Formulating an Optimization Problem by Balancing Tradeoff between Latency and Power Saving in LTE - A Network

Budhaditya Bhattacharyya¹ and Inayat Singh²

¹Associate Professor, Department of Communication Engineering, VIT University, Tamilnadu, India.  
Orcid Id: 0000-0001-6507-763X  
²Student, Department of Communication Engineering, VIT University, Tamilnadu, India.

Abstract

The neoteric evolution of long term evolution (LTE) has helped us to acquire data rate up to 1Gbps but at the expense of Power Hunger amongst User Equipments (UE). Hence there is a need to optimize the usage of battery. Discontinuous reception (DRX) is a method that conserves the battery power but at the expense of increase in latency. This is the reason why distinctive optimizations are executed in the DRX model for different types of traffic modes but in this research an effort has been made to make the DRX mechanism slightly more adaptive by instigating a unique model for all the traffic modes. In this analysis on the basis of traffic intensity the UE in ON state will decide whether to receive the packets in the same states or switch to active state in case of high traffic intensity to minimize the latency. Further another parameter has been introduced i.e. channel strength; how it varies when the distance between UE and base station changes has been illustrated and keeping that into account appropriate modifications have been induced into the design to increase the Quality of Service (QoS) while saving the power.

Keywords: Discontinuous Reception, Active traffic, Background traffic, channel strength, MATLAB.

INTRODUCTION

With the growing demand of Smart Phones and extensive mobile applications, high data rate has almost become a necessity today. With the help of Carrier Aggregation [1], Advanced Coding Techniques and Advanced Multiple Antenna Schemes [2] we are able to achieve data rates as high as 1Gbps in Long Term Evolution Advanced (LTE-A) [1]. However with the increase in demand of high data rate, the mobile devices have also gradually become more Power Hungry.

Mobile phones switch between two states: RRC _Connected and RRC _Idle, in RRC _Connected state; it monitors the UE’s Physical Downlink Control Channel (PDCCH) [1], stays in RRC _Connected as long as the packets are being transferred and switches to RRC _Idle mode if there is no packet activity for a pre-defined time. At the time of design of LTE and the initial WiMAX 802.16e standards [3] (released in 2008/2006 respectively) timeframe, application traffic was largely dominated by Web browsing, Email, File transfer, Voice over IP (VoIP) types of applications [4] and with the development of Diverse Data Applications (DDA) [5]-[7] like Twitter, Whatsapp, Skype, etc which generates small short packets at certain intervals caused the mobile phones to switch between the two modes several times. This in turn caused a huge power loss as in idle state the UE gets disconnected from the Base Station, so when it switches back it has to send a large amount of Overhead Packets to register again.

To avoid huge power loss in RRC _Connected state, an intermediate state within the connected state is introduced called Discontinuous Reception (DRX) [8], [9]. In this state, the UE shall monitor only when there is a transfer of packets and move to sleep when there is no activity within the Connective mode. We will discuss more about DRX mechanism in Section II. We might be able to conserve the battery of UE but this comes at the cost of high latency due to the introduction of Sleep Mode in Connected State. Hence there is a need to optimize the parameters of DRX [10], [11] to balance both the trade-off i.e. Power Saving and Latency.

In the literature survey, we came across various papers in which Authors used different DRX parameters and states for different kinds of traffic i.e. Active and Background Traffic [12]-[14] to reach the optimization for both latency and power saving but in this paper certain adaptations are considered in a single model that shall work in all kinds of Traffic Intensities and have tried to include other parameters like channel strength that will help to improve the QoS [15] while saving the power and hence increase the performance of DRX mechanism.

The rest of the paper is organized as follows. Section II describes the DRX mechanism in LTE-A. Implementation of DRX mechanism using Semi-Markov process is shown in Section III. Section IV introduces a modified DRX mechanism and shows the comparison between former and later mechanism. Section V introduces Adaptive DRX mechanism on the basis of channel strength. The paper is concluded in Section VI.
DRX Mechanism

In LTE-A devices, there are two states: Connected and Idle. The network runs a RRC _inactivity timer to switch it from connected state to ideal state. In idle state, it is disconnected from network and only responds to paging messages which are very small so that it can perform handoffs, selection and reselection of cell. Now, DRX state is introduced within the connected state which decreases the frequency of switching between the former two states and saves power even when connected. The following figures show the pattern of traffic.

**Figure 1: HTTP traffic**

![HTTP traffic pattern](image1)

Fig.1 shows the traffic pattern in HTTP traffic, the protocol used for web browsing. So we can see that there are ON and OFF states i.e. packet activity for a certain time and then no activity.

