iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: https://doi.org/10.22214/ijraset.2022.45161

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

Intelligent-Driven Green Resource Allocation for Industrial Internet of Things In 5g Heterogeneous Network

Shweta Ray[1], Mr. Manish Gupta[2]
[1]*M.Tech Student, Department of Computer Science and Engineering, Buddha Institute of Technology, Gida, Gorkhapur, India*
[2]*Assistant Professor, Department of Computer Science and Engineering Buddha Institute of Technology, Gida, Gorkhapur, India*

Abstract: ***Mobile networks' energy consumption is rising in tandem with the volume of traffic and the number of people utilising mobile technology. To maintain the long-term survival of the next generation of mobile networks, there must be an emphasis on energy efficiency. By minimising the network's power consumption and proposing an energy-efficient network architecture, this thesis addresses the issue of increasing 5G and beyond network efficiency. The first component of this thesis focuses on base stations (BSs), the most energy-intensive part of mobile networks. Mobile network providers offer us with a data set that contains information on the amount of traffic on their system. The poor temporal granularity of mobile network traffic data makes it difficult to train ML systems for sleep mode management choices. Bursty arrivals are taken into consideration while generating mobile network traffic statistics.***

For the purpose of deciding when and how deeply to put BSs to sleep, we offer ML-based methods. There is no assurance of service quality in the present literature on using ML to network management. In order to address this problem, we apply a combination of analytical model-based techniques and machine learning (ML). Risk is quantified using an innovative measure that we have developed. For the purpose of assessing and monitoring the success of machine learning algorithms, we create a digital twin that can behave just like a real BS, including with sophisticated sleep modes. Using the suggested approaches, significant energy savings may be achieved at the expense of a small number of delayed users, according to simulation findings. A cloud native network architecture based on virtualized cloud RAN, which forms the foundations of open RAN, is studied and modelled in the second half of the thesis.

The open RAN design based on hybrid C-RAN, which is the subject of this thesis, was recently agreed upon by major telecommunications providers. The transition to hybrid C-RAN network topologies from typical dispersed RAN designs is energy and cost intensive. With economic feasibility evaluations of a virtualized cloud native architecture taking future traffic estimates into account, we calculate migration costs in terms of both OPEX and CAPEX. However, the infrastructure expenses of fronthaul and fibre links are not immediately apparent when comparing C-RAN with D RAN systems. ILP-based fronthaul optimization minimises migration costs by optimising the fronthaul. We provide an artificial intelligence-based heuristic strategy for dealing with large problems. Dealing with the balance between energy consumption and latency is a significant difficulty in network design and management. To lower the network's energy consumption while simultaneously improving latency, we devised an ILP problem as part of a hybrid C-RAN architecture with several layers. As an added consideration, we weigh network bandwidth use against overall power consumption. It is possible to demonstrate that intelligent content placement not only reduces latency, but also conserves energy. We seek to reduce network energy consumption by constructing logical networks that are tailored to the needs of each services. According to the scientific literature, network slicing has reduced total energy consumption. Only radio access network resources are considered in most studies. Network energy consumption is reduced when bandwidth is increased when the RAN is taken into consideration. This thesis describes a novel model for cloud and fronthaul energy consumption that indicates that increasing bandwidth allocation also increases processing energy consumption in cloud and fronthaul. While the network's energy consumption is lowered, the quality of service (QoS) of slices is maintained using a non-convex optimization problem. Second-order cone programming is used to find an optimal solution to the problem. According to our research, end-to-end network slicing may lower total network energy usage compared to radio access network slicing.

Keywords: 6G, 5G, Energy efficiency, Machine learning, Reinforcement learning, Network architecture, Sleep modes, Mobile networks.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

I. INTRODUCTION

The ICT industry is rapidly advancing toward 5G and beyond networks to fulfil the ever-increasing need for higher data rates and to deliver a broad variety of services with different quality-of-service (QoS). It is a huge issue for future technologies, like as 5G, to implement and manage all of these requirements. 5G networks should also be able to support services that need more bandwidth and lower latency. Increased data traffic and technological advancements necessitate higher investment expenditures and a rise in energy consumption [4]. Network architecture and management techniques must be reevaluated to take full use of better signal processing systems' ability to analyse large volumes of data. As the number of nodes and the scale of modern networks increase, the management of these systems becomes increasingly difficult. AI and ML technologies have the potential to revolutionise network administration, but they are still in their infancy. As a result of the additional difficulties produced and the increasing complexity of 5G and future networks, traditional solutions are no longer suitable.

Aiming to reduce the energy consumption of future mobile networks, we use AI and ML-based techniques after studying energy and delay-aware resource allocation in mobile networks [12–25]. We highlight the potential, limitations, and unsolved difficulties of energy saving at different network levels.. Our focus is on reducing base station power consumption, environmentally friendly network topologies, and approaches for reducing consumption while maintaining the quality of service (QoS) we have promised. While conducting research on base station energy conservation, this thesis will focus on the following: (BSs). This part is focused on minimising the BSs' power consumption while while ensuring optimal performance. Research on base station energy savings features is being conducted in order to provide a framework for assessing and mitigating risk. Since energy reductions are typically accompanied by a drop in performance in the network, this is a critical consideration. Thus, it is essential to keep track of and maintain the lowest possible performance decline.

This is where we examine the whole network architecture and seek for outstanding issues in green network architecture design. It is our contention that the existing network architecture is not capable of supporting a wide range of services, and that we must move on to a new network architecture. There are several features of new network designs that we study, including the cost of implementing them, how many services can be supported, and how network resources might be tailored for certain services. End-to-end network design is the emphasis of this portion of the thesis, where we estimate the costs of moving to new green architectures, and suggest collaborative network resource allocation to improve user and service quality while reducing energy usage. The recommended solutions are tested numerically, and the results show that the offered solutions are effective[1].

II. RELATED WORK

A. Green Networks

Both fixed (44–46) and wireless networks (47–49) have been widely studied in the last decade in terms of green networks. More and more research in both sectors has focused on reducing network activity, lowering network dependability, or employing power management to test several network performance criteria.. According to studies [51], BSs are responsible for a considerable amount of the energy used by the whole network. Reducing BS energy consumption must be done without compromising the QoS clients get, such as delays in data transmission [25, 52], if these networks are to remain operationally viable. The components of the BS transceiver chain may be divided into four groups based on their activation and deactivation times, as depicted in [10]. The use of sleep modes puts a number of devices to sleep all at once for a specified period of time (SMs). If you use SM, you may save BS power at the cost of certain service interruptions. During the day, when network traffic is often more busy, delays and drops may be more evident.

To prevent a reduction in service quality, proper use of the SM is required (QoS). When looking at cell DTX in heterogeneous networks, scientists in [53] show that BS density and traffic load impact the possibility for energy savings. The purpose of the study in [54], which takes into consideration long-term, deep SM, is load-adaptive SM management. [55] provides an analysis of the impact of 4G BS sleep on QoS metrics like dropping/delay. The ON/OFF state of small-cell mm-wave antennas is evaluated [56] based on Tokyo traffic data.

An ASM-based stochastic model is provided in [57] for modifying the BSs configuration. Machine learning (ML) has lately been a prominent tool for managing BS sleep modes (SMs) activities. As opposed to saving energy, the implementation of such strategies would likely result in service interruptions for end users. The quality of service (QoS) supplied to end users must not be negatively affected by the usage of ASMs. Studies at BSs have shown that ML in ASMs saves a lot of energy [58–63]. The authors in [58–60] examine four SMs in depth using a preset (de)activation sequence. They decide how long each SM will be active for, and it is their choice. In [58], a heuristic method for implementing ASMs is offered.

