LJMU Research Online

Memon, ML, Maheshwari, MK, Saxena, N, Roy, A and Shin, DR

 Artificial intelligence-based discontinuous reception for energy saving in 5G networks

http://researchonline.ljmu.ac.uk/id/eprint/25720/

Article

For more information please contact researchonline@ljmu.ac.uk

Artificial Intelligence-Based Discontinuous Reception for Energy Saving in 5G Networks

Mudasar Latif Memon [1] [ID]**, Mukesh Kumar Maheshwari [2], Navrati Saxena [3,*]** [ID]**, Abhishek Roy [4] and Dong Ryeol Shin [3,*]**

[1]   College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, Korea
[2]   Department of Electrical Engineering, Bahria University, Karachi 75260, Pakistan
[3]   College of Software, Sungkyunkwan University, Suwon 16419, Korea
[4]   Advanced Communication Technology, Wireless System Design, MediaTek USA Inc.,
      San Jose, CA 95134, USA
*   Correspondence: navrati@skku.edu (N.S.); drshin@skku.edu (D.R.S.); Tel.: +82-31-299-4676 (N.S.)

Abstract: 5G is expected to deal with high data rates for different types of wireless traffic. To enable high data rates, 5G employs beam searching operation to align the best beam pairs. Beam searching operation along with high order modulation techniques in 5G, exhausts the battery power of user equipment (UE). LTE network uses discontinuous reception (DRX) with fixed sleep cycles to save UE energy. LTE-DRX in current form cannot work in 5G network, as it does not consider multiple beam communication and the length of sleep cycle is fixed. On the other hand, artificial intelligence (AI) has a tendency to learn and predict the packet arrival-time values from real wireless traffic traces. In this paper, we present AI based DRX (AI-DRX) mechanism for energy efficiency in 5G enabled devices. We propose AI-DRX algorithm for multiple beam communications, to enable dynamic short and long sleep cycles in DRX. AI-DRX saves the energy of UE while considering delay requirements of different services. We train a recurrent neural network (RNN) on two real wireless traces with minimum root mean square error (RMSE) of 5 ms for trace 1 and 6 ms for trace 2. Then, we utilize the trained RNN model in AI-DRX algorithm to make dynamic short or long sleep cycles. As compared to LTE-DRX, AI-DRX achieves 69% higher energy efficiency on trace 1 and 55% more energy efficiency on trace 2, respectively. The AI-DRX attains 70% improvement in energy efficiency for trace 2 compared with Poisson packet arrival model for $\lambda = 1/20$.

Keywords: discontinuous deception; multiple beam communications; artificial intelligence; energy efficiency; 5G; wireless communications

1. Introduction

The use of cellular gadgets, like smartphones, notebooks, and tablets has comforted our life. The Ericsson mobility report predicts the rise of cellular traffic to 8.8 billion by 2024 [1]. These extensive growing cellular users require improved data rates with heterogeneous services in next generation networks. 5G expects to deal with various types of traffics including periodic and delay tolerant traffic for IoT devices or burst type of traffic for delay intolerant services [2,3]. 3rd Generation Partnership Project (3GPP) planned the standardization of GHz spectrum (mm-wave) to address the users' demand of high bandwidth. However, communication over high-frequency bands of the mm-wave requires directional air interface and narrower beams to reduce the path loss. In directional air interface, UE has to search for best beam pairs and make adequate beam alignment with the next generation nodeB (gNB) [4]. In addition to beam searching process in 5G networks, massive MIMO, higher order modulation schemes and advanced coding techniques also increase UE energy expenses.

Long term evolution (LTE) networks utilize DRX mechanism to reduce the energy consumption of the UE [2,5]. DRX enables a UE to save energy by switching off the radio circuitry part, in case of no incoming data. LTE-DRX turns off radio part during long and short sleep cycles in order to reduce the energy consumption of UE. The sleep cycles in LTE-DRX are of the fixed time period. If any new packet arrives during the fixed time sleep cycle, the packets will be buffered at evolved NodeB (eNB). The eNB serves the stored packets after the completion of each sleep cycle in LTE-DRX. Hence, the LTE-DRX saves the UE energy at the cost of delay [6,7]. LTE-DRX in its current form is not suitable for energy savings in 5G enabled cellular devices [8,9], due to two main reasons.

- The communications in LTE do not consider beamforming, whereas in 5G networks UE has to align the best beam pair before the start of communications.
- The LTE-DRX mechanism has a fixed length of sleep duration, which increases the energy consumption and delay. Hence, LTE-DRX is not suitable for low latency communications in 5G networks.

Authors in [3] propose DRX for 5G network, which requires UE to search for best beam pairs after completion of each sleep cycle in order to serve the packets. The additional beam searching operation after each sleep cycle enhances the energy consumption of UE. The work in [10] suggests beam aware DRX approach for 5G enabled machine to machine communications, in which UE has prior information about best beam pairs. Similarly, beam aware DRX mechanism for energy saving in the 5G network is proposed in [11]. The beam aware mechanism may not perform well in case of UE mobility and beam misalignment [12]. Kwon et al. [12] show that the probability of beam misalignment increases with an increase in UE velocity. The probability of misalignment is 0.1 and 0.38 if a UE moves with the velocity of 30 Km/h and 60 Km/h, respectively.

The DRX with built-in state of beamforming is proposed by Liu et al. [13]. Authors present the concept of DRX for multiple beam communications and utilize the semi-Markov model to design eleven-states of DRX. Their approach considers the beam training process only in case of beam misalignment and after completion of long sleep duration. Hence, authors save the energy of UE while minimizing the delay for 5G services. However, the fixed duration of sleep cycles in their approach may cause more energy consumption.

On the other hand, Recurrent Neural Network (RNN) in AI has shown incredible results to predict the upcoming value of a given sequence [14]. Long Short-Term Memory (LSTM) is a popular type of RNN that is specially designed to learn long-term dependencies of a sequence for predicting the upcoming value of a sequence [15]. The term long-term dependency refers to the sequence, whose desired/current output values (prediction results) depend on long-sequence of previous input values rather than the only single previous input value.

Motivated by the success of RNN to learn and predict long-term dependent sequence values in various applications [11,14,16], we use RNN to extract the pattern of packet arrival time from real wireless traffic traces and to predict the values of the next packet arrival time. Based on the prediction results, we propose an AI-DRX algorithm that works on a ten-state DRX model to enable energy saving. AI-DRX for multiple beam communications in 5G network saves the UE energy by enabling dynamic short and long sleep cycles, respectively. To be more specific, the following are our key contributions to save UE energy in multiple beam communications scenario of 5G networks.

- We perform the training of the LSTM neural network on wireless traffic. During training, LSTM network extracts the packet arrival time pattern from the wireless traffic trace. The prediction results show that the trained model predicts with minimum RMSE of 5 ms on random test set from trace 1 and 6 ms on random test set from trace 2, respectively.
- We devise DRX as a ten-state model.
- We propose an artificial intelligence based DRX mechanism for multiple beam communications in 5G networks. We suggest AI-DRX algorithm using ten state model to enable dynamic short or long sleep cycles, depending on the prediction results.

- We evaluate the performance of AI-DRX in terms of energy efficiency and mean delay. AI-DRX achieves the energy efficiency of 59% on trace 1 and 95% on trace 2, respectively; while considering the mean delay requirements of different services.

The remaining paper is organized as follows. In Section 2 we present an overview of existing DRX mechanism and introduce the RNN with focus on LSTM neural network. Section 3 proposes AI-DRX algorithm with a ten-state DRX model, to enable dynamic short or long sleep cycles in multiple beams communications of the 5G network. Section 4 presents the performance analysis of AI-DRX in terms of energy efficiency and mean delay. Finally, we conclude our work in Section 5.

