Efficient Service for Next Generation Network Slicing Architecture and Mobile Traffic Analysis Using Machine Learning Technique

Billian Khan Tapan
Marquette University

EFFICIENT SERVICE FOR NEXT GENERATION NETWORK SLICING ARCHITECTURE AND MOBILE TRAFFIC ANALYSIS USING MACHINE LEARNING TECHNIQUE

by

Billian Khan Tapan

A Thesis submitted to the Faculty of the Graduate School,

Marquette University,

in Partial Fulfillment of the Requirements for

the Degree of Master of Science

Milwaukee, Wisconsin

August 2021

ABSTRACT

EFFICIENT SERVICE FOR NEXT GENERATION NETWORK SLICING ARCHITECTURE AND MOBILE TRAFFIC ANALYSIS USING MACHINE LEARNING TECHNIQUE

Billian Khan Tapan

Marquette University, 2021

The tremendous growth of mobile devices, IOT devices, applications and many other services have placed high demand on mobile and wireless network infrastructures. Much research and development of 5G mobile networks have found the way to support the huge volume of traffic, extracting of fine-gained analytics and agile management of mobile network elements, so that it can maximize the user experience. It is very challenging to accomplish the tasks as mobile networks increase the complexity, due to increases in the high volume of data penetration, devices, and applications. One of the solutions, advance machine learning techniques, can help to mitigate the large number of data and algorithm driven applications. This work mainly focus on extensive analysis of mobile traffic for improving the performance, key performance indicators and quality of service from the operations perspective. The work includes the collection of datasets and log files using different kind of tools in different network layers and implementing the machine learning techniques to analyze the datasets to predict mobile traffic activity. A wide range of algorithms were implemented to compare the analysis in order to identify the highest performance. Moreover, this thesis also discusses about network slicing architecture its use cases and how to efficiently use network slicing to meet distinct demands.

ACKNOWLEDGMENTS

Billian Khan Tapan

The completion of my dissertation works would not have been possible without many helps. I would like to express my heartfelt gratitude to my supervisor Professor Dr. Sheikh Iqbal Ahamed, who has given me tremendous support to complete my entire works. He persuasively encouraged me and guided me to right way. His moral support encouraged me during the hard work times and gave me an opportunity for my personal growth and for the achieve results to completion of my work. It would have been challenging without his persistent help.

I would also sincere appreciation to my thesis committee member Dr.Thomas Kaczmarek and Dr. Anik Iqbal and their excellent support help me a lots to complete the project successfully.

The research work involved to study most demandable and about next generation future technology. The technological surge forwards will lead to new products, businesses, industries and even from shelf driving cars to artificial intelligence. The Economists' estimates that 5G or next generation technology deployment contribute $400 billion to $500 billion to US GDP and creates about 1 million jobs from 2020 to 2030.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

IP Internet Protocol

TCP Transmission Control Protocol

UDP User Datagram Protocol

CDMA Code Division Multiple Access

ML Machine Learning

SVM Support Vector Machine

CDR Call Data Record

LTE Long Term Evolution

PDCCH Physical Downlink Control Channel

eNodeB Base Station

DCI Downlink Control Information

TTI Transmission Time Interval

MCS Modulation and Coding Scheme

SDR Software-Defined Radio

UE User Equipment

TBS Transport Block Size

RNTI Radio Network Temporary Identifier

LSTM Long Short-Term Memory

MLP Multilayer Perceptron

ReLU Rectified Linear Unit

RNN Recurrent Neural Network

NN Neural Network

MC Memory Cell

RB Resource Block

HTTP Hypertext Transfer Protocol

GAN Generative-Adversarial Networks

SSD Solid State Storage

HMM Hidden Markov Model

RF Random Forests

AD Anomaly Detection

AE Autoencoder

DL Downlink

EDA Exploratory Data Analysis

MNO Mobile Network Operator

RAN Radio Access Network

MIMO Multiple input and Multiple output

AMF Access and Mobility Function

SMF Session Management Function

UDM Unified Data Management

UPF User Plane Function

AUSF Authorization Server Function

NSSF Network Slice Selection Function

CSCF Call Session Control Function

MGW Media Gateway

BGCF Breakout Gateway Control Function

MGCF Media Gateway Control Function

CHAPTER 1
INTRODUCTION

Internet usage and demand has increased in every aspect of our lives and places of work. Most people around the world use smartphones and other devices. As the number of devices using different applications increases, so does the mobile data traffic. As assessed, 4.1 billion people are using the internet and the global internet penetration rate, defined by the percentage of population that uses the internet, is 59.5 percent. Increased usage creates the demand for affordable internet services. Mobile and wireless networks can give more affordable and uninterrupted services. Figure 1 shows how mobile traffic has increased exponentially over the past decade. Extrapolating from this graph, we can predict that Internet mobile traffic consumption will continue to increase, and take up a continually growing percentage of Internet traffic over non-mobile devices.

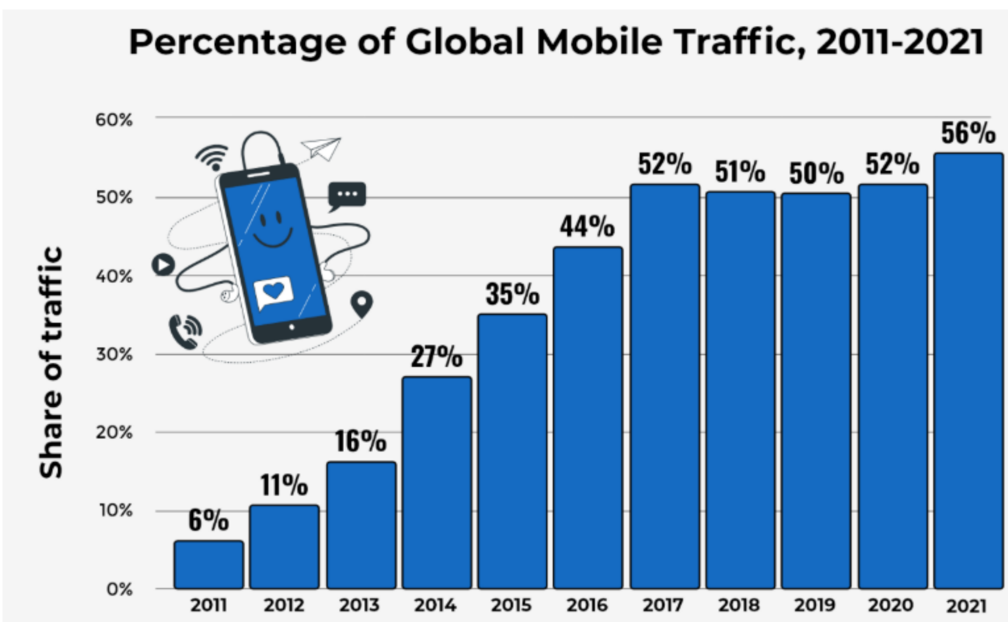


Figure 1.1: Percentage Mobile Traffic (Broadband search)

Internet service providers, especially mobile operators, have developed intelligent network infrastructures to cater to current and future internet penetration growth. 5G is the current next-generation mobile network in development, with deployment already occurring in limited geographies around the globe. The key functional services of 5G or next generation future mobile network technology will reveal a wide range of opportunities, better performance for cellular customer, as well as faster speeds and lower latency. 5G network is new global wireless standard after 1G, 2G, 3G, and 4G networks. The main features of fifth generation technology are to deliver higher multi-Gbps data speeds, lower latency, a more reliable network, vast network capacity and increased bandwidth. Vast quantities of data are constantly being generated from an extensive number of mobile phone devices and IoT devices. It is essential to collect and analyze the data to find behavior patterns that can lead to further work to provide better performance and optimize operation and management in cellular networks.

Advanced machine learning intelligence into the next generation mobile network platform is gaining unparalleled research interest. Machine learning-based solutions help to diagnose and identify problems ranging from radio access technology to network malware or intrusion detection. Machine learning intelligence can also contribute to the development of new wireless networks [5] [6]. Machine learning facilitates analysis of very efficient and systematically useful information from a huge number of traffic data and automatically uncovers correlations that would be very difficult to extract from some complex data by human experts. The main objective of this thesis project was to design efficient services for 5G and next generation network slicing architecture and extensive mobile traffic analysis for improving performance and quality of services. Our work also involved the implementation of machine learning algorithms to detect the anomalies and recognize the mobile traffic patterns for analyzing user experience as well as efficient network operation and management to ensure better performance. Moreover, dynamic selection of network slicing serves the desired purposes of requirements in efficient ways. Derived models of Machine learning techniques are achieved to evaluate the effectiveness of implementation of different algorithms to predict the traffic pattern, network policy, optimization, anomaly detection, ensure security and user classifications.

The thesis is divided into the following chapters:

a) Chapter 2 describes next generation wireless network technology, system design and management, standardization of 5G network and use cases.
b) Chapter 3 discusses mobile traffic, the architecture of next generation cellular network, end to end call flow into the entire network, division of networks, description of nodes in each network, different network layers and protocols to carry signaling and data.
c) Chapter 4 continues the discussion on next generation networks and key performance indicators (KPI).
d) Chapter 5 documents the log analysis and datasets, including deetaidl description about a variety of tools like QUALCOMM QXDM, QCAT, QPST and Wireshark tools we used to capture the network packets in each of the layers and analyze the data for improving network performance. We also discuss testing and automation tools using python-based frameworks like Pytest, Unittest and Robot framework.
e) Chapter 6 discusses the datasets and mobile traffic analysis using machine learning techniques.
f) Chapter 7 provides a short literature review.
g) Chapter 8 is on network security and anomaly detection, including intrusion detection system, Splunk tools, machine learning toolkits and describes possible attacks in any kind of network.

CHAPTER 2
5G TECHNOLOGY, IMPACT & SYSTEM DESIGN

2.1 Technological Diversity

The key functional services of 5G or next generation future mobile network technology will reveal a wide range of opportunities. 5G promises better performance for cellular customers, on top of faster speeds and lower latency. It reduces the latency for more responsive game streaming. This technology is not just about cell phones, it can also be devoted to residential WIFI, robots, VR, self-driving cars and more.

This 5G next generation future technology is different from previous generations based on the following technological concepts:

a) Millimeter waves: 5G uses extremely high frequency radio waves that provide more than gigabit speeds over short distances.

b) Beamforming: High precision antenna arrays capable of covering long distances and directing the wireless signal to individual devices

c) Massive MIMO: This is another new technology that is meant to deliver wireless data to devices at a higher capacity.

2.2 Use Cases

5G and next generation wireless technology will transform the world. It promises opportunity and benefits in every aspect of our lives. Here are some of the use cases of next generation wireless technology that will change our lives for the better.

2.2.1 Gadgets

We are already seeing the results of a smart revolution with self-driving vehicles, smart assistants like Siri, Alexa, Google’s Home assistance, smart devices that monitor our home security, home cleaning robots and many more. Everything is more instantaneous, connected, and accessible. In the era of instantaneous fulfilment, fifth generation will make amplified reality

(AR) and virtual reality(VR) experiences in real time whether it is in gaming or exploring, making everything will feel livelier and more instantaneous. The speed and lower latency of 5G network will eliminate lag in video streaming, especially in VR and AR devices. It makes for more realistic and immersive experiences.

2.2.2 Internet of Things (IoT):

5G cellular next generation wireless network will radically extend the reachability of the Internet. We can see the rise of smart cities, smart home, smart communities, smart education, smart power grid and internet of things (IoT)- together creates a huge number of devices and gadgets across the world that are connected to the Internet. For example, lightbulbs, thermostats, appliances, security cameras, parking meters, streetlamps, smart watches, home security systems and many more application-based devices have sensors that are constantly collecting data, sharing data, and communicating within the networks. 5G and next generation wireless network connect all of these devices for exchanging data. All these sensors will become universal and exchange data in real time that will impact every part of our lives. 5G wireless networks with new technologies will make our communities safer, our environments more convenient, our communications faster, and more connected.

2.2.3 Healthcare

A primary focus of 5G is to achieve real time speed, which will make critical contributions to the health care sector. Just think of a world where a doctor can conduct surgical treatment on a patient from another continent through robots. This scenario first happened in China using 5G network in real time speed, and it was successfully accomplished [46]. 5G network also gives patients an opportunity to obtain services of the highest quality from the best doctor in the world, regardless of where they live. 5G will enable Artificial Intelligence to ascertain the most suitable treatment and a diagnosis for a patient. Moreover, the healthcare industry can organize their tasks in the most efficient ways, like downloading imaging files, real time patient monitoring, gathering data for preventive care, and faster services for critical patients. All of these services will be transformed through new fifth generation wireless technology. 5G will save lives, time, and greatly improve entire health care systems for the global population.

2.2.4 Workplace

Next generation wireless technology and its unseen levels of efficiency will have a huge impact in the workplace. 5G technology, its potential data driven insight, and faster real time speed can make automation of many repetitive day to day activities very efficient, eliminating the need for many human efforts. The workplace will be more productive, and engineers can concentrate more time on research and development tasks to facilitate innovation. One common concern is that increased automation will lead to job losses. By using organized automation, employees can reduce their tedious day to day work that kills valuable time and allow them to dedicate more time to high-level tasks. 5G and next generation wireless network technology will also provide more opportunities to work remotely though real-time high-speed internet service, no matter how far away an employee is from the office.

2.2.5 Manufacturing

Manufacturing is another sector where 5G will have a strong impact. Next generation wireless technology will enable use of the IoT in industries. According to a Mckinsey report, IoT will add $11.1 trillion in value per year to industries by 2025 [47]. IoT will be able to collect data in every step of the manufacturing process. By implementing technological automation and robots, manufacturing processes can improve production quality and quantity and eliminate inefficiencies. With 5G wireless networks and its real time base access, companies will have more insight into their processes and get increased utilization from factories and workers.

2.2.6 Farming and Agricultural Industries

5G will also significantly impact the farming industry by enabling implementation of IoT devices that can monitor livestock and ensure their proper growth and health, or detecting herd diseases. IoT devices, sensors, or drones can provide real time monitoring of crop production, harvesting, water or seed distribution. 5G can also contribute towards collecting data on various factors that impact crop growth including soil testing, humidity and temperatures, water levels and distribution, and insect detection, to ensure proper cultivation to the final stages of crop production. Devices will connect 5G networks to monitor and collect the data fast, and ensure that production quality and collected information are accurate. Population is growing while land use is decreasing proportionally. In this situation, using the latest 5G technology helps farmers to produce the largest possible crop yields and utilize their most out of their land. IoT devices will help farmers to collect real time data to grow their crops using less water, pesticides, and fertilizers.

2.2.7 Smart Grid Automation

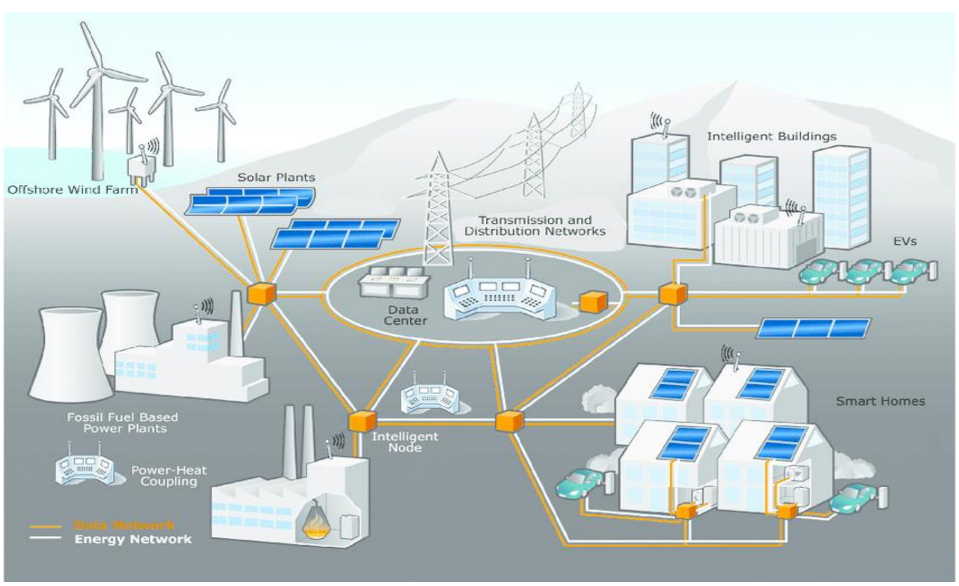


Figure 2.1: 5G use case in Smart grid automation

Primary benefits of the implementation 5G wireless technology in the energy and utilities sector include real time management and automation of the smart electricity grid. The aim of this solution is optimizing the operation and maintenance activities, quick fault detection and resolution, increase of overall efficiency, cost reduction, minimization of losses, and real time control and automation. Nokia, ABB and Kalmar have successfully established 5G-based URLLC for smart electricity grid.

2.2.8 5G wireless technology in intelligent transportation system

Figure 2.2: 5G in Intelligent and safety traffic system

5G wireless technology is expected to have a huge impact in transportation domains by sharing real time information about traffic and road conditions. IoT or sensors will collect and share real time data for intelligent and safety traffic management. In South Korea, SK telecom is building a 5G-enabled safety traffic management and intelligent transport system in Seoul.

2.3 The Future of Connectivity

Mobile internet traffic is expected to grow 40% a year through 2025. 75% of internet traffic will be used for video consumption and 60% of the worldwide population will be online by

2025. The trend of the last 5 years seems to indicate that internet traffic will go beyond expectations, and SIM connections are likely to double. Humans are competing with intelligent machines for internet bandwidth. Huge volumes of data contain valuable information about customer usage patterns, consumer behaviors, consumption rates, etc…

5G has a critical role in recognizing the full prospects of the IoT, Internet and other evolving technologies. 5G is expected to expand 350% over the the next 5 years. Figure 2.3 shows how 5G drives this evolution.