©IJRASET: All Rights are Reserved

The energy-saving potential of ASMs is influenced by the synchronisation signalling frequency. In [59], a Q-learning-based method for determining the optimum SM is presented. In this case, they're using a Q-learning algorithm to create a codebook that relates traffic load to certain actions.

The authors of these studies assume a preset sequence of activities, which may not result in the best possible savings in electricity. [61] suggests an algorithm for picking SMs that takes traffic into consideration. In order to train their network, however, they use a planned traffic pattern.

For deeper SMs, such as SM4, or for longer SM durations, [62] recommends that the control and data planes of 5G BSs be separated. Control signalling may limit the energy savings of ASMs. Where basic coverage cells transmit all periodic control signals, but capacity-enhancing ones have a higher ability to sleep and conserve energy, this separation may also help cocoverage scenarios. The term "states" is used here to distinguish between "high" and "low" loads for the purposes of [63]. While the study focuses on tabular techniques, it only examines two load levels, making it unable to reveal long- and short-term temporal correlations.

B. Network Architecture Design

Adding additional features and services to mobile networks means consuming more energy, but it also means utilising more of the network's limited resources. When it comes to energy consumption and capacity, it is presently incapable of handling a wide range of services in a dispersed network design. It is thus necessary to consider a range of elements, such as energy consumption and cost as well as operation and maintenance time delays as well as handovers. When the number of services on a network grows, so does the amount of energy it consumes.

There's been a lot of buzz recently about C-RAN technology. With C-RAN, all digital baseband computing and cooling is centralised at chosen cluster centres, leaving behind just RF and analogue processing at the cell locations for cost and energy consumption reduction. "" The capacity to offer coordinated multipoint (CoMP) technology, which permits millisecond-scale inter-cell coordination, is another benefit of centralization, as is the possibility to increase resource efficiency and EE, especially for users at the cell's edge. Technology based on the CoMP protocol. Several advantages of C-RAN designs need to be addressed in the literature, as do certain disadvantages. All processing units are centralised, which might cause issues with connection capacity between radio devices and the central cloud during certain periods of the day. The media used to link radios to the cloud must be scalable, low-cost, and able to simultaneously transmit raw I/Q signals to the cloud. Time-sensitive applications are made more challenging due to the network's added transit latency. To keep its environmental effect to a minimum, the C-RAN architecture has developed throughout the years in an effort to provide as many services as feasible. The addition of an additional processing layer to the C-RAN architecture may increase performance by distributing processing responsibilities between the central cloud and an additional layer known as the edge cloud. Our hybrid C-RAN (H-CRAN) architecture, as proposed in SooGREEN, is an adaptive and semi-centralized system. Power consumption and midhaul bandwidth are both reduced by adjusting the physical layer network function divides in H-CRAN in conjunction with one another. This architecture expands on the previous C-RAN structure by splitting processing tasks between CC and EC functionally in three layers. The evolution from DRAN to H-CRAN is seen in Figure 1. Each radio unit in DRAN has a digital unit and is connected to the main network through a backhaul connection. In order to link all of the C-RAN digital units and RUs together, fronthaul connections are utilised. H-CRAN offers an extra layer of processing that may be distributed across RUs and DU pools (central cloud). There have been several efforts to develop a viable H-CRAN. H-three CRAN's layers are EC, CC, and cell layer. The tightly packed cell layer of RUs provides service to user equipment (UEs). They are intertwined. An EC gathers the data from a group of REs, which is typically used as an aggregator. It is widely accepted that MmWave lines are the principal mode of communication between the RUs and the ECS. In addition, the ECs use midhaul to transfer the aggregated data to the CC. Function processing may be performed on requested material in this architecture utilising DUs in the CC and ECs. These DUs can assist any connected RUs by pooling their computer resources. Midhaul connections may be reduced in bandwidth by partially processing cell-to-cell traffic at the EC while CC handles the rest. An EC has more units than a CC, making it less efficient at using energy. So the CC can save more energy because of the shared infrastructural resources. Reduce bandwidth and delay performance by spreading activities at the ECs, or take use of the natural power savings that are associated with centralising all operations at CC. Cloud computing may be used to save energy in H-CRAN by providing delay-sensitive services at the EC, while applications that are more tolerant of delays can be serviced at the CC. We further distribute the processing activities to meet the problem of capacity constraint in the fronthaul in light of the delay limits.

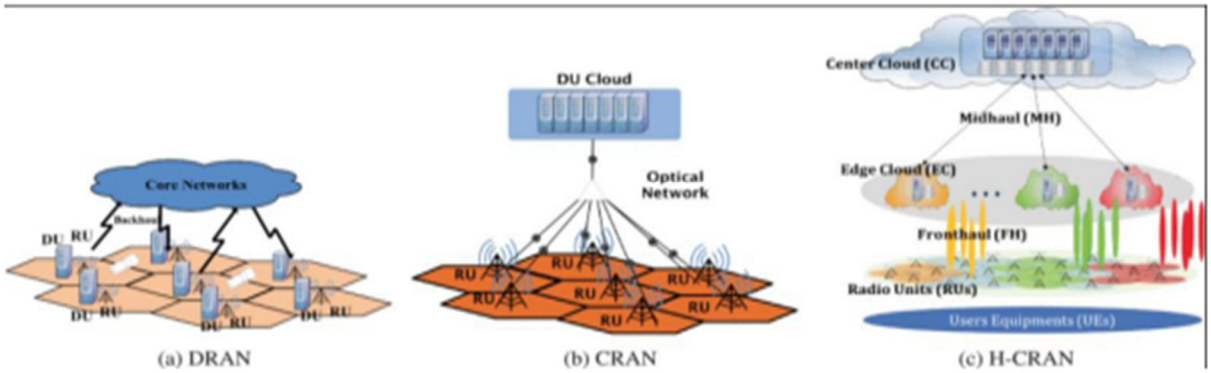


Figure 1: Evolution of RAN (DRAN/C-RAN/H-CRAN) [4].

1) *Discrete Functionality:* A number of challenges must be solved before the C-RAN idea can be put into practise in the real world. If all physical layer processing, including channel coding, modulation, and Fast Fourier Transform (FFT), is completely centralised, the bandwidth of the fronthaul signal may be considerably increased in comparison to its earlier DRAN architecture by several orders of magnitude. Because of this, the effect of such technologies as massive MIMO may be much more limited. It affects both the upstream and downstream of the supply chain. Fronthaul bandwidth must be increased for CoMP joint reception, for example, because of the necessity for improved signal resolution (finer quantization). An intermediate functional split between C-RAN and cell edge users is needed to retain the C-RAN scaling benefits while reducing bandwidth congestion and facilitating effective CoMP coordination. This is shown in Figure 2 to give you an idea of the functional split of the baseband processing chain for both cells and users. Baseband processing services include cell processing (CP) and user processing (UP). The CP, a component of the physical layer, houses the cell's signal processing functions. CP features include things like Serial to Parallel Encoding, FFT, Cyclic Prefix, and Resource Mapping. UP upper layer carries out signal processing for each user in a cell in the same way as physical layer does, and it includes a number of functions. Some examples of UP functions in operation include antenna mapping and forward error correction. Anywhere between Split 1 and Split 7 may be the point at which the split occurs, as shown by Fig. 1.7. At Split 1, all functions are centralised at the CC, resulting in C-RAN. When Split 7 occurs, all functions are consolidated at EC, resulting in DRAN. Functions above and below a split are denoted by CC and EC, respectively.

