5G Emancipation: A Review of the Panacea for an Efficient Communication Growth in the Evolution of Cellular Networks

Ms. Blessing Chinemerem Dike, Prairie View A&M University, TX USA

Dike, Blessing Chinemerem is a PhD student of Electrical and Electronic Engineering in Electrical and Computer Engineering Department at Prairie View A&M University (PVAMU), TX, USA. Her research interest is in Broadband Communication, Dynamic Spectrum Access and Management, Wavelet and its Applications to Image and Signal processing, and Computer Security. She is a graduate Research Assistant at CECSTR.
5G Emancipation: A Review of the Panacea for an Efficient Communication Growth in the Evolution of Cellular Networks

Blessing C. Dike
Center of Excellence for Communication Systems Technology Research (CECSTR)
Electrical and Computer Engineering Department
Prairie View A & M University, Prairie View Texas

Cajetan M. Akujuobi, P.E.
Center of Excellence for Communication Systems Technology Research (CECSTR)
Electrical and Computer Engineering Department
Prairie View A & M University, Prairie View Texas

Abstract

Wireless communication aims at providing a reliable and high quality communication. The evolution of cellular communication has been a step in the right direction from 1979 to date. However, each generation of cellular network has some requirements like delay, throughput and QoS that must be considered to provide an effective communication. The growth in the consumption of mobile services has resulted to an overload in the cellular networks. This has opened up challenges for resource management in future mobile networks. Therefore, there is need for an effective resource scheduling and sharing schemes to cope with the available bandwidth. This paper provides a review of the different generations of networks to date with reference to their efficient communication resource sharing and scheduling schemes. We will explore possible deployment of effective data driven AI and machine learning algorithms for Radio Access Network (RAN) slicing, which provides reduced latency and an overload reduction in 5G networks. We will also proffer solutions to the new problems encountered by the 5G RAN slices.


List of Nomenclatures
ITU-T – International Telecommunication Union- Telecommunication
3GPP – 3- Group Partnership Project
CDMA- Carrier Division Multiple Access
EDGE - Enhanced Data rate for GSM Evolution
SMS - Short Message Service
MMS – Multimedia Services
GSM- Global System for Mobile Communication
GPRS- General Packet Radio Services
IMT-2000 – Internal Mobile Telecommunication – 2000
DARPA- Defense Advanced Research Project Agency
OFDM – Orthogonal Frequency Division Multiplexing
UDM- Unified Data Management
AMF- Access Mobility Management
SMF- Session Management Function
PCF- Policy Control Function
AUSF- Authentication Server Function

Introduction

Wireless communication aims at providing a reliable and high quality communication and has achieved this over the years by the introduction and implementation of standards by the ITU-T and industry standard group-3GPP, respectively. The mobile cell phone technology started its evolution journey as far back as late 70’s. The evolution was from 1G to 5G, and each of the generation is defined by a set of standards with details of its technological specifications and implementation of the use cases like bandwidth, speed, and latency\(^1\). The positive impact of cellular network evolution has led to an increase in the consumption of mobile services and the continued request for the services often overwhelms the base station. This opens up challenges for resource management in future mobile networks\(^2\). The resource scheduling schemes for the 1G through 4G could not cope with the influx of request for cellular consumption, therefore, the need for 5G and beyond.

The introduction of 5G system came with the requirements that can support industry 4.0 and other wide range of application scenarios such as smart cities, high-tech manufacturing, and remote surgery\(^2\). These requirements are domiciled on network functionalities such as security, mobility, and over load management. 5G’s expected performances are peak data rate above 10Gbps, latencies below 1ms, and 500 km/h mobility target, etcetera\(^2\). These demands cannot be met through a common network settings, hence the exigency to introduce network slicing in 5G, as one of the needs for dynamic spectrum management. Network slicing allows operators to manage logical networks built over a single physical network\(^3\). The beauty of RAN slicing is that it brought improvements in network performance and encourages greater utilization of network resources by dynamically placing the activated network slice with the appropriate amount of resources to meet with their distinct requirements. However, the introduction of 5G New Radio brought about the complexity of the Radio Access Network (RAN), because of the increase in the number of band combinations that would be managed and at the same time extend the capabilities of the network to support multiple network slices with different characteristics\(^3\). Indeed, slicing the RAN, managing RAN resources, and sharing them among network slices is a tedious task\(^3\).

To effect a control over the slice networks, which depends on the aforementioned use cases, there is need for a data driven AI algorithm that can assist in scheduling and control of resources and User equipment needs in other not to overwhelm the gNodeBs (5G base station)\(^4\).