2. Related Work

2.1. DRX in LTE

LTE networks use DRX mechanism to minimize the power consumption of UE. The energy expenditure of UE can be reduced by switching off radio components during the unavailability of incoming data packets [17]. LTE-DRX is configured and controlled by evolved node B (eNB). The eNB informs the UE to turn off its radio components in case of no data in the buffer. The LTE-DRX can be regulated by radio resource control (RRC) layer at eNB. RRC sends the packets' information to UE via a physical downlink control channel (PDCCH). RRC operates in two modes: (1) RRC_connected mode; (2) RRC_idle mode, after a UE is turned on [18]. The RRC_idle mode is only responsible for paging operations, UE neither receives nor transmits the data, but only monitors the paging signals during the paging occasion. Whereas, all the data exchange between UE and eNB takes place during RRC_connected mode. Since all transmissions take place during the RRC_connected mode and this mode is responsible for more energy consumption of UE, our work focuses on the DRX in RRC_connected mode. Figure 1 shows the LTE-DRX in RRC_connected mode that works on two types of states. These states are:
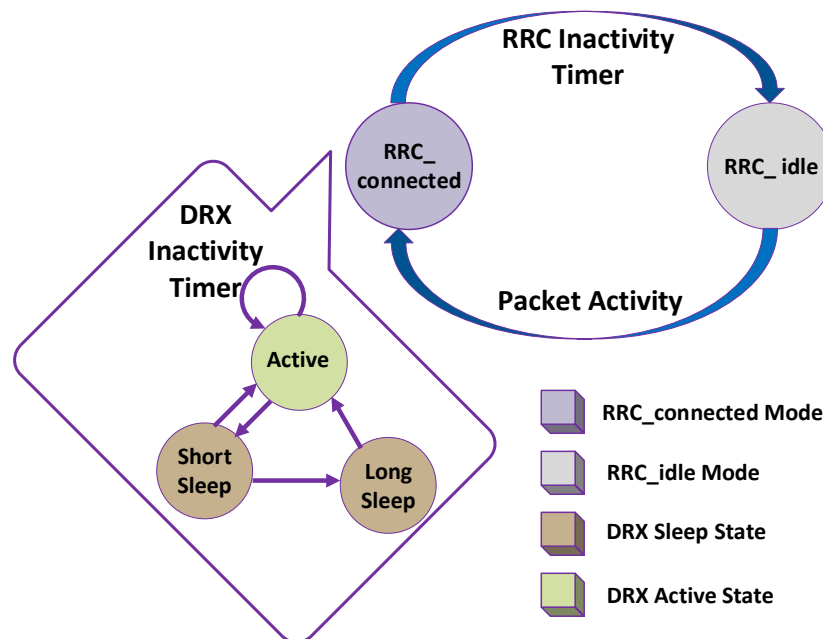
- Active State
- Sleep State

Figure 1. LTE-DRX in RRC_connected mode.

Based on the above two states, different DRX parameters are configured in RRC_connected mode while considering various services' delay requirements.

2.1.1. Active State

A UE sends or receives the data packets during DRX active state. This state does not allow UE to save the energy because UE has to receive the data packets during this state. Moreover, a UE has to turn on the radio circuitry throughout the active state. The active state of DRX can be controlled by two parameters.

- Inactivity timer
- Active time

All the data packets are transmitted/received during the active time. The inactivity timer is a countdown timer in DRX active state. This timer re-starts every time a new data packet arrives at eNB and then eNB serves the received packets to UE. In case of no new packet arrival, the inactivity timer gets expired and the UE switches to sleep state for a certain time period.

2.1.2. Sleep State

A UE turns off the radio components during sleep state in order to save power. UE cannot receive or send the data packets during DRX sleep state but can only monitor the PDCCH for any incoming data. DRX in a sleep state is controlled by the following parameters:

- ON time
- Short sleep cycle
- Short sleep timer
- Long sleep cycle
- Long sleep timer

During ON time, a UE monitors the PDCCH. ON time always starts after completion of each sleep cycle. A short sleep cycle is a small duration of time that saves the UE energy by switching off the transceiver part. A short sleep cycle is repeated up to N_{sc} number of short sleep cycles. After the expiry of N_{sc}, a UE transits to the long sleep cycle. The long sleep cycle and long sleep timer are similar to short sleep cycle and short sleep timer but have a longer time period than the counterpart, respectively.

Figure 2 delineates the timing diagram of LTE-DRX. As shown in Figure 2, all the data transmission and reception take place during the active state. After reception of each new packet, inactivity timer restarts. In the case that no new packet reaches and inactivity timer finishes the countdown timer, then UE switches to short sleep cycle. After every short sleep cycle, a UE monitors the PDCCH for any incoming packet during the ON time. If no new packet arrives before the completion of short sleep timer, the UE switches to the long sleep cycle and remains there until intimation of the new packet is received. UE transits from a long sleep to idle state if no new packets arrive and the long sleep timer expires.
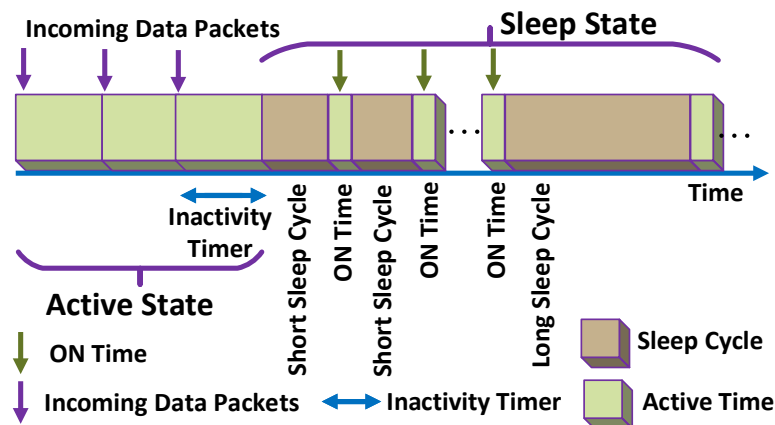
Figure 2. LTE-DRX timing diagram.

2.2. Non-Compatibility of LTE-DRX for 5G Networks

The LTE-DRX mechanism in its existing form is not suitable for 5G networks. The main reason for non-suitability is the directional communications in 5G networks, which require beamforming operation prior to data transmission [4]. Whereas, LTE-DRX mechanism does not include beam searching operation [19,20]. Figure 3 shows the concept of directional communication with multiple beams in 5G network. The gNB transmits K number of beams and a UE has L number of beams. A UE needs to search and align the best beam pair form $K \times L$ beam pairs, prior to the start of communications [4]. The process of beam searching and beam alignment in 5G is not included in LTE-DRX. Moreover, the process of beam searching in 5G network causes a UE to remain in the active state more than that of the LTE-DRX, which also increases the power consumption of UE [21]. Furthermore, the beamforming process in 5G requires additional time for beam searching and aligning best beam pairs [22]. The additional beam searching and beam alignment time cause more delay than the delay in LTE-DRX. Hence, directional communications in 5G is one of the reasons, which makes LTE-DRX mechanism in its existing form, not a suitable solution for power saving in 5G networks [21,22].

The second main reason for non-compatibility of existing LTE-DRX mechanism with 5G networks is the fixed length of sleep cycles in LTE-DRX [11]. LTE-DRX uses fixed length short and long sleep cycles to economize the power consumption of UE. Whereas, 5G network is expected to deal with different types of services, simultaneously [4]. These services may have different size of packet lengths, variable packet arrival time and variable transmission time interval (TTI) [23]. The fixed length of sleep cycles in LTE-DRX mechanism may not be suitable for 5G services as these cycles may under-utilize or over-utilize TTI. Moreover, fixed length sleep cycles in 5G may be responsible for least power savings in 5G enabled devices [23]. However, DRX design can still be used for 5G [9,24,25].
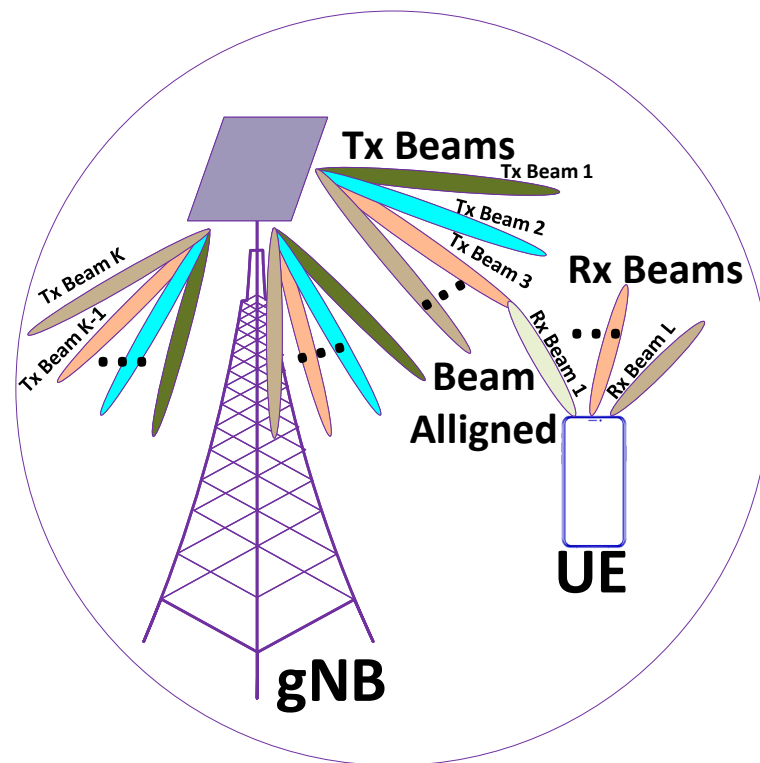
Figure 3. Directional communication with multiple beams in 5G network.

2.3. Recurrent Neural Network for Predicting Sequential Data

Recurrent Neural Network (RNN) is a part of AI that is widely used for predictions over sequential data [26]. Long short-term memory (LSTM) network is one of the popular kinds of RNN that performs well while predicting the long-term dependent time sequences [15]. LSTM is different from other kinds of RNN due to its gated structure and internal cell state in every single unit, as shown in Figure 4. LSTM neural network utilizes various previous inputs to decide the values of internal gates (forget gate, input gate, input modulation gate, output modulation gate), which contributes to cell state. Cell state helps the LSTM to remember or forget the impact of past inputs while deciding the prediction results. This is the main reason LSTM has better performance for predicting the long-term dependent sequences [14,15]. Table 1 shows the notations used in our work.