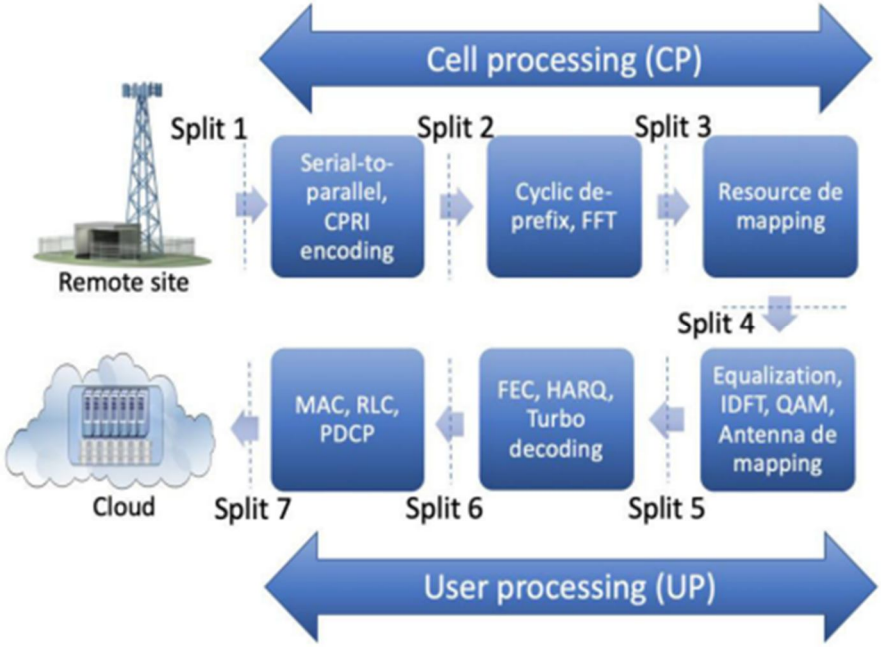


Figure 2: Functional split model [5].

2) *Network Architectures and Migration Costs:* Even though C-RAN has numerous advantages, it is difficult for operators to deploy. C-progress RAN's and challenges are discussed in detail in [64]. C-RAN may need 160 Gbps of throughput with a latency of 10 to 250 microseconds if all network operations are consolidated in the base band unit (BBU). These conditions are more liberal as long as lower physical layer (PHY) functions are kept near to the RUs. Lower-level PHY functions, such as split 4 in Figure 2 (about 933Mbps), have fronthaul transport connection needs that are only about 40% of the size of those needed in a fully centralised situation. Figure 2's baseband processing capabilities include cellprocessing (CP) and userprocessing (UP). These include signal processing and physical layer tasks such as serial to parallel encoding and resource mapping. UP is in charge of all per-user signal processing, including equalisation and turbo coding. Functional splitting in C-RAN [4] allows for the division of this processing chain. Depending on where the functional divide occurs, several types of fronthaul are required. When all functions are centralised in the cloud, as with conventional C-RAN, the greatest amount of energy savings and virtualization may be obtained. Alternative split sites supported by the most cost-effective fronthaul are preferable [66], however this may not be feasible due to the limited fronthaul capacity. For fronthaul connections, as illustrated in Figure 2, Table 1 summarises the capacity requirements, and Table 1 displays possible split positions [7, 66]. For fronthaul transmission, passive optical networks are the only viable option [67] since they satisfy these requirements. Consequently, new design issues have arisen. In order to centralise processing, factors such as pool location, fronthaul design, and network cost must be addressed.

Table 1: Fronthaul capacity requirement for different split points

	Capacity requirements (Mbps)						
Split point:	7	6	5	4	3	2	1
Downlink	151	151	173	933	1075	1966	2457.6
Uplink	48	49	452	903	922	1966	2457.6

III. AI ASSISTED GREEN MOBILE NETWORKS

A. Risk Aware Sleep Mode Management

A risk-aware sleep management framework is proposed in this section. Algorithms based on artificial intelligence (AI) or machine learning (ML) are vulnerable to poor performance due to abnormal or unfamiliar input. Re-training may be required for AI systems if they fail to perform as expected when compared to the amount of traffic they get. The DT in Figure 3 obtains network data, connects with and simulates the actual system's performance, including the AI module, and analyses risk. We make a decision to employ AI, retrain the network, or temporarily disable the AI module based on this evaluation and the risk in the actual network. Detailed explanations of this method are provided below. To begin, we'll discuss the potential dangers of using BS sleeping algorithms. The DT model for BS ASM risk management is explained here. Figure 4 depicts the framework for risk-aware BS sleep mode management, which we integrate with the intelligent sleep mode management and DT.

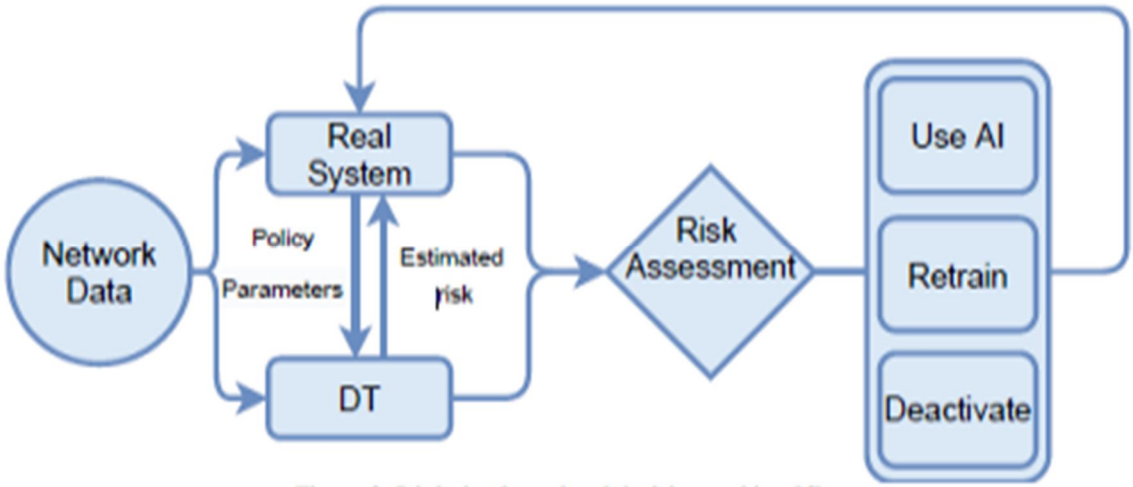


Figure 3: Digital twin assisted decision making [6].

1) *Identifying and Reducing Potentially Danger:* An example of a risk in this research is when an SM management algorithm makes sleep choices that delay connection setup times for new users. The greater the number of people facing delays, the greater the danger. Therefore, BS can prevent delaying a significant number of users if the danger might be assessed in advance. The DT idea is used as a virtual representation of the BS sleeping process for this purpose. Here, we provide an example of a DT aided risk-aware SM framework. For example, sleep length, ON/OFF state durations, sleep times, and arrival rates are all provided by the physical network in this design approach. utilising Equation, a virtual network updates its model parameters and predicts RDM based on observed data (2.11). Thus, the DT may use the Markov model and its updated parameters to anticipate the risk. With the help of the update parameter step and the Markov model, the virtual network is put together. The virtual network's output is utilised to estimate the risk. The operator sets a risk threshold, and this risk is compared to the actual risk (which will be available in future). The SMs should be disabled if the anticipated danger exceeds the threshold. The algorithm may need to be retrained if the estimated risk is larger than the real one. Otherwise, BS may activate the SMs and take use of the sleep mode management algorithms of BS if RDM is below the preset level. Paper 3 outlines and explains the details of this method. Operators may ensure that their sleep mode management algorithm does not make sleep choices when there are user arrivals or misses the chance to conserve energy when it is safe to do so.
2) *Managing the Risks:* A risk-aware sleep management framework is proposed in this section. To begin, we'll discuss the potential dangers of using BS sleeping algorithms. On the basis of a suggested hidden Markov model for BS ASMs, we estimate the risk and formulate it accordingly. Our solution for BS ASM management to quantify the risk follows. An integrated approach combining intelligent sleep mode management with hidden Markov model-based risk estimate is proposed as the last step in our risk-aware BS sleep mode management framework.