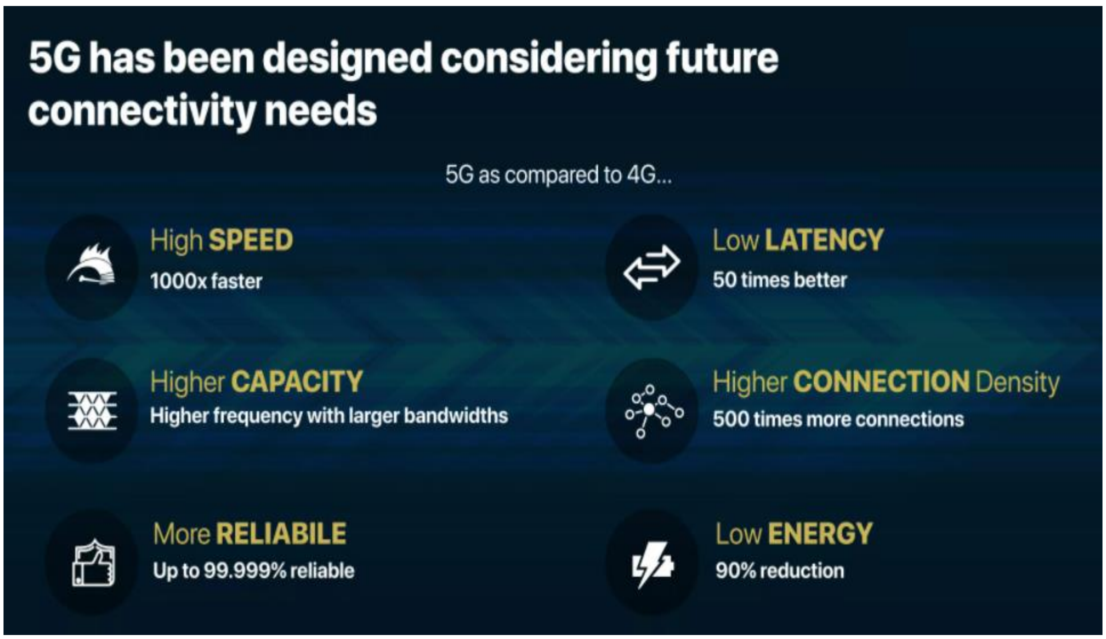


Figure 2.3: 5G Technology and future connectivity

5G technology or next generation future wireless technology is envisaged 50 times faster than 4G while it is using higher frequency spectrum and has been augmented though small cell. Other features 5G include:

- Higher Speed: 5G is expected to be 1,000 times faster, facilitating to download a entire HD film in less than 10 seconds.

- Lower latency: 5G is expected to 50 times better than 4th generation that can reveal the potential of augmented/virtual reality (AR/VR) and autonomous driving technology.
- High Capacity: 5G technology will attain higher frequencies and capabilities for higher bandwidth.
- High density connection: Fifth Generation supports higher densities of connection capable of connecting massive IoT devices.
- More Reliable: 5G wireless network is up to almost 100% reliable, which is essential for providing time sensitive data.
- Less Energy consumption: 5G consumes less power, leading to longer battery life (almost 90% reduction) in low power IoT devices.

Fifth generation wireless technology can be used in both consumer-led and enterprise-led systems. Consumer-led systems can serve mobile internet, broadband internet, vehicular communication systems, entertainment-gaming, self-driving cars and mHealth, all of which require high speeds, higher bandwidths, lower latency, and reliable connections for mobility. On the other hand, the use cases of enterprise-led systems such as remote surgical treatments, critical remote-control devices, smart power grids, smart communities, smart homes, smart cities, human-machine interaction (HMI), and smart industrial units would require a massive number of IoT devices, higher data rates, high quality of services, higher mobility, and high bandwidth to manage all these connected sensors and devices.

2.4 The Fifth-generation Architecture

The traditional mobile network was designed to meet the requirements for voice and typical broadband internet services, and lacks the capabilities to serve multi-connectivity technologies, orchestration of virtual network functions, and manage deployed applications. 5G wireless network architecture is designed to be more flexible and meet the requirements of mobile diversified services very efficiently. Figure 2.3 shows how the 5G network is divided into 5 layers of services.

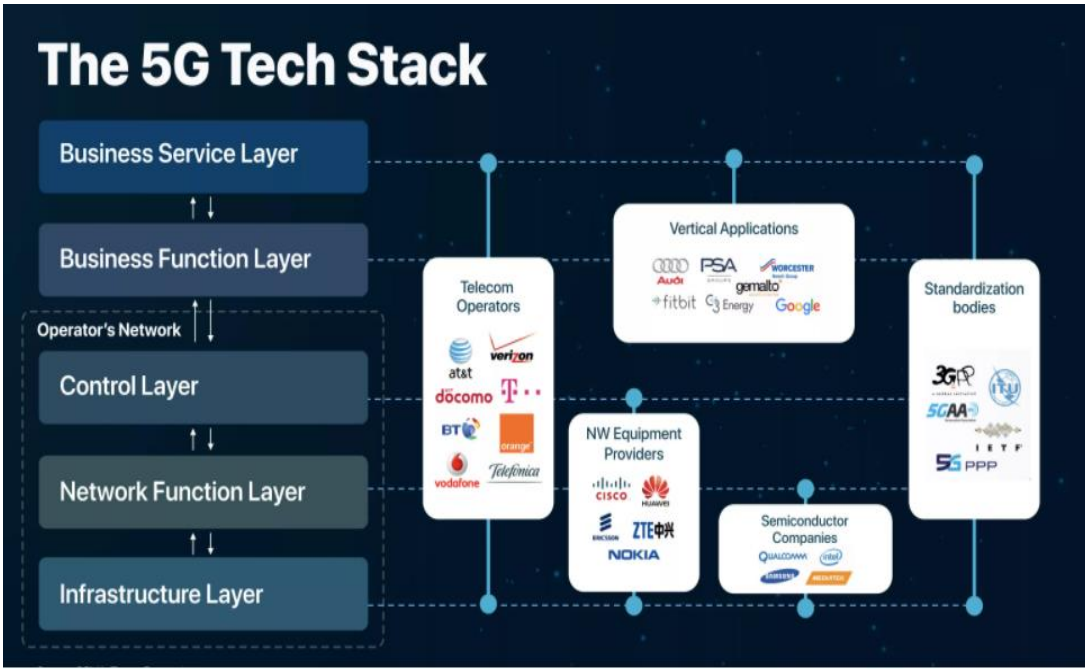


Figure 2.4: The 5G tech Stack

The 5G network layers are grouped by operators. The top layers are business function and service layers. Business service layers provide mobile broadband and vertical services to customers, and business function layers serve application-based virtual function repositories. The operator's networks consist of three layers—control layer, network function layer and infrastructure layer. The control layer allocates computing and network resources to meet required services. The network function layer provides computing and networking functions like software defined network (SDN) and network functions virtualization (NFV), also called "network slicing." The infrastructure layer includes the physical tower, radio access network and optical core network.

2.5 Summary of How 5G is Different From Its Predecessors

5G implementation is completely unlike previous cellular generations. 5G network features:

- Cloud native-based application
- Service based architecture
- Network functions virtualization
- Software defined network
- Network softwarization
- Design patterns are microservice based
- DevOPS, CI/CD methodologies
- Network slicing architecture for proper and better utilizing physical network resources
- Edge computing-based mobile access for delivery and lower latency
- high density connections
- The control plane and user plane are separated for faster communication

2.6 Three Essential Technologies in 5G Network

5G brings 3 major technology-enabled benefits: 5G promises substantial speed, a quantum leap capacity, capablities to manage critical systems in mobile technology, healthcare, transportation, industry, and in network dense environments such as smart homes and apartments with billions of connected devices, sensors. Even the clothes we wear may be on the network. The emerging technologies in the 5G domain include:

- **Enhanced Mobile Broadband (eMBB):** Enhanced mobile broadband provides high data transmission rates with wide coverage areas.
- **Ultra-reliable low latency communications (URLLC):** Increased low latency, the amount of time a signal sent receive data, is needed for reliability in mission critical communications such as remote surgery, autonomous vehicles.

- **Massive Machine Type Communications (mMTC):** Massive machine type communications can support a large number of devices in small area remarkably. To support large number of devices, the 5G network works on an expanded radio frequency. In 4G networks, the radio frequency was below 6 GHz, whereas 5G network radio frequency is more than 28 GHz. Another important technology of 5G network is base station antenna, which is called massive multiple input and Multiple output. This technology eliminates the problem in mmWave technology where communications do not travel in densely populated areas due to obstacles, such as a building that interferes with the signals.

2.7 5G Network Materials and Components

To enable 5G and all of its potential applications, electronic manufacturers and software developers are working hard to develop high level components for 5G infrastructure. These components will have super potential to overcome a number of challenges like delivering high performance at very low power, and small encapsulated material size capable to endure high thermal temperatures. A number of high-level components are being used to build the 5G infrastructure.

5G Antenna: In simple definition, antenna is an electronic device converting electric power into radio wave. There are many types of antennas in cellular market like monopole and dipole, microstrip, MIMO and patch antenna. MIMO is becoming more popular in terms of powerful, greater capacity and long-distance coverage. 5G network using most power antennas that heave greater capacity and wider wireless spectrum. These are the antennas made quite advanced ceramic materials that include a mix of materials like silicon dioxide, yttrium oxide and barium carbonate. With mixing all of these raw materials with appropriate ratio and then synthesized the desire outputs.

Antenna Radome: Radome is a shielding component that protect antenna from extreme weather condition like rain, heat and ice. This Radome is made by combination of different kind of plastic. Radome is not only protecting the extreme weather it also saves the maintenance cost and reduced fabrication.

Fifth generation Microwave circuits: 5G microwave circuits are built by using super conductor and ceramic materials. This microwave circuits can supply electrical signals like current and voltage, stipulate better stability, greater frequency and nominal temperatures. 5G networks use millimeter wave (mmWave) spectrum for increasing wireless data transfer. Another component is called Absorvers used to control the electromagnetic interferes in mmWave, counter the effect of radiation signal and performance of overall microwave circuits.
Power Amplifier: Power amplifier are used in 5G network to enhance the power level of a given input signal. This is one of the important components in 5G network is made of semiconductor gallium nitride and polybutylene PBT. This power amplifier can run at higher frequencies with very minimal power rates. Indium Phosphide is also used to enhanced performance at higher frequencies while consuming nominal power rates.

5G cable: 5G networks are consider as wireless network but infrastructure still requires high quality cable, Cat 6A 5G SxTP cable are designed particularly to provide application in both indoor and outdoor. These cables are high density tinned copper braids that can support for voice and data with higher rate. These copper cables are sealed using flame retracing, heat conducting, foaming and other sealing materials. These cables are very less weight and sealing materials includes liquid silicon, rubber-based glue, a heat conducting powder, a foaming and auxiliary agent.

Circuit board substrates: 5G network use circuit board substrates and alkali-free glass fiber cloth and polytetrafluoroethylene (PTFE) are used as raw materials. Some other materials that also used in circuit board substrates like polyether ether ketone (PEEK), modified epoxy, polyphenol ether (PPE), phenolic resins, glass and flexible ceramics. These circuit board substrates are used in 5G network that can provide many benefits like light values of dielectric constant (Dk), a dissipation factor (Df) and moisture absorption. It is also cheap and low manufacturing costs.

Fifth Generation Network Slicing: 5G network slicing is the use of network virtualization to divide single network connections into multiple distinct virtual connections that provide different amounts of resources to different types of traffic. Network slicing is a type of virtual networking architecture in the same family as software-defined networking (SDN) and network functions virtualization (NFV) — two closely related network virtualization technologies that are moving modern networks toward software-based automation. SDN and NFV allow far better network flexibility through the partitioning of network architectures into virtual elements.

CHAPTER 3
MOBILE TRAFFIC AND CALL FLOW

3.1 Fifth Generation Network Diagram and Call Flow

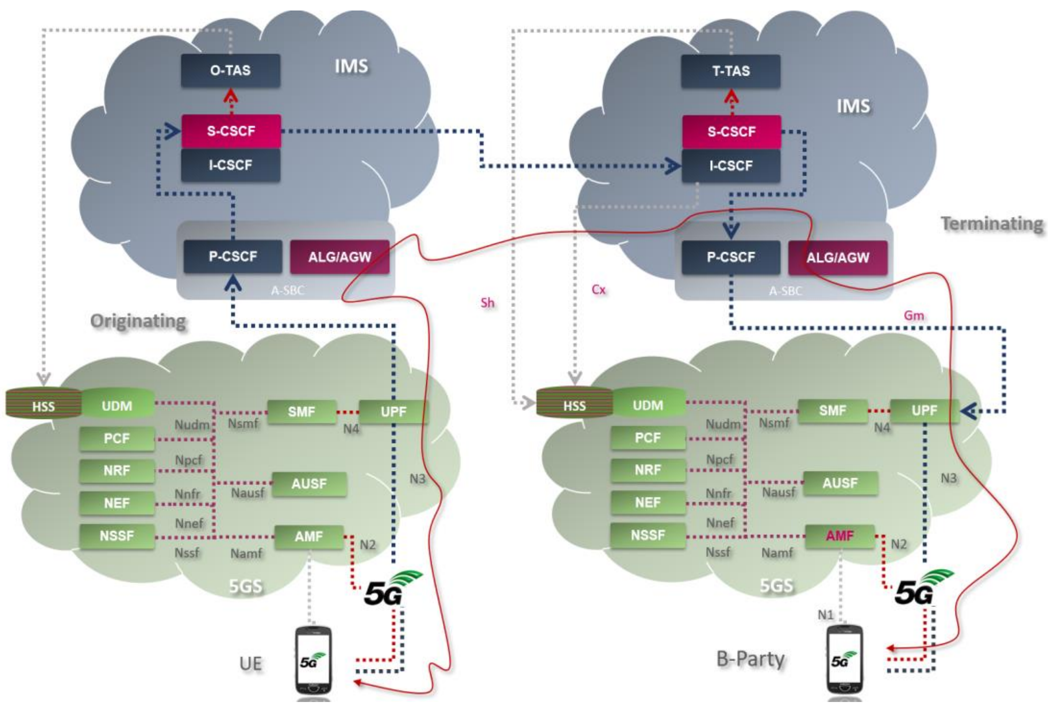


Figure 3.1: 5G Network Architecture and Call Flow

Figure 3.1 shows a complete next generation network architecture documenting how users connect with the network to get the different kinds of desired services, like audio or video calls, text messages, and Internet connections. As an example, User A (called A-party) can place a call to B-party. Both subscribers can use the same networks or different networks at any place in the world. When a call is originating from UE from A-party, first the call is received to eNB, or radio base station, the the call goes through the Mobile Core network. The core network consists of different elements and each of the elements has a unique function. In the core network, the first entry node is access mobility function (AMF), then the authentication process selection of next node session management function. There are other elements that deliver network functions, like

PCF for “Policy and Control Function,” or NSSF for “Network Slice Server Function.” Most importantly, UDM and AUSF contain all kinds of subscriber information, including the location. Finally, the call connects to the IMS network, called “IP multimedia platform.” In this network, call session control function is the core element, and other network elements also perform unique functions to complete the call from originating to connect terminating network. Terminating network consists of the same network elements of the originating network, only the call is terminating to B-party from IMS to core network then B-party user equipment. Similarly, if any subscriber wants to connect through text message or internet access, the same networks are used to connect their desired services.

End to End Call flow in 5G network:

Access Network

- User Equipment’s: Mobile Phones, Tablets, IoT devices etc.
- 5G gNB

5GC Core Network

- 5G Core Network Elements: AMF, SMF, UPF etc.

IMS core network

- IMS network: P=CSCF, ICSCF, CSCF, TAS, ALG/AGW etc.

3.1.1 Description of call flow in each network

Access Network UE: User equipment where call registration starts. UE can be mobile phones, tablets or any kind of device. A party UE initiates the call and B party UE receives the call. Figure 3.2 shows how UE performs interactions during standalone registration procedure.

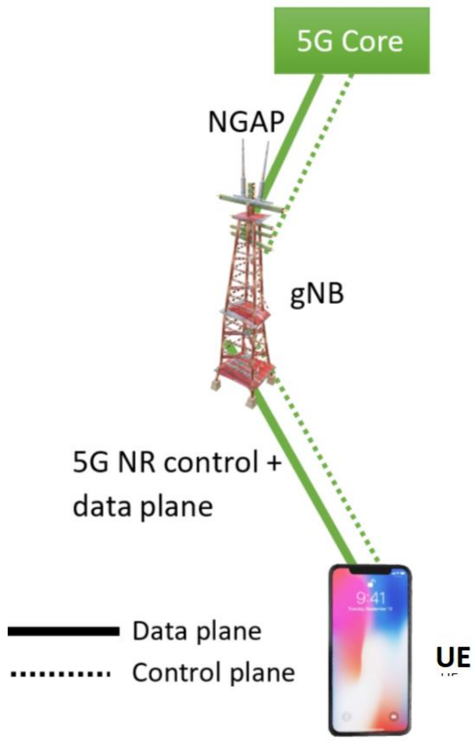


Figure 3.2: Call Flow from Access Network to Core Network

Figure 3.2 describes how a call connects from UE to the mobile core network. The step by step call flow in different network protocol layers includes:

1. First, perform the random-access procedure to initiate the communication with gNB.
2. Next, set-up an RRC connection with gNB. A SRB or signaling bearer is set up at this point. UE also sends a registration request to 5G core via NAS protocol.
3. After that, perform NAS level authentication and initiate ciphering for the NAS message between UE and 5G core network.
4. Complete the AS security procedure with gNB.
5. Handle RRC reconfiguration from the gNB. This message sets up the default PDU session. The message also may add secondary cells.
6. Finally, UE accomplishes the registration procedure. Then data is flowing in the downlink and uplink directions.

