5G High Performing Stand Alone Heterogeneous Network Deployment Employing Machine Learning Techniques

A thesis submitted towards partial fulfillment of the requirements for the degree of

Master of Engineering in Electronics and Telecommunication Engineering

Submitted by

ARCHI MUKHERJEE

Exam Roll No. M4ETC23010

Registration No. 160205 of 2021-22

June 2023

Under the Guidance Of

Prof. Iti Saha Misra

Department of Electronics and Telecommunication Engineering

Faculty of Engineering and Technology

Jadavpur University

Kolkata 700032

M.E.(Electronics and Telecommunication Engineering) course affiliated to **Faculty of Engineering and Technology**

Jadavpur University, Kolkata, India

CERTIFICATE OF RECOMMENDATION

This is to certify that the dissertation entitled "**5G High Performing Stand Alone Heterogeneous Network Deployment Employing Machine Learning Techniques**" has been carried out by **ARCHI MUKHERJEE(**ExamRollNo.:M4ETC23010,University Registration No.: 160205 of 2021-22**)**, under the guidance and supervision of Prof. Iti Saha Misra and be accepted in partial fulfillment of the requirements for the degree of **Master of Engineering in Electronics and Telecommunication Engineering**. The research results presented in the thesis have not been included in any other place submitted for any degree to any other University or Institute.

Iti Saha Misra 12/06/2023

DR. ITI SAHA MISRA
PROFESSOR
Dept. of ETCE
JADAVPUR UNIVERSITY
Kolkata - 700 032

(Prof. Iti Saha Misra) SUPERVISOR

Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata-700032

Manotosh Biswas 12/06/23

(Prof. Manotosh Biswas) HOD

MANOTOSH BISWAS
Professor and Head
Electronics and Telecommunication Engineering
Jadavpur University

Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata-700032

Ardhendu Ghoshal 12/06/23

(Prof. Ardhendu Ghoshal) DEAN

Faculty Council of Engineering and Technology, Jadavpur University, Kolkata-700032

DEAN
Faculty of Engineering & Technology
JADAVPUR UNIVERSITY
KOLKATA-700 032

M.E. (Electronics and Telecommunication Engineering)
course affiliated to

Faculty of Engineering and Technology

Jadavpur University, Kolkata, India

CERTIFICATE OF APPROVAL **

This foregoing thesis is hereby approved as a credible study of an engineering subject carried out and presented in a manner satisfactorily to warranty its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not endorse or approve any statement made or opinion expressed or conclusion drawn therein but approve the thesis only for purpose for which it has been submitted.

---------------------------- ---------------------------

External Examiner **Internal Examiner**

** Only in case the thesis is approved

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that this thesis contains literature survey and original research work by the undersigned candidate, as part of her **Master of Engineering in Electronics and Telecommunication Engineering** studies during academic session 2022-2023.

All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by this rules and conduct, I have fully cited and referred all material and results that are not original to this work.

NAME: **ARCHI MUKHERJEE**
EXAM ROLL NUMBER: **M4ETC23010**
CLASS ROLL NUMBER: **002110702005**
UNIVERSITY REGISTRATION No.: **160205 of 2021-22**
THESIS TITLE: **5G High Performing Stand Alone Heterogeneous Network Deployment Employing Machine Learning Techniques**

Archi Mukherjee
SIGNATURE

12.06.2023
DATE

Acknowledgement

I hereby convey my deepest gratitude towards my supervisor and guide Prof. Iti Saha Misra for her constant help, encouragement, supervision and valuable guidance that helped me to prepare this thesis paper. Her courageous attitude to choose challenging topics has helped me to explore innovative ideas. She constantly encourages one's growth as a future advanced researcher.

I would like to convey my sincere gratitude to Mr. Arijeet Ghosh, my senior and Phd Scholar of ETCE Department of Jadavpur University. The important discussions that we had enhanced the quality of the results of this thesis.

I would like to thank the entire ETCE department who directly or indirectly helped me to accomplish and fulfill this major goal of my life.

I am grateful to my family, my parents, for their constant support and encouragement at every step of my entire academia.

Place: **Kolkata**

Date: 12.06.2023

Dept.: **Electronics and Telecommunication Engineering**

Jadavpur University

Archi Mukherjee

(ARCHI MUKHERJEE)

Abstract

Fifth Generation mobile Wireless Network practically aims not only to connect everyone but everything as well, such as the IoT devices, self-driven cars etc. virtually on the same platform. The demand for higher data rates, lower latency, increased network capacity and uninterrupted user experience marks the definition for 5G wireless network. An important characteristic of 5G network is the Heterogeneous Networks. Any practical environment will include open outdoor – a direct Line Of Sight environment to office buildings – portraying the Non-Line Of Sight environment as well. Also there will be static users and dynamic users demanding for varying ranges of data rates. All these characterize the concept of Heterogeneous Network. Any new generation usually ensures backward compatibility to previous generations. 5G (IMT 2020) also ensures backward compatibility with 4G(IMT-Advanced) or 3G (IMT 2000) devices. Any 4G or 3G device will work seamlessly even when 5G is fully operational across the globe. As the name suggests multiple types of access nodes in the wireless network acts as the core of any Heterogeneous Network which include densely deployed femto/small cells under laid traditional macro cell network. This is one of the leads to the promising solution of high data rate, more capacity, lower latency etc. of the 5G communication. But small cell deployment is no easy task. Incorrect deployment can lead to poor data quality and higher interference. Machine Learning techniques can induce accurate and automated approach to small cell activation. This can resolve the issue of any incorrect deployment. Also Automation will lead to Rapid-to-Market of 5G services. In our current thesis work we thus implement the Kohonen's Self-Organizing Map (SOM) as an automatic data-analysis method. It is extensively applied to clustering problems and data exploration in industry. This unsupervised neural network is popular for its topological character mapping

where input data can be multi-dimensional. Hence we present how Kohonen's SOM can be efficiently used to proceed with the small cell deployment with maximum achievable SINR. Mobile user of wireless services increases each day. Hence the concept of Ultra Dense Network ~ which characterizes the much denser small cell network as compared to actual number of active users ~ brings an important solution to the seamless connectivity to the increasing tele density. We show that with increasing user density, on applying SOM for the small cell deployment leads to enhanced Area Spectral Efficiency targeting for Maximum Coverage. We also demonstrate the concept of Macro User Offloading which essentially leads to traffic offloading from Macro Cell network to Small Cell Network leading to lesser load on existing Macro base stations with increasing tele density. A popular unsupervised algorithm is K-Means. In order to establish the better characteristics of Kohonen's SOM we also establish a comparative study with the popularly implemented K-Means unsupervised learning algorithm. 5G cell edge spectral efficiency is represented by 5 percentile spectral efficiency. This is an important feature of any system's performance index. We hence produce the comparison of system performance when deployed using SOM and while using K-Means. This further demonstrates the better clustering characteristics of SOM. Our entire analysis of SOM's performance is extended from static to mobile user-equipment exhibiting stable and enhanced system performance with increasing successful handover percentage. K-Nearest Neighbors Algorithm is a supervised learning algorithm. We utilized K-NN intelligently in our work in between the training phases of our unsupervised learning techniques to extract with ease the interfering base stations while calculating the SINR of the system.

TABLE OF CONTENTS

Abbreviations

SOM	Self Organizing Map
KNN	K-Nearest Neighbors
IMT	International Mobile Telecommunications
ITU	International Telecommunications Union
eMBB	Enhanced Mobile Broadband
mMTC	Massive Machine-type Communications
uRLLC	Ultra-reliable and Low Latency Communications
IoT	Internet of Things
3GPP	Third Generation Partnership Project
gNB	Next Generation Node B
eNB	Evolved Node B
BSs	Base Stations
UEs	User Equipment
PPP	Poisson Point Process
IoT	Internet Of Things
SINR	Signal to Interference and Noise Ratio
UMi	Urban Micro
UMa	Urban Macro
LOS	Line of Sight
HetNet	Heterogeneous Cellular Network
QoS	Quality of Service
ANN	Artificial Neural Network
NN	Neural Network
uSL	Unsupervised Learning
OAM	Operations, Administration and Maintenance
LTE	Long Term Evolution
UTs	User Terminals
d_{2D}	2D distance
d_{3D}	3D distance
d'_{BP}	Breakpoint distance
h_{BS}	height of base station
h_{UT}	height of user terminal
σ_{SF}	Shadow fading standard deviation

List of Tables and Algorithms

List of Figures

Chapter 1 Introduction

1.1 Preface

Wireless Communication System has amazing effect in our daily lives. It leads to enhanced standard of living at minimal cost. Hence the thrive from one generation to another continues in wireless mobile communication, each targeting for more improvement and betterment over the other. The ITU (International Telecommunication Union) has always tried to provide specific improved requirements for IMT (International Mobile Telecommunication) for each generation such that 3GPP (Third Generation Partnership Project) or other technical bodies can gracefully implement the same for the improved living of human beings across the globe. IMT-2000 is the global standard for 3G wireless communication as defined by ITU. Similarly for 4G it's IMT-Advanced and for 5G it's IMT-2020 [1]. Among everything each standard aims for more and more data rates with improved lower latency, higher spectral efficiency and deeper coverage. The COVID-19 pandemic became an eye-opener for the entire world that how seamless connectivity, higher speed with lower latency has now become a necessity for the world and not a luxury. Though it had affected the speed of 5G deployment but had induced a positive thrust for the quick and correct deployment of 5G meeting the IMT-2020 standards. As per 3GPP Release 15 document [2] the peak data rate is supposed to be 20 Gbps which is 20 times more than 4G(1 Gbps as specified for 4G). Latency is also expected to be 10 times lesser in comparison to 4G (i.e 1ms [2]). Conclusively 5G is bound to bring better spectral efficiency and coverage in comparison to any previous generations. Seamless connectivity is expected starting from static users, pedestrians to high speed objects ~ as high as 500 km/hour [2]. According to IMT-2020, 3 broadly classified usage scenario for 5G are :- eMBB, mMTC and uRLLC [3]. eMBB covers high user density(like city/urban area) requiring increased data rate, supporting seamless connectivity in high mobility. mMTC is as e.g. for IoT devices, requiring lower power consumptions supporting higher connection density. uRLLC cater for critical applications (like military applications or remote surgical applications) where lower latency is super important along with reliability. Hence the 5G network technologies are intended to accommodate varied innovative services – which connect every object, appliance or device to the Internet. This includes large influx of data traffic with enhanced user experience in terms of quality.

In order to meet 5G requirements of higher data rate along with lower latency, heterogeneous networks (HetNets) have been introduced. This involves deployments of additional low power nodes under the existing conventional high-power nodes within the coverage area. As mentioned in [4], a future wireless heterogeneous network (HetNet) was already envisioned to have macro base stations (BSs) providing a coverage blanket, along with small BSs like pico BSs and low powered user-deployed femtocells. Fig.1.1 demonstrates various applications of 5G wireless communication in heterogeneous networks in day-to-day life [5][6]. As was mentioned previously, post COVID-19 people urge for rapid deployment of 5G networks. Also the large influx of data created due to varied services requires intelligent handling for reliable communications. Machine Learning plays an important role in this field. Machine Learning is broadly split into Supervised, Unsupervised and Reinforcement Learning. While Supervised Learning gets trained with labeled data to provide clear classification, Unsupervised Learning is meant to act on stochastic unlabeled data set to provide intelligent clustering. Reinforcement Learning can work even in complete unknown environment and gets trained via rewards in its trial-error process. Under Unsupervised Machine Learning, SOMs are categorized as unsupervised neural network, first proposed by Kohonen [7]. In [8] Self-Organizing Map (SOM) had been used as a clustering tool for database operations. SOM follows a competitive learning process in its training period resulting in winning neuron ruling over the input data. Unsupervised learning feature of SOM and its capability of handling large influx of data due to its clustering feature persuaded us to explore its capability in our thesis work.

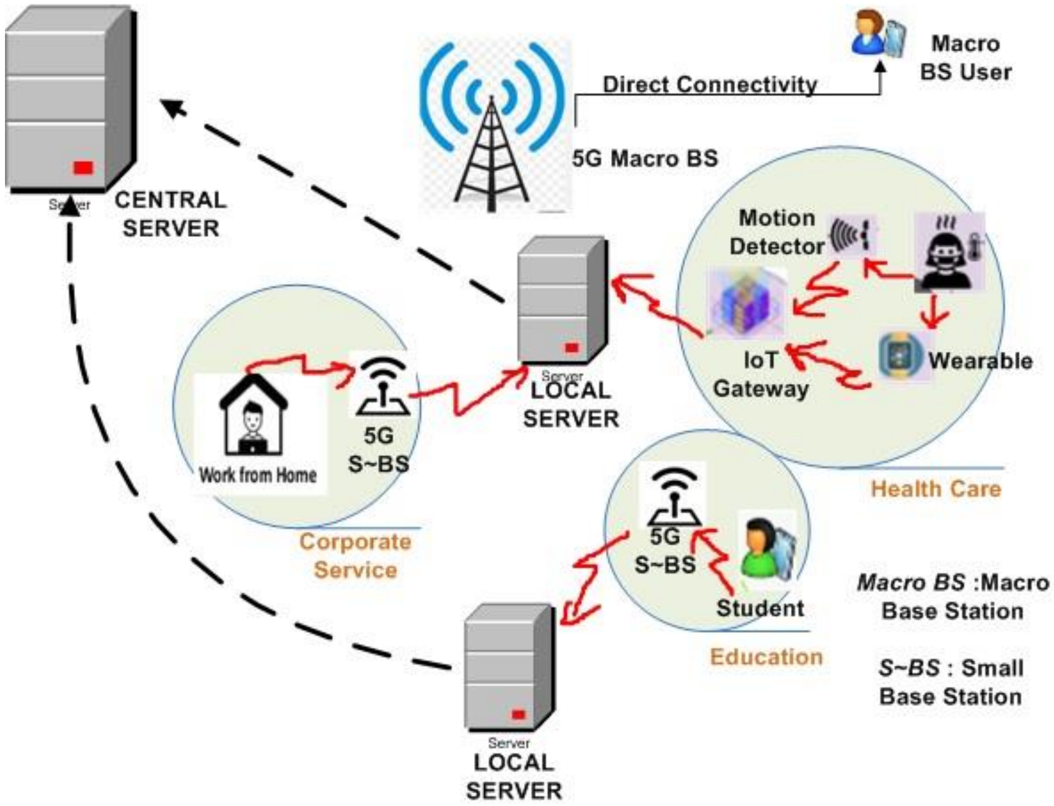


Fig.1.1. Applications of 5G

Another popular unsupervised learning algorithm K-Means was included in our work to establish a comparative study between the two exhibiting enhanced performance of SOM over K-Means.

1.2 Objective

According to the Literature Survey in 1.3, SOM has extensive capabilities that are yet to be explored in Wireless Communication. Hence in this thesis, we explore the performance of the unsupervised neural network, Kohonen's SOM in Wireless Communication field. Our model is based on 5G FR1 spectrum. But the prototype is flexible to be extended to 5G FR2 spectrum as well post careful calculation of path loss is performed.

At first we will provide a comparison between the two Unsupervised Clustering Approach, K-Means and Kohonen's SOM, in our work while imposing each clustering technique for efficient deployment of the base stations in the heterogeneous wireless network. This will be specifically done to prove the enhanced performance of SOM over a popular UL technique like K-Means. We have already seen how K-Means have been applied in terms of modeling of 5G HetNet [9] under Literature Survey. To get the complete picture we also study the computational complexity difference between K-Means and SOM. This gives an idea of faster performance with respect to increasing tele density as well.

Further we induce mobility following Random Walk Model, among the UEs of a specific coverage area to study how the SOM distributed HetNet model is capable of handling successful handoff at different SINR threshold in the system. This has been studied by extending across varying mobility, velocity and tele-density. Study based on tele-density gives a new horizon reflecting the ultra-dense 5G network.

Along with this, we also study the behavior of SOM distributed HetNet model by varying the macro to small base station ratio in terms of Area Spectral Efficiency, Coverage Probability and Load Per Tier for different SINR thresholds as required by any 5G system model. Load per Tier demonstration provides a clear idea of the macro offloading benefits in 5G HetNets.

1.3 Literature Survey

The impact of COVID-19 pandemic sweeping across the world affecting the timelines of the network operators for 5G deployment is well stated at IEEE Innovation at Work [10]. As stated in [10] already 3GPP had to delay its two standards – 3GPP Release 16 Stage 3 and 3GPP Release 17 which cover 5G specifications for standalone network, vehicle-to-everything systems and IoT. In 3GPP Release 18 [11], 5G new radio air interface were published. Starting with FR1(Frequency Range 1) ranging from 410 MHz to 7.12 GHz. The other being

FR2 (Frequency Range 2) from 24.25 GHz to 71 GHz – this signifies the millimeter wave application [11][12]. Our very own Indian physicist, Sir Jagadish Chandra Bose had discovered mm wave (one can refer Wikipedia for details). But still developing countries like ours, India, suffers with the 5G spectrum cost as an additional burden along with time constraint. These two major factors motivated us to apply Machine Learning techniques for faster and economical rollout of 5G services. Along with rollout of 5G services backward compatibility of the existing 4G wireless network architecture is inclusive [13][14].

As elaborated in [15] Machine Learning techniques can be broadly categorized as Supervised Learning (SL), Unsupervised Learning (UL) and Reinforcement Learning (RL). UL is based on unlabeled data set and hence uses functions to deal with unknown input patterns [15]. In [16][17] we find the clear mathematical demonstration of stochastic characteristics of wireless channel. Further, in [18] we see that in presence of fading, shadowing and interference, deterministic approach of placing the base stations at any new area is not a good idea. Hence, UL proves to be a good option to be applied in wireless communication which is well equipped with unlabeled varying data. Among the UL techniques, ANN is an efficient and popular method having applications in various domains like image recognition, speech recognition, medical diagnosis etc. [19]. An artificial neuron's operating principle is similar to biological nerve cell [19]. Similar to the dendrites of the biological neuron, the artificial neuron receives the information in its input layer, processes the same in the processor and calculates the weights to generate the output layer, which is further forwarded to another neuron just like the synaptic communication between two biological neurons [19]. According to [20] once trained an unsupervised ANN doesn't need any reprogramming and can produce the required results. Kohonen's SOM is a special type of unsupervised ANN based on competitive learning [7]. It not only acts as unsupervised clustering algorithm but also a powerful visualization tool performing a powerful nonlinear mapping from a high-dimensional input layer to a low-dimensional output layer of neurons [21]. In SOM the learning process considers both the similarity between the input data and also the distance between the neurons [7]. This is an added advantage in wireless mobile communication, since it has the ability to provide the accurate position of the base stations considering the interference from the neighboring base stations as well. Another popular UL technique used is K-Means algorithm. It uses an iterative process to cluster the unknown input data [22]. But this methodology requires K clusters and initial K centroids as input for clustering of data [22]. Also since it's not a neural network hence the idea of distance between the K centroids doesn't come into play by default, as can be seen in case of SOM. Another aspect is there and will be large influx of varied data when 5G is operational at full mode across the globe. As mentioned earlier SOM is also a strong visualization tool. Hence if SOM is utilized for the deployment of small base stations in HetNets of 5G, it's visualization characteristics can also be utilized as an add on benefit. Such benefit cannot be expected from K-Means on the other hand.

K-Means clustering is a popular UL technique and several related works applying K-Means have been published in wireless communication. Among them we find that for uneven distribution of sensor nodes in Wireless Sensor Networks, [22] demonstrates the use of optimized K-Means for the clustering of sensor nodes. Also in [23] K-Means is used for network modeling followed by analysis of downlink HetNet. To maximize throughput in 5G HetNets optimized K-Means is used in [24]. In another work [9], K-Means is used to find the optimal location of small cell BSs.

Application of the efficient yet simple and highly visual neural network, Kohonen's SOM, on efficient deployment of small base stations for maximum throughput in 5G HetNets, is hard to find. Though we find its application in various other fields. Under smart grid applications [25], it has been observed that due to frequent climate variations maintaining uninterrupted power supply is a challenge. Hence for uninterrupted power supply to critical loads in distribution system operators' networks, SOM is used as a leverage in this work [25]. In [8], SOM had been used as a clustering tool for database operations. A detailed survey can be observed in [21] where implementation of SOM found in hardware [Application-specific integrated circuit (ASIC), field-programmable gate array (FPGA)] is analyzed. The substantial parallelism feature of SOM algorithm makes it suitable as high energy efficient algorithm – the main reason why SOM is targeted for hardware implementation. In medical field there has been a critical study for diagnosis of spinal column patients, where comparison between SOM and K-Means has been established [26]. In [26], we find that SOM outperforms K-Means while detecting patients with spinal problems. For further improvement of K-Means' performance in recommendation system of smart TV, in [27], SOM has been used. SOM has been opted for the initial clustering phase to obtain the required central points and its output has then been used in K-Means for its improved performance. In telecommunication the use of SOM for pattern analysis in 3G mobile networks can be observed in [28]. In order to create a more optimized and cost-effective task for any mobile operator, SOM has been used to find near-behavior cells based on two correlated characteristics: uplink noise and frame error rate in [28]. Clustering of channel modeling in wireless communication using SOM is demonstrated in [29][30]. As in [31] cell clustering is proposed based on SOM with the help of OAM collected data. SOM algorithm in [31] is used to analyze the LTE radio access network's performance. The LTE cells are clustered based on downlink features. Earlier in [32], the clustering capability of Self Organizing Map is utilized for the detection of anomaly in mobile communication system which is under surveillance.

In the thorough survey conducted on clustering [33], a clustering technique is designed to support Load Balancing, Throughput, Coverage, Connectivity, Mobility, Inter-Cluster routing etc [33]. Based on our discussion till now, we already find SOM's extensive capability in various other fields. We thus intend to explore how efficiently SOM can also provide these clustering characteristics to the benefit of Wireless Communication.

1.4 Outline of the Thesis

- Chapter 2 provides the background study of 5G Wireless Communication and various machine learning techniques. Along with this it portrays how KNN Supervised Learning has been intelligently used throughout the thesis work along with Unsupervised Learning application. The simulation results for urban and rural areas on applying Unsupervised Neural Network SOM for the deployment of the base stations is also presented. The UEs are notably distributed by Poisson Point Process. This idea is applied across the thesis work.

- Chapter 3 deals with the comparative performance analysis of the popular Unsupervised Learning Algorithm K-Means and Unsupervised Neural Network Kohonen's SOM on various performance parameters like Area Spectral Efficiency, Coverage Probability, Load Per Tier and Computational Complexity. The entire comparison is based on Static UEs.

- Chapter 4 introduces the concept of Random Walk Model which helps in imposing speed in static UEs. The SOM based deployed System Model is then applied to the mobile UEs to carefully study its behavior and efficiency with regards to handover. Successful handoff percentage is studied by varying velocity as well as mobility percentage in a particular geographical area.

- Chapter 5 derives the conclusion of the results that we have obtained in our thesis work and also includes the future scopes of research of the work.