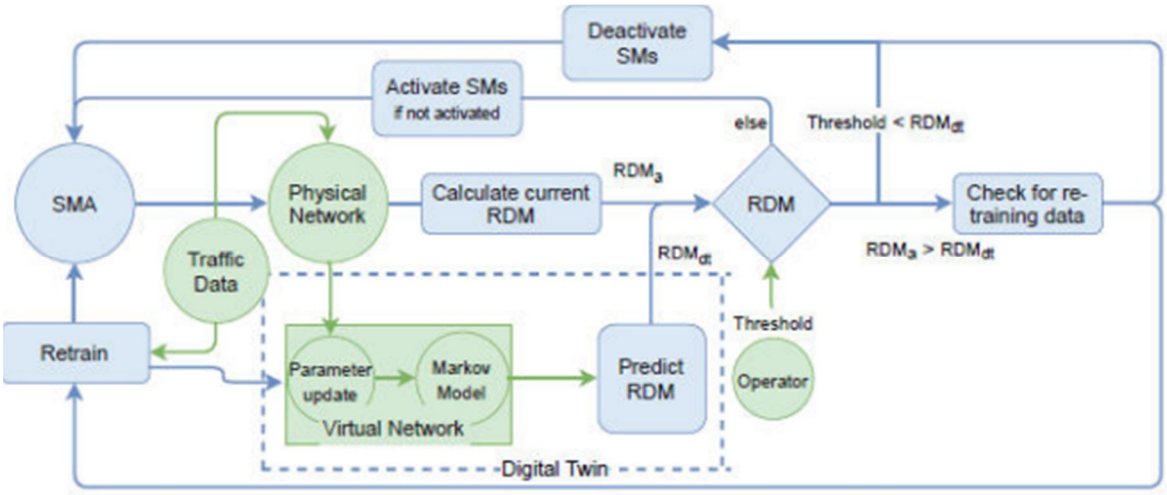


Figure 4: Digital twin based risk aware sleep mode management [6]

In this study, the risk of the SM management algorithm slowing the time it takes for new users to establish a connection is assessed. The more individuals who face delays, the higher the risk. If the hazard can be evaluated in advance, then BS can avoid delaying a considerable number of users. We employ the hidden Markov model, which in this case depicts BS sleeping, as a virtual representation of a process. Using machine learning and reasoning, the model is constantly updated based on real-time information. BS sleeping performance parameters like as sleep duration, customer delay risk, and risk for each network state may all be better understood and anticipated as a result. The BS sleep mode management strategy uses a Markov model, which is detailed in detail below. This is a digital twin in action. There are three main components to this research's digital twin: parameter updates, network models, and forecasts. The virtual network is a representation of the physical network, which is shown in the first two. The model is regularly updated using machine learning and reasoning based on real-time data. We apply it in the prediction phase to estimate and anticipate future performance indicators including sleep duration, user delay probability, and risk for each network state as a hidden Markov process. The BS sleep mode management Markov model is further explained in this section.

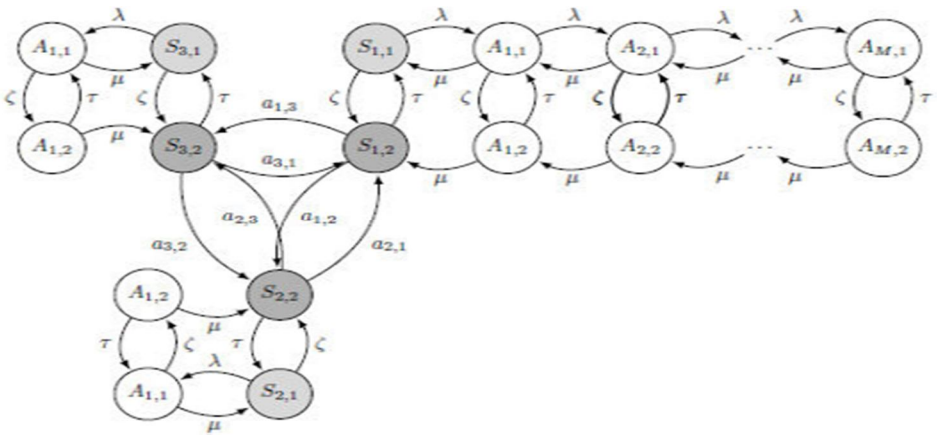


Figure 5: Markov model for advanced sleep mode management [6]

3) *Virtual Network Model:* A Markov Chain That Wasn't Noticed Analytical models are required to accurately predict the actual behaviour of ML-based sleep mode regulation in the form of a virtual network. According to Figure 5, we employ a hidden Markov model1 where states, traffic and model parameters are retrieved from the real network, as seen in Figure 5. A person's environment has an effect on two things: the level of stress on the body and the quantity of sleep they obtain. The next parts cover the hidden Markov model, information requirements, performance metrics, and a risk management technique.

IV. AI ASSISTED NETWORK ARCHITECTUREDESIGN AND MANAGEMENT

A. *Cloud RAN Network Architecture Power and Delay Models*

Improved cooperative solutions, greater load-balancing, and the possibility to share RANs are all advantages of centralised RAN operations. On the other hand, migration costs, supplying low-latency services, and requiring big bandwidth at the transport layer may pose significant challenges. To get around these issues, many network topologies and strategies have been devised. Processing processes must be divided and scattered over two layers of processing to avoid centralization, while certain lower-layer network operations may remain close to the RU in order to reduce fronthaul bandwidth demands. It is thus necessary to have a distinct level of network processing for the non-integrated service processing. This layer is connected to the central cloud through fast, low-latency transport. As a result, the network's transportation links are critical components. It is possible for an X-haul transport link to be either front or mid-haul. Passive optical network (PON), Ethernet network, or time-sensitive network may all be used to build this connection, which employs CPRI or eCPRI to transfer data to the cloud. To varying degrees, each of these innovations is preferable in its own right, depending on the specifics of the situation.

1) C-RAN Architecture as It Is Currently Known As indicated in Figure 6, the digital units (DUs) are located in a central cloud (CC) and are linked to the RUs through fronthaul connections in standard C-RAN architectures. It is a network that links the RUs to the central cloud through Fronthaul. There is a central cloud that handles baseband processing. Due to insufficient fronthaul connections, this design may not be able to carry I/Q signals between RUs and DUs. Because of this, we need to reduce the fronthaul's workload. Addition of additional processing layer between the RUs and central clouds known as an edge cloud may accomplish this.

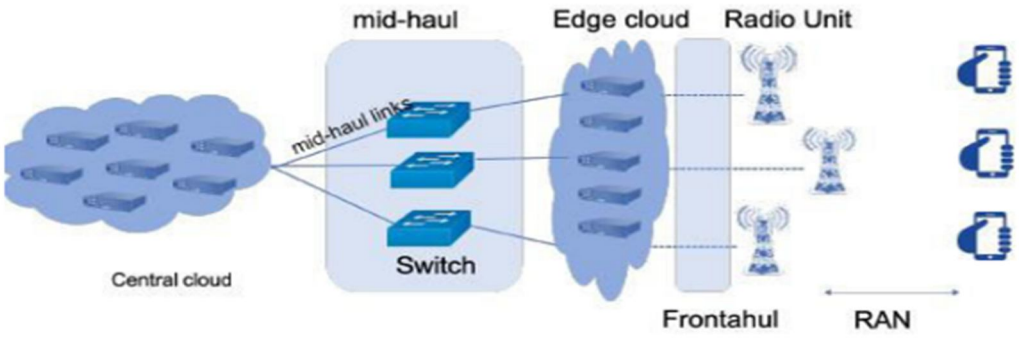


Figure 6: Basic C-RAN architecture.