Above steps are the first steps to initiate the call registration between UE and gNB. gNB acts as a channel for communication between UE and 5G Core as follows:.

1. gNB handles a random access request from UE and allocates resources for setting up the RRC connection.
2. Establish RRC connection with signaling bearer one called SRB1. Also, gNB starts assigning downlink and uplink resources to the UE through PDCCH.
3. In this step, gNB transfers the registration request from UE to AMF. AMF is a border node of the 5G core network.
4. NAS signals authentication message carried by gNB and establishes a ciphered link between UE and AMF.
5. The 5G core initiates the PDU session setup and also receives a registration accept request from UE.
6. The gNB also enables security in between UE and gNB
7. After security setup, gNB activates the default PDU session though RRC reconfiguration message. It also sends registration complete message to AMF.
8. At this stage, downlink and uplink data stream flow is created between UE and the internet.

Figure 3.3 shows the flow chart of call registration from UE to RAN to 5G core network, and the final stage activation of the default PDU session. This diagram shows how call or data connections flow among all network elements with different network protocols layers.

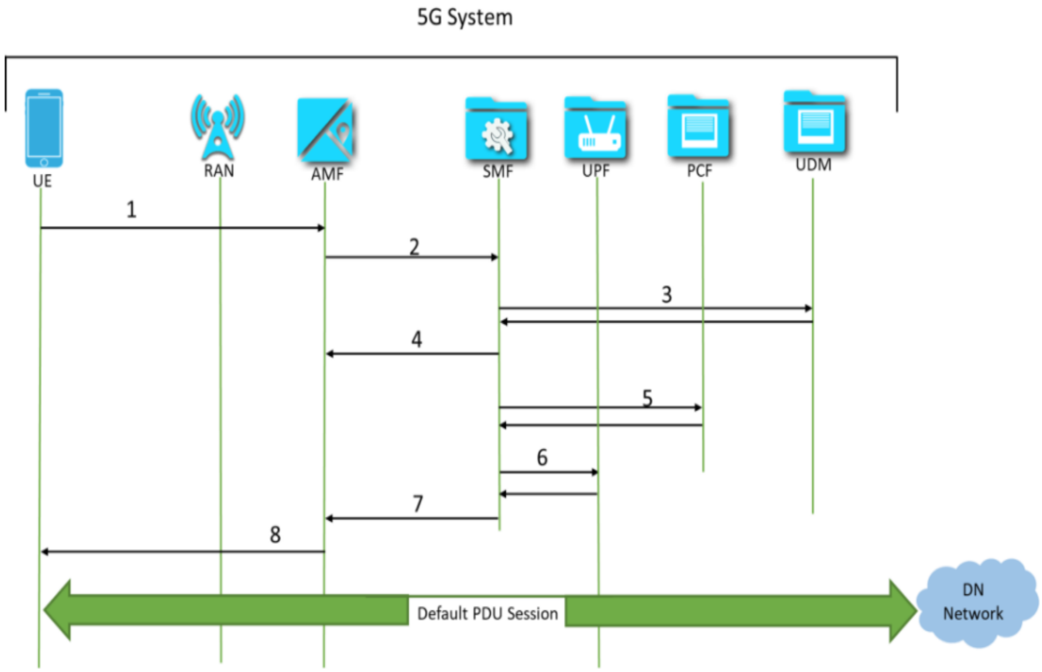


Figure 3.3 Default PDU session establishment

3.2 5G core interactions

The first node of the 5G core network is Access and Mobility Function (AMF). AMF is the first border node acting as a bridge between access and 5G core networks. It synchronizes the message exchange for the call registration process as follows:

1. UE sends registration request message with some user parameters via gNB to AMF. AMF initiates the UE message from gNB.
2. AMF communicates with AUSF for authentication request. After that, NAS level security is established between the UE and AMF.
3. After getting the reports from AUSF, AMF checks the 5G-EIR to verify the validity of registration.
4. AMF then attains subscription data from UDM.

5. AMF creates policy and checking validity with PCF
6. AMF sends initial context request to SMF for activating the default PDU session.

AUSF: Authentication Server Function is a home network and performs authentication by obtaining the UE authentication information from the UDM.

UDM: User Data Management provides authentication directions during the registration process. It contains all subscriber data management information.

PCF: Policy Control Function registers with AMF so that it can notify on events like location change and communication failure.

SMF: Session Management Function serves as control plane entity for session management. SMF performs the following interactions:

- OLD AMF deregisters with the session management function node.
- The new AMF sends session management context, then SMF assigned IP address and tunnel ID to send uplink data.
- The SMF also selects the UPF for this session.
- SMF also communicates to new AMF when the session is ready for uplink and downlink data transfer.

UPF: User Plane Function is a data plane component that handles user data. SMF uses the packet flow control protocol to update the data plane.

3.3 SIP/IMS Registration

After UE is attached to the network and PDU sessions are completed, UE must register with the IMS network. The IMS registration is via SIP protocol, and includes IMS authentication and security negotiation between UE and IMS. IMS SIP/registration procedure will be completed in 2 attempts. In the first attempt, IMS challenges user agent in UE and in the seond attempt, UE will be registered in the IMS network. The following flow chart contains first and second attempts in the SIP/IMS network.

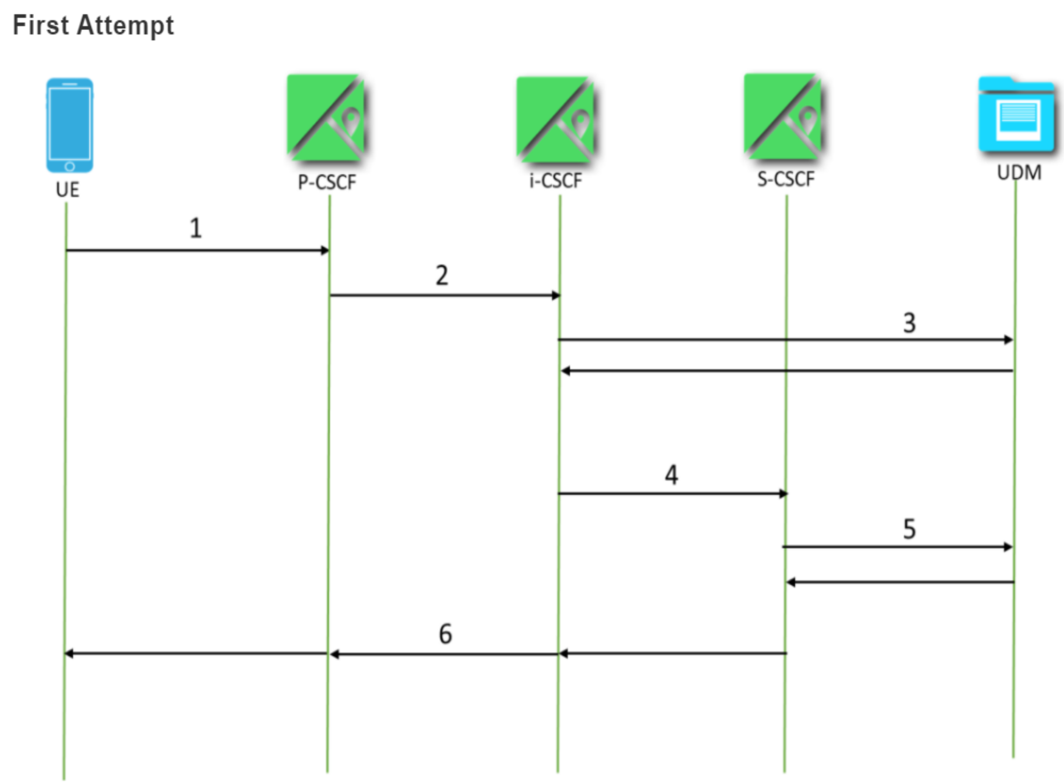


Figure 3.4: First attempts SIP registration without authentication.

Figure 3.4 shows how calls are processed in IMS platform with SIP protocol and how the elements of the IMS network interact.

1. First UE sends SIP registration request to PCSCF with IMPI, IMPU and home network domain SIP URI.
2. P-CSCF selects the I-CSCF and forwards the SIP registration message to I-CSCF.
3. I-CSCF then starts queries to UDM for authentication and selects the SIP registration CSCF. When UDM responds queries to I-CSCF with authentication information and CSCF for SIP registration.
4. CSCF obtains authentication vector and subscriber data from UDM. In the first attempt, the user is not authenticated and therefore S- CSCF will reject the

SIP/registration request with 401 error code and challenge to UE for verifying the authentication with some parameters.

5. I-CSCF and P-CSCF will forward the responses from CSCF to UE for the verification with authentication parameters.
6. In the second attempt, after completing the established IP-sec connectivity with P-CSCf, UE sends back a message with authentication parameters to P-CSCF.
7. Same process occurs and after cross checking parameters I-CSCF and S-CSCF with UDM, S_CSCF sends 200 ok and creates dedicated bearer.

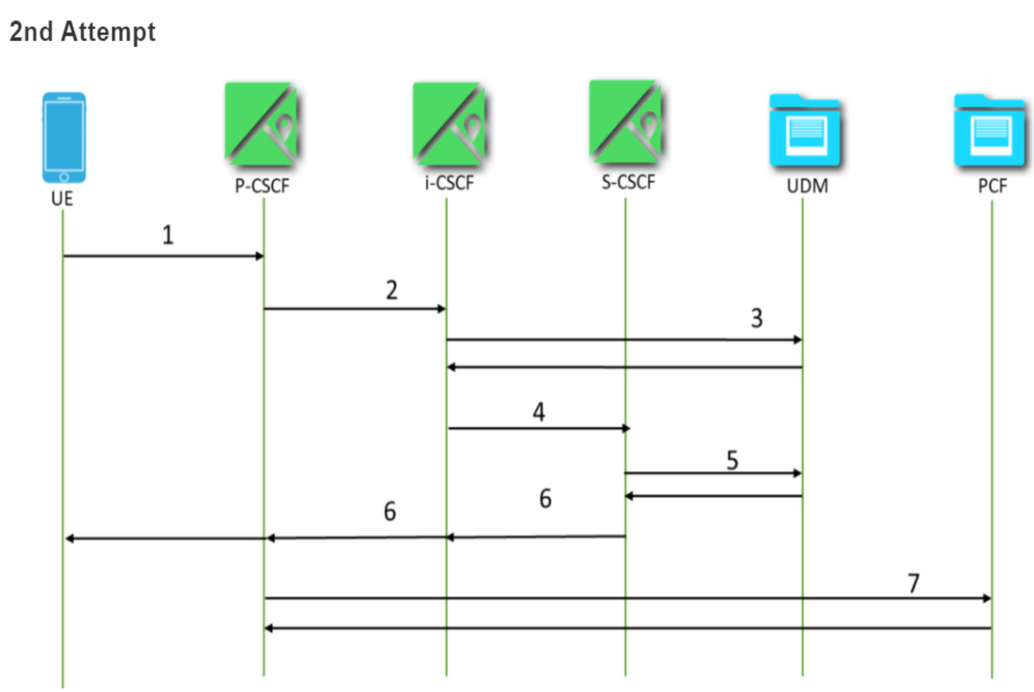


Figure 3.5: Second attempts SIP/registration procedure with authentication parameters

Figure 3.5 shows call connection using SIP registration procedure with authentication parameters.

3.4 A Complete Call Flow in IMS Platform

Figure 3.6 displays how a call is set up between two users in 5G VoNR and registered in the IMS network. The diagram also shows how calling party and called party will contact the policy control function to activate the dedicated bearer towards UE via new radio interface.

A step by step SIP/registration call flow between calling party and called in 5G users to 5G users involves:

1. First, the calling party initiates the voice call by sending SIP invite to P-CSCF, then P-CSCF to S-CSCF. Each of the node's responses with 100 trying. In the initial SIP, invites include SDP protocol containing bandwidth, codec information and all other parameters.
2. B party receives the SIP invite with SDP information and verifies all information. B party then gets back 183 sessions in progress message.
3. Both A party and B party create dedicated bearer QC1=1.
4. A party sends PRACK, which contains final information with codec information. B-party acknowledges the PRACK by 200 ok.
5. B party starts ring message to UE by sending 180vringing message. When B party picks up the call, RTP protocol starts, and call conversation starts.
6. The same process follows when calling a party from 5G network and called party in LTE network pr PSTN network.

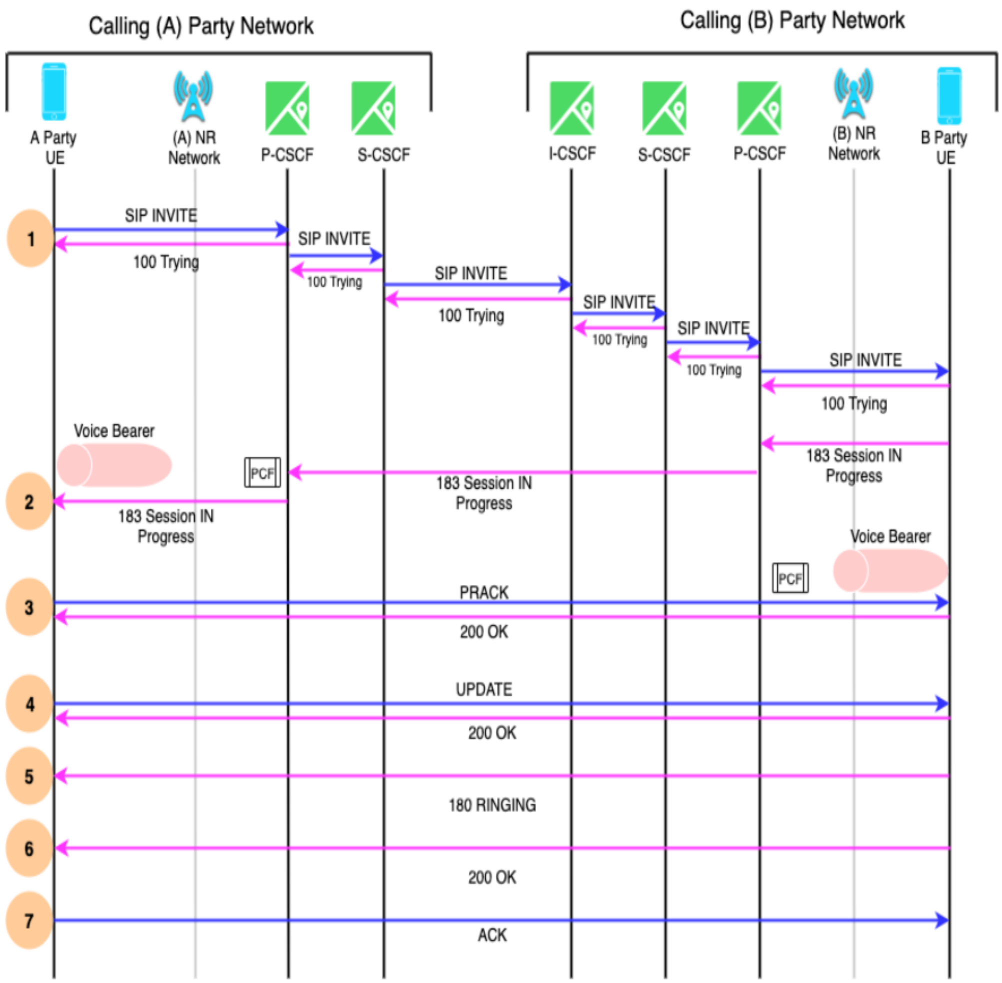


Figure 3.6: A complete call flow in 5G-to-5G user

3.5 5G Protocol Stack

There are two divided protocol stack planes, the control plane and the user plane. The value of user plane and control plane are separated in 5G network. User plane is also called the data plane, which carries network traffic. User plane or data plane refers to the functions and process that forwards network packets from one interface to another network, whereas control plane carries signaling traffic, controls the packets for routing, switching from one route to another route, and run authentication control element. In 5G network control plane and user plane are separated for better management and are developed independently. Software control of the network can develop independently of the hardware.

3.5.1 Control plane

In the figure 3.7, we can see different network layers with different protocols. Each of the network layers serves a unique function in the network. Control plane generally carries signaling traffic, as well as control operations like network attaches, security control, authentication, bearer setup and mobility management. Control plane carries signaling regarding whether the phone is in idle or service modes. Control plane protocol stack between UE and gNB includes radio resource control layer (RRC) which is responsible for configuring lower layers. In the control plane RRC protocols carry the signaling message between mobile phones and base transceiver stations, and exchanges signaling between gNB or base station to Core network using NG-AP signaling protocol.