Chapter 2 Background Study of 5G Wireless Communication and Various Machine Learning Techniques

2.1 Introduction

The evolutionary process of Wireless Communications spanned over a few decades. In 1894 Guglielmo Giovanni Maria Marconi first experimented an end-to-end radio wave based wireless long distance transmission using large built in transceiver system. This discovery was a major boost to the generations of Wireless Communications which were primarily dependent on the radio wave spectrum. It all started in 1980s with the advent of 1G analogue cellular system. 2G evolved in 1990 to upgrade to digital network and to the popular GSM (Global System for Mobile Communication). However demand for high data rates led to the inception of 3G in 2001. To support further modern services (like video conferencing etc) LTE (Long Term Evolution) in 2009 and LTE-Advanced (4G) came into limelight [34]. But as Plato, the Greek Philosopher, states "Necessity…the mother of invention" and hence generation evolution couldn't stop at 4G. The extension of wireless mobile services from men to things in industries led to 5G in 2019 [34]. Interestingly the invention of millimeter wave which was made as early as 1895 by our very own Indian physicist Sir Jagadish Chandra Bose, became a core aspect after a few decades in 5G communication. We will study in this chapter that why these generations were required and why 5G is such an important generation whose deployment can lead to an impressive upgrade to standard of living of human being.

Now higher the generation, denser and more heterogeneous cellular network of wireless communication becomes. At the first step if the deployment of the network is done manually then this might lead to incorrect deployment resulting in higher interference, lower QoS, lower coverage etc. Also this will result in cost of planning and maintenance of the network post installation and configuration of the network. The problem of incorrect deployment and the cost can be subsequently reduced if human intervention is reduced. This can be achieved only when Machine Learning is applied as the relevant solution. The field of Machine Learning allows systems to learn and improve based on experience with reprogrammable feature installed in it. Subsequent correct deployment of small base stations under laid a macro cell network will lead to higher coverage, lower latency, lower interference and hence the required QoS of our latest generation i.e 5G. Thus, in this chapter we first begin with the concepts of Machine Learning and specifically why the unsupervised neural network, SOM (Self Organizing Map) has been employed for our thesis work is explained.

2.2 Evolution of Generations in Wireless Communication From 1G to 5G

Wireless Communication evolved from generation to generation with increasing number of users, services, mobility, deeper coverage etc. Hence we will now discuss briefly the major component evolution that has taken place across generations to support the specific needs and services of mankind. We have referred the survey discussion in [35] while analyzing this evolution.

2.2.1 First Generation (1G)

1G also known as Advanced Mobile Phone Services (AMPS) in the US, aimed towards supporting voice services for mobile users. Analog signals along with Frequency Modulation technique was used in the Circuit Switched Network**.** Carrier frequency used were 30 kHz and a single user was assigned an entire channel whether required or not. Operating Frequency was 800-900 MHz.

Hexagonal cells along with frequency re-use approach were used to accommodate large coverage area. The Mobile Telephone Switching Office (MTSO) connected the Base Transceiver Station (BTS) to the PSTN.

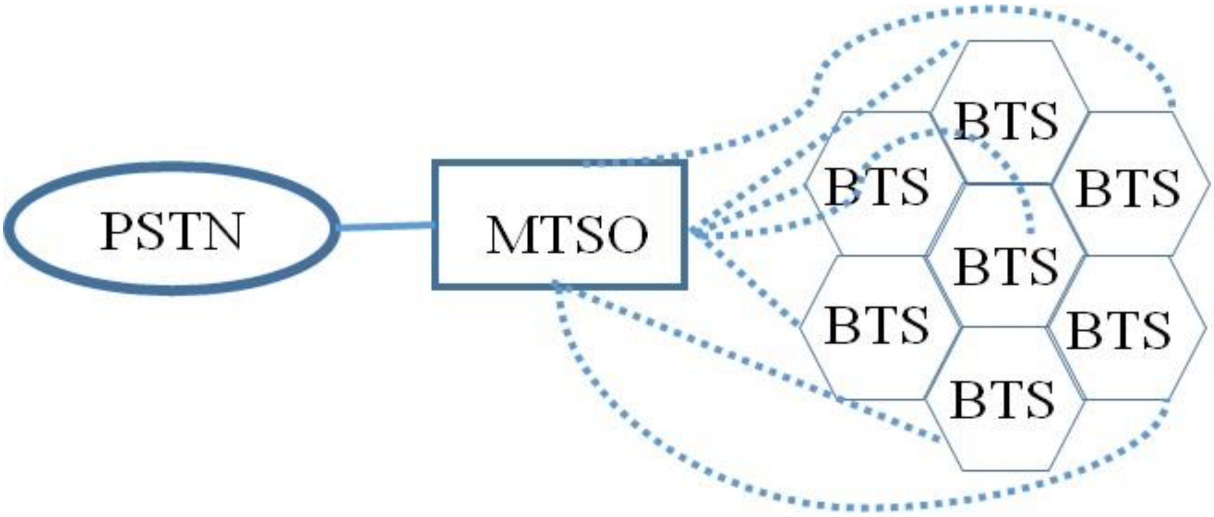


Fig.2.1. Block Diagram of 1G Network Architecture

Note: PSTN (Public Switched Telephone Network)

2.2.2 Second Generation (2G)

2G mobile cellular network were developed to support both voice and data services. Digital modulation came into play and circuit switched (CS) network were extended to support packet switched (PS) network as well. Global System for Mobile Communications (GSM), the popular 2G network, advanced towards General Packet Radio Services (GPRS) in 2.5 G and then towards Enhanced Data Rate for GSM Evolution (EDGE) in 2.75 G, advancing in data rate from 10 Kbps to 200 Kbps. Operating Frequency used were 900 MHz, 1800 MHz/1900 MHz. Carrier frequency used were 200 KHz while channel bandwidth employed were 25 MHz.

The Base Transceiver Stations (BTSs) and the Base Station Controller (BSC) comprised of the Base Station Subsystem. Operation and Support Subsystem (OSS) governed the GSM network. The Mobile Switching Center

(MSC) is responsible for registration, authentication of any subscriber there by supporting further for subscriber's call routing, location update etc.

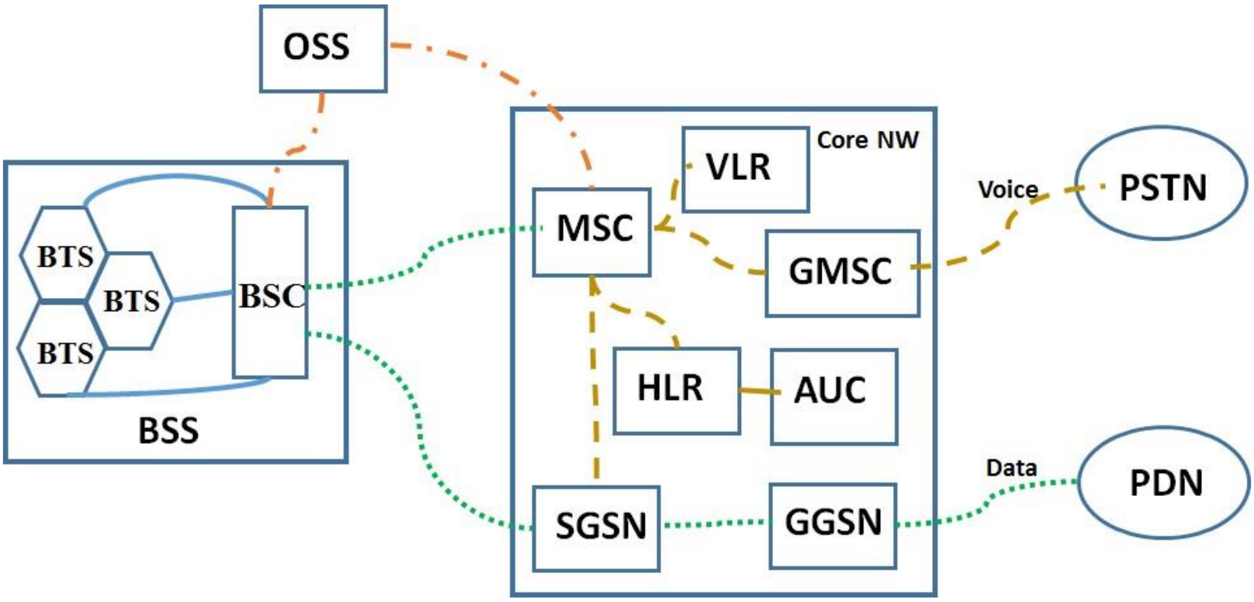


Fig.2.2. Block Diagram of 2G Network Architecture

Home/Visitor Location Register (HLR/VLR) comprised of the database of subscriber information. In order to direct calls to the correct MSC, HLR provides the required subscriber data to the Gate Mobile Switching Center (GMSC). For packet routing in order to support data routing, serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN) plays the major role. In 2G, other than GSM (also known as IS-136 in the US), CDMA was another popular evolving network, also known as IS-95 in the US.

2.2.3 Third Generation (3G)

The demand for video services along with voice and increasing data rates marked the beginning of 3G digital cellular technology. CS and PS services were already provided but the multiple access technology had to be changed from Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) to Wideband CDMA (commonly known as UMTS: Universal Mobile Telecommunication Service) and CDMA2000 respectively.

The Radio Network Controller (RNC) replaces the BSC functions, Node B replaces the BTS in 3G WCDMA network. Node B and RNC forms the UMTS Terrestrial Radio Access Network (UTRAN) of 3G network.

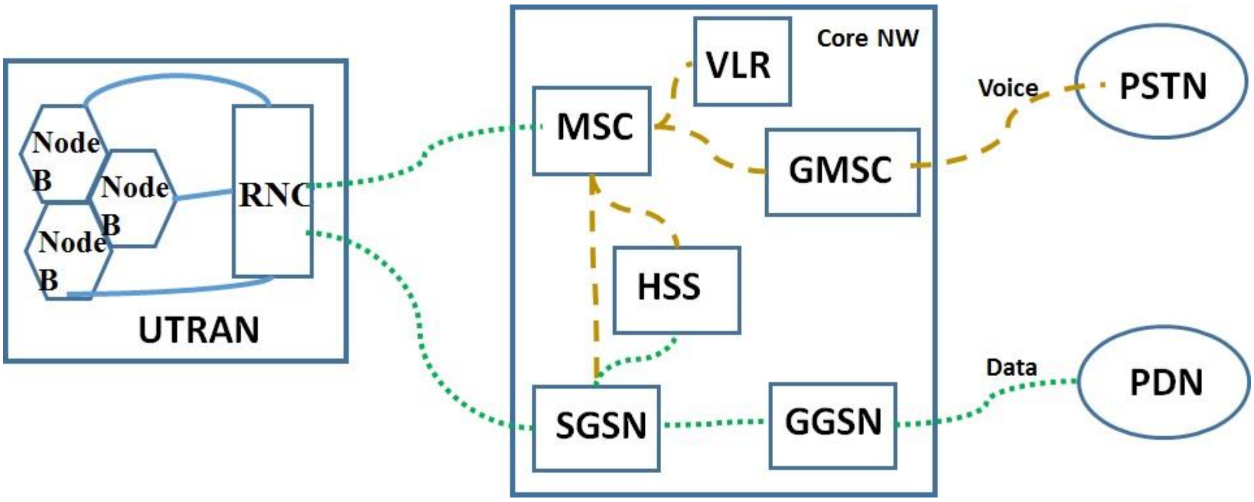


Fig.2.3. Block Diagram of 3G Network Architecture

High Speed Packet Access (HSPA) induced the idea of shared channel leading to efficient packet delivery supporting both downlink (HSDPA) and uplink (HSUPA). The HSS (Home Subscriber Server), concatenation of HLR (Home Location Register) and the AuC (Authentication Center). To deliver voice, video and text messages over IP network, IP Multimedia Subsystem (IMS) plays an important role. Channel Bandwidth supported in 3G was 15 to 20 MHz with operating frequency as 2.1 MHz.

2.2.4 Fourth Generation (4G)

Demand for varied services like IP telephony, Smart TV, Video Conferencing, 3D Television etc. , induced the need for more data speed leading to the invent of Long Term Evolution (LTE) and then 4G VoLTE (Voice Over LTE) with higher speed of communication.

To achieve this Orthogonal Frequency Division Multiple Access (OFDMA) was selected to handle FDMA multiple access efficiently. The entire spectrum was split into sub-channels each assigned to a subcarrier with a spacing of 15 kHz. MIMO is implemented in 4G which helps in increased data capacity with decreased interference. Channel bandwidth up to 100 MHz is supported with operating frequencies at 600 MHz, 700 MHz, 2.1 GHz, 2.3 GHz and 2.5 GHz.

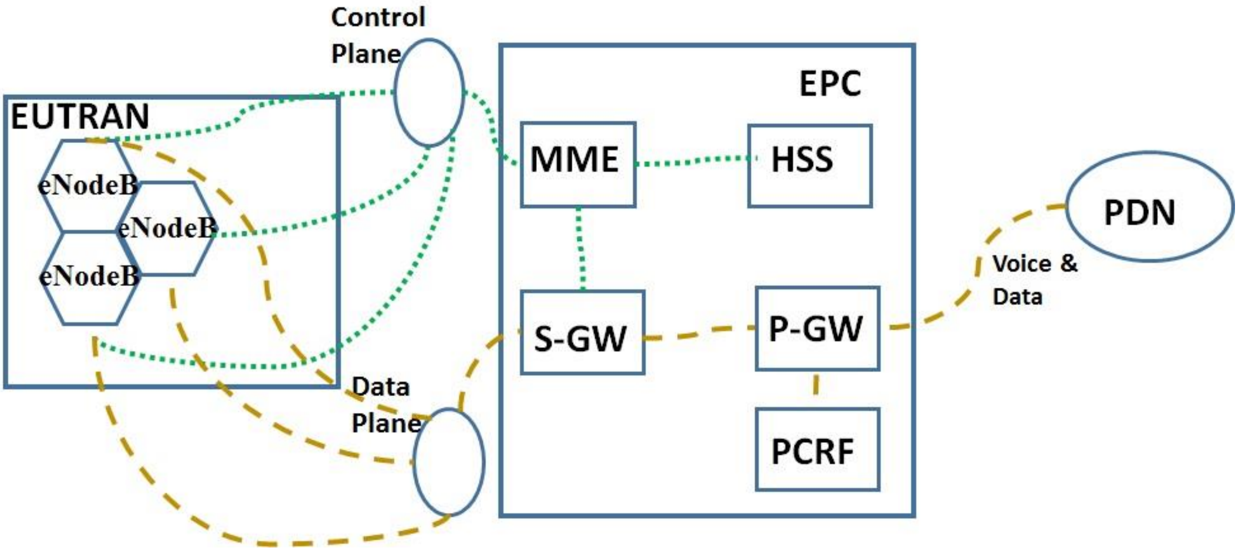


Fig.2.4. Block Diagram of 4G Network Architecture

The NodeB functions gets replaced by the eNodeB which forms the evolved UTRAN (EUTRAN). Evolved Packet Core (EPC) performs the core network functions. The control plane rules over the control signals which includes activities like checking the signal strength, start, stop or change the UE's data session etc. The user plane rules over the actual call flow between two UEs, the actual data that gets exchanged between the UE and the PDN. Both planes support signal flow across EUTRAN and EPC components which forms the core of Evolved Packet System (EPS). 4G marked the inception of varied data rates. HSS continued to be the central database containing the subscriber information. Data from the serving base station gets routed to the Serving Gateway(S-GW) to the Packet Data Network (PDN) via the PDN

Gateway (P-GW). PDN is the connect to the Internet. Each PDN is identified by an access point name (APN). S-GW is similar to SGSN while P-GW is similar GGSN of 2G and 3G networks. All policy control decisions and charging features are executed by Policy Control and Charging Rules Function (PCRF) based on the P-GW's Policy Control Enforcement Function (PCEF). To support backward compatibility to previous generations, Circuit Switched Fall Back (CSFB) is employed that allows GSM or other circuit switched network to connect to the 4G network architecture. LTE is all-IP-based network and hence requires CSFB for enabling circuit switched activities.

2.2.5 Fifth Generation (5G)

The IMT-2020 is the standard that defines the 5G requirements. 5G services are segregated into eMBB, uRLLC and mMTC. While eMBB services demand high data rate and spectral efficiency, uRLLC require lower latency and mMTC demands high connection density [35]. To fulfill all 5G requirements, multiple technologies come into play. Massive MIMO, mmWave, NOMA, cloud Radio Access Network (cloud-RAN), ultra dense networks, wireless caching etc. In the 5G NR cellular network architecture eNB gets replaced by "next generation NodeB" or "gNB" which basically intends to connect not only everyone but also everything. Distributed Unit (DU) and Centralized Unit (CU) are the two groups that form the NG-RAN. EPC of 4G is replaced by the 5G Core network i.e 5GC. We now briefly describe the components of 5GC. Unified Data Management (UDM) handles the Subscription management, user registration, user data and mobility management. The Access and Mobility Management Function (AMF) is responsible for data integrity, ciphering, user authentication and authorization along with mobility management. The session management function (SMF) helps in allocation of UE IP address and maintenance and also selecting the UPF for roaming and controlling tasks. The Policy Control Function (PCF) enforces the required policy rules as part of the

Control Plane (CP) functionality while the Application Function (AF) handles the policy framework. The external connection points to the data network are controlled by the User Plane Function (UPF).

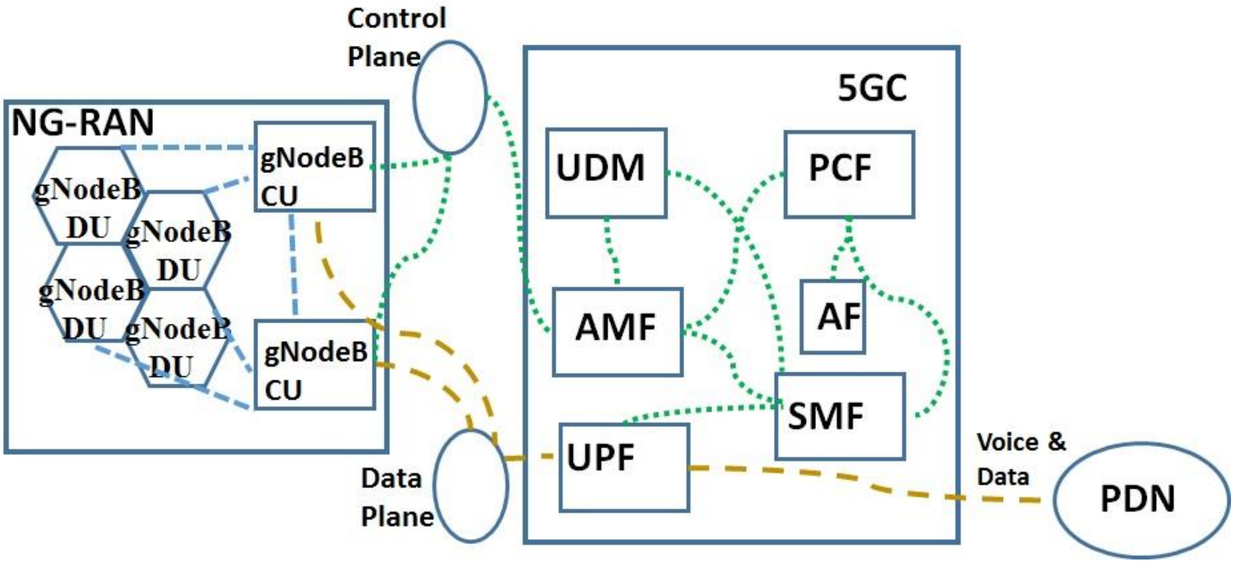


Fig.2.5. Block Diagram of 5G NR Network Architecture

2.3 Detailed Overview of the 5G System Architecture

The entire 5G system is based on densely connected wireless UEs and devices from diverse applications like health, transportation, industry etc. which demands certain challenging requirements. To support more effectively, the 5G services have been broadly classified into three generic services: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC) and ultra-reliable low latency communication (uRLLC). In very brief description, eMBB targets for stable connections while maintaining very high data rates and moderate data rates for cell-edge users. mMTC includes a massive number of Internet of Things (IoT) devices which are intermittently active, based on requirements, and transmit low-weight data payloads. Any kind of low latency transmissions of small but highly reliable payload targeted for limited terminals, form the uRLLC services. In 3GPP Release 18 [11], 5G new radio air interface were published. Frequency Range 1(FR1) ranges from 410 MHz to 7.12 GHz, while Frequency Range 2(FR2) ranges from 24.25 GHz to

71 GHz – this signifies the millimeter wave application [11][12]. The frequency ranges of 3.3 GHz to 4.2 GHz, 4.4 GHz to 5 GHz and 700 MHz are of special interest in 5G network deployment. As per ITU vision document 2083 (ITU-R M.2083-0) we provide a comparative study of requirements between 5G and 4G in table 2.1 for a general Bandwidth requirement of at least 100 MHz to 1 GHz for FR2 frequencies. Note: DL means Downlink in Table 2.1.

TABLE 2.1: Requirement Comparison between 4G and 5G

PARAMETER NAME	5G	4G
Peak Data Rate (DL)	20 Gbps	1 Gbps
User Experienced Data Rate(DL)	100 Mbps	10 Mbps
Spectrum Efficiency	3x	1x
Mobility	500 km/hr	350 km/hr
Latency	1 ms	10 ms
Connection Density (devices/km^2 – for mMTC)	10^6	10^5
Network Energy Efficiency	100x	1x
Area Traffic Capacity(Mbps/m^2)	10	0.1

2.3.1 5G Non-Stand Alone and Stand-Alone Architecture

The framework of 5G is flexible to support a vast variety of applications. The 5G architecture is functionally divided into Core Network (CN) and Radio Access Network (RAN). The concepts of Network Slicing (NS), Software Defined Networking (SDN) and Network Function Virtualization (NFV) enables the 5G architecture to remove the constraint of always requiring the base stations (BS) to be in close proximity. The 5G Mobile System Architecture consists of the following components as per the 3GPP standard, as portrayed in Fig. 2.6 [36].

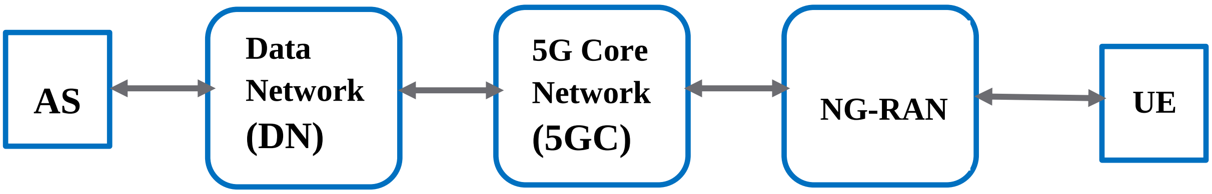


Fig.2.6. 5G System End-to-End Architecture

The New Radio (NR) interface is the key characteristics of the 5G mobile network, standardized by 3GPP. It is the 5G air interface expected to be more faster supporting various mobile technologies across industries. The NR will support communications between transceivers functioning at frequency bands of 52.6 GHz. 5G System Architecture includes both Non-Standalone (NSA) architecture as well as Standalone (SA) architecture [36].