2) Hybrid C-RAN The graphic depicts a central cloud, many edge clouds, and DUs deployed at both the core cloud and the edge cloud (EC). Central and edge clouds may be employed to distribute baseband processing as a consequence. In addition, both CC and EC may be equipped with RAM that can be utilised to store frequently accessed material. H-three-layer CRAN architecture consists of a cell layer, an edge cloud, and a centre cloud. Each cell in the cell layer serves a different UE, hence the density of the cell layer increases as more UEs are added.

The edge cloud may be formed by a DU and an RU, which are basically a DU and RU. In terms of baseband processing, the DU is in charge of some of it, while the CC takes care of the rest. EC and RU are connected via a fronthaul cable, which may be either coaxial or fibre optic. TWDM-PON or Ethernet base networks are possible options for the midhaul connection between EC and CC. Fig. 3 shows how this design works in practise. The common protocol radio interface refers to the midhaul/fronthaul links that link EC and CC (CPRI). Although these links are limited in capacity, demand for midhaul/fronthaul capacity in 5G networks is rising. Because of this, there will be an increase in traffic on the RU-DU links. It is being used to ensure that the bandwidth requirements for 5G are not surpassed by eCPRI (evolutionary CPRI). In order to meet the demands of the 5G network, this standard provides greater efficiency. Two independent functional zones are allowed for PHY of the BSs in the eCPRI standard, unlike CPRI. As a result, flexible radio data transmission may be achieved using a packet-based network such as IP or Ethernet. When compared to more typical WDM networks, packet-based Ethernet midhaul/fronthaul may reduce network expenses by aggregating and switching traffic at a lower cost. [111] Transport architectures based on packets are gaining traction in the industry. In order to fulfil the latency, throughput, and reliability needs of advanced 5G applications, the eCPRI protocol may be used. Many carriers may utilise the eCPRI since it is an open standard, improving global connection and speed by working together. [112]. eCPRI offers a number of advantages over CPRI, including the following:

- Ten-fold reduction of required bandwidth
- Required bandwidth can scale according to the to the user plane traffic.
- Ethernet can carry eCPRI traffic and other traffic simultaneously, in the same switched network.
- A single Ethernet network can simultaneously carry eCPRI traffic from several system vendors.

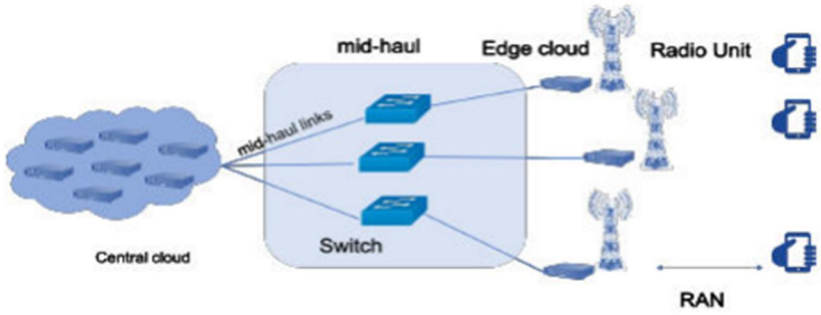


Figure 7: Two layer CRAN architecture.

In H-CRAN, the EC and RUs may coexist. With the use of an EC, you may connect more than one remote unit ECs are used as aggregation points for a group of cells/RUs. mmWave lines may be used to connect a cell to an EC through wireless communications. For midhaul connectivity, TWDM-PON technology is used. Containers for virtualized functions (DUs), containers for virtualized functions, are utilised at the periphery and in the central cloud. Due to a higher number of DUs in the CC, EC is less efficient than the CC. This design may be seen in Figure 8.

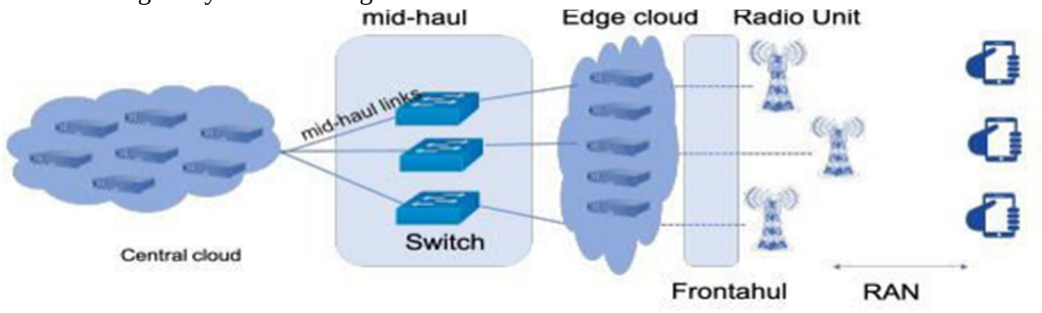


Figure 8: Multi layer C-RAN architecture

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

V. SIMULATION RESULTS

A. AI Assisted Green Mobile Networks

The proposed adaptive algorithm and its baseline counterparts are both executed in the event-based numerical simulation created using Python. A comparison of the SMA algorithm's efficiency with that of the fixed sleep mode, which can only use SM1, is made based on OBS as a benchmark for energy savings goals. These baselines are made available to all customers promptly and without any associated delay. [63] In addition, results from the SARSA algorithm are compared. When it comes to how long new users spend in OBS, there are no surprises. If BS scans all inactive times, the relevant SM may be used to fill up the gaps. If you want to follow along, you'll need to read [63]. Before we can use this data in our simulations, we must first process it. To illustrate this, we plotted both the generated load and a normalised average load from the original data set (Fig 9a). According to a comparison of the produced data to the original data, its typical daily behaviour is maintained. In Fig. 9b, the arrivals are shown on a five-second time scale. The reward function's delay or power savings are given precedence by the argument. In Fig. 10, we show how the suggested algorithm, SMA, performs in terms of the. Higher, the less importance is given to preserving energy. As an upper limit on energy savings gain, we compare SMA's performance to that of the OBS algorithm, as specified in [63]. Since future information is always available, OBS is always the best and non-causative choice possible. In this case, the entire compensation includes solely energy savings. SMA has the potential to save the most energy, but at the expense of causing inconvenience to the end user. In situations when latency is critical, the BS is more cautious when selecting deeper SMs and favours shallower SMs. As a result, less energy is conserved while a smaller amount of time is incurred. Fixed SMA beats SMA in situations when only SM1 is being used, which prevents BS from taking use of deeper SMs and therefore missing out on potential energy savings. As a result of its ability to better identify long- and short-term relationships in traffic data, SMA outperforms SARSA in terms of energy savings. It is shown in Fig. 11 that the SMA consumes more energy when it is used for sleeping than when it is not used at all, or when it is just used for SM1. Because the SMA prefers to use the shallowest SM during rush hour, the performance is extremely similar to that of the single SM1. Furthermore, saving energy isn't an option at these times, thus it's important to be extra cautious about delaying users and instead let the BS sleep or utilise the SM1. Figure 12 displays the SMA and OBS's normalised 24-hour energy-saving performance. The OBS algorithm's maximum energy savings are taken into account when calculating the figures. The greatest time of day to save the most energy is between the hours of 2:00 and 6:00 in the morning. At about 18:00, the best time to save electricity is when the sun is setting. It's important to note that the data being analysed pertains to a region characterised mostly by industrial structures. Depending on the kind of place, traffic patterns and hence energy-saving practises change.