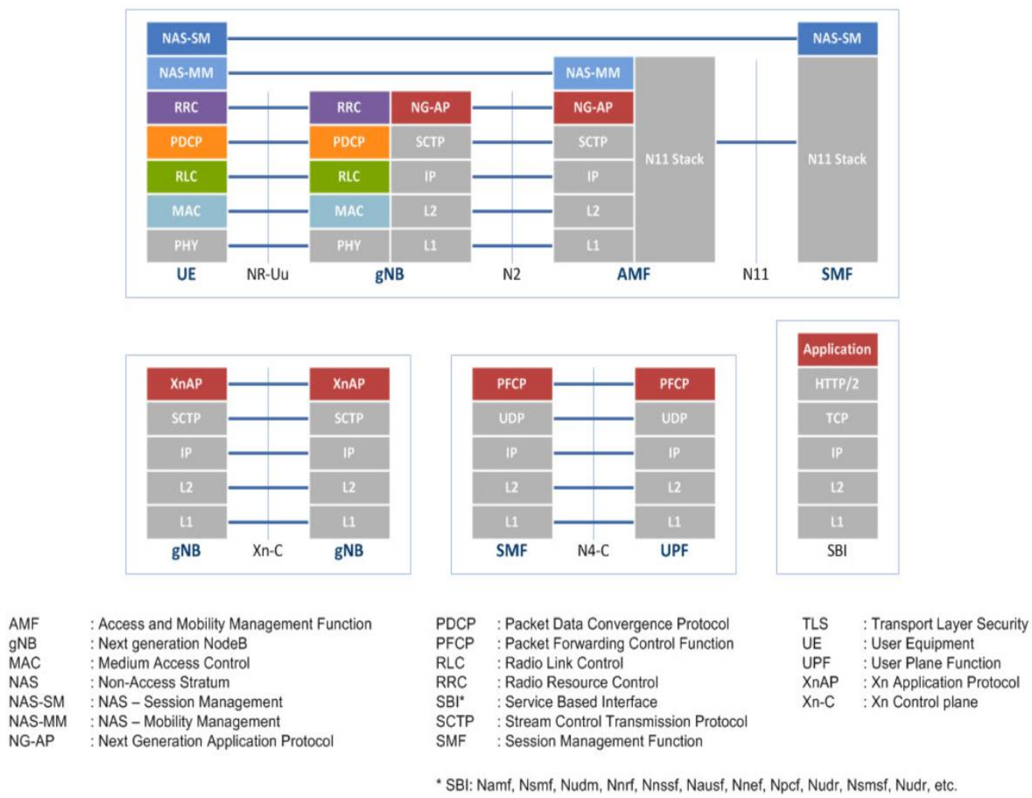


Figure 3.7: Control plane 5G network protocol stack

3.5.2 User plane

Figure 3.8 shows different network layers with protocols. User plane carries network user traffic. The user plane is also called data plane. When a user connects the calls or data, the user plane is responsible to carry the traffic. The user plane protocol stack between the UE and gNB comprises sublayers like service data adaption protocol (SDAP), packet data convergence protocol (PDCP), radio link control (RLC) and medium access control (MAC). In the user plane, data packets are processed by protocols like TCP, UDP and IP. GPRS tunnel protocol establishes the IP based communication protocols.

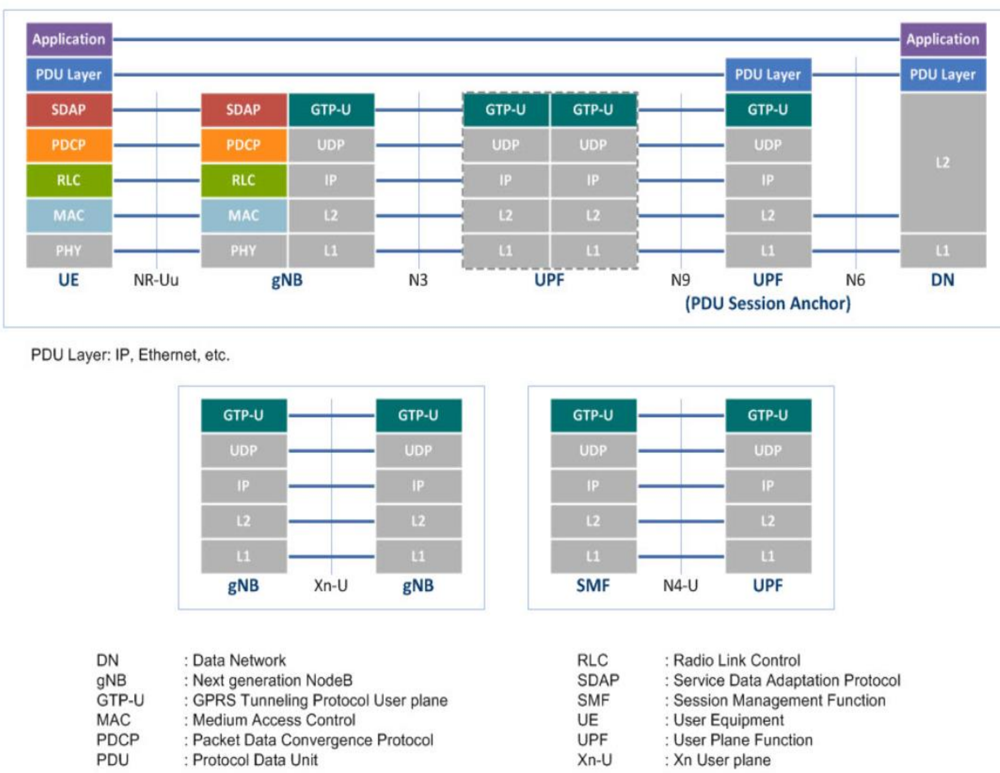


Figure 3.8: User plane 5G network protocol stack

3.6 Description of Each Protocol Layers

5G network user plane and control plane protocol stack are the layers that involve communication from UE to gNb, from gNB to 5G core network, and vice versa. User plane and control plane protocol stack are the same in certain levels of layers. They are separated in 5G network. 5G network protocol stack divided by three layers:

1. 5G new radio layer 3 is called radio resource control (RRC) protocol. The RRC protocol is signaling exchanged between mobile and evolve Node base station. The main services and functions of RRC layers are:
    - Broadcast the information between AS and NAS system
    - Initiate the paging by 5GC or NG-RAN.
    - RRC connection between UE and NR-RAN including establishment, maintenance, and releases
    - Security functions
    - Signaling radio bearer and data radio bearer establishment, configuration, maintenance, and releases
    - Management and control UE cell selection and reselection, mobility function like handover and context transfer
    - Inter RAT mobility
    - Quality of management functions
    - Reporting and control of UE measurement
    - Detection and recovery radio link failure
    - NAS message transfer to/from UE

    5G NR layer 2: The layer 2 of 5G new radio split into following sublayers:

    - Service data adaption protocol (SDAP)
    - Packet data convergence protocol (PDCP)

- Radio link control (RLC)
- MAC (Medium Access Control)

The main function and services of SDAP protocol are:

- It offers 5GV QoS flow
- Maps between QoS flow and a DRB
- Makes QoS ID in both uplink and downlink packets
- SDAP is also configured for each individual PDU session

PDCP sublayer function and services as follows in both control and user plane:

- Sequence Numbering
- Header compression
- Transfer of user and control plane data
- Ciphering, deciphering and integrity protection
- Retransmission of PDCP SUDs
- Duplicate detection and duplication of PDCP PDUs

The function of RLC layer:

- Transfer of upper layer PDUs
- Error correction, segmentation, and DE segmentation
- RLC SDU discard
- RLC re-establishment

2. MAC layer is in the lower level of Layer 2 Protocol stack. The function of MAC layer follows:
    - Plotting between logical and transport channel
    - Multiplexing and demultiplexing of MAC SDUs
    - Scheduling information

- Error correction
- Priority handling of UEs
- Priority handling between logical channel
- Padding

3. Physical (PHY) layer is last layer of the protocol structure. The main function and services involve communication channels between UE and 5G core network. Also, modulation and beamforming are done by this layer.

CHAPTER 4
MOBILE NETWORK KEY PERFORMANCE INDICATORS (KPI)

4.1 Categories of KPI for 5G Network Performance

Key performance indicator (KPI) is a detailed analysis of the performance and evaluation of network quality. KPIs describe how effectively a network is serving users. Much like a doctor can affirm the fitness and wellness of a human being by measuring some parameters like temperature, blood pressure, heart rate, sugar levels, etc...in a similar way, network performance or fitness can also be measured by data like network attach success rate, PDU session rate, Average UL/DL data rate, mobility and other network performance-based parameters. 5G network is designed as a high performance with low latency. greater capacity, and faster speed. 5G KPI is measured under following categories:

- Accessibility
- Retainability
- Availability
- Mobility
- Integrity
- Utilization

3GPP Specification TS 28.554 Management and orchestration defines 5G end to end key performance indicators mainly describes for Accessibility, Integrity and Utilization.

4.1.1 Accessibility KPI

If the user can not register to the network slice instance, they cannot access the network services in the network slice instances. In the 5G network, subscribers register to the network or network slice instances through Access Mobility Function (AMF) node. KPI parameters are considered as follows:

- **Number of users through AMF**

Name: Number of subscribers registered to the single network slice instances through AMF

Description: Total number of subscribers that are registered to the network slice instances

Logical formula: Counting total number of subscribers registered in the network slice instances through AMF

Physical formula: $$RSSNSI = \sum_{AMF} RegisteredSubNbrMean$$

Number of subscribers through UDM

Name: Number of subscribers that are registered in network slice instances through UDM

Description: Total number of subscribers register in the network slice through UDM

Logical formula: Counting the subscribers in UDM

Physical formula: $$RSSNSI = \mathrm{Re}\,gisteredSubUDMNbrMean$$

Registration success rate pf one single network slice instance

Logical formula: KPI is calculated by successful registration procedures divided by attempt registration procedures.

Physical formula: $$\mathrm{RSR} = \frac{\sum_{Type} \mathrm{AMF.5GSRegisSucc.}Type}{\sum_{Type} \mathrm{AMF.5GSRegisAtt.}Type} * 100\%$$

4.1.2 Integrity KPI

One of the main targets in 5G wireless network service is lower latency. End to end latency in 5G wireless network and Integrity KPI in 5G network describes end to end packet , latency in 5G network, Transmission is part of 5G network and is also used to analysis and evaluate the utilization of end-to-end networks. The KPI is calculated end to end latency of UE IP packets from UE through the N6 interface in the 5G network. N6 interface is considered in between UPF and DN. In summary, Integrity KPI is concerned with E2E latency, 55Gs, integrity, time interval in each point, and MEAN.

Downlink latency in gNB-Du

Here KPI is calculated downlink latency for IP packet from UE to gNB in split scenario. Since gNB is part of packet transmission in 5G network, it calculates and evaluate the gNB latency to contribute the total packet latency.

Logical formula: This KPI is the arithmetic mean of time from reception of IP packet to gNB until the first part that packet transmits over the air interface. Also consider the packet arriving time when there is no previous data in queue for transmission to the user equipment.

Physical formula: DownlinkLat = DRB.RlcSduLatencyDl or optionally DownlinkLat.QoSx = DRB.RlcSduLatencyDl.QoSx where QOS identifies the target quality of service class.

Unit and type of the KPI: Time interval (Millisecond) and MEAN

Use cases: gNB-DU latency KPI use case, end to end latency is an important and challenging performance parameter in fifth generation wireless network. If there is insufficient end-to-end latency in 5G network, customer will not get guaranteed network performance. Therefore, it is necessary to measure and evaluate the end-to-end latency to ensure the satisfaction of customer using their network.

Upstream Throughput for network and network slicing instance

Logical formula: The KPI is calculated by quantity of upstream data in the N3 interface from NG-RAN to UPF associated with single network slice instances.

Physical formula:

$$UTSNSI = \sum_{AMF} GTP.InDataOctN3UPF$$

Unit and type of KPI: kbit/s and collective measurement

Downstream throughput for a single network slice instance

Logical formula: The KPI is calculated by quantity of downstream data in the N3 interface from NG-RAN to UPF related with single network slice instances.

Physical formula:

$$UTSNSI = \sum_{UPF} GTP.OutDataOctN3UPF$$

Unit and Type of KPI: kbit/s and cumulative measurement

Downstream throughput at N3 interface

Downstream GTP data throughput at N3 interface. The KPI is calculated the total number of all downstream GTP data packets on the N3 interfaces generated by GTP-U protocol during granularity period.

Logical formula: The KPI is calculated by total quantity of downstream GTP data on N3 interface from UPF to NG-RAN.

Physical formula: DGTPTS=SUM (GTP.OutDataOctN3UPF)/timeperiod) at UPF

Upstream throughput at N3 interface

Upstream GTP data throughput at N3 interface. The KPI is calculated the total number of all upstream GTP data packets on the N3 interfaces generated by GTP-U protocol during granularity period.

Logical formula: The KPI is calculated by total quantity of upstream GTP data on N3 interface from UPF to NG-RAN.

Physical formula: UGTPTS=SUM (GTP.InDataOctN3UPF)/time period) at UPF

Unit and type of KPI: kbit/s and MEAN

RAN UE throughput

A KPI shows how next generation RAN influences the service quality to an end subscriber.

Logical formula: The KPI is calculated the volume of payload data on RLC level per elapsed time unit on the air interface, for restricts the transfer data by air interface.

Physical formula: RAN UE Throughput DL = DRB.UEThpDl and RAN UE Throughput UL = DRB.UEThpUl

or

optionally RAN UE Throughput DL for single mapped 5QI or QCI = DRB.UEThpDl.QoS and RAN UE Throughput UL for single mapped 5QI or QCI = DRB.UEThpUl.QoS

The values of DRB.UEThpDl, DRB.UEThpUl, DRB.UEThpDl.QoS, DRB.UEThpUl.QoS

Unit and types of KPI: kbit/s and MEAN

Use cases of UE throughput KPI in NG-RAN: The KPI of UE throughput in NG-RAN is potential parameter for key performance in fifth generation wireless mobile network. If you see that UE throughput of new radio cell cannot meet the performance requirement, you can fine-tune by reconfiguration and increases the capacity. When network slicing supports by the Next generation RAN network, the UL and DL UE throughput for each network slice plays very important role to measure the specific performance problem.

4.1.3 Utilization KPI

This is another important KPI category in 5G wireless network. The utilization of KPI depends on “mean number of PDU session in the network and network slice instance. Utilization KPI considered on following parameters:

Average Number of PDU session of the single network slice instance. The KPI refers to mean number of PDU session that are successfully found in the network slice instance.

Logical formula: The KPI is calculated by the establishment of successful PDU session measures in SMF associated to each network slice instance.

Physical formula: PDUSesMeanNbr=Sum (SM.SessionNbrMean.SNSSAI) over SMFs.

Some other KPI parameters are: PDUSessionNum, 5GS, UtilizationInteger MEAN

Use case: It necessary to evaluate the MEAN number of PDU session to measure the system level load or capacity. If the MEAN number of PDU session is high system level capacity also increase proportionally.

Network slice virtualized resource utilization

Virtualized resource utilization of network slice instances. The KPI is calculated the utilization of virtualization resources like processors, memory, disk etc. that are allotted to a particular network slice instance.

Logical formula: The KPI is calculated by the utilization of virtual resources divided by the system level capacity that are allocated to the network slice instance.

Physical formula:

$$VRU_{Processor} = \frac{MeanProcessorUsage}{System\ Capacity_{Processor}} * 100\%$$

$$VRU_{Memory} = \frac{MeanMemoryUsage}{System\ Capacity_{Memory}} * 100\%$$

$$VRU_{Disk} = \frac{MeanDiskUsage}{System\ Capacity_{Disk}} * 100\%$$

KPI also depends on Some other parameters such as MeanProcessorUsage, MeanMemoryUsage, MeanDiskUsage; System capacity (the amount of virtual resources that are allocated to the network slice instance; 5GS; utilization; KPI unit=percentage, and KPI type: Ratio.

CHAPTER 5
NETWORK LOG AND DATASET COLLECTION

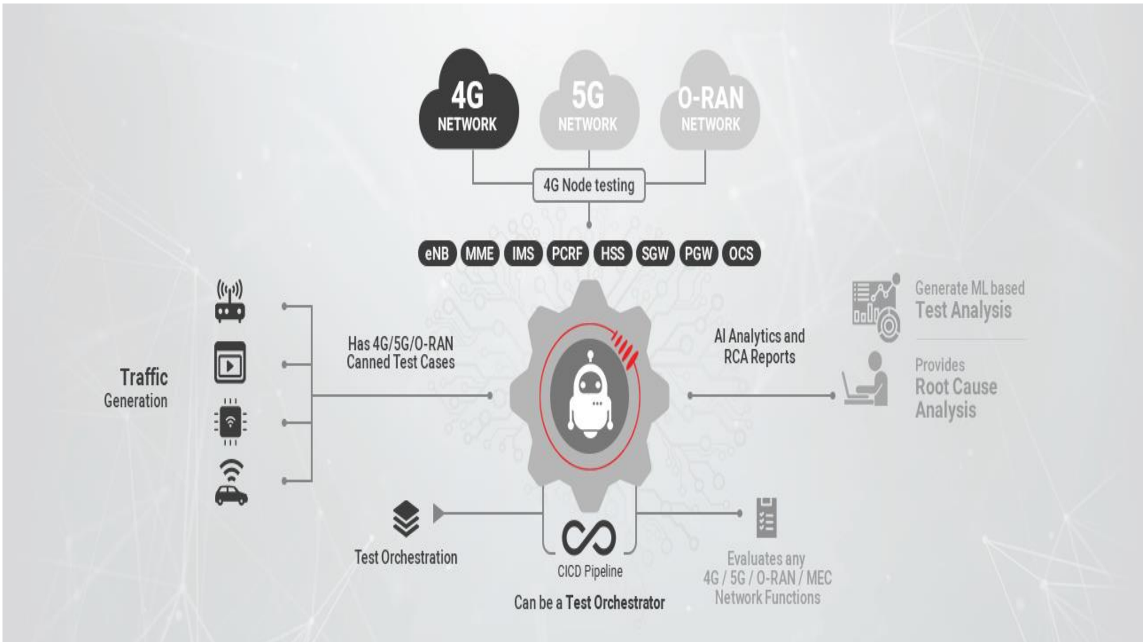


Figure 5.1: Traffic generation to analysis the data

In Figure 5.1, we can see how traffic is generated and data collected. Traffic is generated from different IoT devices using different wireless technologies like 4G/5G and O-RAN technologies. This thesis seeks to analyze and predict these traffic patterns using machine learning techniques.