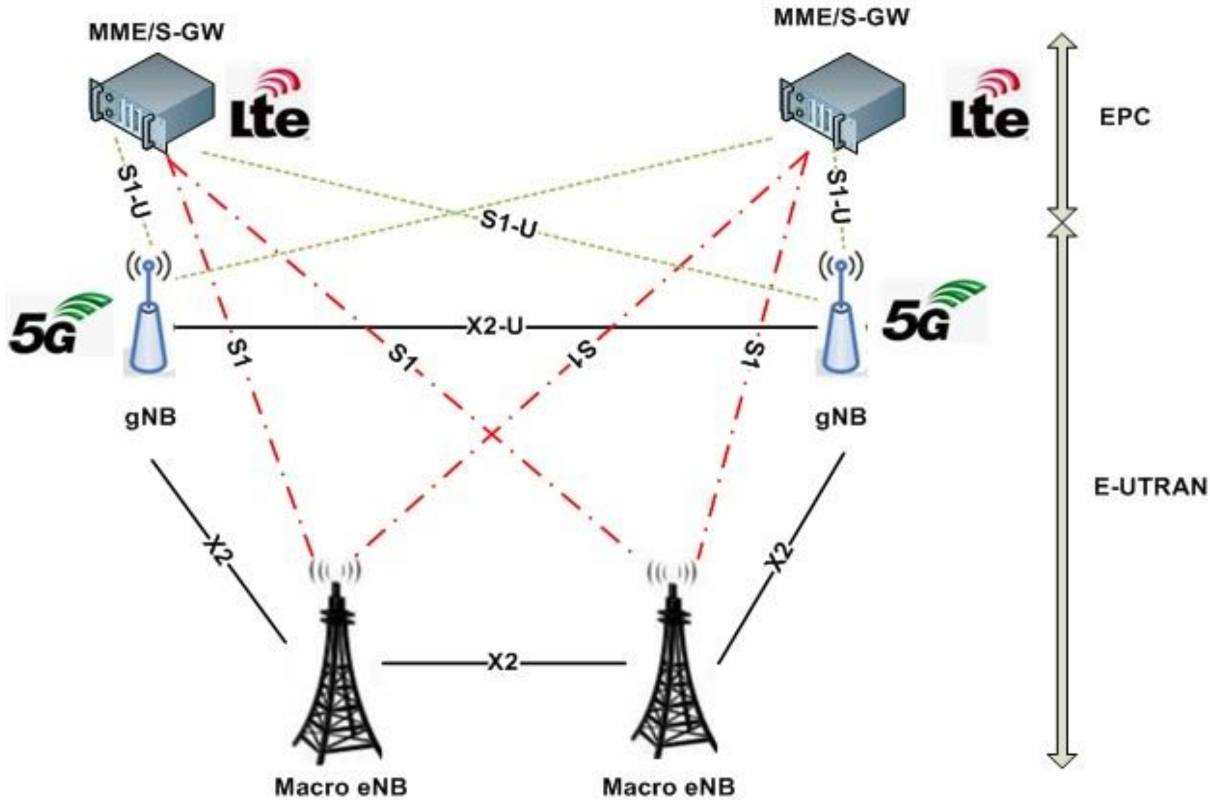


Fig.2.7. 5G Non-Stand Alone Architecture

For backward compatibility NSA suffice the requirement. It can connect NG-RAN to both 4G (LTE) and 5G core network. SA suffice the pure 5G requirements. It connects the 5G Core Network and NG-RAN alone [36].

We provide a comparative study between NSA and SA Architecture [36], in table 2.2, wherein both architectures have been diagrammatically portrayed in Fig. 2.7 and Fig. 2.8. In our thesis work 5G SA architecture has been referred where 5G RAN and LTE services are solely provided by the 5G System Model prototype.

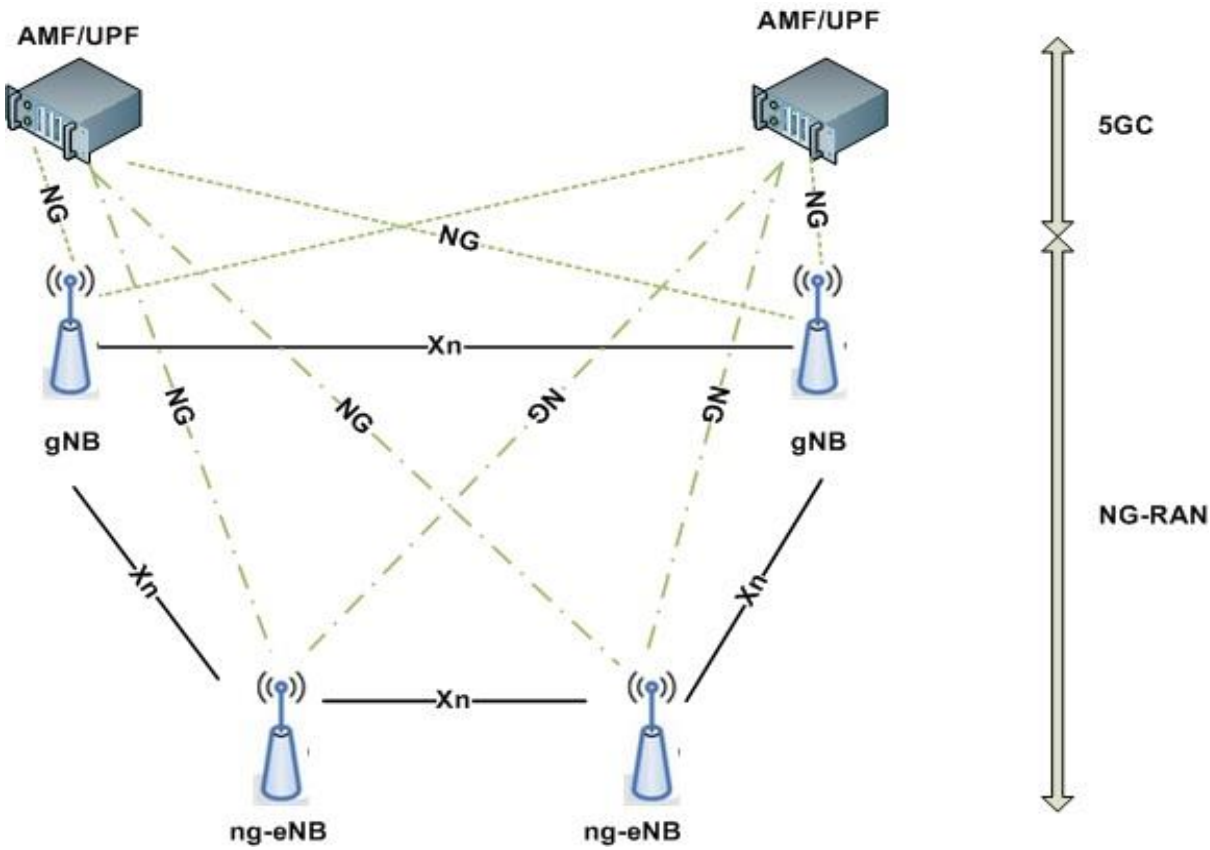


Fig.2.8. 5G Stand Alone Architecture

TABLE 2.2: Comparative Study between 5G NSA and SA

5G Non-Stand Alone Architecture	5G Stand Alone Architecture
1) To ensure 5G NR technology without immediate network replacement, NG-RAN and NR interface is used along with existing 4G Access Network (AN) and 4G Core Network (CN)	**1)** 5G NG-RAN alone will be providing both LTE and NR services.
2) The BS of 5G NR i.e. gNB connects to the BS of 4G-LTE (eNB) through the X2 interface. The Control Plane (CP) and User Plane (UP) functions of LTE towards UE will be fulfilled by eNB while gNB will be delivering the CP and UP functions of 5G-NR.	**2)** The 5G NG-RAN node consists of gNB and next generation eNB (ng-eNB). The Control Plane (CP) and User Plane (UP) functions of LTE towards UE will be fulfilled by ng-eNB while gNB will be delivering the CP and UP functions of 5G-NR.

3) The S1 interface connects the eNB to the MME/S-GW of EPC (as discussed in section 2.2.4) while S1-U interface connects the gNB to the EPC. X2-U is the interface that connects gNB to gNB.	**3)** The Xn interface interconnects both gNBs as well as ng-eNBs. The NG interface connects ng-eNBs and gNBs to the 5GC. NG-C is the interface between 5GC and the Access and Mobility Management Function (AMF) while NG-U is the interface between the 5GC and the UP function.

2.4 Machine Learning Techniques

Machine Learning (ML) models learn using the discriminative features about a system and there by establish an inherent computational model. These learning can be based on classification, regression, clustering or interaction between an agent and environment [15]. Once accomplished the model can effectively work on any unknown data. In context of wireless communication, ML can thus provide proper network configuration with enhanced QoS and throughput having the maximum coverage for any new and unknown geographical area. The field of ML covers three important aspects:- a) Supervised Learning b) Unsupervised Learning and c) Reinforcement Learning. We will be discussing each aspect in sections 2.4.1, 2.4.2 and 2.4.3 respectively.

2.4.1 Supervised Learning

The ML model that gets trained using labelled input training data and the model can estimate the required output. Labelled input data means it is already mapped to the correct output data. When from a given set of data prediction of target numerical value is required then Supervised Learning (SL) is best. Based on prior experience, Supervised Learning's output can be optimized and hence best suited for problems of classification and regression. But using SL unknown information cannot be extracted. Also with increasing data sets its accuracy decreases. Fig.2.9 gives the visual aspect of Supervised Learning.

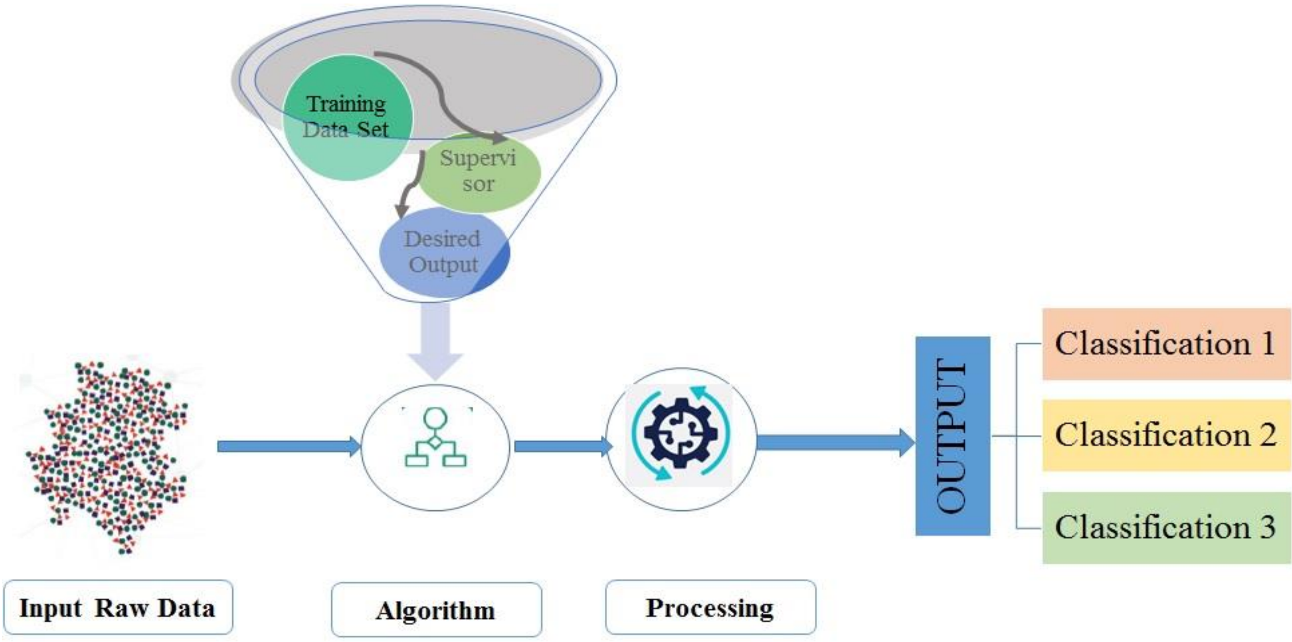


Fig.2.9. Supervised Learning

Mathematically supervised learning [15] is illustrated in table 2.3.

TABLE 2.3: Mathematical Model of Supervised Learning

Consider Training data Set of N samples $\tau = \{(x_i, y_j)\}_{i,j=1}^{N}$
$f(\tau)$ is the SL model which produces output 'O' for each τ: (x_i, y_j)
Input Raw Data τ' is applied to $f(\tau)$ such that $O(\tau) \longleftarrow$ M = $f(\tau') \longleftarrow \tau'$ $O(\tau)$ is returned with highest score value and hence is the required output

Few examples of SL techniques are Decision Trees (DTs), Support Vector Machine (SVM), K-Nearest Neighbors (KNN) etc. In our thesis work we have intelligently applied SL technique KNN while calculating the interference of our wireless system model. Hence, we put forward a brief discussion on KNN.

2.4.1.1 K- Nearest Neighbors (KNN)

The KNN algorithm is the SL technique which groups individual data point based on proximity. It is a typical classification algorithm. It studies the similarity between any new data point and other existing data points and classifies to the most suitable category.

Algorithm 1: *KNN Algorithm* [19]

Step 1: Specify the number *k* of the neighbours
Step 2: Based on distance measurement (as eg:- Squared Euclidean Distance, Euclidean Distance etc.), calculate the required distance for *k* number of neighbours
Step 3: On receiving the output in step 2 consider *k* nearest neighbours
Step 4: Once step 3 is done, count the number of the data points for each category.
Step 5: Any new data point is assigned to that category for which the number of the neighbor is maximum.
Step 6: This methodology , Step 1 to Step 5 , keeps applying for all future data points

- **Application of KNN in System Model**

As mentioned in section 2.4.1 we have used KNN intelligently in our thesis work while calculating the interference in our system model. To understand the same, let us consider a random user at random location marked by (x,y) co-ordinates in Euclidean plane positioned at a certain distance from its serving BS in 5G HetNet. Other BSs will be treated as non-serving interferers and will be part of interference towards the particular user.

As per our system model, Interference Management will include co-tier and cross-tier interference [37]. Cross-tier interference involves the interference caused by the co-channel bearer small BSs and macro BSs. Thus, 5G Heterogeneous Network includes small cells for several benefits like lesser power consumption, lesser cost, reduced load on Macro Base Stations etc. but at the additional cost of cross tier interference. Thus, as an example for a macro user, nearby small base stations will add on to this interference. Small base stations located at larger distance will not contribute any significant interference due to long distance. [37][38]

Co-tier interference involves the interference caused by the co-channel bearer of base station of same class (like small BS or macro BS) [37]. As example, for a small BS user, small BSs deployed in adjacent macro cells will contribute to this interference as part of co-channel interference [37][38]. Both these types of interferences have been depicted pictorially in Fig.2.10. These play a major role in the calculation of the system SINR which is the governing parameter of the system's performance. It is the cross-tier interference where KNN plays a vital role in mathematical calculation. Once the system has been trained based on primary parameter Euclidean distance, in a particular coverage sector, specifically which small/femto BSs fall in the same sector as the macro BS can be easily derived by KNN as shown in table 2.4.

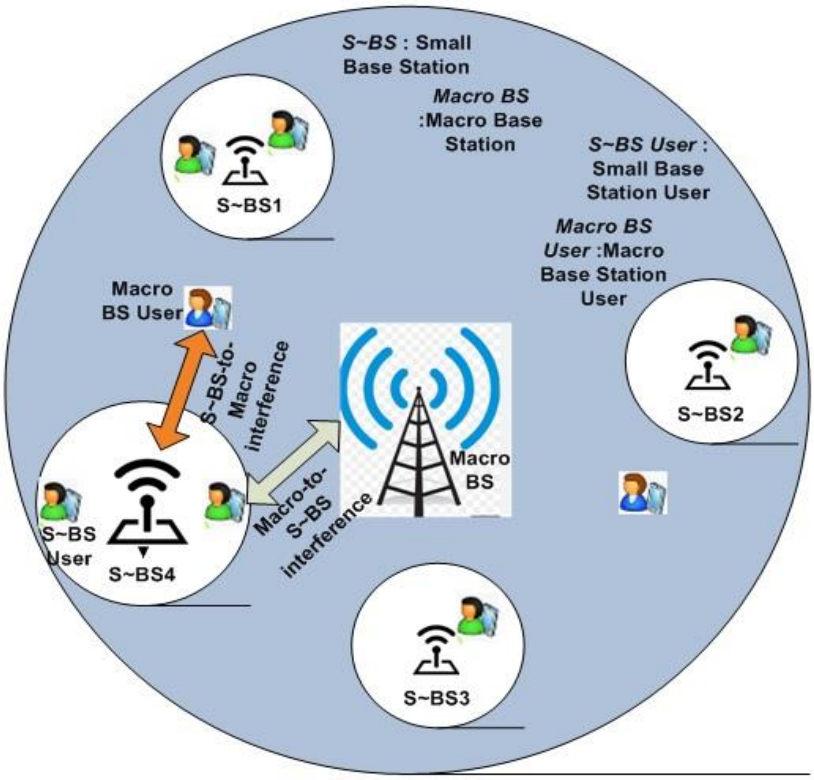


Fig.2.10. 5G Heterogeneous Wireless Network exhibiting interference

2.4.2 Unsupervised Learning

The ML model that acts on the raw input data without any supervision or guidance and groups the unsorted information or data into clusters based on features without any prior training, is termed as Unsupervised Learning (uSL). Clearly it doesn't have access to any labelled data to learn. uSL works with real time stochastic data which makes it very suitable for Wireless Communication where the channel is not at all deterministic. uSL is extensively used for Cluster Analysis for identifying the hidden patterns [15]. It is undue to state that uSL will be computationally complex. Diagrammatically uSL is demonstrated in Fig.2.11.

TABLE 2.4: Mathematical Model of Applied KNN Supervised Learning in Thesis Work

Step 1: After Training Phase 1 based on only Euclidean distance by UL, let the co-ordinates of Macro Base stations be $\{M(x_i, y_j)\}$ and that of Small BSs be $\{S(x_i, y_j)\}$
Step 2: $f_{KNN}(\tau)$ is the KNN SL model which produces output 'O' for each τ: (x_i, y_j) for every value of $\{M(x_i, y_j)\}$ and $\{S(x_i, y_j)\}$ based on distance measurement (as eg:- Squared Euclidean Distance or Euclidean Distance)
Step 3: On receiving output $O_{KNN}(\tau)$ from Step 2 , k nearest neighbours gets evaluated i.e. small BSs which are closest to the Macro BSs respectively

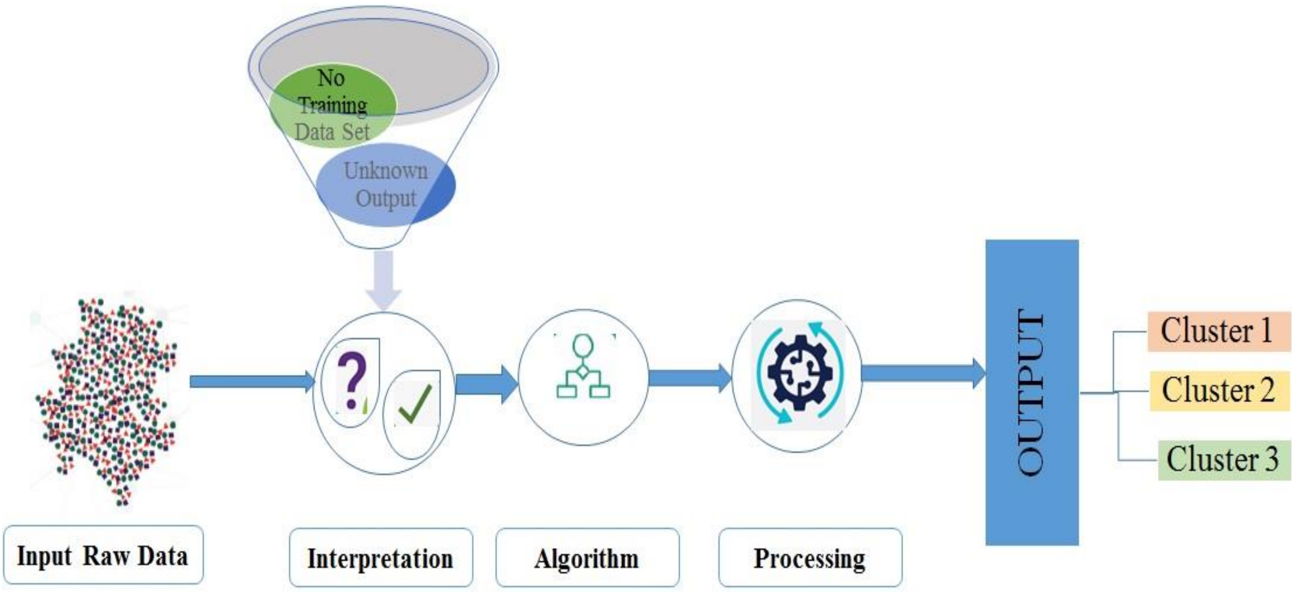


Fig.2.11. Unsupervised Learning

uSL techniques include K-Means Algorithm, Principal Component Analysis (PCA), Self Organizing Map (uSL neural network) etc. We have extensively used uSL technique in our thesis work because of its capability to work in unknown stochastic environment resembling the practical wireless channel in wireless communication.

Mathematically Unsupervised learning [15] is illustrated in table 2.5.

TABLE 2.5: Mathematical Model of Unsupervised Learning

Consider Input raw data Set of N samples $\tau_{raw} = \{(x_i, y_j)\}_{i,j=1}^{N}$
$f(\tau_{raw})$ is the function of the uSL model (M) which considers 'a' (constant) as the starting point of the algorithm. $f(\tau_{raw}) = M\ considering\ 'a' \rightarrow\ starting\ point\ of\ the\ algorithm$
If 'O' is the desired output then the expected output will then be representing the cost function of the algorithm defined as :- Expected Output = $E[(O - f(\tau_{raw}))^2]$

Among others particularly we have chosen the uSL neural network Kohonen's Self Organizing Map (SOM) for our thesis work. The reason behind it will be discussed now. Its enhanced performance is further proved when we did a comparative study between SOM and K-Means. As already illustrated in Literature Survey under section 1.3 that K-Means is a popular uSL algorithm. Hence any performance beyond K-Means will result in significant attention.

2.4.2.1 Kohonen's Self Organizing Map (SOM)

Teuvo Kohonen in 1980s introduced SOM and thus it is also sometimes called a Kohonen map or Kohonen network. Its unsupervised learning is based on competitive learning methodology resembling that of the human biological neuron. Other than clustering SOM is also extensively used in dimensionality reduction functionalities where to visualize complex problems lower dimension is a necessity. SOM is unique both in its architecture as well in its algorithmic property when compare with any typical Artificial Neural Network. At the base architecturally it comprises of a single layer of linear 2D neuron grid as opposed to series of layers. The data input vector is connected to all the nodes on this grid, but no node is connected to each other. At each iteration the grid organizes itself as a function of the input vector. At the end of learning each node has its own (x,y) co-ordinate. The grid itself is the MAP here.

In terms of algorithm or learning property, Self-Organising Map uses competitive learning to adjust the weights connected to the input as opposed to error-correction learning. When the feature of an instance of the input vector is presented to the neural network, all nodes compete among themselves to become the chosen node or the Best Matching Unit (BMU), selected based on similarity. Eventually the entire grid matches the complete input data set with similar node of neurons mapped together towards one area.