The rest of the paper is divided into sections. Section II discussed the generation of cellular networks, section III treated NG-RAN slicing in 5G networks, and also discussed the data-driven AI algorithm as an efficient method for resource sharing and scheduling of the 5-G use cases. Finally, a conclusion and recommendation for future work was given.
II. Cellular Network Generation

Comparative Analysis of 1G-5G

Nippon Telegraph and Telephone (NTT) commercially launched the first cell phone technology – 1G, in Japan in 1979. Although the technology was analog, it introduced the era of mobility in the telecommunication industry\(^1\). The transmission standard technology used was Advanced Mobile Telecommunication System (AMTS) and the access system was FDMA, which was of low capacity and as such allocated frequencies to each user. It offered only voice calls, which had a poor signal strength and had frequency range of 30 kHz, bandwidth of 2kbps and speed of 2.4kbps. The handoff capability was poor and its major security challenge was eavesdropping.\(^1\)

In 1991, Radiolinja in Finland commercially launched 2G networks on the GSM standard. It implemented the concept of CDMA, whereby multiple users can share band of frequencies through multiplexing. GSM uses FDMA and TDMA technologies for the radio access network. The radio access network is the air interface that aids wireless connectivity between the base station and the cellular phones. It is a digital system and operates at the frequency of 1.8GHz, bandwidth of 14.4-64kbps and speed of 64kbps with GPRS. The 2.5G (EDGE) offers SMS and MMS services. However, its major drawback was that CDMA phones could not transmit both voice and data at the same time. It required strong signal as there was no network coverage in some areas because of poor signal strength and it could not handle complex data like video\(^1\).

The 3G mobile networks was developed between 1990 and 2002 by 3G Partnership Project (3GPP). The smartphone technology was introduced in 3G and offers services such as web browsing, email, video conferencing, video on-demand, gaming, and picture sharing. These services comply with IMT-2000 specifications. It utilizes a new technology called Universal Mobile Telecommunication System (UMTS), as its core network architecture and offered speed of up to 1.92 Mbps on UTMS Wideband CDMA (WCDMA) and currently operates at a downlink speed of 14 Mbps, achieved on a High speed Datalink Packet Access (HSDPA). However, the problem of roaming persisted and it requires the base station to be closer to the user for an improved services. The increased bandwidth and protocol necessitated the upgrade of the cellular device and the base station, which in effect added more cost to the carrier \(^1,\(^2\).

The 4-G mobile network was developed from 2000-2010 and built on technologies such as DAPRA-OFDM, packet switched network and MIMO. The purpose of 4G is to provide high security, lower cost data rate and high-speed data. The maximum speed is 1Gbps. It offers gaming services, IP telephony, high definition mobile TV (HDTV), video conferencing and cloud computing. The important standard for 4G are WIMAX and LTE. Long Term Evolution (LTE) is an upgrade to the existing UMTS\(^1\). The upgrade of 4G such as 4.5-G and 4.9-G got the LTE to the stage called LTE-Pro, which gets us more MIMO and device-to-device communication and a push nearer to 5G requirements. 4G drawback is that cell tower needs new hardware component and data service cost is still high for users. LTE networks are not simple upgrades to existing infrastructure, so they are deployed in parallel and on separate spectrum from existing 3G infrastructure. 4G-LTE offers up to 20 times the capacity of 3G but without any network improvement, 4G network must reach its maximum.

The 5G was introduced in the late 2010 as new generation radio (NGR) or simply new radio (NR) and deployed by cellular phone companies worldwide in 2019. It has up to 400 MHz channels and uses OFDM based air interface technology. It was on a high demand for data subscription and...
data consumption. 5G networks are expected to coexist with 4G LTE and thus require strong integration \textsuperscript{3,4}. The deployment mode for 5G are stand-alone and non-stand-alone networks. 5G stand-alone is the complete 5G network that makes use of its cloud-native core while the non-stand-alone version of 5G network uses the existing 4G LTE core network to convey 5G speed via dual connectivity. 5G bandwidth ranges from 50Mbps to over 1Gbps, although the millimeter wave band is expected to reach 4Gbps with carrier aggregation before the year 2030. 5G networks operate at a low frequency (2G) of 0.6GHz-3.7GHz, Mid-frequency (3G and 4G) of 3.7-24GHz, high frequency (5G) of 24GHz-100GHz, and 26GHz-300GHz which is the millimeter wave (MMW) frequency. It is expected to operate at a bandwidth of 1Gbps or higher and a latency rate of less than 1ms or none. The different generations of the cellular network are summarized in Table 1.