**Figure 2: FTP traffic**

![FTP traffic pattern](image2)

Same can be inferred from Fig.2 and Fig.3, so Power management i.e. DRX management can be set up in connected state only with appropriate parameters to get suitable output.

**Figure 4: DRX Scenario**

As shown in the Fig.4 there are different states in DRX i.e. inactivity timer (Ti), ON duration timer (Ton), Long sleep cycle (Tls), Short sleep cycle (Tss) and short cycle timer (Nsc) [16]-[18]. UE’s states are described verbally in section III. Inactivity timer is set to monitor PDCCH. If no activity occurs in that certain time, it moves to other sleeping state and if it does occur then the timer gets restart. There is Ton state during which UE remains in touch with network so that if any packet activity occurs it can switch the UE to S1 state. There can be multiple short sleep cycles to delay the arrival of long sleep cycles by assigning the value to Nsc before going to S4 state. The DRX state timer is comparatively less than the RRC _inactivity for idle state so that it doesn’t go to idle state in between DRX mode. Instead of Uplink, the focal point of this research paper shall be on the analysis of Downlink Traffic.

**Figure 5: DRX States**

![DRX States](image5)

**DRX Mechanism as Semi-Markov Process**

Semi-Markov process [19] representing DRX mechanism is shown in Fig.5 which indeed shows the various states of UE equipment [12].
S1 (Active state): In this state the UE monitors the PDCCH for a certain time (Ti). If any packet arrives the UE restarts the Ti timer, if not then it moves to the S2 state on expiry of Ti timer. It is the highest power consuming state as it continuously monitors the PDCCH and receives the packets.

S2 (DRX on state): This is a small awake timer that comes in between the sleep time so that if any packet comes during sleep time then it can send the UE to the S1 state to receive it and if no packet comes it again goes to short sleep S3 or long sleep S4 depending on the Nsc value.

S3 (DRX short sleep): In this state the UE does not monitor PDCCH and at the end of sleep time decides whether to go to state S1 or S2 depending on the packet activity. If packet arrives it moves to S1 and if it doesn’t then it moves to S2.

S4 (DRX long sleep): If there is any activity during S4 then after its termination the UE goes back to state S1 and if there is no activity, then it moves to S2. In this state it neither receives nor monitors PDCCH. Consequently, in this state we are able to save a lot of power and the sleep time period becomes very high but at the same time these actions take place at the cost of latency.

Author [12] suggests a large value of Nsc during Active traffic as there is a need of minimum amount of latency or end to end delay and remove long sleep cycles. Whereas during background traffic the emphasis is more on power saving, hence remove short sleep cycle and introduce long sleep cycle i.e. no S3 state and after S2 it moves to S4 state. This process was implemented in MATLAB. All the states were considered as a single unit and ran for several times with outer loop value of 10. Some changes were made with each outer loop which is discussed later in this section. To make it a realistic scenario input was considered as Poisson distribution with mean arrival time as 60 , therefore the input varies from 40 and 80. This input was used to set threshold for each four state to decide whether the packet has arrived or not. The entire decision making has been done on the basis of these threshold values. Simulation results and the parameter table of existing DRX are shown below.

**Table 1: Specifications Adopted for DRX States**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Value(loops)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti</td>
<td>6</td>
</tr>
<tr>
<td>Ton</td>
<td>2</td>
</tr>
<tr>
<td>Tss</td>
<td>4</td>
</tr>
<tr>
<td>Tls</td>
<td>6, 10, 14, 18, 22, 26, 30, 34, 38, 42</td>
</tr>
<tr>
<td>Nsc</td>
<td>1</td>
</tr>
</tbody>
</table>

With each outer loop the long sleep cycles are increased by 4. To compare the results from all different outer loops the inner loops are limited to 500, so when it exceeds 500 it automatically goes to next outer loop.

Fig. 6 shows that how the number of long sleep cycles are increasing with each outer loop and in correspond to that how the sleep duration is varying as shown in Fig. 6. When long sleep cycles are 70 cycles in 500 cycles then the overall sleep duration is 0.48 and when the long sleep cycles are 346 in 500 cycles then the overall sleep duration is 0.82. Hence long sleep cycles can be placed between the two packet calls in background traffic as the interarrival time is higher i.e. 50 to 80ms.

Mathematical calculations [12] are shown below

\[
\Gamma_{SLEEP} = \frac{\varphi_1E[\Theta_1] + \varphi_2E[\Theta_2]}{\varphi_3E[\Theta_3] + \varphi_4E[\Theta_4]} \tag{1}
\]

Where, \(\varphi_1, \varphi_2, \varphi_3, \varphi_4\) are the probabilities of staying in or coming to the state1, 2, 3, 4 respectively and \(E[\Theta_1], E[\Theta_2], E[\Theta_3], E[\Theta_4]\) are the holding time of the state1, 2, 3, 4 respectively.