The latest generation of cellular mobile technology empowers not only faster speeds than 4^{th} Generation, but it is also enables an incredible amount of capacity in the network. The next generation gives enhanced capacity because it uses high frequency band, above 3GHz, never before used in cellular networks. 5G networks enable high capacity, which increases the mobile data capacity almost 2.7 times more than the 4G network. It is essential to collect and analyze the traffic data to ensure better performance and optimize operation and management. We can use different tools to collect data and analyze the traffic pattern, troubleshoot and ensure performance. A number of existing tools are frequently used in cellular network to analysis and troubleshooting, including ABot, QXDM, QCAT, QPST, Wireshark and Splunk. In this project, we used call details record (CDR) as datasets and then implemented machine learning algorithms to find the mobile traffic pattern. Raw CDR is generated in core network elements called session management function and call session control function elements in the core network. CDR contains all kind of user and network information.

5.1 ABot Test Orchestration Solution

ABot test orchestration solution supports network functions for both LTE and 5G wireless network. It can also perform testing on both control and user planes by simulating network functions in a containerized cloud native and virtualized metal environment. ABot is an analytical tool to analyze the log, events, and KPI from different layers of NFV platform, and also perform anomaly detection and data insight.

Network slicing service is very crucial in maintaining performance of the 5G network. ABot can also validate different 5G use cases through network slicing. It can generate traffic to different types of network slicing and interlink with PCF and NRF. ABot is very efficient and capable to support 4G EPC nodes, 5G SBA core network functions and the gNB N1/N2 interfaces.

5.2 Qualcomm Tools

5.2.1 QXDM analyzing tool.

QXDM is a popular tool and broadly used in wireless LTE and NR 5G technology to capture the packet data based on mobile signaling and data details log for debugging and analyzing the performance. QXDM is the most acceptable tool to carrier acceptance and initial product development tool in both field and lab testing.

Here we share some important logs, containing packet information in wireless technology and how QXDM is used to debug and develop the initial product.

2017 Aug 1 14:20:35.970 [35] 0xB113 LTE LL1 PSS Results

```
Version                    = 22
Number of Half Frames      = 4
Sub-frame Number           = 5
System Frame Number        = 30
Number of PSS Records      = 16
PSS Records
```

#	PSS Peak Value (dB)	Peak Position	PSS Indicies
0	20.954	12171	2
1	6.284	12112	2
2	5.812	12170	2
3	5.593	12230	2
4	5.518	12102	2
5	4.861	12240	2
6	4.393	12172	2
7	1.576	12229	2
8	1.383	12101	2
9	0.746	12113	2
10	0.263	12111	2
11	-0.280	12241	2
12	-0.580	12239	2
13	-0.902	12171	0
14	-0.902	12103	2
15	-1.627	12231	2

2017 Aug 1 14:20:35.998 [37] 0xB176 LTE Initial Acquisition Results

```
Version                                = 3
E-ARFCN                                = 39550
Band                                   = 40
Duplex Mode                            = TDD
Result                                 = Success
Min Search Half Frames                 = 1
Min Search Half Frames Early Abort     = 1
Max Search Half Frames                 = 1
Max PBCH Frames                        = 0
Number of Blocked Cells                = 0
Number PBCH Decode Attemp Cells        = 1
Number of Search Results               = 1
Search Results
```

#	Frame Offset	Sample Offset	Physical Cell ID	CP	Frequency Offset (Hz)	PSS Correlation Result	SSS Power Value
0	Unknown	128240	401	Normal	-24	0	0.000228

PBCH Decode Attempt Cells

#	Frame Offset	Sample Offset	MIB Payload	Updated Frequency Offset (Hz)	Physical Cell ID	Number of Decode Attempts	Decode Result	Number of Tx Antennas
0	Unknown	128240	0xA9680000	0	401	1	Success	2

Figure 5.2a and 5.2b: Initial Access Log Packets Related to Cells Search Procedures

```
2017 Aug  1  14:20:36.018  [B6]  0xB130  LTE LL1 PDCCH Decoding Result
Version                = 21
Sub-frame Number       = 5
System Frame Number    = 362
Demback Mode Select    = 1
Carrier Index          = PCC
Number of Hypothesis = 2
Hypothesis
   -------------------------------------------------------------------------------------------------------------------------------
   |   |                   |           |         |Search|      |              |     |       |       |                          |
   |   |                   |Aggregation|         |Space |DCI   |              |Start|Payload|Tail   |                          |
   |#  |Payload            |Level      |Candidate|Type  |Format|Decode Status |CCE  |Size   |Match  |Prune Status              |
   -------------------------------------------------------------------------------------------------------------------------------
   |  0|0x8ED8504800000000|       Agg4|        0|Common|    1A|        SI_RNTI|    0|     47|  Match|      FAIL_SURVIVOR_SELECT|
   |  1|0x8ED8504800000000|       Agg8|        0|Common|    1A|        SI_RNTI|    0|     47|  Match|            SUCCESS_DCI1A|
```

```
2017 Aug  1  14:20:36.328  [51]  0xB173  LTE PDSCH Stat Indication
Version      = 5
Num Records  = 12
Records
   ------------------------------------------------------------------------------------------------------------------------------------------------------
   |   |        |     |   |      |Num      |       |Transport Blocks
   |   |        |     |   |      |Transport|Serving|     |  |   |      |          |     |Discarded|           |        |   |          |   |
   |   |Subframe|Frame|Num|Num   |Blocks   |Cell   |HARQ |  |   |CRC   |          |TB   |reTx     |Did        |TB Size |   |Modulation|Num|ACK
   |#  |Num     |Num  |RBs|Layers|Present  |Index  |ID   |RV|NDI|Result|RNTI Type |Index|Present  |Recombining|(bytes) |MCS|Type      |RBs|Dec
   ------------------------------------------------------------------------------------------------------------------------------------------------------
   |  0|       5|  362| 20|     2|        1|  PCell|    0| 2|  0|  Pass|        SI|    0|     None|         No|      31|  5|      QPSK| 20|
   |  1|       0|  368| 20|     2|        1|  PCell|    1| 0|  0|  Pass|        SI|    0|     None|         No|      35|  6|      QPSK| 20|
   |  2|       0|  382|  8|     2|        1|  PCell|    0| 0|  0|  Pass|        RA|    0|     None|         No|      10|  0|      QPSK|  8|
   |  3|       8|  384| 12|     2|        1|  PCell|    0| 0|  1|  Pass|    Temp-C|    0|     None|         No|      44|  0|      QPSK| 12|
```

```
2017 Aug  1  14:20:36.020  [89]  0xB0C0  LTE RRC OTA Packet  --  BCCH_DL_SCH / SystemInformationBlockType1
Pkt Version = 8
RRC Release Number.Major.minor = 10.7.2
Radio Bearer ID = 0, Physical Cell ID = 401
Freq = 39550
SysFrameNum = 362, SubFrameNum = 5
PDU Number = BCCH_DL_SCH Message,    Msg Length = 28
SIB Mask in SI =  0x02

Interpreted PDU:
```

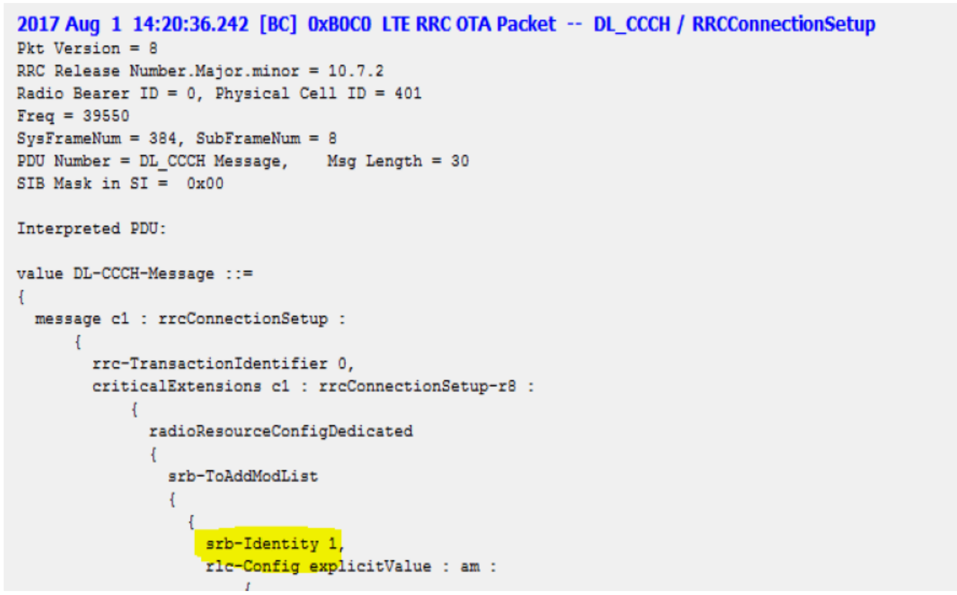
Figure 5.3a, 5.3b, and 5.3c: Initial Message from UE (SIB Decode Logs)

```
2017 Aug  1  14:20:36.091  [AE]  0xB166  LTE PRACH Configuration
Version                    = 3
Logical Root Seq Index     = 20
PRACH Config               = 3
Preamble Format            = 0
Duplex Mode                = TDD
High Speed Flag            = 0
PRACH Frequency Offset     = 2
Max Transmissions MSG3     = 4
Cyclic Shift Zone Length = 12
RA Response Window Size    = 10
```

Figure 5.4: PRACH configuration parameter

```
2017 Aug  1  14:20:36.152  [AC]  0xB061  LTE MAC Rach Trigger
Version = 1
Number of SubPackets = 2
SubPacket ID = 3
SubPacket - ( RACH Config Subpacket )
   Version = 2
   SubPacket Size = 28
   RACH Config V2:
      Preamble initial power = -108 dB
      Power ramping step = 2 dB
      RA index1 = 28
      RA index2 = 56
      Preamble trans max = 10
      Contention resolution timer = 64 ms
      Message size Group_A = 7
      Power offset Group_B = 0 dB
      PMax = 23 dBm
      Delta preamble Msg3 = 4
      PRACH config = 3
      CS zone length = 12
      Root seq index = 20
      PRACH Freq Offset = 2
      Preamble Format = 0
      High speed flag = 0
      Max retx Msg3 = 4
      RA rsp win size = 10 ms
SubPacket ID = 5
SubPacket - ( RACH Reason Subpacket )
   Version = 1
   Subpacket Size = 20 bytes
   RACH Reason V1:
      Rach reason = CONNECTION_REQ
      RACH Contention = Contention Based RACH procedure
      Maching ID = 0x5D, 0x8B, 0x69, 0x2C, 0xFC, 0xE2
      Preamble = 0
      Preamble RA mask = 0xFF
      Msg3 size = 6 bytes
      Group chosen = Group A (0)
      Radio condn = 81 dB
      CRNTI = 0x8E42
```

Figure 5.6: RACH trigger report

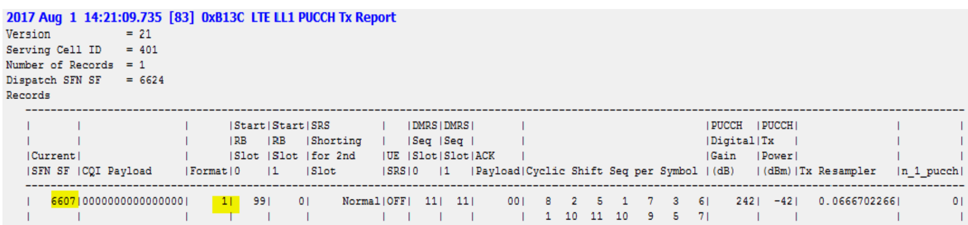


```
2017 Aug  1  14:20:36.242  [BC]  0xB0C0  LTE RRC OTA Packet  --  DL_CCCH / RRCConnectionSetup
Pkt Version = 8
RRC Release Number.Major.minor = 10.7.2
Radio Bearer ID = 0, Physical Cell ID = 401
Freq = 39550
SysFrameNum = 384, SubFrameNum = 8
PDU Number = DL_CCCH Message,    Msg Length = 30
SIB Mask in SI =  0x00

Interpreted PDU:

value DL-CCCH-Message ::=
{
  message c1 : rrcConnectionSetup :
      {
        rrc-TransactionIdentifier 0,
        criticalExtensions c1 : rrcConnectionSetup-r8 :
            {
              radioResourceConfigDedicated
              {
                srb-ToAddModList
                {
                  {
                    srb-Identity 1,
                    rlc-Config explicitValue : am :
```

Figure 5.7: RRCC Connection setup report



```
2017 Aug  1  14:21:09.735  [83]  0xB13C  LTE LL1 PUCCH Tx Report
Version            = 21
Serving Cell ID    = 401
Number of Records  = 1
Dispatch SFN SF    = 6624
Records
   ---------------------------------------------------------------------------------------------------------------------------------------------------------
   |       |                 |      |Start|Start|SRS           |   |DMRS|DMRS|       |                             |PUCCH  |PUCCH|                  |         |
   |       |                 |      |RB   |RB   |Shorting      |   |Seq |Seq |       |                             |Digital|Tx   |                  |         |
   |Current|                 |      |Slot |Slot |for 2nd       |UE |Slot|Slot|ACK    |                             |Gain   |Power|                  |         |
   |SFN SF |CQI Payload      |Format|0    |1    |Slot          |SRS|0   |1   |Payload|Cyclic Shift Seq per Symbol  |(dB)   |(dBm)|Tx Resampler      |n_1_pucch|
   ---------------------------------------------------------------------------------------------------------------------------------------------------------
   |   6607|0000000000000000|     1|   99|    0|        Normal|OFF|  11|  11|     00|  8   2   5   1   7   3   6|    242|  -42|  0.0666702266|        0|
   |       |                 |      |     |     |              |   |    |    |       |  1  10  11  10   9   5   7|       |     |              |         |
```

Figure 5.8: PDCCH Report

5.2.2 QCAT

QCAT is used to analyze the file(log) generated by the mobile capture packet.

5.2.2 QPST

QPST is used for connecting COM port of the mobile and analysis the capture packets in different protocol layers.

5.3 Wireshark Analyzer

Wireshark has open source analyzing tools that can capture packets and filter the packets. It can capture live data from any of 5G wireless networks. Wireshark can filter data flow and be used in attach procedure.