| Table 1: Comparative Analysis of 1G-5G |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Parameters      | 1G              | 2G              | 3G              | 4G              | 5G              |
| Technology      | Analog(AMTS)    | Digital(GSM,CDMA,DAMPS,WIMAX) | Digital(UMTS,WCDMA) | Digital(LTE,WIMAX,MIMO,OFDM) | OFDM,Sub-6GHz,MMW |
| Frequency       | 30kHz           | 1.8GHz          | 1.6-2GHz        | 2-8 GHz         | 24GHZ-100GHz    |
| Maximum Speed   | 2.4Kbps         | 64Kbps          | 144Kbps-2Mbps   | 100 Mbps-1Gbps | 50Mbps-20Gbps   |
| Bandwidth       | 2 Kbps          | 14.4-64 Kbps    | 1Mbps           | 200Mbps-1Gbps  | 10Gbps          |
| Access system   | FDMA            | TDMA/FDMA       | CDMA            | CDMA            | OFDMA           |
| Transmission Standard | AMTS            | Packet-Switched mobile data network | CDMA           | LTE             | OFDMA air interface |
| Core Network    | PSTN            | PSTN            | Packet Network  | Internet        | Internet        |
| Data rate       | < 9.6 Kbps      | 9.6-270 Kbps    | 0.5-5Mbps       | 300Mbps         | 100Mbps-20Gbps  |
| Latency         | <100ms          | 100-1000ms      | 100-500ms       | < 100ms         | >1ms            |
| Transmitted item | Voice only      | Voice/Data      | Enhanced Voice and Data | Packets-only Networks | Voice and Packet Network |
| Data Transmit Mode | No data        | SMS             | SMS and MMS    | SMS and MMS    | SMS and MMS    |

III. 5G Requirements

A. 5G System and Use Cases

5G supports the corresponding use cases: New Radio (NR), New Core Architecture, Network Slicing, and MIMO. It also supports four categories of use cases: Enhanced Mobile Broadband (eMBB), Fixed Wireless Access (FWA), Massive Machine Type Communication (MMTC) and Ultra Reliable Low Latency Communication (URLLC) \textsuperscript{4,5}. In eMBB, people consume more cellular networking video and game streaming and virtual reality content. For FWA, operators have unused spectrum that have already been paid for in the sub- urban area. Operators can meet the world’s need.
of broadband wireless access by using their cellular infrastructure to provide broadband. MMTC supports low data volume, infrequent data, large number of devices, and long battery life. Finally, URLLC supports Ultra reliable and Low latency\textsuperscript{4}.

5G system consists of UE- User Equipment- tablets, laptops, 5G IOT, 5G Routers, for fixed wireless access. The second is the 5G RAN- gNodeBs which supports the new air interface NR (New Radio). The third is Core Network, which contains functions that supports subscriber data storage, mobility management, session management, etcetera. In 5G core, there is a clear separation of user plane (UP) and control plane (CP). 5G is built using modern software development technologies like service based architecture and network virtualization\textsuperscript{4}. Figure 1 is the representation of 5G system.

B. 5G Network Slicing Concept
Each of the network slices is configured specifically for the use case at hand. Functions such as latency, capacity and coverage will be allocated to meet the particular demand of each use case. There is a single physical network but many logical networks running on top of it\textsuperscript{3}. The logical networks are called slices. The concept activated the creation of multiple virtual networks to be kept on top of the common shared physical network called Network Slicing. There can be a slice for MMTC use case that is configured for infrequent data but for very long battery performance. Another slice is for mobile broadband such as smart phone traffic. The next slice is for low latency such as self-driving cars. For instance, an automated car can rely on a particular communication, which requires low latency but not a high throughput. On the other hand, a streaming service watched while in motion will require a high throughput but not low latency. Infact, a user equipment (UE) can connect multiple slices at the same time\textsuperscript{5}. Each slice is optimized for the specific use case and will be completely isolated so that no slice can interfere with the other traffic. This supports migration because new technologies can be introduced on each slice without risking the other slide.

Security impact is high because if a cyber-attack breaches one slice, the other slices are contend. This is because it cannot spread beyond that slice. 5G architecture is service- based; this implies that if one service breaks, the system continues performing. It is flexible, and supports web based interface and development. Figure 2 is the representation of network slicing.