Risk-aware BS sleeping algorithm performance is shown in Figure 13 (below). SMs are temporarily disabled when the risk value is large, as seen in Fig. 10 (below). When SMs are turned off, BS evaluates the hazards and turns them back on when the risk is low, i.e. during a period of time T.

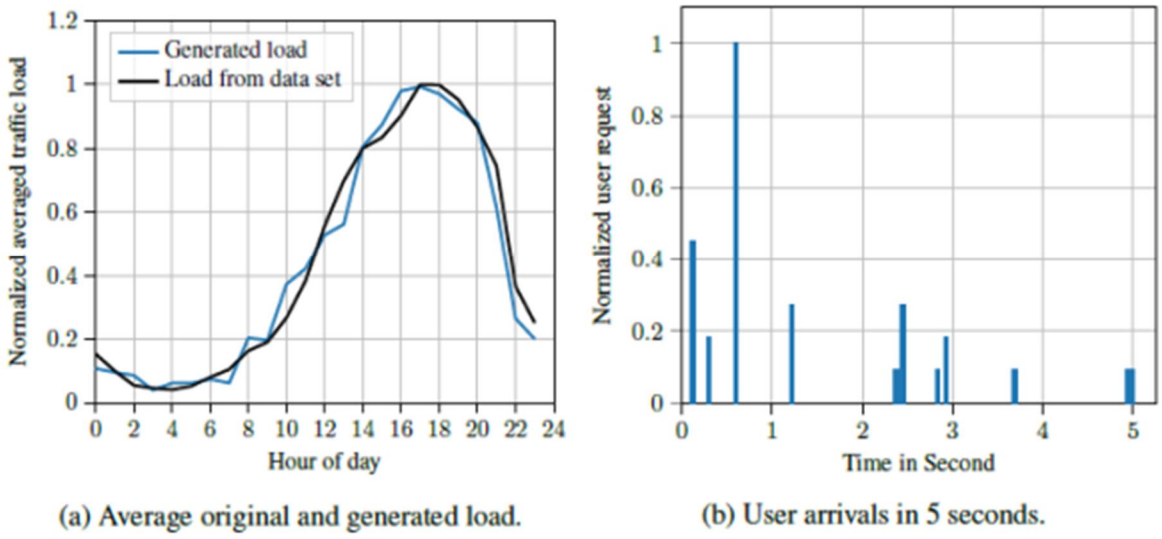


Figure 9: Network load generation as input to the network for $\tau = 0.1$ and $\zeta = 0.5$ [6].

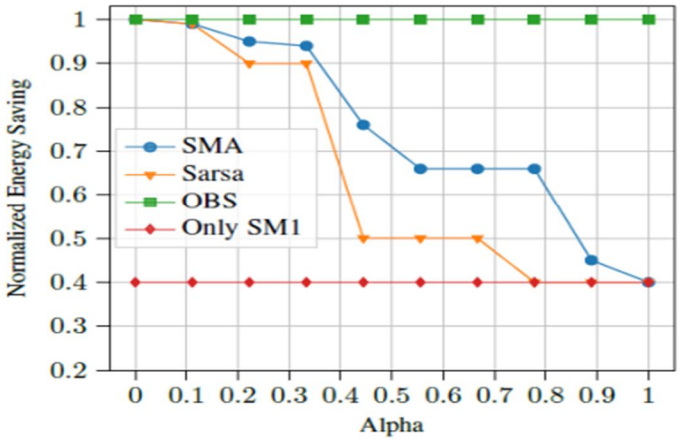


Figure 10: Energy saving vs delay trade-off [6]

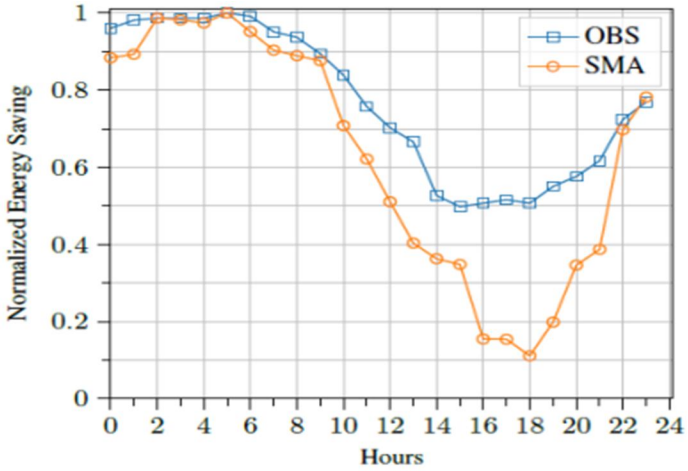


Figure 11: Normalized daily profile of energy saving [6].

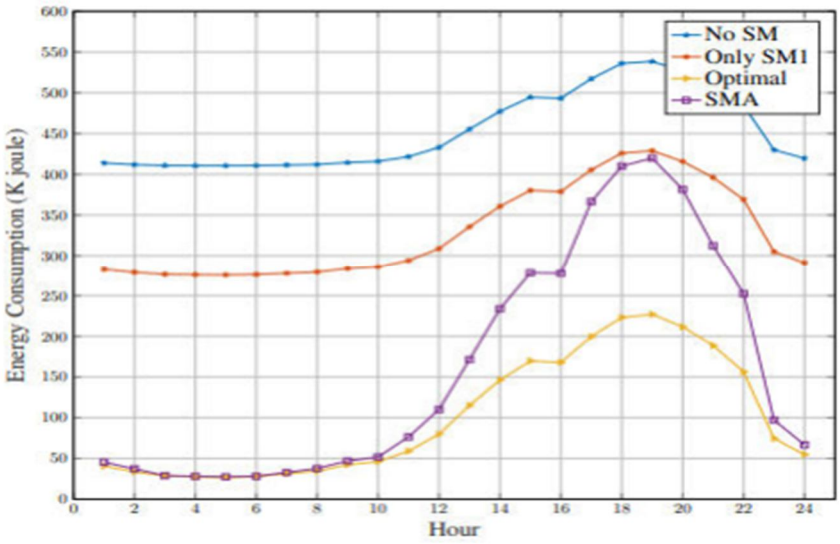


Figure 12: Daily profile of energy consumption.

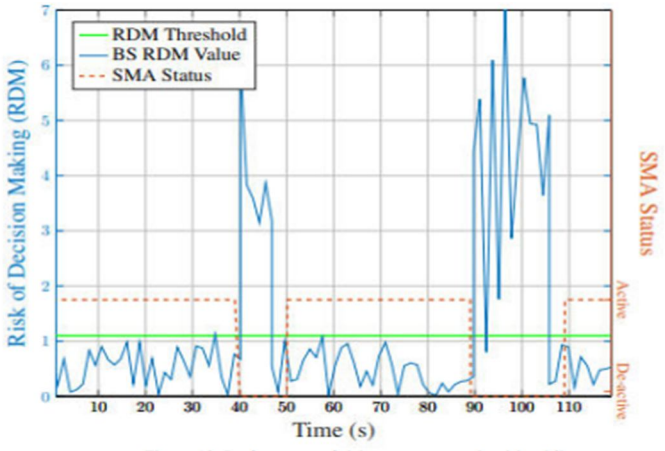


Figure 13: Performance of risk management algorithm [6].

