International Journal of Applied Sciences: Current and Future Research Trends
(IJASCFRT)
ISSN: 2790-3990
https://ijascfrtjournal.isrra.org/index.php/Applied_Sciences_Journal

Optimizing Energy-Efficient 5G Resource Allocation for Machine-Type Communication through Reinforcement Learning

Anum Ali*
Charles V. Schaefer, Jr. School of Engineering and Science,Hoboken, Newyork, US
Email:aali17@stevens.edu

Abstract

With the proliferation of machine-type communication (MTC) devices in 5G networks, there is a growing need to optimize resource allocation to ensure energy efficiency and network performance. The advent of 5G networks has opened up opportunities for massive connectivity of MTC devices in various application domains, such as Internet of Things (IoT) sensors and devices. However, the resource allocation for MTC poses significant challenges due to the dynamic and diverse traffic patterns generated by a large number of devices. Traditional resource allocation methods may not be suitable for handling MTC's unique energy efficiency and scalability requirements. This re- search paper proposes a novel approach using reinforcement learning to optimize energy-efficient resource allocation for MTC in 5G networks. This paper presents a reinforcement learning- based approach to optimize resource allocation in 5G networks for MTC. The proposed approach utilizes reinforcement learning algorithms to learn and adapt resource allocation policies based on real-time network conditions and device demands. By employing a reward-based mechanism, the system maximizes energy efficiency while meeting the quality of service (QoS) requirements of MTC devices. The experimental evaluations demonstrate the effectiveness of the proposed approach in optimizing energy- efficient resource allocation for MTC in 5G networks. The results show significant improvements in energy consumption and network performance compared to traditional resource allocation methods. The reinforcement learning-based approach adapts to varying traffic conditions, effectively balancing resource allocation and minimizing energy waste. The contributions of this research include the development of a novel framework for energy-efficient resource allocation in 5G networks, specifically tailored for MTC. The proposed approach enables intelligent decision-making and adaptation based on real-time network dynamics by leveraging reinforcement learning techniques. The optimization of resource allocation for MTC devices in 5G networks enhances energy efficiency, scalability, and overall network performance.

Keywords: Machine-to-Machine; Energy Efficiency; Re- source Allocation; Reinforcement learning; 5G.

Received: 8/1/2024
Accepted: 10/1/2024
Published: 10/10/2024

----

* *Corresponding author.*

1.Introduction

Mobile communication has been evolved to the Internet of Things (IoT) architecture that not only support analog voice but various other hundreds of services [1]. Machine- type communication (MTC) network is a stable connected network that is in its development phase and aims to provide access to data for millions of users. Therefore, future of wireless communication is of providing services as well as connectivity to users [2]. These services include monitoring systems, health-care, household and many other applications. Due to these benefits, MTC network have a separate set of characteristics and requirements. This network has become a market changing factor for nearly every aspect of life whether it is health care, logistics, traffic control or security [3]. Most of the IoT services are based on monitoring and control without human intervention. For connectivity, IoT uses low-power short-range wireless technologies or cellular networks. MTC smart devices are tightly coupled in cellular infrastructure [4]. Such as tracking of vehicles and providing fleet management or monitoring of power usage such as smart metering. The 5G network is the future of wireless communication that includes LTE-M, WiMAX, GPRS [5]. 5G offers new modulation tech- niques that are much efficient and simpler but still need to be implemented and tested on the front.

The popularity of IoT network lies in the fact that in the whole ecosystem it provides something to each module from sensors to network operators and further on to enterprises. For connectivity, the best option is cellular network due to its large range coverage and economical network operations. But due to unique characteristics of IoT communication, there exist many challenges such as radio resource allocation and scheduling, transmission on physical layer and random access procedures [6], etc. Other important issues are to reduce access delay and saving energy consumption [7] in simultaneous access to the channel by multiple autonomous devices. Interference and other synchronization problems should be handled by the network load sharing capability. In 5G networks, downlink and uplink channels are partitioned into a number of subchannels which are made of time and frequency resources, termed as radio resource blocks (RBs), each RB can be allocated to the devices upon the access requests [9]. A sensible re- source allocation algorithm also plays a fair role in solving interference and energy consumption issues. Uplink traffic is more in use than downlink traffic[8], therefore resource is a critical task in the cellular network. More attention is required on the design of energy efficient resource allocation while making the eNB nodes of the 5G network intelligent. Resource allocation also depends on the quality of experience (QoE) and quality of service (QoS) power profiles. It is better to differentiate between QoE and QoS, furthermore to adopt QoE as criteria for assessing the quality evaluation. First, QoS is applied to handle technical factors regarding service and related power usage which does not include any kind of human usage quality-affecting factors. In this regard, the same QoS level might not give surety that the same QoE level might is be affecting two different users. If not considering the system’s technical properties, other factors such as the user-specific characteristics, the context of use, and the pricing of a service impact significantly on the final perceived QoE as well.