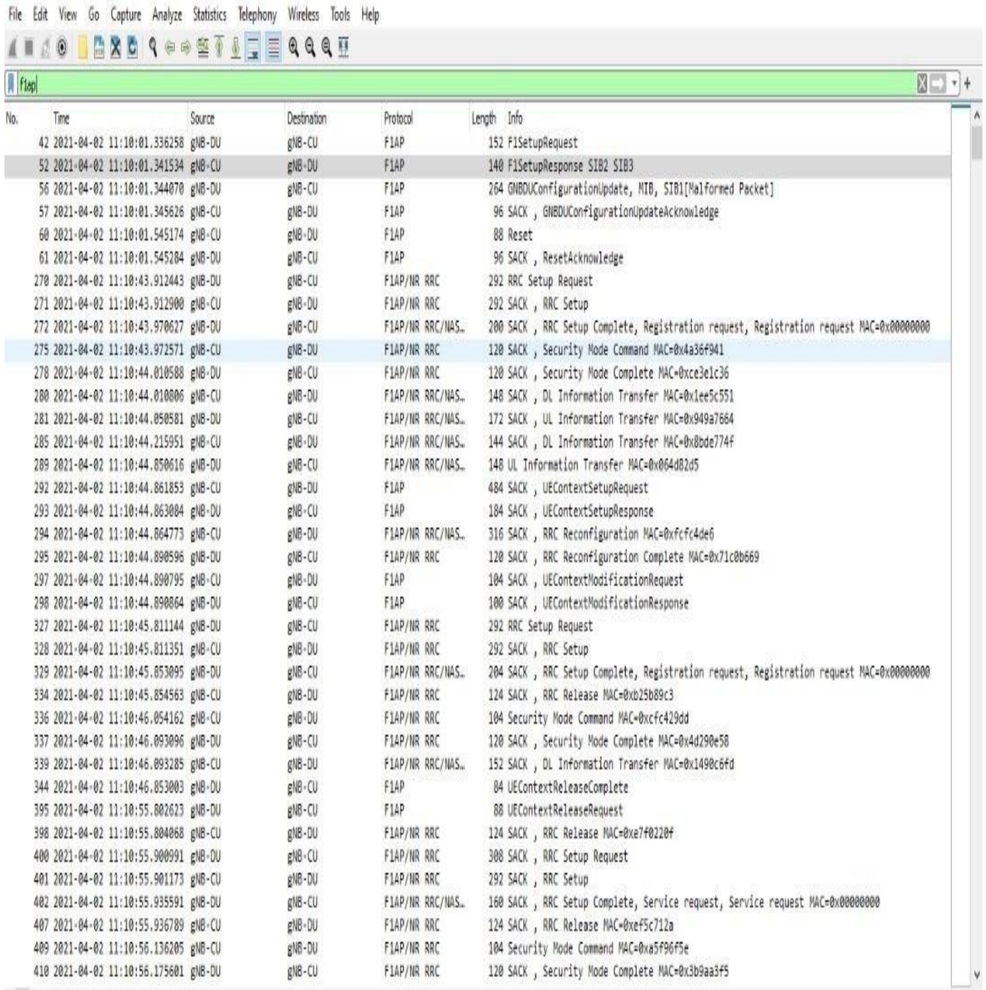


Figure 5.9: Wireshark packets filter

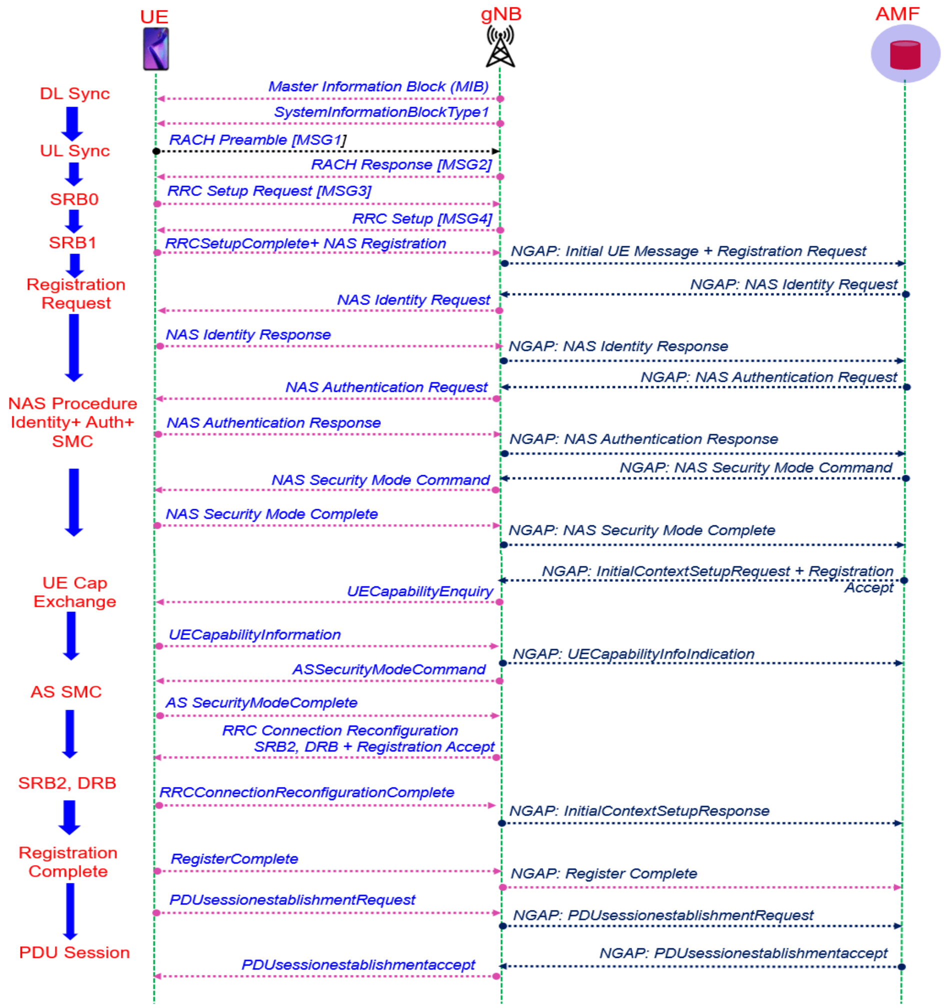


Figure 5.10: Wireshark traces UE to AMF

Table 5.1: Wireshark Nodes wise call traces

Time	Source	Destination	Protocol	Message
18:13:00.476	gNB	AMF	NGAP	NG Setup Request
18:13:00.476	AMF	gNB	NGAP	NG Setup Response
18:13:00.476	gNB	AMF	NGAP/NAS-5GS	InitialUEMessage, Registration request, Registration request
18:13:00.476	AMF	gNB	NGAP/NAS-5GS	DownlinkNASTransport, Authentication request
18:13:00.684	gNB	AMF	NGAP/NAS-5GS	UplinkNASTransport, Authentication response
18:13:00.684	AMF	gNB	NGAP/NAS-5GS	DownlinkNASTransport, Security mode command
18:13:00.703	gNB	AMF	NGAP/NAS-5GS	UplinkNASTransport, Security mode complete
18:13:00.703	AMF	gNB	NGAP/NAS-5GS	DownlinkNASTransport, Registration accept
18:13:00.722	gNB	AMF	NGAP/NAS-5GS	UplinkNASTransport, Registration complete
18:13:01.011	gNB	AMF	NGAP/NAS-5GS	UplinkNASTransport, UL NAS transport, PDU session establishment request
18:13:01.012	AMF	gNB	NGAP/NAS-5GS	InitialContextSetupRequest, DL NAS transport, PDU session establishment accept
18:13:01.122	gNB	AMF	NGAP	UERadioCapabilityInfoIndication
18:13:01.324	gNB	AMF	NGAP	InitialContextSetupResponse

5.4 Python-based Automated Testing Framework.

Here is the most popular python-based testing framework.

- Robot framework
- PyTest
- Unittest
- Doctest

Robot framework is an open-source python base framework. It is quite helpful and simple for automation testing and acceptance testing those overall speeds up the entire testing process. Robot framework test data is defined in different sections listed below.

Table 5.2: Robot Framework testing process

Section	Description
Settings	It is used for importing resource files, libraries, and variable files. Also used for defining metadata for test cases and test suites.
Variables	Used for defining variables that can be used elsewhere in test data.
Test Cases	It is used to create test cases from available keywords
Tasks	Used to create tasks using available keywords
Keywords	Creating user keywords from available lower-level keywords
Comments	Additional comments that are normally ignored by the framework

PyTest: PyTest is also opensource Python based testing framework use generally for all purpose but particularly it is most popular for functional and API testing.

Unittest: Unittest is very first python based automated testing framework. Standard workflow of Unittest:

- First import the Unittest module
- Define own class
- Define a function inside the Class

- Put unittest.main() which is the main method at the bottom of the code to run the test case

CHAPTER 6
DATASETS AND MOBILE TRAFFIC ANALYSIS USING ML

6.1 Dataset for Traffic Analysis in Mobile Network

Understanding network utilization and performance, it is very important to analyze mobile network data generated by subscribers. This chapter describes our dataset and mobile traffic analysis procedure. Network resource utilization and proper optimization of the network resources will make the network more efficient.

However, for research it is very challenging to get access to real time mobile data. Mobile network operators hardly ever release the full datasets of mobile traffic due to privacy. Restriction to access of real time mobile datasets makes it very difficult to analyze the mobile traffic profile and prediction. Mobile network operators share their datasets with limited volume for billing and monitoring purposes. There are a huge number of datasets generated in mobile networks. By analysis of this big amount of data, we can allocate proper network resources and can predict mobile traffic patterns.

For this project we used a dataset available through Kaggle, an online service offering public datasets for scientific analysis. This dataset is a call details record (CDR) suitable for analysis of phone activity. Originating from a call session controller within the network, the CDR is the heart node of a mobile network, which is called SMF in a 5GS core network and call session control function (CSCF) in an IMS network. Subscriber information or events are stored in this node and are accessed through the call details record. Raw CDR generates in the node and contains many features, Operator billing departments collect this raw CDR into their database and then perform mediation process. This dataset is basically composed of one week of call details record (CDR) from a city. This CDR contains some information like time stamp, cell ID, country code, voice or audio calls, text message and Internet usage. Cell ID is a generally unique number carrying the identification of base transceiver station (BTS) or a sector of the BTS within a location area code. Country code is three-digit area code followed by 7-digit local number. For an example US country code is 001. The analysis of this dataset will determine the phone activity in a city within different time frames. We can categorize the phone activity by calls, sms and internet users with hourly, daily and cell wise activity.

6.2 Datasets Characteristic

The characteristic of datasets in mobile networks are mainly divided into three categories based on the three main services in a network - voice/video call, text message/SMS and internet or data usage. Datasets are analyzed based on these three services by conducting mobile traffic analysis and prediction using machine learning techniques.

The raw data can be sorted by using different kinds of tools. Using this tool, we can collect signaling protocols, interfaces, timestamp, user information, attach/call registration information just to name a few. CDR is generated in the soft switch and then it is transfer by push or pull methods to the billing department to mediation, pursing and reporting the CDR. Call details record contains following features:

- Phone number of the subscriber originating and receiving call (A party and B party)
- Time stamp (date & Time)
- The call duration
- A unique sequence number classify the record.
- Additional digit on the called number used to route or charge the call.
- Call type (Voice or sms)
- Cell ID information
- IMSI/IMEI/PLMAN ID and so on.

6.3 Prediction and Modeling the Mobile Traffic

Machine learning techniques can help to analyze and predict phone activity. Datasets mainly contain few important phone activities along with timestamp and cell ID information.

Phone activities are divided into three categories, sms_in/sms_out, call_in/call_out and Internet or data usage. First, we collected the raw CDR, then we filtered and cleaned the data. Figure 6.1 shows the overall workflow to analyze the CDR. For this project we extensively analyzed the mobile activity with different times with all three categories, total volume of activity, cell wise phone activity utilization and so on.

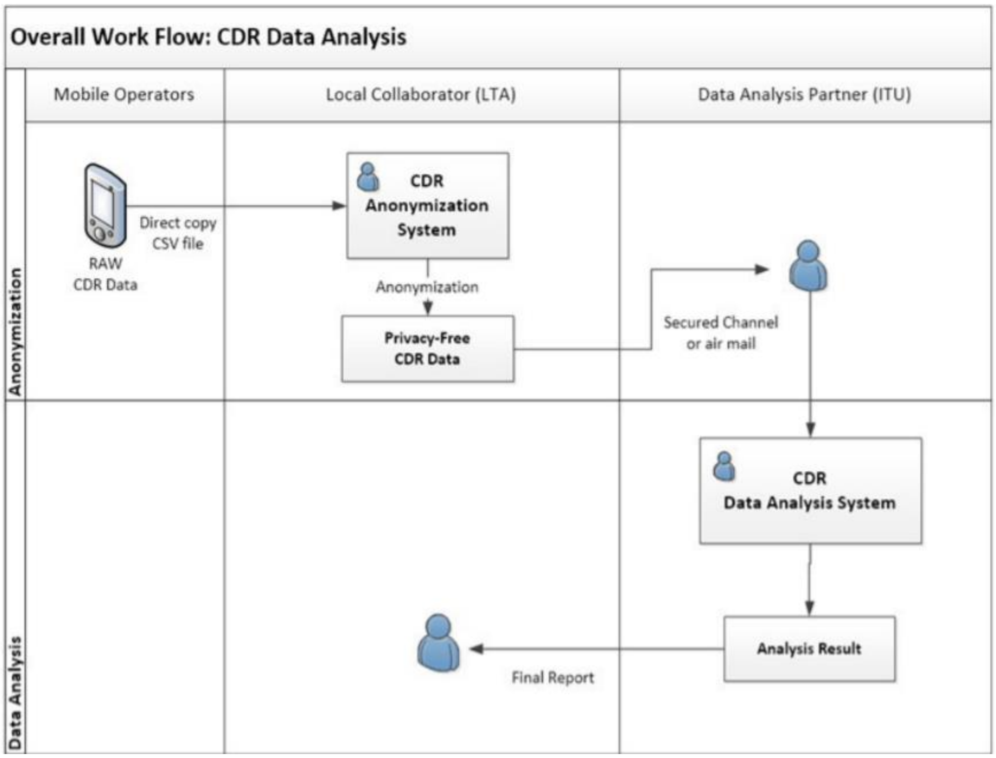


Figure 6.1: Overall Data Analysis Workflow

6.4 Machine Learning Algorithm

Machine learning allows the system to make the decision independently without taking external support. These decisions are made when machine can read the data, learn from the data and understand the fundamental pattern that are contained into it. Figure 6.2 shows types of machine learning and their use cases.

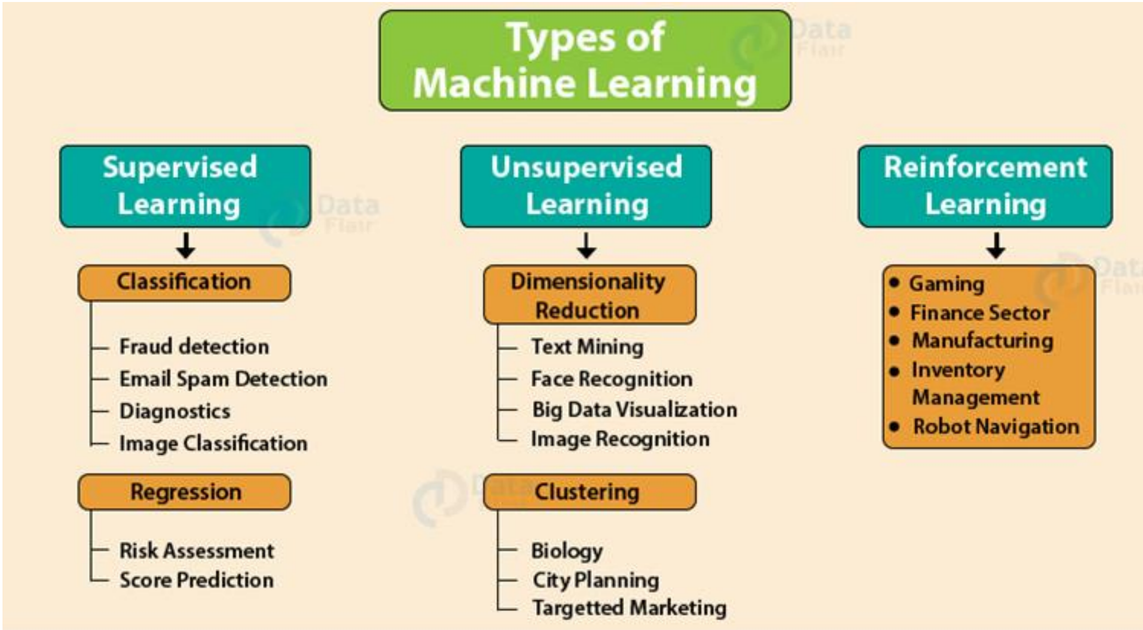


Figure 6.2: Types of Machine Learning

Two popular machine learning algorithms are Linear regression and K-means clustering. These two algorithms were used in this work to explore mobile traffic datasets.

6.4.1 Linear regression

Linear regression is a supervised machine learning algorithm. The methodology of linear regression algorithm is measuring the relationship between two continuous variables. It involves two variables: Independent variable - “X” and dependent variable – “Y". In linear regression the predictor value is an independent value that does not have any fundamental dependency on any variables. The relationship of X and Y are defined as Y =mX + c, where m is the slope and c is the intercept. Based on this relationship equation, we can calculate the predictor or output by the relationship between dependent and independent variables.

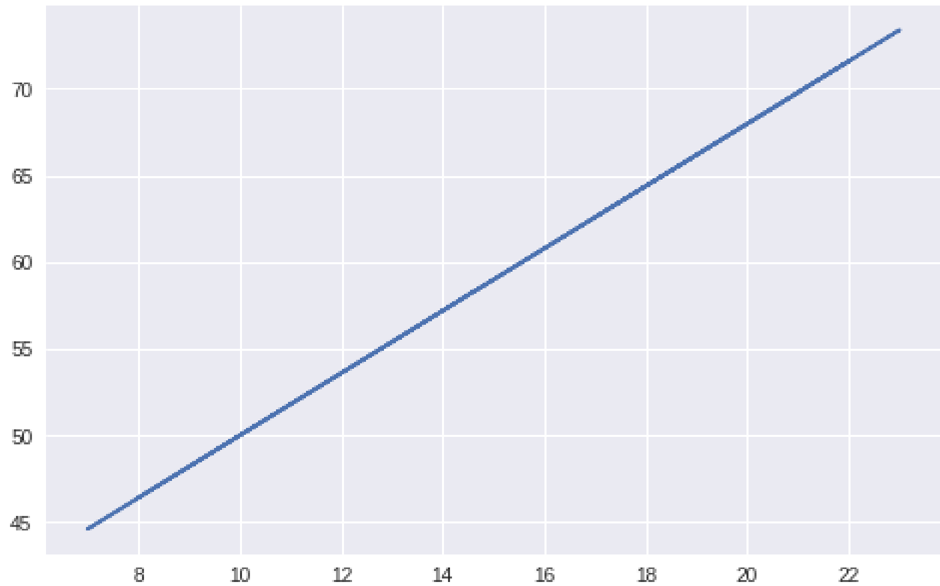


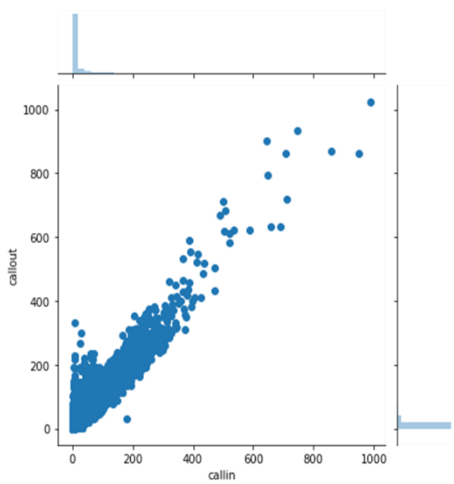
Figure 6.3: Linear Regression

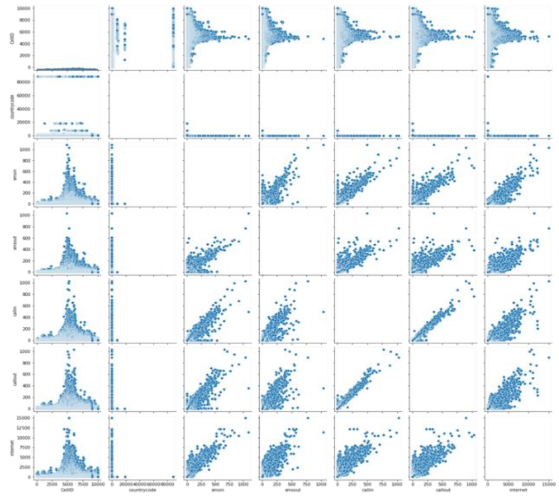
6.4.2 K-means clustering

K-means clustering is an unsupervised machine learning algorithm. It is an iterative algorithm that performs segmentation of the data containing of n values into subsequent k subgroups. Each of the n values with the nearest mean belongs to the K-cluster. The group of objects is partitioned into several subgroups, with the distance of each subgroup related to their centroids. Simply put, k-means is a centroid based algorithm, or a distance-based algorithm that calculates the distances to allocate a particular point to a cluster. Each cluster is then associated with a centroid.