B. AI Assisted Network Architecture Design And Management

Slice-based resource allocation (SBRA) is a suggested approach for slicing resources end-to-end and compares to the baselines in terms of performance. In this research, we assume that the slice manager is given an estimate of the number of users in each slice in advance. Consequently, a solution to the optimization issue is found in order to save resources. A network with four RUs and four slices, namely critical MTC, eMBB delay sensitive, eMBB delay tolerant, and enormous MTC, is considered. The maximum tolerated delays for these slices are 10, 50, 100, 150 ms, respectively. Each slice's file size is 0.1, 2, 2, 0.1 MB for the users. Otherwise, we'll presume the 5G NR numerology index is 0. All of the BSs are equipped with only one RRU/DDU, and we assume that the slice time is T = 1 second. If the BS has the same processing capacity as the CC, we assume that the BS has the same processing capacity as the CC. Furthermore, the BSs in (3.5) distribute processing resources according to the demand on each slice. We compare the suggested method with two cases: RAN slicing and user-based resource allocation (UBRA) (UBRA). With RAN slicing we just assign the bandwidth to each slice and the compute resources are presumed to be constant and supplied. We assume that the assigned compute resources are adequate to service the peak demand and this quantity is fixed for all loads. Comparing with this baseline, we can emphasise the importance of considering computation resource allocation. As part of UBRA, we build a slice for each user and assign resources in the CC and bandwidth in the RAN to that user. We can illustrate how the suggested SBRA may reduce the complexity of network slicing and how near the solution is to optimum solution by comparing it to UBRA. Figure 14 depicts the network's hourly energy usage under SBRA and RAN slicing scenarios. Real network data from BSs in rural Uppsala County is used in this illustration. Each weekday's right-hand axis in Figure 3.13 depicts the number of eMBB delay-sensitive and eMBB delay-tolerant service customers throughout each hour. The enormous MTC and critical MTC services were not included in the data set, thus a single critical MTC user and 100 massive MTC devices were substituted. Due to the fact that it uses less processing resources and hence may switch off superfluous equipment, SBRA operates better during non-peak hours, as seen in this picture. Processing resources can only be idle in RAN slicing, wasting energy. There is a greater disparity in network energy use at peak hours of the day, when both technologies are in use. However, owing to end-to-end resource allocation, SBRA still marginally outperforms RAN slicing. For SBRA, Figure 15a shows the amount of processing power assigned to each user per slice. The SBRA will calculate the average amount of resources needed for the following slice window based on this number. Because of the strictest delay requirements, the important MTC slice needs more processing resources per user than the other slices, as seen in the picture below: Due to their lenient delay limits, massive MTC users use the least amount of processing power. For each slice, Figure 15b shows the amount of processing power that has been allotted to it. Total processing at the CC is determined by the total traffic demand in each slice illustrated in Figure 15c and their delay requirements. There are more big MTC devices in this investigation, hence the overall amount of processing resources required is greater than in previous slices. Please keep in mind that each huge MTC device consumes much less processing power than represented in Figure 15a. The processing requirements of essential MTC devices, on the other hand, are greater than those of eMBB devices, despite their lower load. So that QoS may be met, greater processing resources should be allocated to the important MTC slice, which has a stricter latency requirement As a reminder, the file size of each user in UBRA may vary, and hence the processing resources needed may differ from SBRA. Therefore, as compared to SBRA, fewer processing resources may be provided to one slice of users in UBRA and more processing resources may be allotted to other slices of users in UBRA

Compared to SBRA, UBRA's combined communication and computation resource allocation challenge is more difficult. As a result, the number of optimization variables and constraints for users' Quality of Service (QoS) linearly grows with number of users, but for SBRA the number of constraints is limited to a fixed number of slices. To solve the issue of UBRA, as represented in Figure 16, it takes longer.

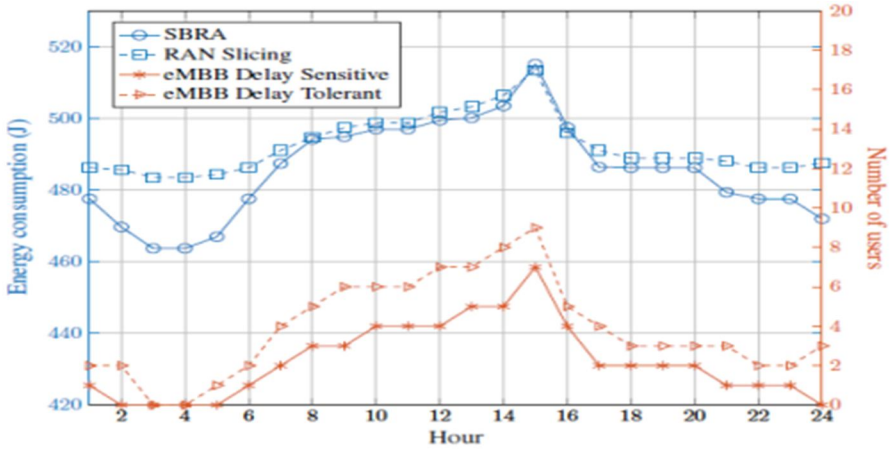
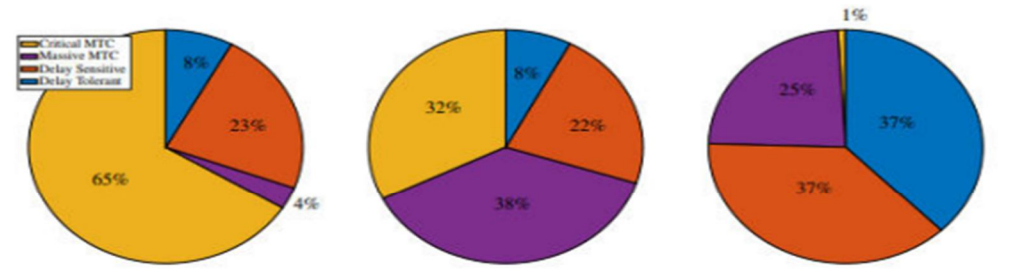


Figure 14: Hourly profile of expected energy consumption and number of users for eMBB delay sensitive and eMBB delay tolerant slices [8].



(a) Allocated processing resources per user. (b) Allocated processing resources per slice. (c) Proportion of traffic demands per slice.

Figure 15: Allocated processing resources to each slice and user [8].

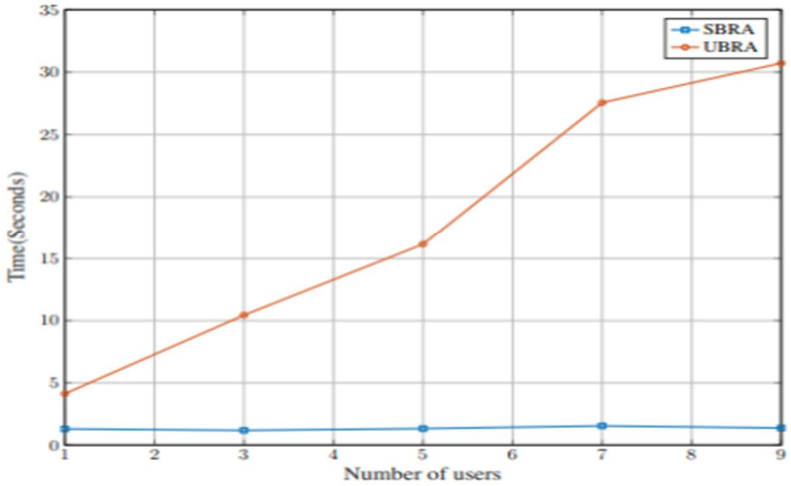


Figure 16: Time complexity performance comparison of SBRA and UBRA [8].