LSTM unit includes forget gate F_T that can be mathematically shown by Equation (1). The output of forget gate varies from 0 to 1, due to sigmoid $Sig = \frac{1}{1+e^{-T}}$ function. The letters w_x and w_h in the Equation (1) represent the weights associated with the current input x_T and previous output h_{T-1}, respectively. The term b_F is the bias for the forget gate. Similarly, the input gate (I_T) and input modulation gate (G_T) can be calculated by Equations (2) and (3), respectively. The hyperbolic tangent function in input and output modulation gates can be mathematically written as $\tanh(T) = 2Sig(2T) - 1$ and ranges from -1 to $+1$. Moreover, b_I and b_C are biases for input gate and cell state, respectively.

$$F_T = Sig(w_x x_T + w_h h_{T-1} + b_F) \tag{1}$$

$$I_T = Sig(w_x x_T + w_h h_{T-1} + b_I) \tag{2}$$

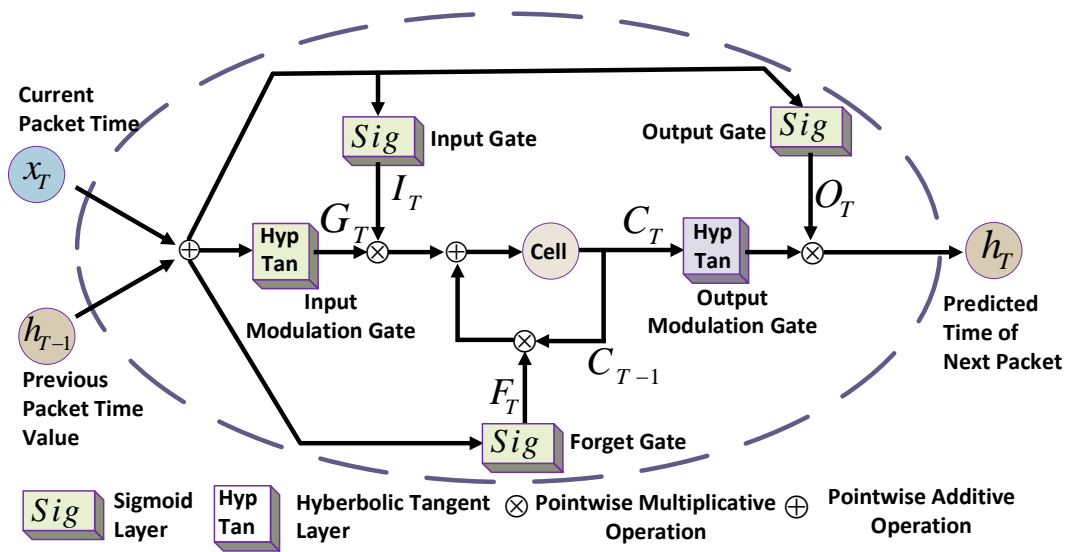$$G_T = \tanh(w_x x_T + w_h h_{T-1} + b_C) \tag{3}$$

Figure 4. Gated structure of single long short-term memor (LSTM) unit.

Table 1. List of abbreviations.

Abbreviation	Notation	Abbreviation	Notation
Forget gate	F_T	DRX active time	T_{Ac}
Input modulation gate	G_T	DRX ON time	T_{ON}
Input gate	I_T	DRX Inactivity timer	T_{IN}
Output layer	O_T	Packet inter-arrival time	λ_κ
Cell state	C_T	DRX Dynamic sleep time	T_{DY}
Output modulation gate	h_T	Weight corresponding to each gate	w_i

Cell state C_T and the output layer O_T can be calculated by Equations (4) and (5), respectively. The final prediction results can be observed by output modulation gate (h_T) using Equation (6).

$$C_T = F_T \otimes C_{T-1} + I_T \otimes G_T \tag{4}$$

$$O_T = Sig(w_x x_T + w_h h_{T-1} + b_O) \tag{5}$$

$$h_T = O_T \otimes \tanh(C_T) \tag{6}$$

During the training process, LSTM learns to extract the relationship between the input and desired output (upcoming value of sequence) by adjusting various weight values in Equations (1)–(3) and (5). Once the model is trained with the least error, the learned weight values can be used in Equations (1)–(6) to calculate the output values of a sequence (prediction). The error between prediction result and observed (actual) value can be computed by the root mean square error (RMSE) and is given as:

$$\text{RMSE} = \sqrt{\frac{\sum_{T=1}^{N}(Predicted_T - Observed_T)^2}{N}} \tag{7}$$

where the notation T represents the number of samples and N shows the maximum number of samples.

3. AI-DRX for Multiple Beam Communications

3.1. System Model

In 5G directional air interface, beam information is a very crucial operation as a UE wakes up to receive data or control packet [3]. Hence, our work considers DRX for multiple beam communications [13], having dynamic long and short sleep cycles instead of fixed length sleep cycles. Dynamic sleep cycles have different values of sleep time in each cycle rather than one static time value in each sleep cycle. The random values of dynamic sleep time can be predicted based on previous values of packet arrival time.

We model DRX as a ten-state model. These states range from S_0 to S_9 and are shown in Figure 5. Let us elaborate on the states below.

- S_0 is the active time. A UE sends and receives the data during S_0 state. UE predicts the dynamic time T_{DY} during S_0 state.
- S_1 state is threshold comparison state. S_1 compares T_{DY} with both threshold values (Th_{Min} and Th_{Max}) to decide whether to remain in the active time or switch to the dynamic sleep cycle.
- S_2 state is dynamic short sleep cycle. It saves the UE power for a short period of time up to T_{DY}. The predicted value of T_{DY} in the dynamic sleep cycle is always lower than that of value in S_3 state.
- S_3 state highlights the dynamic long sleep cycle. It is similar to S_2 state but has a larger sleep period than S_2. The predicted value of T_{DY} always has larger value of dynamic long sleep cycle than that of the dynamic short sleep cycle.
- S_4 state shows beam training during active time. This state allows UE to train and search for best available beam pairs between gNB and UE.
- S_5 state represents feedback after beam training operation in active time.
- S_6 state shows active time after best beam pairs are aligned. S_6 is different from S_0 as S_6 always occurs after beam training and feedback process. Moreover, S_6 enables UE to search beam pairs in active time until best beam pairs are found.
- S_7 state delineates beam training process after the execution of dynamic long or short sleep cycle.
- S_8 state is feedback operation that occurs after beam training and after the execution of dynamic long or short sleep cycle, respectively.
- S_9 state is the ON period, a UE only monitors PDCCH for an incoming packet. UE could not send or receive data packets during ON time.

3.2. Proposed AI-DRX Algorithm

We propose artificial intelligence based DRX (AI-DRX) algorithm for multiple beams communications in 5G network. Our proposed algorithm enables DRX to achieve dynamic long or short sleep cycles and to reduce the power consumption of UE. AI-DRX algorithm works on ten-state model of DRX as shown in Figure 5. Algorithm 1 demonstrates AI-DRX mechanism and is elaborated below:

- AI-DRX algorithm takes ON timer T_{ON}, minimum threshold value Th_{Min} and maximum threshold value Th_{Max} as input (Line 1).
- AI-DRX examines the buffer for any incoming data packets (Line 2).
- If any packet is received in the buffer, the packets will be served and the value of dynamic sleep time T_{DY} will be predicted, simultaneously (Line 4).
- During the active time, if no beam misalignment occurs between UE and gNB, AI-DRX compares the predicted value of dynamic sleep time T_{DY} with threshold values Th_{Min} and Th_{Max}.
- If the predicted value of dynamic sleep time T_{DY} is less than the minimum threshold value Th_{Min} (Line 8), UE continues to remain in the active time (Line 9).