Now to increase the power saving we need to decrease the \(\varphi_1\) as the lesser we switch to the state1 the lesser is power loss so, in order to achieve that we need to minimize the value of \(P_{31}\) and \(P_{41}\) where \(P_{kl}: k, l \in \{1, 2, 3, 4\}\) represent the transition probability from state \(k\) to \(l\).
But decreasing the probability of switching to state 1 will result in latency. In order to find the best tradeoff between power saving and latency a profound effort is made to make the state 2 more adaptive towards different traffic intensities which will be later discussed in Section IV.

![Figure 8: Delay VS increase in long sleep cycles](image)

Fig.8 shows how end to end delay is increasing with the increase in long sleep cycles. With each outer loop Tls is increased by 4 loops so initially the delay is 0.003 and when the maximum Tls is 40 in the 10th outer loop the delay becomes 0.045 and also to run 500 loops the time taken is 0.8s. Now from the above figures it is clear how both the parameters are varying when the long sleep cycles are increased. This is the reason why the author [13] suggested to have only short cycles for active traffic where the least latency is required and only long sleep cycles in background traffic where the power saving is the main motive. But in this paper there is an attempt to amalgamate both the scenarios in single model which has been discussed in Section IV.

**Proposed DRX Mechanism**

The proposed DRX mechanism (MDRX) also consist of the four same states as discussed in Section III but there is a difference in switching of states as shown in Fig.9.

As shown in Fig.9 there is no direct switching between S3 and S1 and similarly for S4 and S1. All the switching occurs via S2 that decides where to send the UE depending on the incoming traffic. So there is no imaginary state of decision making as it was in existing DRX mechanism after the S3 and S4 state that would decide whether to switch to S1 or S2. Here everything is handled by State 2 i.e. ON duration timer.

ON duration is a very small timer as compared to inactivity timer so we are just trying to minimize the probability of going to S1 to save the power till the switching is mandatory. In the proposed DRX the packets are received during Ton which is not the case in the existing DRX mechanism. Hence a threshold was set in Ton, so till that threshold it can receive the packets. If packet intensity is high then it won’t be able to receive all and hence will switch to S1. This threshold can be varied according to the Ton timer size and this threshold can be number of packets or the time limit. In the simulation the threshold was set to single packet, if more than 1 packet arrives during the state 2 then the UE will switch to S1. So precisely if there is high traffic intensity then the UE will behave like exiting DRX model and for moderate traffic sometimes it will switch to S1 and sometimes not. Eventually this shall result in power saving without actually increasing the latency. In this particular scenario the main motive is to switch the model mainly in two states i.e. S2 and S3 and avoid S1 and S4 to get better results and get a better tradeoff between power saving and latency. Following results show the comparison between the two models on the basis of varied traffic intensity.

![Figure 10: High intensity Traffic](image)

![Figure 11: High intensity Traffic](image)
As shown in Fig.10 and Fig.11 the results from both the model are tracing each other. In case of high traffic, the probability of packet arrived is greater then 0.8. Hence S2 state in proposed model tends to switch to S1 as packets are above the threshold value and proposed model behaves like the previous DRX model. So both the latency and sleep duration is shown in Fig.10 and Fig.11 respectively. From the Fig. it can be inferred that in case of high intensity traffic the delay is least i.e. for high Tls value delay is 0.02 when Tls value is 40 for 10th outer loop but the sleep duration also is minimal i.e. 62% in comparison with other scenarios. This is because the UE tends to switch more to the S1 state which is not the case in other scenarios (same as the active traffic in the previous DRX model).

Results for low intensity traffic i.e. P(Packet arrived)>0.2 are shown in Fig.14 and Fig.15. Clearly sleep duration in proposed model is 83% that is better than 77% of existing DRX model. But the latency i.e. 0.04 in this case is more; as in this model S2 state will rarely switch to S1 state due to low traffic intensity whereas the existing DRX model will switch to S1 state if it senses any downlink/uplink packet.