The map retains the calculated relative distance (or any other parameter) between the points. Points closer to each other within the input space are mapped to the nearby map units in Self-Organizing Maps. Self-Organizing Maps can thus serve as a cluster analysing tool. The important aspect of Self Organizing feature map about using competitive learning instead of error correction learning is there in [7].

There are two models of SOM: *Willshaw–Von der Malsburg SOM* and *Kohonen Map or SOM* [7]. Here our discussion is based on Kohonen Map or SOM**.** Kohonen SOM includes the following basic steps, as illustrated in Fig. 2.12[39]:

1. Initialization: All the connection weights are initialized with small random values. The weights in Kohonen SOM is mapped to the input data.

2. Competition: For each input pattern, the neurons compute their respective values of a discriminant function which provides the basis for competition. Hence SOM is competitive UL. The particular neuron with the smallest value of the discriminant function is declared the winner.

3. Co-operation: The winning neuron determines the spatial location of a topological neighbourhood of excited neurons, thereby providing the base for cooperation among neighbouring neurons.

4. Adaptation: The excited neurons decrease their individual values of the discriminant function in relation to the input pattern by suitable adjustment of the associated connection weights. This increases the response of the winning neuron to the subsequent application of a similar input pattern. We will be providing detailed simulation results and analysis based on SOM in Chapter 3 and Chapter 4. But before moving into detailed analysis we will present here a small snippet of Rural-Urban Wireless Tele density after discussing about SOM Algorithm.

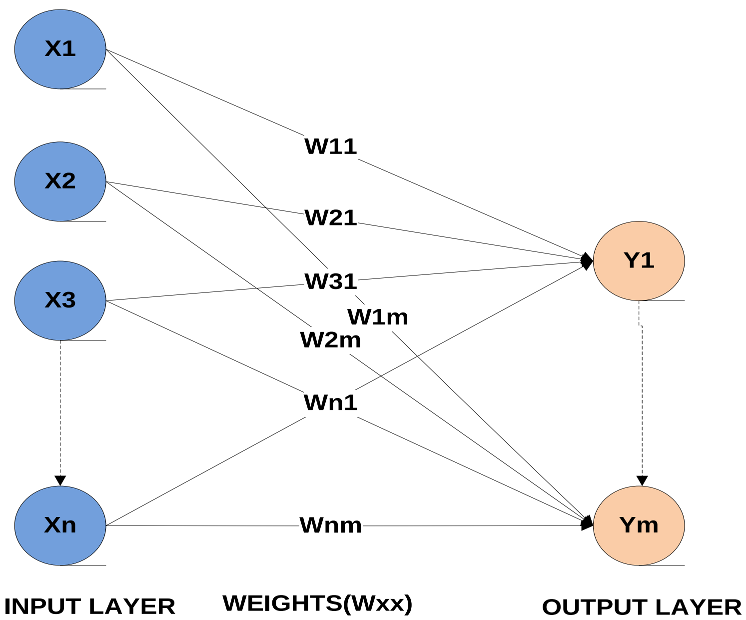


Fig.2.12. Architecture of Kohonen SOM NN

Algorithm 2: *Kohonen- SOM Algorithm* [22][40]

Step 1: Initialization – Choose random values for the initial weight vectors w_k
Step 2: Sampling – Draw a sample training input vector x from the input space.
Step 3: Competition– Find the neurons I(x) that has the weight vector closest to the input vector i.e the minimum value of $$d_k(x) = \sum_{i=1 \ to \ D}(x_i - w_{ik})^2 \quad (2.1)$$ Note: ith data is associated with kth cluster and considering D input units where dk(x) is the discriminant function
Step 4: Cooperation:- If S_{ik} is the lateral distance between winning neuron I(x) and kth neighbouring neuron on the grid of neurons then topological neighbourhood is given by:

$$T_{k,I(x)} = \exp\left(-\frac{S^2_{k,I(x)}}{2\sigma^2}\right) \quad (2.2)$$

Size σ of the neighbourhood decreases exponentially with time

Step 5: Adaptation– Apply the weight update equation

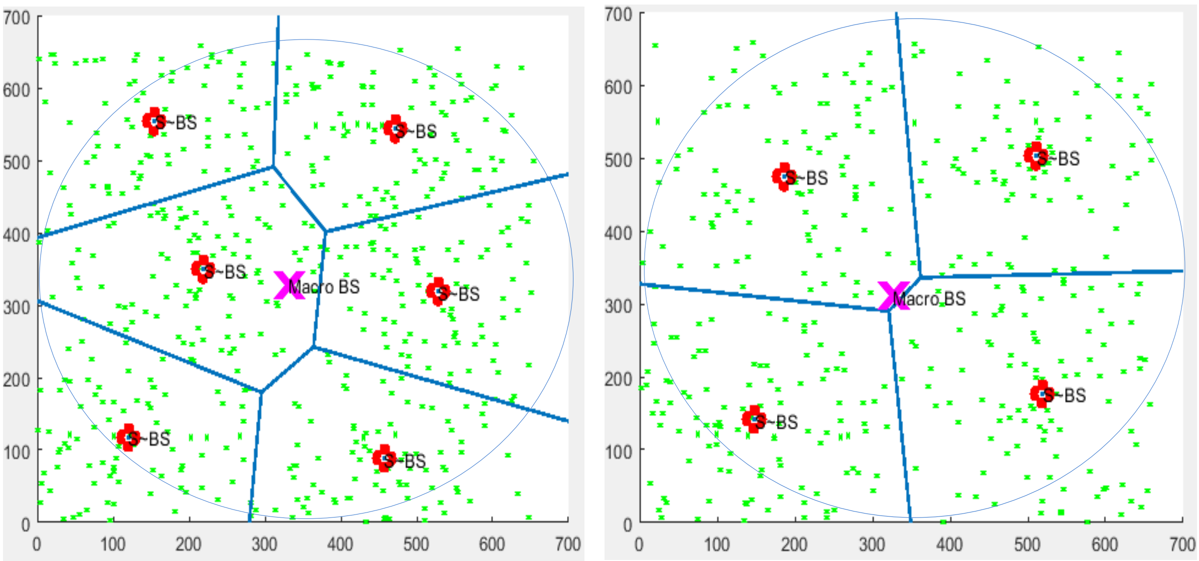
$$\Delta w_{ik} = \eta(t).T_{k,I(x)}\,(x_i - w_{ik}) \quad (2.3)$$

where, η(t) is the learning rate, this decreases exponentially with time; x_i is the input vector and w_{ik} is the current weight vector

Step 6: Continuation – Repeat from step 2 till iteration limit has been reached or weights don't change any further

- **Depiction Of Typical RURAL AND URBAN Wireless Tele-density using SOM**

Before moving to detailed analysis the visualization of tuple distribution of [Macro eNB, Small gNB , UE] over a specific area is important to know.



(a)URBAN (b) RURAL

Small gNB; Macro eNB; UE

Fig.2.13. Simulation of Coverage Regions of HetNet over 700x700 m^2 based on SOM with UE distribution in PPP.

As per Monthly Telecom Scenario [41] as published by Department of Telecom (Govt of India), dated 10th March 2021, 44.78% of wireless phone connections is part of rural India while 55.22% is part of urban India. Applying the percentage distribution on our prototype, the tuple distribution of *[Macro* *eNB, Small gNB ,UE]* is shown below for urban and rural scenarios respectively in Fig. 2.13.

2.4.2.2 K-Means

K-means clustering is an uSL method that aim at partitioning *n* observations/data points (or mobile UEs) into *k* clusters in which each observation belongs to the cluster with the nearest mean called the cluster centres (or Cluster Centroid). We have next stated the K-Means Algorithm followed by Fig.2.14 demonstrating the basic working principle of K-Means.

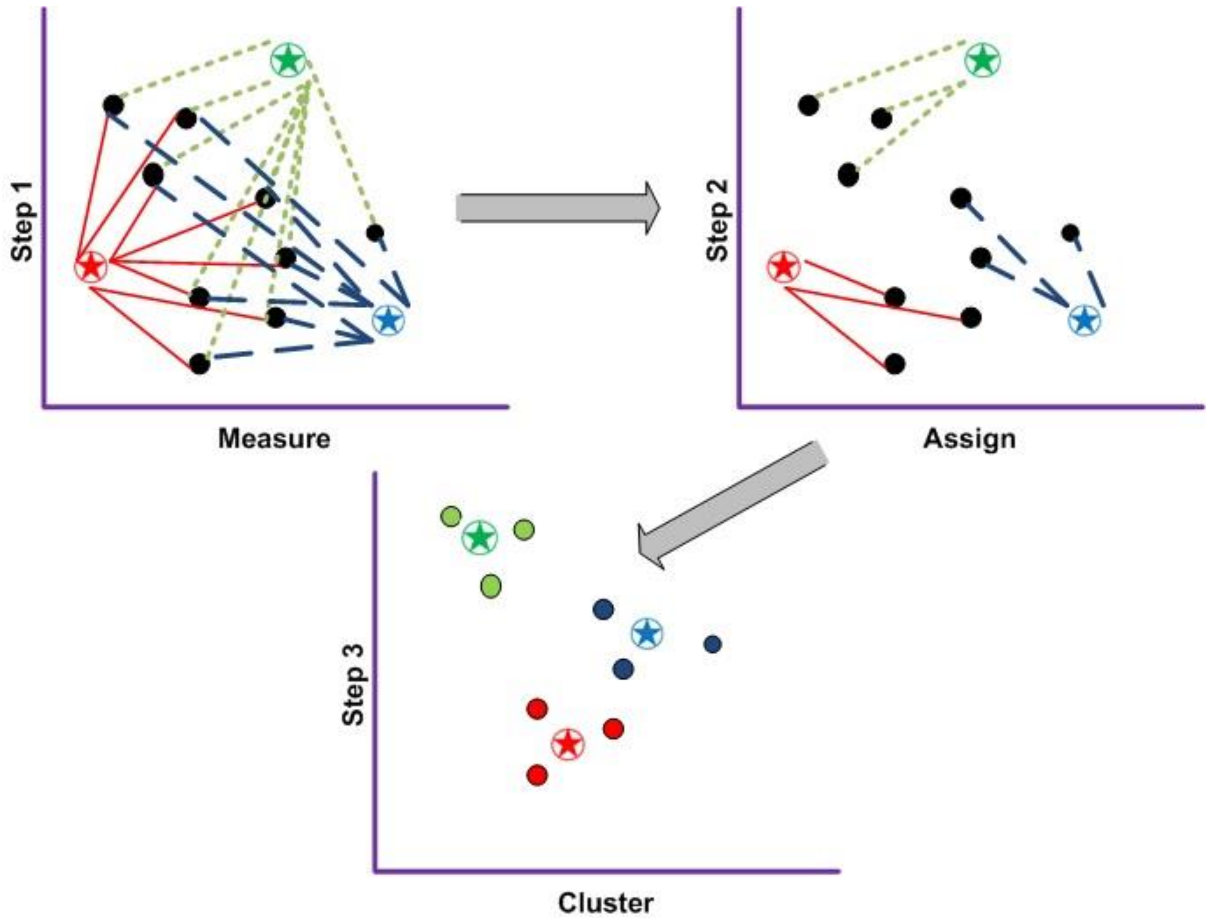


Fig.2.14. Basic Working Principle of K-Means

Algorithm 3: *K- Means Algorithm* [22]

Step 1: Specify the number *k* of clusters to assign
Step 2: Randomly initialize *k* centroids
Step 3: Step 2 is then repeated through few iterations

Step 4: *expectation*: Each point is then assigned to its closest centroid; Similarity grouping in K Means can be done based on distance measurement (as eg:- Squared Euclidean Distance, Euclidean Distance etc.).
Step 5: *maximization* : The new centroid (mean) of each cluster is the recalculated
Step 6: This methodology continues until the centroid positions do not change

2.4.3 Reinforcement Learning

Reinforcement Learning (RL) concentrates on current situation. Then makes decisions pertaining to situation. Based on the decision action is carried out. The action that bears the maximum long-term reward is considered to be the best action. Hence the goal of RL is to maximize the received reward while interacting with the environment [15].

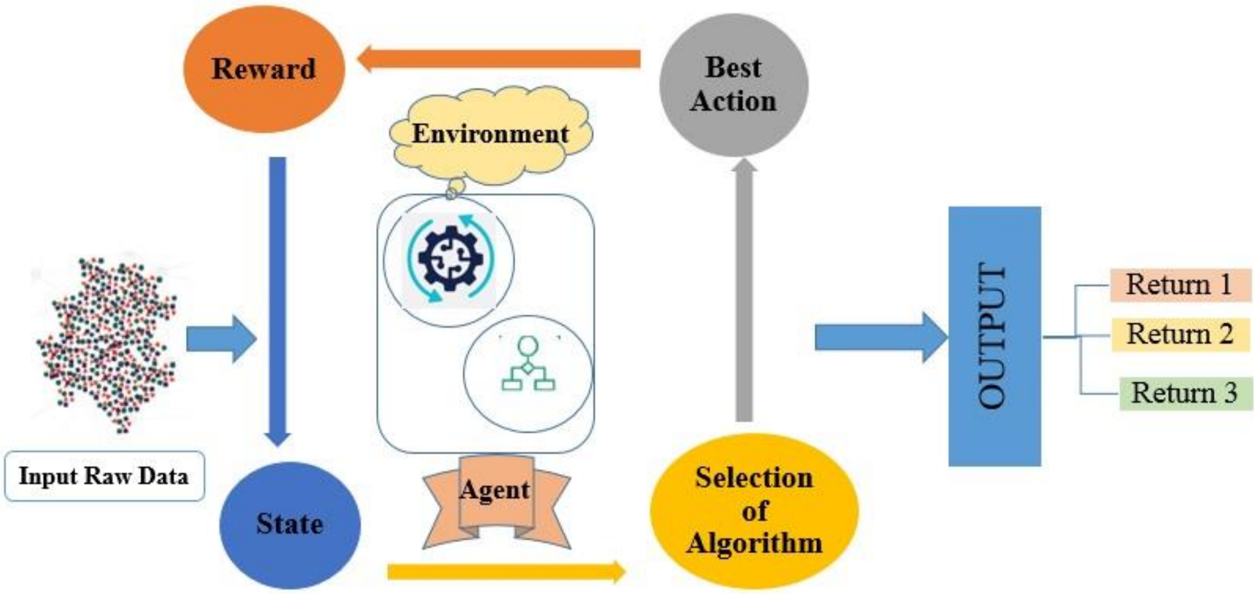


Fig.2.15. Reinforcement Learning

The learning model comprises of the following elements [42] (as also demonstrated in Fig.2.15)

a) Agent :- like the base station or transmitter in wireless communication
b) Environment :- like the channel or receiver or user equipment in wireless communication
c) State Space :- all possible environment states in a finite set

d) Action Space :- all possible actions in a finite set

e) Reward (r) :- The reward function obtained on exercising the action on the environment

RL's goal is to find the suitable policy (π) in a stochastic environment that can maximize the cumulative received reward (R) given in (2.4).

$$maximize_{\pi}\ E[R|\pi] \text{ where } R = \sum_{i=0}^{\infty} r_i \tag{2.4}$$

Mathematically we can represent Reinforcement Learning [15] as demonstrated in table 2.6.

TABLE 2.6: Mathematical Model of Reinforcement Learning

Policy Evaluation(π) : Probability that an action 'a' is taken at environment state 's' for selection of Agents is defined as: π: A X S → [0,1] $\pi(a,s) = P(a_t = a \| s_t = s)$
State-Value Function: State value (v_π) is calculated by the algorithm for estimating the maximum reward in the given state 's'. $v_\pi(s) = E[R] = E \sum_{r=0}^{\infty} \gamma\, r_t \mid s_0 = s$ where γ is the discount factor in range $0 \le \gamma \le 1$ and R = return value
Return (R) : Algorithm calculates 'R' , the return value which can be calculated as follows:- R= $\sum_{r=0}^{\infty} \gamma\, r_t$ where r_t is the reward at step 't' and γ is the discount factor in range $0 \le \gamma \le 1$

2.5 System Model Description

In our thesis work, we produce a model of 5G heterogeneous cellular network of 2 tier (j=2 where j implies number of tiers), where each tier is modeled based on the class of the base station of existing eNBs such as macro BS or gNBs, small cell BS (as eg:- femto cell). The 5G NR system is modeled in FR1 [12] under carrier frequency range of 4.1 GHz and bandwidth of 5MHz. The base stations across tiers may differ in terms of transmit power, coverage area, user handling capacity, path loss models etc. We initially build a base scenario where the macro cell users follow PPP distribution over a specific coverage area with a specific BSs density of macro BSs and small cell BSs with transmitting power as per ETSI v14.3 standard [14].

In our prototype, the mobile UEs are distributed in the Euclidean plane in the j^{th} tier according to PPP (Poisson Point Process) Φ_j of specific user intensity λ_j [18]. In any finite region A $\subseteq R^2$, the no. of points of the process denoted by $N_j(A)$ is defined [9][43] as the Poisson distributed random variable.

$$P\left[N_j\,(A) = p\right] = \frac{e^{\left[-\mu_j(A)\right]} \cdot \mu_j(A)^p}{p!} \tag{2.5}$$

where p = 1,2,3,……

and with mean rate function in (2)

$$\mu_j(A) = \iint_A \lambda_j\,(x, y)\,dx\,dy \tag{2.6}$$

Any mobile UE can only be served by BS of j^{th} tier if and only if the received downlink SINR with respect to that BS is greater than the threshold SINR of the system.

The BSs are laid over the mobile user points using the concept of Self Organizing Map – the unsupervised Neural Network, and clustering of the mobile UEs along with the BSs as centroids are accomplished based on maximum downlink SINR. This leads to maximum coverage to mobile users in a particular area. We have intelligently used the concept of KNN Search (Supervised Learning) while calculating the cross tier interference in our system model. These Machine Learning concepts have been discussed in detail in Sec.2.4.

Chapter 3 Performance Analysis of SOM For Static UEs

3.1 Introduction

A User Equipment (UE) is any device that is used for direct end user communication. It can be a hand held phone or a mobile broadband adapter equipped laptop. As demonstrated in Chapter 2, in Universal Mobile Telecommunication System (UMTS) this UE connects to the base station Node B while in 3GPP LTE to the eNode B as per the ETSI 125/136 series and 3GPP 25/36 specification series respectively. In 5G UE includes mobile phones as well as mobile devices like IoT devices. These will have in-built MIMO antenna technology supporting mmWave frequencies as well.

When a user doesn't move while carrying the mobile equipment or while having a call, then it will count for a Static UE. Also in 5G say IoT sensors detecting the water level status in any farming land, will also be accounted as Static UE. In this chapter we will be studying the enhanced performance of SOM while deploying the base stations (Macro BSs and small BSs) while supporting the Static UEs at different system thresholds and tele-density in a definite geographic area.

3.2 System Model

Our System Model is depicting the 5G wireless network where two tiers have been employed in the Heterogeneous Network. Tier 1 is Macro BSs overlaid on Static UEs while Tier 2 is Small/Femto BSs overlaid on Static UEs but under laid in existing Macro BSs' architecture. The SOM employed deployment of the BSs undergoes two Training Phases.

- ***Training Phase I***: In phase I, the BSs are first deployed based on a very typical but basic distance measurement i.e Euclidean measurement. This accounts for the required 2D distance between the Static UE and the serving BS. This 2D distance serves as the input to the next Training Phase II
- ***Training Phase II*** :

In Training Phase II, we will primarily producing a SOM re-clustering based on SINR (based on 5G Path loss model) and study the performance of SOM on various scenarios. The 2D distance thus obtained in Training Phase I is used to calculate the 3D distance which form the core of 5G Path loss Model as per the 3GPP Release 15 standards [13]. Diagrammatically we represent this final Training Phase in Fig.3.1

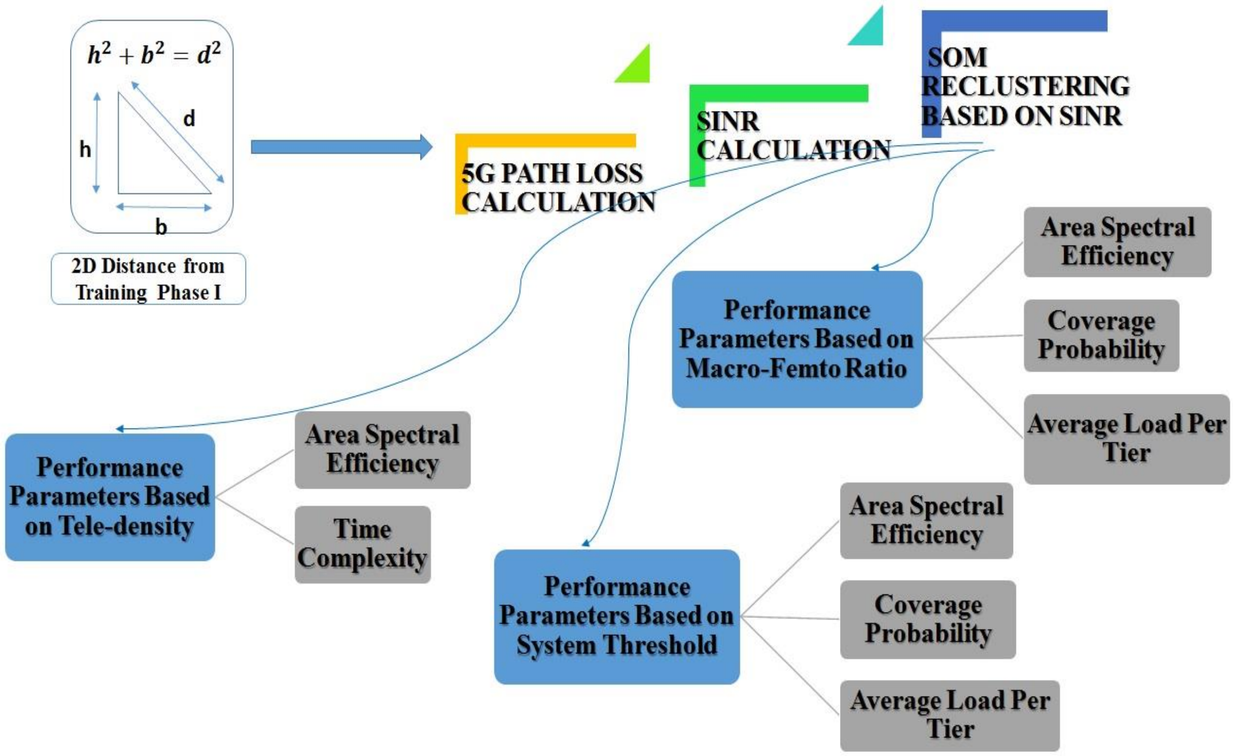


Fig.3.1. SOM Re-clustering Based on 5G SINR and Its Outcomes

3.2.1 5G Path Loss Calculation

As per [14] if the 2D distance (d_{2D}) and height of base station (h_{BS}) and height of user terminal (h_{UT}) is depicted in Fig.3.2 , then 3D distance, d_{3D} can be defined as in (3.1) –

$$d_{3D} = \sqrt{(d_{2D}^2 + (h_{BS} - h_{UT}{}^2))} \tag{3.1}$$

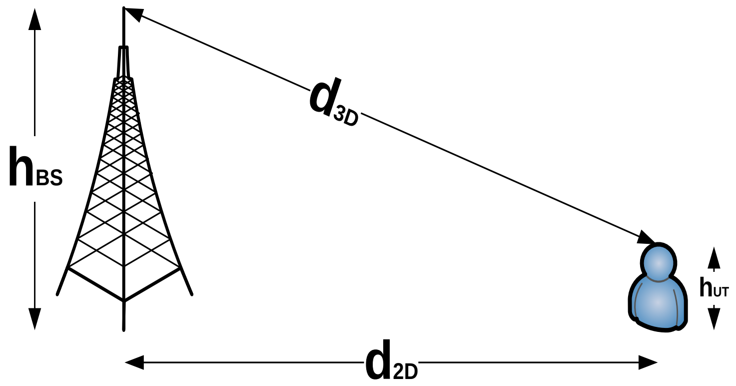


Fig.3.2. d_{2D} and d_{3D} for outdoor UTs

Breakpoint distance(d'_{BP}) as defined in [14] as

$$d'_{BP} = 4\, h'_{BS}\, h'_{UT}\, \frac{f_c}{c} \tag{3.2}$$

where f_c is the center frequency in Hz, c= 3.0×10^8 m/s is the free space propagation speed and h'_{BS} and h'_{UT} are the effective antenna heights of BS and UT defined as :

$$h'_{BS} = h_{BS} - h_E \tag{3.3}$$

and

$$h'_{UT} = h_{UT} - h_E \tag{3.4}$$

where, h_{BS} *and* h_{UT} are the actual antenna heights, while h_E is the effective environment height.