- If the predicted value of dynamic sleep time T_{DY} is greater than or equal to the minimum threshold value Th_{Min} and less than maximum threshold value Th_{Max} (Line 11). UE observes dynamic short sleep cycle (Line 12).
- In case of any beam misalignment after UE wakes up from dynamic short sleep cycle (Line 13), UE starts the beam training process (Line 14) and Feedback (Line 15) followed by ON time (Line 16).
- During ON time, if beam pairs are still misaligned (Line 17), UE searches beam pairs again (Line 18).
- During ON time, if a new packet arrives (Line 21), UE switches to active mode and start receiving data packets (Line 22).
- After completion of ON time, if no new packet arrives, UE continues to sleep for previous predicted time value T_{DY} of the long sleep cycle or short sleep cycle. respectively (Line 24).
- If T_{DY} is greater than Th_{Max} (Line 28) the UE will go to dynamic long sleep cycle (Line 29).
- UE will perform beam training and feedback after completion of each dynamic long sleep cycle (Line 30).
- In the case of beam misalignment during an active time (Line 35), UE performs beam training (Line 36) and feedback (Line 37) and then re-enter active time (Line 38).
- During the active time (Line 38), if UE still find beam misalignment (Line 39), AI-DRX executes beam training process again (Line 40).
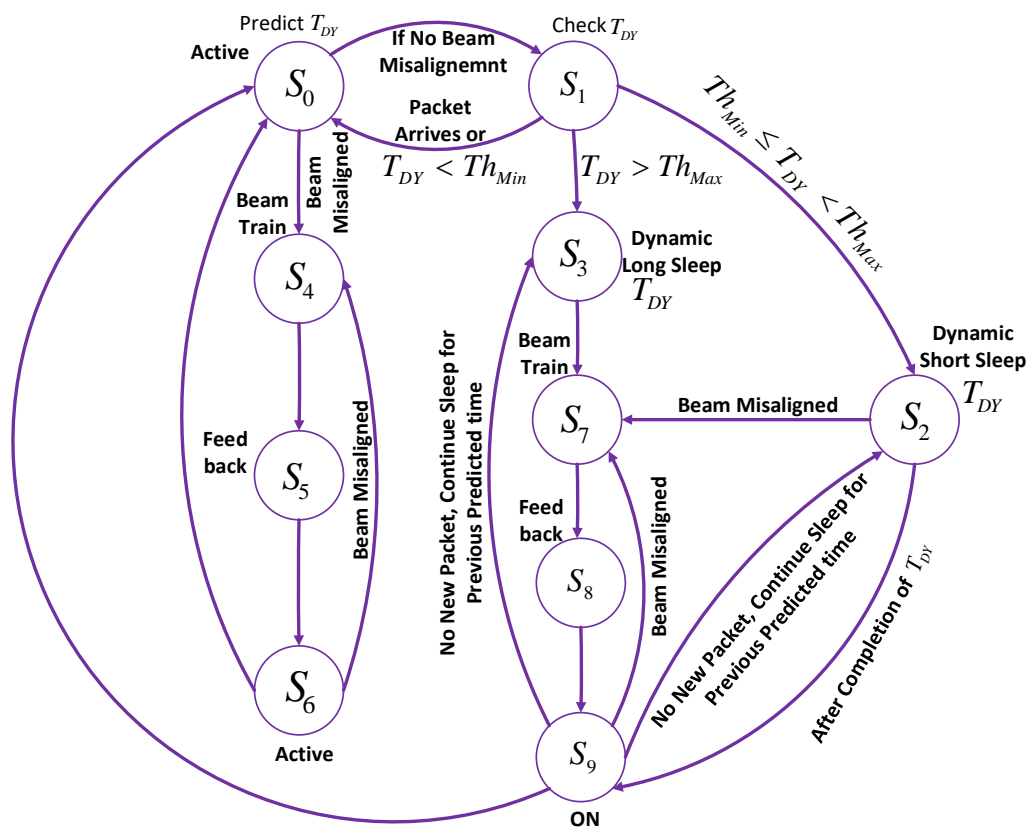- After the alignment of beam pairs, the UE transits to active time (Line 42).

Figure 5. State diagram for AI-DRX.

Algorithm 1 AI based DRX for Multiple Beam Communications.

1: Input T_{ON}, Th_{Min}, Th_{Max}
2: Examine buffer for incoming packets
3: **if** Packets in buffer > 0 **then**
4:     Serve the packets and Predict T_{DY} (S_1)
5: **else**
6:     **if** No Beam Misalignment **then**
7:         Check T_{DY} (S_1)
8:         **if** $T_{DY} < Th_{Min}$ **then**
9:             Go to Step 4 (S_0)
10:        **else**
11:            **if** $Th_{Min} \leq T_{DY} < Th_{Max}$ **then**
12:                Go to Dynamic Short Sleep up to T_{DY} (S_2)
13:                **if** Beam Misalignment **then**
14:                    Execute Beam Training (S_7)
15:                    Feedback (S_8)
16:                    ON (S_9)
17:                    **if** No Beam Aligned **then**
18:                        Go to Step 14 (S_7)
19:                    **end if**
20:                **else**
21:                    **if** New packet arrives before completion of ON Time T_{DY} **then**
22:                        Go to Step 4 (S_9 to S_0)
23:                    **else**
24:                        Go to Dynamic sleep for previous predicted time T_{DY} (S_2/S_3)
25:                    **end if**
26:                **end if**
27:            **else**
28:                **if** $T_{DY} > Th_{Max}$ **then**
29:                    Go to Dynamic Long sleep (S_3)
30:                    Go to Step 14 (S_7)
31:                **end if**
32:            **end if**
33:        **end if**
34:    **else**
35:        **if** Beam Misalignment **then**
36:            Execute Beam Training (S_4)
37:            Feedback (S_5)
38:            Active after Beam Training (S_6)
39:            **if** No Beam Aligned **then**
40:                Go to Step 36 (S_4)
41:            **else**
42:                Go to Step 4 (S_0)
43:            **end if**
44:        **end if**
45:    **end if**
46: **end if**

3.3. AI-DRX for Enabling Dynamic Long and Short Sleep Cycles

Algorithm 1 (AI-DRX) demonstrates the use of artificial intelligence in the implementation of DRX for 5G networks. AI-DRX makes dynamic short and long sleep cycles in DRX. AI-DRX utilizes trained LSTM model to predict the upcoming value of packet arrival time and subsequently to enable dynamic sleep cycles in DRX. The training process is conducted offline on two traces of real wireless traffic acquired from the University of Massachusetts (UMass) trace repository [27] and Crawdad data set repository [28]. Training process learns the packet arrival time pattern of increasing sequence values from both traces. Once the LSTM network is trained offline with least prediction error, the trained model predicts the upcoming packet time of real wireless traffic. AI-DRX algorithm calculates the dynamic sleep cycles by using prediction results of upcoming packet arrival time value.

AI-DRX enables a UE to reduce power consumption in multiple beam communications scenario of 5G networks. AI-DRX saves energy by enabling dynamic short and long sleep cycles, respectively. AI-DRX enables dynamic short sleep cycles or dynamic long sleep cycles based on two threshold values Th_{Min} and Th_{Max}. The values of Th_{Min} and Th_{Max} can be varied according to delay requirements of various services. We discuss three cases of AI-DRX: (1) dynamic short sleep cycles; (2) dynamic long sleep cycles; (3) dynamic inactivity timer.

During an active time of AI-DRX, the packets are served to UE. The trained LSTM model predicts the value of upcoming packet arrival time (T_{DY}) while serving the packets. UE checks for new packets in the buffer via PDCCH during ON state of AI-DRX. If no new packet is observed in the buffer, UE continues to sleep up to T_{DY}. However, if a new packet is observed in the ON state, UE transits to active time and receives the packet(s). Meanwhile, inactivity timer in the active mode is restarted on the reception of every new packet. Furthermore, AI-DRX also considers the case of the active state having empty buffer, the UE counts down for dynamic inactivity timer (third case below) to complete and then transits to sleep cycle for predicted sleep time (T_{DY}).

The first case enables dynamic short sleep cycle if the value of T_{DY} is greater than or equal to Th_{Min} and less than Th_{Max}. The condition of dynamic short sleep cycle gets satisfied and UE sleeps for the predicted time value of T_{DY}. There are very small chances of beam misalignment after a short sleep period [12,13]. Hence, AI-DRX considers the beam training process after dynamic short sleep cycles only if beams are misaligned. Figure 6 shows the timing diagram of dynamic short sleep cycle using AI-DRX algorithm.
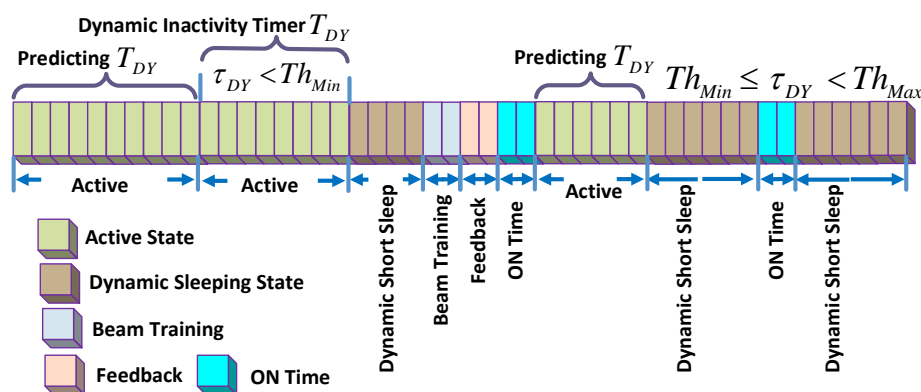

Figure 6. Timing diagram of AI-DRX based dynamic short sleep cycle.

The second case considers the dynamic long sleep cycle up to the predicted value of T_{DY}, if T_{DY} is greater than Th_{Max}. The concept of the dynamic long sleep cycle is depicted in Figure 7. The dynamic long sleep cycle saves more energy than that of short sleep cycle. Moreover, there are more chances of beam misalignment after the dynamic long sleep cycles. Hence, AI-DRX performs beam training and feedback after completion of each dynamic long sleep cycle.