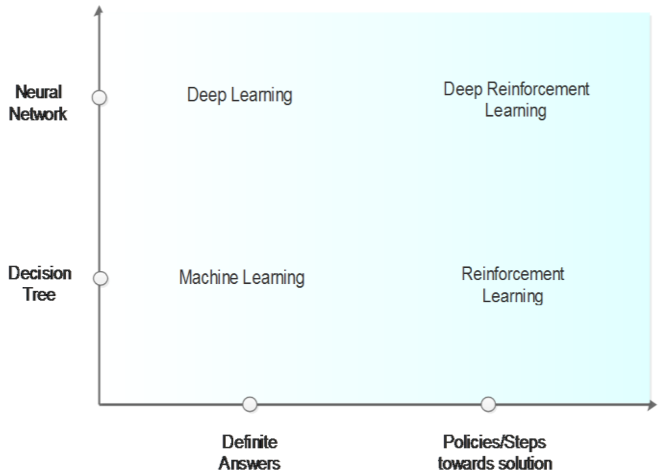

Figure 1: Division between reinforcement learning

1.1. Reinforcement learning in IoT

Reinforcement learning is a technique in which the assigned agent picks up the environment and learn to achieve the desired results [10,12]. Fig.1 shows division between methodologies. In IoT network, most of the devices act without any human in- tervention therefore necessary autonomous techniques should be included for efficient resource allocation. It is derived from machine learning [11]. Mainly the research is changed as the delayed caused due to resource allocation effects the potential and suitability of data analytics and machine learning in the network management, context of services, and systems integration. Furtheron it will provide better and deeper understanding on decision making based on the available operational, largely collected and serviced data. This will increase the rates of avaliable opportunities for improving data analytics algorithms and machine learning methods on aspects such as dependability, reliability, and scalability, as well as explainations of the benefits of these methods in control and management systems. In addition to that, there is an vast area of opportunity to define decisive platforms that can effects the varity of advanced data analysis algorithms and operational data to drive management decisions in data centers, networks, and clouds. By increasing the reinforcement learning rate, the agent sets the reward and accordingly finds a good strategy. It is based on tried-and-error interactions in a dynamic environment. Reinforcement learning can be used for efficient resource allocation that can save device energy and time. Previously also many techniques are used in handling the massive number of devices in IoT infrastructure. Each IoT application has different traffic requirements and patterns. Usually, it is delay-tolerant and time-controlled traffic that may be periodic or synchronized traffic. Wireless resource sharing is a difficult task that can cause a delay while wasting the energy of the MTC devices.

The main contribution of the paper is that, it introduces a reinforcement learning algorithm for IoT communication network using 5G wireless environment, that tragets to in- crease energy efficiency based on IoT periodic communication structure and reduce delay caused due to resource allocation of spectrum carriers. A quality of experience and quality of service power metrics were used accordingly with current available applications power demands. For more feasible re- source allocation reinforcement learning was simulated that selects appropriate threshold on which it makes the selection of subcarriers to individual devices those were basically sensors type devices. The proposed algorithm is intended to be less complex and reduces time taken for allocation that also saves energy consumption.

1.2. Outline

In this paper, we have proposed an energy efficient resource allocation technique while implying reinforcement learning in 5G network. Previously there have been many techniques for resource allocation but they are based on LTE networks using older modulation techniques. Furthermore usage of reinforcement learning till now is limited, the latest related researches are discussed in section II. This gap is fulfilled and simulated the results on Matlab. The remainder of this paper is organized as follows. Section II discusses the related work followed by section III that elaborates the system model and problem formulation. Section IV presents the proposed algorithm. Results are presented in section V and in the end conclusion is presented in section VI.

2.Related work

In paper [13], the authors have proposed a joint massive access control and resource allocation strategy. The algorithm performs device grouping. A coordinator is chosen that selects the group and another coordinator is chosen that perform resource allocation. This coordinator also defines the number of groups under certain transmission schemes such as 2-hop transmission protocol. By grouping, they aim to reduce energy consumption in frequency-selective and flat-fading channels. The algorithm claimed to achieve suboptimal results regarding energy consumption. In article [14] a joint scheduling and power allocation algorithm is proposed and demonstrated. It is based for orthogonal frequency division multiplexing (OFDM) for multiple users. The authors target the stability of device queues and claim to achieve energy efficiency. The queue of the devices is broken in sub-queues and apply the master theorem. The channel models used are frequency-selective and flat fading. Further on utility functions are applied on low and high demands for resource and traffic load. Results are shown as sum-throughput of the system. In literature [15], the authors try to formulate the problem in mixed integer programming (MIP) and apply canonical duality theory. The algorithm is compared with previous techniques in literature such as the best channel quality indicator (BCQI) and round robin (RR). For maximizing energy efficiency results, invasive weed optimization (IWO) technique is applied. The author claimed to achieve better results while fulfilling QoS require- ments. In the thesis [16], resource management techniques are reviewed for machine-to-machine (M2M) devices. In litera- ture, a multi-objective optimization solution is proposed for resource allocation. In algorithm, the devices are clustered and within a cluster, a cluster head is chosen through which the devices communicate. The cluster head acts as a controller and maximizes the throughput. Further on the Q-learning algorithm is used for slot selection in the random access network. The results were dependent on the set of rewards and learning rate. The results are compared with channel- based and ALOHA algorithms. The results show a better performance. through Q-learning. In article [17], the authors propose a dynamic resource allocation technique using LTE- A network for device-to-device communication. For channel allocation, reinforcement learning is used. The outband and inband resources are assigned by eNB nodes to devices pairs. It eliminates a central control link between cellular radio and unlicensed interfaces. eNB node handles the allocation through adjusting the learning rate. The allocation of outband resources is formulated as a dynamic single-player game. The eNB estimates proper joint utility functions based on reinforcement learning (JUSTE-RL). The simulation shows near-optimal results.