For UMi h_E = 1m.

For UMa h_E = 1m with a probability equal to $\frac{1}{(1+C(d_{2D}, h_{UT}))}$ and chosen from a discrete uniform distribution (12,15,…,(h_{UT} - 1.5)) otherwise.

$C(d_{2D}, h_{UT})$ is given by

$$C(d_{2D}, h_{UT}) = \begin{cases} 0 & , h_{UT} < 13\, m \\ (h_{UT} - \frac{13}{10})^{1.5}\, g(d_{2D}), & 13\, m \le h_{UT} \le 23\, m \end{cases}$$

where

$$g(d_{2D}) = \begin{cases} 0 & , d_{2D} \le 18\, m \\ \frac{5}{4}(\frac{d_{2D}}{100})^3\, exp\left(-\frac{d_{2D}}{150}\right), & d_{2D} > 18\, m \end{cases} \quad (3.5)$$

NOTE: A BS site can be a 1 BS or multiple co-locate BSs. Also h_E depends on d_{2D} and h_{UT} and requires to be independently determined for every BS-UT link.

Path Loss formulae in LOS scenario for UMa and UMi as per 3GPP Release 15 [13] are as stated in equations (3.6) and (3.7) respectively.

TABLE 3.1: 5G Path Loss in LOS scenario for UMa and UMi

LOS Path Loss for UMa h_{BS}=25 m; σ_{SF}= 4dB; h_{UT}= 1.5 m $PathLoss_{UMa\text{-}LOS}$ = $$\begin{cases} 28 + 22\, log_{10}(d_{3D}) + 20\, log_{10} f_c, & 10m \le d_{2D} \le d'_{BP} \\ 28 + 40\, log_{10}(d_{3D}) + 20\, log_{10} f_c \\ -9\, log_{10}\left(d'^2_{BP} + (h_{BS} - h_{UT})^2\right), & d'_{BP} \le d_{2D} \le 5km \end{cases}$$ (3.6)
LOS Path Loss for UMi h_{BS}=10 m; σ_{SF}= 4dB; h_{UT}= 1.5 m $PathLoss_{UMi\text{-}LOS}$ = $$\begin{cases} 32.4 + 21\, log_{10}(d_{3D}) + 20\, log_{10} f_c, & 10m \le d_{2D} \le d'_{BP} \\ 32.4 + 40\, log_{10}(d_{3D}) + 20\, log_{10} f_c \\ -9.5\, log_{10}\left(d'^2_{BP} + (h_{BS} - h_{UT})^2\right), & d'_{BP} \le d_{2D} \le 5km \end{cases}$$ (3.7)

3.2.2 Downlink SINR Calculation

Let us consider a random user at random location marked by (x,y) co-ordinates in Euclidean plane positioned at a certain distance from its serving BS in 5G HCN. Other BSs will be treated as non-serving interferers and will be part of Interference towards the particular user.

As per our system model, *Interference Management* will include *co-tier and cross-tier interference* [37].

Cross-tier interference involves the interference caused by the co-channel bearer small BSs and macro BSs. Thus, 5G Heterogeneous Network includes small cells for several benefits like lesser power consumption, lesser cost, reduced load on Macro Base Stations etc. but at the additional cost of cross tier interference. Thus, as example for a macro user, nearby small base stations will add on to this interference. Small base stations located at larger distance will not contribute any significant interference due to long distance. [37][38]

Co-tier interference involves the interference caused by the co-channel bearer of base station of same class (like small BS or macro BS) [37]. As example, for a small BS user, small BSs deployed in adjacent macro cells will contribute to this interference as part of co-channel interference [37][38].

We also define our strategy to be open access strategy. Open Access Strategy is when a typical mobile user is allowed to connect to any tier without restriction [23].

Considering a fully loaded model where every BS's always transmitting, along with open access strategy as defined in above para, a basic generalized form based on binary association indicator [4][44] (either 0 or 1) is realized in order to maximize the load and coverage in presence of power constraint of each class of BS and respective *target SINR threshold (γ_j^{th} in dB)* in the system. SINR threshold values have been referenced from ETSI [13]. Implementing the idea of Macro User Offloading (as referenced from [45]), binary association indicator of the system is 0 till the small BSs gNBs are able to support at a particular SINR threshold. When it goes beyond the capacity of the small BSs gNBs, the indicator is switched to 1 and Macro eNBs comes into play to support the mobile UEs under the target system SINR threshold.

Basically in downlink there are 2 major components: transmitted power (P_{tx} *in dBm*) and Pathloss (*in dB*; Expression of the same is provided in Table 3.1). Along with this Noise Power (η *in db*) and Shadow Fading standard deviation (σ *in dB*) adds on. So for any typical UE at (x,y) received power (*in dBm*) is given by [16][46]:

$$P_{rx}(x,y) = P_{tx}(x,y) - PL(x,y) \tag{3.8}$$

The received SNR then follows as below:

$$SNR(x,y) = P_{rx}(x,y) - \eta - \sigma \tag{3.9}$$

If P_{IN} denotes the average interfering power (*in dBm*) and then interfering SNR can be defined as below

$$SNR_{IN}(x,y) = P_{IN}(x,y) - \eta - \sigma \tag{3.10}$$

Considering equations (10) and (11), SINR can then be evaluated as:

$$SINR(x,y) = SNR(x,y) - SNR_{IN}(x,y) \tag{3.11}$$

3.2.3 Area Spectral Efficiency Calculation

Our system model follows the UMi and UMa LOS scenario [14] with respective transmitter height (10m for Umi & 25m for UMa). With increased *Signal-to-Interference Noise Ratio (SINR)*, provide increased capacity. As per Shannon-Hartlet theorem maximum *Capacity (C in bits per second)* of a communication system with specified bandwidth (B in Hz) in presence of noise(N) and additionally interference (I) can be related [17][38] as :

$$C = B\ log_2(1 + SINR) \tag{3.12}$$

With above equation (3.12) and referencing [16][46], *Spectral Efficiency (SE)* can then be defined as the information rate (in bps) that can be transmitted over a specific bandwidth in Hz.

$$SE = \frac{C}{B} = \ log_2(1 + SINR) \qquad (3.13)$$

As further explained in [46][47], *Area Spectral Efficiency (ASE)* is defined as the maximum throughput per unit bandwidth per unit area supported by cells BS interfered by neighboring BSs deployed in geographical area under same radio frequency. Hence, ASE (bps/Hz/m^2) is given as:

$$ASE = \frac{SE}{Coverage\ Area} = \frac{log_2(1+SINR)}{Coverage\ Area} \qquad (3.14)$$

3.2.4 Coverage Probability & Average Load Per Tier Calculation

When a mobile user manages to connect to at least one BS with SINR above its threshold (γ_j^{th}) then the mobile user is said to be in coverage. With same SINR threshold across tiers, coverage probability is defined as the complementary cumulative distribution function (CCDF) of the effective received SINR, where outage being CDF i.e 1-CCDF [23][9].

The expression [23], [9] of coverage probability (P_c) is given as below:-

$$P_c\left[P_{tx}(x,y), \lambda_j, \gamma_j^{th}\right] = \lambda_j \iint_A P[SINR(x,y) > \gamma_j^{th}] \qquad (3.15)$$

In our prototype we consider that all UEs are active at any instant of time. This is done to demonstrate the maximum capability of the prototype extended for all use cases defined in Chapter 3 and Chapter 4. Since we have two tiers (tier j=2) in our prototype, hence we calculate the number of active users supported at any instant of time by either tier. Also since we incorporate the concept of Macro-User Offloading, hence first the small BS tier steps to support any active user for particular system threshold and next Macro BS tier steps forward when small BS tier is unable to support. Now active UEs might account for both new calls (ϑ_n) as well as hand-off (ϑ_h) calls. The total offered load (ϑ) to any system is given by the sum of new and handoff calls that have arrived [16] :

$$\vartheta = \vartheta_n + \vartheta_h \tag{3.16}$$

If P_B = the call blocking probability where a new call attempt is not accommodated and is given by (Number of new call blocked + Number of new call timed out)/(Number of new call arrived)

and P_F = the forced termination probability defined as (Number of hand-off call blocked + Number of hand-off call timed out)/ (Number of hand-off call arrived) then Grade of Service (GoS) [16] is the penalty cost function of a wireless system which accounts for more penalty if handoff forced termination takes place and lesser penalty if new call gets blocked. It is defined as:

$$\text{GoS} = P_B + 10\, P_F \tag{3.17}$$

As opposed to the offered load in (3.16) the amount of traffic carried by the wireless network is defined as the total carried traffic [16] which is further defined in (3.18)

$$\text{Total Carried Traffic} = [(1 - P_F)\vartheta_h + (1 - P_B)\vartheta_n] \tag{3.18}$$

3.3 Performance Analysis of Simulated System Model

We now put forward the simulation results studied after re-clustering based on SINR was performed, as mentioned in section 3.2. The required parameters and their respective values are provided in Table 3.2.

TABLE 3.2: Simulation Parameter's Values

PARAMETRIC NAME	PARAMETRIC VALUE
Operating Carrier Frequency	4.1 GHz
Bandwidth	5 MHz
Coverage Area	4 km^2
No.of Macro BSs	3
No.of Small/Femto BSs	15
Transmitting power of Macro BS	46 dBm
Transmitting power of Femto BS	20 dBm
No.of Static UEs	Varied by PPP

3.3.1 Simulation Results of SOM Based System on varying Macro : Small BS Ratio in 5G HetNet

An analysis was conducted for specific area on varying the count of femto/small BSs deployment (*using SOM*) under 1 macro BS for a specific UE PPP distribution. This demonstrates how the Area Spectral Efficiency and Load Per Tier along with Coverage Probability varies (*Ratio of Macro eNB to Small gNB* were as follows: *1:4; 1:6; 1:8; 1:10; 1:15*). Simulation Graphs are provided as in Fig.3.3, Fig.3.4, Fig.3.5 and Fig.3.6.

As observed with rise in small BSs, at a particular SINR threshold leads to corresponding decrease in *Area Spectral Efficiency*. This obviously happens since due to rise in small BSs the Coverage Probability also increases. For any specific Macro: Small BSs ratio, Area Spectral Efficiency increases with increasing threshold, which is quite obvious as at high system threshold lesser UEs can be supported. As observed with rise in small BSs, *Coverage Probability* increases at a particular SINR threshold which naturally leads to corresponding decrease in *Area Spectral Efficiency*. Also regarding *Coverage Probability* of Macro vs Small Base Station/Femto, with respect to SINR threshold, with increasing ratio (i.e more small base stations under 1 macro BS), coverage probability of macro BS obviously decreases while of small BS increases since more offloading from macro BS takes place by the small BSs.

Another interesting observation with respect to *Load Per Tier* (tier j=2 in our prototype), at lower SINR threshold small base stations in our system model are able to share the load of macro base station more. But with increasing threshold, this percentage of share decreases and gradually macro base station provide more coverage at higher threshold. This is self-explanatory since transmit power of macro base stations is more than small base stations, hence at higher threshold capacity for carrying more load for small base stations naturally decreases.

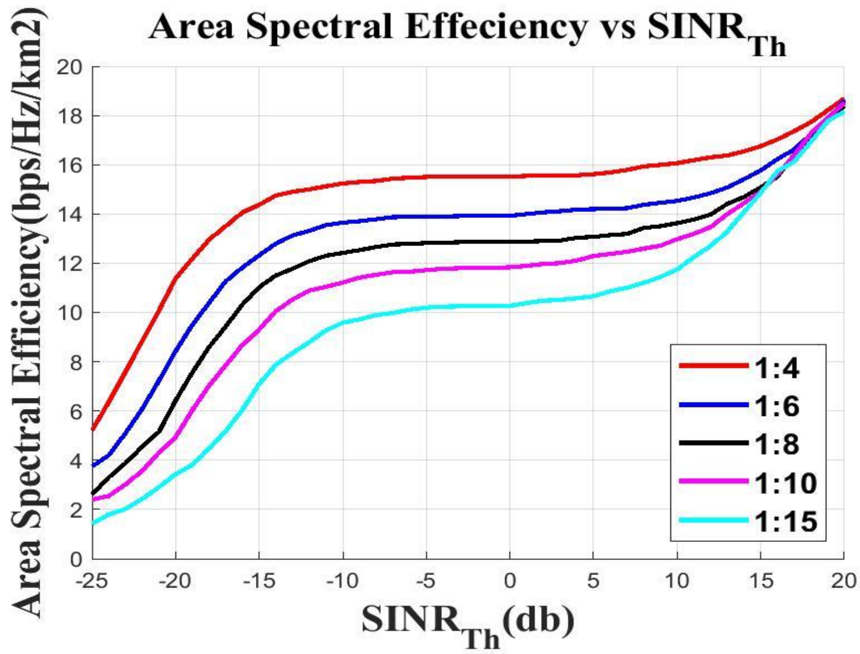


Fig.3.3. Comparison of Area Spectral Efficiency at different Macro to Small Base Station Ratio across target SINR thresholds in Urban India

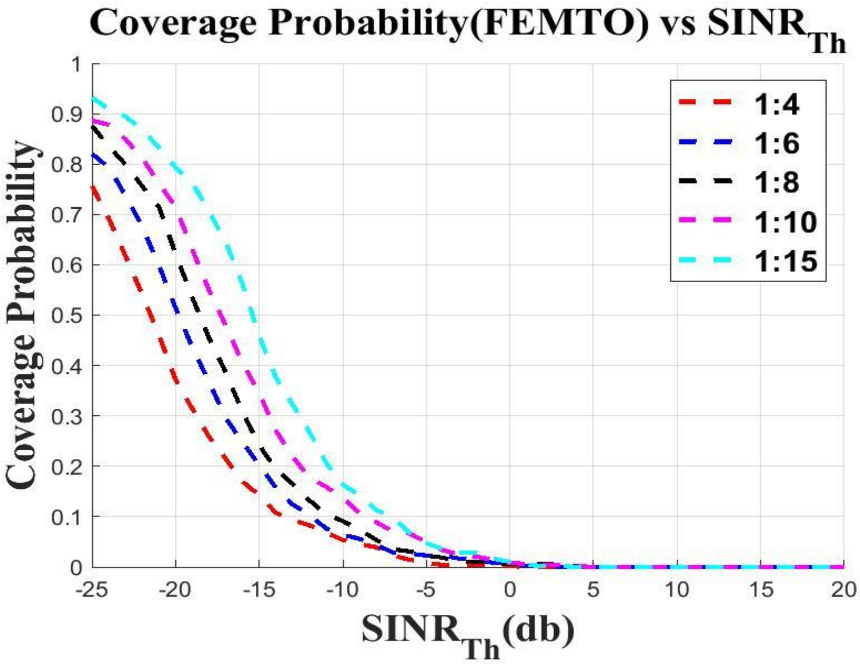


Fig.3.4. Comparison of Coverage Probability of Small Base Stations at different Macro to Small Base Station Ratio across target SINR thresholds in Urban India

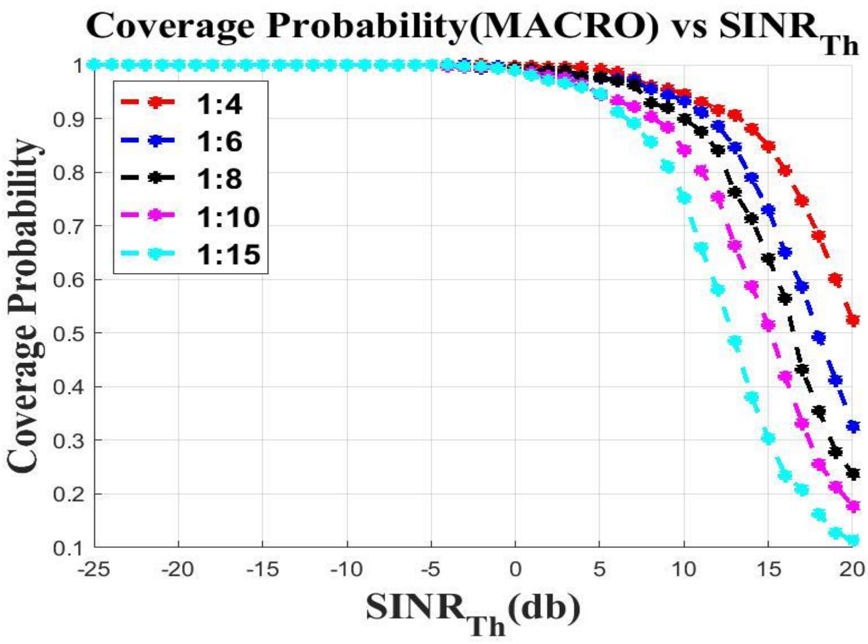


Fig.3.5. Comparison of Coverage Probability of Macro Base Stations at different Macro to Small Base Station Ratio across target SINR thresholds in Urban India

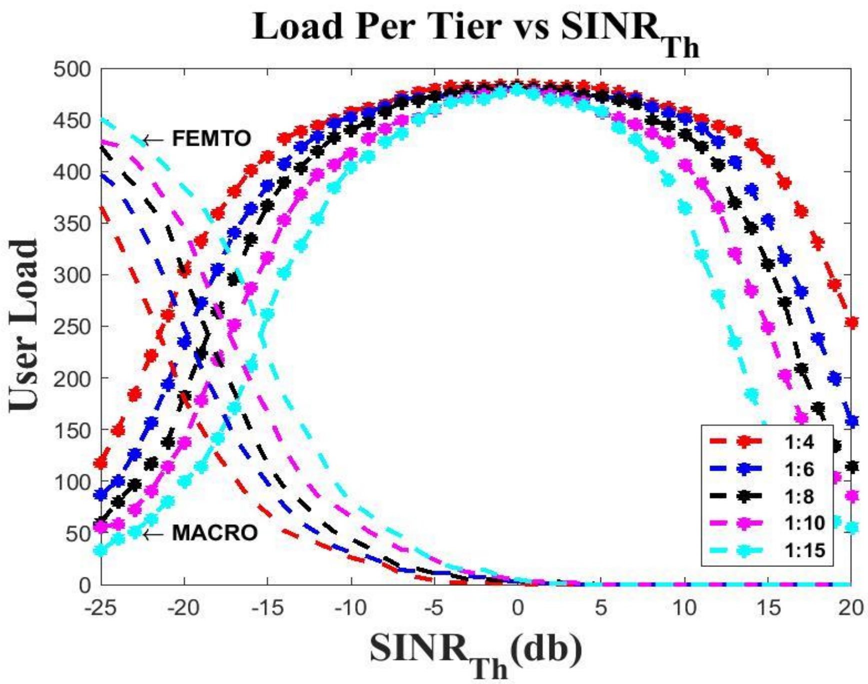


Fig.3.6. Comparison of Load Per Tier at different Macro to Small Base Station Ratio across target SINR thresholds in Urban India

3.3.2 Simulation Results of SOM Based System vs K-Means Based System Employed to Deployment of 5G HetNet

As demonstrated below there will be two models. *Model 1* projects the BS distribution using *SOM* while *Model 2* projects the BS distribution using *K-Means*. The performance parameters being Area Spectral Efficiency, Coverage Probability and Load Per Tier vs System Threshold in dB.

Both Kohonen's SOM and K-Means undergoes the two training phases defined in section 3.2. During Training Phase 1, for the same UE PPP distribution over a specific coverage area in Euclidean Plane, we portray Model 1 followed by Model 2. In this first phase the model is trained based on Euclidean distance.

In Training Phase 2 we generate a re-clustering of the models primarily based on SINR. In order to accomplish Training Phase 2 we need to calculate appropriate path loss for 5G HCN as per ETSI standard [14], described in section 3.2.1.

Considering path loss, noise, shadow fading standard deviation and transmitted power of BSs, downlink SINR is calculated, as discussed under section 3.2.2.

At Training Phase 2, based on UE position and SINR re-clustering is invoked with the help of SOM to calculate the best positions of serving small base stations leading to maximum coverage under existing locations of macro BSs.

It has been observed that in terms of *Area Spectral Efficiency* at various SINR thresholds SOM performs better than K-Means for similar user PPP distribution. Simulation Results are shown in Fig. 3.7.