The third case deals with dynamic inactivity timer. If the predicted value of T_{DY} is less than Th_{Min}, the UE will remain active until the period T_{DY} or any new packet arrives. The concept of dynamic inactivity timer can be seen in Figures 6 and 7. AI-DRX also addresses the problem of beam misalignment during active time. AI-DRX performs beam training and feedback in case of any beam misalignment during active time.

It may be noted that AI-DRX utilizes dynamic short and long sleep cycles instead of static fixed time sleep cycles. Moreover, the inactivity timer value is also dynamic according to the prediction results (using trained LSTM model). Our proposed algorithm keeps updating the predicted time value based on most recent received packets and their packet arrival time.
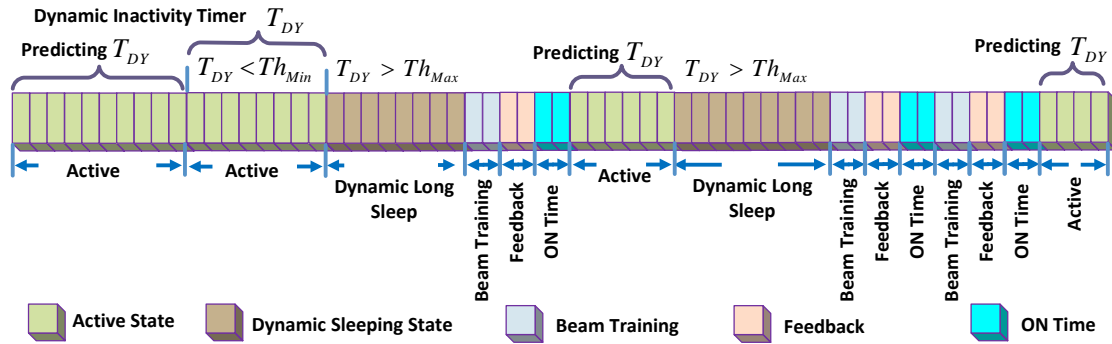
Figure 7. Timing diagram of AI-DRX based dynamic long sleep cycle.

DRX saves the power of UE at the cost of delay. Hence, we considered two performance parameters: energy efficiency (EE) and mean delay. The performance parameter; EE of the UE is the ratio of dynamic short sleep time and dynamic long sleep time to the sum of active time (T_{AC}), ON time (T_{ON}), beam training time, feedback time, dynamic short sleep (T_{DY}) time and dynamic long sleep time (T_{DY}). EE can be calculated by Equation (8). The beam training and feedback processes are considered during active time (T_{AC}) and dynamic short and long sleep cycles take place during (T_{DY}).

$$EE = \frac{T_{DY}}{T_{AC} + T_{ON} + T_{DY}} \tag{8}$$

Similarly, for packet inter-arrival time λ_κ, the mean holding period of active state can be calculated as [5]:

$$T_{AC} = \frac{1 - e^{-\lambda_\kappa T_{IN}}}{e^{-\lambda_\kappa T_{IN}}\left(1 - e^{-\lambda_\kappa}\right)} \tag{9}$$

Furthermore, the packets arrived during the sleep state and ON duration are stored in the buffer until next ON period. The packet arrival events are the random observer to the sleep period and ON state. Therefore, the mean delay is defined as the sum of mean sleep time and ON duration and is given as:

$$\text{Mean Delay} = \frac{T_{DY}}{2} + \frac{T_{ON}}{2} \tag{10}$$

4. Performance Analysis

We use MATLAB 2019a for training and testing of RNN on two different traces (data sets). These traces of burst traffic type are taken from the Crawdad dataset repository (trace 1) [28] and UMass trace repository (trace 2) [27]. The trace 2 shows the traffic pattern of HTTP and video streaming applications. The video streaming trace is used as it is expected by 2024 over three-quarters of mobile data traffic will be video traffic [1]. The trace 1 and trace 2 include seven parameters as shown in [11]. These parameters are: (1) serial number; (2) packet arrival time in seconds; (3) source address of the packets; (4) destination address of the packet; (5) protocol used; (6) length of packets in bytes; and (7) additional information. We have utilized the time parameter from both traces to train the LSTM network. Time parameter contains the information of packet arrival time. Our purpose is to train the LSTM network until it learns the time interval pattern of packet arrival time from a given sequence. Data pre-processing makes the training process simple and avoid training from the divergence [11,29]. Hence, we standardize the data with zero value of mean and unit value of the variance of the training set during the training process. Moreover, we also standardize the test set during the prediction time.

Hyper-parameters in LSTM network are selected manually to make the training process more efficient. These parameters include the number of hidden units, learning rate drop factor, the maximum number of epochs, initial learning rate, and the optimizer used. The best selection of hyper-parameters

may result in the learning with least prediction error. The training process performance can be measured in terms of RMSE. The less the value of RMSE, the better is the prediction results of the trained model on unseen data (test set). We have selected 200 and 125 hidden units in LSTM network during training process over trace 1 and trace 2, respectively. The value of the learning rate drop factor remains 0.2 for both traces. Maximum number of epochs during the training process of trace 1 are considered to be 600 while for trace 2 are 1000 epochs. Initial learning rate during the training process of both traces remains 0.004. Furthermore, the optimizer used during the training process for both traces is Adam optimizer. The above mentioned hyper-parameters reduce the RMSE during training and testing processes.

We have considered the total length up to 130,820 sample values (323.886 s) of trace 1, while trace 2 has 77,470 values of time samples (417.642 s) [11]. We have divided both traces into 10% of samples as the training set and a number of different test sets randomly selected from the remaining 90% of both traces. Each test set has an equal number of samples. Figure 8a shows the RMSE for the initial test set of trace 1 that is as small as 12 ms. Whereas, Figure 8b shows the RMSE value for trace 2 is 10 ms for the first random test set.
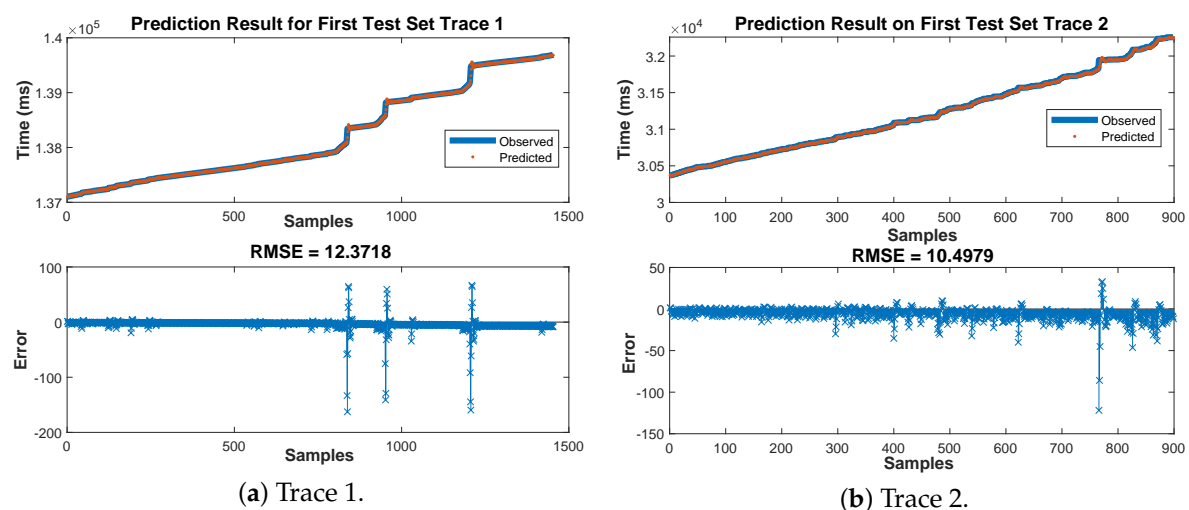

**Figure 8.** Prediction result and root mean square error (RMSE) for first test set.

The RMSE values for random test sets from the remaining 90% samples of trace 1 and trace 2 are shown in Figure 9. It can be seen from Figure 9a,c that the minimum value of RMSE is least at 6 ms and 5 ms on random test sets from samples 70,500 to 71,953 and 90,700 to 92,153 of trace 1, respectively. Whereas, Figure 9b,d show the RMSE value of 6 ms and 8 ms from 10,000 to 10,899 and 65,000 to 65,899 samples of trace 2, respectively.