Hence this framework will focus more on power saving as this kind of traffic belongs to background traffic which consists of single packet. Here Fig.15 shows that latency is more than other scenarios as it directly depends on the switching from S2 to S1. Hence more it switches to S1 state lesser is the latency with less power saving therefore to find the best solution S2 state was made more adaptive towards all kinds of traffic.
Table II: Comparison Between Power Saving For Different Traffic Models

<table>
<thead>
<tr>
<th>Traffic intensity</th>
<th>Power saving (DRX)</th>
<th>Power saving (MDRX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>62%</td>
<td>62%</td>
</tr>
<tr>
<td>Moderate</td>
<td>65%</td>
<td>76%</td>
</tr>
<tr>
<td>Low</td>
<td>77%</td>
<td>83%</td>
</tr>
</tbody>
</table>

Table III: Comparison Between Latency For Different Traffic Models

<table>
<thead>
<tr>
<th>Traffic intensity</th>
<th>Latency(sec) (DRX)</th>
<th>Latency(sec) (MDRX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.034</td>
<td>0.036</td>
</tr>
<tr>
<td>Low</td>
<td>0.036</td>
<td>0.04</td>
</tr>
</tbody>
</table>

in [12] a point is specified that for active traffic the concern is more on minimizing the latency. Hence to implement that there is a need to delay the arrival of long sleep cycles by increasing the value of number of short sleep cycles(Nsc) (Nsc ∈ {1,3,5,7,9}). As the Nsc value increases the sleep duration will decrease i.e. we are preferring latency over power saving for these kind of high intensity traffic (Active traffic). In the proposed DRX model we can include this parameter into the S2 state that has a threshold limit so it can detect what kind of traffic is coming.

Adaptive DRX Over Varying Channel Strength

As discussed before not only the power saving but also the quality of service (QOS) and minimum latency are important [14], [20]. If the UE’s location is not constant and varying with time than with that, channel strength will also vary as channel strength is inversely proportional to cube of distance. Author [21] refers that if the channel strength is high then opt for long sleep time but if it’s low then switch to long awake time to have a good QOS and minimum latency.

If the channel strength is weak then there may be 3 factors that must be consider.

- With increase in distance the propagation delay will also increase.
- Base station chooses to send the packets that are in its buffer to those UEs whose channel strength is good so more delay for those with less channel strength.
- With weak channel strength the probability of proper or correct reception of data decreases as noise remains the same so UE may have to NACK back to eNB to ask for the correct data.

So to implement that with minimum latency we need to increase the inactive timer of DRX model so that it doesn’t move to sleep state before receiving the correct data or else it will lead to high latency which we are trying to avoid in this paper.

Therefore DRX model is made more adaptive to channel strength by considering two parameters: distance from eNB and inactivity timer (Awake time) as shown in algorithm.

Algorithm: Channel strength based DRX configuration

1. Begin
2. Repeat the following for 10 outer loops.
3. Receive the distance between UE and eNB
4. Repeat the following for certain duration of time i.e.20 loops to adjust the changes with each outer loop
5. if distance<dist1&&Ti<Ti1
6. Ti=Ti+1
7. elseif distance<dist1&&Ti>Ti1+delta
8. Ti=Ti-1
9. elseif distance<dist2&&Ti<Ti2
10. Ti=Ti+1
11. elseif distance<dist2&&Ti>Ti2+delta

Therefore if the active traffic is going on it can automatically increase the Nsc value before switching it to the S1 state and if the low intensity traffic is coming it can decrease the Nsc value. As in low intensity traffic the concern is more on the power saving. Results from both the models are shown in Fig.16.

![Figure 16: Sleep Duration VS Increase in Nsc](image.png)
In the above algorithm the location of UE is constantly monitored and according to that the parameters of DRX model are changed. As discussed in the section how channel strength decreases with the distance, several threshold distances are set and according to that how the awake time should increase or decrease is observed. This algorithm was implemented on MATLAB and the results are shown below.

In this paper a profound effort is made to find the best tradeoff between power saving and latency. Following this a single adaptive MDRX is proposed for all types of traffic (active and background traffic). In MDRX it has been shown how S2 state is the core of the whole model and how it works adaptively depending upon the traffic intensity. The results have been compared with the existing DRX model. The results proved the efficiency of MDRX and better outcomes have been achieved. Even an adaptive mechanism is introduced to keep the MDRX states unchanged while the traffic intensity changes. Thereafter it has also been discussed, that how increasing the number of short cycles (Nsc) in high intensity traffic can help to minimize the latency. A detailed explanation on how to increase the Nsc while switching from S2 to S1 in the case of high intensity traffic has also been presented. Further channel strength has been introduced as a new parameter and taking distance and inactive timer as variables a successful attempt is made to set particular inactive timer for a particular threshold distance. From the results it has also been confirmed that the QoS is directly proportional to the inactive timer of MDRX. Hence in order to increase the QoS an increase in inactive timer is required.

REFERENCES