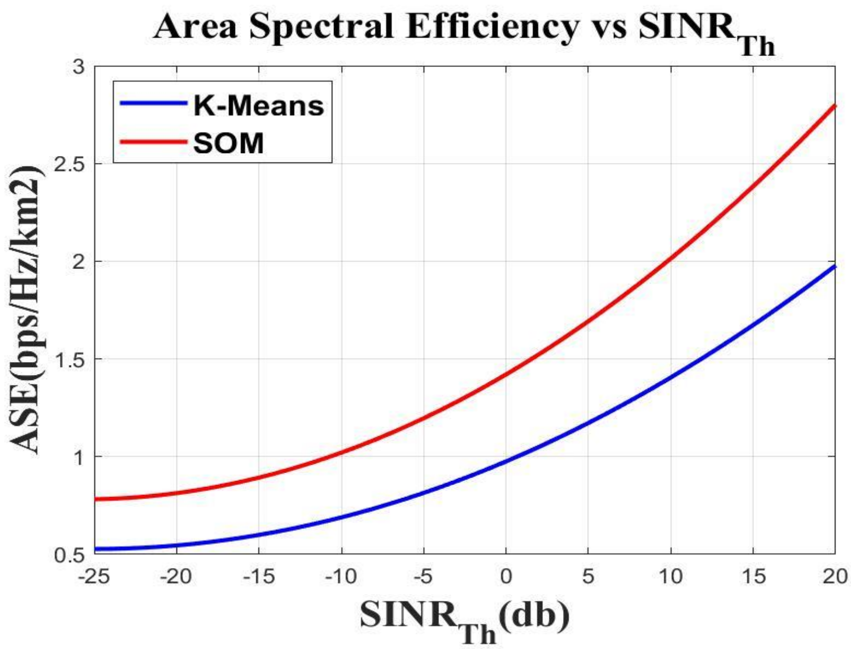


Fig.3.7. Comparison of Area Spectral Efficiency between K-Means & SOM, across different target SINR thresholds

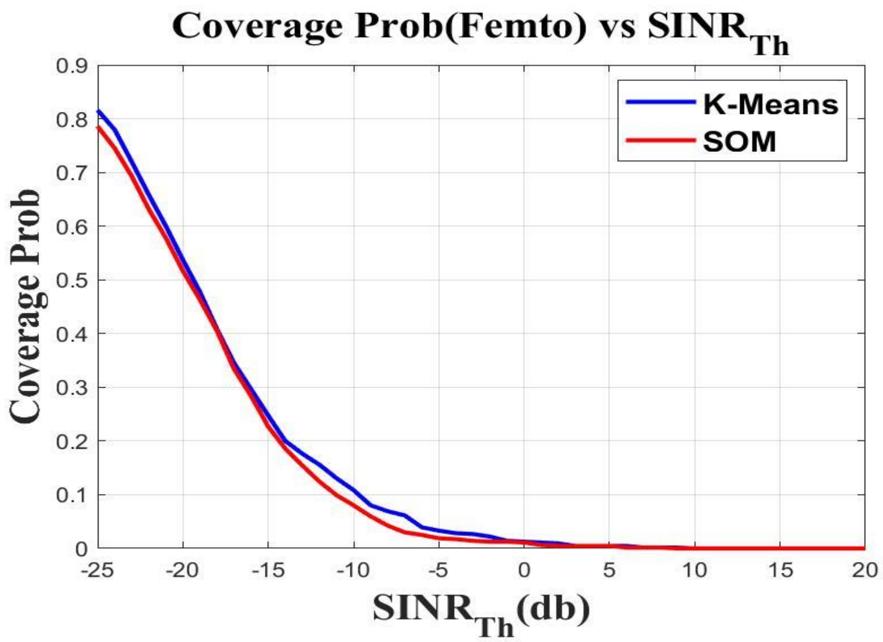


Fig.3.8. Comparison of Coverage Probability of Small BSs between K-Means & SOM, across different target SINR thresholds

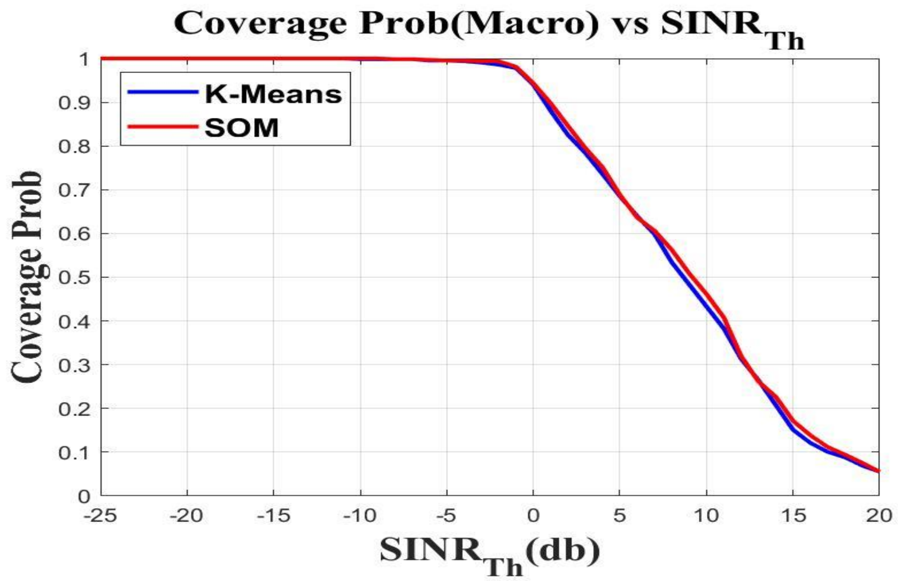


Fig.3.9. Comparison of Coverage Probability of Macro BSs between K-Means & SOM, across different target SINR thresholds

While studying the Coverage Probability of Macro & Small BSs across target SINR thresholds, we observe K-Means and SOM perform hand-to-hand efficiently. As seen in the results portrayed in Fig.3.8 and Fig.3.9, the above demonstration gets cleared. While Fig.3.10 demonstrates the macro user offloading concept.

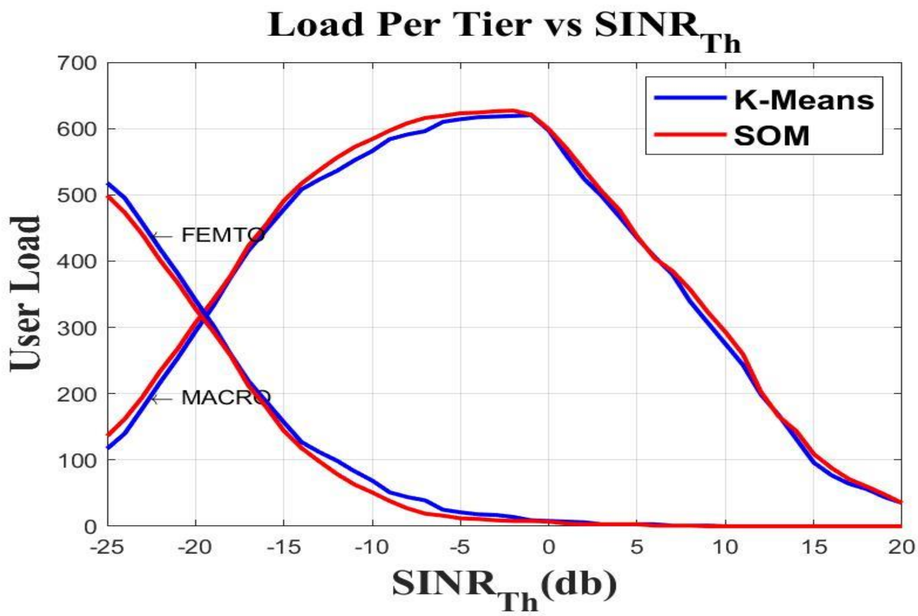


Fig.3.10. Comparison of Load Per Tier between K-Means & SOM, across different target SINR thresholds

We put forward another observation, by varying the tele density in terms of Area Spectral Efficiency. As eg:- Rural and Urban India will exhibit different tele-densities (as shown under [41]). Keeping this in mind we provide the comparison result between K-Means & SOM. Once again SOM outstands K-Means in terms of Mean Area Spectral Efficiency (refer Fig.3.11).

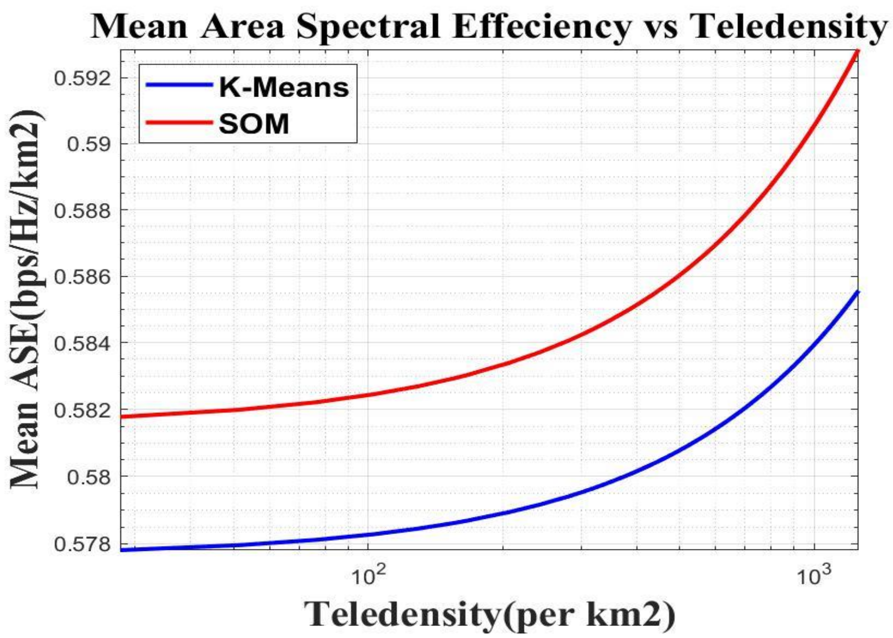


Fig.3.11. Comparison of Mean Area Spectral Efficiency between K-Means & SOM, across different Tele density

But as stated earlier SOM is unsupervised neural network. So, definitely its time computation is higher when compared with K-Means. While studying the *Time/Computational Complexity of K-Means* from [48][49] we get the expression:-

$$Time\ Complexity_{K-Means} = O(I\ K\ N\ d) \qquad (3.19)$$

where I = no. of iterations; K= no. of clusters; N = no. of input points; d = no. of attributes; O = implies the big O notation

As stated in [50], considering a SOM of *'n' neurons*, each being *'p' dimensional* and applied to input data with *d-dimension*, then for t_f learning steps and after 'iter' iterations the complexity is stated as below:-

$$Time\ Complexity_{SOM} = O(iter * t_f(d * n + p * n^2)) \quad (3.20)$$

Here again O = implies the big O notation. As demonstrated in Fig.3.12 and also with above computational expressions for time complexity, it is obvious that SOM has a higher time complexity as K-Means. But it is worth noting that as the tele density increases, the performance gap between SOM and K-Means keeps reducing, also exhibiting the performance consistency of SOM over K-Means even with increasing tele density.

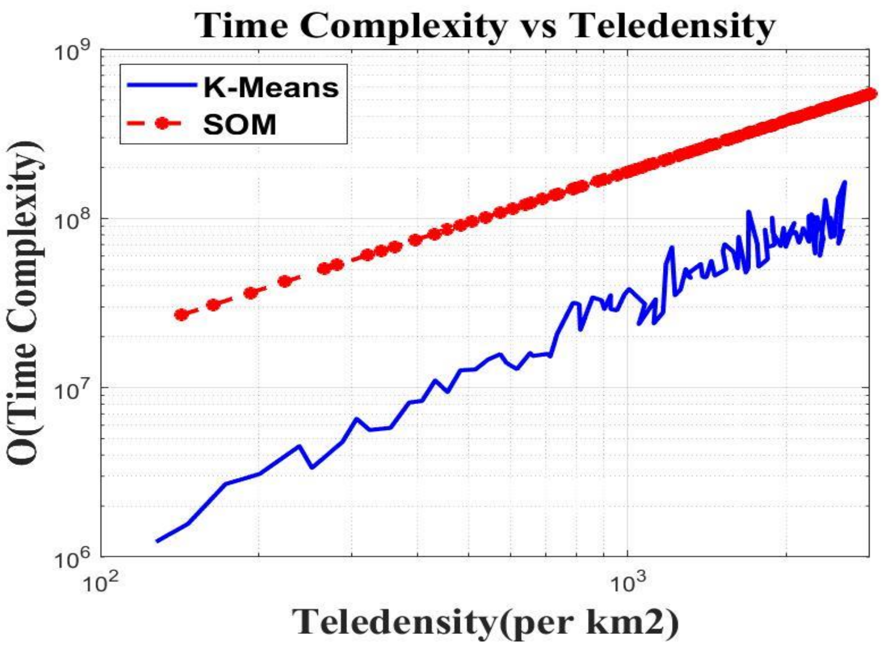


Fig.3.12. Comparison of Time Complexity between K-Means & SOM, across different Tele density

3.3.3 5th Percentile Spectral Efficiency Simulation Results of SOM Based System vs K-Means Based System Employed to Deployment of 5G HetNet Supporting Cell Edge UEs

The 5th percentile point of the cumulative distribution of the normalized user throughput is termed as the 5th percentile Spectral Efficiency. The number of bits accurately received is termed as the normalized user throughput. In a particular cell, UEs close to the base station experience a better SINR as compared UEs at the cell edge. It is also termed as the outage spectral efficiency meaning 95 percent of the time the downlink spectral efficiency is above the required value for proper communication [51].

The 5th percentile point Spectral Efficiency can be derived in 3 simple steps as follows:

Step 1: The spectral efficiencies of all UE points in a cell under a base station is first calculated

Step 2: Now plot of the cumulative distribution function of Step 1 is performed

Step 3: From the plot in Step 2 the 5th percentile point represents the cell edge spectral efficiency

When Effective Bandwidth is multiplied by frequency reuse factor then this gives the Channel Bandwidth. The Effective Bandwidth is appropriately normalized by considering the Uplink and Downlink ratio.

If $R_i(T_i)$ is the 'i' number of bits that are correctly received while T_i is the Active Session Time for ith UE and W is the effective bandwidth then the normalized user throughput r_i is given by [51]:-

$$r_i = \frac{R_i(T_i)}{W.T_i} \tag{3.21}$$

In Table 3.3 we provide the key performance minimum parameter value with respect to 5th Percentile Spectral Efficiency for minimum standard performance of IMT 2020 (5G) as defined by the ITU in ITU-R Report M.2410-0 for a general Bandwidth requirement of at least 100 MHz to 1 GHz for FR2 frequencies:

TABLE 3.3: IMT 2020 KEY PARAMETER VALUE

PARAMETER NAME	PARAMETER VALUE	
5th percentile user spectral efficiency	Dense Urban	Downlink: 0.225 bit/s/Hz
	Rural	Downlink: 0.12 bit/s/Hz
	Indoor Hotspot	Downlink: 0.3 bit/s/Hz

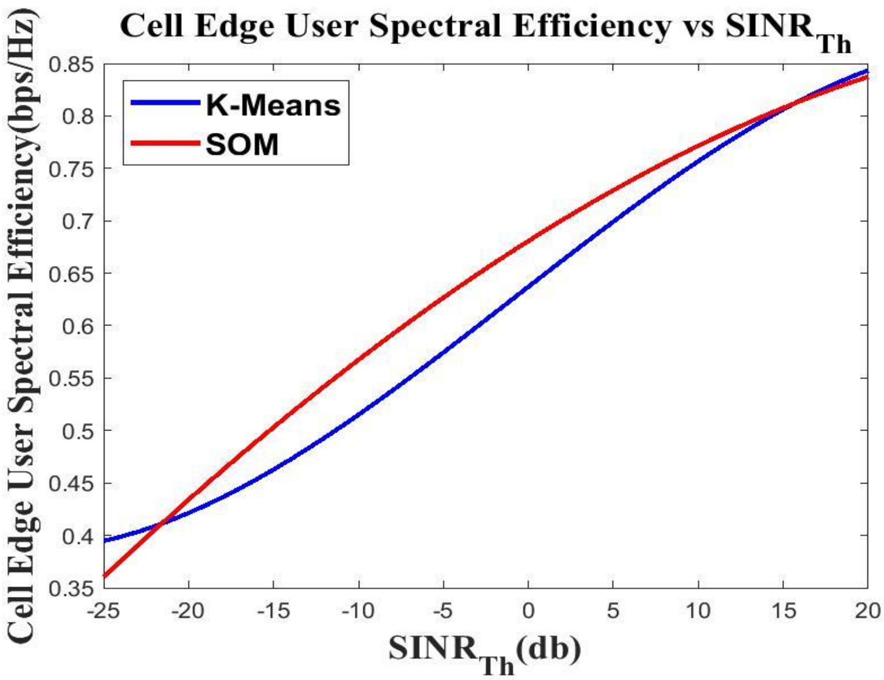


Fig.3.13. Comparison of 5th Percentile User Spectral Efficiency between K-Means & SOM, across different target SINR thresholds

As per Table 3.3 in Dense Urban scenario, minimum downlink 5th percentile spectral efficiency value is expected to be 0.225 bit/s/Hz. When compared with the simulation result in Fig.3.13. this is well met when the 5G HetNet deployment is served by Kohonen's SOM. For further significance we have compared this performance with popular K-Means algorithm. As observed for maximum of the SINR System Threshold, SOM outperforms K-Means while supporting cell edge users as well. This is a very significant and important result since supporting cell edge users marks the standard of any 5G system model.

3.4 Why SOM

Our primary focus have been on 5G eMBB service, as it is the primary service that requires early deployment to support the existing customers in the market. We have produced a comparison between SOM and K-Means primarily because K-Means is quite popular Unsupervised Learning applied in previous works, this has been discussed in detail under section 1.3 Literature Survey of the Thesis. Also as discussed under the same section 1.3, application of unsupervised neural network SOM is quite less. During our exploration of SOM we found that not only SOM performed quite well as compared to K-Means under Area Spectral Efficiency vs SINR Threshold and Area Spectral Efficiency vs Tele density, but also it worked hand-in-hand efficiently when compared with K-Means under Coverage Probability for Small Base Stations, Coverage Probability for Macro Base Stations and Load Per Tier. Also notably, Area Spectral Efficiency is the primary performance ruling factor for 5G eMBB service. Hence it is surprising that exploration of SOM is quite less wherein it can rule over K-Means effectively while training the network for the proper and quick deployment of the base stations in any specific geographic area.

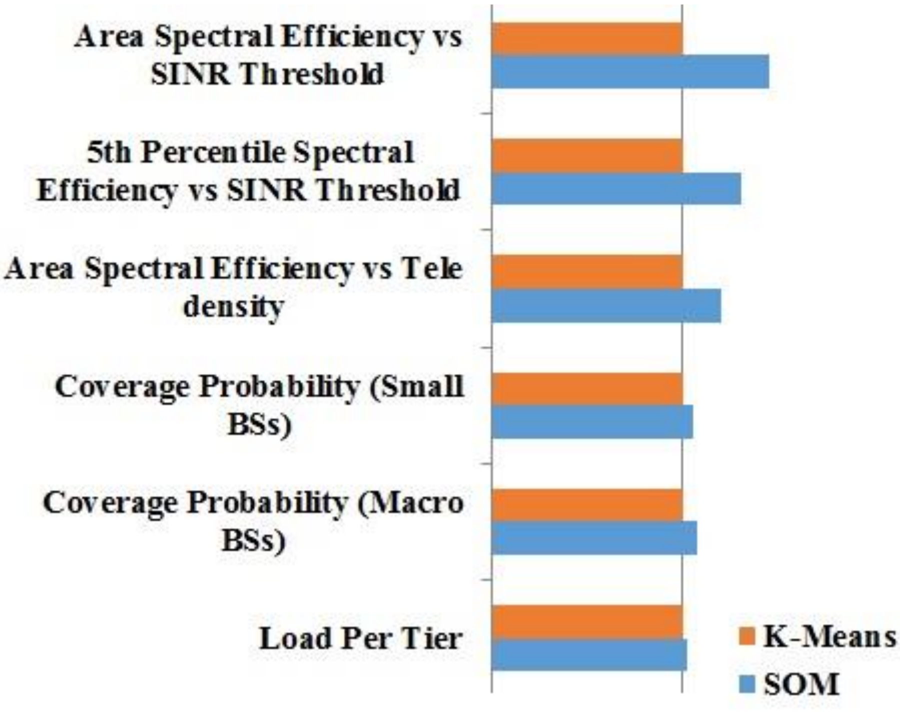


Fig.3.14. Final Comparison Chart between SOM and K-Means

Chapter 4 Performance Analysis of SOM For Dynamic UEs

4.1 Introduction

Initiation of motion for a User Equipment (UE) (either a mobile device or mobile thing like Self Driven Car) will account for any Dynamic UE in a 5G HetNet. While in motion a dynamic UE may cross several cells or coverage area of serving and neighboring base stations. During its movement it might be in an active call or may be in passive mode. If in active call then transferring the call from the serving base station to the neighboring base station is required in order to continue the current communication without any disruption. This process of transfer of an active call from one cell to another is termed as Handoff or Handover [16]. In passive mode when any dynamic UE makes movement it keeps updating its registration in HSS(Home Subscriber Server) to keep the wireless system updated about its latest location [16][17]. In this chapter we will be focusing the Handoff process in the SOM deployed 5G HetNet system.

An exponential growth has been observed in mobile services from 2G to 5G. The ultra-densification concept of 5G is a key factor to support this high data traffic demands. As observed in previous chapters how 5G HetNet concept plays an important role in providing high spectral efficiency. But on the other hand a clear impact of ultra-dense networks is on handover increased rate due to deployment of low power small/femto BSs. So if a system is able to perform the Handover Management efficiently as well for any 5G HetNet, then it stands out to be a standard system capable of maintaining minimum QoS for user experience.

4.2 System Model

The mobility management in wireless network enables the dynamic UE to be mobile within the coverage area and still be serviced. The UE can be in two different modes:

- IDLE_MODE : when the UE is under no active transmission or reception
- CONNECTED_MODE: when the UE is actively transmitting or receiving

Our system model is based on the cell re-selection and handover process when the UE is in CONNECTED_MODE. Before going into 5G NR specific handover management, let's discuss certain basic concepts that sweeps across generations of wireless communication.

When there is UE assisted Network Controlled Handover, the UE will generally measure the signal strength (SS) with reference to a downlink reference signal from the serving base station (S-BS) as well as the neighboring BSs. Based on this measurement report the S-BS starts a handover request to the next target BS. Once done the targeted BS (next S-BS) informs about the successful handover to the current S-BS, which then release any allocated resource to the UE. As example in LTE, the Reference Signal Received Power (RSRP) defined as the average received power excluding interference and noise. Reference Signal Received Quality (RSRQ) defined as the ratio between RSRP and Received Signal Strength Indicator (RSRI), where the total received power along with noise and interference is termed as RSRI. When the RSRP measurements goes below the threshold value, the UE then starts its measurements from S-BS and neighboring BSs.

Now this UE specific reporting can be of below types:-

- Event-triggered Reporting : UE reports about the measurements after the occurrence of a specific event
- Periodic Reporting : UE reports about the measurements at specified time intervals

- On-demand/blind reporting: Immediately when BS sends a request to UE, it reports the measurement. Whenever an entry condition like threshold, hysteresis parameter, is satisfied an event is triggered

In our thesis work we basically initiate On-demand handover. Whenever downlink SINR (with respect to the S-BS) is below the system threshold, immediately handover is initiated based on measurements from neighboring BSs.

In Chapter 3, UE has been static in all use cases. We now induce velocity in UE with the help of Random Walk Model and then study the SOM deployed HetNet's efficiency in handling the handover of the system. But before moving into the Simulation Results let's discuss the Handoff types, Handoff Algorithm and the Random Walk Model that has been employed in our thesis work.

4.2.1 Types of Handoffs

Handover can be of several types with respect to different domains of use cases. We will be broadly discussing the most popular ones [52]:

A) INTER-/INTRA FREQUENCY HANDOVER

When both serving and target BS operates at the same carrier frequency then they are termed as intra-frequency neighbours else they will be called as inter-cell neighbours. In both cases the UE measures the received signal strength as guided by the network. But in case of inter-frequency neighbours the UE needs a measurement gap for performing these measurements on different frequencies. Measurement gap is defined as the time period when no DL/UL transmissions take place. Any UE does these measurements based on order of priority as indicated by the BS since the more the UE measures the more its battery gets consumed.