AI-DRX can be implemented at the gNB of 5G network. During the execution of AI-DRX, we consider packet generation event, active event, dynamic long sleep event, dynamic short sleep event, ON duration event, beam searching event and feedback event. Moreover, our approach enables dynamic sleep cycles in multiple beam communications scenario for 5G networks. The packet generation event in our simulation scans the time column of both traces and produces the data packets of the identical length on the same instant in the respective trace. The generated data packets are collected in the buffer and served to the UE during the active event. UE checks the buffer during ON period, in case of any packet in buffer, the UE switches to the active event, or else continues to sleep. During an active event, the buffered packets are served to UE after getting the beam pairs alignment between UE and gNB. At the same time, packet arrival time is inputed to the trained LSTM model to predict the upcoming packet arrival time. We can obtain the dynamic sleep duration by subtracting the previous packet arrival time value from the predicted dynamic time value T_{DY} (upcoming packet arrival time).
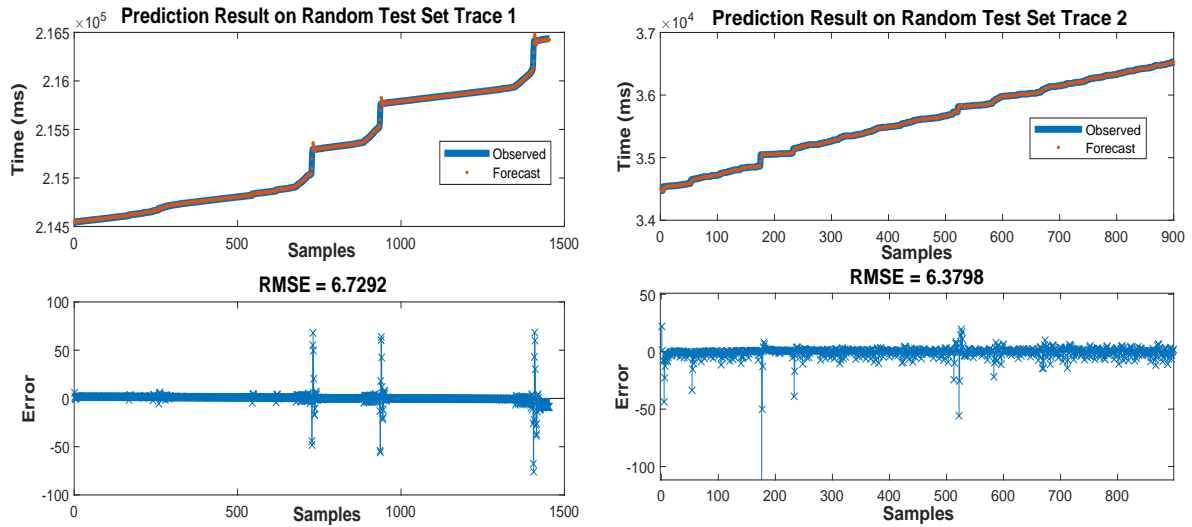
(**a**) Test set from trace 1 (samples: 70,500 to 71,953).

(**b**) Test set from trace 2 (samples: 10,000 to 10,899).
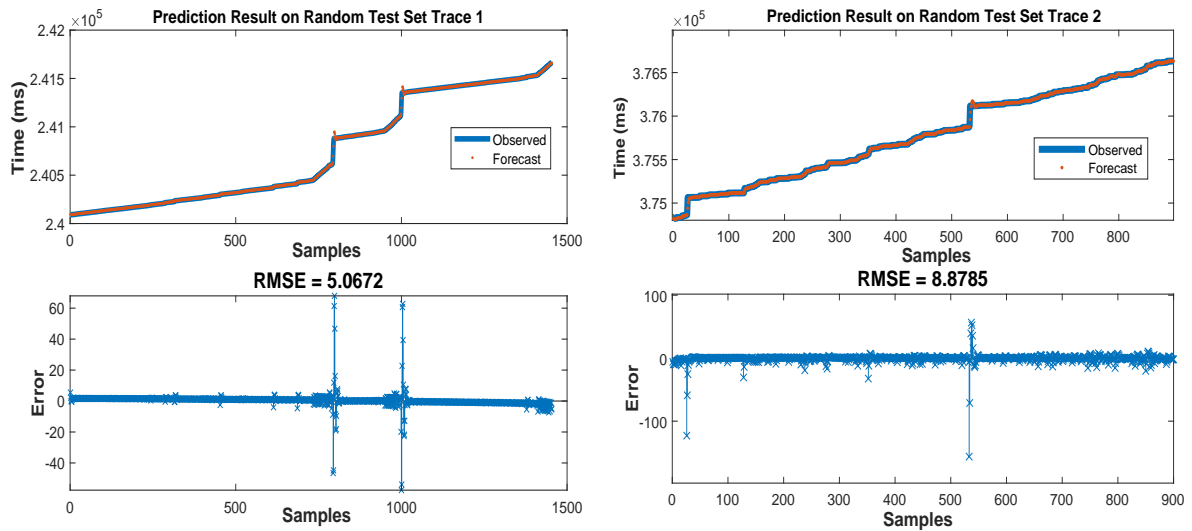
(**c**) Test set from trace 1 (samples: 90,700 to 92,153).

(**d**) Test set from trace 2 (samples: 65,000 to 65,899).

Figure 9. Prediction results and RMSE for random test sets.

Figure 10a,b, demonstrate the energy efficiency and mean delay with varying ON Period (T_{ON}) for AI-DRX and LTE-DRX over trace 1 & trace 2, respectively. In Figure 10, T_{ON} is varied from 1 ms to 160 ms while the values for $Th_{Min} = 20$ & $Th_{Max} = 100$ are considered. Figure 10a highlights the decrease in energy efficiency with increase in T_{ON}. The reason for the drop in energy efficiency is, UE waits for longer period in ON state prior switching to active state (to serve packets). It can be seen from Figure 10b, the mean delay for trace 1 ranges from 150 ms to 194 ms with an increase in ON period. The reason for rise in mean delay lies in the feet that with an increase in ON duration, the time spend by UE in ON state will be higher, which results in higher delay. Moreover, by selecting the optimum value of T_{ON}, delay observed by UE can be minimized. From Figure 10a, it is noticed that the energy efficiency of trace 2 is higher than that of trace 1 due to higher arrival rate of trace 1 as compared to trace 2.

We have compared the performance of AI-DRX with LTE-DRX. To implement the LTE-DRX, we select the value of short sleep cycle to Th_{Min}, the value of long sleep cycle to Th_{Max} and fed trained model with trace 1 and trace 2. It is observed in Figure 10a AI-DRX energy efficiency for trace 1 is 69% higher than that of LTE-DRX, at the cost of higher delay. The reason arises from the fact that AI-DRX

calculates sleep time form real wireless traffic trace based on arrival rate, while LTE-DRX uses Th_{Min} and Th_{Max} to set the short and long sleep time. For AI-DRX, the small value of Th_{Min} and Th_{Max} achieves higher energy efficiency as UE easily transits to long sleep time, which results in a larger delay. Whereas, for small values of Th_{Min} and Th_{Max} in LTE-DRX, a UE sleeps for a short period that results in less energy efficiency and smaller mean delay. AI-DRX achieves 55% higher energy efficiency as compared to that of LTE-DRX using trace 2.
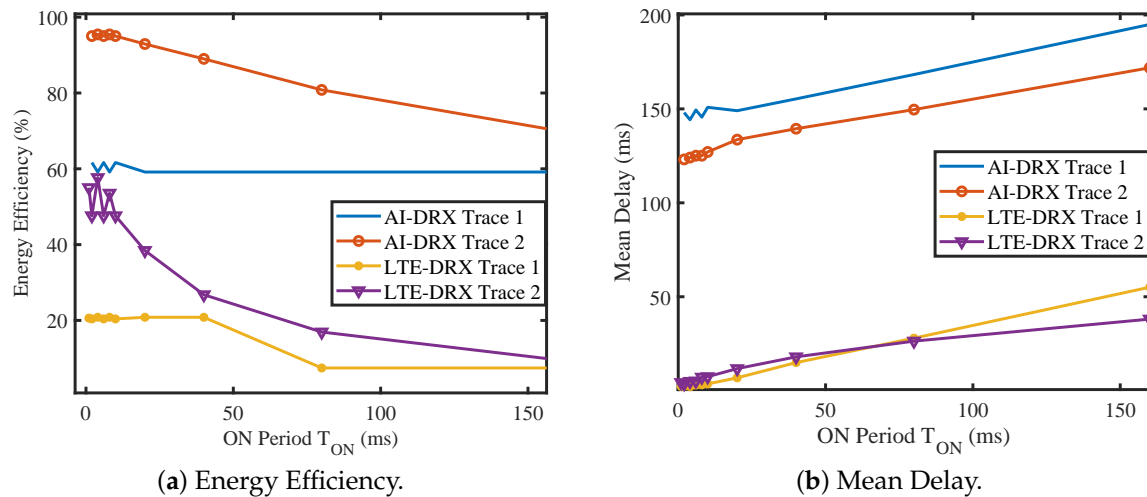

(**a**) Energy Efficiency. (**b**) Mean Delay.

Figure 10. AI-DRX energy efficiency and mean delay with varying T_{ON} ($Th_{Min} = 20, Th_{Max} = 100$).