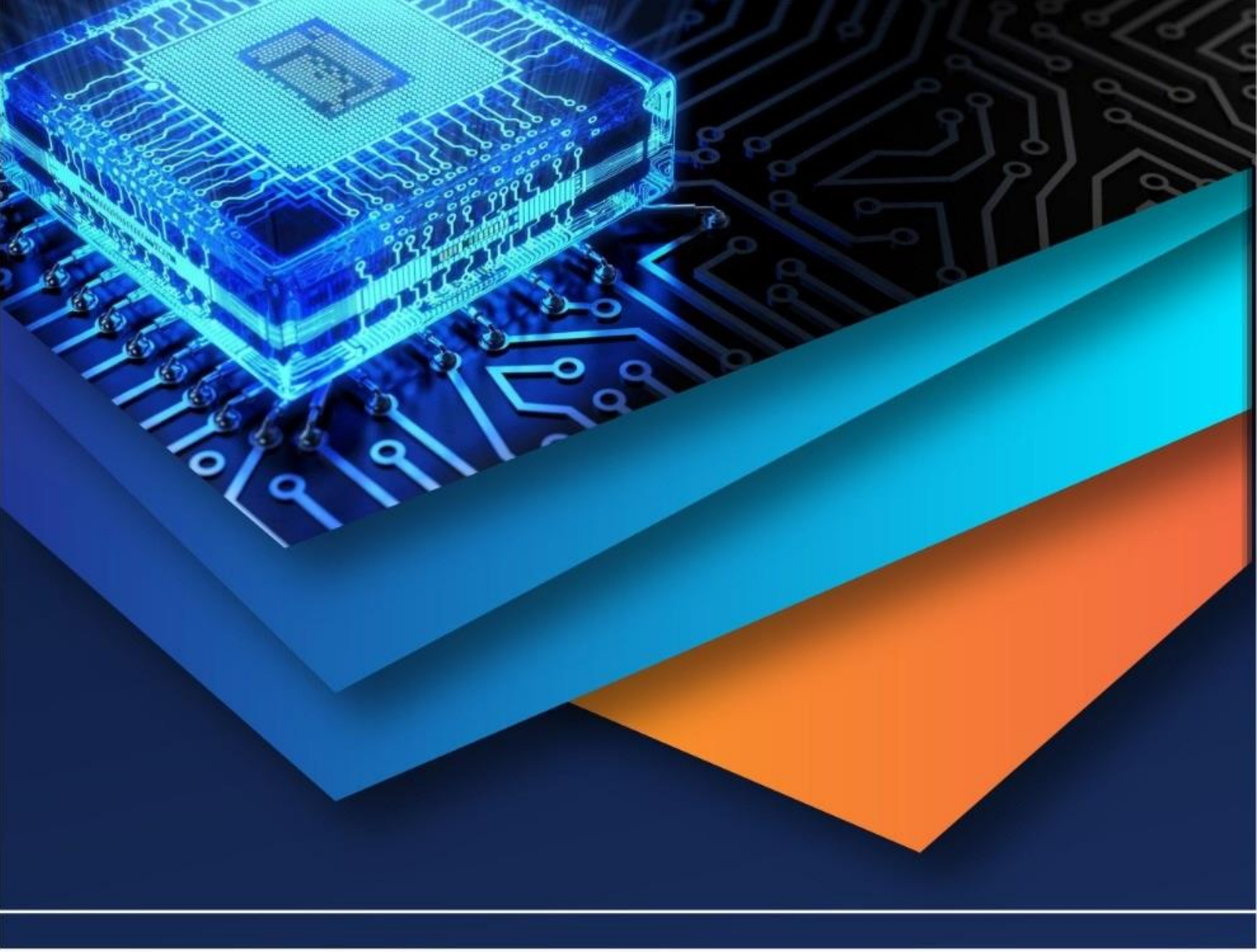
VI. CONCLUSION AND FUTURE WORK

Mobile networks use one to two percent of the world's overall energy usage. Increasing demands on service providers to increase network capacity, expand geographical coverage, and adopt advanced technical use cases in 5G and beyond will lead to a spike in energy usage. The long-term survival of mobile networks depends on their capacity to reduce their energy consumption. In this thesis, we examined how future wireless networks, such as 5G and beyond, may be made to be more energy-efficient. AI-assisted green mobile network, where we reduce the energy consumption of the network's base stations, and AI-assisted network architectural design and management, where we look at end-to-end network layouts and resource allocation to boost energy efficiency in networks are two examples of this type of work.

Energy efficiency will play a major role in the design and maintenance of networks beyond 5G. However, mobile network data has been available in recent years. In order to construct, optimise, and maintain their network, mobile operators and suppliers may greatly benefit from the use of accessible data. Artificial intelligence (AI) approaches may now be used to help network management and enhance performance indicators, due to advancements in the area and processing technology that can handle processing large-scale activities. Many issues of increasing network energy efficiency in mobile networks are discussed in this thesis. There are, however, a few limits to this concept, which may be addressed and expanded upon.

REFERENCES

[1] "Mobile data traffic outlook," https://www.ericsson.com/en/mobility-report/ dataforecasts/mobile-trafficforecast, accessed: 2020.
[2] "5G technology: Mastering the magic triangle," https://www.reply.com/en/ industries/telco-and-media/5gmastering-the-magic-triangle, accessed: 2020.
[3] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka et al., "Scenarios for 5G mobile and wireless communications: the vision of the metis project," IEEE communications magazine, vol. 52, no. 5, pp. 26–35, 2014
[4] M. Masoudi et al., "Green mobile networks for 5G and beyond," IEEE Access, vol. 7, pp. 107 270–107 299, 2019
[5] M. Masoudi, S. S. Lisi, and C. Cavdar, "Cost-effective migration toward virtualized C-RAN with scalable fronthaul design," IEEE Systems Journal, vol. 14, no. 4, pp. 5100–5110, 2020.
[6] M. Masoudi, E. Soroush, J. Zander, and C. Cavdar, "Digital twin assisted risk-aware sleep mode management using deep Q-networks," submitted to IEEE Transactions on Vehicular Technology, 2022.
[7] A. Sriram, M. Masoudi, A. Alabbasi, and C. Cavdar, "Joint functional splitting and content placement for green hybrid CRAN," in 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2019, pp. 1–7.
[8] M. Masoudi, D. Özlem Tugfe, J. Zander, and C. Cavdar, "Energy-optimal end-to-˘ end network slicing in cloud-based architecture," submitted IEEE Open Journal of Communications Society, 2022.
[9] imec, GreenTouch Project, "Power model for wireless base stations." [Online]. Available: https://www.imecint.com/en/powermodel .
[10] B. Debaillie, C. Desset, and F. Louagie, "A flexible and future-proof power model for cellular base stations," in Proc. IEEE VTC-Spring'15, May 2015, pp. 1–7.
[11] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. Da Silva, "Waveform and numerology to support 5G services and requirements," IEEE Communications Magazine, vol. 54, no. 11, pp. 90–98, 2016.
[12] M. Masoudi, "Energy and delay-aware communication and computation in wireless networks," 2020.
[13] M. Masoudi, H. Zaefarani, A. Mohammadi, and C. Cavdar, "Energy efficient resource allocation in two-tier OFDMA networks with QoS guarantees," Wireless Networks, pp. 1–15, 2017.
[14] M. Masoudi, B. Khamidehi, and C. Cavdar, "Green cloud computing for multi cell networks," in Wireless Communications and Networking Conference (WCNC). IEEE, 2017.
[15] M. Masoudi, H. Zaefarani, A. Mohammadi, and C. Cavdar, "Energy efficient resource allocation in two-tier OFDMA networks with QoS guarantees," Wireless Networks, vol. 24, no. 5, pp. 1841–1855, 2018.
[16] M. Masoudi and C. Cavdar, "Device vs edge computing for mobile services: Delayaware decision making to minimize power consumption," IEEE Transactions on Mobile Computing, 2020.
[17] M. Masoudi, A. Azari, and C. Cavdar, "Low power wide area IoT networks: Reliability analysis in coexisting scenarios," IEEE Wireless Communications Letters, 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)