With available solutions in literature, there still exists a number of issues such as complex algorithm implementation and delay induced due to large computations. For effective resource allocation, the probability of service should not add delay and complexity, which have been ignored. This paper fills this gap while testing on 5G which is the future of wireless communication. Moreover, the resource allocation not only adjusted along with QoS metric but also for 5G environment QoE metrics plays an important part which is handled in this paper.

3.System Model and Problem Formulation

IoT communication is considered in the 5G network. The network consists of the massive number of devices within some macro cell [18]. The base station is required to handle multiple demands from different devices using different application profiles. Due to different requirements, an algorithm is proposed that allocate resources according to device power limits adjusted in QoS and QoE metrics. Since IoT communication should be autonomous therefore for result enhancement reinforcement learning is applied, that accordingly select an appropriate threshold that controls the selection of resources. The 5G network is considered to be a heterogeneous wireless network, in which LTE, WiMAX, and GSM can co-exist, the model is shown in Fig. 2. For modulation, FBMC technique is used. There are many waveforms introduced for 5G network including FBMC. But FBMC has gained alot of popularity in 5G network [20]. It is a physical layer concept. For the data link layer, Open Wireless architecture [21] is used in the 5G network model. The resource blocks are distributed in both time and frequency domains. The bandwidth selected from 27 to 29 GHz is divided into subcarriers spaced with 75KHz frequency. All the used variables are listed in the Table I.

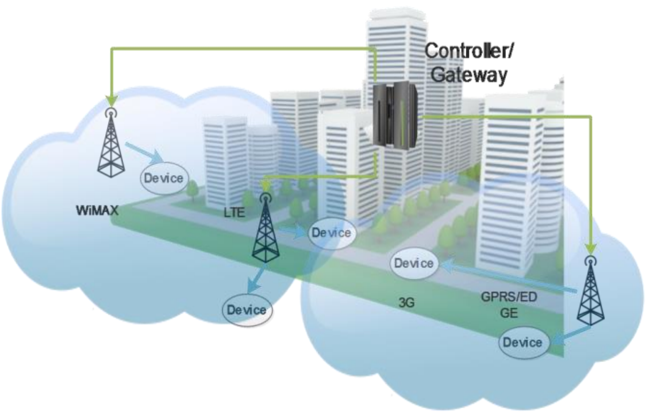

Figure 2: 5G system model

Table 1: Notation s

Symbol	Definition
m_i	M2M i^{th} device
p_i	device power
P_{total}	Total summed power
G_i	channel gain
N_o	channel noise
C_i	shared channel available for i^{th} MTC device
$M[i]$	transmit powers from M2M devices.
$RB[n]$	matrix of resource block
θ	maximum power usage limit
$Psig_{m,k}$	reference signal power
f	number of subcarriers within k RB's
P_{bs}	base station power
$P_{mi,k}$	consumed power by a device
Pl	pathloss
$\phi_{mi,k}$	Signal-to-Noise Ratio (SNR)
γ^2	power of Additive White Gaussian Noise (AWGN)
$h_{mi,k}$	channel fading amplitude
d_{mi}	data rate achieved by ith device
B	effective bandwidth
λ_i	calculated number of carriers for ith device
ϑ_i	calculated power by estimation for ith device
α	computed number of carriers
β	total number of available carriers.
ω	threshold
J_o	constant depends on antenna characteristic
δ	path loss constant
Ψ	Rayleigh random variable

3.1. Problem Description

The problem that has been targeted in this paper is to maximize the energy efficiency during resource allocation while keeping the throughput at maximum in IoT scenario. The total number of available channels can be expressed as C, where C is the total number of channels. The channel capacity for shared channel available for i^{th} MTC device can be expressed as follows.

$$C_i = log\left(1 + \frac{p_i}{N_o}|G_i|^2\right), \forall i = 1, 2, 3, ...I \qquad (1)$$

Here Gi is the channel gain and No is the channel noise. The total throughput available can be expressed as following.

$$C_{total} = \sum_{i=1} C_i \quad (2)$$

Each channel consists of resource blocks (RB), each RB can be expressed as RB[n] in matrix form, where n is the highest total number of subcarriers. Each IoT device (mi) reference signal power $P_{sig_{m,k}}$ [22] can be computed as following equation while using k resource blocks.

$$P_{sig_{m_i,k}} = P_{bs} - 10\log_{10} k \times f \quad (3)$$

Here fk denotes number of subcarriers within k resource blocks and Pbs is the base station power. The consumed power by a device expressed as Pmi,k is further computed by subtracting pathloss Pl from reference signal power Psigm,k in (4).

$$P_{m,k} = P_l - P_{sig} \quad (4)$$

Signal-to-Noise Ratio (SNR) denoted as $\phi_{m_i,k}$ [23] experienced by a MTC device (m_i) is expressed as follows while using k resource blocks.

$$\phi_{m_i,k} = \frac{P_{m,k_i} \times |h_{m,k_i}|^2}{\gamma^2} \quad (5)$$

In equation (5), γ2 is power of Additive White Gaussian Noise (AWGN), Pm ,k presents consumed power by the IoT device (I mi) and hmi,k is channel fading amplitude. The algorithm is based on two phases. In the first phase numbers of sub-carriers are computed based on the QoE and QoS data rate is expressed as dmi [24] that can be available is computed by the following equation. Here total RB assigned to a device is expressed as k, ϕmi,k is effective SNR and B denotes effective bandwidth. Total power Ptotal usage for all devices can be computed as following.

$$d_{m_i} = B \times k \times \log_2 (1 + \phi_{m_i,k}) \quad (6)$$

$$P_{total} = \sum_{i=1}^{M} \sum_{k=1}^{N} P_{m,k} \quad (7)$$

The purpose is to minimize the summed power P_{total} during resource allocation and selecting appropriate RB's to IoT devices. The energy efficiency (EE) can be computed by equation (8).

$$EE = \frac{dm_i}{P_{m_i,k}} \quad (8)$$

Here dmi denotes achieved data rate and Pmi,k is the con- sumed power by the IoT device (mi) while been allocated with k resource blocks.