6.5 Mobile Traffic Dataset Analysis Using Machine Learning

Using linear regression, we employed the following libraries for plotting the data: NumPy, pandas, seaborn and matplotlib. The following graphs display our results.





Figures 6.4a and 6.4b: Mobile Traffic Analysis Using Seaborn Joint Plot and Pair Plot of Phone Activity on SMS, CALL and INTERNET.

Here is the exploratory data analysis with day and hourly activity.

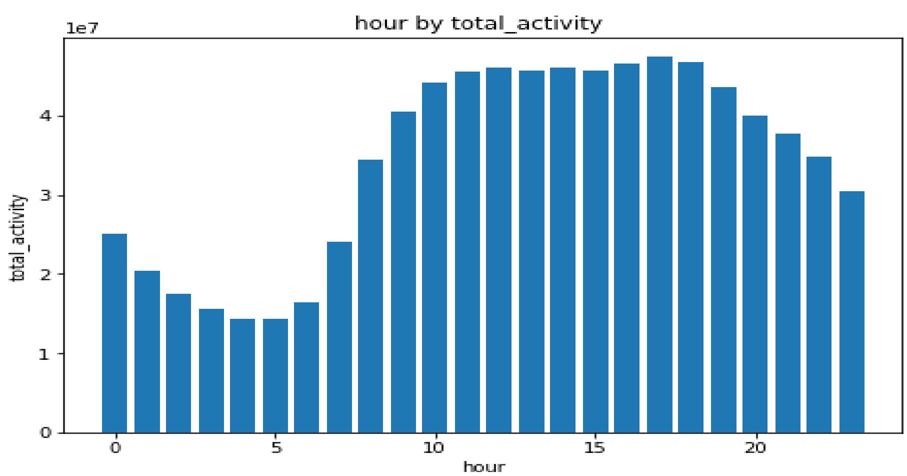


Figure 6.5: Total Activity

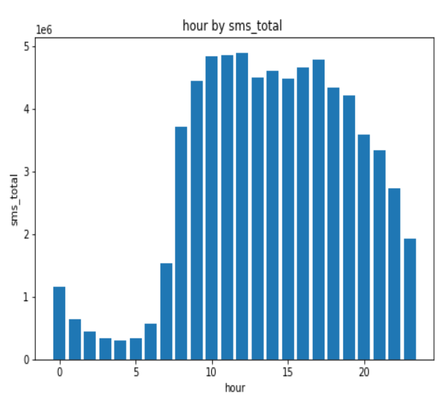


Figure 6.6: SMS Activity

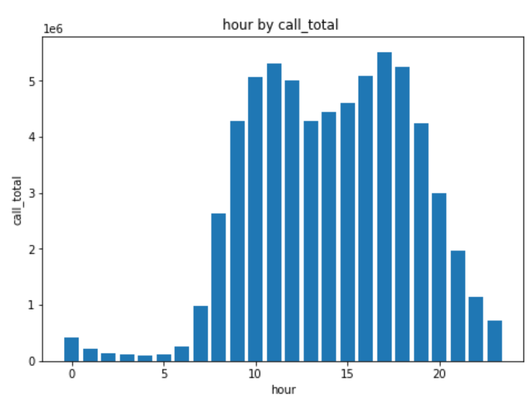


Figure 6.7: Hour by Hour Call Total

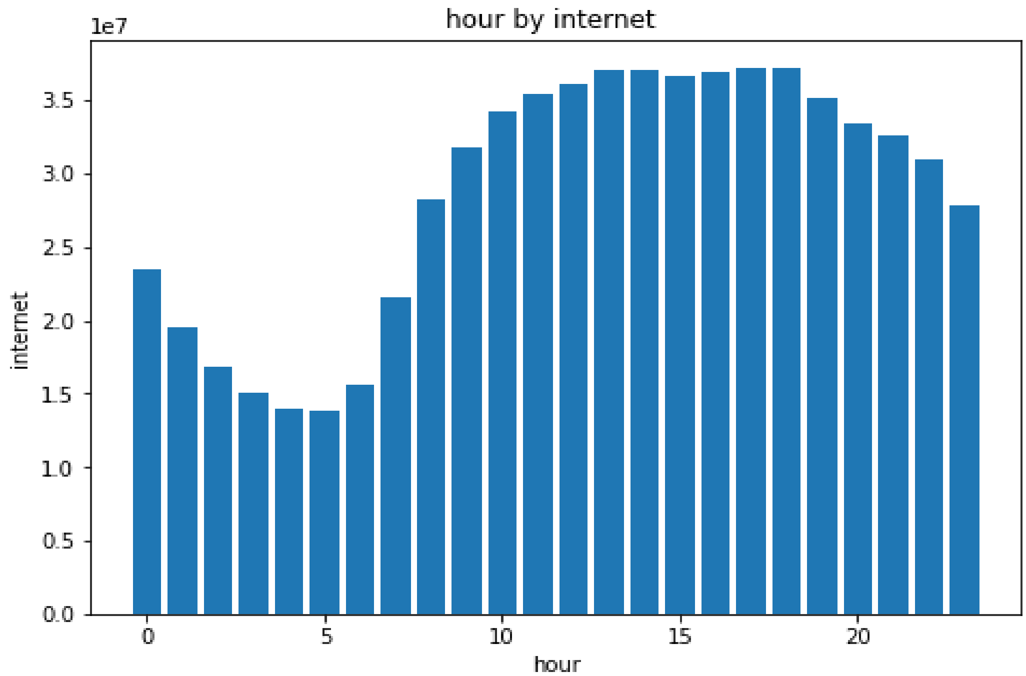
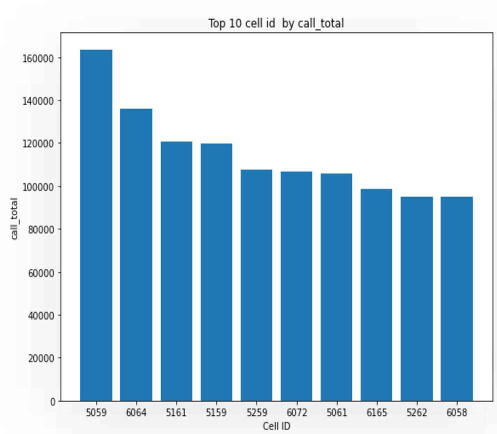


Figure 6.8: Total Internet Activity

These figures show the hourly mobile traffic data analysis. We used matplotlib for plotting the data. Mobile phone activities like sms, call, and internet volumes are calculated hourly so that we can figure it out hourly traffic utilization. Based on the hourly traffic profile, we can optimize and design our network resources. According to the hourly mobile traffic analysis, we can see midnight to 6 AM are the least active hours and 10 AM to 7 PM are the most active hours. Here is the exploratory mobile traffic analysis on top 10 CELL and last 10 CELL.



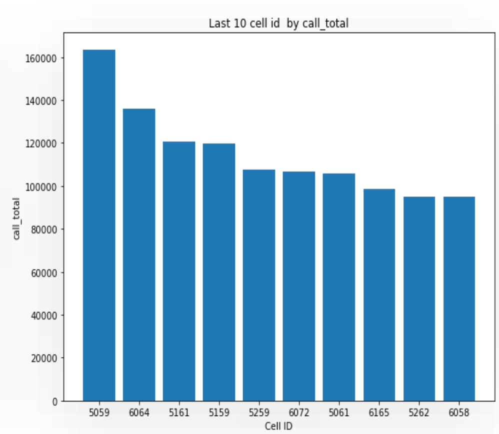
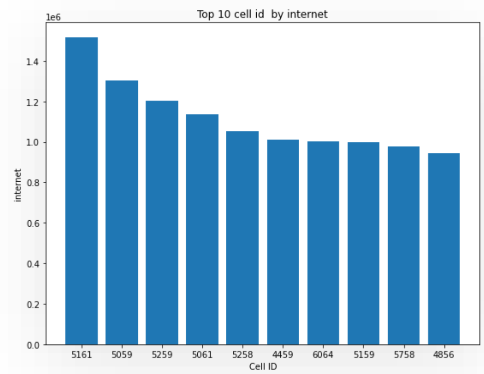


Figure 6.9: Total call activity on top 10 cell

Figure 6.10: Call activity on last 10 cell



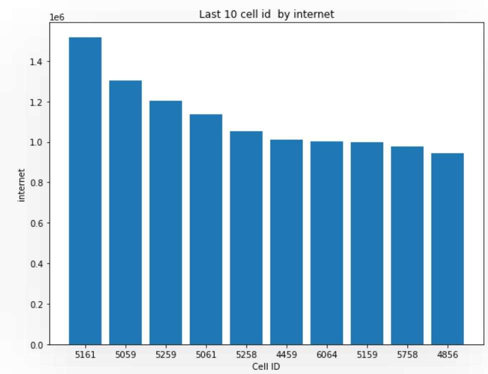
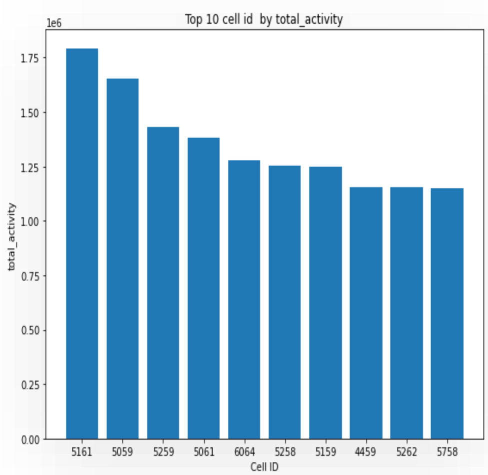


Figure 6.11: Total Internet activity on top 10 cell. Figure 6.12: Internet activity on lat 10 cell



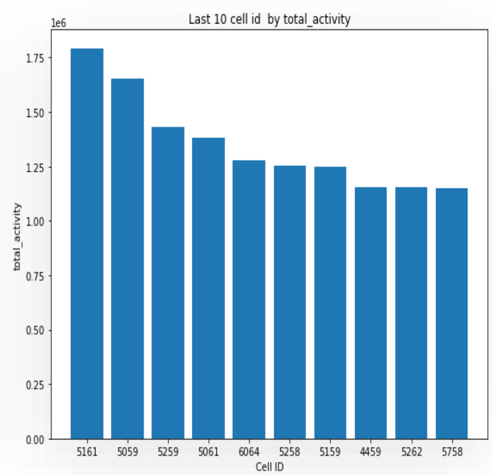


Figure 6.13: Total activity on top 10 cell Figure 6.14 Total activity on last 10 cell

According to the cell ID and phone activity analysis, we can see cell IDs 5000 and 6000 are experiencing maximum utilization. By this analysis result, the engineering team can optimize and plan to proper resource planning.

6.6 K-means Clustering Implementation

We used elbow method to determine the number of clusters. The elbow methods runs k-means clustering of the mobile traffic datasets for a range of values K, for each value of k then computes average score of all clusters. Using elbow method in this dataset, we find 4 clusters.

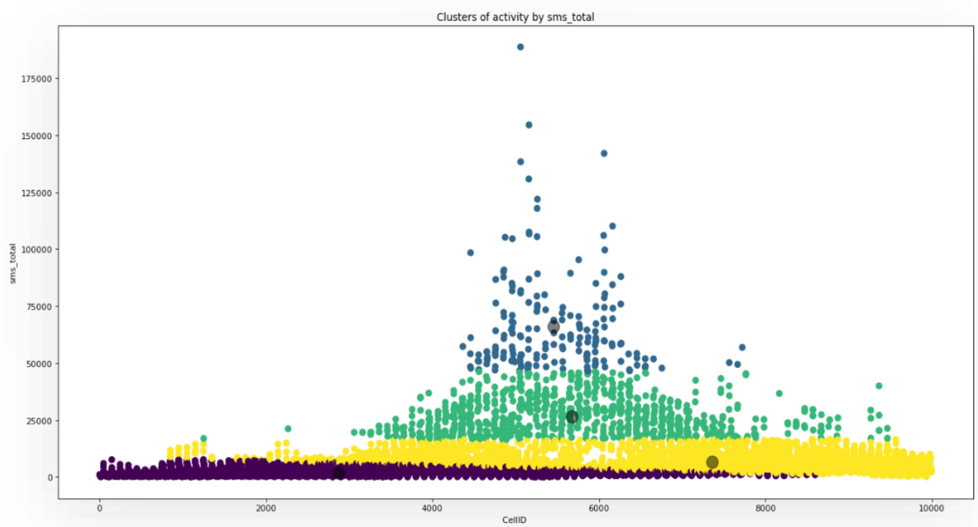
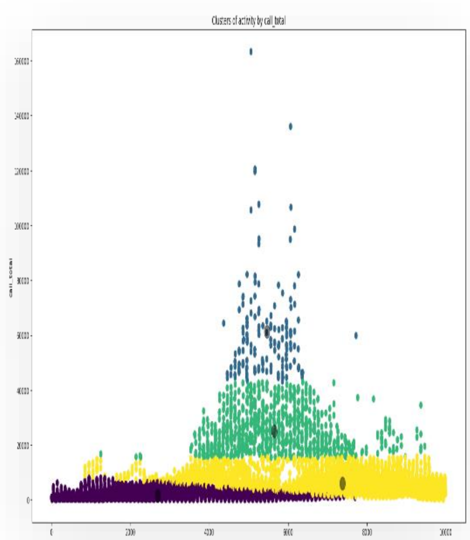


Figure 6.15: Cluster of Total SMS Activity



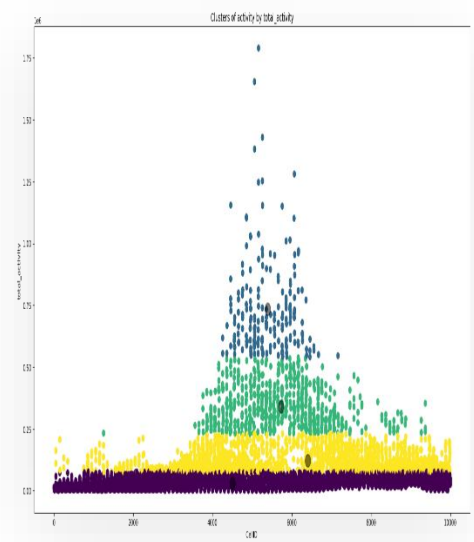


Figure 6.16: Cluster of Call Activity

Figure 6.17: Cluster of Internet Activity

According to mobile traffic analysis by k-means clustering, we can see some cell IDs have very high activity. There are also anomalies in the dataset that need to be explored further in future. Cell Ids from 1 to 3500 are very low utilization. Engineering can identify whether it is due to lower base station or lower subscribers. We can also do some more testing or run k-means clustering again in between the cells 4000 and 6000.

CHAPTER 7
BACKGROUND, MOTIVATION AND RELATED WORK

5G network is the new global wireless standard after 1G, 2G, 3G, and 4G networks. The main features of fifth generation technology is to deliver higher multi-Gbps data speed, lower latency, a more reliable network, vast network capacity and increased bandwidth. It has been observed that a huge number of devices are connected all over the world to the internet using next generation wireless network, generating large amounts of data. It is essential to collect analysis the data and find the behavior and pattern the data for providing better performance, optimizing operation and management in cellular network.

7.1 Related work

In [9], the authors describe the increasing consumption of multimedia services and its demand, and 5G network capable to support future generation vertical applications based on their requirements. They explain the vision of 5G network slicing in different physical layers with multiple isolated logical network layers of varying size and structures, with different types of dedicated services based on the requirements. Network slices are for a massive number of IoT devices, smartphones, autonomous card etc. Software defined networking (SDN) and Network function virtualization in 5G are expected to provide automatic control and management of network resources. They present 5G quality and business requirements followed by descriptions of 5G mobile network softwarization and slicing includes the essential concepts, history, and different use cases. Secondly, they provide 5G network slicing technology including SDN, NFV, MEC, cloud computing, hybrid networks, virtual machine, and contents. Thirdly they broadly survey different industrial initiatives, technologies and projects that are embracing SDN and NFV into 5G network slicing. They also provide various comparison 5G architectural approaches, implementation, adoption technology and deployment strategies [9].

