissions required to update all devices over 100 runs. When few devices receive the multicast data the number of multicast transmissions is around 50% of the number of devices. As the number of devices increase, the number of multicast transmissions falls to around 40% of the number of devices.

ouncements. Essentially, each announced service corresponds to a specific group and each device becomes a part of the group by subscribing to it. In the scheme of [3], this is no longer possible as the eNB needs to fully handle the grouping and synchronization of the devices that need to receive the multicast data.

Past works have studied the problem of device grouping in the context of cellular networks, but have focused on entirely different settings such as grouping based on similar device QoS [15], or grouping based on device priority in the Random Access (RA) process [16], [17]. Moreover, most of them assume static groups and may require inter-group communication which are not applicable in our setting.

Works in [18], [19], [15] aim to optimize different parameters in the uplink communication. In [18] devices are grouped together based on their buffer size to allow for more efficient scheduling in the uplink. In the ACB and ECB [16], [17] schemes of 3GPP as well as [19] the goal is to minimize the RACH collisions and the devices are split into two groups based on the delay tolerance. Finally, in [15] devices are placed into groups based on their QoS requirements.

In [20], [21] devices are placed in groups based on their location and one of the devices is selected as a leader which coordinates the paging of the other devices due to its much shorter DRX. Similarly [22], [23] group devices according to their location and the head of the group is responsible for aggregating data and transmitting it to the network. However, electing a group leader requires short-distance communications (e.g. Bluetooth) which may not be implemented in these low-capabilities devices. Furthermore, it requires extra procedures that further increase the energy consumption.

B. DRX adaptation techniques

Several works analyze the effect of DRX cycle on the energy consumption and latency of the devices. [8], [24] present analysis on the average energy consumption and latency in LTE/LTE-A systems using different DRX cycles values.
Other works attempt to adapt the DRX cycles to achieve various results such as energy consumption minimization or preservation of QoS of received service. However, they do it for each device independently without the need to synchronize different devices. In [25] a DRX adaptation method is presented that aims in minimizing the energy consumption while guaranteeing the QoS of the multicast service being received. Similarly, [26], [27] attempt to fine-tune the DRX configuration to result in optimized energy consumption. Finally, [28], [29] present enhancements to the existing DRX mechanism that compromise the QoS in favor of the energy consumption.

VI. CONCLUSIONS

In this paper we explore three different mechanisms to achieve device grouping for multicast transmissions in NB-IoT: DR-SC, DA-SC and DR-SI. Each of the mechanisms makes different trade-offs between three important aspects; device energy consumption, bandwidth usage and standards compliance. To assess the performance of the three mechanisms in term of device energy usage and network bandwidth utilization, we conducted a thorough experimental evaluation under realistic traffic conditions. Our results show that the DR-SC mechanism results in very high bandwidth usage, and it is not much more efficient than delivering the data with unicast transmissions. Therefore it is not practical for NB-IoT deployments, where the available bandwidth is already limited. The DR-SI mechanism has excellent performance both in terms of energy consumption at the device side as well as bandwidth utilization at the network side. However, it requires protocol changes and may face deployment/adoption challenges. Finally, the DA-SC mechanism introduces slightly higher energy usage compared to DR-SI due to the adaptation of the devices’ DRX cycles, but this increase is very small compared to the actual time spent receiving the multicast data. Given the fact that it does not require any protocol changes, this mechanism offers the best trade-off among the three mechanisms for the target use case of distributing firmware updates.

REFERENCES

[12] 3GPP TSG-RAN WG2 Meeting #57bis, R2-071284, St. Julian’s, Malta, Agenda item:5.2.3, Source: Nokia, Title: DRX parameters in LTE.
[13] 3GPP TSG-RAN WG2 Meeting #57bis, R2-071285, St. Julian’s, Malta, Agenda item:5.2.3, Source: Nokia, Title: DRX parameters in LTE.