3.2. Interference Model

A machine-type multicast service model (MtMS) [19] has opted as an interference model. This model defines the procedures for transmission for handling device-to-device (D2D) communication using multicast traffic. The model consists of the following architecture. The session is initiated by MtMS serving center (MtMS-SC). This is the main source working of MtMS. The anchor point in this architecture is a service capability center (SCS), which trigger sessions to send and receive data to and from the MTC devices. This architecture provides an option for devices to connect with a suitable group of devices that are waiting for service. Through tracking area information that is provided to MtMS coordination entity (MtMS-CE) when the device joins any group. The tracking information is related to joining devices that will be paged through the control interface. MtMS-CE handles the group joining procedure. MtMS gateway (MtMS-GW) handles data delivery after joining the group in certain established MtMS sessions through balanced resource allocation having related transmission properties. On enhancing the procedure, it can be divided into subgroups and devices are paged according to subgroup properties. The time interval of paging and size of subgroup depends on available resources of radio interfaces.

3.3. Energy Efficient Resource Allocation Using Reinforcement Learning

The purpose of this paper is to efficiently allocate resources to achieve maximum energy efficiency. The proposed RTA algorithm is based on two phases. In the first phase numbers of sub-carriers are computed based on the QoE and QoS, power profile metrics. The QoS metric is created according to application power requirements. The algorithm is further optimized in the second phase by including threshold in the algorithm. The value of the threshold is estimated through reinforcement learning that makes the algorithm autonomous with the environment.

Table 2: MAJOR QOE FACTORS

Service	Aspect	Quality Influence Factor
Independent	Mobile Networks	Connectivity loss, delays, power consumption, handovers
	MAC layer	Packet loss, buffer overhead
	Phy layer	Low throughput,low capacity, bandwidth bottleneck
Depedent	Video/Voice	Coding rate, packet lost, noise level, service provider
	Uplink service	Limited power,small data
	Download service	Web browsing, file download

3.4. Proposed Energy Efficient Resource Allocation

In the first phase of the proposed RTA algorithm each device convey its minimum transmit power level through uplink physical channels [25]. According to the power limit, the number of subcarriers are computed that will be assigned to a particular device. The highest power limit is set which is the maximum power usage limit, expressed as θ. This value can be extracted from the QoS metric. This metric is based on the power requirement of different applications. The received power limit of a particular device is compared with the highest power limit θ. The algorithm is expressed mathematically below.

$$\Delta = 1 - \frac{M[i]}{\vartheta} \tag{9}$$

$$\vartheta_i = PL - M[i] \times \frac{1}{\Delta} \tag{10}$$

$$a = \Theta_i - P_{BS} \tag{11}$$

$$\lambda_i = 10^a \tag{12}$$

Through equation (9), a difference Δ is calculated. By using Δ value number of carriers are estimated that should be assigned to the device. There lie two options for carrier estimation, if the power is less assuming it to be the power of small IoT device then more carriers are allocated on uplink thus giving more data rate on the uplink in limited power, otherwise, on downlink, fewer carriers are provided to save the receiving power of the device. The value of the ϑ_i and the number of carriers λ_i can be calculated from equations (10), (11) and (12) respectively. Here in equation (9) $M[i]$ is the transmitted

Table 3: QoS Power Metric

Domain	Highest Power Limit	Priority
Health care	0 - 8 db	High
Surveying	0 - 10 db	Low
Control	5 - 9 db	High
Enterprise	2 - 15 db	Medium

power of ith device, ϑ_i is the calculated power by estimation and Δ is the calculated difference through equation (9). λ_i is the estimated number of carriers in the equation (12) that will be assigned to the device. In case of downlink, the equation (10) will be updated to equation (13).

$$\vartheta_i = PL - M[i] + \Delta \quad (13)$$

The achievable maximum data rate for each device on a subcarrier can be expressed as following

Algorithm 1 Achievable maximum data rate

1: Initialize b;

2: $b = 10^a$

3: **if** $b < 1$ **then**

4: $b = 100 \times b$

5: $\alpha = b \times \beta$ **return** α

Here a is computed from 11, α is the computed number of carriers and β are the total number of available carriers. The above algorithm can be adjusted according to the requirement of the domain. Since each IoT application has different priority levels and relative device power profiles, therefore, QoE metric is shown in Table II and a QoS metric both are required for listing the highest power limit θ value used in equation (9). The QoS metric is shown in Table III. Each resource block is precomputed in groups/cluster according to the defined ranges of power they consume. A set of groups can be defined and resource blocks are assigned to each group according to power range.