7.2 Intelligent Network data analytics

In [45] the authors describe new features of 5G cellular networks such a network data analytics function (NWDAF). This feature gives an opportunity to mobile operators to analyze the data by their own development machine learning algorithm or incorporate with a 3rd party solution to their network. They analyzed the cell-based dataset in 5G network and predicted the anomalies like suddenly increases in the traffic in particular cells, or anomalies within each cell, and also classified the dataset into subscriber category and user equipment. Then, they implemented three machine learning models to predict the mobile traffic pattern and their behavior, anomalies in the network traffic, traffic load in the network. They mainly implemented three ML models like linear regression, Long-short term memory (LSTM) and recursive neural network. For the prediction of traffic load in each cell, they used three different ML models to minimize the mean absolute error. According to the simulation result, it showed neural network ML algorithm was better than linear regression algorithm for prediction of traffic load. On the other hand, tree-based gradient boosting algorithm outstrips than logistic regression for anomaly detection. These analyses are expected to increase the performance of the 5G network through NWDAF. [45]

7.3 Network Slices Resource Allocation

Network slicing concept in 5G network is a new standard. A single network is divided into multiple network slices with different services based on the requirements. Its also requires the forecast their respective demands. This paper presented an analytical tool called DeepCog for the cognitive management of resources in a 5G system. DeepCog predicts the capacity needed to assist the future traffic demands within the individual network slice. DeepCog hinges on deep learning architecture that helps to predict or forecast the future traffic pattern. Moreover, this tools also carried extensive analysis of the tradeoff between capacity over dimension and un serviced demands in adaptive slice network and in presence of real traffic. [46]

7.4 Mobile traffic prediction using LSTM network.

In [47] the authors described the prediction of mobile traffic from raw dataset using Long-short term memory (LSTM) network. They designed a system to for the traffic prediction using recurrent neural network. In this case mobile traffic data was collected from logical control layer like Physical downlink Control channel (PDCCH) using passive tools. This tool can collect all the information at 1 ms resolution from the base station. Their other objective was to minimize the prediction error given the information extracted from PDCCH. [47]

CHAPTER 8
NETWORK SECURITY AND ANOMALY DETECTION

8.1 Types of Attacks

Most cases of network or enterprise attacks involve flooding or overloading the system or network, collecting data and attacking weak points, and inserting information into the network to spread or gain the access into the network. Next generation cellular network will accommodate a huge number of mobile phone devices and IoT devices. Sensors will connect the network and exchange an enormous amount of data. Network security is critical to manage risk and vulnerability. A system or software called IDS acts as a first step to secure the network and prevent the vulnerabilities.

The most common types of cyber-attacks are follows:

Scanning attacks: Scanning attacks involve injecting some packets or information into the network to gather data about the network topology, open and close port information, types of traffic, active hosts on the network and also types of software or applications running in the network. Blind SQL injection attacks can be used in the network to collect the vulnerabilities. Viruses or malicious code can also be inserted into the network open ports for gathering vulnerability points.

Asymmetric Routing: Asymmetric routing occurrs when packets are sending one network to another network and then route back to a different network. The attackers can use the asymmetric routing to transfer unusual packets into the system bypassing security setups like firewall or some other security setups. If the network allows asymmetric routing it could be vulnerable to SYN flood attacks, which is a type of DDOS attack, so it is generally considered better to turn off asymmetric routing for network protection.

Buffer overflow attacks: This type of attack basically happens in the section of memory in a device on the network. The attacker replaces the data with malicious data and then later tries to attack, in most cases as a DDOS attack. Essentially, a buffer section of memory stores character strings or a large array of integers. When more data is written to the buffer beyond the handle capacity results in the data overflowing into adjacent memory. This can be crash the entire system.

Protocol specific attacks: These types of attacks involve ICMP, TCP and ARP poisoning attacks. The ping and traceroute tools use ICMP to communicate. Ping flood is one kind of ICMP attack where the attacker overwhelms the device with ICMP packet floods. In this situation the device gets busy dealing with huge number of malicious packets and cannot respond to the normal packets. ICMP protocol can also use ICMP tunneling attacks. In this state, data is injected while bypassing security tools like firewalls. Another type of ICMP attack is spoof, where basically the address where packets are coming from is spoofed. ICMP error message is also used to detect the open port by port scanning methods.

TCP SYN flood attack: Transmission control protocol (TCP) are also used to perform a specific kind of attack which is called TCP SYN flood attack. This protocol communicates with the device by three -way handshake. For example, a client sends a request to the server by SYN message to make the connection. The server then acknowledges by sending back SYN-ACK message, and the devices respond with an ACK message to complete the connection. The attackers send a huge number of SYN message to different ports on the servers without ACK message. This is called a TCP SYN flood attack. Basically, in this scenario, ports stay open and wait for the ACK message, and during this time attacker injects a large number of SYN messages, overflowing the server and the system crashes or malfunctions.

ARP flooding: Another protocol attack is ARP, where flooding happens by a huge number of packets sent to recipients to make overflow ARP tables. ARP poisoning also known ARP spoofing; this is mainly performed over a LAN by sending malicious ARP packets to a default gateway on a local area network to change the legitimate IP to MAC address table. This protocol can also translate IP address into the MAC address.

Malware attacks: There are different types of malware like worms, trojans, viruses and bots. These software programs are designed to destroy the system. Most of the time malware is accidentally downloaded via email or any other software. Sometime these viruses get into the system by mishandling the browser, software, network, and network devices. Among all these malware, viruses and worms are more dangerous because this type of malware self-replicates and spreads very quickly into a system and crashes them out. Viruses are mainly attached to .exe files and become active when users run the .exe file. As soon as the user runs this file, viruses become active and spread into one device one after another. Another type of malware is trojan. This is basically a software disguised as a normal program , but it is malware.

Bots malware is a type of software application. Bots start infecting the devices, then connect back to the central control server. This central server can infect all connected devices with bots as a botnet. These are high power attacks. Traffic flooding attacks also often happen on the network. This is also called denial of service and distributed denial of service attacks. Traffic flooding happens by one or multiple external network or devices.

8.2 Splunk Tools to Detect Network Threats.

Splunk is is a tool that can be used for searching, monitoring, visualizing, and analyzing the real time machine data. Designed to remove the obstructions between data and all kinds actions, Splunk is a very powerful tool for IT, DevOps, and security teams. The main features of Splunk are:

- **Visibility:** It help us to gather security and non-security data across the entire network, organization, and different cloud environments platform which gives clear picture and better room for investigation and visual presentation of any kind on incidental responses.
- **Efficiency and context:** Splunk tool can handle unstructured data, it can collect any form of data, eliminate the duplicate data, aggregate, and prioritize the threats. It also improves the security investigation and efficiency.

- **Flexibility:** Splunk tool is very flexible to implement in any emprise solution also can be deployed on any cloud platform and hybrid environment.
- **Behavioral analytics:** This is modern platform of big data can use the machine learning tool kits to detect the issue, try to predict the anomalies, reduce the complexity, accelerate the investigation and response to attack or threats.

8.3 Splunk Machine Learning Toolkits

Traditional intrusion detection systems can only detect known attacks, but it is difficult to detect and predict new attacks or zero days attacks by IDS. Machine learning toolkits help to reduce the false positive rate and improve result.

Splunk machine learning toolkit employs machine learning techniques and methods against all kinds of data. MLTK is a plugin application in Splunk tools that enables users to create, validate, manage and operationalize machine learning models through a guided interface. These models are used to conduct different tasks such as predicting the value of a field, projecting future values, detect the patterns in data and identifying the anomalies from new data.

8.4 Discussion and Conclusion

Nowadays everything is online based. People are using smartphones and applications in their personal life, business, trading, online shopping etc... This kind of daily traffic requires a sustainable infrastructure. 5G network is expected to handle this large amount of data. The concept of 5G network is slicing models, which are divided by different network layers to support different services based on their requirements. An intelligent network monitoring and performance management system is required to control the huge demands. Machine learning tools can help to manage mobile and wireless networks efficiently. Deep neural networks rely on massive and high-quality data to achieve good performance. In contrast, deep reinforcement learning does not make strong assumptions about the target system. It employs function approximation, which explicitly addresses the problem of large state-action spaces, enabling reinforcement learning to scale to network control problems that were previously considered hard. Though there are enormous contributions of deep learning in mobile and wireless networks, there are also limitations that we need to work on in the future. Also, there are anomalies in the datasets that we need to investigate further. We can run k-means clustering again to analyze the data and find a clearer picture.

BIBLIOGRAPHY

[1] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," IEEE Access, vol. 3, pp. 1206–1232, 2015.

[2] K. Zheng et al., "Big data-driven optimization for mobile networks toward 5G," IEEE Netw., vol. 30, no. 1, pp. 44–51, Jan./Feb. 2016.

[3] C. Jiang et al., "Machine learning paradigms for next-generation wireless networks," IEEE Wireless Commun., vol. 24, no. 2, pp. 98–105, Apr. 2017.

[4] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Reinforcement learning with network-assisted feedback for heterogeneous RAT selection," IEEE Trans. Wireless Commun., vol. 16, no. 9, pp. 6062–6076, Sep. 2017.

[5] F. A. Narudin, A. Feizollah, N. B. Anuar, and A. Gani, "Evaluation of machine learning classifiers for mobile malware detection," Soft Comput., vol. 20, no. 1, pp. 343–357, 2016.

[6] K. Hsieh et al., "Gaia: Geo-distributed machine learning approaching LAN speeds," in Proc. USENIX Symp. Netw. Syst. Design Implement. (NSDI), 2017, pp. 629–647.

[7] W. Xiao et al., "Tux2: Distributed graph computation for machine learning," in Proc. USENIX Symp. Netw. Syst. Design Implement. (NSDI), 2017, pp. 669–682.

[8] M. Paolini and S. Fili, Mastering Analytics: How to Benefit From Big Data and Network Complexity: An Analyst Report, RCR Wireless News, London, U.K., 2017.

[9] Alcardo alex, Arslan Ahamed, Rashid Mijunbi and Andrew Hinesi, 5G network slicing using SDV and NFV: An Survey Report, University College Dublin, Ireland, 2019.

[10] C. Zhang, P. Zhou, C. Li, and L. Liu, "A convolutional neural network for leaves recognition using data augmentation," in Proc. IEEE Int. Conf. Pervasive Intell. Comput. (PICOM), Liverpool, U.K., 2015, pp. 2143–2150.

[11] R. Socher, Y. Bengio, and C. D. Manning, "Deep learning for NLP (without magic)," in Proc. Tuts. Abstracts ACL, 2012, p. 5.

[12] (2017). IEEE Network Special Issue: Exploring Deep Learning for Efficient and Reliable Mobile Sensing. Accessed: Jul. 14, 2017. [Online]. Available: http://www.comsoc.org/netmag/cfp/exploringdeep-learning-efficient-and-reliable-mobile-sensing

[13] M. Wang, Y. Cui, X. Wang, S. Xiao, and J. Jiang, “Machine learning for networking: WorkFlow, advances and opportunities,” IEEE Netw., vol. 32, no. 2, pp. 92–99, Mar./Apr. 2018.

[14] M. A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, “Mobile big data analytics using deep learning and apache spark,” IEEE Netw., vol. 30, no. 3, pp. 22–29, May/Jun. 2016.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[15] Paul Patras, Chaoyin Zhang, Deep learning in mobile and wireless networking, IEEE, 2019

[16] Boutaba, R.; Salahuddin, M.A.; Limam, N.; Ayoubi, S.; Shahriar, N.; Estrada-Solano, F.; Caicedo, O.M. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities. J. Internet Serv. Appl. 2018, 9, 1–99. [CrossRef]

[17] 5G end-toend KPI, https://www.techplayon.com/5g-end-to-end-kpi-accessibility-integrity-and-utilization/, Accessed: June,8,2019

[18] I Chih-Lin et al. “Toward green and soft: a 5G perspective”. In: IEEE Communications Magazine 52.2 (2014), pp. 66–73.

[19] 5G protocol stack, https://www.netmanias.com/en/post/oneshot/14103/5g/5g-protocol-stack-user-plane-control-plane, Accessed: Feb,08,2019

[20] NGMN Alliance. “5G white paper”. In: Next generation mobile networks, white paper (2015).

[21] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper. 2017. URL: www . cisco . com / c / en / us / solutions / collateral/service-provider/visual- networking- index- vni/mobilewhite-paper-c11-520862.html.

[22] Jessica Moysen and Lorenza Giupponi. “From 4G to 5G: Self-organized network management meets machine learning”. In: Computer Communications 129 (2018), pp. 248–268.

[23] ETSI. “Self-Configuration of Network Elements Integration Reference Point (IRP); Common Object Request Broker Architecture (CORBA) Solution Set (SS)(Release 8)”. In: 3GPP TR 32.503 (2008).

[24] Multazamah Alias, Navrati Saxena, and Abhishek Roy. “Efficient cell outage detection in 5G HetNets using hidden Markov model”. In: IEEE Communications Letters 20.3 (2016), pp. 562–565.

[25] Lorenza Giupponi et al. “A novel approach for joint radio resource management based on fuzzy neural methodology”. In: IEEE Transactions on Vehicular Technology 57.3 (2008), pp. 1789–1805.

[26] Emil J Khatib et al. “Diagnosis based on genetic fuzzy algorithms for LTE selfhealing”. In: IEEE Transactions on vehicular technology 65.3 (2015), pp. 1639– 1651.

[27] Paulo Valente Klaine et al. “A survey of machine learning techniques applied to self-organizing cellular networks”. In: IEEE Communications Surveys & Tutorials 19.4 (2017), pp. 2392–2431.

[28] Qingchen Zhang et al. “A survey on deep learning for big data”. In: Information Fusion 42 (2018), pp. 146–157.

[29] Kritika Walia. “Self Driving Car Market Report| Industry Share, Trends & Growth Analysis To 2025”. In: (2019).

[30] Norman Jouppi et al. “Motivation for and evaluation of the first tensor processing unit”. In: IEEE Micro 38.3 (2018), pp. 10–19.136 Bibliography

[31] Li Deng. “The MNIST database of handwritten digit images for machine learning research [best of the web]”. In: IEEE Signal Processing Magazine 29.6 (2012), pp. 141–142.

[32] Jia Deng et al. “Imagenet: A large-scale hierarchical image database”. In: 2009 IEEE conference on computer vision and pattern recognition. Ieee. 2009, pp. 248– 255.

[33] Vincent D Blondel et al. “Data for development: the d4d challenge on mobile phone data”. In: arXiv preprint arXiv:1210.0137 (2012). [15] OpenSignal.com. https://opencellid.org/.

[34] Nicola Bui and Joerg Widmer. “OWL: a reliable online watcher for LTE control channel measurements”. In: ACM All Things Cellular (2016).

[35] Simon Haykin. Neural networks: a comprehensive foundation. Prentice Hall PTR, 1994.

[36] Christopher M Bishop. Pattern recognition and machine learning. Springer Science+ Business Media, 2006.

[37] Hoang Duy Trinh et al. “Analysis and Modeling of Mobile Traffic Using Real Traces”. In: PIMRC (2017).

[38] Arcangela Rago et al. “Unveiling Radio Resource Utilization Dynamics of Mobile Traffic through Unsupervised Learning”. In: 2019 Network Traffic Measurement and Analysis Conference (TMA). IEEE. 2019, pp. 209–214.

[39] Hoang Duy Trinh, Lorenza Giupponi, and Paolo Dini. “Mobile traffic prediction from raw data using LSTM networks”. In: 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE. 2018, pp. 1827–1832.

[40] Hoang Duy Trinh, Lorenza Giupponi, and Paolo Dini. “Urban Anomaly Detection by processing Mobile Traffic Traces with LSTM Neural Networks”. In: Proceedings of the 2019 IEEE International Conference on Sensing, Communication and Networking(SECON) (2019).

[41] Hoang Duy Trinh et al. “Detecting Mobile Traffic Anomalies Through Physical Control Channel Fingerprinting: A Deep Semi-Supervised Approach”. In: IEEE Access 7 (2019), pp. 152187–152201.

[42] Hoang Duy Trinh et al. “Mobile Traffic Classification through Physical Channel Fingerprinting: a Deep Learning Approach”. In: arXiv preprint arXiv:1910.11617 (2019).

[43] Voice over 5G, https://the5gzone.com/index.php/voice-over-5g/,

[44] 5G standalone Access registration, https://medium.com/5g-nr/5g-standalone-access-registration-fe80aa28d723, Nov,17,2018

[45] Salih Sevgican, Meric Turan, kerim Gokarslan, Birkan Yilmaz, Tuna Tugcu “Intelligent Ntework data analytics function in 5G cellular network using machine learning Algorithm”, , journal of communication and Networks, VOL.22 NO., 3, June, 2020

[46] Zheng Yibing, China starts making 5G hospital network standard, September, 4, 2019, https://news.cgtn.com/news/2019-09-04/China-initiates-making-5G-hospital-network-standard-JI1xOe352E/index.html

[47] Mckinsey Digital, https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-can-we-recognize-the-real-power-of-the-internet-of-things

[48] Broadband search, Mobile Internet usage data, https://www.broadbandsearch.net/blog/mobile-desktop-internet-usage-statistics, 2021

[49] Dario Bega, Marco Gramagilia, Marco Fiore, Albert Banchs, Xaviar Costa “Cognitive Network Management in sliced 5G networks with deep learning” IEE infocom, 2019

[50] Hoang Duy trinh, Lorenza Giupponi, paolo Dini, “Mobile traffic prediction from raw data using LSTM”, IEEE International symposium on personal, indoor and mobile radio communication, 2018