B) INTER-/INTRA-CELL LAYER HANDOVER

As discussed in Chapter 2, HetNets include a multiplicity of cellular layers (including macro, micro , pico, femto cell layers) to provide more capacity and fulfill the coverage loopholes of the Macro BSs. Intra-frequency same flavor cellular handover (as e.g between macro–to-macro handover) is termed as Intra-layer handover. When the received signal strength with respect to small BS results above any system threshold then offloading to small cell takes place (as eg:- macro-to-femto handover). This is termed as Inter-layer handover. An Inter-layer handover can be both Inter- and Intra-frequency handover. In Intra-frequency Inter-Layer handover, interference is the main challenge, since both macro BS and small BS works in the same carrier frequency. In Inter-frequency Inter-Layer handover the main challenge is the UE power consumption since it has to perform periodic inter-frequency measurements at intervals called measurement gap. As will be demonstrated in our thesis work we encounter the Intra-frequency Inter-Layer handover which also plays an important role in the System Interference Management.

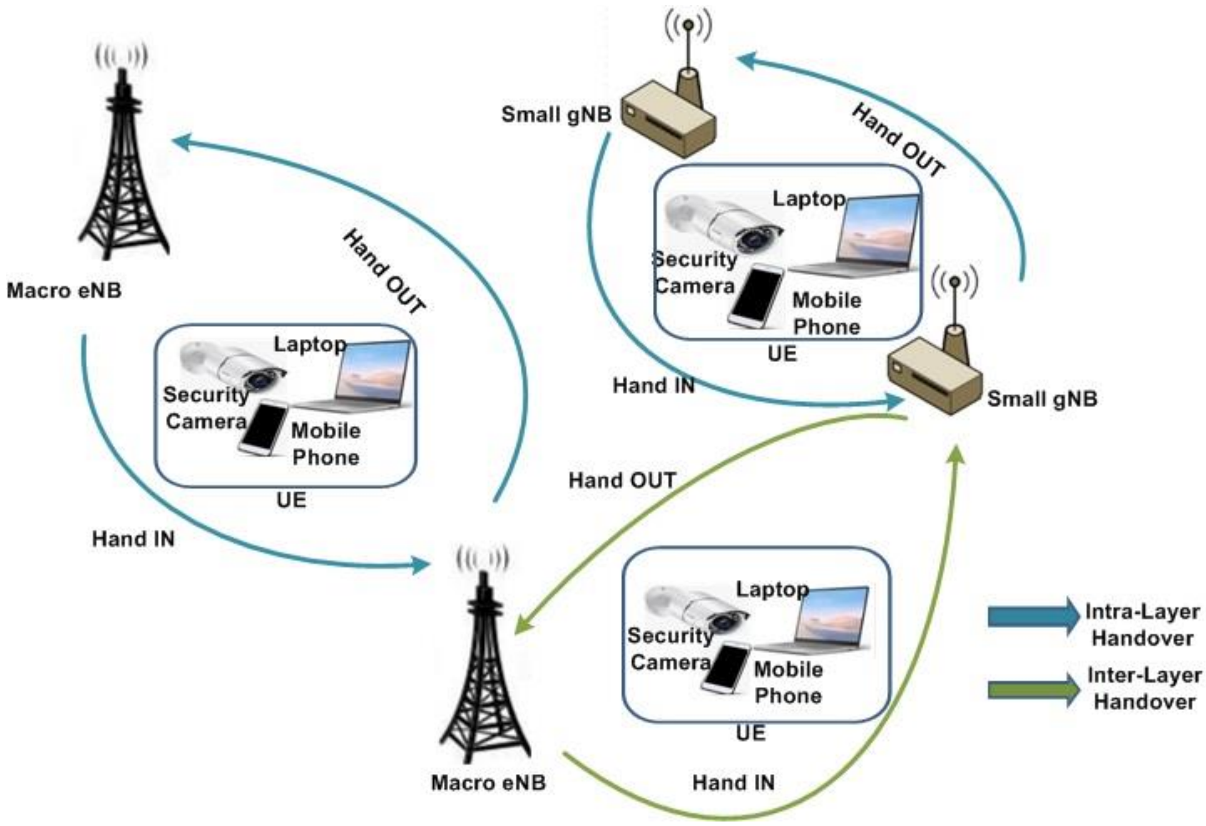


Fig.4.1. Inter-/Intra Cell Layer Handover in 5G HetNet

C) INTER-/INTRA-RAT HANDOVER

The service choice of the user rules over the Radio Access Technology (RAT). The handover that takes place between cells of the same network is termed as Intra-RAT handover while if it occurs between cells of different network then it will be termed as Inter-RAT handover. Inter-RAT handover initiated in HetNet is due to necessity for the service choice opted by the user. This category of handover is still under research in the context of load balancing between RATs.

D) INTER-/INTRA-OPERATOR HANDOVER

In order to provide better service to user, any operator will try to support as many systems and technologies available in the market. Intra-Operator handover accounts for any handover between systems and networks of the same operator while if operators are different then it will be termed as Inter-Operator handover. Biggest advantage of Inter-Operator Handover is when a UE moves from country to country or when UE is roaming. Challenges faced by operators in supporting seamless roaming is to be compatible with other mobile operator's frequency bands, interfaces, protocols etc.

E) BREAK-BEFORE-MAKE VS MAKE-BEFORE-BREAK

Break-Before-Make refers to Hard Handover. In Hard Handover a UE will be connected to only one BS. Thus momentary connection disruption might occur when handover takes place between serving BS and target BS. Make-Before-Break refers to Soft Handover. In Soft Handover, a UE will be able to simultaneously connect to two BSs for some time in order to eliminate any interruption of connection during handover. User Latency experienced was high in LTE's hard handover which is expected to be as low as 4 ms for 5G NR User Plane Latency for eMBB services (as per ITU vision document Report ITU-R M.2410-0). Among others multi-cell connectivity is one solution to achieve this requirement. mmWave can be blocked by building, bricks, human body etc. and hence Multi-Cell connectivity is an important solution. But this also increases the UE reporting and measurement complexities along with battery consumption. This field is still under research for further enhancements.

4.2.2 Handoff Algorithm

There are several different types of handoff algorithms based on received signal strength at UE. The one that has been applied to our thesis work is discussed in this section. When measured downlink SINR with respect to serving BS drops below a threshold γ_L a handoff gets triggered. As example, a handoff could take place between the serving BS (BS_0) and the target BS (BS_1) when the measured downlink SINR with respect to BS_0 drops below the system threshold γ_L and hence handed over to BS_1 whose measured downlink SINR meets with the system threshold and greater than γ_L .This way the scheme reduces the probability of dropped call. Along with this a required average signal strength difference called hysteresis (H) is considered. Excessive handoffs due to “ping-pong effect” between BSs is avoided by handoff hysteresis.

Algorithm 4: *Downlink SINR based Hard Handoff Algorithm* [53]

Step 1: Let $Y_0(n)$ and $Y_1(n)$ are the estimated mean signal strengths (in dBm) of BS_0 and BS_1 respectively $$\mathrm{Y}_0(\mathrm{n}) = \frac{1}{N}\sum_{k=n-N+1}^{n} \lvert r_0(k)\rvert^2 \quad (4.1)$$ $$\mathrm{Y}_1(\mathrm{n}) = \frac{1}{N}\sum_{k=n-N+1}^{n} \lvert r_1(k)\rvert^2 \quad (4.2)$$ where $\lvert r_i(kT_S)^2\rvert$ in (dB) is the kth sample with T_S as sampling period and N as window length
Step 2: Let “H” be the required handoff hysteresis value in dB
Step 3: If the serving BS is BS_0 and γ_L is the system threshold then the handoff is initiated from BS_0 to BS_1 when $\mathrm{Y}_1(\mathrm{n}) > \mathrm{Y}_0(\mathrm{n}) + H\ and\ \mathrm{Y}_0(\mathrm{n}) > \gamma_L$ or when $\mathrm{Y}_1(\mathrm{n}) > \mathrm{Y}_0(\mathrm{n})\ and\ \mathrm{Y}_0(\mathrm{n}) < \gamma_L$
Step 4: If the serving BS is BS_1 and γ_L is the system threshold then the handoff is initiated from BS_1 to BS_0 when $\mathrm{Y}_0(\mathrm{n}) > \mathrm{Y}_1(\mathrm{n}) + H\ and\ \mathrm{Y}_1(\mathrm{n}) > \gamma_L$ or when $\mathrm{Y}_0(\mathrm{n}) > \mathrm{Y}_1(\mathrm{n})\ and\ \mathrm{Y}_1(\mathrm{n}) < \gamma_L$

4.2.3 Random Walk Model

The stochastic process formed by successive summation of independent and identical distributed random variable is defined as the Random Walk. Random Walks are basic models in Applied Mathematics. It is a common example of a Markov Chain. During any stochastic process the fraction of time spent in each state represents the limiting stationary distribution of the Markov Chain. In a finite set of states a random walk is the process that moves from one state to another in a manner that depends only on the last state and the transition from one state to another is associated with transition probability [54][55]. As example, a random walk with few sample trajectories of transitions is shown in Fig.4.2.

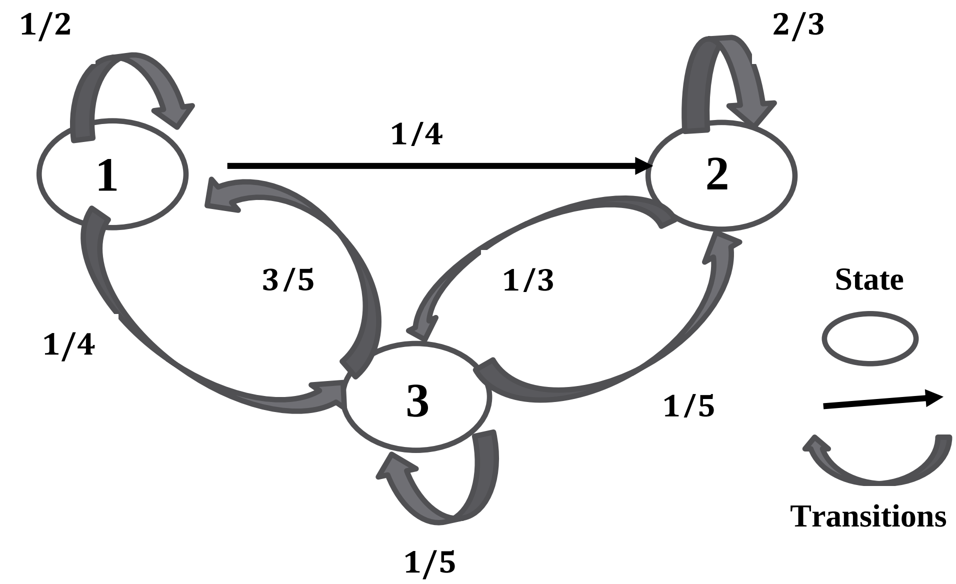


Fig.4.2. A Sample Random Walk associated with Transition Probability

The Random Walk depicted in Fig.4.2. is also known as a Markov Chain. One key property of Markov Chain is the walk is expected to be on average when the process tends towards infinity.

Let's say a random walker (as eg a drunkard) can walk along a line with regularly spaced positions, Δx distance apart. At time 0 the walker is at position 0 while after a fixed time interval Δt the walker either jumps to the right with probability p or to the left with probability q=1-p, for discrete time points $N \geq 0$.

We will be calculating the probability p(m,N) that the walker will be at position 'm' after 'N' steps. This probability follows the binomial distribution and is given by [55]:

$$p(m,N) = \frac{N!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!}\, p^{\frac{N+m}{2}}\, q^{\frac{N-m}{2}} \tag{4.3}$$

The Gaussian approximation for Random Walk probability p(m,N) when N is very large. The position 'm' after 'N' steps can be expressed as the sum of individual positions attained, m_i where $m_i \in \{-l,l\}$ and $m=\sum_{i=1}^{N} m_i$

Mean value averaged over all possible random walks

$$\prec \mathrm{m} \succ = \sum_{i=1}^{N} \prec m_i \succ = N \prec m_1 \succ = N(pl-(1-p)l$$

$$\prec \mathrm{m} \succ = Nl(2p-1) \tag{4.4}$$

Average variance over all possible random walks:

$$\sigma^2 = \prec m^2 \succ - \prec \mathrm{m} \succ^2 = N\sigma_1^2 = N(\prec m^2 \succ - \prec \mathrm{m} \succ^2)$$

$$\sigma^2 = N(pl^2 + (1-p)l^2 - \prec \mathrm{m} \succ^2)$$

$$\sigma^2 = 4Nl^2 p(1-p) \tag{4.5}$$

According to the Central Limit Theorem for large N $p(m,N)$ when approaches towards Gaussian distribution then results in [55]:

$$p(m,N) \approx \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(m-\prec \mathrm{m} \succ)^2/(2\sigma^2)} \tag{4.6}$$

4.3 Performance Analysis of Dynamic Simulated System Model

We now put forward the simulation results studied after performing re-clustering based on SINR, and applied to dynamic UEs, as explained in section 4.2. The required parameters and their respective values are provided in Table 4.1.

TABLE 4.1: Simulation Parameters' Values

PARAMETRIC NAME	PARAMETRIC VALUE
Operating Carrier Frequency	4.1 GHz
Bandwidth	5 MHz
Coverage Area	4 km^2

No.of Macro BSs	3
No.of Small/Femto BSs	15
Transmitting power of Macro BS	46 dBm
Transmitting power of Femto BS	20 dBm
No.of UEs	Varied by PPP
Speed of very slow walker	5 km/hr
Speed of brisk walker	7 km/hr
Speed of Running Female	10 km/hr
Speed of Running Male	13 km/hr

The handover concept is important for any system's QoS. As a known fact [16], the QoS or Grade of Service (GoS) of a wireless network is dependent on P_B *(the probability of new calls being blocked) and* P_F *(the probability of forced termination of ongoing calls).* GoS always has a higher *dependency on* P_F as compared to P_B and any network operator's performance is also more dependent on P_F. Hence the *% of successful handoff* of any wireless network is an important aspect of any system. In our prototype we have induced Random Walk Model while employing motion to our static UE, which is explained under section 4.2.

4.3.1 Simulation Results of Successful Handover in SOM Based Deployed 5G HetNet System with respect to System Threshold

We will be presenting our simulation results of Successful handover percentage in SOM based deployed 5G HetNet system with respect to different system thresholds. We will be analyzing this percentage by varying velocity of users and also, by varying the mobility percentage among users. This will then be helping us to demonstrate that whether varying velocity with increasing mobile users can be supported successfully by the system or prototype.

- **Simulation Result Analysis of Successful Handover % Across Different UE Velocities**

As velocity of mobile UE increases the successful handoff percentage at higher SINR target thresholds decreases but not significantly. Also at lower SINR target thresholds, the percentage reaches to almost 100% across all velocities. Thus handover success rate is quite high and supports well at higher velocity. Fig.4.3 provides the required simulation result in this respect.

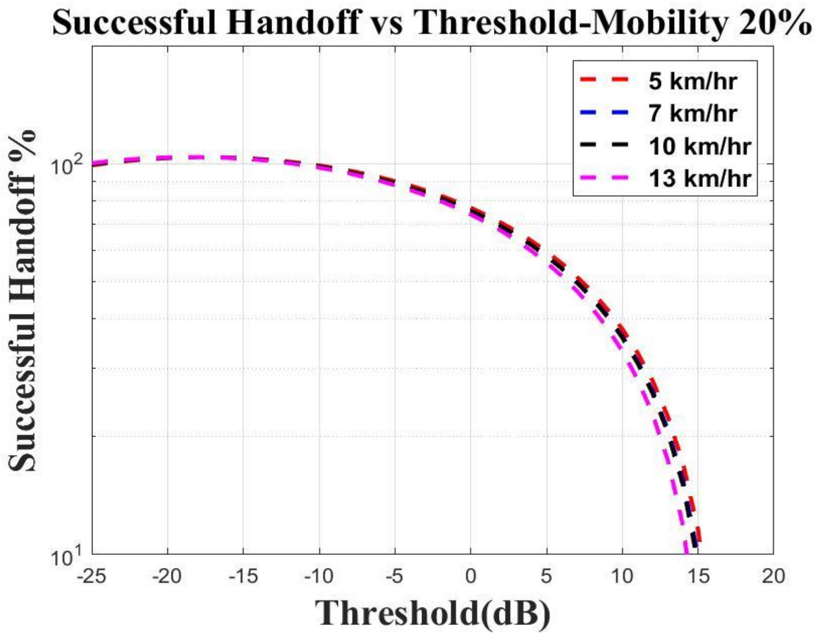


Fig.4.3. Successful handoff % vs Target SINR Threshold across Velocity

- **Simulation Result Analysis of Successful Handover % Across Different UE Mobility**

As mobility percentage of mobile UE increases the successful handoff percentage at higher SINR target thresholds decreases but not significantly. Also at lower SINR target thresholds, the percentage reaches to almost 100%. Thus handover success rate is quite high and supports well at higher mobility. Fig 4.4 provides the required simulation result in this respect. 5G aims at seamless connectivity. We observe that SOM deployed BSs are located at quite favorable positions to provide seamless connectivity with increasing velocity and increasing mobility for different target system SINR thresholds.

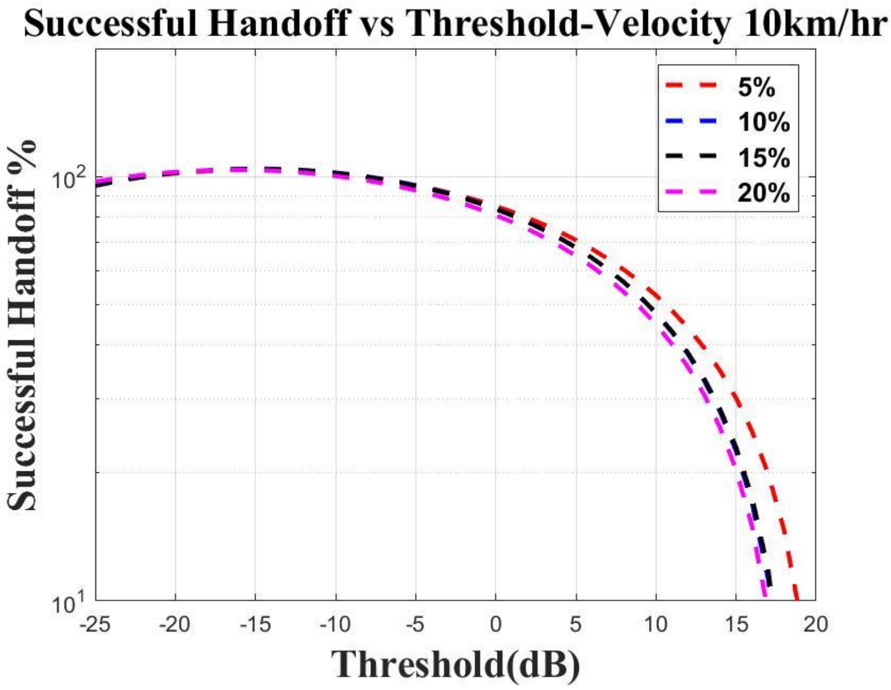


Fig.4.4. Successful handoff % vs Target Threshold across Mobility

4.3.2 Simulation Results of Successful Handover in SOM Based Deployed 5G HetNet System with respect to Tele-density

We will be presenting our simulation results of successful handover percentage in SOM based deployed 5G HetNet system with respect to different Tele-density. We will be analyzing this percentage by varying velocity of users and also by varying the mobility percentage among users. This will then be helping us to demonstrate that whether varying velocity with increasing mobile users can be supported successfully by the system or prototype.

- **Simulation Result Analysis of Successful Handover % Across Different UE Velocities**

As velocity of mobile UE increases the successful handoff percentage at any tele density increases. This successfully demonstrates that the prototype is capable of supporting any mobile UE at any speed. It is natural that as speed increases the chance of handoff increases. Hence successful handoff percentage will be high at higher speed only when the system is able to make successful connectivity at any possible speed. But another aspect is noticeable that as tele density increases naturally at any speed the successful handoff percentage also decreases. This happens since the number of small BSs doesn't increase in our prototype with rise in mobile UE. Fig.4.5 provides the required simulation result in this respect.

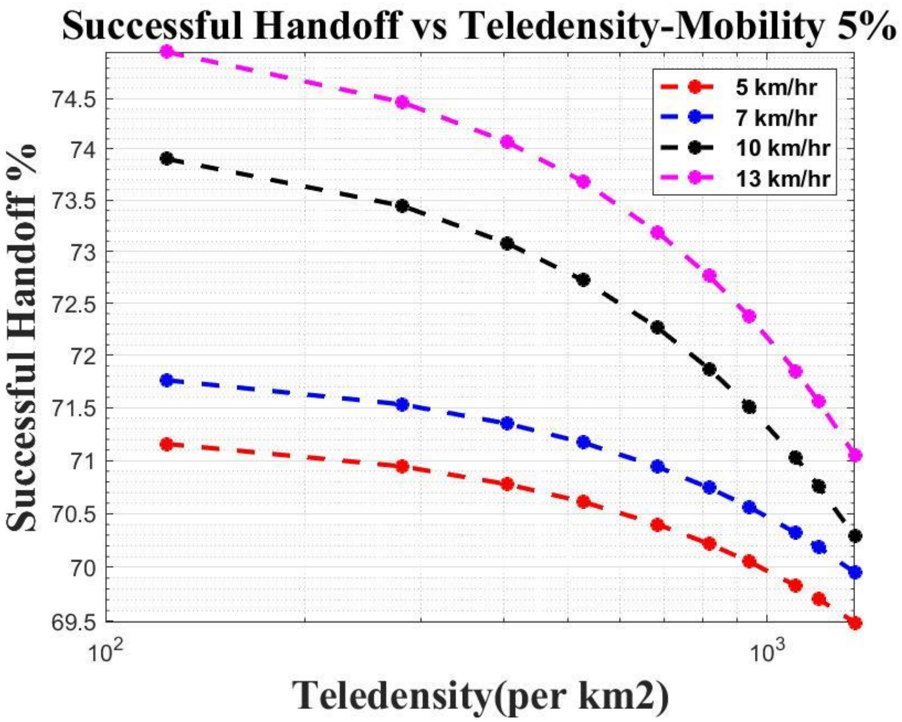


Fig.4.5. Successful handoff % vs Tele density across Velocity

- **Simulation Result Analysis of Successful Handover % Across Different UE Mobility**

Though the prototype shows promising results with respect to increasing velocity in previous section, but for increasing mobility percentage the successful handoff percentage decreases. Thus the system needs to be leveraged further to support high mobility percentage as well. At any constant mobility percentage even for higher tele density, successful handoff percentage decreases since the number of small BSs in our prototype doesn't increase with rise in tele-density. Fig.4.6 provides the required simulation result in this respect.