3.5. Proposing Optimization by Reinforcement learning in algorithm

In a typical system, resource allocation scheme does not cater to maximize energy efficiency which is the major requirement of the IoT network. In the second phase includes optimization of RTA algorithm in which a threshold is induced that is chosen from a reinforcement learning technique. Reinforcement learning techniques let the agents adopt optimal policy through learning and exploring the environment. The response towards optimal policy also depends on the environment and the learning rate. In our algorithm, the chosen an optimal policy depends on optimal value that is the reward.

1) Stochastic Model: In this section, the heuristic proposed system is described which contain modules and policies which are heuristically derived to define the technique that proposes energy conservation. The policy is although tested through simulation. The informally first system model is illustrated in Fig. 3 then definitions and properties of model are analyzed. The system consists of five modules. This system consists of a single source of resources, which entertain requests. The requests are passed through the second module that is to set the threshold. The computation of the threshold is based on re- wards. Rewards are generated through reinforcement learning algorithm 2. Each state is based on rate characterization. The rate is the number of requests per unit time. And the states which have 0 rates are idle states. Requests and services are termed as stochastic processes. The interarrival time is non-deterministic between requests. By this nature, it is shown that a delay can be caused in the system. For choosing the best threshold depends on criteria which is called policy.

2) MDP Model: In this section, we define the MDP model by five elements: decision policy, actions, states, reward function, and transition probability. Following elements are specified.

- Decision Epoch: In our model, the moment of decision is when a device wakeup and sends a request to the network.

- Action: When there is an available resource there can be two possibilities of action for the requesting device during each decision epoch: analysis of power usage and computing resource to allocate. When there is no resource left, the only option will to put the device in wait. Therefore action set function is based on power usage and the number of remaining resources. A(r) = power analysis, compute resources if r $\geq$ 0 allocate if r

=0 then A(r) = wait, reject

- States: The system states are defined as 1) an idle state,

2) resource request state,3) waiting state and 4) resource granted state.

Reinforcement learning main components are states, actions, and rewards. Here is the states those include resource request, resource allocated and waiting states, as are the actions i-e device power analysis and setting of proper threshold and r for reward, its they achieved energy efficiency based on the power consumption by the device. Algorithm 2 summarizes the reinforcement learning based threshold selection algorithm performed for resource allocation to each requested device. In Line 1, the action and state variables are initialized. In Line 5, the policy is set according to the environment such as SNR values, etc. And according to the set policy the actions are chosen, that can be either to analysis power limit or set proper threshold. Rewards are calculated based on the analyzed power and allocated resources. The threshold is updated based on the computed reward in Line 6. Algorithm shown in 2. It allows the system to choose the best threshold value that enhances the resource allocation to different devices having different power limitations according to the environment.

Algorithm 2 Heuristic threshold selection by reinforcement learning algorithm

---

1: Initialize $R(s, a)$ arbitrarily; 2: Repeat (for each allocation), 3: Initialize S

4: Repeat (for each step in allocation);

5: Choose a from s using set policy (e.g greedy, soft, soft- max)

6: Observe reward r i-e energy efficiency

7: **if** $s' < s$ **then**

8: Take action a i-e set threshold (ω);

9: **else**

10: increase threshold (ω);

11: $R(s, a) = R(s, a) + \aleph[r + \gamma_{max\aleph}\ v\ , R(s', a') - R(s, a)]$

12: s = s'

13: **return** ω

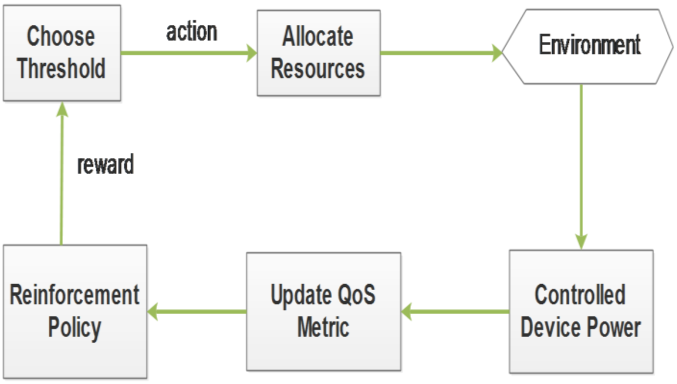

Figure 3: Algorithm flow

The threshold is induced in the mathematical model, before equation 10 equation 14 is computed, which is the following,

$$\Delta = \omega \times \Delta \quad (14)$$

MTC devices are able to choose resources in a self-organizing manner. According to the power limits of an MTC device, the threshold adjusts the allocation of subcarriers in an overlapping area within the presence of multiple devices that can maximize its QoS performance along with the human experience usage through QoE metric. In this case, the eNB node will observe, learn, and adapt independently to the selection decision. The reinforcement learning algorithm is implemented to address the resource allocation selection criteria and specifically used as an environment adjusting technique for selecting the best threshold to distribute MTC devices among the available resources. With the proposed threshold selection algorithm,, MTC devices have the ability to switch to the optimized allocation of subcarriers that provides better energy efficiency while giving smaller delay.