Figures 11 and 12 present the energy efficiency and mean delay with varying ON period T_{ON} for $Th_{Min} = 200$ & $Th_{Max} = 1000$ and $Th_{Min} = 300$ & $Th_{Max} = 1600$, respectively. It is observed from Figures 11 and 12 that the energy efficiency and mean delay of LTE-DRX increase with an increase in Th_{Min} & Th_{Max}, as UE sleep longer. The energy efficiency of AI-DRX decreases with an increase in Th_{Min} & Th_{Max}, as UE will not able to transit to sleep state if $T_{DY} < Th_{Min}$. With a small value of sleep time T_{DY} or higher value of Th_{Min} & Th_{Max}, UE remains in the active state, which results in less energy efficiency.
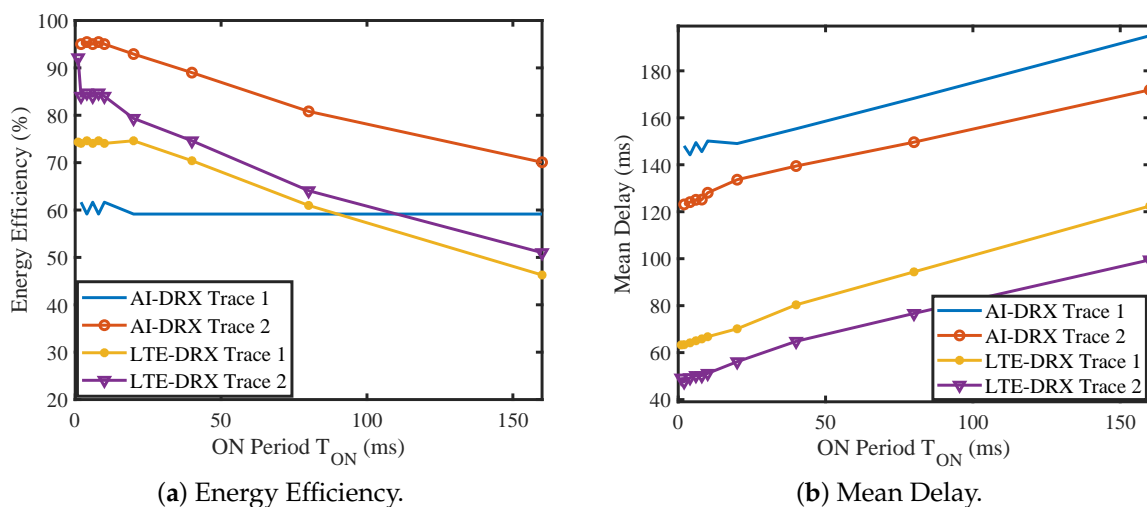

(**a**) Energy Efficiency. (**b**) Mean Delay.

Figure 11. AI-DRX energy efficiency and mean delay with varying T_{ON} ($Th_{Min} = 200, Th_{Max} = 1000$).

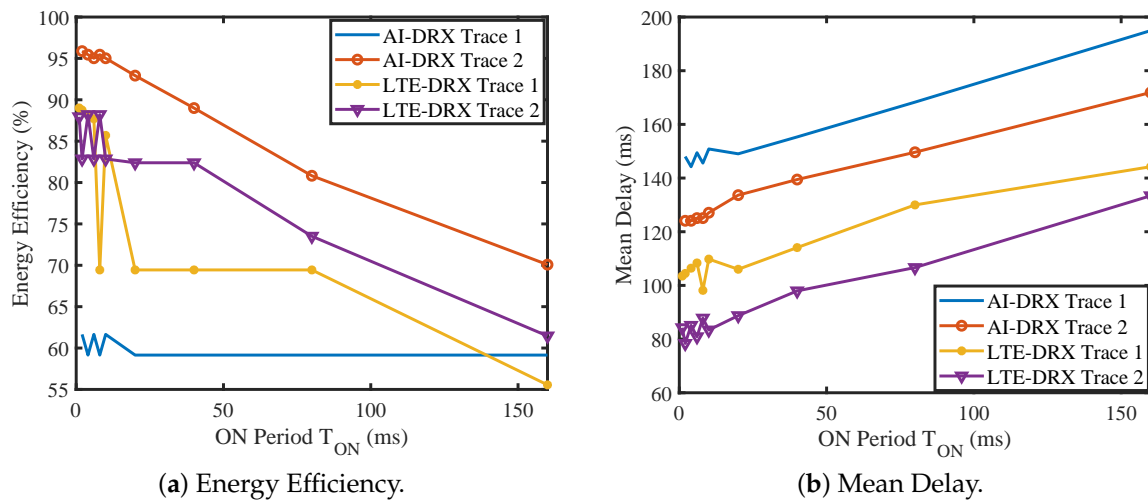**(a)** Energy Efficiency.　　　　　　　　　　**(b)** Mean Delay.

Figure 12. AI-DRX energy efficiency and mean delay with varying T_{ON} ($Th_{Min} = 300, Th_{Max} = 1600$).

To validate our proposal, we have compared our work to traditional Poisson arrival model. We generated three data sets considering Poisson arrival rate (λ) with mean value of $\lambda = 1/20$, $\lambda = 1/10$, and $\lambda = 1$. We have trained model using Poisson arrival rate. The generated traces are fed to AI-DRX algorithm to analyze the energy efficiency and mean delay. Figure 13a,b shows the energy efficiency and mean delay for AI-DRX trace 2 and Poisson arrival with varying T_{ON}. The energy efficiency of AI-DRX is 70% higher than Poisson arrival rate ($\lambda = 1/20$). The corresponding mean delay of AI-DRX is 100 ms (on an average) higher as compared to Poisson arrival for $\lambda = 1/20$. The gain in energy efficiency is achieved as AI-DRX considers real traffic arrival rate for selection of sleep cycles and inactivity timer, while Poisson arrival considers mean arrival rate. The energy efficiency and mean delay both are zero for higher Poisson arrival rate $\lambda = 1$. The reason lies in the fact that for higher arrival rate, UE could not transit to sleep state to save the power.
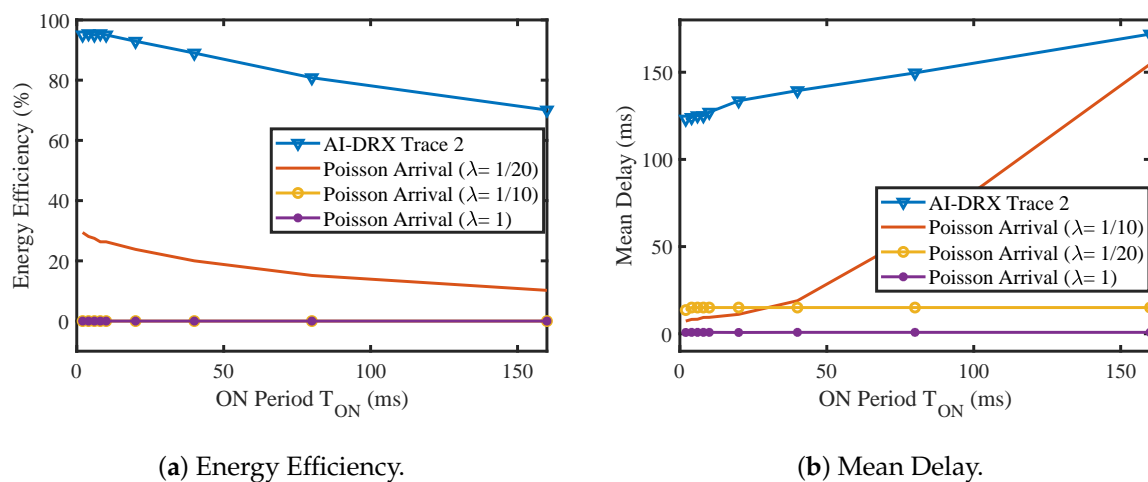

(a) Energy Efficiency.　　　　　　　　　　**(b)** Mean Delay.

Figure 13. AI-DRX energy efficiency and mean delay comparison with Poisson Arrival with varying T_{ON} ($Th_{Min} = 300, Th_{Max} = 1600$).

Various services in a wireless network can tolerate different delay levels while not compromising quality of service (QoS). QoS class identifier (QCI) is a metric that is used to identify the characteristics of traffic. QCI measures the QoS with two parameters; (1) packet loss rate (PLR) and (2) packet delay budget (PDB). PDB can be defined as maximum tolerable waiting time by a packet during its delivery from eNB to UE. Standardized QCI characteristics are shown in Table 2 [6,30,31]. In

various kinds of non-real-time services like email, web browsing, after a certain time period, a UE does not require to monitor PDCCH continuously [6,32]. Hence, these types of services can tolerate higher delays up to 300 ms [6,30,31]. These types of services require DRX with higher values of sleep time ($T_{DY} > Th_{Min}$ & $T_{DY} > Th_{Max}$) and smaller values of ON timer for better energy efficiency. Whereas, the real-time services like voice and live video streaming cannot tolerate delay [33]. Therefore, the delay should be a higher priority than energy saving. If we limit the mean delay to 125 ms, the energy efficiency will be 95.02%. For this energy efficiency we have to select $Th_{Min} = 200$, $Th_{Max} = 1000$, $T_{ON} = 6$ ms. If we increase the T_{ON} the energy efficiency of UE decreases as UE remains in active state. The mean delay also increases with increase in T_{ON} as UE does not receive the data during ON period but only monitors the Physical Downlink Control Channel (PDCCH). The mean delay observed by UE increases to 150 ms when $T_{ON} = 80$, with energy efficiency of 80.2%. The network can maximize the energy efficiency of UE by selecting optimum value of Th_{Min}, Th_{Max}, T_{ON} depending on QCI value of different services.