C. 5G RAN Challenges and some Improvements
5G was standardized to carry services such as Voice, Mobile Broadband- carries many data, and MMTC- Very little data, URLLC- carries time critical data, and fixed wireless access (FWA) - for mobile broadband. These resources enable flexibility in 5G network with demanding requirements in the management of the complex and increased number of band combination in the sliced network. These requirements created more complexities in the resource management. Because of these demands, the radio resource controller (RRC) of the 5G network that mediates between the device in use and the radio base station has more complex work to tackle and needs AI and Machine learning algorithms for solution\textsuperscript{6}.

In the light of these complexities, there was an involvement in resource sharing and scheduling for 5G NR, as shown in the work of Jie Mei et al\textsuperscript{3}. They pointed out that an effective slicing of Radio Access Network (RAN) is a challenging task because of diverse QoS requirements and dynamic conditions in the 6G network. They suggested that a RAN slicing framework, which is structured by step-by-step basis, should integrate the personal-management of network resources to get an adaptive control strategy.
Silem Bakri et al\textsuperscript{4}, in their work, investigated that based on the current estimates of the throughput performances of the users of each network slice derived from the channel Quality Indicator (CQI) values, retrieved from the RAN, the resources are adjusted periodically. This enables the continuous reporting of the CQI information to base station by the User Equipment (UE) to the Slice Orchestrator (SO), and the resultant impact is communication overhead. Therefore, a data-driven algorithm is used to reduce the reporting frequency.

Phoung luu po et al\textsuperscript{5}, in their work, ‘Slicing the Edge : Resource Allocation for RAN Network Slicing’, argued that even though there are complex coupling of the RAN resource allocation and coordination for each respective slice that share the same resources, the design of RAN slicing algorithms can address the coupling issues. Po-Chen Chen et al\textsuperscript{6} proposed a Deep Deterministic Policy Gradient (DDPG)-Based Radio Resource Management for user Interactive Mobile Edge Networks used to schedule resource in an edge network environment. They pointed out that user interactivity- based groups could be worked on when a 3D radio resource structure was fused with computerized Markov decision process (MDP).
D. Data-Driven Mechanism for 5G

Data-driven RAN slicing provides an efficient resource sharing and scheduling schemes, whose capabilities match with the requirement for stable gNodeBs. A data-driven RAN slicing mechanism, which depends on the resource-sharing algorithm running at the slice orchestrator, suit the requirement needed to solving issues of 5G RAN. To this effect, a data-driven RAN slicing framework including channel quality-aware algorithms for slice allocation was adopted as the best methods to tackle RAN slicing issues and complexities. The objective of data driven mechanism for RAN slicing is to reduce the overhead of channel reporting frequency by first classifying the status of each user equipment (UE) channel. In this case, an AI Data-driven algorithm is introduced to serve as a predictive technique.

The Least Short Term Memory (LSTM), a kind of Recurrent Neural Network (RNN), is a Deep Learning technique used to reduce the reporting frequency of the 5G base station (gNodeBs). LSTM can learn the long-term dependencies between time steps in time series, as well as sequence data. It can foretell the values of the future time steps of sequence event by its knowledge of the past and present events. A data driven RAN slicing can attain a channel capacity with enormous accuracy on the stimulated data. Its functionalities and performances on scheduling and resource sharing in terms of throughput and delay for MMTC carries a lot of data and URLLC which carries time critical data respectively and is guaranteed.

Source: 5G: Architecture key Principles by Sathish Jagadeesan at Udemy.com

Figure 2: RAN Slicing Architecture

Source: 5G: Architecture key Principles by Sathish Jagadeesan at Udemy.com
Summary and Conclusion

The paper started by tracking the history of cellular communication from 1979 to present. This necessitated the comparative study done of the generations of the network from 1G-5G. We found out that 5G unlike the other G’s, has more improved system. 5G has attained faster data rates, higher connection density, much lower latency and an improved wireless coverage. 5G also has the capability to tackle the cellular network management issues through the provisioning of network slice in Radio Access Network (RAN). The concept of RAN slicing provided an efficient resource sharing and scheduling schemes for cellular networks. However, because of the ambiguities involve in RAN slicing, a data-driven RAN slicing framework in conjunction with channel quality-aware algorithms for slice allocation was suggested as the best method to tackle RAN slicing issues and complexities. In this case, an AI Data-driven algorithm serves as a predictive technique.

More research on the areas of 5G end-to-end (E2E) communication, efficient energy management and improved overall wireless coverage is encouraged.

References