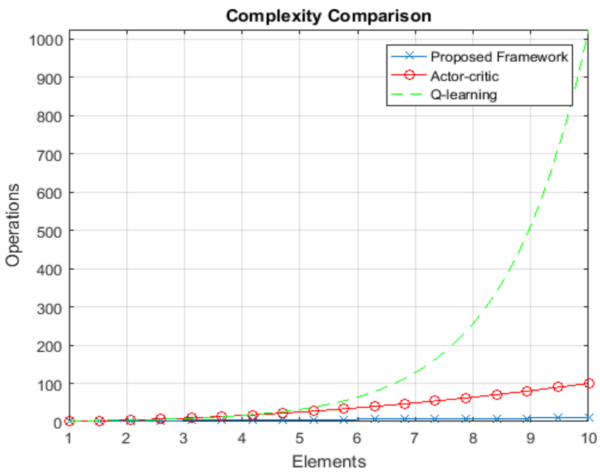

Figure 4: Complexity comparison

3.6. Experimentation

The simulation is tested on the maximal number of IoT devices 500 and simulated in Matlab. In all the simulations, for creating a 5G environment the channels between source and destination have independent distribution. Parameter values for simulation are listed in Table I. Channel gain Gi is modeled as following

$$G_i = \Psi \times J_o \left(\frac{d_o}{d}\right)^{\delta} \qquad (15)$$

where Jo is a constant that depends on the antenna charac- teristic and average channel attenuation, do is the reference distance for the antenna far field, d is the distance between transmitter and receiver, δ is the path loss constant and Ψ is the Rayleigh random variable. Since this formula is not valid in the near field, in all the simulation results, we assume that d is greater than do. In all the results, do = 10m, Jo = 50 and δ = 2. The bandwidth used for simulation is between 27 GHz to 29 GHz. Subcarrier spacing is set to 75 KHz. Rayleigh channel [26] amplitude is set to 7db, 30 is the channel pathloss.

3.7. Complexity of Algorithm

The complexity of this algorithm is in two parts. Part one is related to calculation carriers. Since its linear calcu- lation, therefore, the complexity becomes O(n). Further ahead performing optimization step through reinforcement learning increases complexity. Most of the domains studied in the context of reinforcement learning that consists of additional properties that decrease the complexity of the algorithm in the worst case scenario. The state space complexity has linear upper action limit for all n. Then the complexity of worst case becomes.

$$O(cn^2) = O(n^2) \qquad (16)$$

Here c is the number of states and n is the actions on devices.

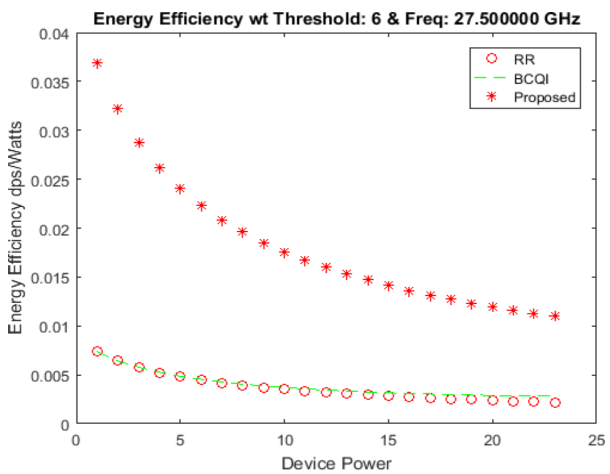

Figure 5: Energy efficiency verse device power

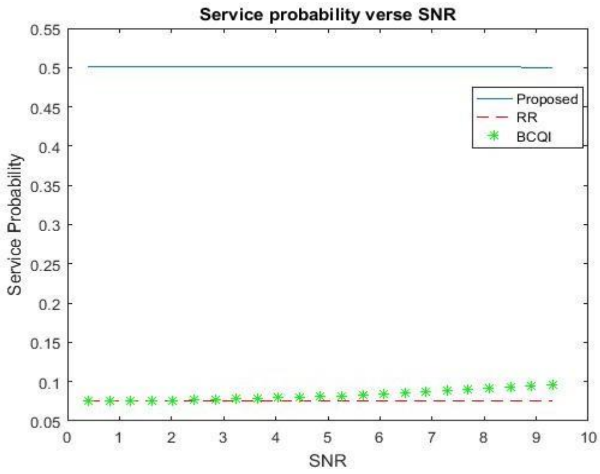

Figure 6: Service probability verse snr

related to resource allocation. The states iterates over action steps and the algorithm consist of two action steps while counting through entire sweep on state space. It matches the convergence test which checks the largest achievable reward r. The complexity comparison graph is shown in Fig. 4.

3.8. Results and Discussion

The results of the simulation are compared with existing resource allocation techniques such as the best channel quality indicator (BCQI) and round robin (RR) algorithms. The pro- posed solution achieves 40 percent better energy efficiency for small powered devices such as monitoring sensors for health- related applications. The results reach to 20 percent better energy efficiency for the rest of the IoT devices, as shown in Fig. 5. It can be seen that at threshold 6, it supports energy optimization to high power devices as well. The Fig. 6 illustrate service probability of the proposed solution. The result shows 50 percent better service probability when compared to RR and BCQI techniques for all types of powered IoT devices. The service probability is steady throughout the different powers of devices. Health-related IoT devices are very prone to restricted.