Table 2. Standardized QoS class identifier (QCI) characteristics in LTE/LTE-A [6,30,31].

Type	Qos Class Identifier	Packet Loss Rate	Packet Delay Budget	Examples
GBR	1	10^{-2}	100 ms	Voice services
GBR	2	10^{-3}	150 ms	Live streaming services
GBR	3	10^{-3}	50 ms	Real time gaming services
GBR	4	10^{-6}	300 ms	Buffered streaming services
Non-GBR	5	10^{-6}	100 ms	IMS Signaling services
Non-GBR	6	10^{-6}	300 ms	TCP based application services
Non-GBR	7	10^{-3}	100 ms	Interactive gaming services
Non-GBR	8	10^{-3}	300 ms	TCP based video services
Non-GBR	9	10^{-6}	300 ms	TCP based video services

5. Conclusions

In this work, we have suggested an AI-based DRX mechanism for energy saving in multiple beams communications scenario. We have modeled DRX as a ten-state model and suggested AI-DRX algorithm depending on these 10 states. AI-DRX algorithm enables dynamic short and long sleep cycles for energy efficiency of UE in the 5G network. We have trained LSTM network, a popular type of RNN, to extract the packet arrival time pattern from real wireless traffic traces. Later, we have utilized the learned model in AI-DRX algorithm for energy saving in 5G enabled devices. AI-DRX economizes power consumption of a UE by enabling dynamic short and long sleep cycles. Extensive training with selected hyper-parameters achieves the least RMSE of 5 ms on a random test set from trace 1 and 6 ms on a random test set from trace 2, respectively. The energy efficiency obtained with AI-DRX is approximately 60% and 95% for trace 1 and trace 2, respectively. AI-DRX achieves 69% higher energy efficiency on trace 1 and 55% more energy efficiency on trace 2 as compared to LTE-DRX, respectively. We also validated the performance of AI-DRX with traditional Poisson packet arrival model. AI-DRX attains 70% more energy efficiency on trace 2 as compared to Poisson packet arrival rate for $\lambda = 1/20$.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ericsson Mobility Report June 2019. Available online: https://www.ericsson.com/assets/local/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf (accessed on 12 June 2019).
2. Fowler, S.; Shahidullah, A.O.; Osman, M.; Karlsson, J.M.; Yuan, D. Analytical evaluation of extended DRX with additional active cycles for light traffic. *Comput. Netw.* **2015**, *77*, 90–102. [CrossRef]
3. Maheshwari, M.K.; Agiwal, M.; Saxena, N.; Roy, A. Directional Discontinuous Reception (DDRX) for mmWave enabled 5G Communications. *IEEE Trans. Mobile Comput.* **2018**. [CrossRef]
4. Agiwal, M.; Roy, A.; Saxena, N. Next generation 5G wireless networks: A comprehensive surve. *IEEE Commun. Surv. Tut.* **2016**, *18*, 1617–1655. [CrossRef]
5. Zhou, K.; Nikaein, N.; Spyropoulos, T. LTE/LTE-A discontinuous reception modeling for machine type communications. *IEEE Wireless Commun. Lett.* **2013**, *2*, 102–105. [CrossRef]
6. Maheshwari, M.K.; Roy, A.; Saxena, N. DRX over LAA-LTE—A New Design and Analysis based on Semi-Markov Model. *IEEE Trans. Mob. Comput.* **2018**, *18*, 276–289. [CrossRef]
7. Ferng, H.-W.; Tseng, Y.-T.; Wang, T.-H. Parameter-Adaptive Dynamic and Adjustable DRX Schemes for LTE/LTE-A. *IEEE Trans. Green Commun. Netw.* **2019**. [CrossRef]
8. Gódor, I. Revitalizing DRX for enhanced mobile sleep modes. *Comput. Commun.* **2019**, *138*, 98–105. [CrossRef]
9. 3GPP TSG-RAN WG2 #94 Tdoc R2-164023 Std., 2016. Available online: https://www.3gpp.org/ftp/tsg_ran/WG2_RL2/TSGR2_94/Docs/R2-164023.zip (accessed on 4 April 2019).
10. Ho, C.H.; Huang, A.; Hsieh, H.-J.; Wei, H.-Y. Energy-Efficient Millimeter-Wave M2M 5G Systems with Beam-Aware DRX Mechanism. In Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, ON, Canada, 24–27 September 2017.
11. Mudasar, L.M.; Maheshwari, M.K.; Shin, D.R.; Roy, A. Deep-DRX: A framework for deep learning–based discontinuous reception in 5G wireless networks. *Trans. Emerg. Telecommun. Technol.* **2019**, *30*, e3579.
12. Kwon, S.-W.; Hwang, J.; Agiwal, A.; Kang, H. Performance analysis of DRX mechanism considering analogue beamforming in millimeter-wave mobile broadband system. In Proceedings of the 2014 IEEE Globecom Workshops (GC Wkshps), Austin, TX, USA, 8–12 December 2014.
13. Liu, D.; Wang, C.; Rasmussen, L.K. Discontinuous Reception for Multiple-Beam Communication. *IEEE Access* **2019**, *7*, 46931–46946. [CrossRef]
14. Wang, J.; Tang, J.; Xu, Z.; Wang, Y.; Xue, G.; Zhang, X.; Yang, D. Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. In Proceedings of the IEEE INFOCOM 2017—IEEE Conference on Computer Communications, Atlanta, GA, USA, 1–4 May 2017.
15. Gers, A.F.; Schmidhuber, J.; Cummins, F.A. Learning to forget: Continual prediction with LSTM. *Neural Comput.* **1999**, *12*, 850–855.
16. Google. Google AI Blog. Available online: https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html (accessed on 1 June 2019).
17. Bontu, C.S.; Illidge, E. DRX mechanism for power saving in LTE. *IEEE Commun. Mag.* **2009**, *47*, 48–55. [CrossRef]
18. Evolved universal terrestrial radio access (E-UTRA); Medium access control (MAC) protocol specification. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2437 (accessed on 2 March 2018)
19. Chang, H.-L.; Tsai, M.-H. Optimistic DRX for Machine-Type Communications in LTE-A Network. *IEEE Access* **2018**, *6*, 9887–9897. [CrossRef]
20. Chang, H.-L.; Lu, S.-L.; Chuang, T.-H.; Lin, C.-Y.; Tsai, M.-H.; Sou, S.-I.Optimistic DRX for machine-type communications. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia , 22–27 May 2016 .
21. Agiwal, M.; Maheshwari, M.K.; Saxena, N.; Roy, A. Directional-DRX for 5G wireless communications. *Electron. Lett.* **2016**, *52*, 1816–1818. [CrossRef]
22. Maheshwari, M.K.; Agiwal M.; Saxena, N.; Roy, A. Hybrid directional discontinuous reception (HD-DRX) for 5G communication. *IEEE Commun Lett.* **2017**, *21*, 1421–1424. [CrossRef]

23. Maheshwari, M.K.; Roy, A.; Saxena, N. Analytical modeling of DRX with flexible TTI for 5G communications. *Trans. Emerg. Telecommun. Technol.* **2018**, *29*, e3275. [CrossRef]

24. Study on new radio access technology Physical layer aspects. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3066 (accessed on 2 March 2018)

25. Study on new radio access technology Radio interface protocol aspects. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3070 (accessed on 2 March 2018).

26. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Learning Internal Representations by Error Propagation*; Technical Report; Institute for Cognitive Science, University of California: San Diego, CA, USA, 1985.

27. UMass Trace Repository. ACM MMSys conference dataset archive. Available online: http://traces.cs.umass.edu/index.php/Mmsys/Mmsys (accessed on2 March 2018).

28. CRAWDAD Data Set Cambridge/Inmotion. Available online: https://crawdad.org/cambridge/inmotion/20051001 (accessed on 2 March 2018).

29. Baughman, M.; Haas, C.; Wolski, R.; Foster, I.; Chard, K. Predicting Amazon spot prices with LSTM networks. In Proceedings of the 9th Workshop on Scientific Cloud Computing, Tempe, AZ, USA, 11 June 2018.

30. Policy and charging control architecture. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=810 (accessed on 2 March 2018).

31. Liang, J.M.; Chen, J.-J.; Cheng, H.-H.; Tseng, Y.-C. An energy-efficient sleep scheduling with QoS consideration in 3GPP LTE-advanced networks for internet of things. *IEEE J. Emerg. Sel. Top. Circ. Syst.* **2013**, *3*, 13–22. [CrossRef]

32. Tseng, C.-C.; Wang, H.-C.; Kuo, F.-C.; Ting, K.-C.; Chen, H.-H.; Chen, G.-Y. Delay and power consumption in LTE/LTE-A DRX mechanism with mixed short and long cycles. *IEEE Trans. Veh. Technol.* **2015**, *65*, 1721–1734. [CrossRef]

33. Huang, B.; Tian, H.; Chen, L.; Zhu, J. DRX-aware scheduling method for delay-sensitive traffic. *IEEE Commun. Lett.* **2010**, *14*, 1113–1115.