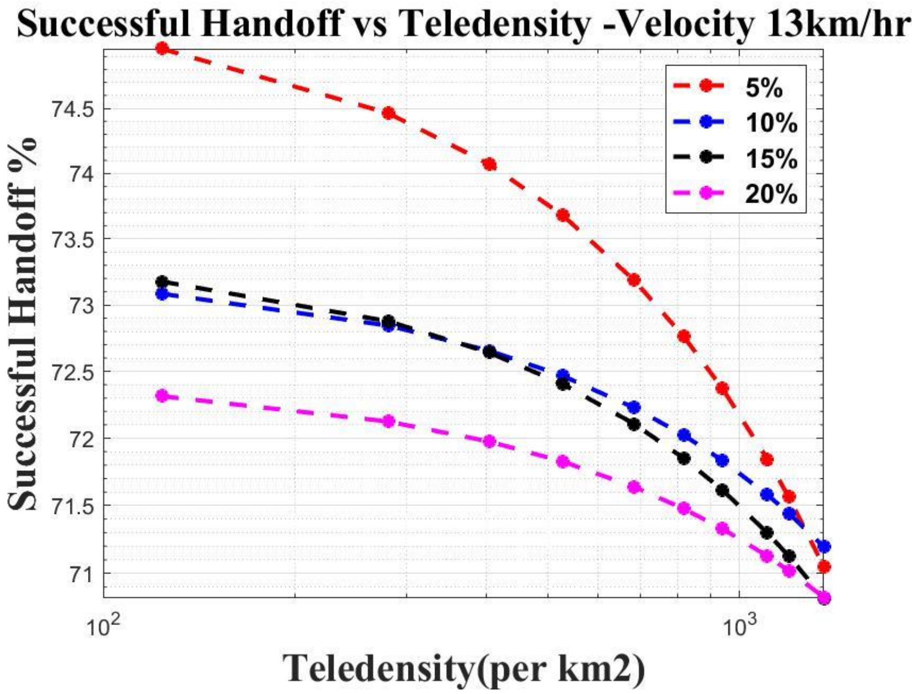


Fig.4.6. Successful handoff % vs Tele density across Mobility

Increasing mobile UEs is the driving factor for 5G network requirements. In terms of deployment, positions of small BSs by SOM provides promising results with respect to increasing velocity & increasing tele density. Providing seamless connectivity with increasing velocity & tele density is definitely the demand of the day in today’s world.

Chapter 5 Conclusion and Future Scope

5.1 Conclusion

In our thesis work, we have explored suitability of Kohonen’s Self Organizing Map (SOM) based downlink HetNet deployment model for 5G along with performance comparisons with K-means. The most useful observation is that mean area spectral efficiency (ASE) for SOM is higher than K-means at the cost of little more time complexity that would help first phase of 5G deployment for eMBB use cases. This demonstration is clearly indicated in Table 5.1 where SOM outperforms K-Means especially for ASE. The primary usage scenario for 5G is eMBB which concentrates on human-centric access to multi-media content, services, data etc. In Fig. 5.1 we portray the key requirements for 5G usage scenarios as referenced from [56]. It clearly displays how spectral efficiency acts as an important parameter for eMBB.

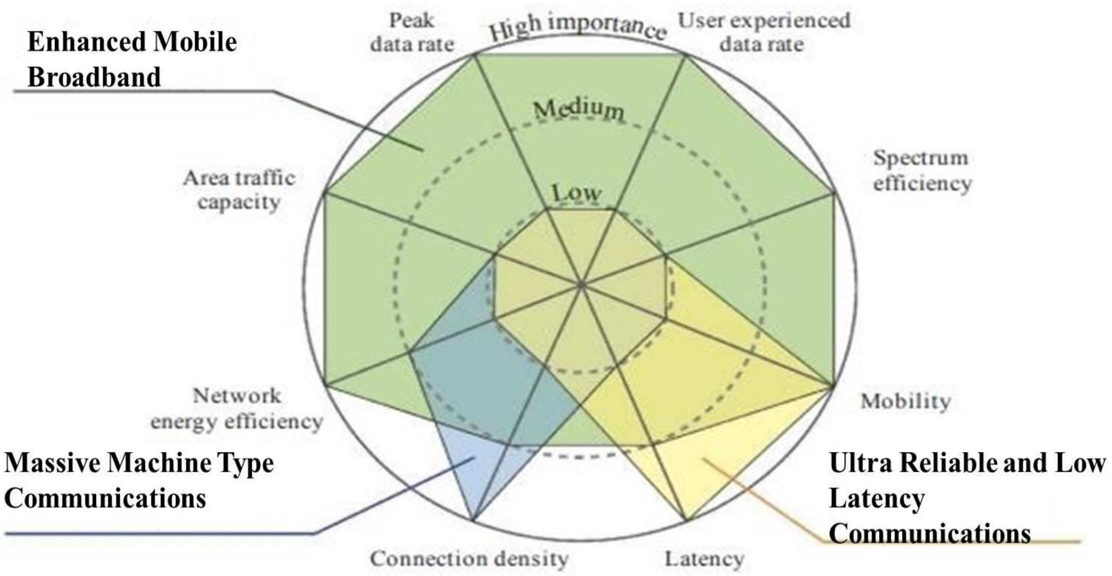


Fig.5.1. Essential Requirements of 5G Use Cases [56]

TABLE 5.1: Comparative Study between K-Means and SOM

Parameter Name	K-Means	SOM	Inference
Mean ASE(bits per km^2) on $SINR_{TH}$	1.017	1.473	45% increase observed in case of SOM
Mean Load Per Tier on $SINR_{TH}$	474	478	0.84% increase observed in case of SOM while considering 520 average tele density in 4 km^2 coverage area. As per 5G IMT 2020 requirements for mMTC [56], Connection Density per km^2 is supposed to be 1 million. For eMBB service even if we consider 10,000 active connections per km^2 still we get the percentage increase to be 16%
Mean Coverage Probability on $SINR_{TH}$	0.89	0.89	Both techniques provide good coverage
Mean ASE(bits per km^2) on Tele density	0.581	0.587	1% increase observed in case of SOM with 638 as an average tele density per km^2. As per 5G IMT 2020 requirements for mMTC [56], Connection Density per km^2 is supposed to be 1 million. For eMBB service even if we consider 10,000 active connections per km^2 still then the percentage increases to 16%

In our prototype, as indicated in Tables 3.2 and 4.1, the coverage area considered is 4 km^2 and operating frequency is 4.1 GHz (part of FR1) and hence bandwidth considered is also the minimum bandwidth, which is 5 MHz. In FR1 allowable bandwidth in 5G is from 5 MHz to 100 MHz while in FR2 it is 50 MHz to 400 MHz [57]. The IMT-2020 5G requirements as stated in [58] is based on minimum bandwidth of 100 MHz. In Table 5.2 we provide a comparative study of these requirements and the results that are obtained when we applied SOM for FR1 for minimum channel bandwidth of 5 MHz considering eMBB use cases – which is the primary use case to be served first. Hence the comparison is based on 20:1 ratio approximately (i.e 100 MHz Channel Bandwidth : 5 MHz Channel Bandwidth).

TABLE 5.2: Comparative Study between 5G IMT-2020 requirements and Results Obtained in Thesis Work

Parameter Name	IMT-2020 5G requirements for 100 MHz Bandwidth	5G Simulated Prototype deployed applying SOM for 5 MHz Bandwidth in LOS Outdoor Environment
5th percentile user spectral efficiency (DL)	0.3 bit/s/Hz (Indoor Hotspot) 0.225 bit/s/Hz (dense urban) 0.12 bit/s/Hz (rural)	0.635 bit/s/Hz
User Experienced Data Rate (DL)	100 Mbps	3.175 Mbps
Average Spectral Efficiency (DL)	9 bit/s/Hz (Indoor Hotspot) 7.8 bit/s/Hz (dense urban) 3.3 bit/s/Hz (rural)	4.06 bit/s/Hz
Peak Spectral Efficiency (DL)	30 bit/s/Hz	7.9 bit/s/Hz

Note: DL means Downlink

In Table 5.3 we produce the SOM based deployed 5G simulated system's performance for the dynamic users with respect to different mobility and speed. The maximum successful handoff % achieved is 100% for different SINR thresholds. This is a major requirement for 5G to achieve seamless connectivity. Also, an average of 70% successful handoff is always maintained for different SINR thresholds and for varying tele density.

It can be concluded that for faster rollout of 5G with higher data rate and spectral efficiency and seamless connectivity, SOM can be effectively applied to get the best position of the base stations in a particular geographical area both for static as well as dynamic UEs.

TABLE 5.3: Performance of SOM Based System Model under Mobility

Parameter Name	Velocity (km/hr)	SOM	Mobility %	SOM	Inference
Mean Successful Handoff % on SINR Threshold	5	68.7	5	67.6	Good system response for increasing velocity and mobility. Maximum 100% was achieved in both cases: while varying velocity as well as mobility
	7	67.8	10	67.3	
	10	67.6	15	67.2	
	13	66.1	20	67	
Mean Successful Handoff % on Tele density	5	73.7	5	76.7	Good system response for increasing velocity and mobility for varying tele density.
	7	74.5	10	75.1	
	10	76.4	15	74.9	
	13	76.7	20	73.6	

We also demonstrate an up gradation over popular K-Means unsupervised algorithm, whose application in Wireless Communication is quite popular as compared to Self Organizing Map. With increasing tele-density in current 5G world, on applying this simple neural network results in enhanced system performance thereby supporting the ultra-dense concept more efficiently.

5.1.1 Establishing the Competency of Self Organizing Map

Post analysis of Tables 5.1, 5.2 and 5.3, we hereby conclude the competitive results as provided by SOM. As per Table 5.1, SOM outperforms the popular unsupervised K-Means algorithm, in the ruling performance parameter of 5G eMBB service i.e Area Spectral Efficiency while competing hand-in-hand with K-Means in other parameters like Coverage Probability, Load Per Tier etc. As per Table 5.2, the 5G system model based on SOM meets the required minimum parameter values of IMT-2020 for 5G. Finally, Table 5.3

demonstrates that the 5G system model based on SOM not only provides promising results with static UE but also is capable of providing 100% successful handoff for maximum SINR threshold values.

Thus clearly, we can conclude that if this less-explored unsupervised neural network, SOM, is utilized for the deployment of base stations in 5G standalone network then not only SOM gives better results but also its implementation is faster and easier. Hence if one has time constraint with respect to implementation, which is an important aspect while launching products to market, then one can easily opt for SOM for completion of the job in time rather than jumping into more complex learning like Reinforcement Learning, for further better results.

5.1.2 Bibliography

Part of our thesis work has been submitted as IEEE Conference Paper in 2023 IEEE 4th Annual Flagship India Council International Subsections Conference, INDISCON 2023. On the title “Leveraging Self Organizing Map For 5G Heterogeneous Network Deployment For Enhanced Performance”, Paper ID: 633.

5.2 Future Scope

- In the entire thesis work 5G FR1 (410 MHz to 7.12 GHz) [11] is considered since this spectrum is cheaper and more in use in developing countries like India. But the thesis can be extended to FR2 (24.25 GHz to 71 GHz) [11] which has been opted in developed countries like the US. This way the entire 5G spectrum implementation can be studied via the prototype.
- In the thesis work the coverage area is bound to a small geographical area. Practically it emulates the private 5G network which is quite trending in current scenario. For large 5G networks, the work can be extended to cover large geographical areas as well with high computational power computers.

- In Chapter 3, it was seen that computational complexity of K-Means is better than SOM. Also with increasing tele-density the performance gap in terms of computational complexity between K-Means and SOM was narrowing (refer to Fig.3.12). Thus if SOM is used along with K-Means and applied for the deployment of the base stations in a standalone 5G network architecture, then even improvement in terms of computational complexity is expected.
- The Thesis work was primarily based on Unsupervised Learning. The third branch of Machine Learning i.e Reinforcement Learning seems to be more promising since it has the capability to work even in complete unknown environment and learn in the process of Action and Reward. Thus for self organizing network it is expected to perform well practically.

References

[**1**]ITU-R, 2014, ITU-R FAQ on INTERNATIONAL TELECOMMUNICATIONS,[Online].Available: https://www.itu.int/en/ITU-R/Documents/ITU-R-FAQ-IMT.pdf

[**2**] ETSI, TS., 2018-07. 122 261 V15.5.0: 5G; Service requirements for next generation new services and markets, 3GPP Std. 3GPP TS 22.261 version 15.5.0 Release 15

[**3**] P. Popovski, K.F. Trillingsgaard, O. Simeone and G. Durisi., 2018. 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view. IEEE Access, 6, pp.55765-55779.

[**4**] S. Singh, H.S. Dhillon and J.G. Andrews, 2013. Offloading in heterogeneous networks: Modeling, analysis, and design insights. IEEE transactions on wireless communications, 12(5), pp.2484-2497.

[**5**] J. Tanveer, A. Haider, R. Ali and A. Kim, 2022. Machine learning for physical layer in 5G and beyond wireless networks: A survey. Electronics, 11(1), p.121.

[**6**] Y. Siriwardhana, G. Gür, M. Ylianttila and M. Liyanage, 2021. The role of 5G for digital healthcare against COVID-19 pandemic: Opportunities and challenges. ICT Express, 7(2), pp.244-252.

[**7**] T. Kohonen, 1990. The self-organizing map. Proceedings of the IEEE, 78(9), pp.1464-1480.

[**8**] A. Ciampi and Y. Lechevallier, Clustering large, multi-level data sets: an approach based on Korhonen self organizing maps. In Principles of Data Mining and Knowledge Discovery: 4th European Conference, PKDD 2000 Lyon, France, September 13–16, 2000 Proceedings 4 (pp. 353-358). Springer Berlin Heidelberg.

[**9**] A. Ghosh, I. Saha Misra and A. Kundu, Coverage and rate analysis in two-tier heterogeneous networks under suburban and urban scenarios. Transactions on Emerging Telecommunications Technologies, 30(12), p.e3648, 2019.

[**10**] IEEE, Innovation at Work, Developments-and-Challenges-Influencing-The-Advancement-of-5g,[Online],Available: https://innovationatwork.ieee.org/developments-and-challenges-influencing-the-advancement-of-5g

[**11**] X. Lin, An overview of 5G advanced evolution in 3GPP release 18. IEEE Communications Standards Magazine, 6(3), pp.77-83, 2022.

[**12**] R. Dilli, March. Analysis of 5G wireless systems in FR1 and FR2 frequency bands. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 767-772). IEEE2020.

[**13**] ETSI, TS., 2018-10. 138 133 V15.3.0: 5G; NR Requirements for support of radio resource management, 3GPP Std. 3GPP TS 38.133 version 15.3.0 Release 15.

[**14**] ETSI, T., 2018. 138 901 V14. 3.0: 5G; Study on channel model for frequencies from 0.5 to 100 GHz, 3GPP Std. 3GPP TR 38.901 version 14.3. 0 release 14.

[**15**] J. Kaur, M.A. Khan, M. Iftikhar, M. Imran and Q.E.U. Haq, Machine learning techniques for 5G and beyond. IEEE Access, 9, pp.23472-23488, 2021.

[**16**] I.S.Misra, 2013. Wireless communications and networks: 3G and beyond. McGraw Hill Education (India) Pvt Ltd; ISBN: 978-1-25-906273-5.

[**17**] P. Nicopolitidis, M.S. Obaidat, G.I. Papadimitriou and A.S. Pomportsis, 2003. Wireless networks. John Wiley & Sons.

[**18**] J.G. Andrews, F. Baccelli and R.K. Ganti, 2011. A tractable approach to coverage and rate in cellular networks. IEEE Transactions on communications, 59(11), pp.3122-3134.

[**19**] A. Konar, 2006. Computational intelligence: principles, techniques and applications. Springer Science & Business Media; ISBN:13. 978-3540208983.

[**20**] H.U. Dike, Y. Zhou, K.K. Deveerasetty and Q. Wu, 2018, October. Unsupervised learning based on artificial neural network: A review. In 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS) (pp. 322-327). IEEE.

[**21**] S. Jovanović and H. Hikawa, 2022. A survey of hardware self-organizing maps. IEEE Transactions on Neural Networks and Learning Systems.

[**22**] A. Sun, K. Zhu, J. Du and H. Cao, 2021, December. PSO-Based K-means Algorithm for Clustering Routing in 5G WSN Networks. In 2021 IEEE Globecom Workshops (GC Wkshps) (pp. 1-6). IEEE.

[**23**] H.S. Dhillon, R.K. Ganti, F. Baccelli and J.G. Andrews, 2012. Modeling and analysis of K-tier downlink heterogeneous cellular networks. IEEE Journal on Selected Areas in Communications, 30(3), pp.550-560.

[**24**] E. Balevi and R.D. Gitlin, 2018, May. A clustering algorithm that maximizes throughput in 5G heterogeneous F-RAN networks. In 2018 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.

[**25**] K.Utkarsh and F. Ding, 2022. Self-organizing map-based resilience quantification and resilient control of distribution systems under extreme events. IEEE Transactions on Smart Grid, 13(3), pp.1923-1937.

[**26**] N.A.M Riveros,. B.A.C. Espitia and L.E.A. Pico, 2019. Comparison between K-means and self-organizing maps algorithms used for diagnosis spinal column patients. Informatics in Medicine Unlocked, 16, p.100206.

[**27**] Z. Ma, Y. Yang, Wang, F., Li, C. and Li, L., 2014, November. The SOM based improved k-means clustering collaborative filtering algorithm in TV recommendation system. In 2014 Second International Conference on Advanced Cloud and Big Data (pp. 288-295). IEEE.

[**28**] K. Raivio, O. Simula, J. Laiho and P. Lehtimaki,2003, March. Analysis of mobile radio access network using the self-organizing map. In IFIP/IEEE Eighth International Symposium on Integrated Network Management, 2003. (pp. 439-451). IEEE.

[**29**] L. Yu, X. Zhao, F. Du, Y. Zhang, Z. Fu and S. Geng, 2021, December. Multipath Component Clustering Based on Improved Self-organizing Feature Map for 5G Millimeter Wave Radio Channels. In 2021 13th International Symposium on Antennas, Propagation and EM Theory (ISAPE) (pp. 1-3). IEEE.

[**30**] F. Du, Y. Zhang, X. Zhao, S. Geng, Z. Fu, X. Wang, L. Yu and Q. Li, 2022. Millimeter Channel Clustering by Self-Organizing Maps with Time-

Varying Topological Structure. IEEE Transactions on Antennas and Propagation.

[**31**] S. Wang and X. Zhang, 2020, April. Analysis of self-organizing maps (SOM) methods for cell clustering with high-dimensional OAM collected data. In 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA) (pp. 229-233). IEEE.

[**32**] Frota, R.A., Barreto, G.A. and Mota, J.C., 2007. Anomaly detection in mobile communication networks using the self-organizing map. J. Intell. Fuzzy Syst., 18(5), pp.493-500.

[**33**] A. Shahraki, A. Taherkordi, Ø. Haugen and F. Eliassen, 2020. A survey and future directions on clustering: From WSNs to IoT and modern networking paradigms. IEEE Transactions on Network and Service Management, 18(2), pp.2242-2274.

[**34**] M.Z. Asghar, S.A. Memon and J. Hämäläinen, 2022. Evolution of wireless communication to 6g: Potential applications and research directions. Sustainability, 14(10), p.6356.

[**35**] A.S. Shah, 2022, January. A survey from 1G to 5G including the advent of 6G: Architectures, multiple access techniques, and emerging technologies. In 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 1117-1123). IEEE.

[**36**] O.O.Erunkulu, A.M.Zungeru, C.K.Lebekwe, M.Mosalaosi and J.M.Chuma, 2021. 5G mobile communication applications: A survey and comparison of use cases. IEEE Access, 9, pp.97251-97295.

[**37**] B. Agarwal, M.A.Togou, M. Ruffini and G.M. Muntean, 2022. A Comprehensive Survey on Radio Resource Management in 5G HetNets: Current Solutions, Future Trends and Open Issues. IEEE Communications Surveys & Tutorials.

[**38**] J. Ghosh, D.N.K. Jayakody and M. Qaraqe, 2018. Downlink capacity of OFDMA-CR based 5G femtocell networks. Physical Communication, 29, pp.329-335.

[**39**] B. Castellani and J. Castellani, 2003. Data mining: qualitative analysis with health informatics data. Qualitative Health Research, 13(7), pp.1005-1018.

[**40**] Simon Haykin,2008.Neural Networks and Learning Machines, 3rd Edition. ISBN: 9789332586253, 933258625X

[**41**] No. 18-06/2016-STT(T) Pt.II; Government of India; Ministry of Communication; Department of Telecom; Dated 10th March, 2021

[**42**] R. Amiri, 2020. Reinforcement Learning in Self Organizing Cellular Networks.

[**43**] S. Mukherjee, 2014. Analytical modeling of heterogeneous cellular networks. Cambridge University Press.

[**44**] J.G. Andrews, S. Singh, Q.Ye, X. Lin and H.S. Dhillon, 2014. An overview of load balancing in HetNets: Old myths and open problems. IEEE Wireless Communications, 21(2), pp.18-25.

[**45**] M.F.Khan, F.A. Bhatti, A. Habib, S. Jangsher, M.I. Khan, I. Zafar, S.M. Shah, M.A. Jamshed and J. Iqbal, 2017, November. Analysis of macro user offloading to femto cells for 5G cellular networks. In 2017 International Symposium on Wireless Systems and Networks (ISWSN) (pp. 1-6). IEEE.

[**46**] A. Goldsmith, 2005. Wireless communications. Cambridge university press; ISBN: 9781139445849, 1139445847

[**47**] H.A. Hou and L.C. Wang, 2017, April. Area spectral efficiency for cellular networks with small reuse distance: An algebraic approach. In 2017 26th Wireless and Optical Communication Conference (WOCC) (pp. 1-6). IEEE.

[**48**] P.I. Dalatu, 2016. Time complexity of k-means and k-medians clustering algorithms in outliers detection. Global Journal of Pure and Applied Mathematics, 12(5), pp.4405-4418.

[**49**] M.K. Pakhira, 2014, November. A linear time-complexity k-means algorithm using cluster shifting. In 2014 international conference on computational intelligence and communication networks (pp. 1047-1051). IEEE.

[**50**] L. Besson, 2016. Self-Organizing Maps (SOM) and Dynamic SOM, models of cortical plasticity.

[**51**] O.O.Erunkulu, A.M.Zungeru, C.K.Lebekwe, M.Mosalaosi and J.M. Chuma, 2021. 5G mobile communication applications: A survey and comparison of use cases. IEEE Access, 9, pp.97251-97295.

[**52**] M. Tayyab, X. Gelabert, and R. Jäntti, "A survey on handover management: From LTE to NR," IEEE Access, vol. 7, pp. 118907_118930, 2019.

[**53**] G.L. Stüber and G.L. Stèuber, 2001. Principles of mobile communication (Vol. 2, p. 752). Boston: Kluwer academic.

[**54**] F. Spitzer, 2001. Principles of random walk (Vol. 34). Springer Science & Business Media.

[**55**] G.F. Lawler and V. Limic, 2010. Random walk: a modern introduction (Vol. 123). Cambridge University Press.

[**56**] IMT Vision – "Framework and overall objectives of the future development of IMT for 2020 and beyond", ITU-R Recommendation M.2083-0

[**57**] M.Fuentes, J.L. Carcel, C.Dietrich, L.Yu, E.Garro, V.Pauli, F.I. Lazarakis, O. Grøndalen, O. Bulakci, J. Yu and W. Mohr, 2020. 5G new radio evaluation against IMT-2020 key performance indicators. IEEE Access, 8, pp.110880-110896.

[**58**] Minimum requirements related to technical performance for IMT-2020 radio interface(s); Report M.2410-0 (11/2017)