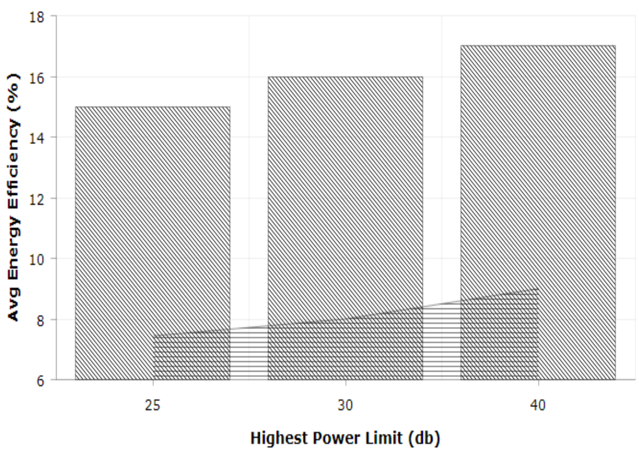

Figure 7: Increased EE graph

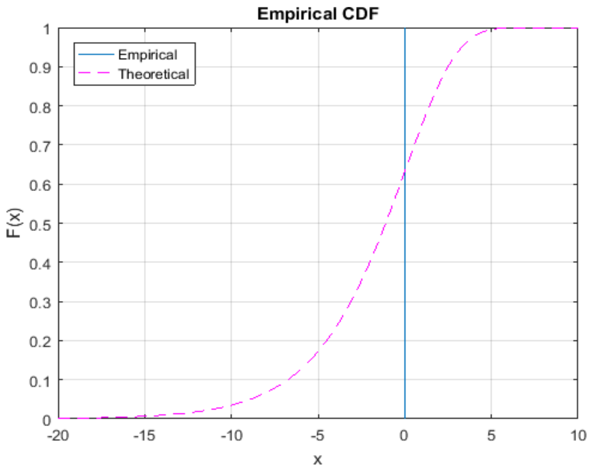

Figure 8: CDF plot

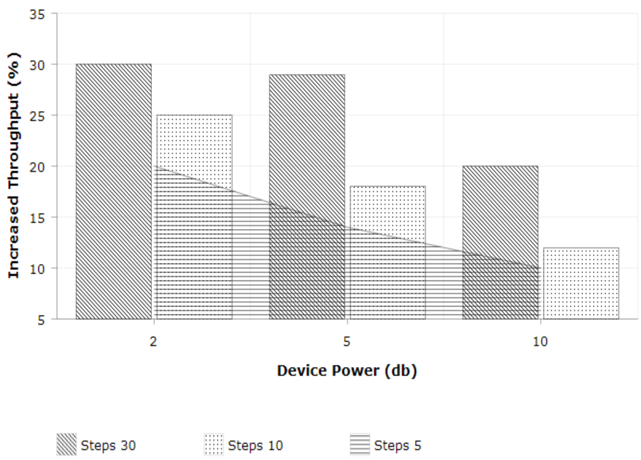

Figure 9: Throughput graph\

power and send data in small chunks periodically. For this type of devices especially the proposed solution shows better throughput while using a less complex algorithm and without any time delay. The Fig. 7 shows increased energy efficiency reaching 20 % for devices belonging to classes having the highest power limit of 40 dB, power profile class having highest power limit to 30 dB achieves 16 % increased energy efficiency whereas highest power limit of 25 dB class reaching to 15 % increased energy efficiency. In Fig. 8, a cumulative distribution function (CDF) plot is shown of performance of the proposed RTA algorithm. The achieved throughput can be illustrated in Fig. 9, where priority is given to low powered devices for their emergency data messaging. The low power devices that are from 2 dB to 5 dB devices achieve about 30% increased throughput. Whereas 10 dB devices reach to 20 % increased throughput. The graph shows the difference between increased steps (recursion) in the reinforcement algorithm.

3.10. Conclusion

This paper introduces an RTA algorithm for IoT communication network using 5G wireless environment, to increase energy efficiency and reduce delay caused due to resource al- location. A QoE and QoS power metric were used accordingly with in-demand applications power profiles. For more optimal resource allocation reinforcement learning was implemented that selects appropriate threshold which controls the selection of subcarriers to individual devices. The proposed algorithm is less complex and reduces time delay that also saves energy consumption. The algorithm shows better results in simulation after comparing with previous popular literature solutions.

References

[1].Chen, S., Ma, R., Chen, H. H., Zhang, H., Meng, W., & Liu, J. (2017). ”Machine-to-machine communications in ultra-dense networks,A survey”. IEEE Communications Surveys & Tutorials, 19(3), 1478-1503.

[2].Shariatmadari, H., Ratasuk, R., Iraji, S., Laya, A., Taleb, T., Jantti, R., & Ghosh, A. (2015). ”Machine-type communications: current status and future perspectives toward 5G systems”. IEEE Communications Magazine, 53(9), 10-17.

[3].Anum Ali, Ghalib A. Shah, Muhammad Omer Farooq, Usman Ghani, ”Technologies and challenges in developing Machine-to-Machine ap- plications: A survey,“ Journal of Network and Computer Applica- tions, Volume 83, 1 April 2017, Pages 124-139, ISSN 1084-8045, http://dx.doi.org/10.1016/j.jnca.2017.02.002.

4.Ericsson White paper, ”Cellular networks for massive IoT“, Jan 2016

[2]. Alonso-Zarate, Jesus, and Mischa Dohler. ”M2M Communications in 5G.“, 5G Mobile Communications. Springer International Publishing, 2017. 361-379.

[3]. Song, Qipeng, Loutfi Nuaymi, and Xavier Lagrange. ”Survey of radio resource management issues and proposals for energy-efficient cellular networks that will cover billions of machines.“ EURASIP journal on wireless communications and networking 2016.1 (2016): 140.

[4]. Anum Ali, Ghalib A. Shah, Junaid Arshad, ”Energy efficient techniques for M2M communication: A survey,“ Journal of Network and Computer Applications, Volume 68, June 2016, Pages 42-55, ISSN 1084-8045, http://dx.doi.org/10.1016/j.jnca.2016.04.002.

[5]. H. Yao, S. Ma, J. Wang, P. Zhang, C. Jiang and S. Guo, ”A Continuous-Decision Virtual Network Embedding Scheme Relying on Reinforcement Learning,” in IEEE Transactions on Network and Ser- vice Management, vol. 17, no. 2, pp. 864-875, June 2020, doi: 10.1109/TNSM.2020.2971543.

[6]. Song, Qipeng, Loutfi Nuaymi, and Xavier Lagrange. ”Survey of radio resource management issues and proposals for energy-efficient cellular networks that will cover billions of machines.“, EURASIP journal on wireless communications and networking 2016.1 (2016): 140.

[7]. Moon, Jihun, and Yujin Lim. ”A Reinforcement Learning Approach to Access Management in Wireless Cellular Networks.“ Wireless Commu- nications and Mobile Computing 2017 (2017).

[8]. Zincir-Heywood, N., Casale, G., Carrera, D., Chen, L. Y., Dhamdhere, A., Inoue, T., ... Samak, T. (2020). Guest editorial: Special section on data analytics and machine learning for network and service

management–parti. IEEE Transactions on Network and Service Management, 17(4), 1971- 1974.

[9]. Gosavi, Abhijit. "A tutorial for reinforcement learning.", Department of Engineering Management and Systems Engineering (2011).

[10]. Ho, Chieh Yuan, and Ching-Yao Huang. "Energy-saving massive access control and resource allocation schemes for M2M communications in OFDMA cellular networks." IEEE Wireless Communications Letters 1.3 (2012): 209-212.

[11]. Sahand Golnarian, "Energy-efficient and queue-aware resource alloca- tion in uplink OFDM systems for wireless M2M communication.", thesis submitted to University of Notre Dame, 2014.

[12]. A. Aijaz, M. Tshangini, M. R. Nakhai, X. Chu and A. H. Aghvami, "Energy-Efficient Uplink Resource Allocation in LTE Networks With M2M/H2H Co-Existence Under Statistical QoS Guarantees," in IEEE Transactions on Communications, vol. 62, no. 7, pp. 2353-2365, July 2014. doi: 10.1109/TCOMM.2014.2328338

[13]. Fatima Hussain, "Resource allocation in clustered M2M networks: A Q-learning approach", thesis submitted to Ryerson University, 2016.

[14]. Asheralieva, Alia, and Yoshikazu Miyanaga. "Dynamic Resource Allo- cation with Integrated Reinforcement Learning for a D2D-Enabled LTE-A Network with Access to Unlicensed Band." Mobile Information Systems 2016 (2016).

[15]. Bockelmann, C., Pratas, N., Nikopour, H., Au, K., Svensson, T., Ste- fanovic, C.,& Dekorsy, A. (2016). "Massive machine-type communica- tions in 5G: Physical and MAC-layer solutions". IEEE Communications Magazine, 54(9), 59-65.

[16]. M. Condoluci, G. Araniti, T. Mahmoodi and M. Dohler, "Enabling the IoT Machine Age With 5G: Machine-Type Multicast Services for Innovative Real-Time Applications," in IEEE Access, vol. 4, no. , pp. 5555-5569, 2016. doi: 10.1109/ACCESS.2016.2573678

[17]. F. Schaich and T. Wild, "Waveform contenders for 5G - OFDM vs. FBMC vs. UFMC," 2014 6th International Symposium on Communica- tions, Control and Signal Processing (ISCCSP), Athens, 2014, pp. 457-460. doi: 10.1109/ISCCSP.2014.6877912

[18]. 5G PPP Architecture Working Group. "View on 5G architecture." White Paper, July (2016).

[19]. Leonhard Korowajczuk, "LTE Measurements: What they mean and how they are used", CelPlan International, Inc.

[20]. Stanczak, Slawomir, Marcin Wiczanowski, and Holger Boche. "Fun- damentals of resource allocation in wireless networks: theory and algo- rithms". Vol. 3. Springer Science & Business Media, 2009.

[21]. Rath Vannithamby, Shilpa Talwar, "Towards 5G: Applications, Require- ments and Candidate Technologies", John Wiley & Sons, 30-Jan-2017, Technology & Engineering, 472 pages

[22]. Verizon 5G TF, "Air Interface Working Group; Verizon 5th Genera- tion Radio Access; Physical channels and modulation (Release 1)", TS V5G.211 V1.3 , June 2016

[23]. A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHZ and 38 GHZ millimeter-wave bands," in IEEE Communications Magazine, vol. 52, no. 9, pp. 78-86, September 2014. doi: 10.1109/MCOM.2014.6894456

Dr Anum Ali (IEEE member since 2009) received her bachelors degree from the Sir Syed University of Engineering and Technology, Pakistan, in 2004, and M.Sc. degree from the University of Lancaster, U.K., in 2006. She got her Ph.D. degree from the University of Engineering and Technology, Lahore in 2020. She got experience of twenty four years in academia, research and development. She was an Assistant Professor in Department of Software Engineering of Lahore Garrison University, Pakistan. Currently associated with Stevens Institute of Technology, USA. Her research interests include in space and satellite communication modules and wireless sensors communication including M2M and V2X evolving towards autonomous systems for Cyber Security.